THE UNIVERSITY *of* LIVERPOOL

DEPARTMENT *of* ENGINEERING

ENTROPY-BASED DESIGN OPTIMIZATION OF WATER DISTRIBUTION NETWORKS

Thesis submitted in accordance with the requirements of

The University of Liverpool

for the degree of

Doctor in Philosophy

by

Yohan SETIADI, B.Eng.

April 2006

ENTROPY-BASED DESIGN OPTIMIZATION OF WATER DISTRIBUTION NETWORKS

Yohan Setiadi

ABSTRACT

Entropy has been suggested to be a good surrogate measure for the reliability of water distribution networks. The problem of entropy-based design optimization of water distribution networks has also been formulated. This thesis presents further investigations into the above issues in order to substantiate the suitability of entropy as a surrogate reliability measure. Aspects of design that may have influence on the relationship between entropy and reliability are examined and the characteristics of the maximum entropy designs are studied. Finally, the applicability of the entropy-based design optimization method is brought one step closer to the real water distribution networks.

The novelty and originality of the present research are presented next.

A comprehensive study on the entropy-reliability relationship is carried out in this thesis. The possible influence of layouts, flow directions, cost functions and modelling errors on the relationship is investigated. The results support the previous conviction that entropy is a good surrogate measure for the reliability and that the influence of the above-mentioned aspects of design on the relationship is negligible. The maximum entropy designs are shown to be highly reliable and their behaviour is more hydraulically predictable than other designs. The hydraulic predictability examination is carried out under two critical network conditions, i.e. link failure and fire loading. Under these conditions, the locations of the critical links and nodes in the maximum entropy designs are more intuitively predictable than in other designs. This beneficial characteristic can assist in the decision making during the design process and operation of the distribution network.

Head Dependent Analysis method is used in the present research. The analysis shows that in general the results are better compared to the previous Demand Driven Analysis results. The entropy-reliability relationship appears stronger, making it easier to evaluate designs with different entropy values. The locations of the critical links and nodes in the maximum entropy designs are more intuitively consistent when the Head Dependent Analysis method is used. Finally, the method also compliments the maximum entropy approach to the layout optimization of water distribution networks. The set of designs that represent the trade off between the cost and reliability can be identified accurately.

Previous entropy-based designs have all been generated using continuous pipe diameters. In this research, the entropy-based design optimization method is applied to generate more realistic designs with discrete pipe diameters. Genetic Algorithms are used in the optimization procedures. The results indicate that the Genetic Algorithms parameters are quite sensitive to the size of the problem and the cost penalty functions also affect the accuracy of the method. However, the results also show that once the sensitivity problem has been solved, the method is efficient and the optimization by Genetic Algorithms always leads to designs with optimum flow directions. This feature may help the search towards finding the global optimum solutions.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor Dr. Tiku Tanyimboh for his support and guidance throughout the course of the PhD program. His invaluable advice and constructive criticisms are greatly appreciated. Dr. Tanyimboh has also played a major role in securing the financial assistance for which I am really grateful. I also would like to thank Professor Richard Burrows, who is acting as my second supervisor, for his advice and guidance throughout the course of the writing up of this thesis.

My sincere appreciation goes to all the academic and non academic staffs of the Department of Engineering at the University of Liverpool who contributed directly or indirectly to this study. In particular, I would like to thank Mr. John Wilkinson for the technical advice I received throughout the development of the computer program in this study. I also would like to thank my friends, colleagues and former colleagues who have helped and supported me through this long journey of the PhD program such as Dr. P Kalungi, Dr. K Whyte, Dr. I Bay, Dr. M Shareef, Hayuti, Yingshi, Grace, Devina, Maria and Isong. Also, special thanks to Ling Siew Tong for his generous help. My sincere gratitude also goes to my counsellor Dr. Andrew Reeves who has given me valuable advice and guidance throughout difficult times.

Additionally, I would like to thank the British Government for the financial assistance that I received through the Overseas Research Students' Award Scheme. I am also grateful for the financial support from the University of Liverpool and the Department of Engineering.

My most sincere and heartfelt gratitude goes to my Mum Lilianti and my Dad Yo Tjoan Keng who have given me parental and material support throughout this study and to my brother Yodi and sisters Noviani and Meita Muliani for their support and encouragement. In a special way, I would like to thank my wife Vita for her moral support, encouragement and prayers throughout this PhD study.

Yohan Setiadi

CONTENTS

LIST OF FIGURES

CHAPTER 7

CHAPTER 8

LIST OF TABLES

NOTATION

α	dimensionless conversion factor.
α_e	ratio of the path flow probabilities in a network.
α_{ei}	ratio of the path flow probabilities related to node i.
α_p	parameter associated with the pheromone concentration.
α_n	a constant for node n.
β	parameter that control the importance of the visibility.
β_n	a constant for node n.
ΔH_{max}	maximum shortage in pressure in a distribution network.
Δq_{ij}	change in q in link ij.
Δq_l	change in q in loop l.
$\Delta \tau^k_{m(n)}$	change in pheromone concentration for option $b_{m(n)}$ at cycle k and iteration t.
$\Delta \underline{x}$	change in $\underline{x}$.
ε	price per unit cost of energy.
γ	cost coefficient for pipes.
γ_d	cost per unit length of pipe whose diameter size is d.
μ_j	Lagrangean multipliers.
μ_{mk}	expected rate of failure of segment k in link m.
$\eta_{m(n)}$	visibility or a guiding factor for the search towards options with smaller "local" costs.
ρ	density of water.
ρ_{pe}	pheromone persistence to evaporation.
τ_{mk}	expected rate of repair segment k in link m.
$\tau_{m(n)}(t)$	concentration of pheromone for option $b_{m(n)}$ at iteration t.
a_{jn}	effective number of independent paths to node n through link jn.
a_m	availability or the probability that link m is available.
a_{mk}	availability of segment k in link m.
$b_{(m)n}$	option m at a decision point n in Ant Colony Optimization Algorithms
c_p	pump coefficient.
$c_{m(n)}$	cost associated with option $b_{(m)n}$.

C	total cost of pipes.
C_e	energy cost for water distribution network.
C_{ij}	Hazen-Williams coefficient for pipe *ij*.
C^k	cost of a trial solution at cycle *k*.
d_k	number of paths in which link *k* is used.
D	set of all demand nodes.
D_i	set of all demand nodes supplied by source node *i*.
D_{ij}	internal diameter of pipe *ij*.
D_{ijm}	diameter of segment *m* in pipe *ij*.
D_{max}	maximum pipe diameter.
D_{min}	minimum pipe diameter.
D_D	set of commercially available discrete pipe diameters.
DSR	demand satisfaction ratio.
e	exponent of the diameter in the cost function for pipes.
e_p	pump curve exponent.
E	total dissipated energy of a pipe network.
f_{ij}	friction factor of pipe *ij*.
$\underline{F}(\underline{x})$	vector of the values of simultaneous equations at the point $\underline{x}$.
$< F_j >$	expected value of the *j*th expectation constraint.
FT	failure tolerance of water distribution networks.
g	acceleration due to gravity.
h_f	head loss across fitting.
h_{ij}	head loss in link *ij*.
h_{ijm}	head loss in segment *m* of link *ij*.
h_{ijr}	head loss in link *ij* for the *r*th flow regime.
h_p	known head loss for path *p*.
h_{pr}	head loss in path *p* for the *r*th demand pattern.
H_i	total heads at nodes *i*.
$H_{max,n}$	maximum desirable head at node *n*.
$H_{min,n}$	minimum desirable head at node *n*.
H_o	shutoff head of a pump.
H_p	head difference across a pump.
H_{prv}	setting of a PRV.
H_s	total head at source node *s*.

H_n^{min}	nodal head at node *n* below which there would be no outflow.
H_s^{min}	head at the source above which outflow just begins at any node in the network.
H_n^{des}	desired head at node *n*.
H_s^{des}	head at the source above which all the demands would be fully satisfied.
I	set of all supply nodes.
I_j	set of all source nodes supplying node *j*.
IJ	set of all the links in a network.
IJ_l	set of all links in loop *l*.
IJ_p	set of all links in path *p*.
J	Jacobian.
k	cycle number.
K	arbitrary positive constant.
K_{ij}	actual pipe resistance for pipe *ij*.
$K_{ij}^{'}$	modified pipe resistance for pipe *ij*.
K_f	coefficient of fitting.
l_{ij}	setof all loops sharing link *ij*.
L_e	equivalent length of a fitting.
L_{ij}	length of link *ij*.
L_{ij}^d	length of pipe *ij* whose diameter size is *d*.
L_{ijm}	length of segment *m* in link *ij*.
n_{ij}	Manning's coefficient for pipe *ij*.
n_n	head-outflow coefficient.
nl_{jn}	total number of links in all the paths supplying node *n* that use link *jn*.
np_{jn}	number of paths to node *n* that use link *jn*.
N	number of outcomes or events.
NC	number of constraints.
ND_n	set of all nodes immediately downstream of, and including any external outflows at, node *n*.
ND_{nr}	set of all nodes downstream of node *n* for the *r*th flow regime.
$ND_n^{\sim}$	set of nodes or links outflows at *n* which are part of a loop.

N_{ij}	number of segments in link *ij*.
NJ	number of expectation contraints.
NL	number of links in a network.
NLP	number of loops in a network.
NN	number of nodes in a network
NP	number of paths whose head losses are known.
NP_{ij}	number of paths from source *i* to demand node *j*.
NR	number of demand patterns.
NS	number of sources in a network.
NU_n	set of all the nodes or links immediately upstream of node *n*.
NU_{nr}	set of nodes upstream of node *n* for the *r*th flow regime.
o_i	*i*th outcome or event.
p_i	probability of *i*th outcome or event.
p_{ij}	probability of the event identified by *ij*.
p_{0n}	fraction of total flow at node *n* provided by q_{0n}.
p_{n0}	fraction of total flow at node *n* satisfying q_{n0}.
$p_{p,ij}$	probability that path flow $q_{p,ij}$, supplied by source node *i*, reaches demand node *j*
$p(0)$	probability that all pipes in a network is available.
$p(m)$	probability that only link *m* is unavailable.
P_{0n}	probability that flow in a network enters the network from source node *n*.
$p_{m(n)}(k,t)$	is the probability that option $b_{m(n)}$ is chosen at cycle *k* and iteration *t*.
P_{n0}	probability that flow in a network leaves the network at demand node *n*.
P_n	probability of flow arriving at node *n*.
P_{pher}	pheromone penalty factor.
PC	penalty cost multiplier.
PR	pheromone reward factor.
q_{0n}	external inflow at node *n*.
$\bar{q}_{i0}$	average demand at node *i*.
q_{ij}	flow in pipe *ij*.
q_{ijr}	pipe flow rate in link *ij* associated with the *r*th demand pattern.
$q_{jk,i}$	flow in link *jk* in the sub-network SNK_i.

q_n	external inflow or outflow at node *n*.
$q_{n,r}$	external inflow or outflow at node *n* for the *r*th pattern.
q_{n0}	external outflow at node *n*.
q_n^{avl}	actual outflow delivered by a network.
q_n^{req}	required outflow or demand.
$q_{p,ij}$	path flow from source node *i* to demand node *j*.
Q_0	sum of network link flows.
Q_n	sum of link flows at node *n*.
$Q_n^{\sim}$	total link inflows and link outflows at node *n*.
Q_p	flow delivered by a pump.
R	network reliability.
R_n	nodal reliability.
R_{ni}	nodal reliability of node *i*.
R_{sw}	system reliability as the weighted mean of all nodal reliabilities.
S	entropy of a finite scheme.
S_n	entropy of node *n*.
S'_n	is the modified entropy for node *n*.
S_0^d	entropy of the distribution demands.
S^i	network entropy based on inflows.
S_n^i	entropy of node *n* associated with the inflows at the node.
S^o	network entropy based on outflows.
S_n^o	entropy of node *n* associated with the outflows at the node.
S^p	network entropy based on path flows.
S_i^p	conditional entropy of path flows for source node *i*.
S_0^s	entropy of the distribution source supplies.
$S_{\min}$	minimum desired entropy value.
$S_n^{\sim}$	entropy of node *n* considering both nodal inflows and outflows.
SNK_i	sub-network connected to source node *i*.
t	iteration number.
t_{nj}	transmissivity of entropy or redundancy from node *j* to node *n*
$T(0)$	(= T_0) total supply or demand.
$T^*(m)$	maximum total outflow when link *m* is unavailable.

T_n	total flow reaching or leaving node *n*.
u_m	unavailability of link *m*.
v_{ij}	mean velocity of the flow in pipe *ij*.
v_{min}	minimum velocity.
v_{max}	maximum velocity.
x	a general variable.
Z_j	elevation of demand node *j*.

CHAPTER 1 INTRODUCTION

1.1 COMPLEXITY IN THE DESIGN OF OPTIMUM WATER DISTRIBUTION NETWORKS

Water distribution networks form major parts of water supply systems. They are required to deliver water at specific pressure and quantity every day and, therefore, any disruptions that may affect the service should be kept to a minimum. The design problem of water distribution networks is often viewed as an optimization problem with the capital cost of the network to be minimised. This approach, which is usually a trial-and-error process, puts emphasis on short term savings gained from the reduction of the construction cost. It does not take much consideration of the performance of the network under many different normal and critical operating conditions, which will have considerable influence on the operation, maintenance and repair costs of the network. Furthermore, there is no guarantee that the resulting design is a minimum cost solution (Mays, 1989).

During its design lifetime, a water distribution network is overshadowed by many uncertainties associated with its operation which often have adverse effects on its performance. Although the advancement in technology has provided means for engineers to deal with the operational and maintenance problems much more easily, the solutions to the problems could nonetheless be expensive and do not justify the savings gained from the reduction of the design cost. As a result, many researchers and engineers alike try to incorporate some kind of performance measure in the design optimization procedure to gauge how well the network would behave under abnormal conditions. The problem of designing optimum water distribution networks is, in itself, a very complex problem due to the large number of design components and their interactions that have to be accounted for. The introduction of performance measure in the optimization process only adds to the complexity of the problem. It turns the design problem into a multi-objective optimization problem with the objective functions usually in conflict with one another.

Other issue is the choice of which performance measure is best to use in the optimization procedures. Many researchers agree that reliability is a good performance measure for water distribution systems. However, many researchers also agree that accurate computation of the reliability values is highly expensive (Provan and Ball, 1983; Mays, 1989). On top of that, there is no comprehensive and generally acceptable measure of reliability of water distribution systems currently available either in the industry or in the research community (Mays, 1989, 2000; Tanyimboh, 2003). These issues complicate the design problem even further. Consequently, although some researchers have tried to incorporate the reliability measure explicitly into the optimization models (Fujiwara and De Silva, 1990; Su et al., 1987), there is no such model with general application has been developed at present.

In the attempt to overcome the above issues, several researchers have proposed indirect or surrogate measures of reliability. Entropy is one such measure. The strength of entropy stems from its ease of computation and incorporation into the design optimization procedure and early research shows that there is apparent correlation between entropy and reliability. The idea of using entropy in obtaining reliable water distribution systems was introduced by Awumah et al. (1989) who proposed the first entropy function for water distribution systems. The function was later refined by Tanyimboh (1993) and Tanyimboh and Templeman (1993a) who also have shown that highly reliable design can be obtained by maximizing the entropy value of the system (Tanyimboh and Templeman, 1993b). Tanyimboh (1993) and Tanyimboh and Templeman (1993c) also proposed a simple algorithm to obtain maximum entropy flows in distribution systems with one source supply without the need for complicated mathematical computations. Their algorithm was later generalised for multi-source multi-demand networks by Yassin-Kassab et al. (1999). The use of entropy in the design of optimum (in terms of cost and reliability) water distribution systems has also been explored further by Tanyimboh and Templeman (2000) and Tanyimboh and Sheahan (2002).

The present research looks into the suitability of entropy as a surrogate measure of water distribution system reliability at a deeper level. Details of the research scope are presented in the next section.

1.2 SCOPE OF THE PRESENT RESEARCH

The outcomes of the design optimization of water distribution systems are influenced by many factors. These factors will in turn influence the entropy and reliability values as well as their apparent relationship. The present research examines the strength of this apparent relationship between entropy and reliability and investigates issues that may have influence on the relationship. Although the study in the present research does not address all the issues that may affect the relationship, it is by far the most comprehensive study that has been carried out to date. The issues investigated in this study are: possible influence of layouts, flow directions, cost functions and modelling errors.

Another aspect of the research is the use of head dependent analysis method in the analysis of networks under normal and abnormal conditions in order to obtain the reliability values. This method has been suggested to be superior to the demand driven analysis, especially for analysing networks in critical conditions (Ackley et al., 2001; Tanyimboh et al., 2003). The analysis of entropy-based designs of water distribution networks has been carried out using the demand driven analysis in the past (Tanyimboh, 1993; Tanyimboh and Templeman, 1993b, 2000; Tanyimboh and Sheahan , 2002; Tanyimboh et al., 2002). The use of the head dependent analysis in the present study is intended to provide more evidence that the apparent relationship between entropy and reliability is not attributed to the method of analysis used and that the relationship holds true in general.

Finally, the issue of the design of water distribution networks using discrete pipe diameters is addressed. This design approach produces more realistic results since the pipe diameters can be selected from the set of discrete pipe diameters available in practice. However, in a discussion of Quindry et al. (1981), Templeman (1982) has pointed out that the problem of obtaining optimum designs of water distribution network using members selected from a discrete set is NP-hard. He stated further that the NP-hardness of the problem means that finding an optimum design using a rigorous algorithm is practically impossible. Recently, stochastic search methods have been applied to water distribution systems optimization with discrete pipe diameters with satisfactory results (Simpson et al., 1994; Loganathan et al., 1995; Savic and Walters, 1997; Maier et al., 2003). In the present study, the entropy-based design optimization using Genetic Algorithms, which is one of the stochastic methods, is investigated.

The analyses in the present research are carried out on hypothetical networks obtained from the literature. The study is also limited to gravity networks with pipes as the only component. Therefore, the words link(s) and pipe(s) are interchangeable throughout this thesis, unless stated otherwise. Some information regarding pumps, valves and storage tanks is given only very briefly. The analysis of the networks is also limited to steady state analysis, in which it is carried out in a very short period of time with constant demand values. The extended period analysis, which is done over a longer period, is not covered in this thesis. Also, only one reliability function proposed by Tanyimboh (1993) is used in establishing the correlation between entropy and reliability. The function is chosen mainly due to its simplicity and ease of interpretation and it is considered sufficient for the present research.

1.3 OBJECTIVES OF THE PRESENT RESEARCH

The objectives of the present research are as follows.

1. To show that the correlation between entropy and reliability of water distribution systems is strong and is not sensitive to the changes in factors that may influence the values of entropy and reliability individually.
2. To use the head dependent analysis method in the analysis of the entropy-reliability relationship.
3. To illustrate beneficial properties of the maximum-entropy-based designs of water distribution networks.
4. To demonstrate the maximum entropy-based design optimization of water distribution networks with discrete pipe diameters using the revolutionary Genetic Algorithms.

1.4 LAYOUT OF THESIS

The main literature review and background study in this thesis are arranged in the next four chapters. Chapter 2 provides introduction to Shannon's entropy function (Shannon, 1948), on which the entropy of water distribution network is based. Brief description of the maximum entropy formalism of Jaynes (1957) is also given in Chapter 2 together with some examples of its applications in Civil Engineering. The entropy function for water distribution networks is presented in Chapter 3. Methods for calculating maximum entropy flows in single and multiple source networks are also detailed in Chapter 3. In Chapter 4, the problem of analysis of water distribution networks is reviewed. Chapter 5 gives brief review on the performance measure and the problem of obtaining minimum cost design of water distribution networks.

The main studies carried out in the present research are presented in the remainder of the chapters as follows. In Chapter 6, results of the sensitivity studies on the relationship between entropy and reliability are presented and discussed. Chapter 7 looks at the beneficial properties of maximum entropy designs of water distribution network in which the network's behaviour is more hydraulically predictable when subjected to critical conditions, such as link failures and fire loadings, in comparison to other designs. The possibility of maximum entropy-based design approach for layout optimization is also explored in Chapter 7. Throughout Chapters 6 and 7, head dependent analysis method is used and the results are analysed in comparison to the previous demand driven analysis results. In Chapter 8, the design of optimum water distribution networks with maximum entropy constraint and discrete pipe diameters is investigated. Finally, the summary and conclusions of the present research are presented in Chapter 9 together with some suggestions for future studies.

CHAPTER 2 THE MAXIMUM ENTROPY FORMALISM

2.1 INTRODUCTION

In this chapter Shannon's entropy and the maximum entropy formalism of Jaynes (1957) are presented. Some of the materials are summarised from Templeman and Li (1985), Li (1987), Tanyimboh (1993) and Yassin-Kassab (1998).

The concept of *entropy* was first introduced by Clausius in classical thermodynamics. It is concerned with the macroscopic states of matter, which can be observed experimentally. Clausius' entropy is non-probabilistic in nature and is known as the *classical entropy.*

Entropy evolved further in statistical mechanics, in which it was used in a probabilistic sense. Boltzmann was the first to emphasize the probabilistic meaning of the entropy. He noticed that the entropy of a physical system can be considered as a measure of 'disorder' in the system and that in a system with many degrees of freedom, the number measuring the disorder of the system also measures the 'uncertainty' about individual micro-states. Boltzmann's entropy is known as the *statistical entropy.*

Shannon (1948) presented entropy in a different context other than thermodynamics. His entropy is related to information theory, which is therefore referred to as the *informational entropy.* It measures the amount of information or uncertainty in a probability distribution quantitatively. This measure has wider applicability than the statistical entropy. Due to its direct relevance to the present research, Shannon's entropy will be described shortly in more detail.

Jaynes (1957) proposed a groundbreaking use of Shannon's measure in generating a probability distribution that would have the greatest amount of information or entropy. Before Jaynes' work, Shannon's entropy was merely a measure of uncertainty or information in a probability distribution in which all the probabilities are known. In the case of insufficient information, some assumptions must be made which may lead to biased results. Jaynes' groundbreaking work is termed the *maximum entropy formalism* and ensures that no arbitrary assumptions are introduced in obtaining the probability distribution to fit the available data.

The entropy in water distribution systems is based on Shannon's entropy measure. It is therefore necessary to present Shannon's entropy before concentrating on the application of the maximum entropy formalism in water distribution systems. Shannon's informational entropy is described next along with its properties. Jaynes' maximum entropy formalism is then presented along with its applications with emphasis on the applications in civil engineering.

2.2 SHANNON'S INFORMATIONAL ENTROPY

Uncertainty in a probabilistic scheme cannot be avoided. The degree of uncertainty often differs greatly between one scheme to another. Consider, for example, the probabilistic experiment of tossing a fair-faced coin. The probability of obtaining a head is equal to the probability of obtaining a tail, which is 0.5. Now, suppose for the second trial the coin is transformed in such a way that the probability of obtaining a tail is equal to 0.99, and the probability of obtaining a head is 0.01. Obviously the uncertainty in the second experiment is very much reduced. If the experiment is then carried out using a fair-faced dice (cube), the probability of obtaining any face of a dice is 1/6, which means that the uncertainty of the outcome is even greater than the first experiment.

In the field of information theory, Shannon tried to measure quantitatively the degree of uncertainty in any finite probability scheme. Before Shannon's measure is presented in this section, a finite probability scheme is first defined.

2.2.1 FINITE PROBABILITY SCHEME

In probability theory, a *complete system* is obtained when the set of *events* or *outcomes* is *exhaustive*, which mean that one and only one of the outcomes must occur at each trial. The events or outcomes are said to be *mutually exclusive* since only one of them can occur at each trial. The events of such a complete system together with their corresponding probabilities form a *finite scheme* (Khinchin, 1953).

Let us denote the events or outcomes of a finite scheme by o_i and the corresponding probability by p_i, $i = 1,...,N$, where N is the number of events or outcomes. The resulting finite scheme O is given by:

$$O = (o_i, p_i) \quad i = 1,...,N \tag{2.1}$$

The probabilities of a finite scheme are non-negative and satisfy the normality condition, i.e.,

$$p_i \geq 0, \quad \forall i \tag{2.2}$$

and,

$$\sum_{i=1}^{N} p_i = 1 \tag{2.3}$$

2.2.2 ENTROPY OF FINITE SCHEMES

From the experiments of tossing a fair-faced coin and a fair-faced dice, it was found that the corresponding probabilities are $\left(\frac{1}{2}, \frac{1}{2}\right)$ and $\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$, respectively. It is clear that there is more uncertainty associated with the latter experiment. In his attempt to measure how much information or uncertainty is conveyed by different finite probability schemes, Shannon put forward the following function

$$S = -K \sum_{i=1}^{N} p_i \log p_i \tag{2.4}$$

in which S is the entropy or the amount of uncertainty in the probability distribution; K is an arbitrary positive constant; and the logarithms can take any suitable fixed base. Also, it is defined that 0log0 = 0 to ensure continuity for all the probabilities (see, for example Jones, 1979, and Khinchin, 1953). In this thesis, the value of K is set to unity and the natural logarithms are used throughout. It is axiomatic that the probabilities p_i, $i = 1,\ldots,N$, which represent a finite scheme are non-negative, exhaustive, mutually exclusive and satisfy the normality condition of Equation (2.3). Shannon's entropy, Equation (2.4), is a measure of uncertainty, or conversely, a measure of information depending on the measurement being taken. Once an experiment has been done, the actual outcomes are known and the uncertainty about the results of the experiment is removed. Therefore, the information gained from the experiment is equal to the amount of uncertainty removed. See, for example, Guiasu (1977) and Kapur (1989).

2.2.3 SOME PROPERTIES OF SHANNON'S ENTROPY

Presented below are some properties of Shannon's entropy. These properties are to be expected from a reasonable measure of uncertainty and are presented here without proof. For more details and other properties, which are mostly mathematical derivations, the interested reader may consult Khinchin (1953), Guiasu (1977), Kapur and Kesavan (1987) and Kapur (1989).

1. $S_N(p_1,\ldots,p_N) \geq 0$

   The function takes the value of zero if, and only if, one of the probabilities is unity and the rest is zero. Such a scheme obviously contains no uncertainty.

2. $S_N(p_1,\ldots,p_N) = S_{N+1}(p_1,\ldots,p_N,0)$

   This property is to be expected since an impossible outcome does not affect the amount of uncertainty in any scheme.

3. $S_N(p_1,\ldots,p_N) \leq S_N(\frac{1}{N},\ldots,\frac{1}{N})$

   The equality holds if, and only if, all the probabilities are equal, which will produce the maximum value of the function S. This agrees with one's expectation.

4. The entropy function S is continuous and is invariant with respect to the positional changes in the $p_i, \forall i$, i.e. S is a symmetric function of its arguments.

5. When all the probabilities are equal, i.e. $p_i = \frac{1}{N}$, $\forall i$, then S is a monotonic increasing function of the number of outcomes N. If these probabilities are substituted in Equation (2.4), the maximum value of S/K is $\ln N$.

6. S is a concave function and, therefore, its maximum value of $\ln N$ is a global maximum.
7. The joint entropy of two *mutually dependent* schemes is the entropy of one scheme plus the conditional entropy of the other, i.e.

$$S(O_1O_2) = S(O_1) + S(O_2 \mid O_1) \tag{2.5}$$

Also,

$$S(O_1) + S(O_2 \mid O_1) = S(O_2) + S(O_1 \mid O_2) \tag{2.6}$$

in which $S(O_1O_2)$ is the joint entropy of two mutually dependent schemes, O_1 and O_2, whose entropies are $S(O_1)$ and $S(O_2)$ respectively. $S(O_2|O_1)$ and $S(O_1|O_2)$ are, respectively, the conditional entropy of scheme O_2 provided that O_1 has occurred and vice versa. The proof of Equation (2.5) can be derived using a chain rule (see, for example, Cover and Thomas, 1991). The general form of Equation (2.5) for any number of finite schemes is (Tanyimboh, 1993)

$$S(O_1O_2...O_M) = S(O_1) + S(O_2 \mid O_1) + ... + S(O_m \mid O_1O_2...O_{m-1}) + ...$$

$$+ S(O_M \mid O_1O_2...O_{M-1}) \text{ , } \quad M = 2,3,... \text{ ; } 2 < m \in Z^+ < M \tag{2.7}$$

in which $S(O_1O_2...O_M)$ is the joint entropy of M finite schemes; Z^+ represents the set {0, 1, 2, 3, ...}.

In the special case where O_1 and O_2 are *mutually independent*, the outcomes in one scheme have no effect on the occurrence of the other. Therefore, the joint entropy of two mutually independent schemes is the sum of their separate entropies, i.e.

$$S(O_1O_2) = S(O_2O_1) = S(O_1) + S(O_2) \tag{2.8}$$

It was found that the only measure of uncertainty that satisfies all the above properties is in the form of Equation (2.4) (see, for example, Khinchin, 1953). This uniqueness of Shannon's entropy is stated next as a theorem.

2.2.3.1 THE UNIQUENESS THEOREM

Let S be a function defined for any integer N and for all values of p_i, $i = 1,...,N$, such that p_i satisfies the non-negativity and the normality conditions of Equations (2.2) and (2.3) respectively. Suppose that S is a continuous function with respect to all its arguments and this function satisfies the following three of the basic properties for the measure of uncertainty (Kinchin, 1953), i.e.

1. For a given N and for $\sum_{i=1}^{N} p_i = 1$, the function takes its maximum value for $p_i = 1/N$, $\forall i$.
2. $S(O_1 O_2) = S(O_1) + S(O_2|O_1)$
3. $S_N(p_1,...,p_N) = S_{N+1}(p_1,...,p_N,0)$

The only possible function that satisfies the above properties is equal to Shannon's entropy function, which is

$$S = -K\sum_{i=1}^{N} p_i \ln p_i \tag{2.4}$$

This well-established theorem shows that Shannon's entropy for a finite scheme is the only measure that has general properties necessary for a measure of uncertainty or information. The proof of this theorem is outside the scope of the present research. The interested reader could refer to Khinchin (1953), Jones (1979), etc.

2.3 THE MAXIMUM ENTROPY FORMALISM

Laplace's *principle of insufficient reason* states that all outcomes of a finite probability scheme should be considered to be equally likely if there is no reason to think otherwise. Therefore, the uniform distribution should be adopted whenever there is no information based on which a different distribution may be selected. This uniform distribution will in turn produce the maximum value of Shannon's entropy as explained in Section 2.2.

However, some information about the probability scheme is often available. In these situations, the principle of insufficient reason has no means of dealing with the available information and the uniform distribution does not fit due to the presence of the additional information. For example, consider an observable probabilistic process in which a discrete random variable x can take any discrete value x_i, $i = 1,...,N$, with the corresponding probabilities $p(x_i) = p_i$, $\forall i$. Suppose also that some information is available in the form of

$$\sum_{i=1}^{N} p_i F_{ji}(x) = < F_j >, \ j = 1,...,NJ \tag{2.9}$$

in which $F_{ji}(x)$ and $< F_j >$, $\forall j$ and $\forall i$, are known. If the number of equations NJ together with the normality condition of Equation (2.3), is less than the number of possible x values, i.e. $NJ+1<N$, clearly an infinite number of distributions can satisfy Equations (2.3) and (2.9). It is therefore virtually impossible to find the unique distribution that best represents the above scheme by using the principle of insufficient reason alone.

Jaynes (1957) recognized that every probability distribution that fits the available information has different value of Shannon's entropy measure of uncertainty. He postulated that the distribution that has the maximum value of entropy within the limitation of the available information must correspond to the maximum uncertainty and is, therefore, noncommittal to the missing information, unbiased and best describes the finite scheme subject to the information available. Jaynes stated:

> In making inference on the basis of partial information we must use that probability distribution which has maximum entropy subject to whatever is known. This is the only unbiased assignment we can make; to use any other would amount to arbitrary assumption of information which by hypothesis we do not have.

The above method of inference is known as the *maximum entropy formalism* and can be mathematically represented as a maximization of Shannon's entropy function as detailed below.

Problem 1

$$\underset{\forall p_i}{Maximize} \quad S/K = -\sum_{i=1}^{N} p_i \ln p_i \tag{2.4}$$

Subject to:

$$p_i \geq 0, \quad \forall i \tag{2.2}$$

$$\sum_{i=1}^{N} p_i = 1 \tag{2.3}$$

$$\sum_{i=1}^{N} p_i F_{ji}(x) = \langle F_j \rangle, \quad \forall j \tag{2.9}$$

The analytical solution to Problem 1 can be found by examining the stationary of its Lagrangean. Templeman and Li (1985) showed that Problem 1 is a convex programming problem and, as such, there is a unique solution that corresponds to the global maximum probability distribution which is called the *maximum entropy distribution.* The solution to Problem 1 is (see e.g. Li, 1987)

$$p_i = \frac{\exp\left[\sum_{j=1}^{NJ} \mu_j F_{ji}\right]}{\sum_{i=1}^{N} \exp\left[\sum_{j=1}^{NJ} \mu_j F_{ji}\right]}, \quad i = 1, \ldots, N \tag{2.10}$$

in which μ_j, $\forall j$ are the Lagrangean multipliers. Rather than solving the above problem directly, Templeman and Li (1985) have shown that the Lagrangean multipliers may be calculated conveniently using its dual form, which is an unconstrained optimisation problem. Problem 1 is the classical maximum entropy problem. It may not always be possible or easy to formulate the problem of inferring the least biased probabilities in the form of Problem 1 (Tanyimboh, 1993).

If the entropy is maximized subject to the normality condition only, the result is $p_i = 1/N$, $\forall i$. This may be seen by examining the stationary of the Lagrangean of this special case. This result concurs with Laplace's principle of insufficient reason, which can therefore be seen as a special case of the maximum entropy formalism. Thus, it can be deduced that the maximum entropy distribution has the property of being the most uniform distribution that satisfies the constraints of the system. Any gain in information about the system leads to an extra constraint in the maximum entropy formalism and consequently reduces the entropy value of the system. Conversely, any gain in entropy means loss of information.

2.4 THE CONTINUOUS CASE OF THE MAXIMUM ENTROPY FORMALISM

Problem 1 and its solution are written in terms of discrete random variables. In a situation where a random process is continuous, the maximum entropy formalism still applies (Li, 1987; Tanyimboh, 1993). In general, the theory and the process of obtaining the solution remain unaltered. However, in the continuous case, integrations are used in place of the summations and probability density functions must be used instead of the discrete probabilities. The continuous case of the maximum entropy formalism may therefore be stated as follows (Tanyimboh, 1993).

$$\underset{\forall f(x)}{\mathit{Maximize}} \quad S/K = -\int_a^b f(x)\ln\big(f(x)\big)\, dx \tag{2.11}$$

Subject to

$$\int_a^b f(x)dx = 1 \tag{2.12}$$

$$\int_a^b F_j(x)f(x)dx = \langle F_j \rangle, \quad j = 1, \ldots, NJ \tag{2.13}$$

in which x is a continuous random variable; F_j is a function of x; $< F_j >$ is the expected value of the function F_j; NJ is the number of expectation constraints; and $f(x)$ is the probability density function. The range of the integrals $[a, b]$ may be extended to $[-\infty, +\infty]$.

Tanyimboh (1993) has pointed out that the above integrals may not exist, and the entropy can be negative because the probability density function $f(x)$ can be greater than unity. Also, because the entropy is defined in terms of a probability density rather than a probability, the entropy may not be invariant to a change of variable. Moreover, the limits of negative and positive infinity are strictly not the limits of the discrete entropy whose properties therefore cannot be extended to the continuous case. This research, however, focuses only on discrete probabilities where the random processes are discrete and yield distinct probabilities with finite limits.

2.5 THE MAXIMUM ENTROPY FORMALISM IN CIVIL ENGINEERING

The maximum entropy formalism has been widely used in the areas of science and engineering due to its simplicity and efficiency in generating solutions to a wide range of problems where the available information is incomplete. In this section, however, only applications in the civil engineering area are presented. For other applications in different fields, the interested readers could refer to Jones (1979), Guiasu (1977), Levine and Tribus (1979), Kapur and Kesavan (1987), Templeman and Li (1987, 1989), Kapur (1989). The use of entropy in optimization processes, however, is mentioned in this section due to its close relevance to the present research.

Basu and Templeman (1984) used the maximum entropy formalism in their search to find the most suitable probability distribution to fit the available data. They argued that in most engineering problems, the problems of selecting the probability distributions are based on *ad hoc* selection criteria, which are not always justifiable and may introduce bias into the calculations. They showed that from the investigation using a wide range of different distributions the maximum entropy distribution was the nearest to the actual distribution of the data being examined.

Also, Li (1987) and Templeman and Li (1987, 1989) have used the maximum entropy formalism in optimization processes in an attempt to develop a different method for solving constrained nonlinear programming problems. They looked at the problem in a probabilistic context and incorporated the principle of maximum entropy into the process to improve the convergence towards the optimum solution. This new entropy-based approach was applied by Li and Templeman (1988) to an optimum truss sizing problem and the method was found to be very effective and encouraging to use in more difficult structural optimization problems.

Basu and Templeman (1985) have also used the maximum entropy formalism to estimate the failure probability of a structure, in which they used the maximum entropy probability distribution to represent random loads and strengths in structural reliability analysis. They demonstrated that the entropy-based approach produced a more logical and rigorous method to generate accurate failure probabilities, and therefore, the use of some known analytical distribution to represent such random data is inappropriate. In another application, Munro and Jowitt (1978) used the maximum entropy formalism in decision-making analysis in the ready-mixed concrete production industry. The problem is concerned with making optimal decisions under uncertainty about future orders. They stated that the estimated probability distribution associated with the orders for each mix should take into account all prior knowledge but this knowledge must not affect the decision in obtaining the probability distribution. Their findings indicate that the entropy-based method can produce good decisions in difficult decision-making situations.

In the field of traffic engineering, there has been a considerable amount of applications of entropy in which the primary concern is in the estimation of the origin-destination matrices or the so-called trip matrices from limited data. A typical transportation problem is to minimize the total travelling cost between origins and destinations subject to the available information about the total flows leaving each origin and entering each destination. Erlander (1977) added an entropy constraint to the problem in order to preserve the level of accessibility between all origins and destinations. He stated that a network with a high value of entropy has a high level of accessibility, and vice versa. Van Zuylen and Willumsen (1980) and Bell (1983) have used entropy to estimate the origin-destination trip matrix from traffic counts.

The methods are based on information minimisation and entropy maximisation principles (Van Zuylen and Willumsen, 1980). The interested reader may refer to their papers for more details. Recently, a procedure to enhance the estimation of multi-class trip matrices by means of entropy maximization is proposed (Wong et al., 2005). Solving multi-class rather than single-class problem has the advantage of eliminating any inconsistency in the estimation of different vehicle classes. The case study on the Hong Kong highway network demonstrates the effectiveness of the proposed entropy-based method. The results show that the error in the estimation is much reduced compared to the previously obtained non-entropy-based model. Also, in a series of papers Mountain et al. (1983a, 1983b, 1986a, 1986b) used entropy inference to estimate traffic turning flows at road junctions. It was shown that the entropy-based approach leads directly to the *gravity model* which is well known in roundabout turning flow problems.

In open channel flow studies, Chiu (1987, 1988, 1989, 1991) applied the entropy approach in the modelling of the velocity and shear stress distribution and suspended sediment concentration across the channel. He claimed that the uncertainty associated with the distributions due to the inherent randomness and insufficient information can be overcome by maximizing the entropy of the probability density function of the distributions subject to the constitutive laws for flows in open channels.

2.6 SUMMARY AND CONCLUSIONS

Shannon's entropy function and its properties have been presented in this chapter. The function has been shown to be the only measure that has general properties necessary for a measure of uncertainty. The maximum entropy formalism of Jaynes' has also been described in this chapter. It is capable of producing the most unbiased probability distribution by maximising the entropy function of Shannon's subject to the available information. Finally, some applications of entropy in Civil Engineering were presented. Although the list of applications is clearly incomplete, it serves to highlight the use of entropy in generating solutions to a wide range of problems where the available information is not complete or where probability distribution is required but cannot be determined by analysis.

CHAPTER 3 ENTROPY FUNCTION FOR WATER DISTRIBUTION NETWORKS AND ITS APPLICATIONS

3.1 INTRODUCTION

In recent years, some research has been carried out to study the use of entropy in water distribution networks. There are several areas in which entropy has been applied; first, as a possible surrogate measure for the reliability of water networks; second, in estimating the most likely pipe flow rates in looped water networks in which the available data is insufficient to uniquely determine the pipe flow distribution; third, in calibrating water distribution systems to find the most likely pipe characteristics in the system, where the roughness coefficients of the pipes have been lost or have changed with time; and fourth, in the design optimization of water distribution networks. The first application usually goes hand in hand with the last.

In the case where there is not enough information to determine the pipe flow rates in a looped water network, for example in a buried old distribution network where the pipe characteristics have been lost or have changed over time, physical measurements are possible but they would be very expensive and time consuming. Faced with these difficulties, a simple method to quickly estimate the pipe flow rates in such situation would be very useful. Some researchers have proposed an entropy-based method as an answer to this problem. The approach uses the maximum entropy formalism to estimate the most likely flow distribution in the network. The strength of this approach stems from the ease of computation and minimal data requirement. Due to its direct relevance to the present research, this entropy-based approach will be described fully in this chapter.

The present research involves mainly the first and the last of the above mentioned applications of entropy in which it is used as a surrogate measure of reliability in the design optimization of water networks. The main issue that drives the use of surrogate performance measures for water distribution networks lies in the difficulty of determining the reliability itself. Although it has been accepted that reliability is an important issue to consider in design and management of water distribution systems, to date there is no comprehensive and generally acceptable method to calculate the reliability values. Some researchers have proposed several models for calculating the reliability. However, problems arise due to the complexity of each model in dealing with the uncertainties surrounding the network. Such uncertainties are component failures, time between failures, duration of repairs or replacements, sufficiency of pressure, variations in demands and supplies, etc. Water distribution networks are usually designed to cope with these uncertainties to a certain extent, and thus their reliability values reflect how well the networks perform under such a situation. The problem of obtaining the reliability values for water networks is also computationally very expensive and therefore it is often impractical for large networks. A surrogate measure, such as the entropy measure, which is relatively easy to compute and can be easily incorporated directly into the network-optimization design model is very desirable. This application of entropy, as well as the other applications mentioned earlier, will be described briefly later in this chapter. First, the entropy function for water distribution networks is presented.

3.2 WATER DISTRIBUTION NETWORK ENTROPY

The drive that initiates the use of entropy in water distribution networks is its ability to measure flexibility in a system. Yao (1985) and Kumar (1987), for example, used entropy as a measure of flexibility within manufacturing systems in which flexibility is referred to as the ability of the whole production system to overcome the failure of one of the units in the system without significantly affecting the production capacity of the whole system. By looking at the similarity between the manufacturing systems and water distribution networks, Awumah et al. (1989) pioneered the use of entropy as a measure of flexibility or redundancy in looped water networks. Using the flow rates in a probabilistic way as required by Shannon's entropy to obtain the network entropy values, they showed that the performance of water distribution networks can be measured comparatively. Their entropy-based function, however, violates two of the basic properties of Shannon's entropy (Tanyimboh, 1993).

Tanyimboh (1993) and Tanyimboh and Templeman (1993a) were the first to propose the correct function of entropy for water distribution networks. Their entropy functions are rigorous and form the basis of the network entropy analysis throughout this thesis. In a series of papers, Tanyimboh and Templeman (1993a, 1993b, 1993c) explored the feasibility of using network entropy to estimate the most likely flows in distribution networks where the required data to uniquely determine the pipe flow rates is unavailable. Their algorithm for maximizing the entropy of single-source networks was generalised for multi-source multi-demand networks by Yassin-Kassab et al. (1999). The possibility of using the entropy to design water distribution networks has also been explored by Tanyimboh and Templeman (1993b, 2000), Tanyimboh and Sheahan (2002) and Kalungi (2003).

In this section, the flow entropy functions for water distribution networks by Awumah et al. (1989, 1990, 1991) are presented next. The flow entropy functions developed by Tanyimboh and Templeman (1993a, 1993b, 1993c) then follow together with their detail derivations due to their direct relevance to the present research. Also, the path entropy functions of Yassin-Kassab et al. (1999) will be presented as alternative but equivalent approach to the flow entropy functions of Tanyimboh and Templeman. The maximum entropy flows in a network calculated by Tanyimboh and Templeman (1993a) are then described followed by the path-based approach for calculating maximum entropy flows in single (Tanyimboh and Templeman, 1993c) and multiple (Yassin-Kassab et al., 1999) source networks in the next section. Finally, brief outline of the previous applications of entropy in water distribution networks is presented.

3.2.1 FLOW ENTROPY FUNCTION OF AWUMAH, GOULTER AND BHATT

In trying to cast pipe flow rates in a water distribution network in a probabilistic form as required by Shannon's entropy, Awumah et al. (1989, 1990, 1991) started investigating the flows in links incident at node n. They argued that the probability quantities in Shannon's entropy of Equation (2.4) may be regarded as the fraction of total flows into node n carried by each link incident at that node. Therefore the following function could be used as an entropy measure of node n in a network after setting K in Equation (2.4) to unity.

$$S_n = -\sum_{j \in NU_n} \frac{q_{jn}}{Q_n} \ln \frac{q_{jn}}{Q_n}, \quad \forall n \tag{3.1}$$

in which S_n is the entropy of node n, $\forall n$; NU_n, $\forall n$ represents the set of nodes on the upstream ends of links incident on node n; q_{jn} is the flow in link jn, $\forall j \in NU_n$; and Q_n is the sum of link flows entering node n, i.e.,

$$Q_n = \sum_{j \in NU_n} q_{jn}, \quad \forall n \tag{3.2}$$

Since there is no real need to assign a value to K in the formal expression (Equation 2.4), it is therefore assumed to be unity in the rest of this thesis and, as such, does not appear in most of the equations which follow.

It has been mentioned previously that Awumah et al. (1989, 1990, 1991) looked at the entropy measure from redundancy view point. They used Equation (3.1) to measure the redundancy of node n and they expanded the equation to the whole network (see Awumah et al., 1990, 1991, and Awumah and Goulter, 1992). They argued that in assessing the overall network performance, the relative importance of a link to the total flow is the important parameter instead of the relative importance of a link to the local flow. Therefore, they suggested that q_{jn}/Q_n in Equation (3.1) be replaced by q_{jn}/Q_0 and the resulting equation is

$$S = -\sum_{n=1}^{NN} \sum_{j \in NU_n} \frac{q_{jn}}{Q_0} \ln \frac{q_{jn}}{Q_0} \tag{3.3}$$

in which S is the entropy or redundancy of the whole network; NN is the number of nodes in the network; and Q_0 is the sum of all link flows in the network, i.e.,

$$Q_0 = \sum_{ij \in IJ} q_{ij} \tag{3.4}$$

in which IJ is the set of all the links in the network.

Since the value of Q_0 is the sum of all the link flows in the network, as such there is double counting of some of the link flows when the pipes are connected in series and therefore the sum of all the ratio q_{jn}/Q_0, $\forall j \in NU_n$, $\forall n$ will not add up to unity. Hence, Equation (3.3) violates one of the fundamental requirements of Shannon's entropy.

Despite the above violation, Awumah et al. (1990, 1991) and Awumah and Goulter (1992) transformed Equation (3.3) by substituting the following equation.

$$\frac{q_{jn}}{Q_0} = \frac{q_{jn}}{Q_n}\frac{Q_n}{Q_0}$$

The transformed equation for the entropy or redundancy of the network is therefore

$$S = \sum_{n=1}^{NN} \frac{Q_n}{Q_0} S_n - \sum_{n=1}^{NN} \frac{Q_n}{Q_0} \ln \frac{Q_n}{Q_0} \tag{3.5}$$

in which S_n is the entropy of node n and is given by Equation (3.1).

Awumah et al. (1990) realised that the node entropy measure given by Equation (3.1) treats the node in isolation without considering the connectivity of the node to the rest of the network. Water from the source to a demand node may travel through several paths and these paths may have some links in common. To account for these alternate paths, Awumah et al. (1990) introduced the following function for nodal entropy.

$$S_n = -\sum_{j \in NU_n} \frac{q_{jn}}{Q_n} \ln \frac{q_{jn}}{a_{jn} Q_n} \quad , \quad \forall n \tag{3.6}$$

in which a_{jn}, $\forall j \in NU_n$, $\forall n$, is the effective number of independent paths to node n through link jn, and its value is given by

$$a_{jn} = np_{jn}\left[1-\frac{\sum_{k=1}^{nl_{jn}}(d_k-1)}{\sum_{k=1}^{nl_{jn}} d_k}\right], \quad \forall n,\ \forall j \in NU_n \tag{3.7}$$

in which np_{jn}, $\forall n$, $\forall j \in NU_n$, is the number of dependent or independent paths to node n through link jn; nl_{jn}, $\forall n$, $\forall j \in NU_n$, is the number of links in the np_{jn} paths; d_k is the number of paths in which link k is used. Awumah et al. (1990) provides more details on Equations (3.6) and (3.7). Tanyimboh (1993) has pointed out that the calculation of the path parameter a_{jn} can be computationally expensive for large networks since it relies on path enumeration.

In an attempt to account for the interaction between adjacent nodes in a network, Awumah et al. (1991) used the following equation.

$$S'_n = S_n + \sum_{j \in NU_n} t_{nj} S'_j, \quad \forall n \tag{3.8}$$

in which S'_n is the modified entropy for node n; S'_j is the entropy of node j, $\forall j \in NU_n$. Therefore, to calculate the modified entropy of any node, the modified entropy of its upstream nodes must all be calculated first. Also, t_{nj} is termed the transmissivity of entropy or redundancy from node j to node n, which is approximated by the ratio of the flow entering node n from j to the total flow entering node j, i.e.,

$$t_{nj} = \frac{q_{jn}}{Qj}, \quad \forall n, \forall j \in NU_n \tag{3.9}$$

If S'_n is used in Equation (3.5) instead of S_n, the modified network entropy therefore becomes

$$S = \sum_{n=1}^{NN} \frac{Q_n}{Q_0} S'_n - \sum_{n=1}^{NN} \frac{Q_n}{Q_0} \ln \frac{Q_n}{Q_0} \tag{3.10}$$

Awumah et al. (1991) observed that S'_n, $\forall n$, give higher values of entropy for the network than S_n, $\forall n$. However, there is little evidence that the former relate better to the conditions in water distribution networks.

Finally, it may be noted that all nodal entropy functions presented so far have been defined in terms of link inflows only. No consideration was given to the outflow links which may become inflow links to the node being considered in the event of a link failure. Obviously, this may only happen to outflow links which are part of a loop since flow reversal cannot occur in a link which does not belong to a loop. To allow for such situation, Awumah et al. (1990) proposed the following equation for node entropy in which all the links incident at the node are included rather than simply those which supply the node under normal condition.

$$S_n^{\sim} = -\sum_{j \in NU_n} \frac{q_{jn}}{Q_n^{\sim}} \ln \frac{q_{jn}}{a_{jn} Q_n^{\sim}} - \sum_{k \in ND_n^{\sim}} \frac{q_{nk}}{Q_n^{\sim}} \ln \frac{q_{nk}}{a_{nk} Q_n^{\sim}}, \quad \forall n \tag{3.11}$$

in which $S_n^{\sim}$ is the entropy of node n, $\forall n$; $ND_n^{\sim}$ is the set of nodes immediately downstream of node n, $\forall n$, which belong to a loop containing node n; $Q_n^{\sim}$ is the total flow entering and leaving node n from the set of nodes contained in NU_n and to the set of nodes contained in $ND_n^{\sim}$ respectively, i.e.,

$$Q_n^{\sim} = \sum_{j \in NU_n} q_{jn} + \sum_{k \in ND_n^{\sim}} q_{nk}, \quad \forall n \tag{3.12}$$

Equation (3.11) also can be substituted for S_n in Equation (3.5) to obtain the entropy of the network. It has been mentioned earlier that Equation (3.3) and hence Equation (3.5) are incorrect theoretically from the entropy viewpoint since there is double counting in the quantity of Q_0 and thus the probability-like terms in the equation are not mutually exclusive. Moreover, the entropy functions of

Awumah, Goulter and Bhatt do not directly account for the external inflows and outflows in the network. Although the external inflows and outflows may be known, they have to be considered in the entropy functions because there is uncertainty surrounding the contributions of the source supply at each node to the total flow reaching that node. Tanyimboh and Templeman (1993a) recognised the above weaknesses and proposed alternative and more rigorous flow entropy function which is presented next.

3.2.2 FLOW ENTROPY FUNCTION OF TANYIMBOH AND TEMPLEMAN

In general looped water distribution networks, the flow entering or leaving node n, $\forall n$, depends on whether the flow has reached the node n. In other words, the probability of the flow entering and leaving node n, $\forall n$, are both conditional upon the probability that the flow has reached the node n. Tanyimboh (1993) and Tanyimboh and Templeman (1993a) recognised this fact and used the conditional entropy formula of Khinchin (1953) and the multiple probability space model to formulate a rigorous entropy function for general looped water distribution networks in which the directions of the flow in the network are known and each node in the network may have either an external inflow or outflow. They introduced two conditional finite probability schemes for each node; the first represents the flow entering (inflow) and the other for the flow leaving (outflow) the node n, $\forall n$.

Based on the outflows, the conditional flow probability is given by

$$p_{nk} = \frac{q_{nk}}{T_n}, \quad \forall n, \ \forall k \in ND_n \tag{3.13}$$

in which p_{nk} is the conditional probability that flow leaving node n, uses link nk; T_n represents the total flow leaving node n, i.e.,

$$T_n = \sum_{k \in ND_n} q_{nk}, \quad \forall n \tag{3.14}$$

It should be noted that ND_n in Equations (3.13) and (3.14) not only represents the set of immediate nodes downstream of node n but also include any external outflows or demands. Defining q_{n0} as the demand at node n, Equation (3.13) includes the following probabilities.

$$p_{n0} = \frac{q_{n0}}{T_n} \, , \quad \forall n \in D \tag{3.15}$$

where D is the set of demand nodes in the network and p_{n0} is the fraction of the total demand consumed at node n.

The way in which Equation (3.13) defines the probability set ensures that it satisfies the normality condition without the need for separate normality constraints. Non-negativity of the probabilities is also ensured provided that link flows are always in the direction defined and never become negative. The nodal entropy can therefore be formulated as follows.

$$S_n^o = -\sum_{k \in ND_n} p_{nk} \ln p_{nk} \, , \quad \forall n \tag{3.16}$$

Tanyimboh (1993) and Tanyimboh and Templeman (1993a) have stated that the finite probability scheme represented by Equation (3.13) is conditional upon the probability that flow reaches node n. Therefore, by applying the conditional entropy function of Khinchin (1953), Tanyimboh and Templeman (1993a) were able to define the nodal entropy for a general network as the sum of the nodal entropy multiplied by the probability of flow arriving at the node through all possible paths, i.e.

$$S'^o_n = -P_n \sum_{k \in ND_n} p_{nk} \ln p_{nk} \, , \quad \forall n \tag{3.17}$$

in which S'^o_n is the entropy of node n associated with the outflows from that node; the apostrophe is used to indicate that the above equation represents the overall contribution of node n. ND_n is the set of all nodes immediately downstream of, and including any external outflows at, node n; P_n is the probability of flow arriving at node n. In Tanyimboh and Templeman (1993a) the nodal probability P_n, $\forall n$, is found by adding the probability of flow arriving at the node by each path. However, the probability of flow reaching node n can be seen as the probability of an event in a repeated experiment and can be interpreted in terms of the relative frequency as the frequency or the number of occurrences of that event divided by the sum of the frequencies of all the events in the experiment. Therefore, the following probability may be defined.

$$P_n = \frac{T_n}{T_0}, \quad \forall n \tag{3.18}$$

in which

$$T_n = \sum_{j \in NU_n} q_{jn}, \ \forall n \tag{3.19}$$

and T_0 is the total supply or demand, i.e.,

$$T_0 = \sum_{n \in I} q_{0n} = \sum_{n \in D} q_{n0} \tag{3.20}$$

Having defined the conditional entropy of node n, $\forall n$, the conditional entropy for any number of nodes in a water distribution network can then be defined by following the general form of conditional entropy of Equation (2.7). The first term of Equation (2.7) represents the entropy of an absolute rather a conditional finite scheme. Such an absolute scheme in a water distribution network is a scheme representing the fraction of the total supply provided by source node n, $\forall n \in I$, i.e.,

$$P_{0n} = \frac{q_{0n}}{T_0}, \quad \forall n \in I \tag{3.21}$$

Therefore, in accordance with Equation (2.7), Tanyimboh and Templeman (1993a, 1993b, 1993c) introduced the following entropy function for general water distribution networks.

$$S^o = S_0^s + \sum_{n=1}^{NN} S_n^o \tag{3.22}$$

in which S^o is the network entropy based on the outflows; S_n^o is the conditional entropy of outflows, including any demand, at node n, $\forall n$, as given by Equation (3.16);

$$S_0^s = -\sum_{n \in I} P_{0n} \ln P_{0n} \tag{3.23}$$

is the entropy of the distribution of T_0 amongst the sources and P_{0n} is given by Equation (3.21).

The above network entropy function of Equation (3.22) is based on the conditional finite probability scheme of Equation (3.13), which represents the outflows from node n, $\forall n$. Tanyimboh (1993) produced a similar network entropy function based on the inflows at node n, $\forall n$, which uses the following conditional flow probability function.

$$p_{jn} = \frac{q_{jn}}{T_n}, \quad \forall n, \ \forall j \in NU_n \tag{3.24}$$

The above equation also includes the probabilities for the supplies at source nodes, i.e.

$$p_{0n} = \frac{q_{0n}}{T_n}, \quad \forall n \in I \tag{3.25}$$

in which I is the set of supply nodes in the network; p_{0n}, $\forall n \in I$, is the probability representing the fraction of the total supply provided by source node n, $\forall n \in I$.

The network entropy based on the inflows is therefore

$$S^i = S_0^d + \sum_{n=1}^{NN} S_n^i \tag{3.26}$$

in which S^i is the network entropy; S^i_n is the conditional entropy of inflows, including any source supply, at node n, $\forall n$, and is given by

$$S^i_n = -P_n \sum_{j \in NU_n} p_{jn} \ln p_{jn} \ , \quad \forall n \tag{3.27}$$

where P_n is given by Equation (3.18) and p_{jn} is given by Equation (3.24). Also, in Equation (3.26), S^d_0 is the entropy of the distribution of T_0 amongst the demand nodes, i.e.,

$$S^d_0 = -\sum_{n \in D} P_{n0} \ln P_{n0} \tag{3.28}$$

in which P_{n0} is the fraction of the total demand consumed at node n, $\forall n \in D$, and is given by

$$P_{n0} = \frac{q_{n0}}{T_0} \ , \quad \forall n \in D \tag{3.29}$$

It follows from Equations (2.5) and (2.6) that the entropy of the outflows must equal the entropy of the inflows, i.e.,

$$S^i \equiv S^o \tag{3.30}$$

However, it may be noted that in general $S^i_n \neq S^o_n$, $\forall n$, just as $S(O_2 \mid O_1) \neq S(O_1 \mid O_2)$ (Tanyimboh, 1993). The network entropy of the outflows, S^o, is used throughout this thesis, and the superscript o is therefore dropped hereafter.

3.2.3 PATH ENTROPY FUNCTION OF YASSIN-KASSAB, TEMPLEMAN AND TANYIMBOH

In their attempts to formulate a procedure to calculate the maximum entropy flows in multi-source multi-demand networks, which will be discussed later in this chapter, Yassin-Kassab et al. (1999) formulated a slightly different method to determine the entropy of water distribution networks. Instead of using the link flows as in the method used by Tanyimboh and Templeman (1993a), Yassin-Kassab et al. (1999) used the path flows from the source node(s) to each demand node. Considering any demand node in the network of Figure 3.1 served by more than one path from a source, in accordance with the maximum entropy formalism of Problem 1, the flow supplied by a source to a demand node should be distributed equally amongst all the paths supplying that node from that source. Therefore, all the paths from the source to the demand node should have the same flow probability if there is no further information about those paths. However, the proportion of the flow for a demand node supplied by each source is unknown, and the relationship between path flows for each demand node supplied by different sources is consequently unknown. Defining $q_{p,ij}$ to be the path flow from source node i to demand node j, Figure 3.2 shows the unknown equal path flows from each source to each demand node reachable from that source. For demand node 6, there are three paths from each source in the network, these being 1-3-5-6, 1-3-4-6 and 1-4-6 from source node 1 (Figure 3.2a), each carrying the path flow $q_{p,16}$, and 2-3-5-6, 2-3-4-6 and 2-4-6 from source node 2 (Figure 3.2b), each carrying the path flow $q_{p,26}$. Demand node 5 receives two paths, one from each source node. The flow from source node 1 (Figure 3.2c) is equal to $q_{p,15}$, and from source node 2 (Figure 3.2d) equal to $q_{p,25}$. Demand node 4 receives two paths from source node 1 (Figure 3.2e), each carrying a flow equal to $q_{p,14}$, and two paths from source node 2 (Figure 3.2f), with each path carrying a flow equal to $q_{p,24}$. Finally, there are two paths supplying demand node 3, one path from each source (Figure 3.2g and Figure 3.2h, respectively), where the path from source node 1 carries a flow equal to $q_{p,13}$ and the path from source node 2 carries $q_{p,23}$.

Consequently, there are eight unknown path flows for the network of Figure 3.1, two for each demand node. However, four node equilibrium or nodal continuity equations, which states that the inflows to a node must balance the outflows from that node, i.e.,

$$\sum_{j \in NU_n} q_{jn} - \sum_{k \in ND_n} q_{nk} = q_n \,, \quad \forall n \qquad (3.31)$$

in which q_n is the external inflow or outflow at node n, can be constructed, one for each demand node. The above nodal equilibrium equations can be expressed in terms of the path flows by equating the path flows supplying each demand node to the demand of that node. The expressions can be generally written as

$$\sum_{i \in I_j} NP_{ij} q_{p,ij} = q_{j0} \,, \quad \forall j \in D \tag{3.32}$$

in which NP_{ij} is the number of paths from source i to demand node j; I_j and D are the set of all source nodes supplying node j and the set of all demand nodes in the network respectively; and q_{j0} is the demand at node j.

Also, two source equilibrium equations can be set up, one for each source node. These equations can be generally expressed as

$$\sum_{j \in D_i} NP_{ij} q_{p,ij} = q_{0i} \,, \quad \forall i \in I \tag{3.33}$$

in which D_i is the set of all demand nodes supplied by source node i; q_{0i} is the supply at source node i.

Following this relative frequency interpretation, Yassin-Kassab (1998) and Yassin-Kassab et al. (1999) obtained the probabilities of path flows by normalizing each path flow by its individual source flow, i.e.,

$$p_{p,ij} = \frac{q_{p,ij}}{q_{0i}} \,, \quad \forall i \in I; \ \forall j \in D_i \tag{3.34}$$

in which $p_{p,ij}$ is the probability that path flow $q_{p,ij}$, which is supplied by source node i, reaches demand node j. Substituting Equation (3.34) into Equation (3.33) gives the following normality condition at the source nodes.

$$\sum_{j \in D_i} NP_{ij} p_{p,ij} = 1 \,, \quad \forall i \in I \tag{3.35}$$

Equation (3.34) represents *NS* sets of path probabilities, where *NS* is the number of sources in the network, each set corresponding to a source. For each set, the path probabilities are mutually exclusive and they sum to unity [Equation (3.35)]. Therefore, each set of path probabilities represents a finite scheme. However, each scheme is dependent upon the condition at the corresponding source, i.e. each set of the path probabilities is dependent upon the probability representing the fraction of the total supply in the network provided by that source. There are *NS* such probabilities in the network and their expressions have been given in Equations (3.21). Following Khinchin (1953), the conditional entropy of the path flows for each demand node can be written as

$$S_i^p = -P_{0i} \sum_{j \in D_i} Np_{ij} p_{p,ij} \ln p_{p,ij} , \quad \forall i \in I \tag{3.36}$$

where S_i^p is the conditional entropy of path flows for source node i.

By using the concept of compound scheme presented in Chapter 2, the network entropy in the context of path flows can be formulated. This is done by summing the entropy of the source supplies to the conditional entropy of the path flows from each source to each demand node, i.e.

$$S^p = S_0^s + \sum_{i \in I} S_i^p \tag{3.37}$$

in which S^p is the network entropy based on path flows; S_0^s is the entropy of the distribution of the total supply amongst all sources and is given by Equation (3.23); S_i^p is given by Equation (3.36).

3.3 CALCULATING MAXIMUM ENTROPY FLOWS IN NETWORKS

Consider a case of a buried water network in which the available information is not sufficient to uniquely determine the pipe flow rates in the network. Information on pipe lengths, diameters and roughness is assumed to be unavailable. However, source flow rates, demand flow rates and the topology of the network with the flow directions are assumed to be known. Under such circumstances, how can the most likely pipe flow rates in the network be estimated?

It has been suggested that the flow distribution that has the maximum entropy and satisfies the available information must be used in accordance with Jaynes' Maximum Entropy Formalism. Having defined the appropriate entropy function for network flows, Tanyimboh and Templeman (1993a) calculated the maximum entropy flows in a looped network by maximizing the network entropy of Equation (3.22) subject to non-negativity of all the link flows and flow equilibrium at each node in the network. The optimization problem can therefore be presented as in Problem 2 below.

Problem 2

$$\underset{\forall q_{nk}}{\text{Maximize}} \qquad S = S_0^s + \sum_{n=1}^{NN} S_n \tag{3.38}$$

Subject to:

$$\sum_{k \in ND_n} \frac{q_{nk}}{T_n} = 1, \quad \forall n \tag{3.39}$$

$$\sum_{j \in NU_n} q_{jn} - \sum_{k \in ND_n} q_{nk} = q_n \,, \quad n = 1, \ldots, NN-1 \tag{3.34}$$

$$\frac{T_n}{T_0} \geq 0, \quad \forall n \tag{3.40}$$

$$\frac{q_{nk}}{T_n} \geq 0, \quad \forall n, \ \forall k \in ND_n \tag{3.41}$$

Problem 2 is a convex programming problem because the objective function of Equation (3.38) is concave since it is the sum of a set of concave functions of the form $-\sum p_i \ln p_i$, and also the linear constraints in Problem 2 represent a convex set. Therefore, Problem 2 has a unique global maximum point which can be obtained using any standard constrained non-linear programming algorithm. Tanyimboh and Templeman (1993a), however, solved Problem 2 as an unconstrained optimization problem after eliminating the non-negativity constraints of Equations (3.40) and (3.41) by arguing that the maximum entropy flows are expected to be as uniform as possible without any being equal to zero. Moreover, the network entropy of Equation (3.38) will be undefined in the infeasible region, thus satisfying the non-negativity constraints implicitly in the objective function.

3.4 PATH-BASED ALGORITHM FOR CALCULATING MAXIMUM ENTROPY FLOWS IN PIPE NETWORKS

3.4.1 SINGLE SOURCE NETWORKS

The proposed method by Tanyimboh and Templeman (1993a) to calculate the maximum entropy flows, and presented earlier as Problem 2, involves non-linear programming. In their subsequent paper, Tanyimboh and Templeman (1993c) proposed a simpler non-iterative path-based approach to calculate the maximum entropy flows in single source networks.

Following the maximum entropy formalism, Tanyimboh and Templeman (1993c) argued that when a demand node is served by more than one path from the source, the demand of that node should be divided equally amongst all paths supplying it if there is no further information about those paths. This approach is demonstrated in Appendix A1. Tanyimboh and Templeman (1993c) presented several algorithms for this method. They are node numbering algorithm, node weighting algorithm and flow distribution algorithm.

A single source network shown in Figure 3.3 is used to demonstrate the above algorithms. First, all the nodes in the network are numbered according to the node numbering algorithm. The source node is given the number 1, and then the rest of the nodes are numbered in an ascending sequence starting with any node whose upstream nodes have all been numbered. The numbering of nodes 4 and 5 is arbitrary and may be interchanged.

Once all the nodes in the network have been numbered, the number of paths from the source to each demand node can be calculated using the node weighting algorithm as shown in Figure 3.4. This is done by assigning a weight of 1 to the source node and, in ascending node numbering sequence, the weight of each demand node is equal to the sum of the weights assigned to all nodes immediately upstream of it. Consequently, the node numbering algorithm ensures that all nodes immediately upstream of the node being considered have been weighted. For the network in Figure 3.3, the weight of node 2 is equal to the weight of node 1 since node 1 is the only node upstream of node 2. The weight of node 3 is the sum of the two upstream nodes, which are node 1 and node 2, and is therefore equal to 2. Similarly, the weight of node 4 is equal to 3, which is the sum of the weights of nodes 1 and 3, and the weight of node 5 is also equal to 3 resulting from the weights of node 2 and 3.

Finally, the flow distribution algorithm is used to determine the maximum entropy flows in the network. It operates in descending node number order starting from the terminal node, i.e. the most downstream node in the network. Therefore, for the network in Figure 3.3, the algorithm can start from either node 5 or node 4. The flow distribution algorithm ensures that the total outflow at a node is shared among the inflow links incident at the node in proportion to the upstream nodal weights. Hence, starting with node 5, the flow in link 2-5 is obtained by multiplying the total outflow at node 5, which is 24 units, by the ratio between the weights of nodes 2 and 5, which is 1/3. Similarly, the flow in link 3-5 is equal to 24 multiplied by the ration 2/3, which is the ratio between the weights of nodes 3 and 5. The next step is to choose any node immediately upstream of node 5, whose link outflows have all been calculated. However, both nodes 2 and 3 have outflow links whose flows have not been calculated. In consequence, they cannot be treated yet. At this point, the procedures stops, and re-starts at any terminal node that has not been dealt with. For the network considered, it is node 4. If the same procedures as explained for node 5 are applied to node 4, the flow in links 1-4 and 3-4 are 5 units and 10 units, respectively. At this stage, the flow for the links incident at node 3 can be calculated. The total outflows leaving node 3 is 36 units. Consequently, the links 1-3 and 2-3 share the total outflows at node 3 equally since the nodes immediately upstream of these links have equal weights. The flow in link 1-2 can then be obtained as the sum of the outflows from node 2, including the demand at node 2. The resulting maximum entropy flows are shown in Figure 3.5.

The above algorithms are rigorous for single source networks. They produce identical results to those given by solving Problem 2 in a much simpler and quicker method since it does not involve linear or non-linear programming and the procedures are non-iterative. Tanyimboh and Templeman (1993c) attempted to extend the above algorithms to multiple source networks by means of a super-source concept. However, Walters (1995) pointed out that the concept was actually incorrect. As an answer to this problem, Yassin-Kassab et al. (1999) proposed relatively simple algorithms for calculating maximum entropy flows for general multi-source networks based on the path concept. These simple algorithms, which also do not involve linear nor non-linear programming, will be presented next.

3.4.2 GENERAL MULTIPLE SOURCE NETWORKS

In a single source network the demand at any node is numerically known and can be supplied only from one source. It is therefore easy to allocate the path flows equally among all paths from the source to the demand node. In the case where there are several sources in the network, although the total demand at a node is numerically known, the proportion of the flow received from each source may not be known and cannot be allocated numerically among the available paths. Yassin-Kassab et al. (1999) concluded that the key to solving the problem is in determining the proportions of the flow received by a demand node from each of the sources.

To demonstrate the path-based method on multiple source networks, Yassin-Kassab et al. (1999) used the two-source network, shown here as Figure 3.6, whose maximum entropy flows had first been determined by solving Problem 2 for the network. The resulting flows from the optimization are shown in Figure 3.7 with a maximum entropy value of 2.3885315. As for the single source networks, the multiple paths from each source to a demand node in multiple source networks must carry the same flow. Figure 3.8 shows the unknown equal path flows from each source to each demand node. By equating the total flow in each link of Figure 3.7 to the sum of all path flows passing through that link in Figure 3.8, the following equations can be obtained.

$2q_{p,15} + q_{p,14} + q_{p,13} = 20.061912,$	for link 1-3	(3.42a)
$q_{p,15} + q_{p,14} = 9.938088,$	for link 1-4	(3.42b)
$2q_{p,25} + q_{p,24} + q_{p,23} = 16.268405,$	for link 2-3	(3.42c)
$q_{p,25} = 3.731595,$	for link 2-5	(3.42d)
$q_{p,25} + q_{p,15} + q_{p,24} + q_{p,14} = 17.996982,$	for link 3-4	(3.42e)
$q_{p,25} + q_{p,15} = 8.333335,$	for link 3-5	(3.42f)
$q_{p,25} + 2q_{p,15} = 12.935070,$	for link 4-5	(3.42g)

Solving the above equations, all the path flows can be obtained. Also, by normalizing each path flow by its individual source flow, the path flow probabilities $p_{p,ij}$ can be obtained as given by Equation (3.34). By investigating the path flow probabilities further, Yassin-Kassab et al. (1999) noticed that for each demand node the ratio of the path flow probabilities from each pair of sources are identical, i.e., for the network of Figure 3.6,

$$\frac{p_{p,13}}{p_{p,23}} = \frac{p_{p,14}}{p_{p,24}} = \frac{p_{p,15}}{p_{p,25}} = \alpha_e \qquad (3.43)$$

Hence, it appears that all demand nodes receive their flows from the two sources with the same proportion. Based on this finding, Yassin-Kassab et al. (1999) presented the following principle:

> The maximum entropy flows in multiple source networks are such that the ratio of the probabilities of path flows from any pair of sources to a demand node reachable from those sources is the same for every demand node supplied by those sources in the network.

Following the above principle, Yassin-Kassab et al. (1999) formulated the solution for finding the maximum entropy flows in multiple source networks starting from the equilibrium equations at each demand node in the network. For the network in Figure 3.6, the following nodal continuity equations based on the path flows in Figure 3.8 can be obtained.

$$3q_{p,15} + 3q_{p,25} = 25 \tag{3.44a}$$

$$2q_{p,14} + q_{p,24} = 15 \tag{3.44b}$$

$$q_{p,13} + q_{p,23} = 10 \tag{3.44c}$$

By normalizing the path flows by their individual source flows to obtain the path flow probabilities and substituting them in the above nodal continuity equations, the following equations are obtained.

$$90p_{p,15} + 60p_{p,25} = 25 \tag{3.45a}$$

$$60p_{p,14} + 20p_{p,24} = 15 \tag{3.45b}$$

$$30p_{p,13} + 20p_{p,23} = 10 \tag{3.45c}$$

Substituting the path flow probability ratios of Equation (3.43) into the above equations gives

$$p_{p,15} = 25\alpha_e /(60 + 90\alpha_e) \tag{3.46a}$$

$$p_{p,14} = 15\alpha_e /(20 + 60\alpha_e) \tag{3.46b}$$

$$p_{p,13} = 10\alpha_e /(20 + 30\alpha_e) \tag{3.46c}$$

$$p_{p,25} = 25 /(60 + 90\alpha_e) \tag{3.46d}$$

$$p_{p,24} = 15 /(20 + 60\alpha_e) \tag{3.46e}$$

$$p_{p,23} = 10 /(20 + 30\alpha_e) \tag{3.46f}$$

Using the normality condition of Equation (3.35) for source node 1, the following equation is obtained.

$$NP_{15}p_{p,15} + NP_{14}p_{p,14} + NP_{13}p_{p,13} = 1 \tag{3.47}$$

Substituting the probabilities of the path flow of Equations (3.46) into Equation (3.47) yields the following equation

$$75\alpha_e/(60+90\alpha_e)+30\alpha_e/(20+60\alpha_e)+10\alpha_e/(20+30\alpha_e)=1 \tag{3.48}$$

which can be solved to give the value of α_e. Note that the normality condition at source node 2 can be used to check the value of α_e. Back-substituting α_e into Equations (3.46) gives the path flow probabilities and hence path flows. Finally, the link flows can be obtained by summing all the path flows passing each link in turn, which turn out to be the same as those obtained earlier by solving Problem 2 for the network. It should be noted that for a network with *NS* sources, there will be *NS* normality condition equations available for the network, one for each source. Any (*NS*-1) of these equations can be used to determine the value of α_e, and hence the path flow probabilities in the network. The remaining normality condition can be used for checking purposes.

Yassin-Kassab et al. (1999) presented several algorithms for the above procedures. These algorithms are a global node numbering algorithm, a source reachability algorithm, a demand node reachability algorithm, a local node numbering algorithm, a node weighting algorithm, an Alpha algorithm and a flow distribution algorithm.

The global node numbering algorithm allocates a number to each node in the network starting from the source nodes in ascending order followed by any demand node also in ascending order until all the nodes in the network have been numbered. The global node numbering for the network in Figure 3.6 is given here in Figure 3.9a. The source reachability algorithm and the demand node reachability algorithm are then applied to construct sub-networks by identifying which demand nodes are supplied by which sources. Therefore, for networks with *NS* sources, there will be *NS* sub-networks that have to be constructed, one for each source. Once the sub-networks have been constructed, each of the nodes in the sub-network should be renumbered according to the local node numbering algorithm. Therefore, working on each of the sub-networks in turn, the source node in a sub-network should first be allocated the number 1. The demand nodes can then be numbered starting from the node whose upstream nodes have all been numbered.

Once all the nodes in all sub-networks have been numbered, the number of paths from the source to each demand node in the sub-network can be calculated using the node weighting algorithm. The node weighting algorithm is applied in the same way as that for single source networks. The only difference is that in multiple source networks, it is applied to each of the sub-networks in turn instead of the whole network.

Yassin-Kassab et al. (1999) used the Alpha algorithm to calculate the path flows in network. The algorithm defines the path flow probability as follows.

$$p_{p,ij} = \frac{q_{j0}\alpha_{ei}}{\sum_{i \in I_j} NP_{ij} q_{0i} \alpha_{ei}}, \quad \forall i \in I_j \tag{3.49}$$

in which α_{ei} is the ratio of the path flow probabilities related to node i. The above equation is obtained by substituting Equation (3.34) into Equation (3.32) and adding the α_{ei} terms. The values of α_e are obtained by constructing (NS-1) normality conditions of Equation (3.35) and solving the equations. However, due to the way the path flow probabilities are defined, there are consequently NS values of α_e. Since only (NS-1) are needed, the value of α_{e1} can be set to unity. At this stage, the network entropy can be calculated by using the path entropy function of Equation (3.37). Also, using Equation (3.34) the path flows in the network can be obtained.

To calculate the maximum entropy flows in the network, the flow distribution algorithm first calculates the link flows in each of the sub-networks. It operates in descending local node number order starting from the terminal node and continuing to the node whose link outflows have all been calculated. The algorithm ensures that the total outflow at a node in a sub-network is shared among the inflow links incident at the node in proportion to the upstream nodal weights. However, the total outflow at each demand node must first be expressed in terms of the path flow instead of the actual outflow, i.e.,

$$T_{ji} = NP_{ij} q_{p,ij} + \sum_{k \in ND_j \subset SNK_i} q_{jk,i}, \quad \forall j, \ \forall i \tag{3.50}$$

in which T_{ji} is the total outflow at local node j; $(NP_{ij}q_{p,ij})$ is the local demand at node j supplied by source node i; SNK_i indicate the sub-network related to source node i; $ND_j \subset SNK_i$ represents the set of immediate downstream nodes of node j, provided that these downstream nodes are in the sub-network SNK_i; and $q_{jk,i}$ is the flow in link jk in the sub-network SNK_i. Note that the actual demand at the node whose local node number in the sub-network SNK_i is j is not included in T_{ji}.

Once all the outflows in the sub-networks have been expressed as in Equation (3.50), the link flows in a sub-network are calculated by multiplying the total outflows at the local demand node by the ratio between the weight at the node to the weight at the immediate upstream node of the link considered. The final maximum entropy flows in each link in the network are obtained by summing the flows from the corresponding link in the sub-networks.

The above procedures are rigorous and have been applied to the network of Figure 3.6 by Yassin-Kassab et al (1999). The results of this application are shown in Figures 3.9b and 3.9c.

3.5 BRIEF OUTLINE OF THE PREVIOUS WATER DISTRIBUTION NETWORK ENTROPY APPLICATIONS

The difficulties faced in direct quantification of the performance of water distribution networks have motivated researchers to find a suitable surrogate performance measure. This is the early motivation of using the entropy in water distribution networks (Awumah et al., 1989, 1990, 1991; Tanyimboh, 1993; Tanyimboh and Templeman, 1993a, 1993b, 1993c). Investigations by Tanyimboh and Templeman (1993a, 1993b, 1993c) have shown that there is a correlation between entropy and reliability in which the increase in the entropy value is followed by the increase in the reliability of the network. Recently, Ang and Jowitt (2003) have explored the relationship between entropy and energy loss in water distribution network to help gain a deeper understanding of the properties of entropy. Their results support the previous conclusion by Tanyimboh and Templeman (1993b) that the importance of a pipe in a water distribution network can be related to the amount of energy that the network dissipates following the removal or closure of that pipe.

The ease of calculating the entropy value of water distribution networks and its minimal data requirements have also been exploited by Tanyimboh and Templeman (1993b, 2000) by incorporating the entropy directly into the optimization process. Details of the way in which entropy is integrated into the optimization procedures are presented in Chapter 5. Recently, Tanyimboh and Sheahan (2002) have proposed a maximum entropy-based approach to the layout optimization of water distribution networks. The problem of optimizing the layout and pipe sizes of water distribution networks is extremely difficult as will be explained later in Chapter 5. Tanyimboh and Sheahan (2002) have demonstrated that their maximum entropy-based approach is quite robust and efficient. This approach is studied in more detail later in Chapter 7 of this thesis.

Finally, Templeman and Yassin-Kassab (2002) have proposed an entropy-based approach to the calibration of computer models of loop water networks with limited data. The approach has been successfully applied to hypothetical networks to predict the most likely pipe characteristic in the network. Nevertheless, problem with the accuracy of the calibration was noticed. More studies are therefore required to deal with this issue. The problem of selecting the type of information required in the calibration to make the results more accurate needs to be addressed.

3.6 SUMMARY AND CONCLUSIONS

The entropy function for water distribution networks has been detailed in this chapter. The problem of obtaining the network flows that correspond to the maximum entropy of the network was also presented in this chapter. Simple algorithms for calculating the maximum entropy flows were described and the example of their application to a hypothetical network was given. Finally, several previous entropy applications in water distribution networks were presented.

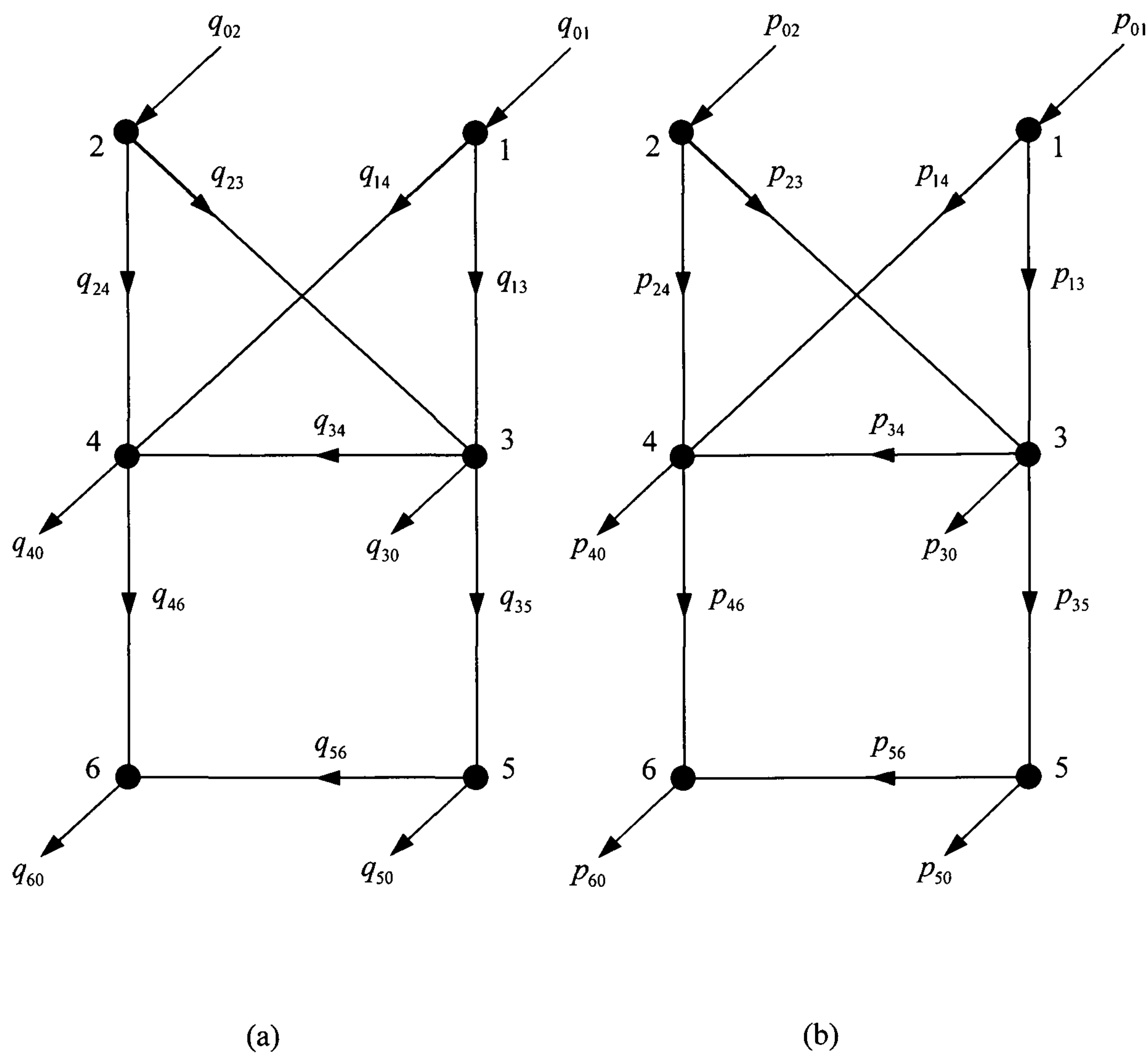


Figure 3.1 Water Supply network adapted from Tanyimboh and Templeman (1993a).

(a). Supply, demand and pipe flow definitions.

(b). Flow probability definitions.

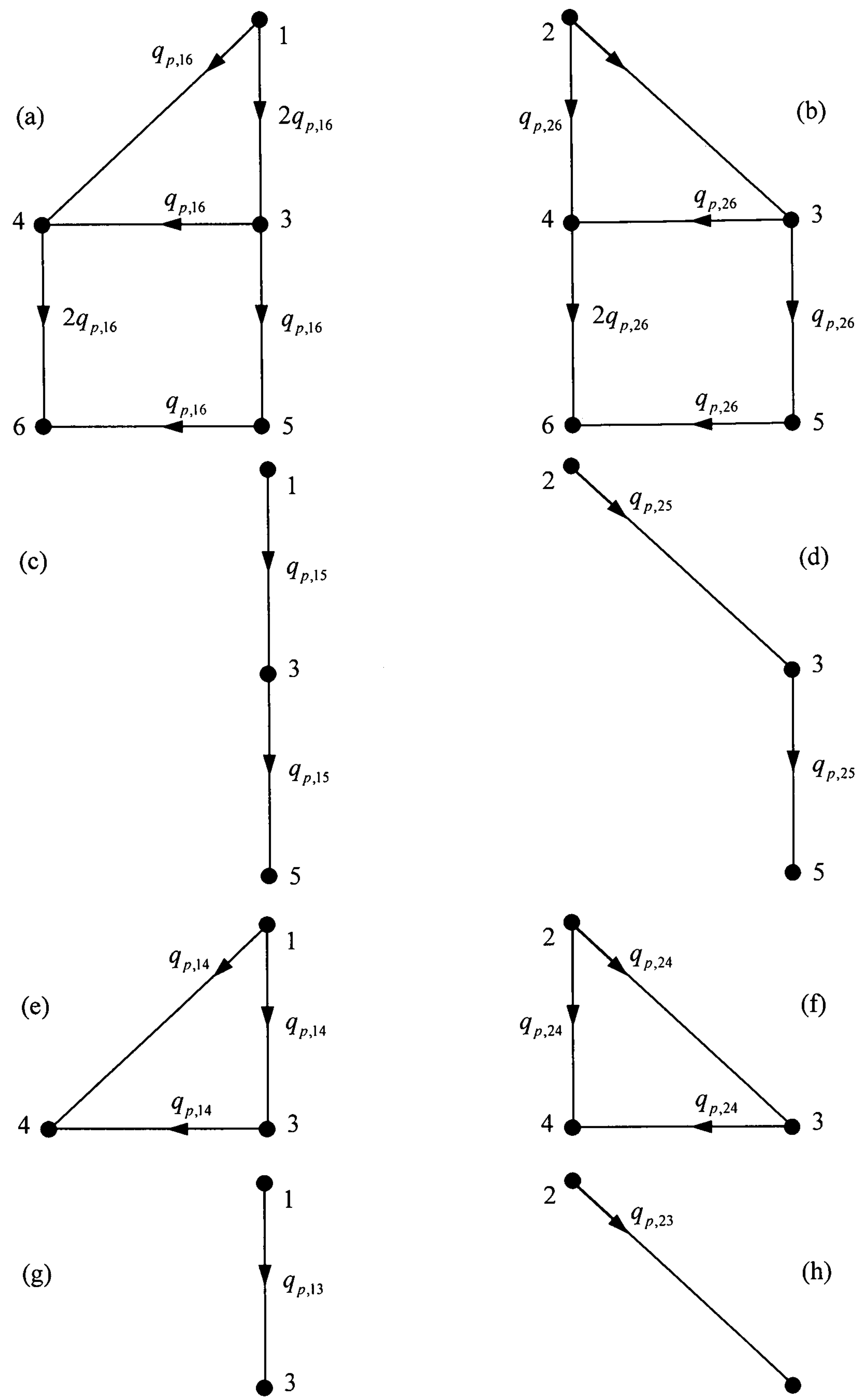


Figure 3.2 Equal path flows from each source to each reachable demand node (Tanyimboh and Templeman, 1993a).

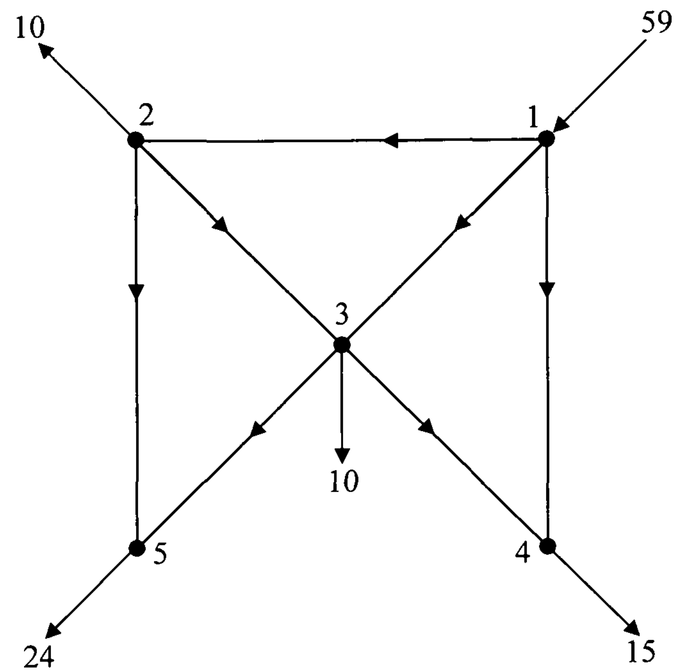


Figure 3.3 Single source network taken from Tanyimboh and Templeman (1993c).

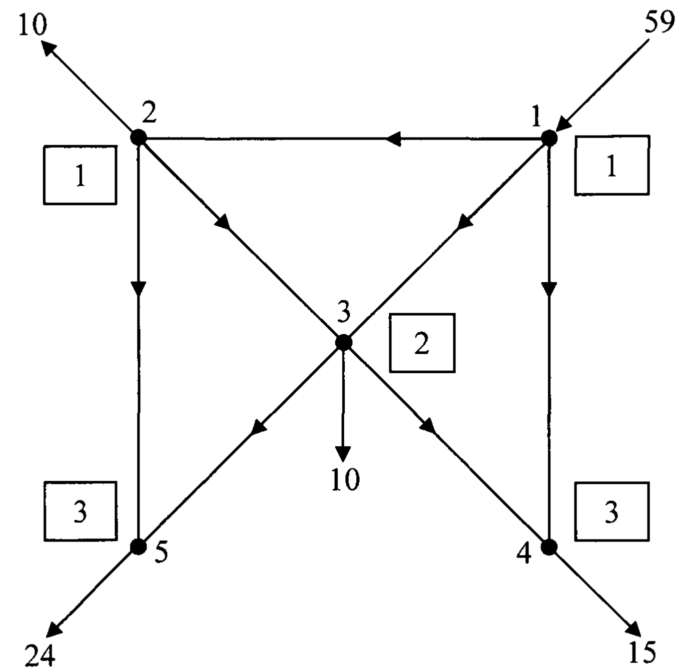


Figure 3.4 Number of paths to each node for the network of Figure 3.3 (Tanyimboh and Templeman, 1993c).

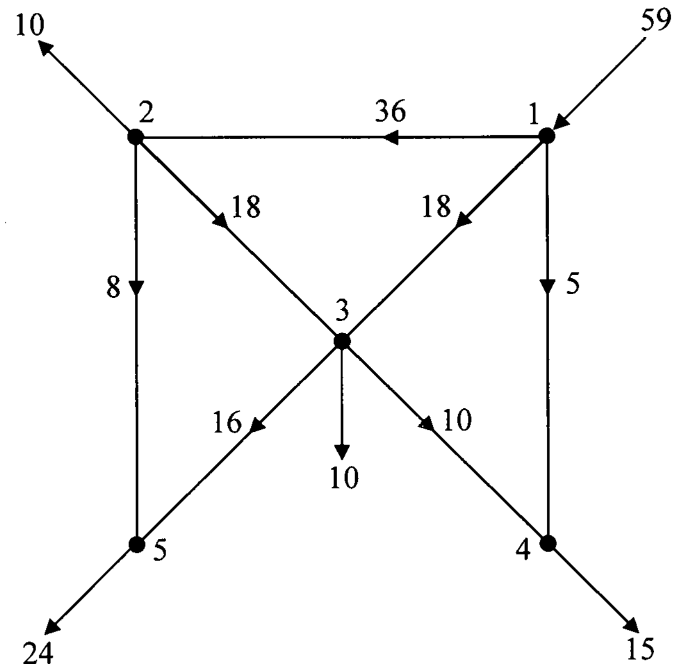


Figure 3.5 Maximum entropy flows for the network of Figure 3.3 (Tanyimboh and Templeman, 1993c).

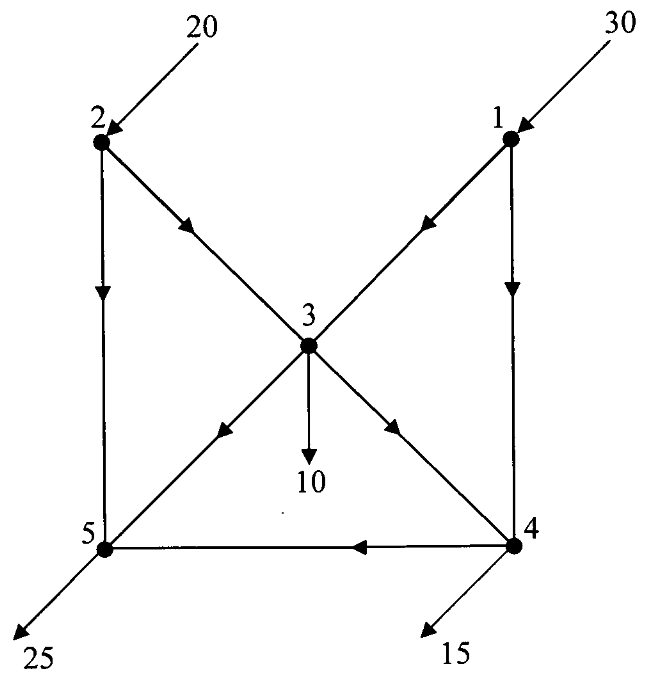


Figure 3.6 Two source network adapted from Yassin-Kassab et al. (1999).

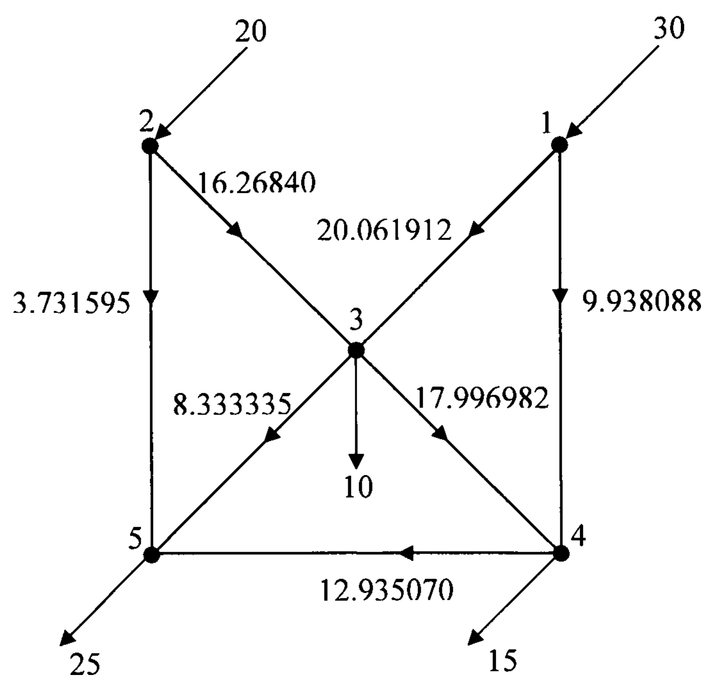


Figure 3.7 Maximum entropy flows for the network of Figure 3.6, obtained by solving Problem 2 for the network (Yassin-Kassab et al., 1999).

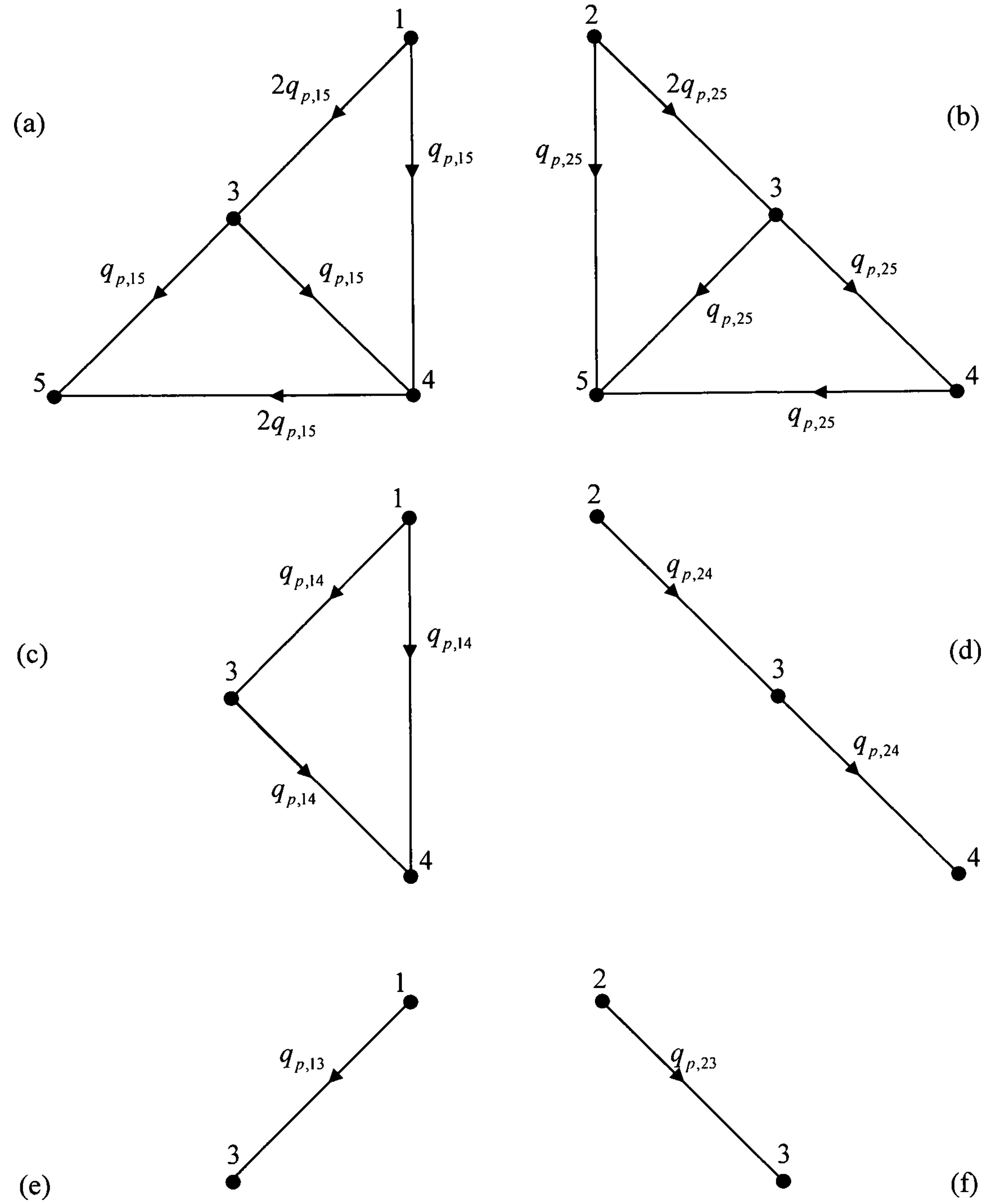


Figure 3.8 Equal path flows from each source to each demand node for the network of Figure 3.6 (Yassin-Kassab et al., 1999).

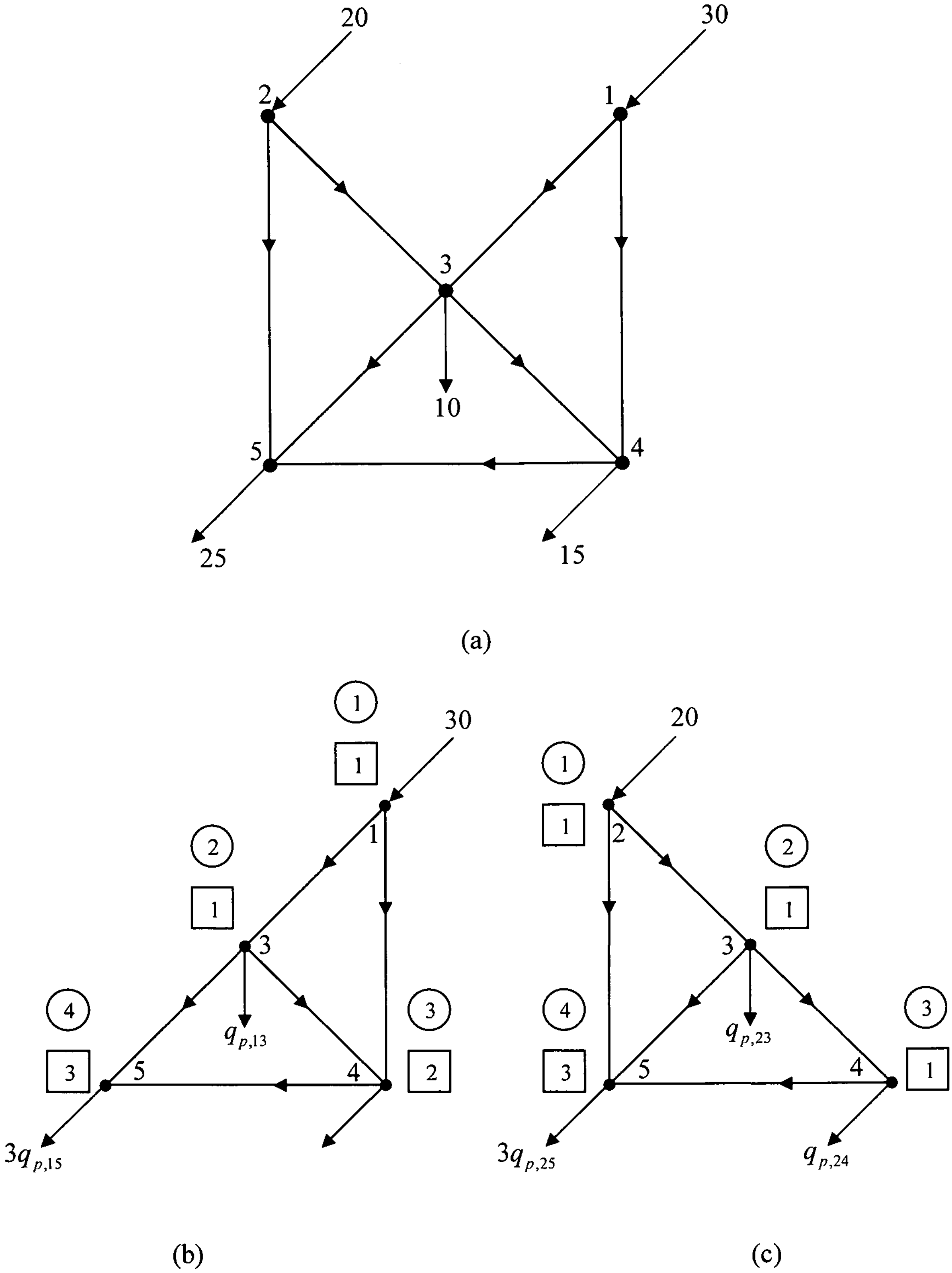


Figure 3.9 Maximum entropy flows for the network of Figure 2.6 (Yassin-Kassab et al., 1999)

(a) Global node numbering, (b) Sub-network 1, (c) Sub-network 2.

CHAPTER 4 HYDRAULIC ANALYSIS OF WATER DISTRIBUTION NETWORKS

4.1 INTRODUCTION

Water distribution networks are required to distribute water to customers at the desired quantity and pressure head. These requirements are normally satisfied amidst the constant change in the network flows which is often gradual as a result of the fluctuation of water consumptions, which usually follows a certain pattern. However, sudden drop in pressure or increase in consumption also occurs occasionally as a result of component failure or fire fighting requirements. In these situations, the network may be over-stressed and unable to deliver enough water at enough pressure to some parts of the network.

In order to maintain service, Engineers must be able to analyse the network under normal and abnormal conditions so that necessary measures can be taken to minimise the negative impact that may be experienced by the network in emergency situations. Nowadays, many network analysis softwares are available in the market. The basic principles behind these softwares are the same, which consist of the constitutive equations or the governing laws for flow of water in pipe networks. Details of the constitutive equations are given next. Three possible formulations of these equations are then presented followed by several methods for flow analysis in looped water networks.

4.2 CONSTITUTIVE EQUATIONS

Pipe flows in water distribution networks are subject to a loss in energy. The energy loss per unit weight is called a *head* loss. This energy loss is caused by frictional resistance along the pipe wall, which acts in the opposite direction to the flow. Energy loss also occurs when there is a change in the flow momentum, for example at a bend or at sections with fittings, such as flow measuring devices, valves, etc. Head losses other than that caused by the pipe friction are called *minor losses.* Although the loss of energy in pipe flows is unavoidable, the quantity of the flow, however, must remain the same unless there is an abstraction somewhere along the line. This is in accordance with the flow equilibrium or the conservation of mass requirement. The constitutive equations for flow in pipe networks, therefore, consist of the flow equilibrium equations, also often called the *continuity* equations, head loss equations, and the conservation of energy equations, which consist of loop and path equations (Bhave, 1991; Walski et al., 2003).

4.2.1 CONTINUITY EQUATIONS

The continuity equations are usually applied at each node in the network. The equations have been presented earlier in Chapter 3 as Equation (3.31) and restated below as Equation (4.1) for completeness.

$$\sum_{j \in NU_n} q_{jn} - \sum_{k \in ND_n} q_{nk} = q_n \ , \quad n = 1, \ldots, NN - 1 \qquad (4.1)$$

in which q_{jn} and q_{nk} are the flows in link *jn* and *nk* respectively; q_n is the external inflow or outflow at node *n*; the set NU_n consists of all the immediate upstream nodes of node *n*; and the set ND_n consists of all the immediate downstream nodes of node *n*.

4.2.2 HEAD LOSS EQUATIONS

Head loss equations consist of the pipe head loss and minor head loss as mentioned earlier. However, friction is usually the predominant cause of head loss and therefore minor losses are not considered explicitly throughout this thesis as explain later in this sub-section.

4.2.2.1 PIPE HEAD LOSS

There are several pipe head loss equations available. The foremost of these equations is the Darcy-Weisbach equation given below.

$$h_{ij} = 4 f_{ij} \frac{L_{ij}}{D_{ij}} \frac{v_{ij}^2}{2g}, \quad \forall ij \in IJ \tag{4.2}$$

in which L_{ij} and D_{ij} are the length and internal diameter of pipe ij respectively; h_{ij} is the head loss, which is positive in the direction of flow; v_{ij} is the mean velocity of the flow in the pipe; g is the acceleration due to gravity; f_{ij} is the friction factor, which is dimensionless and depends on the flow rate and the roughness of the pipe; IJ is the set of all the links in the network. In general, f_{ij} cannot be written explicitly in terms of the flow rate and the roughness of the pipe. Consequently, some iterative scheme is usually needed for its determination. Jeppson (1976, pp. 30) has tabulated several equations for f_{ij} for various flow conditions.

Empirical approximate equations, which are easier to use than the Darcy-Weisbach equation, are available. One of these equations is Manning's equation stated below.

$$h_{ij} = \frac{\alpha L_{ij} (n_{ij} q_{ij})^2}{D_{ij}^{5.333}}, \quad \forall ij \in IJ \tag{4.3}$$

in which n_{ij} is Manning's coefficient; q_{ij} is the flow rate, which is positive in the direction of flow; and α is a dimensionless conversion factor, which is equal to 10.29 in S.I. units.

Another empirical approximate head loss equation, which is frequently used, is the Hazen-Williams equation.

$$h_{ij} = \alpha L_{ij} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ij}^{4.87}}, \quad \forall ij \in IJ \tag{4.4}$$

in which C_{ij} is the Hazen-Williams coefficient and α here equals to 10.67 in S.I. units. The Hazen-Williams equation is used throughout this thesis. However, other head loss equations may be used if it is considered appropriate.

4.2.2.2 MINOR LOSSES

Since minor losses do not contribute much to the energy loss, the concept of *equivalent length* is used to account for their effect in the analysis of the pipe networks in this thesis. The equivalent length is the length of pipe of the same diameter as the pipe with the fitting, which would cause the same head loss as the fitting. In general, the head loss across any fitting depends on the flow rate in the pipe with the fitting. Using equivalent length, Equation (4.2) can be written as

$$h_f = 4f_{ij}\frac{L_e}{D_{ij}}\frac{v_{ij}^2}{2g} = K_f\frac{v_{ij}^2}{2g} \tag{4.5}$$

where h_f is the head loss across the fitting; L_e is the equivalent length of the fitting; K_f is a coefficient of the fitting. From Equation (4.5)

$$L_e = \frac{K_f D_{ij}}{4f_{ij}} \tag{4.6}$$

Some typical values of K_f can be found in many hydraulic text books, for example Walski et al., (2003, pp.41).

The effective length of a pipe for the head loss calculations is the sum if its physical length and the equivalent length of all its fittings. It is assumed that all pipe lengths are effective lengths throughout this thesis.

4.2.3 CONSERVATION OF ENERGY EQUATIONS

4.2.3.1 LOOP EQUATIONS

Conservation of energy requires the net loss of energy around a loop to be zero. Therefore, the loop equations can be written as

$$\sum_{ij \in IJ_l} h_{ij} = 0 \ , \quad l = 1,...,NLP \tag{4.7}$$

in which IJ_l is the set of all links in loop l and NLP is the number of loops in the network, which must satisfy the following equation.

$$NL = NN + NLP - 1 \tag{4.8}$$

where NL and NN are the number of links and nodes in the network respectively.

4.2.3.2 PATH EQUATIONS

The total head loss along any path must equal the difference in head between the end points of that path. An equation may therefore be written for each path along which the head loss is known.

$$\sum_{ij \in IJ_p} h_{ij} = h_p \ , \quad p = 1,...NP \tag{4.9}$$

in which IJ_p is the set of all links in path p; h_p is the known head loss for path p; NP is the number of paths whose head losses are known. In general, NP will be, at most, one less than the number of constant path head losses in the network. It may also be noted that a path may contain only one link. Furthermore, to ensure linear independence of the path equations, the NP paths must be specified such that none of the paths duplicates information contained in any other path.

Finally, head loss in a link may also be obtained from the difference in head at the nodes at each end of that link, i.e.

$$h_{ij} = H_i - H_j \ , \quad \forall ij \in IJ \tag{4.10}$$

in which H_i and H_j are the total heads at nodes i and j respectively. The total head at a node is the sum of the elevation and the pressure head at that node.

For simplicity, the velocity head is considered negligible throughout the present study.

4.2.4 STORAGE TANKS AND MULTIPLE DEMAND PATTERNS

A storage tank is a boundary node in a distribution network that can supply and accept water in large quantity. In designing a storage tank, the decision variables could be the elevation, volume of the tank in terms of the water level or combination of these. The location of the storage tank is usually pre-specified by the designer, although, it can also be incorporated into the optimization procedure as another variable. To make sure that the proposed storage tank operates in a satisfactory manner, more than one demand pattern, i.e. multiple loading conditions, need to be considered in the design. This is important since by definition a storage tank has to act as a buffer for the sources, i.e. to fill at times of low demands and empty when demands are at their peak. The capital cost of the tank has to be included in the objective function in the optimization process for each tank added. This thesis, however, is not particularly concerned with the design of storage tanks. Hence, details of the design procedures are not presented here. Interested readers may consult, for example, Alperovits and Shamir (1977).

In the absence of a storage tank in a network, many different loadings or demand patterns can still occur in 24-hour period, for example, peak hour demand, daily average demand, daily maximum demand, fire demand, periods of low demand (usually at night), etc. For problems with multiple demand patterns there will be a set of corresponding nodal heads, pipe flow rates and head losses. The design procedure involving multiple demand patterns will be described later in this chapter.

4.2.5 VALVES AND PUMPS

The locations of line valves in a network are usually pre-specified at strategic places, for example, at junctions so that sections of pipes may be isolated when necessary, near washouts for flushing purposes, etc. These valves are not usually included in a design model. Their existence can be accounted for as minor losses. However, flow control valves are often required in water distribution networks and they are often included mathematically in the design optimization model, for example, a non-return valve which is usually fitted at the downstream end of a pump to prevent back water effect that can damage the pump. The expression for this type of valve is as follows (Bhave, 1991).

$$q_{ij} = \begin{cases} \dfrac{H_i - H_j}{K_f^{0.54} \left| H_i - H_j \right|^{0.46}} & H_j \leq H_i \\ 0 & H_j > H_i \end{cases} \tag{4.11}$$

Another type of flow control valve is Pressure Reducing/Regulating Valve (PRV). It can be mathematically presented as (Bhave, 1991)

$$q_{ij} = \begin{cases} \dfrac{H_{prv} - H_j}{K_f^{0.54} \left| H_{prv} - H_j \right|^{0.46}} & H_j \leq H_{prv} \leq H_i \\ \dfrac{H_i - H_j}{K_f^{0.54} \left| H_i - H_j \right|^{0.46}} & H_j < H_i < H_{prv} \\ 0 & H_j > H_{prv} \end{cases} \tag{4.12}$$

in which H_{prv} is the PRV setting.

A network may also need additional energy to satisfy the minimum head requirement at the downstream end of the network. For this purpose, a pump may be included in the design optimization model and the following *Power* equation may be used to represent the pump in the model (Walski et al., 2003).

$$H_p = H_o - c_p Q_p^{e_p} \tag{4.13}$$

in which H_p is the pump head or the head difference across the pump; H_o is the shutoff head or the pump head at zero flow; Q_p is the pump flow; c_p and e_p are coefficients, which are usually specified by the pump manufacturer to describe the shape of the pump curve, i.e. the relationship between the pump head and pump flow. The location of the pump may be pre-specified by the designer or incorporated into the optimization procedure.

When considering flow control valves and pumps in the design optimization model, the problem becomes one of the design-operation types. Multiple loadings need to be considered since the designer has to decide the status of the valves and pumps at each of the loading conditions. The above expressions for valves and pumps are presented in this thesis for completeness. However, the present research is not particularly concerned with these components.

4.3 NETWORK ANALYSIS

In the analysis of water distribution networks, the constitutive equations can be formulated in several ways. The analysis problem typically has three kinds of variables, which are the pipe flow rates q_{ij}, the nodal heads H_i and the pipe head losses h_{ij}. The constitutive equations, therefore, can be set up and solved in terms of one kind of variable, i.e. the pipe flow rates, the nodal heads, or the corrective loop flow rates, which may then be used to express the other variables.

4.3.1 SYSTEMS OF EQUATIONS

4.3.1.1 PIPE FLOW RATES AS UNKNOWNS

The continuity and head loss equations in Equations (4.1) and (4.4), respectively, are expressed with the flow rates, q_{ij}, as the independent variables or unknowns. It follows that the unknowns in the loop and path equations in Equations (4.7) and (4.9), respectively, are the q_{ij}. Following Jeppson (1976) and Bhave (1991), these equations will be called q-equations or the q-system of equations.

4.3.1.2 NODAL HEADS AS UNKNOWNS

The Hazen-Williams equation can be expressed in terms of the pipe head loss.

Substituting h_{ij} with the nodal heads from Equation (4.10), Equation (4.4) may be written as

$$q_{ij} = \frac{\alpha C_{ij} D_{ij}^{2.63} (H_i - H_j)^{0.54}}{L_{ij}^{0.54}} \;, \quad \forall ij \in IJ \tag{4.14}$$

in which, here, $\alpha = 0.2785$ in S.I. units. However, it is perhaps better to determine the sign of q_{ij} according to whether H_i is greater or less than H_j, in which case Equation (4.15) below may be used instead.

$$q_{ij} = \frac{\alpha C_{ij} D_{ij}^{2.63} sign(H_i - H_j) \,|\, H_i - H_j \,|^{0.54}}{L_{ij}^{0.54}} \;, \quad \forall ij \in IJ \tag{4.15}$$

Using Equation (4.15), the continuity equation, Equation (4.1), may be written in terms of the nodal heads.

$$\alpha \sum_{j \in (NU_n \cup ND_n)} \frac{C_{nj} D_{nj}^{2.63} sign(H_n - H_j) \,|\, H_n - H_j \,|^{0.54}}{L_{nj}^{0.54}} = q_n \;, \quad n = 1, \ldots, (NN-1) \tag{4.16}$$

These equations, which are based on the nodal heads, will be called H-equations. Since the head loss in every pipe is considered explicitly, the H-equations describe the flow in a pipe network completely, hence the loop or path equations are no longer needed. Also, there will be as many continuity equations as unknown nodal heads; usually at least one nodal head will be constant and known.

4.3.1.3 CORRECTIVE LOOP FLOW RATES AS UNKNOWNS

To express the pipe flow rates in terms of a corrective flow rate around each loop in any iterative scheme, the following equations are used.

$$q_{ij}^{(n)} = q_{ij}^{(n-1)} + \sum_{l \in l_{ij}} \Delta q_l^{(n)} \;, \quad \forall ij \in IJ \tag{4.17}$$

in which $q_{ij}^{(n-1)}$ in an estimated flow rate; $\Delta q_{ij}^{(n)}$ is a correction to be applied, taking into account the direction of the flow, to all flows in loop l; $q_{ij}^{(n)}$ is the corrected flow rate. The bracketed superscripts indicate the number of iteration. Finally, l_{ij} consists of all loops sharing link ij. The unknowns in Equations (4.17) are therefore Δq_l. Equations (4.17) may be inserted in the head loss equation to give a complete set of equations based on the Δq_l, which are called Δq-system of equations.

$$h_{ij}^{(n)} = \frac{\alpha L_{ij} \left[q_{ij}^{(n-1)} + \sum_{l \in l_{ij}} \Delta q_l^{(n)} \right]^{1.852}}{C_{ij}^{1.852} D_{ij}^{4.87}}, \quad \forall ij \in IJ \tag{4.18}$$

This system contains *NLP* equations, one for each loop, and each Δq_l corresponds to a loop. Also, like the q-system, this system requires the loop and path equations but does not directly involve the continuity equations. To set up the Δq equations, the initial flow estimates, $q_{ij}^{(0)}$ must satisfy continuity. Thereafter, successive iterates in the analysis will also satisfy continuity.

4.3.2 ANALYSIS METHODS

There are several analysis methods for water distribution networks. Based on the relationship between outflows and pressure in the system, analysis methods are divided into two main groups namely Demand Driven Analysis and pressure or Head Dependent Analysis. These methods are presented next.

4.3.2.1 DEMAND DRIVEN ANALYSIS

The Demand Driven Analysis (DDA) method has been widely used in the water industry for many years. The method assumes that the demands in the system are fully satisfied regardless of the pressure in the system. Three main numerical approaches to the network analysis problem based on the DDA method are presented next.

HARDY-CROSS METHOD

The Hardy-Cross method is the oldest method for systematic solution of water distribution networks (Jeppson, 1976, pp.145-150). It is the most commonly used for hand computations due to its simplicity, although many computer programs exist for its execution by digital computers. This method solves the constitutive equations sequentially in each iteration. Also, each equation is solved for a single variable only, while keeping the other variables fixed. The method is demonstrated here for the Δq-equations.

Equation (4.17) can be written for each loop.

$$q_{ij}^{(n)} = q_{ij}^{(n-1)} + \Delta q_l^{(n)} , \quad \forall l, \ \forall ij \in IJ \tag{4.19}$$

When considering the links in loop l, the corrections due to other loops sharing the same links are neglected. The resulting head loss equations are therefore

$$h_{ij}^{(n)} = \frac{\alpha L_{ij} \left[q_{ij}^{(n-1)} + \Delta q_l^{(n)} \right]^{1.852}}{C_{ij}^{1.852} D_{ij}^{4.87}} , \quad \forall l, \ \forall ij \in IJ \tag{4.20}$$

and the loop equations are

$$\sum_{ij \in IJ_l} h_{ij}^{(n)} = 0 , \quad \forall l \tag{4.21}$$

Using the first order Taylor's series expansion of the head loss equation, the corrective loop flow rates are therefore

$$\Delta q_l^{(n)} = \frac{-\sum_{ij \in IJ_l} h_{ij}^{(n-1)}}{1.852 \sum_{ij \in IJ_l} \frac{h_{ij}^{(n-1)}}{q_{ij}^{(n-1)}}} , \quad \forall l \tag{4.22}$$

in which

$$h_{ij}^{(n-1)} = \frac{\alpha Lij(q_{ij}^{(n-1)})^{1.852}}{C_{ij}^{1.852} D_{ij}^{4.87}} \tag{4.23}$$

The resulting values of $\Delta q_l^{(n)}$ of Equation (4.22) are then used in Equation (4.19) to obtain the new flow estimates. Thus completes one iterative cycle. A new cycle can be started by setting up and solving Equations (4.21) again. This iterative cycle is repeated until the loop correction values become insignificant and the loop and path equations are satisfied. See, for example, Jeppson (1976, pp.147) for a step-by-step implementation of the Hardy-Cross method.

NEWTON-RAPHSON METHOD

Martin and Peters (1963) were the first to propose the application of the Newton-Raphson method for the analysis of water distribution network. The Newton-Raphson method is an iteration procedure for finding the root(s) of a function, F. For example, for a single variable function $F(x) = 0$, the value of x can be found by using an *additive* correction, Δx, in the iteration so that the function can be written as

$$F(x) = F(x + \Delta x) = 0 \tag{4.24}$$

Using the first order Taylor's series expansion, Equation (4.24) becomes

$$F(x^{(n)}) + F'(x^{(n)})\Delta x^{(n)} = 0 \tag{4.25}$$

in which the superscript indicates the nth iteration. The higher order of the expansion is neglected assuming that the Δx correction is relatively small compare to x and therefore its higher order will be smaller still, hence insignificant. Equation (4.25) can be rearranged to solve for Δx as shown below.

$$\Delta x^{(n)} = -\frac{F(x^{(n)})}{F'(x^{(n)})} \tag{4.26}$$

The value of x for the next iteration is obtained from

$$x^{(n+1)} = x^{(n)} + \Delta x^{(n)} \tag{4.27}$$

For functions with more than one variable, Equation (4.25) can be extended for a system of equation in the form of

$$\underline{F}(\underline{x}^{(n)}) + J^{(n)}\Delta\underline{x}^{(n)} = 0 \tag{4.28}$$

in which $\underline{F}$ is the vector of the function values for the system of simultaneous equations at the point $\underline{x}$; $\underline{x}$ is the vector of the variables; J is the Jacobian matrix, which is the matrix of the first partial derivatives of each F with respect to each of the x's; $\Delta\underline{x}$ is the vector of the $\underline{x}$ correction. By rearranging Equation (4.28), the corrections for each of the x's are therefore

$$\Delta\underline{x}^{(n)} = -(J^{(n)})^{-1}\underline{F}(\underline{x}^{(n)}) \tag{4.29}$$

Usually, inversion of J is computationally expensive and is avoided by pre-multiplying both sides of Equation (4.29) by $J^{(n)}$, which gives

$$(J^{(n)})\Delta\underline{x}^{(n)} = -\underline{F}(\underline{x}^{(n)}) \tag{4.30}$$

Equation (4.30) is solved for $\Delta\underline{x}^{(n)}$ using, for example, the Gaussian elimination technique, whose values are then used to find the values of each x for the next iteration, i.e.

$$\underline{x}^{(n+1)} = \underline{x}^{(n)} + \Delta\underline{x}^{(n)} \tag{4.31}$$

The symmetry of J may be exploited for greater computational efficiency. The Newton-Raphson method described above may be used to solve the H-, q-, Δq or H-q equations in which the heads at demand nodes and the link flows are calculated at the same time. The convergence rate of the Newton-Raphson method is much faster compare to the Hardy-Cross method since calculations are done simultaneously rather than sequentially. Also, it can be shown that in the

(n+1)th iteration, the error is proportional to the square of the error in the nth iteration. Each subsequent reduction in error in the Newton-Raphson method is proportional to the square of the previous error, and hence the convergence property is *quadratic* (Bhave, 1991, pp. 233).

The method described by Martin and Peters (1963) was applied to networks having pipes and reservoirs only. Shamir and Howard (1968) have described a generalised formulation of the Newton-Raphson method to include other elements like pumps and valves. Interested readers can refer to their paper for further details on the Newton-Raphson method.

LINEAR THEORY METHOD

The linear theory method was developed by Wood and Charles (1972). They suggested the Hazen-Williams equation [Equation (4.4)] can be approximated linearly as follows.

$$h_{in}^{(n)} = K_{ij}(q_{ij0}^{(n-1)})^{0.852} q_{ij}^{(n)} = K'^{(n)}_{ij} q_{ij}^{(n)} \ , \ \ \forall ij \qquad (4.32)$$

in which (n) is the iteration number; q_{ij0} is the approximate discharge in pipe ij; K_{ij} and K'_{ij} are the actual and the modified pipe resistance for pipe ij, respectively. When the value of q_{ij0} approaches the actual discharge, Equation (4.32) is the exact expression of the head loss. K_{ij} is given by

$$K_{ij} = \frac{\alpha L_{ij}}{C_{ij}^{1.852} D_{ij}^{4.57}} \ , \ \ \forall ij \qquad (4.33)$$

Once all the loop equations have been set up using the above approximate equations together with the nodal continuity equations, the iteration begins by setting the value of q_{ij0} to unity so that $K'^{(1)}_{ij} = K_{ij}^{(1)}$ and solving the nodal continuity and loop equations simultaneously to find the values of the initial approximate discharge, $q_{ij0}^{(1)}$. The values of $K'^{(2)}_{ij}$ are calculated from

$$K'^{(2)}_{ij} = K_{ij} q^{(1)}_{ij0} \ , \quad \forall ij \tag{4.34}$$

The results are substituted back into the nodal continuity and loop equations to obtain the value of pipe flow rates for the next iteration. Wood and Charles (1972) found that the average of two successive trials gave a result very close to the final value of flow rate. They, therefore, suggested that the value of the discharge and the modified pipe resistance, $K^{'}_{ij}$, for the subsequent iterations are computed using the average value of the previous two sets of flow rates, i.e.

$$q^{(n)}_{ij0} = \frac{q^{(n-1)}_{ij} + q^{(n-2)}_{ij}}{2} \ , \quad \forall ij, n = 3,4,5,... \tag{4.35}$$

Wood and Charles (1972) concluded that the above method has a better convergence rate than the Newton-Raphson method. Also, they found that iteration always converge to a solution. Their Linear Theory method described above is based on q-equations. Isaacs and Mills (1980) applied the method using the H-equations, but concluded that the H-equations are better suited for networks with some known heads, whereas the q-equations applies better to networks with known external flows.

4.3.2.2 HEAD DEPENDENT ANALYSIS

The DDA method does not take into consideration the relationship between the nodal outflows and the pressure within the system. It is perhaps necessary to differentiate between *outflow* and *demand* at this point. Outflow is the actual amount of water yielded by the network, while demand is the required amount of water to be extracted from the network. The DDA method is satisfactory when the network is under normal condition, i.e. when the pressure in the network is sufficient. However, when the pressure drops below the required level, network analysts would have no reliable information to determine how much outflow would be delivered by the system under the available pressure regime. In this situation some customers would receive reduced supplies and, in the worst scenario, they might not receive any supply at all (Ackley et al., 2001; Tanyimboh et al., 2003). The drop in pressure in the distribution network can be triggered by many factors. Excessive abstraction at one demand node, for example, in a fire fighting situation, may cause the pressure in the neighbouring abstraction points to drop below the required level.

Head Dependent Analysis (HDA) has been suggested to be superior to the DDA method, particularly for networks under subnormal operating conditions. It is well known that outflows from a water distribution network are dependent upon the pressure within that system and, therefore, the DDA assumption that demands are always satisfied regardless of the pressure in the system is often inappropriate. HDA takes into consideration the pressure dependency of nodal outflows, and in consequence, the results are more realistic. Nevertheless, this method is not yet commonly used in the water industry since more research and verification of the true relationship between network pressure and nodal outflows are still necessary.

Bhave (1981) was probably the first to consider the nodal heads and flows simultaneously in the analysis of deficient water distribution network. In his study, however, no head-outflow relationship is given and the analysis is carried out using the DDA analysis iteratively. The demand nodes are given a set of criteria based on the result of the DDA analysis and the outflows are updated based on these criteria before they are used as input data for the next iteration. Once certain criteria have been satisfied, the iteration is stopped and the result of this approach, termed as Node Flow Analysis (NFA) method, is reached. Ackley et al. (2001), on the other hand, used a mathematical programming formulation based on the maximization of the nodal outflows.

Other researchers have proposed several assumed head-outflow relationships for the HDA analysis. For example, Wagner et al. (1988b) and Chandapillai (1991) suggested the following parabolic function

$$\frac{q_n^{avl}}{q_n^{req}} = \left(\frac{H_n - H_n^{\min}}{H_n^{des} - H_n^{\min}} \right)^{\frac{1}{n_n}} \ , \quad H_n^{\min} \le H_n < H_n^{des} \tag{4.36}$$

in which $\frac{q_n^{avl}}{q_n^{req}}$ is the demand satisfaction ratio (DSR) of node n, q_n^{avl} and q_n^{req} are the actual outflow that can be delivered by the system and the required outflow or demand, respectively. H_n is the actual head at node n, H_n^{min} is the nodal head at node n below which there would be no outflow and H_n^{des} is the desired head at node n, above which the outflow would be equal to the demand. The DSR is set to zero if H_n is less than H_n^{min} or 1.0 if H_n reaches H_n^{des}. Values of the exponent parameter, n_n, are thought to lie between 1.5 and 2 (Gupta and Bhave, 1996).

Another example of the head-outflow relationship is the following function proposed by Fujiwara and Ganesharajah (1993).

$$\frac{q_n^{avl}}{q_n^{req}} = \frac{\int_{H_n^{\min}}^{H_n} \left(H - H_n^{\min}\right)\left(H_n^{des} - H\right) dH}{\int_{H_n^{\min}}^{H_n^{des}} \left(H - H_n^{\min}\right)\left(H_n^{des} - H\right) dH}, \quad H_n^{\min} \le H_n < H_n^{des} \tag{4.37}$$

Another HDA approach termed the *source head method* was proposed by Tanyimboh and Templeman (1998). Instead of considering each demand node individually, the method approximates the total outflow delivered by the entire system. Hence, the following head-outflow relationship is used in this method.

$$\frac{q_s^{avl}}{q_s^{req}} = \left(\frac{H_s - H_s^{\min}}{H_s^{des} - H_s^{\min}}\right)^{\frac{1}{2}}, \quad H_s^{\min} \le H_s < H_s^{des} \tag{4.38}$$

In the above expression, q_s^{avl} and q_s^{req} are the total available flows supplied by the network and the required flows that must be provided by the source, respectively; $H_s^{\min}$ is the head at the source above which outflow just begins at any node in the network or can be set to the minimum ground level elevation for demand nodes while H_s^{des} is the head at the source above which all the demands would be fully satisfied. H_s is the actual head at the source.

The method is capable of simulating the behaviour of networks under deficient conditions. However, the approximation tends to underestimate the total outflow delivered by the reduced network. This is probably caused by the negative heads from the DDA analysis that are used in the approximation producing a high value of H_s^{des} which in turn underestimate the value of the total flow delivered.

Recently, Tanyimboh and Templeman (2004) suggested the following nodal outflow function.

$$\frac{q_n^{avl}}{q_n^{req}} = \frac{\exp(\alpha_n + \beta_n H_n)}{1 + \exp(\alpha_n + \beta_n H_n)} \tag{4.39}$$

The values of the parameters α_n and β_n are determined by relevant field data for the node under consideration. In the absence of field data, Tanyimboh and Templeman (2004) suggested default values for the DSR as 0.01 and 0.999 for situations when the H_n is less than H_n^{min} and when H_n reaches H_n^{des}, respectively. These DSR values give two simultaneous equations whose solution gives the following expressions for the parameters α_n and β_n.

$$\alpha_n = \frac{-4.595 H_n^{des} - 6.907 H_n^{\min}}{H_n^{des} - H_n^{\min}} \tag{4.40}$$

$$\beta_n = \frac{11.502}{H_n^{des} - H_n^{\min}} \tag{4.41}$$

Unlike the previous head-outflow relationships, Equation (4.39) does not need additional conditions for the case when the nodal head, H_n, is less than or equal to H_n^{min} or when H_n reaches or exceeds H_n^{des}.

The HDA method requires the constitutive equations to be expressed in terms of the heads at the nodes, i.e. H-systems of equations. The nodal continuity equations are modified to include the head-outflow relationship as depicted by Equations (4.42) below and the set of equations can then be solved using one of the numerical approaches mentioned above, e.g. the Newton-Raphson method.

$$\alpha \sum_{j \in (NU_n \cup ND_n)} \frac{C_{nj} D_{nj}^{2.63} sign(H_n - H_j) | H_n - H_j |^{0.54}}{L_{nj}^{0.54}} - q_n^{avl} = 0, \; n = 1, ..., (NN - 1) \quad (4.42)$$

In the above equations, q_n^{avl} is the head-dependent outflow.

In this thesis, the HDA analysis is carried out using a Fortran program called PRAAWDS (Program for the Realistic Analysis of the Availability of Water in Distribution Systems) (Tahar et al., 2002; Tanyimboh et al., 2003), which calculates the actual flow delivered under normal and subnormal pressure conditions. The program allows the user to choose which head-outflow relationship to be used in the analysis and, apart from the Source Head method head-outflow relationship, all of the above mentioned relationships are available to choose. PRAAWDS can also perform a DDA analysis. The DDA analysis in this thesis, however, is carried out using EPANET (Rossman, 2000).

4.3.3 STEADY STATE ANALYSIS AND EXTENDED PERIOD SIMULATION

Based on the period or duration, analysis methods are divided into two categories, i.e. steady state analysis, which is carried out in a very short period of time with constant demand values, and extended period analysis, which is done over a longer period, usually over 24 or 48 hours under varying demand conditions. In extended period analysis the analysis is divided into several time intervals (typically with duration of 15 minutes to 1 hour) and a sequence of steady state analysis are performed at these intervals. Extended period analysis is important for most real water distribution systems, for example in analysing networks with storage tank(s). A storage tank in a network discharges water during peak time and acts as a demand node when the demands in the network are low and the tank is filling up. Therefore, the water level in the tank needs to be checked to make sure that it is back to its original level at the end of the 24- or 48-hour, ready for the next period. Also, pump and valve settings are likely to change within the 24-hour period. This thesis, however, is concerned only with steady state analysis. Interested readers may consult, for example, Bhave (1991) for a methodology of performing extended period analysis.

4.4 SUMMARY AND CONCLUSIONS

The equations that govern the flow in pipe networks have been detailed in this chapter. Several ways in which these equations can be formulated have also been presented in this chapter. Two network analysis methods have been discussed briefly and the limitation of the commonly used DDA method is highlighted. Although the HDA seems to provide a better alternative to the DDA method, more research are required to further validate its applicability to the real water networks. A more general expression of the actual head-outflow relationship is required. Nevertheless, the available head-outflow relationships proposed by several researchers are considered appropriate for use in the present study.

CHAPTER 5 PERFORMANCE MEASURES AND LEAST COST DESIGN OF PIPE NETWORKS

5.1 INTRODUCTION

The design of water distribution networks is often viewed as a least-cost optimization problem with pipe diameters being the decision variables. Most traditional design methods put the emphasis on economy and hence the resulting designs are tree-type or branch networks. Suppose that an upstream pipe in such network fails and has to be taken out of service for repair or replacement. It is obvious that tree-type networks do not have the capabilities to sustain supply in such an event. The supply to the nodes downstream of the pipe will therefore cease completely since there are no alternative routes for the water to go through to reach these nodes.

In recent years, interest in the optimum design of water distribution networks has increased. Most of the studies are concern with the optimization of looped water networks. These networks are highly desired in urban water supply systems since the addition of redundant links in the network increases its flexibility and helps maintain the network performance in emergency situations such as pipe failure and fire fighting. However, the redundant links in a loop network also contribute to the high cost of the network. Therefore, several researchers have proposed that the requirements for the optimum design of water distribution networks are not only the economical aspect of the design but also the high level of performance of the distribution network. The optimization of this problem is highly complex. This is partly due to the difficulty in quantifying the performance of water distribution networks.

In this chapter, several measures for water distribution network performance proposed by several researchers are presented and discussed. Several methods for optimizing the network performance are also examined in this chapter. From the discussion of these methods, it will become clear that the inclusion of performance measure in the optimization of water distribution networks create complications and make the design problem extremely difficult to solve. Finally, problem formulation for the least cost design of pipe networks is presented followed by several of its solution methods.

5.2 PERFORMANCE MEASURES FOR WATER DISTRIBUTION NETWORKS

Adding loops to a distribution network is one way of increasing the flexibility of the network as mentioned earlier. These loops reduce the possibility of some demand nodes being completely cut off from the rest of the network, hence improving the performance of the network. Following a failure in a looped network, the behaviour of the reduced network is, in general, not predictable without hydraulic simulation of the reduced network. This problem contributes to the difficulty in quantifying the performance of water distribution networks. Also, Mays (1989) and Fujiwara and Ganesharajah (1993) have shown that the nodal heads and flows should be considered simultaneously for more accurate assessment of deficient-network performance. However, thanks to the development of the HDA method, this objective is now achievable.

Unavailability of components may be due to maintenance or failure. Failures in water distribution systems, according to Mays (2000), may be classified into two major categories, i.e. performance (hydraulic) failure and component (mechanical) failure. Performance failures occur when there is a shortage in pressure or flow at one or more demand nodes. This type of failure may arise due to component failure. Excessive abstraction somewhere in the network, for example, for fire fighting, may also lead to a reduction in pressure throughout the network. External factors such as power supply and availability of water at the source are also involved in wider performance assessment (Tanyimboh, 1993). This thesis, however, is concerned only with failures due to excessive demands or pipe failures.

Reliability is probably the most commonly used parameter in quantifying the performance of water distribution networks. Due to the nature of failures discussed above, the reliability can also be differentiated into mechanical and hydraulic reliability. Mechanical reliability measures the probability that the components in the system are operational at any time. It is not a true measure of the performance of water distribution networks since it does not take into account the effect of failure in terms of the amount of water delivered by the network. Hydraulic reliability, on the other hand, quantifies the performance of water distribution network by considering the probability that the system can deliver the right amount of water at the right pressure. Also, since the hydraulic performance of a network is influenced by the performance of its individual components, the hydraulic reliability should include a measure of mechanical reliability in the system too (Mays, 1989).

Unfortunately, there are no universally accepted definitions for reliability of water distribution systems (Mays, 2000). Many researchers have proposed several measures of reliability. Some of these measures are discussed next followed by some ways of optimizing their values.

5.2.1 SOME RELIABILITY MEASURES

Fujiwara and De Silva (1990) used the minimum total shortfall in the flow delivered to measure water distribution system reliability. Following Carey and Hendrickson (1984), they defined the system reliability, R, as follows

$$R = 1 - \frac{\text{Expected minimum total shortfall in flow}}{\text{Total demand}} \tag{5.1}$$

They limited their study to gravity networks in which the network consists only of pipes. The Linear Programming Gradient (LPG) method of Alperovits and Shamir (1977) was employed in the design optimization process (this design method is discussed in brief later in this chapter) and, as a result, the links in the resulting designs may consist of more than one segment with different pipe sizes. The expected minimum total shortfall in flow for the evaluation of reliability was obtained by considering single link (including all its segmental pipes) failure events. Fujiwara and De Silva (1990) argued that since the probability that any link in the network available at any one time was very high, the probability of more than one link failed simultaneously was therefore very small.

The value of the expected minimum total shortfall in flow was calculated as the complement of the expected maximum total outflow, which was obtained using an optimization procedure with nodal continuity constraint. This procedure was favoured over network simulation due to its capability to produce results in a shorter period of time. However, this advantage was achieved at the expense of the hydraulic consistency of the network, i.e. violation of the conservation of energy around each loop in the network and violation of the pressure requirements at each demand node. As a solution, Fujiwara and De Silva (1990) introduced flow capacity, whose value was equal to the optimal link flow obtained from the LPG method in the design process, as the maximum flow that could be carried by each link. The definition of the flow capacity constraint, however, is not clear. The justification for its use is therefore questionable. Although, its application ensures that the pressure requirements at all demand nodes are satisfied, the energy constraint around each loop is still not accounted for.

Details of Fujiwara and De Silva (1990) reliability function are as follows. For each link, they assumed that the probability of that link is available is given by the following function.

$$a_m = \prod_{k:L_{mk}>0} a_{mk}, \quad \forall m \tag{5.2}$$

In the above expression, a_m is the availability or the probability that link m is available; a_{mk} is the availability of segment k in link m; L_{mk} is the length of segment k in link m. The probability that segment k is available is given by the following function

$$a_{mk} = \frac{\tau_{mk}}{\tau_{mk} + \mu_{mk}} \tag{5.3}$$

in which τ_{mk} and μ_{mk} are the expected rate of repair and failure of segment k, respectively. However, it is not clear how Fujiwara and De Silva (1990) obtained the values of these expected rates. From Equation (5.2), it follows that the availability of all pipes in the network, $p(0)$, is given by

$$p(0) = \prod_{m=1}^{M} a_m \tag{5.4}$$

and the probability that only link m is unavailable, $p(m)$, is

$$p(m) = (1 - a_m)\frac{p(0)}{a_m} \tag{5.5}$$

in which $1 - a_m = u_m$ is the unavailability of link m. Therefore

$$p(m) = p(0)\frac{u_m}{a_m} \tag{5.6}$$

Let $T^*(m)$ be the maximum total outflow when link m is unavailable as obtained by the optimization procedure. The ratio of the minimum total shortfall in flow to the total demand when link m is unavailable is given by

$$1 - \frac{T^*(m)}{T(0)} \tag{5.7}$$

in which $T(0)$ is the total demand. Therefore, the ratio of the expected minimum total shortfall in flow to the total demand over the entire failure scenarios is

$$\sum_{m=0}^{M} p(m)\left(1 - \frac{T^*(m)}{T(0)}\right) \tag{5.8}$$

Finally, the reliability of the network is given by

$$R = 1 - \sum_{m=0}^{M} p(m)\left(1 - \frac{T^*(m)}{T(0)}\right) \tag{5.9}$$

Bao and Mays (1990) defined nodal reliability as a function of the probabilities of the heads at demand nodes being equal or above the minimum required level. Their work was based on the DDA method and therefore the demands in the network are always considered satisfied. They take into account the uncertainties of pipe roughness coefficient, nodal demand and the required nodal head at demand nodes and their random values were generated using Monte Carlo simulation by first specifying the type of probability distribution for each of them. The interested readers may refer to Bao and Mays (1990) for the type of probability distributions used. The value of the available nodal head, on the other hand, was obtained using network simulation. The effects of mechanical failures were not considered in the hydraulic reliability calculations of Bao and Mays (1990). This may cause overestimation of the reliability values. Three nodal reliability functions were proposed and one of these is reproduced below as Equation (5.10).

$$R_n = p\left(H_n > H_{\min,n}\right) = \int_{H_{\min,n}}^{\infty} f\left(H_n\right) dH_n \,, \quad \forall n \tag{5.10}$$

H_n, $H_{\min,n}$ and $f(H_n)$ are the pressure head, the minimum required head and the probability density function of the pressure head at node n, respectively. R_n is the nodal reliability. In order to establish the distribution of the nodal pressure head, the iterative process of random sampling of demand, required pressure head and pipe roughness and the subsequent hydraulic simulation must be repeated a large number of times to ensure a reasonable accuracy of the results. The computational requirements could therefore be very large even for simple networks. There is also a difficulty in selecting the type of probability distribution to be used in generating the random values since reliability data for water distribution networks are usually very limited. On top of that, it is difficult to estimate the parameters, e.g. coefficient of variation, etc, for the distribution regardless of the type of distribution used. Bao and Mays (1990) proposed an alternative approach by setting the values of the nodal demand, required pressure head and pipe roughness as constants and set the nodal reliabilities to one if the result of the hydraulic simulation satisfied the required pressure head requirements and zero if not. This approach, however, may underestimate the nodal reliability since the situation in which the demand is partly satisfied due to a reduced nodal head is not accounted for.

For the system reliability, Bao and Mays (1990) proposed three heuristic definitions, i.e. system reliability as: the minimum nodal reliability in the system, the arithmetic average of all the nodal reliabilities, or the weighted average of all nodal reliabilities weighted by the demand at the node. The latter of these definitions is presented below.

$$R_{sw} = \frac{\sum_{i=1}^{I} R_{ni}\bar{q}_{i0}}{\sum_{i=1}^{I} \bar{q}_{i0}} \tag{5.11}$$

in which R_{sw} is the system reliability as the weighted mean of all nodal reliabilities; R_{ni} is the nodal reliability of node i; and $\bar{q}_{i0}$ is the average demand at node i. The definition of system reliability as the minimum nodal reliability may prove too conservative since deterioration in a network can occur locally with insignificant effect to the rest of the network. Bao and Mays (1990) have stated that the value of the system reliability obtained as the arithmetic average of the nodal reliabilities may be significantly different than the system reliability as the weighted average of nodal reliabilities in real networks. However, they did not offer any suggestion as to which is more appropriate to use.

Throughout this thesis the reliability measure proposed by Tanyimboh (1993) is used. It is defined as the time-averaged value of the ratio of the flow delivered to the flow required (Tanyimboh and Templeman, 2000, Tanyimboh and Sheahan, 2002). By assuming a constant demand value, the reliability function can be written as

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

$$+ \frac{1}{2}\left(1 - p(0) - \sum_{m=1}^{M} p(m) - \sum_{m=1}^{M-1}\sum_{n=m+1}^{M} p(m,n) - \cdots \right) \tag{5.12}$$

in which R is the reliability; $p(0)$ is the probability that no links is unavailable; $p(m)$ is the probability that only link m is unavailable and $p(m, n)$ is the probability that only links m and n are unavailable. $p(0)$ and $p(m)$ are calculated from Equations (5.4) and (5.6), respectively, while $p(m, n) = p(0)(u_m / a_m)(u_n / a_n)$. $T(0)$, $T(m)$ and $T(m, n)$

are the respective total flows supplied with no links unavailable, only link m unavailable, and only links m and n unavailable. Finally, M is the number of links while T represents the total demand. Equation (5.12) can be used to calculate nodal or system reliability. For nodal reliability, $T(0)$, $T(m)$ and $T(m, n)$ becomes the nodal outflow with no links unavailable, only link m unavailable, and only links m and n unavailable, while T represents the demand at the node in question. It may be noted that T is not always equal to $T(0)$, especially for old distribution networks where the increasing demands exceeds the network capacity. This condition is also found in a fire fighting situation where the large abstraction often cause the actual outflows to be lower than the demand.

There are two main terms in Equation (5.12), which correspond to the two pairs of large parentheses. The first term of the equation corresponds to the basic definition of hydraulic reliability as stated above. The second term is a correction function whose value approaches zero as more and more multiple-component failure simulations are included in the first part. The derivation of this Equation is provided in Tanyimboh and Sheahan (2002) and reproduced in this thesis in Appendix A2.

It also needs to be mentioned that Equation (5.12) is applicable to pipes, pumps and valves in general. In this thesis, however, only pipes are considered and throughout this thesis the Cullinane et al. (1992) function for estimating the availability of a pipe is used. The function is an approximation based on the data presented by Mays (1989) and Walski and Pelliccia (1982) and is in the form of

$$a_m = \frac{MTBF_m}{MTTR_m + MTBF_m} = \frac{0.21218 D_m^{1.462131}}{0.00074 D_m^{0.285} + 0.21218 D_m^{1.462131}}, \quad \forall m \qquad (5.13)$$

in which $MTBF_m$ and $MTTR_m$ are the mean time between failures and the mean time to repair of pipe m, respectively; D_m is the diameter of pipe m. It needs to be emphasized that the above availability function is an approximation based on limited data. Its use in this thesis is for easier comparison between results generated in this thesis and available results in literature. Other availability functions, e.g. Fujiwara and De Silva (1990), may be used if considered appropriate. Tabesh (1998) also has listed several pipe availability functions.

To simplify the reliability calculations, only single-pipe failure situations are considered in this thesis. The justification for this approach is that for a repairable component such as pipe, the probability of having two or more pipes fail at the same time is very small (Fujiwara and De Silva, 1990). In reality, several pipes may need to be taken out of service depending on the availability and arrangement of valves in the vicinity of the broken pipe. The reliability of Equation (5.12) allows any number of pipes to be isolated simultaneously if necessary.

Equation (5.12) also requires the outflows in the reduced networks to be calculated. These outflows can be obtained using the HDA method. However, if the DDA method was used for analysing the reduced networks, the calculation could be approximated using the source head method as described in the previous chapter. Appraisal and comparison of the source head method for calculating reliability has been carried out by Tanyimboh et al. (2001). Although a slight underestimation is present in the source head method, the appraisal shows that the method is comparable to other reliability methods in the sense that it is capable of distinguishing networks with various levels of reliability.

In this thesis the reliability of water distribution networks is calculated using a constant demand value. This approach is considered sufficient for the present research since steady state analysis is used in the hydraulic simulations in this study. The inclusion of variations in demands in reliability assessment is currently an area of active research. Some researchers, e.g. Bao and Mays (1990), Gargano and Pianese (2000), have tried to incorporate the random nature of demands in the reliability analysis by using the Monte Carlo simulation to generate the random demand data based on some given probability distribution. The shortcomings of this method have been explained earlier in this section. Surendran et al (2005) used the peaking factor, which is the ratio of the peak demand to the average demand, to account for the variation in demands in the reliability calculation. They also used statistical modelling of demands to assess the confidence levels of the reliability value by predicting the probability for which the critical demands used in the reliability assessment will be exceeded. More studies, however, are required to determine the appropriate probability distribution of demands to be used in the statistical model. Interested readers may refer to their original paper for further details.

In addition to the variation in demands, the random nature of failure of the components in the network and the random occurrence of fire demands and their locations all affect the reliability of the water distribution networks. All of these issues need to be considered in order to obtain accurate reliability values. The process is therefore very complicated and tedious for large networks. The use of the Monte Carlo simulation to generate the random data is intended to help reduce the complexity of the calculations. The Minimum Cut Set method (e.g. Su et al., 1987) is another technique to help simplify the reliability calculation. The minimum cut set is a set of components in the system which causes the system to fail only when all of the components within the set fail. To obtain an accurate reliability value, all the minimum cut sets in the system must be identified. This could be very tedious for large networks. The failure probability of each set is then obtained as the product of the failure probability of each component in the set. While the failure probability of the system is obtained by adding the failure probability of all the minimum cut sets in the system. The reliability of the system is then obtained as the complement of the system probability of failure. This technique, as such, does not take into account the random nature of normal and fire demands in the distribution system.

Many other performance measures are available in addition to those described above. Interested readers may refer to Goulter (1995), Mays (1989, 2000) and the references therein for a list and details of other measures.

5.2.2 RELIABILITY OPTIMIZATION

Due to the above mentioned problems in calculating the reliability value, optimizing the value of the reliability is an even more difficult challenge. Yet, obtaining the optimum value of reliability for water distribution networks remains highly desirable. Many researchers have tried to optimize the value of reliability indirectly because of calculational effort involved in determining the reliability value. Morgan and Goulter (1985), Loganathan et al. (1990) and Afshar et al. (2005), for example, tried to optimize the reliability by optimizing the layout of the distribution network. Other researchers, e.g. Fujiwara and De Silva (1990) and Fujiwara and Tung (1991), optimized the pipe diameters in the network to get the optimum value of the reliability. However, their proposed methods are somewhat unsatisfactory as discussed next.

Morgan and Goulter (1985) optimized the layout of water distribution networks by first considering all the candidate links in the network. They used linear programming to obtain the optimum set of pipe diameters and linked the optimization program to Hardy-Cross network solver to obtain the heads and the corresponding link flows in the network. Within the linkages between the optimization and the network analysis phase, they introduce an algorithm to remove uneconomical links from the network. Each link is assigned a weight, which is proportionate to the ratio between the flow in the link and the total flow at the downstream node of that link. The weight is calculated prior to the development of the pressure constraint in the optimization phase. At the end of each optimization phase, the minimum diameter pipe with the lowest weighting is removed from the network and the resulting layout is reanalysed to obtain the new sets of nodal pressure heads and link flows. If the layout resulting from the removal of this pipe is more economical than the previous solution then it becomes the new best result and the procedure continues as before. If the new layout is more expensive then the pipe is retained and the old best layout is used instead. This process of removing pipes continues until the lowest weighting is greater than a specified value. Once this condition is reached, the pipe optimization and flow distribution phases based on the new layout continue as before. Final optimum solution is considered to have been reached when the linear program cannot reduce the size of any pipe in the network and the minimum weighting in any link in the network is greater than the specified value. Several demand patterns corresponding to several critical operating conditions may be considered to increase the performance of the resulting design. However, the performance of the resulting design depends largely on the value of the minimum weighting specified by the designer since no actual measure of performance is used in the model. Also, the true optimum layout and the corresponding pipe sizes may not be achieved since the minimum weighting that corresponds to this layout is not known.

Afshar et al. (2005) proposed a different method of optimizing the layout of water distribution networks. The approach, however, is the same as that of Morgan and Goulter (1985) in the sense that both methods start by considering the whole set of the candidate links in the network. Continuous pipe diameters are used in the design optimization phase. Once a design based on the full set of links is obtained, one link in the network is "floated" so that the configuration would yield a cheaper layout. A link is floated by relaxing the minimum pipe diameter constraint, i.e. by setting the minimum pipe diameter value of zero, so that in the design optimization phase in the next iteration the link can be eliminated. The criteria used in choosing which link to be floated are based on the assumption that the link is the most hydraulically unimportant but economically important link in the network. This assumption, as Afshar et al. (2005) admitted, cannot be proved and justified. Once a link has been floated, the method continues with the linear programming phase to design the network. Only this time the linear program can decide whether or not to eliminate the floated link by choosing a zero or nonzero value for the diameter. Convergence of the iteration is assumed to have been reached when there are no more links in the network to be floated, i.e. floating one or more of the remaining link(s) would lead to violation of the design constraints or the specified reliability criteria. Afshar et al. (2005) used the concept of reliability as a measure of independent paths from the source node(s) to each of the demand nodes. As such, no quantified measure of the reliability was actually used. The performance of the resulting design is therefore difficult to quantify and questionable since multiple connections between a demand node and the source node(s) do not guarantee that the node would have sufficient supply when one or more of these connections fail (Wagner et al., 1988a).

Loganathan et al. (1990) used an inverse strategy to Morgan and Goulter (1985) and Afshar et al. (2005). Instead of starting the algorithm by considering all the candidate links in the network, their approach starts by obtaining an optimum tree-type design and proceeds by adding loop-forming redundant links to increase flexibility and performance of the network while keeping the rise in cost to a minimum. The addition of the redundant links is done in such a way that all the demand nodes are connected to the source by at least two independent paths. All the redundant links are first assumed to have a minimum available diameter. Network analysis is then carried out on the augmented design and the pipe diameters are adjusted accordingly to satisfy the pressure requirements. Changes on the core tree links are kept to a minimum. The performance of the resulting design is, however, questionable since the cost of the design is kept down by using mostly minimum diameter pipes for the redundant links.

Therefore, the bulk of the flow is carried by the initial core tree links. This raises question of how useful the redundant links are. When a redundant link fails, the network may be able to cope well with the situation, but when a core-tree link fails, the reduced network may struggle to meet the demands in some areas.

Fujiwara and De Silva (1990) tried to improve the reliability of a water distribution network by increasing the size of some pipes in the network. The algorithms start with initial link flows and a least cost design of a pipe network is obtained based on these flows. Each link is then assigned a length, which is defined as a function of marginal reliability and marginal cost with respect to the flow increment. The reliability of the network is improved by increasing the link flows, hence the pipe sizes, along the longest path from the source to a demand node in such a way that higher reliability is achieved with only minimal cost increment.

Fujiwara and Tung (1991) have listed several weaknesses of the Fujiwara and De Silva (1990) approach. These weaknesses are:

1. The model used for the assessment of network reliability does not consider hydraulic consistency along loops or the head requirements at demand nodes.
2. The increase of flow along the selected longest path does not represent an improvement in the reliability.
3. The proposed method does not necessarily produce a symmetric design when the network is symmetric.

Fujiwara and Tung (1991), therefore, proposed an improved method for optimizing the reliability of a water network by increasing the diameter of the pipes in the network. Their algorithms begin by obtaining an arbitrary initial design that satisfies the constitutive equations, i.e. continuity equations, conservation of energy, etc. They used continuous pipe diameter and each link consists of a single pipe diameter. The reliability of the network is then increased by increasing the diameter of a pipe by a fixed quantity called a step size in such a way that the ratio of the total increase in reliability to the total increase in cost with respect to pipe size change is maximized. If the reliability of the new design is still below the previously obtained upper bound, the process is repeated until this value is exceeded. Once it is exceeded, the pipe size is decreased by half the size. This process of increasing and decreasing a pipe diameter to achieve the best local improvement is called the *greedy algorithm* (see e.g. Cormen et al., 2001) and may be repeated until a satisfactory level of network reliability is achieved, i.e. roughly equal to the upper bound. For the next iteration, a new pipe diameter that maximizes the increase in reliability relative to the increase in cost is obtained. A new upper bound of the reliability value is then calculated for that particular iteration and the greedy algorithm is repeated. The above process is repeated until the pre-specified value of network reliability is reached.

The reliability measure adopted by Fujiwara and Tung (1991) is similar to that of Fujiwara and De Silva (1990) and is defined as the ratio of the expected maximum total water supplied to the total demand. Similar to Fujiwara and De Silva (1990), the maximum total outflow in critical operating conditions, i.e. single and multiple pipe failures, must first be obtained. However, unlike Fujiwara and De Silva (1990), Fujiwara and Tung (1991) used a nonlinear maximum flow model to obtain the maximum total water supplied under critical conditions, therefore, the hydraulic consistency as well as the pressure requirements are accounted for. Also, in obtaining the maximum total outflow, pipe flow capacity was used. The value was specified for each pipe based on practical considerations and was defined as the flow that occurs when a hydraulic gradient is of maximum allowable value (Wagner et al., 1988a). However, from their sensitivity study it was found that the calculated maximum total outflow, hence the reliability value, was shown to be quite sensitive to the value of the pipe flow capacities specified. The method also required high computational times even for a moderate sized network (4-loop network with 12 links in their study) since all possible link failure conditions were considered for reliability assessment each time pipe diameter(s) was changed.

5.3 HYDRAULIC REDUNDANCY

Redundancy is another measure used to assess the robustness of water distribution networks. Redundancy means having extra components or having large components in order to sustain the flow of water in the event of failure. This concept of failure tolerance is exploited by Tanyimboh and Templeman (1998) who suggested that it can be used to quantify the redundancy of water distribution networks. Tanyimboh et al. (2001) recognised that under normal condition, the distribution networks are expected to perform satisfactorily. Therefore, the network performance under deficient conditions is probably more important to analyse.

The redundancy measure proposed by Tanyimboh and Templeman (1998) is used in this thesis and it can be written as

$$FT = \frac{R - DSR(0)\,p(0)}{1 - p(0)} \tag{5.14}$$

in which *FT* is the failure tolerance of water distribution networks as a measure of redundancy; *R* is the system reliability; *DSR*(0) is the demand satisfaction ratio when all the pipes are available; and *p*(0) is the probability that all the pipes in the network are available. Assessment by Kalungi and Tanyimboh (2003) has shown water distribution networks that have similar performance under normal conditions may behave differently when they are subjected to critical conditions. They therefore suggested that the above function may be used for quantifying the redundancy or the failure tolerance of the distribution networks.

Equation (5.15) defines failure tolerance as the expectation of the proportion of the demand in the network that is satisfied during periods in which some components are out of service (Tanyimboh and Templeman, 1998). Its calculation is very straight forward once the value of *R*, *DSR*(0) and *p*(0) have been obtained in the process of determining the reliability value. As such, it does not add unnecessary burden in the assessment of the performance of water distribution systems. However, the ease of obtaining its value depends largely on the ease of computation of the reliability.

5.4 SURROGATE PERFORMANCE MEASURE

Several surrogate performance measures for water distribution networks are also available. Entropy is one such measure. Its use as a surrogate measure for the reliability has been outlined in Chapter 3. Another surrogate measure used in this thesis is the total energy dissipated by the pipe network. Following Rowell and Barnes (1982) who suggested that the efficiency of a pipe can be measured from the rate at which it dissipates energy, Tanyimboh and Templeman (1993b) have used this approach to assess and compare alternative designs for water distribution networks. For a given set of flows, the total energy dissipated by a pipe network, E, can be calculated using the following function

$$E = \rho g \sum_{ij \in IJ} q_{ij} h_{ij} \tag{5.15}$$

in which ρ is the density of water; g is the acceleration due to gravity; q_{ij} and h_{ij} are the flow rate and head loss in link ij, respectively; IJ represents the set of links in the full or reduced network as appropriate. For the same rate of outflows, the more energy dissipated by the network indicates higher stress levels experienced by the network. Hence, the use of the above expression as a surrogate measure of performance of water distribution networks seems justifiable.

5.5 LEAST COST DESIGN OF PIPE NETWORKS

In this section, the optimum design of water distribution networks with pre-specified layout is described. Also, external flows, i.e. supplies and demands, are assumed known together with the length of each pipe. The problem is presented for a single demand pattern and then extended for multiple demand patterns. The problem with multiple demand patterns is important especially in designing networks with storage tanks or service reservoirs since the demand patterns when the tank is filling up will obviously be different to the one when the tank is discharging. The formulation of the problem consists of minimizing the capital cost of the pipes, which means determining the cheapest set of pipe diameters subject to the constitutive equations presented earlier in Chapter 4 and to other constraints due to practical considerations, which will be described later in this section.

5.5.1 OBJECTIVE FUNCTIONS

The cost per unit length of a pipeline is usually given by the following function.

$$F(D_{ij}) = \gamma D_{ij}^{e} \tag{5.16}$$

in which D_{ij} is the diameter of pipe ij; γ and e are user specified coefficients whose values depend on the units of D_{ij}. For problems considering continuous pipe diameters, the value of e typically lies between 1.0 and 2.5 (Fujiwara and Khang, 1990). In practice, however, the available pipe diameters are discrete and standardised and, therefore, the value of the cost per unit length of the diameters are empirically known and specified.

The total cost of pipes for problem with continuous diameter is therefore

$$C = \gamma \sum_{ij \in IJ} L_{ij} D_{ij}^{e} \tag{5.17}$$

in which L_{ij} is the length of pipe ij. Whereas, for discrete-pipe-diameter problem the total cost of the pipes is given by

$$C = \sum_{ij \in IJ} \gamma_d L_{ij}^{d} \tag{5.18}$$

where the script d is the discrete pipe diameter index; γ_d and L_{ij}^{d} are the cost per unit length of pipe and the length of pipe ij whose diameter size d.

For a design problem involving other components, e.g. pumps and valves, the objective function must also include the capital costs of the components. These components are usually modelled as links in the network and their corresponding head gains or losses must be included in the constraint set. In addition, energy cost required to drive water through the network also needs to be considered when designing a network with pump(s). Awumah and Goulter (1992), for example, have used the following function to calculate the energy cost for water distribution networks.

$$C_e = \varepsilon\left[\sum_{ij\in IJ} q_{ij}h_{ij} + \sum_{j=1}^{D} q_{j0}\left(H_j - Z_j\right)\right] \tag{5.19}$$

ε in the above equation is the price per unit cost of energy; q_{ij} and h_{ij} are the flow and head loss in link ij, respectively; IJ and D are, respectively, the sets of all the links and nodes in the network; q_{j0} is the demand at node j; H_j and Z_j are the total head at and elevation of demand node j.

Kalungi and Tanyimboh (2002) have looked into the optimization of design and upgrading of water distribution networks. The objective of their optimization model is to minimise the present value of the total future costs subject to some projected constraints. The cost function in the optimization problem comprises the pipeline life-cycle costs (i.e. including the cost for installing, paralleling, maintaining and replacing the pipes), the cost for setting up the plants and machinery at the start of each construction phase and the cost that varies in association with the network capacity, e.g. treatment, transmission, etc.

The primary concern of the present research, however, is on designing new gravity networks and, therefore, only the capital cost of pipes is considered. The interested readers could refer to the original publications mentioned above for details of other cost functions.

5.5.2 CONSTRAINTS

The constitutive equation constraints have been presented in detail in the previous chapter and therefore will not be reproduced here. The constraints due to practical requirements, which include flow velocity constraints, nodal pressure constraints, pipe diameter constraints and non-negativity of flow constraints, are presented next.

FLOW VELOCITY CONSTRAINTS

$$v_{\min} \le v_{ij} = \frac{4q_{ij}}{\pi D_{ij}^2} \le v_{\max}\ , \quad \forall ij \in IJ \tag{5.20}$$

in which v_{ij} is the velocity of flow in pipe ij; and v_{min} and v_{max} are the lower and upper bounds on the velocity respectively. Rearranging Equation (5.18) becomes

$$\frac{\pi v_{\min}}{4} \le \frac{q_{ij}}{D_{ij}^2} \le \frac{\pi v_{\max}}{4} , \quad \forall ij \in IJ \tag{5.21}$$

NODAL PRESSURE CONSTRAINTS

$$H_{\min,n} \le H_n = H_s - \sum_{ij \in IJ_n} h_{ij} \le H_{\max,n} , \quad \forall n \tag{5.22}$$

in which $H_{min,n}$ and $H_{max,n}$ are the lower and upper bounds on the nodal head H_n, respectively; IJ_n consists of all links along a specified path from a selected source to node n.

Rearranging Equation (4.44) becomes

$$H_s - H_{\max,n} \le \sum_{ij \in IJ_n} h_{ij} \le H_s - H_{\min,n} , \quad \forall n \tag{5.23}$$

PIPE DIAMETER CONSTRAINTS

$$D_{\min} \le D_{ij} \le D_{\max} , \quad \forall ij \in IJ \tag{5.24}$$

where D_{min} and D_{max} are, respectively, the lower and upper bounds on the pipe diameters. The above constraints are for continuous pipe diameters. If discrete pipe diameters are used then the pipe must be selected from the set of the available discrete pipe sizes. Hence, the pipe diameter constraint becomes

$$D_{ij} \in D_D \tag{5.25}$$

in which D_D is the set of the available pipe diameters.

NON-NEGATIVITY OF FLOWS

$$q_{ij} \geq 0 \ , \quad \forall ij \in IJ \tag{5.26}$$

This equation is required on designing water networks using entropy as explained later. Having defined all the necessary equations, the problem formulation of minimizing the cost of pipe networks are now brought together as Problem 3. However, objective functions that include costs other than pipes can be used if considered appropriate.

Problem 3

$$\underset{\forall D_{ij}}{Minimize} \quad C = \gamma \sum_{ij \in IJ} L_{ij} D_{ij}^{e} \tag{5.17}$$

(for continuous pipe diameters)

Or

$$\underset{\forall D_{ij} \in D_D}{Minimize} \quad C = \sum_{ij \in IJ} \gamma_d L_{ij}^{d} \tag{5.18}$$

(for standardised discrete pipe diameters)

Subject to:

$$h_{ij} = \alpha L_{ij} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ij}^{4.87}} \ , \quad \forall ij \in IJ \tag{4.4}$$

$$\sum_{j \in NU_n} q_{jn} - \sum_{k \in ND_n} q_{nk} = q_n \ , \quad n = 1,\ldots,NN-1 \tag{4.1}$$

$$\sum_{ij \in IJ_l} h_{ij} = 0 \ , \quad l = 1,\ldots,NLP \tag{4.7}$$

$$\sum_{ij \in IJ_p} h_{ij} = h_p \ , \quad p = 1,\ldots NP \tag{4.9}$$

$$\frac{\pi v_{\min}}{4} \leq \frac{q_{ij}}{D_{ij}^{2}} \leq \frac{\pi v_{\max}}{4} \ , \quad \forall ij \in IJ \tag{5.21}$$

$$H_s - H_{\max,n} \leq \sum_{ij \in IJ_n} h_{ij} \leq H_s - H_{\min,n} \ , \quad \forall n \tag{5.23}$$

$$D_{\min} \le D_{ij} \le D_{\max}\ , \quad \forall ij \in IJ \tag{5.24}$$

(if continuous pipe diameters are used)

$$D_{ij} \in D_D \tag{5.25}$$

(if discrete pipe diameters are used)

Problem 3 is for a single demand pattern only. In reality, however, water distribution networks are subject to many different demand patterns. To obtain more rigorous designs, more than one demand pattern may be considered in the design optimization problem. For each of the demand patterns there will be a set of constitutive equation, velocity and nodal pressure constraints. The pipe diameter constraints are included only once since they are not a function of the flows. For each demand pattern, an additional subscript *r* is introduced in this thesis to identify the corresponding variables. The problem formulation of Problem 3 for multiple demand patterns is stated below as Problem 4.

Problem 4

$$\underset{\forall D_{ij}}{Minimize} \quad C = \gamma \sum_{ij \in IJ} L_{ij} D_{ij}^{e} \tag{5.17}$$

(for continuous pipe diameters)

Or

$$\underset{\forall D_{ij} \in D_D}{Minimize} \quad C = \sum_{ij \in IJ} \gamma_d L_{ij}^{d} \tag{5.18}$$

(for standardised discrete pipe diameters)

Subject to:

$$h_{ijr} = \alpha L_{ij} \left(\frac{q_{ijr}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ij}^{4.87}}\ , \quad \forall ij \in IJ, \forall r \tag{5.27}$$

$$\sum_{j \in NU_{nr}} q_{jnr} - \sum_{k \in ND_{nr}} q_{nkr} = q_{n,r}\ , \quad n = 1,\ldots, NN-1,\ \forall r \tag{5.28}$$

$$\sum_{ij \in IJ_l} h_{ijr} = 0\ , \quad l = 1,\ldots, NLP,\ \forall r \tag{5.29}$$

$$\sum_{ij \in IJ_{pr}} h_{ijr} = h_{pr}\ , \quad p = 1,\ldots NP,\ \forall r \tag{5.30}$$

$$\frac{\pi v_{\min}}{4} \le \frac{q_{ijr}}{D_{ij}^2} \le \frac{\pi v_{\max}}{4}, \quad \forall ij \in IJ, \forall r \tag{5.31}$$

$$H_s - H_{\max,n} \le \sum_{ij \in IJ_{nr}} h_{ijr} \le H_s - H_{\min,n}, \quad \forall n, \forall r \tag{5.32}$$

$$D_{\min} \le D_{ij} \le D_{\max}, \quad \forall ij \in IJ \tag{5.24}$$

(if continuous pipe diameters are used)

$$D_{ij} \in D_D \tag{5.25}$$

(if discrete pipe diameters are used)

q_{ijr} and h_{ijr} in the above problem are, respectively, the pipe flow rate and the head loss in link ij associated with the rth demand pattern, $\forall ij \in IJ$, $r = 1,...,NR$, where NR is the number of demand patterns; h_{pr}, $p = 1,...,NP$, $\forall r$, is the head loss in path p for the rth demand pattern; $q_{n,r}$ is the external inflow at supply node and outflow at demand node n for the rth pattern. Finally, the sets NU_{nr} and ND_{nr} consist of the nodes upstream and downstream of node n, respectively, for the appropriate flow regime.

5.5.3 LEAST COST DESIGN OF PIPES USING ENTROPY CONSTRAINT

To obtain a more reliable design of water distribution network, entropy can be incorporated directly into the optimization problem. The maximum value of the entropy must first be calculated. The algorithm described in Chapter 3 can be used to obtain the maximum entropy value provided that the external inflows and outflows in the network, layout of the network as well as the flow directions are known. Once the value is found, it can then be incorporated into the optimization procedure as described next.

To illustrate the entropy-based optimum design problem, Problem 3 is restated below as Problem 5 with the entropy constraint added. The problem then becomes an entropy-constrained cost minimization (Tanyimboh, 1993).

Problem 5

$$\underset{\forall D_{ij}}{Minimize} \quad C = \gamma \sum_{ij \in IJ} L_{ij} D_{ij}^e \tag{5.17}$$

Subject to:

$$h_{ij} = \alpha L_{ij} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ij}^{4.87}} , \quad \forall ij \in IJ \tag{4.4}$$

$$\sum_{j \in NU_n} q_{jn} - \sum_{k \in ND_n} q_{nk} = q_n , \quad n = 1,...,NN-1 \tag{4.1}$$

$$\sum_{ij \in IJ_l} h_{ij} = 0 , \quad l = 1,...,NLP \tag{4.7}$$

$$\sum_{ij \in IJ_p} h_{ij} = h_p , \quad p = 1,...NP \tag{4.9}$$

$$\frac{\pi v_{\min}}{4} \leq \frac{q_{ij}}{D_{ij}^2} \leq \frac{\pi v_{\max}}{4} , \quad \forall ij \in IJ \tag{5.21}$$

$$H_s - H_{\max,n} \leq \sum_{ij \in IJ_n} h_{ij} \leq H_s - H_{\min,n} , \quad \forall n \tag{5.23}$$

$$D_{\min} \leq D_{ij} \leq D_{\max} , \quad \forall ij \in IJ \tag{5.24}$$

$$S \geq S_{\min} \tag{5.33}$$

$$q_{ij} \geq 0 , \quad \forall ij \in IJ \tag{5.26}$$

in which S is the entropy and $S_{\min}$ is the minimum desired entropy value. Equation (5.33) will ensure that the entropy of the network does not fall below the specified value of $S_{\min}$, which can take any value between zero and the maximum entropy value for the network under consideration. Equation (5.26) is necessary to ensure that the flow directions in the resulting design do not change since entropy cannot be defined when the flows have negative values. This is a major limitation of the current entropy-based design approach since the optimum set of flow directions is not generally known in advance. A possible solution to this problem is presented in Chapter 8 of this thesis. All other symbols have previously been defined and those definitions are unchanged. The above problem formulation is consistent with the current studies on entropy-based design of water distribution networks in which continuous pipe diameters are used. The problem is non-linear and may have many local minima. It is therefore difficult to obtain the global optimum to the above problem and even more so when discrete pipe diameters are used.

Problem 5 (and Problem 3) may be reduced in size using the following procedures (Tanyimboh, 1993). When continuous pipe diameters are used, it is logical to assume that the lower bound of the pipe diameter constraint is more significant than the upper bound since the optimization procedure tends to reduce the size of the diameters in the network. Also, the minimum nodal pressure constraint is more important than the maximum since the reduction in the pipe sizes tend to eliminate the residual head at demand nodes. Furthermore, the minimum nodal pressure constraint is more likely to be critical at the terminal node(s), i.e. the node(s) furthest from the source with no other nodes downstream of it. The optimization procedure also makes the maximum velocity constraint more likely to be binding than the minimum constraint. Due to these considerations, some of the constraints can therefore be eliminated from the optimization process to reduce the size of the calculations. Hence, the maximum diameter constraints and minimum velocity constraints can be excluded. Also, the minimum nodal pressure constraint can be applied only at the terminal node(s).

At the end of the optimization, however, verification must be carried out to ensure that the constraints that have been excluded are not violated. Any omitted constraints that have been violated must be reinstated into the optimization procedure and the problem is re-solved. This process is repeated as many times as necessary.

5.6 SOLUTION METHODS

Several solution methods to the above problems are described next. The most widely used method is probably the Non-linear Programming, which uses continuous pipe diameters in the calculations. In practice, however, the pipe diameters are discrete and standardised. Another method is the Linear Programming Gradient method. The resulting designs of this method have discrete pipe diameters, which are selected from the set of standardised diameters available. However, several links in the resulting design may consist of more than one segment with different pipe sizes, which in practice is undesirable due to construction consideration.

Both of the above methods are deterministic in nature. As such, a single solution is obtained at the end of the process. Alternatively, stochastic search methods are now used in the optimization of water distribution networks. They are capable of producing optimum and near optimum results from a single run. These methods, amongst others, are the Genetic Algorithms and Ant Colony Optimization Algorithms. Another advantage of using the stochastic methods in water network optimization is that they work well with discrete variables and therefore the pipe diameters for the networks can be selected directly from the set of discrete standard pipe sizes available in practice. Although, this makes the design problem more difficult to solve, several studies have demonstrated that these methods can produce optimum design fairly quickly and effectively

The Ant Colony Optimization Algorithms (ACOA) is a relatively new search method formulated by Dorigo et al. (1996). Its application in water distribution systems was introduced by Maier et al. (2003). The method is based on the behaviour exhibited by ant colonies in their search for food and using pheromone trails to generate better solutions at each succeeding iteration. Genetic Algorithms (GA) search methods, on the other hand, are based on the natural selection of the survival of the fittest in which the process of generating trial solutions is governed by crossover probabilities and mutation. Its use in the optimization of pipe networks has been studied quite comprehensively in the past few decades e.g. Goldberg and Kuo, 1987; Simpson et al., 1994; Halhal et al., 1997; Savic and Walters, 1997; Vairavamoorthy and Ali, 2000). This method has also been recognised as the most capable optimization procedure for design and rehabilitation of water distribution networks (Walski et al., 2003). More details of these stochastic methods are presented next following the brief descriptions of the deterministic search methods mentioned above.

5.6.1 NON-LINEAR PROGRAMMING METHODS

Problems 3, 4 and 5 are formulated as non-linear constrained optimization. Yates et al. (1984) have shown that the requirement for discrete pipe diameters makes the problem extremely difficult to solve. Simplification can be achieved, however, by using continuous, as opposed to discrete pipe diameters, in the optimization and solving the problem by any suitable algorithm for constrained non-linear programming, e.g. sequential quadratic programming method.

The pipe diameters obtained from the continuous problem are often not available in practice. This problem can be solved by replacing the pipe in the real network by two commercially available pipes in series so that the total head loss in the replacement pipes is the same as that in the pipe (with continuous diameter) being replaced. This is to ensure that the head requirements at the critical node(s) remain satisfied. Also, in order for the actual network to remain close to optimal, the replacement pipes are selected from D_D, such that, they are the closest diameter above and below the computed (continuous) value. Furthermore, the total length of the replacement pipes must equal the length of the pipe being replaced.

To determine the length of the replacement pipes, the following simultaneous equations can be set up and solved (Tanyimboh, 1993).

$$L_{ij} = L_{ij1} + L_{ij2} \tag{5.34}$$

in which L_{ij1} and L_{ij2} are the lengths of the first and second replacement pipes, respectively, for link ij.

$$h_{ij}^{*} = h_{ij1} + h_{ij2} = \alpha \left[\frac{L_{ij1}}{D_{ij1}^{4.87}} + \frac{L_{ij2}}{D_{ij2}^{4.87}} \right] \left[\frac{q_{ij}^{*}}{C_{ij}} \right]^{1.852} \tag{5.35}$$

in which the superscript * is used to indicate the optimum value of the variable; h_{ij1} and D_{ij1} are the head loss and diameter for the first replacement pipe for link ij and similarly for h_{ij2} and D_{ij2}, which correspond to the second replacement pipe for link ij. The diameters D_{ijm}, m = 1, 2, are selected and therefore known. Also, the values of all the variables in Equation (5.35) are known except L_{ij1} and L_{ij2}. The above equations are linear in the unknown lengths and can be solve to give

$$L_{ij1} = \frac{h_{ij}^{*} - k_2 L_{ij}}{k_1 - k_2} \tag{5.36}$$

$$L_{ij2} = -\frac{h_{ij}^{*} - k_1 L_{ij}}{k_1 - k_2} \tag{5.37}$$

in which the coefficients k_1 and k_2 are

$$k_1 = \frac{\alpha\left(q_{ij}^{*} / C_{ij}\right)^{1.852}}{D_{ij1}^{4.87}} \tag{5.38}$$

$$k_1 = \frac{\alpha\left(q_{ij}^{*} / C_{ij}\right)^{1.852}}{D_{ij2}^{4.87}} \tag{5.39}$$

In Equations (5.35) to (5.39), the roughness coefficient is assumed constant, i.e. $C_{ij1} = C_{ij2} = C_{ij}$. However, different values may be used if necessary without any difficulty.

The above split-pipe solution to the continuous pipe diameter problem is nevertheless not favoured by practicing engineers since some of the pipe segments may be too short for their construction to be justified. As an alternative, the continuous pipe diameters may be rounded up or down to the nearest discrete pipe sizes available in practice. This exercise, however, may lead to some constraint violations or make the end result a sub optimal design.

5.6.2 LINEAR PROGRAMMING GRADIENT METHOD

The Linear Programming Gradient (LPG) method was developed by Alperovits and Shamir (1977). The method is iterative and is divided into two sub problems. The first solves a linear programming problem to find the sizes of the network components which correspond to the minimum cost of the network based on some specified initial flow rates. The second sub problem uses the dual variables obtained from the solution of the linear programming to adjust the flow rates in the network in such a way that the cost of the network is reduced. Another linear programming problem is then solved based on the new set of flow rates producing another set of dual variables which is then used to further modify the flow rates. This cycle continues until the reduction in cost becomes insignificantly small and therefore negligible.

It has been shown earlier that the problem of designing water distribution networks, i.e. Problem 3 or 4, is non linear. For a given flow rate, the head loss equation, Equation (4.4), is non linear in the diameters only. It follows that for a set of specified flow rates, Problem 3 (or 4) is non linear in the diameters only. However, the problem can be linearised by selecting the lengths as the decision variables instead. Since the length of each link is fixed, the pipe in a link has to be divided into segments with different diameters, which must be selected from a set of discrete pipe sizes. The total length of the pipe segments in a link must equal the length of the link, i.e.

$$\sum_{m=1}^{N_{ij}} L_{ijm} = L_{ij} \; , \quad \forall ij \in IJ \tag{5.40}$$

$$L_{ijm} \geq 0 \; , \quad \forall ijm \tag{5.41}$$

in which L_{ijm} is the length of segment m in link ij; N_{ij} is the number of segments specified for link ij. Alperovits and Shamir (1977) have stated that at the optimum solution, each link will contain two segments at most. The list of candidate diameters for each link may be different and it is obtained by setting the allowable minimum and maximum value of the hydraulic gradient. For a set of flow rates, these limiting gradients will yield a maximum and a minimum diameter admissible for each link, which can be chosen from the standardised discrete pipe diameters available in practice. It should be noted that the specified minimum and maximum values of the hydraulic gradient introduces an implicit constraint into the optimization problem since the gradient reduces the candidate diameters for each link and hence reduces the number of variables in the optimization.

The objective function in LPG is given by

$$C = \gamma \sum_{ij \in IJ} \sum_{m=1}^{N_{ij}} L_{ijm} D_{ijm}^{e} \tag{5.42}$$

where D_{ijm} is the diameter of segment m of link ij. Hence, C is linear in the segmental lengths, given the segmental diameters.

For the continuity equation to hold, the flow rate in each segment of link ij must equal the pipe flow rate q_{ij} of that link. Finally, the head loss in link ij is the sum of the head loss in each segment of that link, i.e.

$$h_{ij} = \sum_{m=1}^{N_{ij}} h_{ijm} \text{ , } \quad \forall ij \in IJ \tag{5.43}$$

in which h_{ijm} is the head loss in segment m of link ij. The head loss in a segment is given by the head loss equation below.

$$h_{ijm} = \alpha L_{ijm} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ijm}^{4.87}} \text{ , } \quad \forall ijm \tag{5.44}$$

The velocity constraints of Equations (5.20) are not affected by the process of link segmentation described above. Hence, all the segments in a link will have the same minimum and maximum velocity constraints. These equations can be used to determine the lower and upper bound of the segmental pipe diameters in each link.

Once the variable transformation has been completed as explained above, the problem is then linear in the L_{ijm} and is stated below as Problem 6 for a single flow regime. The LP formulation for multiple loadings is not done in this thesis. The derivation follows from Problem 6 in the same way that Problem 4 is derived from Problem 3. The continuity equations, Equations (4.1), are omitted from Problem 6 since they are used for specifying the initial pipe flow rates for the linear program.

Problem 6

$$\underset{\forall L_{ijm}}{Minimize} \; C = \gamma \sum_{ij \in IJ} \sum_{m=1}^{N_{ij}} L_{ijm} D_{ijm}^{e} \tag{4.63}$$

Subject to:

$$\alpha \sum_{ij \in IJ_l} \sum_{m=1}^{N_{ij}} L_{ijm} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ijm}^{4.87}} = 0 \ , \quad l = 1, \ldots, NLP \tag{5.45}$$

$$\alpha \sum_{ij \in IJ_p} \sum_{m=1}^{N_{ij}} L_{ijm} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ijm}^{4.87}} = h_p \ , \quad p = 1, \ldots, NP \tag{5.46}$$

$$\alpha \sum_{ij \in IJ_n} \sum_{m=1}^{N_{ij}} L_{ijm} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ijm}^{4.87}} \le H_s - H_{\min,n} \ , \quad \forall n \tag{5.47}$$

$$\alpha \sum_{ij \in IJ_n} \sum_{m=1}^{N_{ij}} L_{ijm} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ijm}^{4.87}} \ge H_s - H_{\max,n} \ , \quad \forall n \tag{5.48}$$

$$\sum_{m=1}^{N_{ij}} L_{ijm} = L_{ij} \ , \quad \forall ij \in IJ \tag{4.61}$$

$$L_{ijm} \ge 0 \ , \quad \forall ijm \tag{4.62}$$

The decision variables for the above problem, i.e. the segmental lengths, are continuous and the problem can be solved by any suitable linear programming algorithms. The solution to the problem will yield the minimum cost of the network for the given flow rates, C^*, the length of all the segmental pipes, L^*_{ijm}, and the dual variables that correspond to each loop, path and maximum and minimum head constraints. These dual variables are used in the gradient phase of the LPG method to find the gradient of the cost, C, with respect to the change in the link flow rates, Δq_l. Knowing the gradient, it is then possible to alter the flow rates in each link in such a way that the solution of the next linear program will have a lower network cost, i.e.

$$C^* \left(\underline{q} + \underline{\Delta q} \right) \le C^* \left(\underline{q} \right) \tag{5.49}$$

where $\underline{q}$ is the vector of the link flow rates. The details of the gradient phase are not given in this thesis. The interested readers can refer to Alperovits and Shamir (1977). Tanyimboh (1993) also has described this method in minute detail.

Qundry, Brill and Liebman (1981) presented an alternative but similar approach to the LPG method described above. The main difference between their method and that of Alperovits and Shamir (1977) is in the way the linearization is brought about. Quindry, Brill and Liebman (1981) method is base on the H-equations and they used continuous variables for the diameters with the assumption that each link consist only of a single pipe of uniform diameter. Details of the method, however, is not presented here, interested readers may refer to the original publication for further information or Templeman (1982) for a comprehensive discussion of the method.

The LPG method is complex. It involves the solution to two optimization problems since finding the value of the change in the link flows in the gradient phase also requires an optimization process. Although Kessler and Shamir (1989) have shown that the above LPG technique will converge to an optimum solution, the segmental pipes in the resulting design or the alternative continuous pipe solution may cause problems in the construction phase as mentioned earlier.

5.6.3 ANT COLONY OPTIMIZATION ALGORITHMS

The optimization problem formulation using ACOA starts by representing the problem in terms of a graph $G = (A, B, C)$, in which $A = (a_1, a_2,..., a_m)$ is the set of points at which decisions have to be made, $B = (b_{m(1)}, b_{m(2)},..., b_{m(n)})$ is the set of options j available at each decision point i and $C = (c_{m(1)}, c_{m(2)},..., c_{m(n)})$ is the set of costs, each associated with each option in B (Maier et al., 2003). For a water distribution network, each decision point in A is associated with each link in the network. Assuming that the network consists only of pipes, the set of the available pipe diameters corresponds to the set of options in B. C represents the set of the capital costs associated with each link m for which the chosen pipe diameter is n, i.e. $c_{m(n)}$ is a function of link diameter and its length.

Once the problem has been formulated, a number of ants are set out on the journey to construct trial solutions and find the one with the minimum cost. The number of ants specified for the problem, NA, corresponds to the number of trial solutions to be generated. These trial solutions are constructed one at a time in a cycle as follows. As an artificial ant arrives at a decision point, i.e. at a link in a water distribution network, a diameter for the link is chosen from B based on the concentration of pheromone for that option left by the previous ants. At the start of the search, this pheromone concentration is set to a very small positive value (Dorigo et al., 1996). Once all the decision points have been covered, the cost of the trial solution is calculated as in Equation (5.17). The generation of a trial solution with its associated cost corresponds to a cycle, k, and the above process is repeated until $k = NA$. A completion of NA cycles is referred to as one iteration, t. At the end this iteration the pheromone trails are updated so that the artificial ants can choose better options for the trial solutions in the next iteration. The above procedures are repeated until a certain stopping criteria are met, e.g. the completion of a specified number of iterations.

The process of choosing the option m at a decision point n based on the pheromone intensities associated with that option $\tau_{m(n)}$ is now explained followed by the procedure for updating the pheromone trails. At each decision point, the option for the trial solution is chosen stochastically using the following formula (Dorigo et al., 1996).

$$p_{m(n)}(k,t) = \frac{\left[\tau_{m(n)}(t)\right]^{\alpha_p}\left[\eta_{m(n)}\right]^{\beta}}{\sum_{b_{m(n)}}\left[\tau_{m(n)}(t)\right]^{\alpha_p}\left[\eta_{m(n)}\right]^{\beta}} \qquad (5.50)$$

In the above equation $p_{m(n)}(k,t)$ is the probability that option $b_{m(n)}$ is chosen at cycle k and iteration t; $\tau_{m(n)}(t)$ is the concentration of pheromone for option $b_{m(n)}$ at iteration t; $\eta_{m(n)}$ is a guiding factor for the search towards options with smaller "local" costs and also referred to as "visibility" (Dorigo et al., 1996). Its value is calculated from $\eta_{m(n)} = 1/\, c_{m(n)}$; α_p and β are parameters that control the importance of the pheromone and the visibility.

At the end of each iteration, the pheromone trails can be updated as follows (Dorigo et al., 1996)

$$\tau_{m(n)}(t+1) = \rho_{pe}\tau_{m(n)}(t) + \Delta\tau_{m(n)} \qquad (5.51)$$

in which $\tau_{m(n)}(t+1)$ is the concentration of pheromone for option $b_{m(n)}$ at iteration $t+1$; ρ_{pe}, whose value is less than one, represents the pheromone persistence to evaporation; $\Delta\tau_{m(n)}$ is the change in pheromone associated with option $b_{m(n)}$ at iteration t.

The value of the change in pheromone can be calculated by (Dorigo et al., 1996)

$$\Delta\tau_{m(n)} = \sum_{k=1}^{NA} \Delta\tau_{m(n)}^{k} \tag{5.52}$$

or by (Stützle and Hoos, 2000)

$$\Delta\tau_{m(n)} = \Delta\tau_{m(n)}^{k^*} \tag{5.53}$$

in which $\Delta\tau_{m(n)}^{k}$ is the change in pheromone concentration for option $b_{m(n)}$ at cycle k and iteration t; and k^* represents the cycle number with best result during iteration t, i.e. corresponds to the ant with the best solution. The function of Stützle and Hoos (2000) has an advantage that it reduces the number of evaluations needed at each iteration, hence reducing the computer time.

Dorigo et al. (1996) proposed the following function to calculate the value of $\Delta\tau_{m(n)}^{k}$

$$\Delta\tau_{m(n)}^{k} = \begin{cases} \dfrac{PR}{C^k} & \text{if option } b_{m(n)} \text{ is chosen at cycle } k \\ 0 & \text{otherwise} \end{cases} \tag{5.54}$$

in which PR is the pheromone reward factor and C^k is the cost of the trial solution at cycle k. The above functions ensure that the trial solutions with lower costs are awarded with larger concentration of pheromone so that the probabilities of them to be chosen in future cycles are greater.

Equation (5.54) is a general formulation for an unconstrained optimization problem. In the optimization of water distribution networks, however, many constraints, including loop, path and nodal pressure constraints, have to be satisfied in the final outcome. Most of these constraints are a function of the pressure head. Therefore, Maier et al. (2003) proposed the following modifications to Equation (5.54) to include the head constraint violations in the pheromone updating process.

$$\Delta\tau^{k}_{m(n)} = \begin{cases} \dfrac{PR}{C^{k}} - P_{pher} \times \Delta H_{max} & \text{if option } b_{m(n)} \text{ is chosen at cycle } k \\ 0 & \text{otherwise} \end{cases} \tag{5.55}$$

in which P_{pher} is the pheromone penalty factor and ΔH_{max} is the maximum shortage in pressure in the distribution network obtained from the hydraulic analysis of the network. Each trial solution corresponds to a trial network. Hence, the hydraulic analysis has to be carried out as many times as the number of the trial solutions. This is the main cause of the high computer time in any stochastic search methods such as the ACOA.

To accommodate for the violation of the head constraints in the calculation of the network costs, a penalty cost is introduced in the cost function whenever violation occurs. Thus, Equation (5.17) is modified as follows

$$C = \sum_{m \in M} \gamma_d L^{d}_{m} + PC \times \Delta H_{\max} \tag{5.56}$$

in which L^{d}_{m} is the length of link m whose diameter is $d \in D_D$; γ_d is the cost per unit length of pipe with diameter d; M is the set of all the links in the network; PC is the penalty cost multiplier (unit cost/m head violated). When there is no violation of the constraints, the value of ΔH_{max} is set to zero.

The optimization of water distribution networks using ACOA can therefore be formulated as Problem 7 below.

Problem 7

$$\underset{\forall D_{ij} \in D_D}{Minimize} \quad C = \sum_{m \in M} \gamma_d L^{d}_{m} + PC \times \Delta H_{\max} \tag{5.56}$$

Subject to:

$$h_m = \alpha L_m \left(\frac{q_m}{C_m} \right)^{1.852} \frac{1}{D_m^{4.87}} \ , \quad \forall m \in M \tag{4.4}$$

$$\sum_{j \in NU_i} q_{ji} - \sum_{k \in ND_i} q_{ik} = q_i \ , \quad i = 1, \ldots, NN - 1 \tag{4.1}$$

$$\sum_{m \in M_l} h_m = 0 \ , \quad l = 1, \ldots, NLP \tag{4.7}$$

$$\sum_{m \in M_p} h_m = h_p \ , \quad p = 1, \ldots NP \tag{4.9}$$

$$H_s - H_{\max,i} \le \sum_{m \in M_i} h_m \le H_s - H_{\min,i} \ , \quad \forall i \in NN \tag{5.23}$$

$$D_m \in D_D \tag{5.25}$$

In the above problem, the flow velocity constraints have been excluded. However, they can be incorporated in the optimization quite easily and a new set of pheromone and cost penalty multipliers are needed to account for their violations.

The above problem has been solved successfully by Maier et al. (2003). However, some research is still needed to improve the application of the method in water distribution systems. For example, there seem to be no clear guidance as to what value should be used for the parameters α_p, β and ρ_{pe} in Equation (5.50). Dorigo et al. (1996) have tested several values for these parameters and they found that the optimum value for ρ_{pe} is 0.5 while several combinations for the values of α_p and β can be used to obtain good performance of the ACOA. Maier et al. (2003), on the other hand, used two completely different sets of values for the above parameters obtained by generating trial solutions to the two problems they studied. The interested readers can refer to the original publications for the values of the parameters used and their combinations.

Other issues are the determination of the number of ants, the pheromone reward value and the problem specific pheromone and cost penalty factors. Large number of ant population means a large search space for the optimization. However, it also contributes to a large computing time as mentioned earlier. Maier et al. (2003) used the value of 100 for the number of ants in their studies. They also have shown that the values of PR, P_{pher} and PC seem to be proportional to the scale of the problem. However, trial solutions were also required in determining their values. Finally, better stopping criteria are needed to increase efficiency. This would constitute the identification of an optimum solution which would lead to the termination of the search process even before the specified maximum number of iteration is reached. Despite the above weaknesses, the use of ACOA in water distribution networks seems promising.

5.6.4 GENETIC ALGORITHMS

The theory behind GA was first proposed by Holland (1975) and developed further by Goldberg (1989). Several researchers have successfully applied the search algorithms to water distribution network optimization problems (e.g. Simpson et al., 1994; Halhal et al., 1997; Savic and Walters, 1997; Vairavamoorthy and Ali, 2000). To implement GA, the set of decision variables must first be represented by a set of *chromosomes* where each decision variable corresponds to a single chromosome. Binary alphabets, i.e. 1 and 0, are usually used to represent a decision variable in a chromosome. Therefore, a chromosome may consist of a string of binary bits in which a single bit is referred to as a *gene*. A trial solution in GA consists of a string of chromosomes and a set of trial solutions is referred to as a *population*, which is equivalent to the number of ants in the ACOA. Also, the full conception of all the individual trials with their corresponding costs represents one *generation* (i.e. one iteration in ACOA).

The search process in GA is similar to that in the ACOA in the sense that both methods generate a population of trial solutions. However, the modification of the trial solutions in GA is carried out by modifying the binary bits inside the chromosome strings. Therefore, the decision variables are not used directly in the search process once they have been coded. There are three main operators in GA for generating and modifying each trial solution; these are selection, cross-over and mutation. These operators use a pseudo-random number generator to operate. It needs to be pointed out that GA uses a random number generator extensively in its operation (see e.g. Gentle, 2003, for methods for generating pseudo-random numbers). The search process is therefore stochastic and random but, to a great extent, directed towards good solution.

The process involved in the optimization of water distribution networks using GA is described next. Assume that a pipe network is to be optimized and there are 8 candidate diameters available. Each candidate diameter can be represented by a 3-bit binary chromosome and since $2^3 = 8$, each candidate of the 3-bit binary string corresponds to a single decision variable, i.e. a candidate pipe diameter. A slight problem may occur when there are 6 candidate diameters for example. These diameters cannot be represented by 2-bit binary string (since $2^2 = 4$) and there will be 2 redundant binary strings when 3-bit binary representation is used. This problem may be overcome by using a fixed remapping (Savic and Walters, 1997) in which a particular redundant binary string is associated with a specific value of the available diameter. Alternatively, real coding can be used in which the decision variables are represented by real number 0, 1, 2 and so on.

Once all the decision variables have been coded, e.g. using 3-bit binary strings, the initial population can be generated. The size of the population to be generated is specified by the designer. The random number generator generates each individual gene in the chromosome string, i.e. 1 or 0. Assuming that there are 4 links in the network, 12 genes have to be generated for one trial solution requiring 12 runs of the random number generator. The minimum total number of runs is therefore 12 times the number of population. The cost for each individual trial solution can then be obtained. To account for the constraint violations, a cost penalty factor is introduced as explained in the ACOA method. The total cost of each individual trial solution therefore signifies the *fitness* of the corresponding solution and can be calculated using Equation (5.56). The complete set of initial trial solutions with their corresponding fitness represents the first generation in GA.

To generate trial solutions for the second generation, a selection process is carried out on the population from the first generation. A number of selection procedures are available (see e.g. Mitchell, 1999). All of these procedures utilise the fitness of each individual trial so that individuals with higher fitness are more likely to be selected as *parents*. This is to ensure that the new individuals in the next generation as a result of *reproduction* from these parents will have a higher fitness level on average. In a tournament selection, for example, two individuals are chosen at random from the population. A random number between 0 and 1 is then obtained. If this number is less than a pre-specified parameter, the individual with higher fitness is selected for reproduction; otherwise the less fit individual is selected. The two are then returned to the population and can be selected again (Mitchell, 1999).

For every two selected individuals, a cross-over can be carried out to produce two *offspring* as follows. A random number is generated between 0 and 1. If this number is less than a pre-specified parameter, a cross-over is carried out; otherwise it is not. For every cross-over process, a position in the chromosome string is chosen at random. The parts of the strings of the two parents after the cross-over position are exchanged to produce two new individuals. For example, if there are two individuals with strings of chromosome 110111 and 101001 and the cross-over point is at position 3, the result of the cross-over would be two individuals with chromosome strings 110001 and 101111. This process is repeated until a certain number of offspring has been produced.

Following the cross-over procedure, mutation can be done on the new individuals with very low probability of occurrence. If the probability of mutation is set too large, the search will then become a random process and convergence to the optimal solution will never take place. For every mutation, a random number between 0 and 1 is first generated. If this number is less than the specified parameter, a mutation is carried out by selecting a single gene in the chromosome string at random and flipping its value form 0 to 1 or vice versa; otherwise no mutation occurs. This process is also repeated for every new trial solution in the current generation. The above selection, cross-over and mutation procedures are repeated until a specified number of maximum generation has been reached. Alternative convergence criteria may also be used, for example, by comparing the best solution at the current generation to the best solutions in the last several generations.

To determine the value of the parameters used in GA, i.e. the probability of selection, cross-over and mutation, as well as the number of population and the penalty factors, generation of trial solutions is required. Several researchers have proposed different combinations of the parameter values but none seems generally applicable (Mitchell, 1999). The value of the penalty factors are also case specific. Nevertheless, all the

GA studies on water distribution networks have shown that optimum solutions are always found despite the above slight shortcomings. Finally, the problem formulation for the optimization of water distribution networks using GA has the same form as that in Problem 7.

5.7 SUMMARY AND CONCLUSIONS

In this chapter, several performance measures of water distribution networks have been presented. The chapter has also discussed several methods for optimizing the performance of water networks and highlighted the difficulties related to this issue. These difficulties are partly due to the fact that the problem of quantifying the performance of water distribution networks is extremely complex. The problem of obtaining the optimum design of water distribution networks using entropy constraint has also been formulated in this chapter. The entropy-based design problem presented in Section 5.5 does not seem to increase the complexity of the optimization problem significantly. Therefore, provided that the relationship between entropy and reliability holds true in general, designing a network using entropy constraint may reduce the difficulty in obtaining inexpensive yet reliable designs. Finally, several methods for solving the design optimization problem were presented and their strengths and weaknesses are highlighted.

CHAPTER 6 SENSITIVITY ANALYSIS OF THE RELATIONSHIP BETWEEN ENTROPY AND RELIABILITY

6.1 INTRODUCTION

Early studies have shown that, in general, as the entropy of a water distribution network increases, the network becomes more and more reliable. This relationship between entropy and reliability has been shown to be quite strong (Tanyimboh, 1993; Tanyimboh and Templeman, 1993a, b, c; Tanyimboh and Templeman, 2000; Tanyimboh et al., 2002; Tanyimboh and Sheahan, 2002). Considering the ease of computation of entropy and its ability to lend itself directly into the optimization problem, the use of entropy as a surrogate measure of reliability is highly advantageous. However, the determination of the reliability of a network using its entropy value is not a straight forward process since each entropy value does not correspond exclusively to only one value of reliability. In water distribution networks, entropy is a function of flows in the network. There are many factors affecting the design of water distribution networks, which in turn affects the flow of water in the network. These factors will therefore influence the value of the entropy as well as the reliability of the network and the relationship between the two may also be affected.

As described earlier in Chapter 2, entropy is a measure of uncertainty. For water distribution networks, it measures the uncertainty of the distribution of flows in the network. Hence, it applies more to looped water networks since there is no uncertainty regarding the distribution of flows in tree-type branch networks. It follows therefore that all analyses in this thesis are carried out on networks with the links connected to form closed loops. However, looped networks with several branching pipes can still be analysed by omitting the branching pipes and adding the demands along these pipes to the nearest upstream demand node located within a loop.

In this chapter, the suitability of entropy as a surrogate measure for the performance of water distribution system is examined further. Factors that may affect the correlation between entropy and the performance of the distribution systems, i.e. the reliability, are investigated critically in a sensitivity study. These factors are: choice of layouts, the sets of flow directions chosen for the designs, different cost functions used in the design process and the slight modelling errors produced by the rounding of values, especially in the diameters when continuous diameters are used in the design. Although attempts have been made to investigate the sensitivity in the entropy-reliability relationship (e.g. Tanyimboh and Sheahan, 2000; Tanyimboh et al., 2002), no study has ever been carried out to examine the sensitivity of the relationship comprehensively. For this reason, a more thorough investigation is deemed necessary. In addition, all the previous studies were based on the DDA method. In this study and for most of this thesis the HDA analysis method is used and the results are compared to the previous DDA results.

The investigations were carried out on a hypothetical network shown in Figure 6.1, which was taken from Tanyimboh and Sheahan (2002). The source in the network has a piezometric head of 100 m while all demand nodes have elevations of 0 m. The desired nodal service head for fully satisfactory performance, H^{des}, is 30 m and the nodal head corresponding to zero nodal outflow is 0 m, this being the elevation of the nodes. All pipes are 1000 m long with a Hazen-Williams coefficient of 130 (Equation 4.4). To simplify the optimization, continuous pipe diameters are used in the design with the lower and upper bounds taken as 100 mm and 600 mm, respectively. The design optimisation was carried out using a Fortran program called PEDOWDS (Program for Entropy-constrained Design Optimization of Water Distribution Systems; Tanyimboh, 1993), with the cost as the objective function to be minimized. The program is based on the NAG library routine E04UCF (NAG Ltd., 1995), which is a routine for constrained non-linear programming. It uses sequential quadratic programming and requires gradients of the objective and constraint functions. For every design, the design optimization program PEDOWDS was run several times with different starting points, i.e. different initial diameters. This was to ensure that the same optimum design with the lowest cost was achieved several times from different starting points to increase the chances of finding a global minimum. A hydraulic analysis of the network was then carried out using EPANET to validate the result of the optimisation by verifying that no constraints were violated. A simple function in Microsoft Excel known as coefficient of determination (R^2) with linear regression is used to examine the entropy-reliability correlation in the various investigations.

6.2 POSSIBLE INFLUENCE OF LAYOUTS

The choice of layout in the design of a water distribution network will have an effect the performance of the network. Networks with different layouts can have the same value of entropy. The performance of these networks, however, may not necessarily be the same when they are subjected to abnormal conditions like fire fighting and link failure. It is therefore interesting to learn the possible effect of different layouts on the correlation between entropy and hydraulic reliability of the network.

Tanyimboh and Sheahan (2002) have tried to investigate this very issue. They generated 65 different layouts, which correspond to 65 different maximum entropy designs, based on the network in Figure 6.1 (see Figure A3.1). The flow directions for each layout were selected based on the shortest path from the source to the demand nodes. The optimum designs were generated using values of γ of 800 and e of 1.5 in Equation (5.16). A summary of the outcomes from the analysis by Tanyimboh and Sheahan (2002), which was based on the DDA method, is given in Appendix 3. In this thesis, the HDA analysis method is performed on the same designs and comparison is made to the results from the previous DDA-based study. The parabolic function of Wagner et al. (1988b) given in Chapter 4 as Equation (4.36) was used in the analysis.

Figure 6.2 shows the results of the HDA analysis performed in the present study. The graph shows that the correlation between the entropy and reliability is very strong with its R^2 value of 0.89. This is very encouraging since a high performance can be expected from a network that has a high maximum entropy value irrespective of its layout configuration. It can also be deduced that the influence of layout on the relationship between entropy and reliability is negligible. Comparing the HDA results with their DDA counterparts in Figure A3.2, it seems that the use of the HDA method in the analysis can produce a better result, i.e. stronger relationship between entropy and reliability, which leads to a more definite interpretation of the outcomes.

6.3 POSSIBLE INFLUENCE OF FLOW DIRECTIONS

The value of the entropy of a network is very much influenced by the flow directions in the network. Designs with the same layout but different flow directions may have different entropy values. On the other hand, considering the complex way in which the flow direction changes following the change in the network conditions, e.g. change in demands or in emergency situations, the set of flow directions chosen in the design process dictates the performance of the resulting design. In this section, analysis is carried out to investigate whether the change in the performance of the network due to the different sets of flow directions chosen for the design is followed by the change in the value of the network entropy with the same tendency.

The network in Figure 6.1 is again used in this study. Three different layouts were chosen for the designs, each corresponding to different number of loops in the network, i.e. 2-, 3- and 6-loop configurations. The 2- and 6-loop layouts were selected to represent the groups of designs with the least and the most possible combinations of flow directions, respectively, while the 3-loop layout represents the middle range the two extremes. Figures 6.3(a) and 6.3(b) respectively show the 2- and 3-loop layouts used in this study, while the 6-loop layout is based on the full set of links shown in Figure 6.1. The details of the network and the process of obtaining and analysing the designs have been described in the previous sections. It needs to be pointed out that the chosen flow directions for the three layouts are all hydraulically feasible in the sense that for a given network configuration the flow directions in the network *may* follow one of these arrangements. However, some of the directions are not sensible. Water always travels in the path with the least resistance. In the present study, when the length of each pipe and their roughness coefficient are the same, the path with the least resistance is the shortest path from the source to each demand node. However, some of the chosen flow directions are forcing the flows in the network to travel in longer paths. In this study, both DDA and HDA methods are used in the analysis. However, for consistency and to improve readability, the results of the DDA method are shown in Appendix A4 and this section concentrates on the HDA results only. The results from each layout are given separately first before they are combined together for a more general interpretation.

For the 2-layout, 31 different sets of flow directions were generated (Figure A4.1). Because of the way in which the flow directions were specified, two out of the 31 designs were excluded from the subsequent analyses on the grounds that their terminal nodes are too close to the source, i.e. one at node 2 and the other at node 4 (see Figure A4.1 numbers 30 and 31). Due to the minimum diameter requirement, the pressure head constraints at the terminal node for these two designs are not binding; hence the resulting designs are sub-optimal. The diameters for all the remaining 29 optimum designs are shown in Table A4.1.

24 maximum-entropy designs were generated based on the 3-loop layout and their flow directions are shown in Figure A4.2. Two of the designs, however, were excluded since their terminal nodes are very close to the source and therefore, as explained above, the resulting designs are sub-optimal (Figure A4.2 numbers 22 and 23). One other design is excluded since after several runs of the program PEDOWDS, the optimality criteria for the routine E04UCF (NAG Ltd., 1995) are not met even after slight relaxation of the entropy constraint (Figure A4.2 number 24). The entropy constraint was relaxed by 0.0015 at the most. This is to ensure that the entropy value of the resulting design is as close as possible to the maximum since comparing maximum-entropy designs against designs with lower entropy may give a misleading result considering the increase in the entropy value is followed by an increase in the

performance of the network. The diameters of the 21 designs analysed in this study are given in Table A4.2.

35 different sets of flow directions were generated based on the six-loop layout in Figure 6.1. Only 25 maximum entropy designs, however, were analysed in detail and the results are presented in Figure 6.6. The rest of the designs were excluded from further analysis since the optimality criteria of the routine E04UCF for the resulting designs were not met even after relaxing the entropy constraint as explain above. The flow directions for all the 6-loop designs are shown in Figure A4.3 and the diameters of the 25 analysed designs are given in Table A4.3.

Figures 6.4, 6.5 and 6.6 show the relationships between entropy, cost and the HDA-reliability for the 2-, 3- and 6-loop designs, respectively. When analysed separately, some of the results seemed somewhat inconclusive. One possible explanation for the weak relationship between entropy and reliability is due to the specified flow directions that lead to larger pipe diameters in some part of the network in order to satisfy the head requirement at the terminal node(s) under normal condition, which make the designs more expensive than necessary. Under pipe failure situation, however, the flow directions change in accordance with the nature of the flow, i.e. following the path with the least resistance. In this situation, the available large diameters help improve the network performance by preventing large increase in head loss in the network.

On the other hand, the network entropy does not seem to follow the network reliability value since, under normal condition, the specified flow directions lead to a non-uniformity of flows in the network. Optimum value of entropy leads to uniform pipe flows and hence pipe diameters as will be seen later in this chapter. The specified flow directions, therefore, prevent the optimum value of the network entropy to be obtained hence affecting its relationship with the network reliability. On top of that the narrow ranges in the entropy and reliability values seemed to illuminate the scatter in the plots and weaken the relationship between them. There is, however, still a hint that an increase in the entropy is followed by an increase in the reliability value. This would suggest that the different sets of flow directions chosen for the designs do not affect the entropy-reliability relationship in a considerable way. Also, considering that many of the sets of flow directions would not *naturally* occur in practice, the investigation shows that the entropy-reliability relationship still holds under *unlikely* conditions. This idea is reinforced by the strong relationships between entropy and reliability of the combined plots of all the designs shown in Figure 6.7. Also, it is quite conclusive from the analyses in this section that the results of the HDA analysis give a stronger correlation between entropy and reliability compared to the DDA results shown in Figure A4.7.

6.4 POSSIBLE INFLUENCE OF COST FUNCTIONS

The analysis in this section was also carried out using the network in Figure 6.1 whose details have been given in the previous section. Continuous values of the diameter were used in the design and the analysis was done by varying the values of the coefficients γ and e in the cost function of Equation (5.16). Nine groups of designs, each corresponding to different values of γ and e, were generated and in each group the entropy values of the designs range from the smallest to the maximum. The first group consists of 24 designs and were produced using values of γ and e of 400 and 1.0, respectively. The second group has 22 designs and were generated with $\gamma = 800$ and $e = 1.5$ while the third group with 21 designs used γ value of 1600 and e value of 2.0. There are 20 designs in the fourth group corresponding to $\gamma = 3200$ and $e = 2.5$. The fifth group consist of 19 designs with $\gamma = 3840$ and $e = 2.6$. 10 designs form the sixth group with $\gamma = 4480$ and $e = 2.7$ while 8 designs form the seventh group with γ and e values equal to 5120 and 2.8, respectively. The next group has 6 designs with $\gamma = 5760$ and $e = 2.9$ and the final group consists of 14 designs with $\gamma = 6400$ and $e = 3.0$. The user specified coefficient γ can be considered as a scaling factor. It does not have any real effect on the relationship between entropy and reliability as will be seen later. It may be noticed that between the fourth and the final group, the value of e was increased in smaller increment to show the reduction in the range of the entropy value. Also, there is a larger increment within the sixth group between the second and third designs due to the inability of the design optimization program to find the optimum solutions with the specified entropy values.

The minimum entropy designs were obtained by using the design program PEDOWDS but without the entropy constraint. The program optimized the cost of the design, which guides the search towards a tree-type network. However, due to the requirement of the minimum pipe diameter, no links may be eliminated from the design and some of the links are therefore merely loop completing links which have minimum pipe diameter size. The entropy of the resulting design is therefore at its minimum. The maximum value of the entropy, on the other hand, was calculated using the algorithm described in Chapter 3. Once the minimum and maximum entropy values for the network had been obtained for each group, a range of designs with varying entropy values from the minimum to the maximum were produced. The difference in the number of designs in each group is influenced by the minimum entropy value as will be seen later in this section.

The results of the analysis are presented next. Both DDA and HDA methods were used in the analysis but only the HDA results are presented in this chapter. The results of the DDA analysis can be found in Appendix A5. The head-outflow relationship of Wagner et al (1988b) of Equation (4.36) was used in the HDA analysis. The complete results of the analysis are shown in Figure 6.8. In general, the strength of the relationship between entropy and reliability is much stronger than the analyses in preceding sections. All of the designs in the previous analyses are maximum entropy designs while, in this section, the entropy of the designs ranges from the minimum to the maximum with more or less equal intervals. The fact that the increase of entropy is followed by an increase in the reliability of the network clearly suggests that entropy *is* a good surrogate measure of reliability.

The plots of entropy against the HDA-reliability from the 9 groups of designs in Figure 6.8 almost overlap on top of each other, which suggest that the influence of the cost function on the entropy-reliability relationship is negligible. Meanwhile, the cost plots show a flatter slope as the value of the exponent e increases. This detail can be seen more clearly in Figure 6.9 in which the value of γ is set to be equal to 1 for *all* the designs, which indicate that higher e value results in lower increase in the network cost as the entropy and reliability of the network increase. The fact that the trend of the relationship between cost and reliability does not change when the same or different values of γ are used indicates that the coefficient does not seem to have any effect on the relationship at all. Therefore, it follows that the correlation between entropy and reliability is not affected by the value of γ.

Figure 6.10 shows that the average size of the pipe diameters increases as the entropy value of the network becomes higher. This may contribute to the strength of the relationship between entropy and reliability since larger pipe diameters are in general more reliable. The average value of the pipe diameters also seems to increase with the rise of the e value. Meanwhile, higher entropy value also corresponds to designs with more uniform pipe diameters as shown in Figure 6.11. This assessment is considered appropriate in the present study because all the pipes in the network have equal length. Figure 6.11 also shows that the pipe diameters become more uniform as the e value increases. One possible explanation for this phenomenon is that the high value of the exponent e prohibits the search towards tree-type solutions since, for example, having four smaller pipes in a loop may be cheaper than having three larger pipes with a minimum diameter loop-completing pipe. As a result, the optimization process tries to reduce the size of the large diameter pipes and increase the size of the smaller ones producing more uniform pipe sizes in the resulting designs. This explains the increase in the minimum values of entropy and reliability as the value of e increases since more uniform pipe diameters means more uniform flow distributions, which corresponds to higher entropy and reliability values as depicted in Figure 6.11 and 6.12, respectively. Also, the increase in the minimum entropy value leads to a narrow range of the entropy values since the maximum entropy of the network remains constant. Therefore, when the range of the entropy values in each group was divided into equal intervals, the number of designs that can be generated is reduced as the range of the entropy values decreases.

6.5 POSSIBLE INFLUENCE OF MODELLING ERRORS

In this section the possible influence of modelling errors on the relationship between entropy and reliability is assessed. These modelling errors, which were brought about by several factors, produced redundancy or insufficient capacity in the resulting designs in the form of small surpluses or deficits in head at the critical nodes. The errors are seemingly small and at first glance would appear to be insignificant. However, any impact they may have on the relationship between the entropy and reliability has never been studied. The importance of this issue lies in the fact that reliability values for alternative designs of water distribution systems are generally high which means that the differences in the values tend to be very small. This very small range of variation in reliability leaves open the possibility that seemingly small design and modelling errors may have a significant or disproportionate impact on the calculated relationship between the entropy and reliability.

The analysis was carried out on 74 designs based on the network in Figure 6.1. 43 designs analysed in this section were generated in the present study and have been used previously in the analysis of the possible influence of the cost function on the entropy-reliability relationship. 19 designs were taken from Tanyimboh and Templeman (2000) while the rest of the designs were taken from Tanyimboh and Sheahan (2002). All of the designs use continuous pipe diameters whose rounded-off values produced "small" surpluses or deficits in head at the critical nodes with the largest being –0.4m. The programs used in the design optimisation process (PEDOWDS) and the subsequent hydraulic simulations (PRAAWDS) also have slight differences in some of their coefficients. These differences contributed to the discrepancies in the resulting designs. It may be worth noting that the difference in the HDA-reliability values between the most and the least reliable designs in this section is only 0.000180. The head-outflow relationship of Tanyimboh and Templeman (2004) given in Equation (4.38) was used for the HDA analysis in this section. Its use was intended to show that the strong correlation between entropy and reliability seen in the previous sections was not attributable to the head-outflow relationship used.

Figures 6.13 and 6.14 show the total outflows delivered by the full network and the distribution of the demand satisfaction ratio, respectively. From these two graphs it seems that, under normal operating conditions, the total outflow delivered by most of the designs is approximately equal to the total demand required. This would appear to suggest that the accuracy of the results is acceptable and that the small surpluses or deficits in heads and outflows are insignificant. The plots of reliability and entropy against surplus head for all the designs in this section are shown in Figures 6.15 and 6.16, respectively. The plots suggest that there is no correlation between reliability or entropy and surplus head and thus it may be deduced that any influence of the small surpluses or deficits in heads at the critical nodes upon the entropy-reliability relationship is insignificant. The same conclusions can be drawn from the DDA-based analysis, which can be seen in Appendix A6.

6.6 PERFORMANCE OF DESIGNS WITH EQUAL MAXIMUM ENTROPY VALUES

It has been mentioned at the beginning of this chapter that one of the issues concerning the use of entropy as a surrogate measure for the reliability is that one entropy value does not correspond to a single value of reliability. For entropy to be a good surrogate measure for reliability of water distribution networks, different designs with equal maximum entropy value should have equal or at least similar level of performance compared to other designs with different maximum entropy values so that valid comparisons can be justified. Tanyimboh and Sheahan (2002) have carried out some preliminary investigation on this issue. They carried out the study on the 65 designs with different layouts based on the network in Figure 6.1. The DDA network analysis method was used to perform the hydraulic simulations in their study. The result of their investigation is very promising in the sense that the variations in the reliability between designs with equal maximum entropy values are very low in comparison to designs with different entropy.

In this study, the same approach to that used by Tanyimboh and Sheahan (2002) was employed. However, in addition to the 65 designs used in the layout analysis study (Tanyimboh and Sheahan, 2002), 72 designs that were previously used in the flow directions study (Section 6.3) were used in this section making a total of 137 designs analysed. The DDA and HDA methods were used in this investigation. However, only the results from the HDA analysis are presented in this chapter. The DDA results can be found in Appendix A7. The designs with equal maximum entropy values are grouped together and there are 29 equal maximum entropy groups (EMEGs) in total as depicted in Figure 6.17. The figure also shows the values of coefficient of variation of the reliability (CVR) within the EMEGs together with four possible comparators. These comparators are (Tanyimboh and Sheahan, 2002):

1. The potential range of CVR of all the 137 designs, which is the CVR of the most and least reliable designs. The reliability values of these designs form the upper and lower bounds of the reliability for all the 137 designs and the two reliability values indicate the extent to which the reliability values in this particular study could differ.
2. The CVR of all the 137 designs.
3. The CVR of all the designs outside the EMEGs. There are 60 designs in total.
4. The CVR of all the designs within EMEGs. There are 77 designs in this category.

The weighted average of the CVR of all the EMEGs is also presented in Figure 6.17 to show that the variations in the reliability values of the designs in the EMEGs are on average very low compared to other designs.

The set of results in Figure 6.17 seems to confirm the finding of Tanyimboh and Sheahan (2002) in which the level of performance of designs with equal maximum entropy values are similar. This can be seen in the CVR values of the 29 EMEGs which, in general, are much lower than the CVR of the potential range of the reliabilities and the CVR of the full set of 137 designs. Also, the value of the weighted average of the CVR from all the EMEGs is very much lower than the four possible comparators, which substantiate the above conclusion.

6.7 SUMMARY AND CONCLUSIONS

Sensitivity analysis on the relationship between entropy and hydraulic reliability of water distribution networks has been carried out in this chapter. The study critically assessed several aspects of water distribution networks that may have effects on the entropy-reliability relationship. These aspects are: the chosen layouts and flow directions in the design of the distribution networks, cost functions used in designing the network and minor errors produced in the design and modelling process. The other issue investigated is the similarity in performance of designs with equal maximum entropy values. The study employed the HDA network analysis methods, which been suggested to be superior to the DDA method, particularly for analysing networks under subnormal operating conditions. As a result, the outcome of the analysis should be more realistic. However, the method is relatively new and requires more research and development. Hence, the use of the HDA method in this study is for comparison purposes and to show that the entropy-reliability relationship still holds true when different methods are used in the analysis.

The general conclusions drawn from this chapter are as follows. The results of the assessment reveal that the correlation between entropy and reliability is strong and higher value of entropy corresponds to better network performance. The increase in the network costs with increasing entropy values seems quite modest relative to the increase in the reliability values. The influence of the above-mentioned aspects of water distribution networks on the entropy-reliability correlation is insignificant. The increase in the entropy value of a network leads to a larger average pipe diameters. Higher network entropy also corresponds to more uniform pipe diameters, which contributes towards the increase in the reliability level of the network. Also, the level of similarity in the performance of designs with equal maximum entropy values is very high. Finally, the use of the HDA network analysis method leads to a stronger correlation between entropy and reliability in general.

Further deductions may be drawn from the investigations in this chapter regarding the reason why entropy is a good surrogate measure for the reliability of water distribution networks and how the entropy value of a network under normal condition is related to the performance of the network under abnormal/deficient conditions. It has been shown that by designing a network with maximum/high entropy value, the distribution of flows hence the pipe diameters in the network are as uniform as possible subject to constraints. This condition enables the network to have high degree of flexibility to cope with failure conditions compare to other designs. As a result, the extra head loss due to pipe failure may be kept to a minimum since the flow may be rerouted to other path with no great difficulty. The strong relationship between entropy and reliability in the investigation of the impact of layout demonstrates the ability of entropy to represent the range of flexibility of networks with different failure conditions since each layout is in fact a *failure mode* of the fully connected network. Also, the similarity in the performance of designs with equal maximum entropy values justifies the use of entropy as a surrogate performance measure.

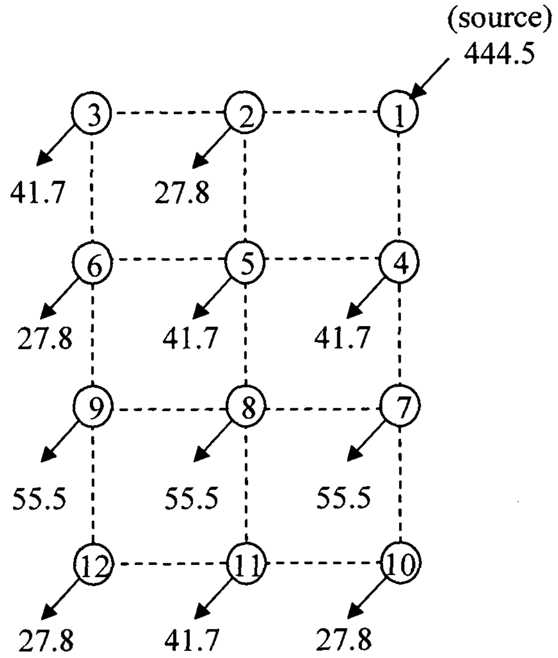


Figure 6.1. Network of supply and demand nodes with demands in litres per second.

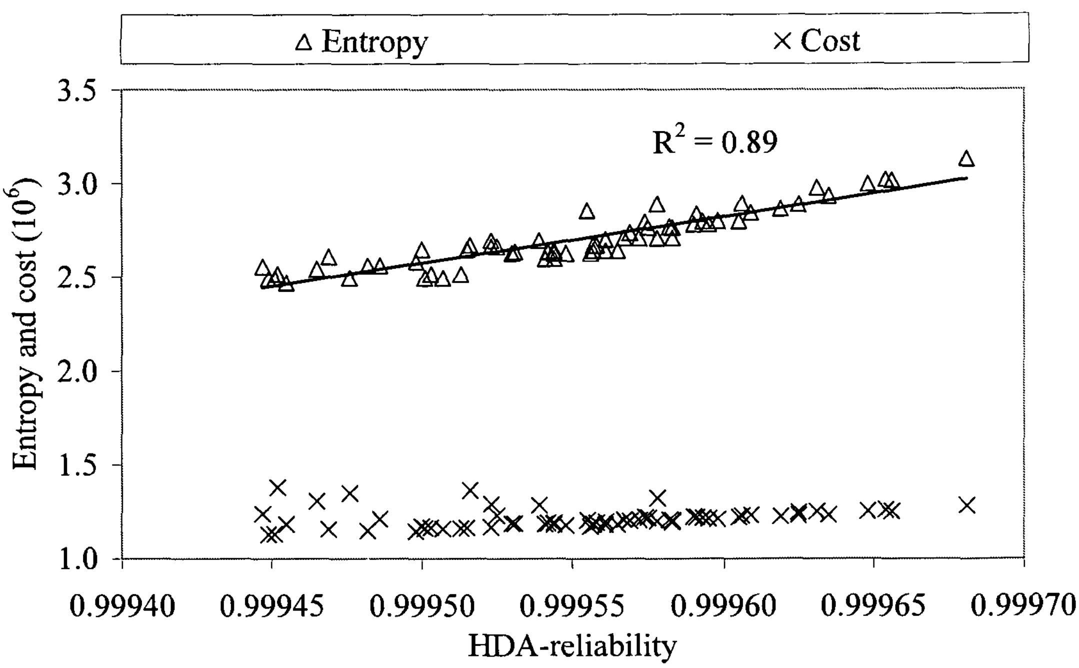


Figure 6.2. Influence of layout on the entropy-reliability relationship analysed by the HDA method.

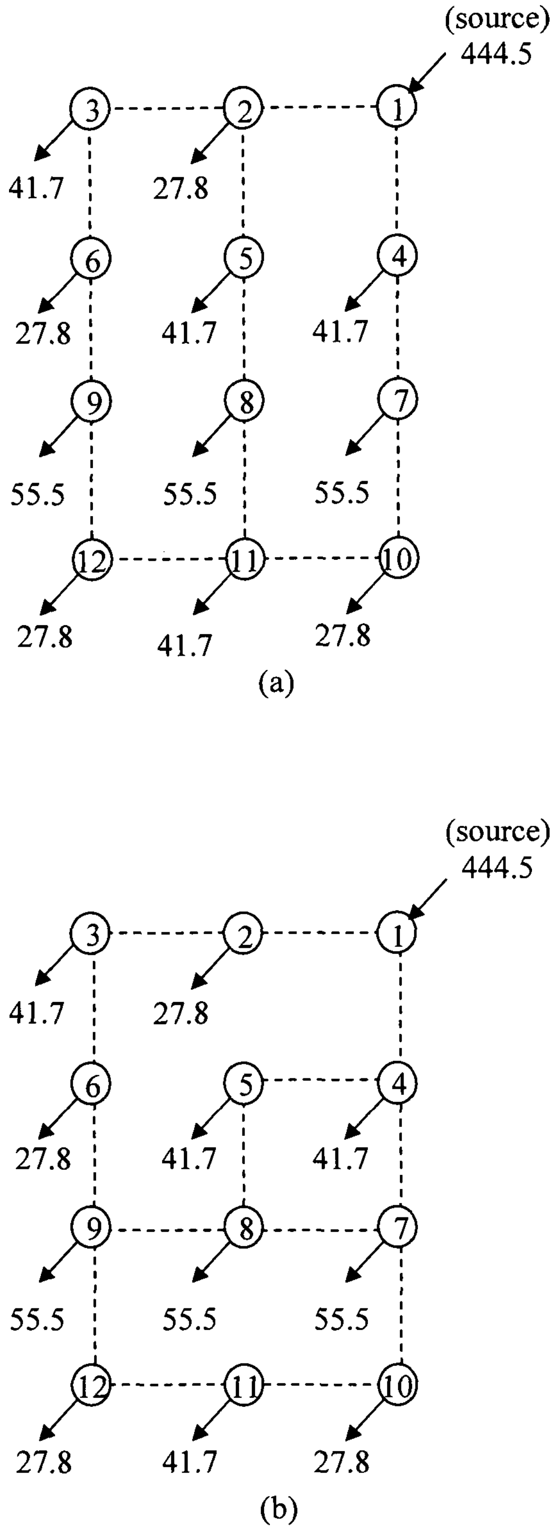


Figure 6.3. Networks of supply and demand nodes (a) 2 loops and (b) 3 loops - with demands in litres per second.

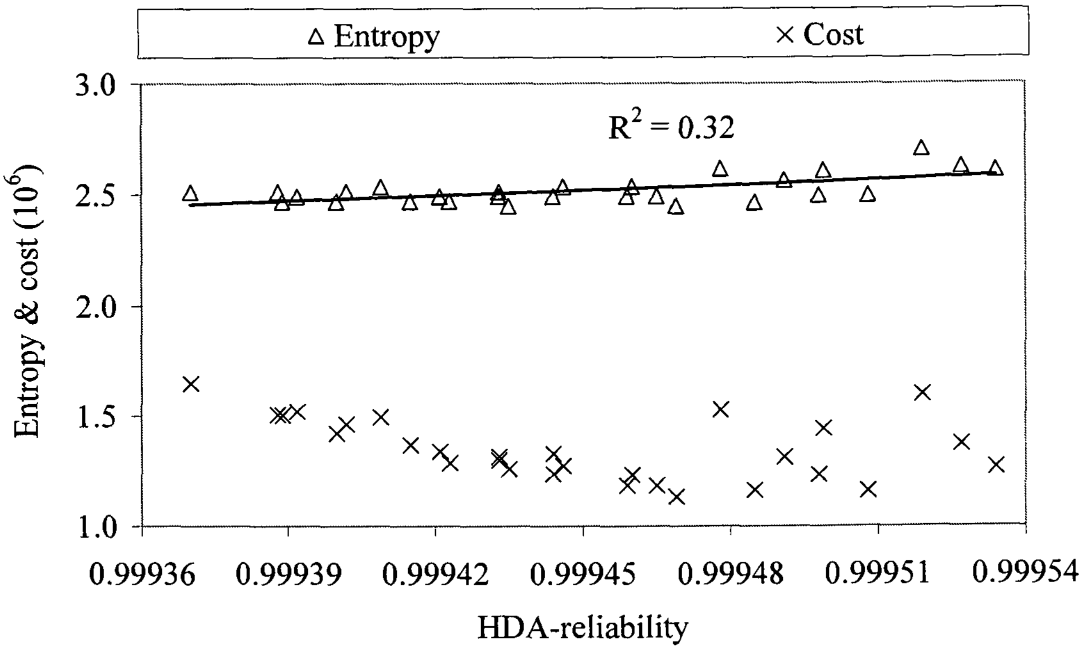


Figure 6.4. Influence of flow direction on the relationship between entropy and HDA-reliability for the 2-loop designs.

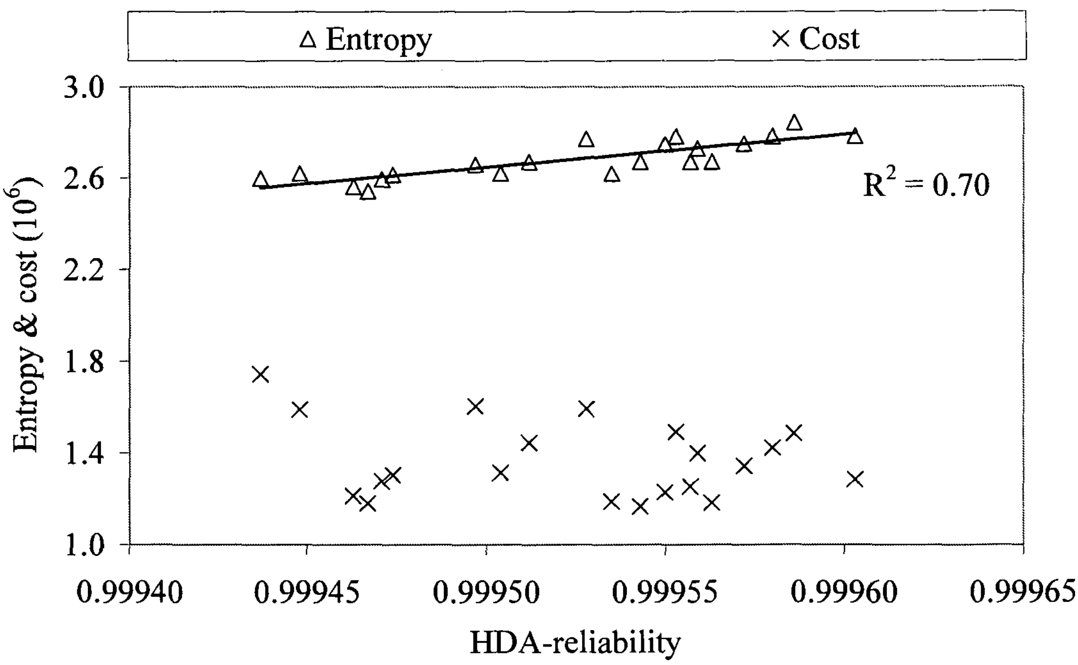


Figure 6.5. Influence of flow direction on the relationship between entropy and HDA-reliability for the 3-loop designs.

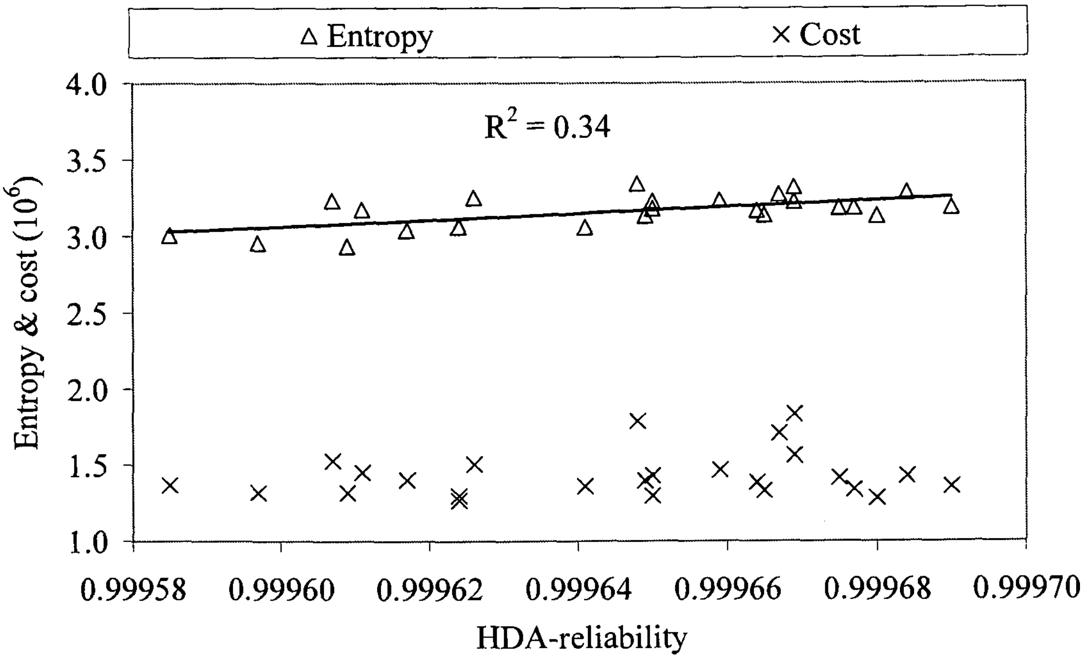


Figure 6.6. Flow direction effect on the relationship between entropy and HDA-reliability for the 6-loop designs.

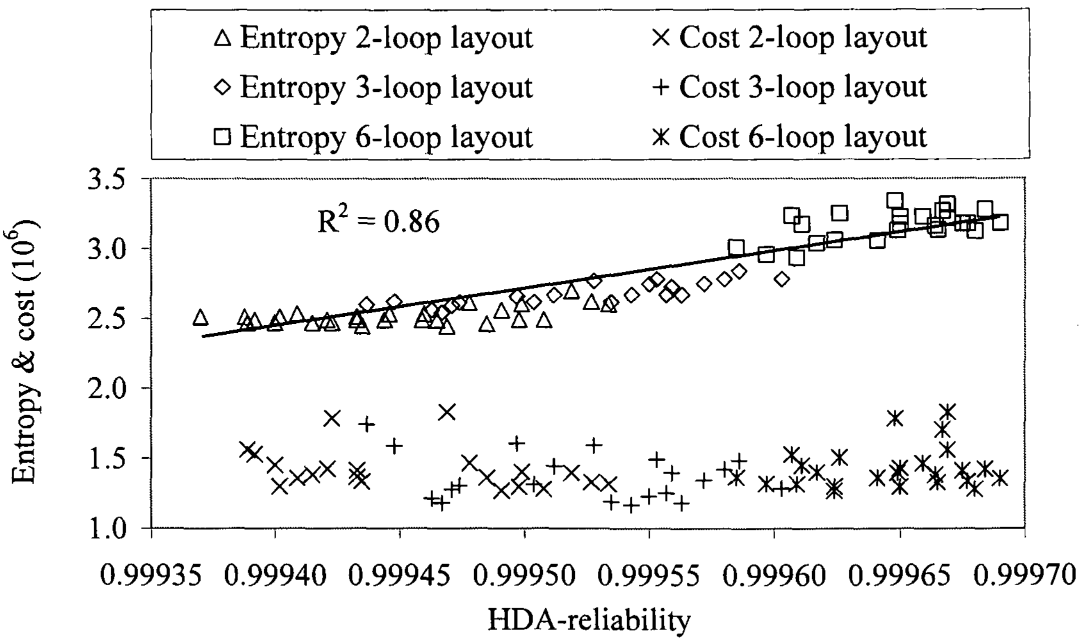


Figure 6.7. Analysis of the possible effect of flow directions on the relationship between entropy and cost against HDA-reliability for the 2-, 3- and 6-loop designs.

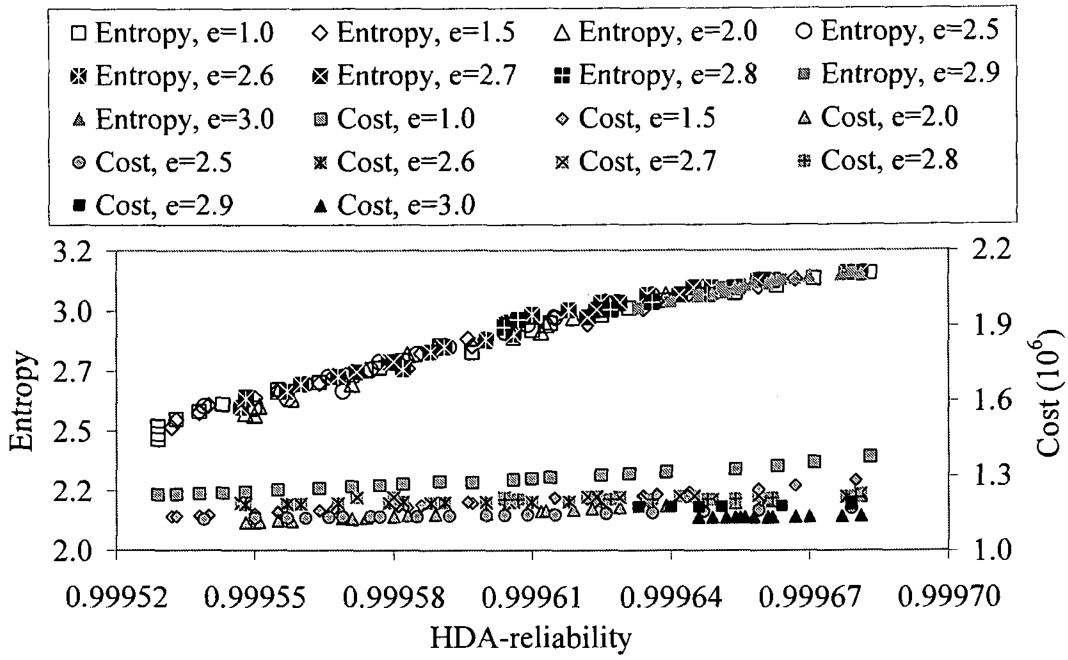


Figure 6.8. Influence of cost function on the correlation between entropy and HDA-reliability.

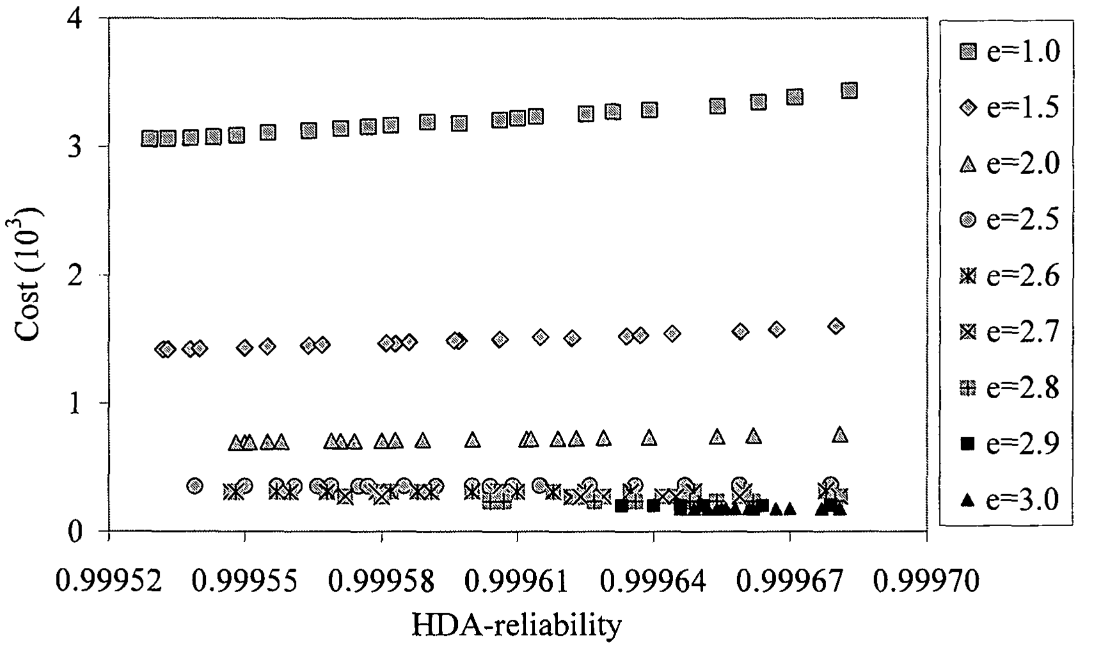


Figure 6.9. Plot to emphasize the slope of the cost functions, all with an arbitrary $\gamma = 1$.

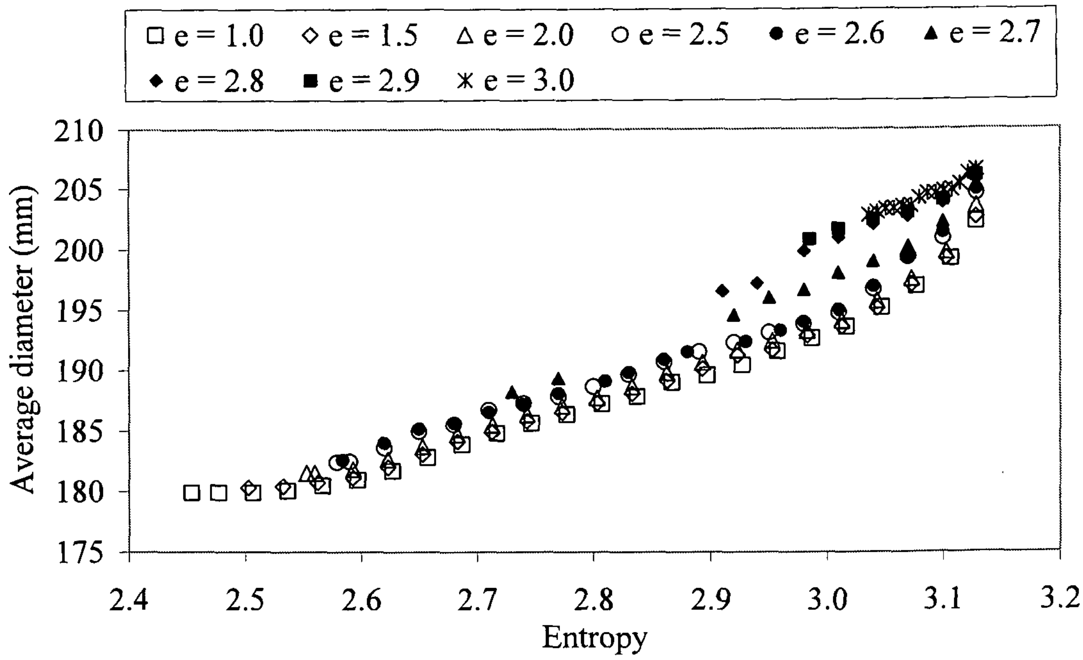


Figure 6.10. Influence of entropy and cost function *e* on the average size of the pipe diameters.

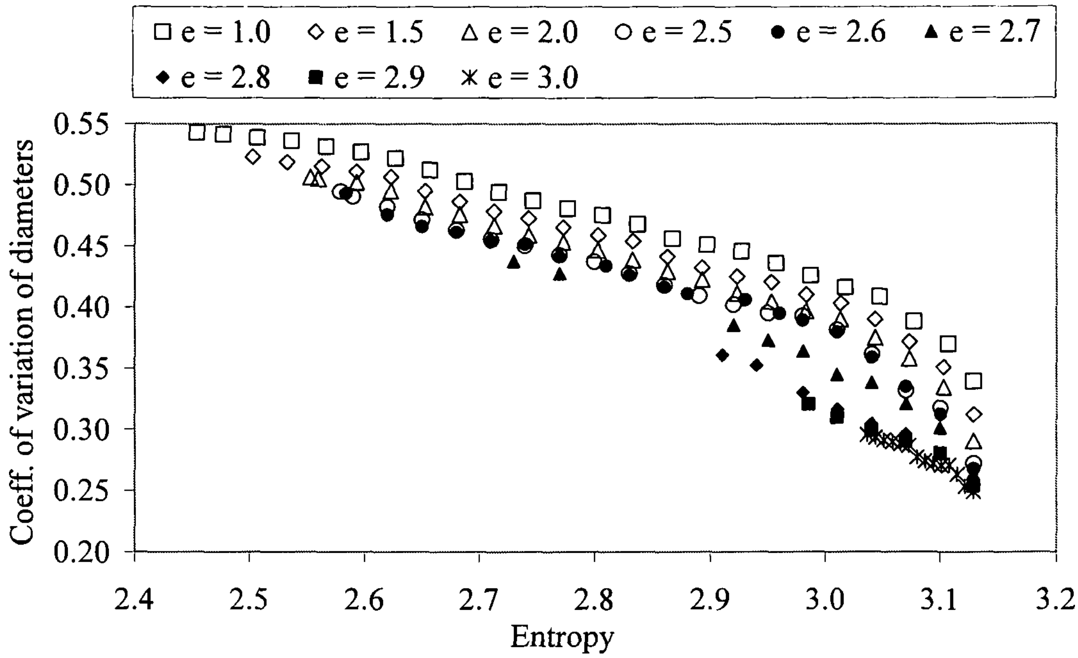


Figure 6.11. Influence of entropy and cost function *e* on the size variation of the pipe diameters.

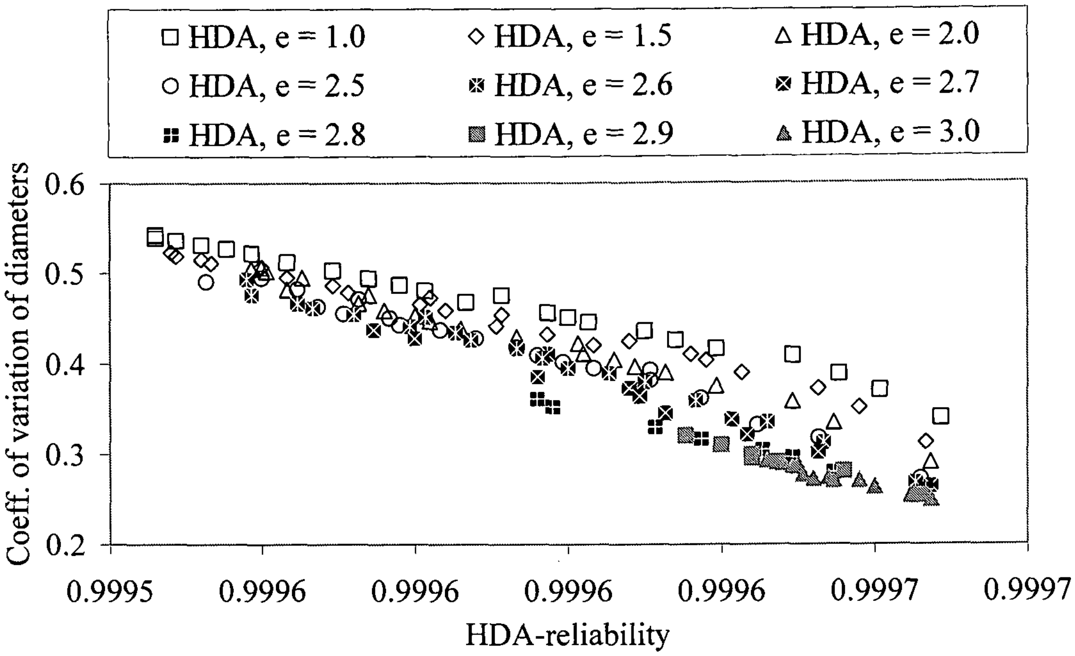


Figure 6.12. Relationship between coefficient of variation of diameters and reliability for designs in the analysis of the cost function effect.

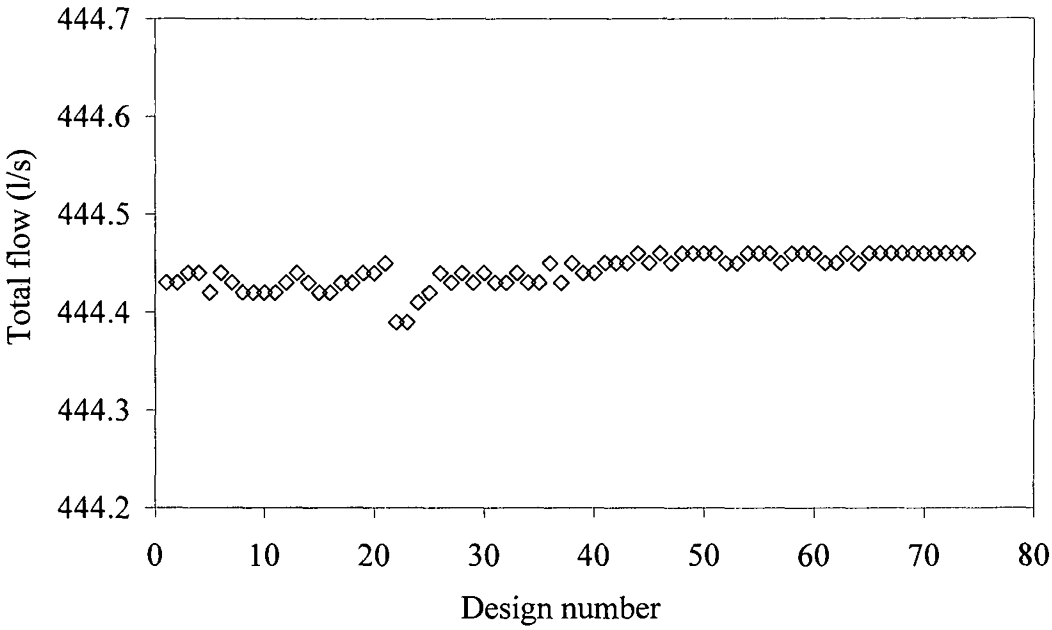


Figure 6.13. Total outflows delivered by the networks under normal condition.

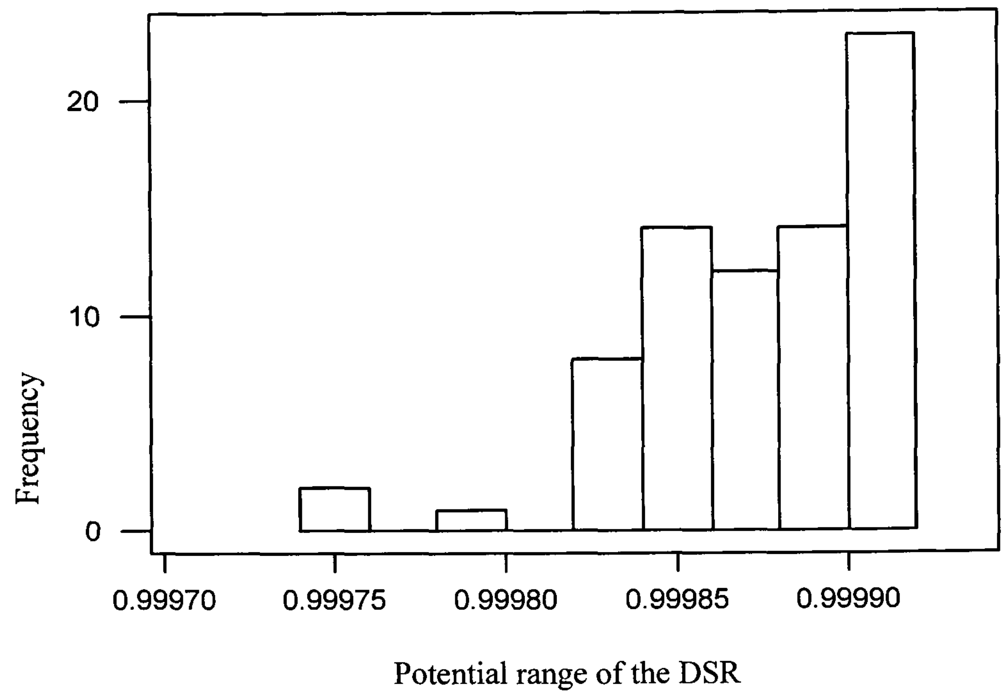


Figure 6.14. Distribution of demand satisfaction ratios.

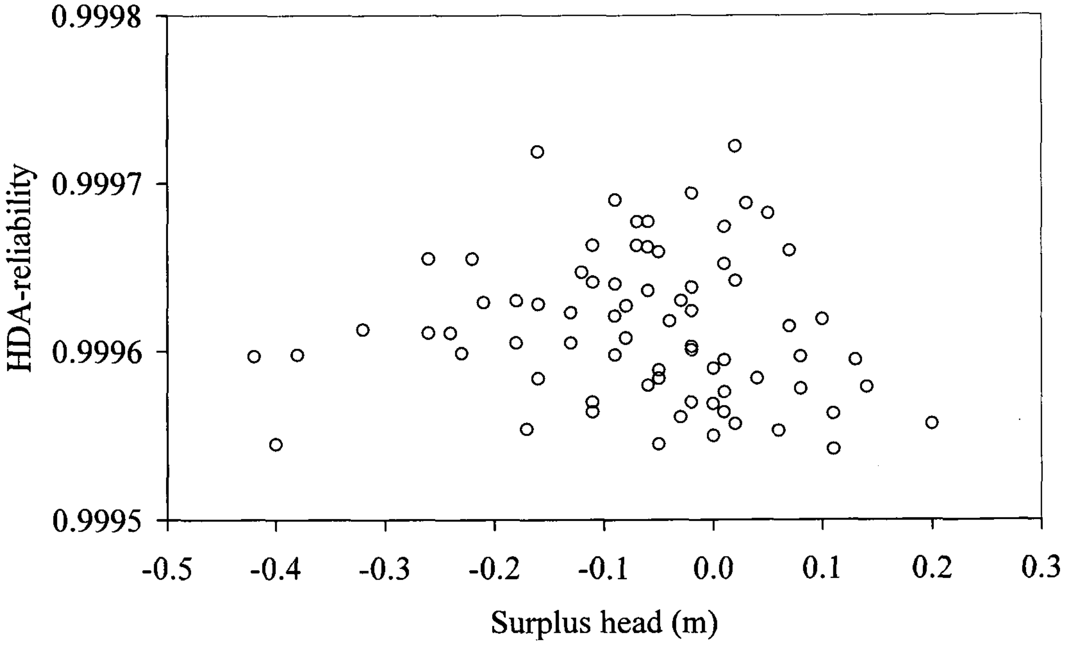


Figure 6.15. HDA-reliability against surplus head.

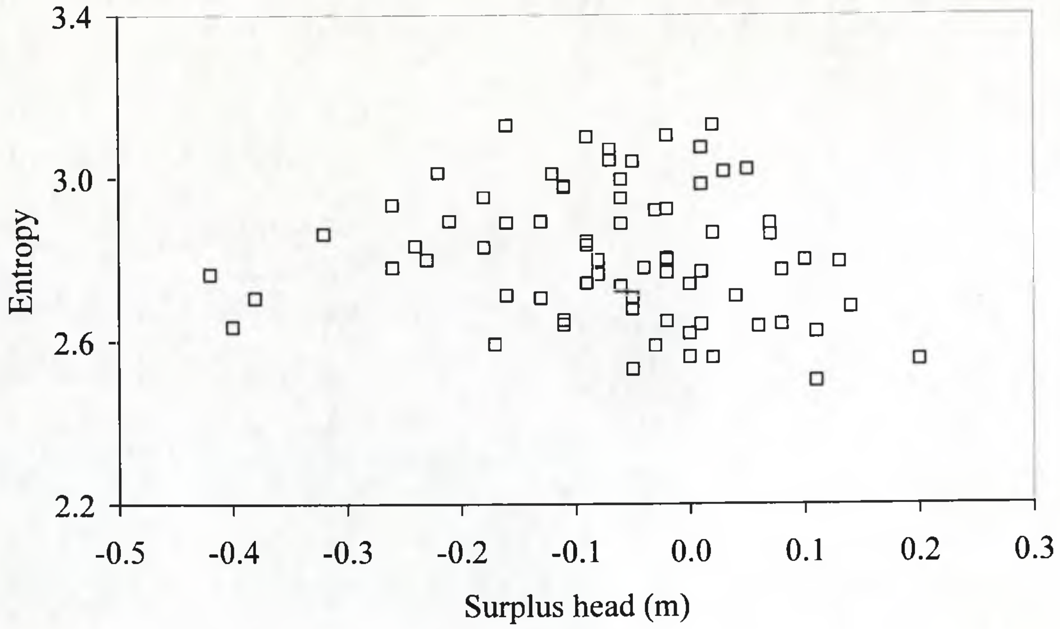


Figure 6.16. Entropy against surplus head.

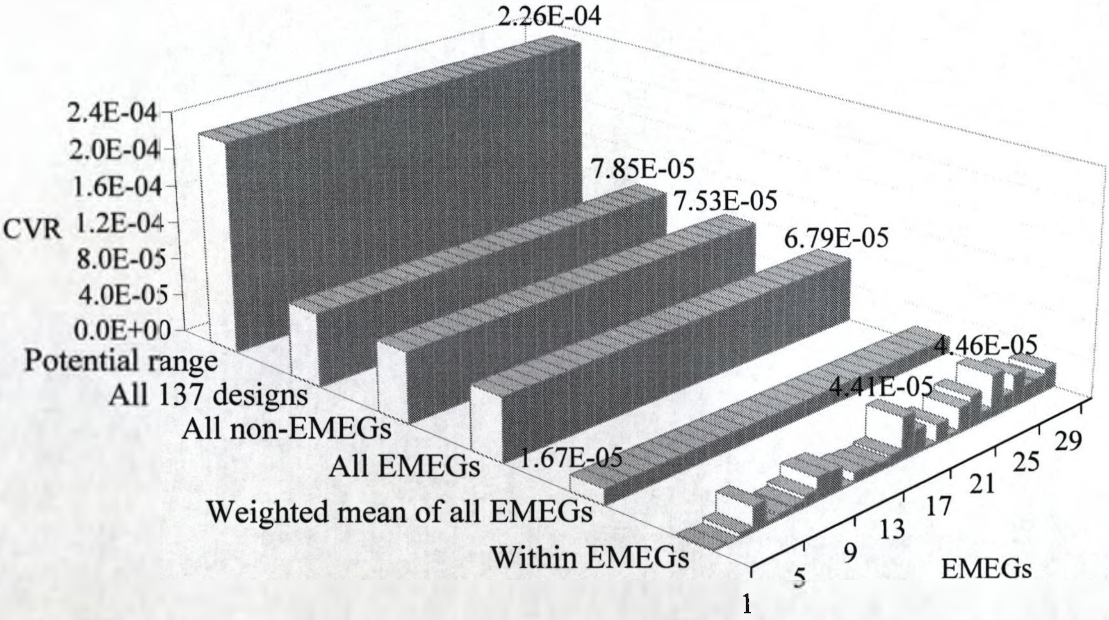


Figure 6.17. Coefficients of variation of the HDA-reliabilities.

CHAPTER 7 HYDRAULIC PREDICTABILITY AND LAYOUT OPTIMIZATION OF WATER DISTRIBUTION NETWORKS

7.1 INTRODUCTION

It is recognised that the layout of a water distribution network has a significant influence on the level of performance of the network (Morgan and Goulter, 1985; Loganathan et al., 1990; Awumah and Goulter, 1992; Afshar et al., 2005). However, layout optimization of water distribution networks is highly complex, especially for large highly-looped networks. This is because there are more than one closely interrelated objectives involved in the process, i.e. selecting the optimum layout, which may involve a large number of potential candidate layouts for looped networks, and selecting the sizes of the components which correspond to an economical design without compromising the level of network performance. On top of that, there is the difficulty of quantifying the performance of water distribution networks. As such, most design optimization algorithms do not consider layout and performance optimization simultaneously. Although the above mentioned researchers have proposed several procedures to overcome this very problem, their approaches are somewhat unsatisfactory since most of the procedures do not actually quantify the reliability of the layout, while others use only connectivity as a measure of the reliability, which has been shown to be misleading (Wagner et al., 1988a).

Another important aspect in the design of water distribution systems is the identification of the probable failure points. Several design optimization models require the critical nodes and links in the network to be identified prior to the formulation of the design problem (Xu and Goulter, 1999; Tolson et al., 2004). This, however, is a very complex issue since the behaviour of water distribution networks, especially those with loops, is in general very difficult to predict. The flow directions in the network can change constantly due to the continuous spatial and temporal variations in the nodal demands. Also in a critical operating condition such as pipe failure or fire fighting situation, the flows in the network are rerouted in complex ways and therefore it is impossible to determine with complete certainty which areas are worst affected without modelling and simulating the deficient network. In the design process, the preliminary designs are usually checked to see whether they can cope with the many emergency situations that may occur during their design life. Depending on the size and complexity of the network, however, it is often impractical to consider all the potential critical operating conditions during the design phase. Hence, it is highly desirable to consider optimum design techniques which can produce designs with less complex and more predictable hydraulic behaviour.

Entropy has been shown to be a suitable surrogate measure for the reliability of water distribution networks. Also, incorporating entropy in the design optimization problem is relatively simple and uncomplicated. In this chapter, the potency of entropy-based designs of water distribution networks in the optimum layout selection and in predicting the critical components in the network are investigated. Preliminary investigations on the two issues have been carried out by Tanyimboh and Sheahan (2002) and Tanyimboh (1993) using the DDA method. The results of their investigations were very encouraging. The present studies are intended to provide more evidence and strengthen the previous contentions on the issues. On top of that, the more realistic HDA method is employed in the present studies.

7.2 HYDRAULIC PREDICTABILITY OF ENTROPY CONSTRAINED DESIGNS

The hydraulic predictability of a network in this study is assessed in terms of the location of the critical links and critical nodes in the network. These critical links and nodes are identified by analysing the performance of the network under a range of normal and abnormal operating conditions. The following three critical operating conditions are considered.

1. The unavailability of individual pipes. As explained in Chapter 5, the isolation of a pipe or group of pipes depends on the number and location of valves in the

network. In this study it is assumed that each pipe can be isolated individually, which in practice may not be the case.

2. A large demand of 0.25 m^3/s at each node in turn, in place of the normal demand at that node, with all other nodal demands at their normal design values. This situation would be akin to a fire occurring under a locally ideal (i.e. zero) background demand condition.
3. A large demand of 0.25 m^3/s at each node in turn, in addition to the normal demand at that node, with all other nodal demands at their normal design values. This situation represents a fire-fighting demand occurring under a more adverse background demand condition. This scenario is also more realistic since the nodal demands in real water networks will not be zero whenever there is a fire occurring.

In a link failure situation, the critical pipe is identified as the pipe the removal of which causes the greatest deterioration in the network performance, while the worst affected node as a result of the removal of the critical pipe. In a fire fighting scenario, the critical node is the node at which the occurrence of fire-fighting demand would generate the most adverse effect on the network. The way in which the deterioration in the network performance is measured is explained shortly.

The criteria used to measure the performance of water distribution networks in the HDA analysis are slightly different from those in the DDA. In the HDA analysis, the network performance is measured by using the outflow and the *DSR* value of the whole network as well as at each demand node individually. The hydraulic reliability and hydraulic redundancy for each demand node and for the entire network are also calculated using Equations (5.12) and (5.14), respectively. Another measure used to assess the performance is the total amount of energy dissipated by the network given in Equation (5.15). Although the 'head loss', 'flow delivered' and 'energy dissipated' are not unrelated, the properties they measure appear to differ in subtle ways and sometimes they lead to different results as will be seen later in this section. To improve readability and for consistency with the previous chapter the DDA results are presented in Appendix A8.

In the HDA analysis the outflows change following the change in pressure in the network. Also, the change in link flows leads to the variation in the energy dissipated by the network. Consistency in the comparisons between the HDA-based results is achieved using a re-defined measure for the HDA method, i.e. the energy dissipation rate per unit flow delivered by the network is used in the HDA analysis as opposed to the total dissipated energy alone. For the same rate of water supply, the amount of energy dissipated would increase as the head loss or stress on the network increases. Therefore, the rate of energy dissipation per unit flow, taken over the entire water distribution network, would appear to represent a measure of hydraulic performance.

Experienced engineers, however, often use rules of thumb to identify critical pipes in water distribution networks. These criteria may include the following:

1. The location of the pipe relative to the source: Pipes which are far from the source are likely to be less critical.
2. The location of the pipe relative to the major demand nodes: Pipes which are not in the vicinity of major demand nodes are likely to be less critical.
3. The number of alternative supply pipes incident on the major demand nodes: A single pipe supplying a major demand node is likely to be more critical than each pipe in a group of pipes supplying another node with a similar demand. This also holds true for pipes connected to the sources.

Also, the criteria for critical demand nodes may include:

1. The distance of the node from the source: Nodes which are far from the source are likely to be more critical.
2. The number of paths supplying the node: Nodes with multiple supply paths are likely to be less critical.
3. The normal demand at the node: Nodes with large demands are likely to be more critical.
4. The increase in demand at the node during a fire-fighting situation: Nodes with relatively large increases in demand can be more critical.

Due to the complexity of water distribution networks, their behaviour is not always in tune with these *hydro-spatial* expectations. Therefore, water distribution networks may be classified as hydraulically predictable if their characteristic behaviour follows the above criteria.

Also, thanks to the ability to calculate the actual amount of water delivered at each demand node using head-dependent modelling, the hydraulic reliability and redundancy of these nodes can now be assessed and compared. Tanyimboh et al. (2001) have provided evidence which suggests that, if other factors are equal, the hydraulic reliability and redundancy of a demand node decrease as its distance from the source increases. The unavailability of a pipe has an adverse effect on downstream rather than upstream nodes. Also, in general, the pipes in a water distribution network reduce in size progressively from the source towards the more remote parts of the network. Another relevant issue is that, under both normal and critical conditions, the nodal pressures tend to decrease as the distance from the source increases. Based on these considerations, it seems reasonable to expect the nodal hydraulic reliability and redundancy to decrease as the distance from the source increases. However, in general, the scale and complexity of water networks are such that there is no guarantee that this property will hold true in any given situation. Consequently, a distribution network in which the nodal reliability and redundancy values decrease with the distance from the source may be said to be spatially predictable with respect to the hydraulic performance of its nodes. Conversely a distribution network may be said to be less spatially predictable with respect to its nodal hydraulic performance, if the correlation between the proximity to the source and the nodal reliability and redundancy is not high.

Preliminary studies based on the DDA analysis by Tanyimboh (1993) indicate that designs carrying maximum entropy flows are more hydraulically predictable than the conventional optimum designs. The analysis was carried out on two simple networks reproduced here in Figure 7.1. All links in both networks have a length of 1000 m and Hazen-Williams coefficient of 130. The designs based on these networks were generated using continuous pipe diameters and the value of γ and e for the cost function in Equation (5.16) were 900 and 2.4, respectively (Tanyimboh, 1993). The minimum and maximum pipe diameters were taken as 0.1 m and 0.6 m, respectively. The minimum required residual head at demand nodes was 30 m for both networks and the piezometric head corresponding to zero flow was 0 m, which was the elevation of the demand nodes. The two-loop network was taken from Alperovits and Shamir (1977) with the pipe between the reservoir and the first demand node eliminated for simplicity (Tanyimboh,1993). The supply at node 1 was therefore the net flow and the piezometric head at this node was assumed to be 50 m to achieve residual head of 30 m at the terminal nodes under normal operating conditions. The four-loop network, which was taken from Fujiwara and de Silva (1990), on the other hand, has a piezometric source head of 53.5 m. For each network, the design optimization program PEDOWDS was used to produce alternative designs. For each of the designs, the program was run several times with different starting points to ensure the global and not local optimum solution was achieved. The entropy value of the designs ranges from the minimum to the maximum. While the designs for the 2-loop network were generated to achieve a uniform coverage of the entropy range, the entropy values for the 4-loop designs were selected to achieve a greater coverage of the upper end of the entropy range to provide a comparison of near-maximum-entropy designs (Tanyimboh, 1993). To facilitate the comparison between DDA and HDA results, the HDA analysis was performed on the same networks and the results are presented and discussed next.

7.2.1 TWO-LOOP NETWORK

For the two-loop network, six designs with a range of entropy values were generated and their pipe diameters are shown in Table 7.1. The table also shows the average, standard deviation and coefficient of variation of the diameters as well as the total cost of the pipe networks in million pounds for all the designs (Tanyimboh, 1993). Looking at the standard deviations and the coefficient of variations in the table, it is quite clear that the higher the entropy value of a network, the more uniform the pipe diameters are. This apparent relationship strengthens the conclusion drawn earlier in the previous chapter. The average of the diameters, however, seems to increase with the increase of the entropy value, which causes the rise in the total cost of the pipes. This rise in cost is compensated by the increase in the network performance as will be seen and discussed later in this section. The nodal demands under normal condition are given in Figure 7.1a while the fire-fighting demands in place of and in addition to the normal demands are listed in Table 7.2. It is worth stating that the analysis of the two networks (Figure 7.1) under fire-fighting load in which the fire demand is added to the normal demand was not carried out by Tanyimboh (1993). However, the DDA results of this analysis, generated in the present study, are presented in the appendix alongside other DDA results of Tanyimboh (1993) for ease of reference.

In the HDA analysis, the critical link is the link which, when fails, will cause the highest reduction in the total network outflow. The determination of the critical node in the HDA analysis is somewhat context dependent. In the case of the failure of a pipe, the critical node could be considered as the node with the smallest post-failure demand satisfaction ratio or the largest post-failure absolute reduction in the flow delivered. The use of the nodal demand satisfaction ratio together with the actual shortfall might help ensure that due consideration is given to all demand nodes. For example, a demand satisfaction ratio of zero is unlikely to be missed, even if the normal demand at the node in question is small. Similarly, a large shortfall in the flow delivered would not be overlooked, even if the node in question has a high DSR.

The results from the HDA analysis in Table 7.3 are in line with the expectation that pipe 1-3 is the critical pipe in the network for all the designs. Pipe 1-3 is connected to the source and, under normal operation condition, the pipe lies in the only path supplying the node with the largest demand in the network, i.e. node 5. Table 7.3 also shows that the flow and energy performance indicators select pipes 3-5 and 1-2, respectively, as the next critical link. The reason for the discrepancy is not immediately obvious. However, each result seems reasonable since pipe 3-5 is connected to Node 5 which has the largest demand in the network while pipe 1-2 is connected to the source. Based on the DSR value at each demand node (Table 7.4), only two designs with the highest entropy values agree with the expectation in which the corresponding critical node due to the critical link failure in the network is node 5. However, when the nodes are assessed using their actual shortfall in flow delivered, node 5 is shown to be the critical node for all the designs. There is also a disagreement in the determination of the next critical node. Based on the nodal DSR, the next critical node is node 3 while the assessment using actual shortfall at each demand node shows that the next critical node is node 4. Both results are justifiable given that node 3 is supplied by a single path under normal conditions and node 4 has the second largest design demand after the critical node. This reinforces the idea that it may be necessary to use both measures for a more thorough assessment.

For the 2-loop network in Figure 7.1a, node 6 should intuitively be the worst location for the fire-fighting load to occur. Node 6 is the terminal node and therefore should have the lowest residual head in the network. In the case of fire-fighting situation with a zero background demand, node 6 experiences high increase in demand since its design demand is relatively low. Although nodes 2 and 3 experience larger absolute increase in their demands in a similar situation, their distances from the source node are shorter, hence they should have much higher residual heads under normal circumstances which should help reduce the effect of fire-fighting demands occuring at those nodes. When fire load is superimposed to the design demand, the critical node can be either node 6 or node 5. Node 5 has the largest design demand in the network. Hence, large additional load occurring at the node may stress the network considerably. The reason that node 6 may be critical when fire demand occurs at the node is as has explained above.

The determination of the critical node in the HDA analysis for networks under fire fighting conditions could be either in absolute terms (i.e. the actual shortfall at the node) or in terms of the nodal demand satisfaction ratio. Another interpretation of the term 'critical node' might also be the node at which a very large increase in demand results in the smallest system DSR or the largest absolute shortfall in the flow delivered by the entire network. In assessing the critical nodes as a result of the fire-fighting demand replacing the design demand, the assessment based on the system DSR and the total flow delivered yielded different outcomes since the total demand changes as a result of fire demand occurring at different nodes. This, however, does not occur when fire demand was superimposed on the design demand. Tables 7.5 to 7.7 show broad agreement between the assessment criteria and the results are more or less as expected with the high entropy designs being more predictable in terms of their response than the low entropy designs. For example, for the designs with smaller entropy values, node 2 is the critical node in the network. However, for the designs with higher entropy values, the critical node is node 6. This suggests that the designs with the highest entropy values are more hydraulically predictable than the designs with the smallest entropy values.

The results of the analysis in Figure 7.2 indicate that as the entropy of the network increases, the average and the minimum network outflows also increase. From the assessment of the designs under three different critical conditions, the design with the maximum entropy value of 1.915 in Figure 7.2seems to have a better performance than the others. Figure 7.3 tells a similar story. Looking at the figure in conjunction with

Figure 7.2, it seems that high entropy designs are able to deliver more water without a high increase in energy compared to designs with smaller entropy. This condition is more apparent when the designs are subject to the two fire fighting conditions. From the link failure analysis, however, the average and the maximum values of the total dissipated energy seem to increase with the increase of the entropy value of the network. A possible explanation is that, in link failure situation, since the network seems to be able to deliver more water as its entropy value increases, the head losses in the network also escalate despite the slight increase in the average size of the diameters. It is well known that the relationship between flow and head loss in a pipe is $h \approx kq^2$, in which k is the pipe coefficient. Hence, the increase in pipe head losses as a result of an increase in pipe flow rates follows the exponential relationship. Also, by considering the way the dissipated energy in a pipe network is calculated, i.e. Equation (5.15), the increase in the average and maximum values of the total dissipated energy in Figure 7.3 is therefore justifiable. Thus, it can be seen that, even after normalising the energy dissipated by dividing by the total flow to obtain the energy dissipated per unit flow, the energy per unit flow increases at a faster rate than the flow rate.

Figure 7.4 shows clear improvements in the hydraulic reliability as the entropy increases, for each node and the entire network. Figure 7.5 shows that, from the smallest to the maximum entropy value, the hydraulic redundancy is maintained at more or less the same level, except for node 2 where a large improvement is observed. With regard to the spatial hydraulic predictability explained earlier, Figure 7.4 shows that the reliability of node 2 is low compared to node 5 for the design with the smallest entropy value. This would appear to be counterintuitive given that node 2, which has a smaller demand, is closer to the source than node 5 (see Figure 7.1a). By contrast, the reliability of node 2 for the designs with higher entropy values is higher than the rest of the nodes. This would appear to suggest that, with respect to space, the entropy-constrained designs are hydraulically more predictable when compared to the traditional minimum-cost design whose entropy value is 1.578.

Figure 7.4 also shows that node 4 has smaller reliability values than node 5 for the two designs with the smallest entropy values while node 4 has higher reliability values than node 5 for the rest of the designs. Node 4 has two supply paths and a demand of 75 l/s while node 5 has only one supply path and a larger demand of 92 l/s (Figure 7.1a). In general, the head loss associated with a small demand is less than a large demand. These two factors, i.e. more supply paths and a smaller demand for node 4, suggest that it would not be unreasonable to expect node 4 to have a higher reliability than node 5. The four designs with the highest entropy values reflect these considerations while the two designs with smaller entropy values do not, suggesting that the spatial hydraulic predictability increases as the entropy value increases.

Comparing nodes 2 and 3, which are equidistant from the source, node 2 has a smaller demand, so its redundancy could be expected to be higher than node 3. Figure 7.5 shows that node 2 has a smaller redundancy than node 3 in the two designs with the smallest entropy values. By contrast, the redundancy of node 2 is higher than node 3 in the four designs with the highest entropy values. Similar arguments apply to the corresponding reliability values in Figure 7.4. These results would appear to reinforce the idea that high entropy values bring about higher levels of spatial hydraulic predictability.

It may be recalled that the hydraulic redundancy parameter used herein is a measure of the fraction of the demand satisfied on average during operating conditions with a partial system failure (Equation 5.14). Figure 7.5 shows that, for the two designs with the smallest entropy values, node 5 is more 'failure tolerant' than node 4. This seems counterintuitive as node 4 has two supply paths and a smaller demand than node 5, which has only one supply path. For the four designs with the highest entropy values, node 4 has a higher hydraulic redundancy than node 5 unlike the two designs with the smallest entropy values. This would appear to support the idea that higher entropy values correspond to higher levels of spatial hydraulic predictability under abnormal operating conditions.

7.2.2 FOUR-LOOP NETWORK

The discussion for the four-loop designs is much shorter than that for the two-loop designs since the results are broadly similar. The findings are intended to provide evidence that the conclusions drawn based on the two-loop network are not network specific. The pipe diameters for the four-loop designs are shown in Table 7.8 together with their statistics and total costs (Tanyimboh, 1993). Similar to the designs of the two-loop network, Table 7.8 shows that as the entropy value of the network increases the pipe diameters in the network become more uniform in size. Also, it is clear from the table that maximum entropy designs distribute the flows in the network as equal as possible. For the symmetrical network of Figure 7.1b, the distribution of flows and hence the diameters in the network are also symmetrical Table 7.8 column 6). Fire-fighting demands for the four-loop designs are shown in Tables 7.9 and 7.10 while the normal demands are shown in Figure 7.1b.

For a symmetrical network such as that in Figure 7.1b, it is logical to expect that the critical link should be either link 1-2 or 1-4 since they are the closest to the source and therefore carry the largest flows. The results of the HDA analysis in Table 7.11 show that designs with higher entropy values comply with this assumption. Also, when either of the two pipes fails, node 9 was found to be the critical node as would be expected since it is the farthest from the source (Table 7.12). The selection of nodes 6 and 8 as the critical nodes when fire-fighting demand replaces the normal demand in the HDA analysis (Tables 7.13 and 7.14) might seem odd at first glance. It seems that the critical node for the maximum-entropy design based on the HDA analysis should be at the farthest node from the source in terms of its reachability, which is node 9 (the terminal node). But since the increase from normal to the fire-fighting demand at that node is less that at any other nodes, the critical node is therefore shifted to the next farthest node, which is node 6 or 8. Again, when the fire demand is superimposed to the design demand the intuitive idea that the critical node is located at node 9 is confirmed by the performance criteria (Table 7.15).

The results of the four-loop network in Figures 7.6 and 7.7 confirm the deduction drawn from the two-loop network that as the entropy value of the network increases the performance of the network under critical conditions also improves. Figure 7.8 shows that the reliability of node 3 for the design with the smallest entropy value is low relative to its location with respect to the source. For the three designs with the highest entropy values, the rank order of the nodal reliabilities (largest to smallest) is as follows: nodes 2 and 4; node 5; nodes 3 and

7; nodes 6 and 8; node 9. This order mirrors the respective distances of the nodes from the source exactly, except that node 5 has twice as many supply paths as nodes 3 and 7 and so node 5 has a higher reliability than nodes 3 and 7. It may be noted that node 9 has twice as many supply paths as nodes 6 and 8. However, in addition to its greater distance from the source, the demand at node 9 is three times larger. Figure 7.8 also shows that the three designs with the highest entropy values are symmetrical while the two designs with the smallest entropy values are not. Also, a clear overall improvement in reliability as the network entropy value increases can be observed.

Figure 7.9 shows that the three designs with the highest entropy values have nodal redundancy values which reflect the symmetry of the network, the respective distances from the source, the number of paths supplying each node and the magnitudes of the demands. On the other hand, the two designs with the smallest entropy values lack symmetry and the relative magnitudes of some of their redundancy values are not easy to explain. The figure also shows a small overall improvement in the hydraulic redundancy as the entropy increases. The above analysis would appear to suggest that designs with high entropy values are hydraulically more predictable in a spatial sense.

7.2.3 CONCLUSIONS

It may be noticed that, at first glance, all the graphs in the above investigations do not show strong apparent relationships for definite conclusions to be drawn. However, closer observation shows that each graph provides small evidence which indicates that higher entropy corresponds to a network with better performance. On top of that, all the evidence from the various assessments from the two networks supports one another with no contradiction. Based on this premise, several conclusions can therefore be drawn from the above analyses. Firstly, the performance level of maximum entropy designs seems higher in comparison to other designs. This high level of performance is closely related to the high level of uniformity in the pipe diameters. Secondly, the HDA analysis method is superior to the DDA method in the sense that it is able to measure the performance of individual nodes in a water distribution network more accurately by using the nodal DSR and the actual amount of water delivered. Finally, with regard to the hydraulic predictability of water distribution networks, maximum entropy designs appear to be more spatially predictable than other designs.

7.3 LAYOUT OPTIMIZATION USING MAXIMUM ENTROPY APPROACH

Tanyimboh and Sheahan (2002) have demonstrated the method for finding the optimum layout of water distribution networks using the maximum entropy approach. Their investigation was based on the DDA method and the analysis was carried out on the six-loop network shown earlier in Chapter 6 (Figure 6.1). The approach starts by identifying all the layout candidates, which have a close-looped configuration, and calculating their maximum entropy values. Also, the flow directions need to be specified prior to the calculation of the maximum entropy values. 65 layouts were identified and are shown here in Appendix A3. For these layouts, Tanyimboh and Sheahan (2002) specified the flow directions intuitively based on the shortest path from the source to each demand node to give minimum network costs. Branch layouts are excluded on the grounds of flexibility, hence reliability. This process may at first seem tedious and time consuming, especially for large networks. However, the procedure may be automated on a digital computer without too much difficulty. Once all the layout candidates with the corresponding maximum entropy values have been obtained, maximum entropy designs and their associated costs based on these layouts can be generated. The process of generating the designs has been explained in Chapter 6. The next step is to identify the designs that belong to the set of cost-entropy *Pareto* optimal designs. To improve readability, the definition of a Pareto optimality condition is given below within this section.

A Pareto optimality condition is usually used in the process of finding the trade-off curve in a multi-objective optimization problem where the objectives are usually in conflict with one another. A Pareto optimal solution is a solution of a vector of multi-objective optimization problem in which each individual outcome in the solution vector is optimum in at least one of its objectives and there must be at least one optimum solution with its objective better than the other. It is perhaps easier to explain this definition by means of a mathematical illustration (Marti, 2005) and with reference to a graph (Figure 7.10). Assume that $\underline{F}(\underline{x}, \underline{y})$ is a vector of multi-objective optimization problem to be minimized whose decision variables are contained in the set D_x and D_y. The vector optimization problem is therefore

$$\underset{\substack{\forall x \in D_x \\ \forall y \in D_y}}{Minimise}\ \underline{F}\left(\underline{x}, \underline{y}\right) = \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \\ \vdots \\ F_m(x,y) \end{pmatrix} \tag{7.1}$$

The above equation may be translated into a graph such as that in Figure 7.10. Each point in the figure represents the solution of one of the optimization problems in the above vector. The vector $(\underline{x}, \underline{y})^*$ is a Pareto optimal solution to the problem (represented by the star points in Figure 7.10) which should satisfy the following conditions

$$F_i\left(x,y\right)^* < F_i\left(x,y\right), \quad i = 1,2,\ldots,m \tag{7.2}$$

and

$$F_j\left(x,y\right)^* < F_i\left(x,y\right)^* \quad \text{for at least one } j, \quad 1 \le j \ne i \le m \tag{7.3}$$

Equation (7.2) ensures that each component in the Pareto Optimal solution is an optimum solution while Equation (7.3) provides a condition that at least one of the objectives in one of the optimum solutions is better than the objective in the rest of the optimums so that the Pareto optimal curve can be plotted. The cost-entropy Pareto optimal designs are therefore the maximum entropy designs and their corresponding costs, which satisfy the above conditions.

Once all the designs in the cost-entropy Pareto optimal set have been identified, in which each design corresponds to a single layout configuration, the reliability values of these designs are calculated. This approach has the advantage over other layout optimization methods in the sense that the reliability of the candidate layouts are quantified and the reliability calculations are carried out only for the layouts included in the cost-entropy

Pareto optimal set. Tanyimboh and Sheahan (2002) have shown that only a small fraction of the candidate layouts would be cost-entropy Pareto optimal. Therefore, the reliability calculations would not affect the efficiency of the approach significantly.

Once all the reliability values of the cost-entropy Pareto optimal layouts have been calculated, a set of cost-reliability Pareto optimal layouts can be identified from the cost-entropy Pareto optimal set. These cost-reliability Pareto optimal layouts represent the trade off between the cost and reliability of the network in question and the optimum layout can therefore be selected by considering the budget constraint. To check the effectiveness of the approach, Tanyimboh and Sheahan (2002) also calculated the reliability values for the full set of the 65 layout candidates so that the true cost-reliability Pareto optimal layouts can be identified. Their results are presented in Appendix 9 (Figure A9.1).

The above approach is applied in the present study to the same network. However, the HDA method is used in analysing the proposed designs and in obtaining the corresponding reliability values instead of the DDA. The algorithm used for selecting the cost-entropy Pareto optimal layouts and the corresponding cost-reliability Pareto optimal is summarised below. Once all the designs based on the candidate layouts have been generated, the cost-entropy Pareto optimal set is obtain as follows. Let n be a counter, CEPO and CRPO the sets of cost-entropy and cost-reliability Pareto optimal designs, respectively.

1. Set n to 0.
2. Increase n by 1.
3. Rank the available designs according to their costs in ascending order.
4. Select the cheapest design from the list. If there is more than one design with the same smallest cost, choose the one with the highest entropy value. Add the design to the set CEPO. Note that designs that have been selected are no longer included in the subsequent selections.
5. Rank the designs (excluding the selected designs) according to their entropy values.
6. Select all the designs that have entropy value higher than that of the CEPO(n) and discard the rest of the designs from subsequent selections.
7. If there are no designs that have higher entropy than CEPO(n), exit. Otherwise, go to step 2.

After the reliability values for all the designs contained in CEPO have been obtained, the above algorithm can also be used in selecting the members of CRPO but with the reliability in place of the entropy. The HDA based results generated in the present study are presented in Figure 7.10. From the figure, it seems that the use of the HDA method in the current approach to layout optimization results in a better outcome compared to the previous DDA-based approach (Figure A9.1 – Tanyimboh and Sheahan, 2002). For the cost-entropy Pareto optimal, the set of layouts obtained by using the HDA method seems to be more concentrated towards the true cost-reliability Pareto front while the DDA set is not. Also, the superiority of HDA can be observed by looking at the two sets of cost-reliability Pareto optimal, which were obtained from the cost-entropy Pareto optimal layouts, from the two network analysis methods. The HDA analysis was able to identify 9 out of 18 true cost-reliability Pareto optimal layouts, i.e. 9 out of 10 cost-reliability Pareto optimal layouts in the HDA analysis are true Pareto optimal. While the DDA analysis only managed to identify 4 out of 13 layouts.

The above maximum entropy approach to layout optimization is easy to implement and seems quite robust, especially when the HDA method is used in the analysis. A wider application of the approach was then investigated by looking at different flow directions. Flow directions in water distribution networks are generally not predictable without a simulation of the 'existing' network. However, many design methods, including the one used in this study, require the flow directions in the network to be specified prior to the design formulation. These specified flow directions would affect the resulting design in terms of cost, performance and, of course, entropy value. The identification of optimum design under a wide range of layouts and flow directions is therefore highly desirable. To investigate this issue, all the designs generated in the flow directions study in Chapter 6 were used together with the 65 original designs mentioned above. It needs to be pointed out that the designs in the flow directions study in Chapter 6 were generated based on three different layouts. However, their flow directions are not based on the shortest path principle. Therefore, their inclusion would assess the robustness of the above layout optimization method. The DDA and HDA methods were used in this investigation and, for consistency, the results of the DDA analysis are shown in the appendix (Figure A9.2) while the HDA results are summarised in Figure 7.11.

The result of the analysis seems to confirm the above conclusion that the method is robust and the HDA method yields a better performance than DDA. Figure 7.11 shows that there are 20 designs that belong to the true cost-reliability Pareto optimal set as identified by the HDA method. The layout optimization approach was able to identify 10 out of the 20 designs and 7 of these had identified earlier before the inclusion of the designs from the flow directions study. It would therefore suggest that the Pareto optimal front provided by the HDA approach is more or less identical to the true cost-reliability Pareto front.

7.3.1 CONCLUSIONS

The maximum entropy approach to layout optimization described above seems to perform well and efficiently. The method required the reliability of the layouts to be quantified explicitly; therefore, the optimum layout can be selected with confidence. The selection of designs that belong to the cost-entropy Pareto optimal set prior to the calculation of the reliability values acts as a filter in the approach and contributes greatly to the efficiency of the method. Also, a better performance in terms of the accuracy of the method was observed when the analysis was carried out using HDA in comparison to the DDA.

7.4 SUMMARY AND GENERAL CONCLUSIONS

In this chapter, the entropy-constrained approach to the design optimization of water distribution networks has been shown to have the potential to generate designs with much higher levels of spatial hydraulic predictability than conventional minimum-cost designs. In particular, maximum-entropy designs appear to have the highest levels of spatial hydraulic predictability. Two simple networks taken from the literature have been used to demonstrate the study. The results of the HDA analysis would appear to support the preliminary DDA-based study by Tanyimboh (1993). Issues considered in the assessment of the spatial hydraulic predictability included the locations of the critical nodes and pipes, based on individual critical operating conditions and the more holistic hydraulic redundancy and reliability measures. The results also support the conclusion drawn in Chapter 6 that the pipe diameters in a water distribution network become more uniform and the overall network hydraulic performance improves as its design entropy value increases. Therefore, the analyses in this chapter support the recommendation by Tanyimboh (1993) and Tanyimboh and Templeman (1993c) that water distribution networks should be designed to carry maximum-entropy flows.

The fact that entropy can be used to influence the hydraulic properties of a water network may have profound design and reliability implications. For example, in the discussion of Quindry et al. (1981), Templeman (1982b) stressed that the use of optimization techniques in the design of water distribution networks tends to remove redundancy by optimizing out any capacity that is not required for the particular loading being considered. Therefore, to ensure network reliability, resilience and flexibility, Templeman (1982b) suggested that the design procedure should consider explicitly the ability of the network to serve fire fighting demands at several nodes simultaneously. Morgan and Goulter (1985) have pointed out that the occurrence of simultaneous fires at all or even some demand nodes is not considered probable. However, they mentioned that a wide range of fire fighting demand patterns still need to be considered in the design process, i.e. demand patterns related to a fire at each node. By designing the network to carry maximum entropy flows, the above problem can be simplified considerably. Maximum entropy designs have been shown to have high reliability. Also, if the location of the critical link(s) and critical node(s) in the network can be pre-determined with high degree of certainty as seen in this chapter, assessment can be concentrated on these critical areas, which may reduce the amount of calculations required in the design process and network assessment.

In addition, the layout optimization technique for maximum entropy designs demonstrates another advantage of designing water distribution networks to carry maximum entropy flows. The method compares favourably with other approaches (Morgan and Goulter, 1985; Loganathan et al., 1990; Awumah and Goulter, 1992; Afshar et al., 2005) in terms of accuracy and efficiency since the reliability of the layouts are quantified and the calculations are carried out on the selected layouts (i.e. cost-entropy Pareto optimal) only. As has been shown in the previous chapter, different designs that have the same entropy values are likely to have similar performance. Therefore, the use of entropy as a preliminary filter for the reliability seems justifiable. The present study also shows that the maximum entropy approach to layout optimization is robust and the use of the HDA method complements the technique quite considerably.

Table 7.1. Pipe diameters for all the designs of the two-loop network.

Link	Pipe diameters for network with entropy value indicated (m)					
	1.578	1.600	1.700	1.800	1.900	1.915
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1 - 2	0.157	0.165	0.203	0.224	0.263	0.261
1 - 3	0.401	0.401	0.390	0.384	0.365	0.367
2 - 4	0.100	0.100	0.165	0.191	0.238	0.235
3 - 5	0.338	0.337	0.337	0.329	0.281	0.294
4 - 6	0.100	0.100	0.100	0.151	0.250	0.234
3 - 4	0.237	0.237	0.213	0.215	0.247	0.234
5 - 6	0.263	0.262	0.262	0.249	0.152	0.185
Mean	0.228	0.229	0.239	0.249	0.257	0.259
σ_{n-1}	0.116	0.115	0.100	0.081	0.063	0.058
$\frac{\sigma_{n-1}}{mean}$	0.510	0.504	0.418	0.325	0.246	0.224
Cost (£10^6)	0.250	0.251	0.254	0.259	0.261	0.263

Taken from Tanyimboh (1993).

Table 7.2. Fire-fighting loads for the two-loop designs.

Node	Fire Fighting Loads (m^3/s)									
	In place of the normal demands					In addition to the normal demands				
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 1	Case 2	Case 3	Case 4	Case 5
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	-0.506	-0.501	-0.459	-0.442	-0.478	-0.506	-0.501	-0.459	-0.442	-0.478
2	0.250	0.028	0.028	0.028	0.028	0.250	0.028	0.028	0.028	0.028
3	0.033	0.250	0.033	0.033	0.033	0.033	0.250	0.033	0.033	0.033
4	0.075	0.075	0.250	0.075	0.075	0.075	0.075	0.250	0.075	0.075
5	0.092	0.092	0.092	0.250	0.092	0.092	0.092	0.092	0.250	0.092
6	0.056	0.056	0.056	0.056	0.250	0.056	0.056	0.056	0.056	0.250

Column 1 to 6 taken from Tanyimboh (1993).

Table 7.3. Critical links for single-link failures for the two-loop network based on the HDA analysis.

	Assessment criteria			
	Total flow supplied (system DSR)		Total dissipated energy per unit flow	
Network entropy (1)	Critical link (2)	Next critical link (3)	Critical link (4)	Next critical link (5)
1.578	1 - 3	3 - 5	1 - 3	1 - 2
1.600	1 - 3	3 - 5	1 - 3	1 - 2
1.700	1 - 3	3 - 5	1 - 3	1 - 2
1.800	1 - 3	3 - 5	1 - 3	1 - 2
1.900	1 - 3	3 - 5	1 - 3	1 - 2
1.915	1 - 3	3 - 5	1 - 3	1 - 2

Table 7.4. Critical nodes for single-link failures for the two-loop network based on the HDA analysis.

	Assessment criteria			
	Nodal DSR		Actual shortfall at individual node	
Network entropy (1)	Critical node (2)	Next critical node (3)	Critical node (4)	Next critical node (5)
1.578	6	5	5	4
1.600	6	5	5	4
1.700	6	5	5	4
1.800	6	5	5	4
1.900	5	3	5	4
1.915	5	3	5	4

Note: All the critical and next critical nodes in the above table are associated with critical link 1-3 for all the designs.

Table 7.5. Critical nodes for nodal fire-fighting demands replacing design demands for the two-loop network based on the HDA analysis. Assessment is done at individual node.

Network entropy (1)	Assessment criteria: Nodal DSR — Critical node (2)	Nodal DSR — Next critical node (3)	Actual shortfall at individual node — Critical node (4)	Actual shortfall at individual node — Next critical node (5)
1.578	2	4	2	4
1.600	2	4	2	4
1.700	6	5	6	5
1.800	6	5	6	5
1.900	6	4	6	4
1.915	6	4	6	4

Table 7.6. Critical nodes for nodal fire-fighting demands replacing design demands for the two-loop network based on the HDA analysis. Assessment is done on the entire network.

Network entropy (1)	Assessment criteria: System DSR — Critical node (2)	System DSR — Next critical node (3)	Total flow supplied — Critical node (4)	Total flow supplied — Next critical node (5)	Total dissipated energy per unit flow — Critical node (6)	Total dissipated energy per unit flow — Next critical node (7)
1.578	2	6	2	6	6	4
1.600	2	6	2	6	6	4
1.700	6	2	6	4	6	2
1.800	6	2	6	4	2	6
1.900	6	5	6	5	6	2
1.915	6	4	6	5	6	2

Table 7.7. Critical nodes for superimposed nodal fire-fighting demands for the two-loop network based on the HDA analysis.

	Assessment criteria							
	Nodal DSR		Actual shortfall at individual node		Total flow supplied (system DSR)		Total dissipated energy per unit flow	
Network entropy (1)	Critical node (2)	Next critical node (3)	Critical node (4)	Next critical node (5)	Critical node (6)	Next critical node (7)	Critical node (8)	Next critical node (9)
1.578	2	4	2	4	2	6	5	6
1.600	2	4	2	4	2	6	5	6
1.700	6	5	6	5	6	4	5	6
1.800	6	5	6	5	6	4	5	4
1.900	6	4	6	4	6	5	5	4
1.915	6	4	6	4	6	5	4	5

Table 7.8. Pipe diameters for all the designs of the four-loop network.

Link (1)	Pipe diameters for network with entropy value indicated (m) 2.170 (2)	2.500 (3)	2.750 (4)	2.775 (5)	2.800 (6)
1 - 2	0.201	0.309	0.293	0.294	0.294
2 - 3	0.156	0.161	0.175	0.183	0.201
1 - 4	0.349	0.273	0.293	0.294	0.294
2 - 5	0.100	0.267	0.239	0.234	0.221
3 - 6	0.100	0.100	0.124	0.137	0.164
4 - 7	0.156	0.155	0.175	0.183	0.201
5 - 8	0.286	0.152	0.227	0.222	0.207
6 - 9	0.100	0.272	0.216	0.216	0.216
4 - 5	0.317	0.222	0.239	0.234	0.221
5 - 6	0.151	0.286	0.227	0.222	0.207
7 - 8	0.100	0.100	0.124	0.137	0.164
8 - 9	0.271	0.100	0.216	0.216	0.216
Mean	0.191	0.200	0.212	0.214	0.217
σ_{n-1}	0.092	0.080	0.055	0.050	0.041
$\frac{\sigma_{n-1}}{mean}$	0.484	0.401	0.259	0.233	0.187
Cost ($£10^6$)	0.277	0.282	0.289	0.290	0.292

Taken from Tanyimboh (1993).

Table 7.9. Nodal fire-fighting loads replacing design demands for the four-loop designs.

Node (1)	Fire Fighting Loads (m^3/s)							
	Case 1 (2)	Case 2 (3)	Case 3 (4)	Case 4 (5)	Case 5 (6)	Case 6 (7)	Case 7 (8)	Case 8 (9)
1	-0.4373	-0.4373	-0.4373	-0.4373	-0.4373	-0.4373	-0.4373	-0.3956
2	0.2500	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208
3	0.0208	0.2500	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208
4	0.0208	0.0208	0.2500	0.0208	0.0208	0.0208	0.0208	0.0208
5	0.0208	0.0208	0.0208	0.2500	0.0208	0.0208	0.0208	0.0208
6	0.0208	0.0208	0.0208	0.0208	0.2500	0.0208	0.0208	0.0208
7	0.0208	0.0208	0.0208	0.0208	0.0208	0.2500	0.0208	0.0208
8	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208	0.2500	0.0208
9	0.0625	0.0625	0.0625	0.0625	0.0625	0.0625	0.0625	0.2500

Taken from Tanyimboh (1993).

Table 7.10. Superimposed nodal fire-fighting loads for the four-loop designs.

Node (1)	Fire Fighting Loads (m^3/s)							
	Case 1 (2)	Case 2 (3)	Case 3 (4)	Case 4 (5)	Case 5 (6)	Case 6 (7)	Case 7 (8)	Case 8 (9)
1	-0.4581	-0.4581	-0.4581	-0.4581	-0.4581	-0.4581	-0.4581	-0.4581
2	0.2708	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208
3	0.0208	0.2708	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208
4	0.0208	0.0208	0.2708	0.0208	0.0208	0.0208	0.0208	0.0208
5	0.0208	0.0208	0.0208	0.2708	0.0208	0.0208	0.0208	0.0208
6	0.0208	0.0208	0.0208	0.0208	0.2708	0.0208	0.0208	0.0208
7	0.0208	0.0208	0.0208	0.0208	0.0208	0.2708	0.0208	0.0208
8	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208	0.2708	0.0208
9	0.0625	0.0625	0.0625	0.0625	0.0625	0.0625	0.0625	0.3125

Table 7.11. Critical links for single-link failures for the four-loop network based on the HDA analysis.

Network entropy (1)	Assessment criteria			
	Total flow supplied (system DSR)		Total dissipated energy per unit flow	
	Critical link (2)	Next critical link (3)	Critical link (4)	Next critical link (5)
2.170	1 - 4	4 - 5	1 - 4	4 - 5
2.500	1 - 2	5 - 6	1 - 2	1 - 4
2.750	1 - 2, 1 - 4	2 - 5, 4 - 5	1 - 2, 1 - 4	2 - 5, 4 - 5
2.775	1 - 2, 1 - 4	2 - 5, 4 - 5	1 - 2, 1 - 4	2 - 5, 4 - 5
2.800	1 - 2, 1 - 4	2 - 3, 4 - 7	1 - 2, 1 - 4	2 - 3, 4 - 7

Table 7.12. Critical nodes for single-link failures for the four-loop network based on the HDA analysis.

Network entropy (1)	Assessment criteria			
	Nodal DSR		Actual shortfall at individual node	
	Critical node (2)	Next critical node (3)	Critical node (4)	Next critical node (5)
2.170	9	(1 - 4)7	9	(1 - 4)7
2.500	9	(1 - 2)3	9	(1 - 2)3
2.750	9	(1 - 2)3, (1 - 4)7	9	(1 - 2)3, (1 - 4)7
2.775	9	(1 - 2)3, (1 - 4)7	9	(1 - 2)3, (1 - 4)7
2.800	9	(1 - 2)3, (1 - 4)7	9	(1 - 2)3, (1 - 4)7

Note: In Columns (3) and (5) in the above table, the next critical nodes are as a result of the failure of the links indicated in the brackets.

Table 7.13. Critical nodes for nodal fire-fighting demands replacing design demands for the four-loop network based on the HDA analysis. Assessment is done at individual node.

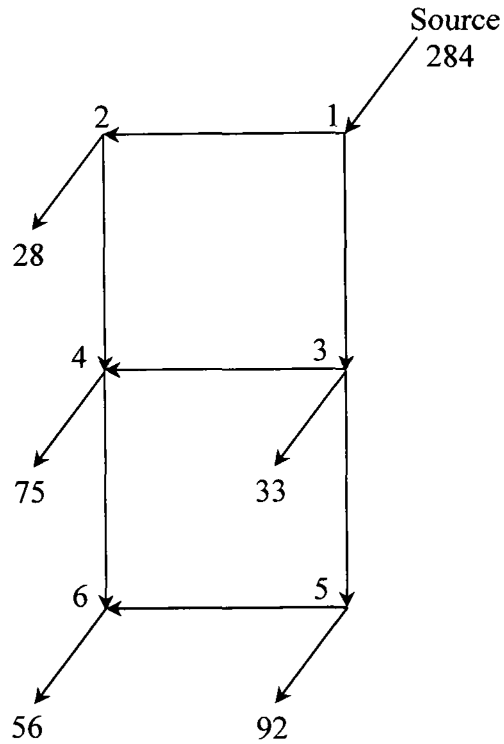
Network entropy (1)	Assessment criteria: Nodal DSR — Critical node (2)	Nodal DSR — Next critical node (3)	Actual shortfall at individual node — Critical node (4)	Actual shortfall at individual node — Next critical node (5)
2.170	3	6	3	9
2.500	7	8	7	9
2.750	3, 7	9	3, 7	9
2.775	3, 7	9	3, 7	9
2.800	6, 8	9	6, 8	9

Table 7.14. Critical nodes for nodal fire-fighting demands replacing design demands for the four-loop network based on the HDA analysis. Assessment is done on the entire network.

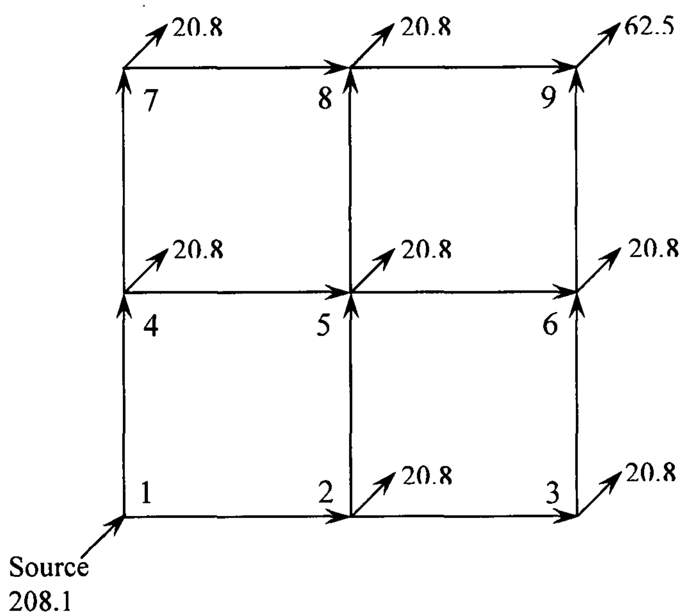
Network entropy (1)	Assessment criteria: System DSR — Critical node (6)	System DSR — Next critical node (7)	Total flow supplied — Critical node (6)	Total flow supplied — Next critical node (7)	Total dissipated energy per unit flow — Critical node (8)	Total dissipated energy per unit flow — Next critical node (9)
2.170	3	6	3	6	8	5
2.500	7	8	7	3	6	5
2.750	3, 7	6, 8	9	3, 7	6, 8	5
2.775	3, 7	6, 8	9	3, 7	6, 8	5
2.800	6, 8	3, 7	9	6, 8	6, 8	5

Table 7.15. Critical nodes for superimposed nodal fire-fighting demands for the four-loop network based on the HDA analysis.

Network entropy (1)	Assessment criteria							
	Nodal DSR		Actual shortfall at individual node		Total flow supplied (system DSR)		Total dissipated energy per unit flow	
	Critical node (2)	Next critical node (3)	Critical node (4)	Next critical node (5)	Critical node (6)	Next critical node (7)	Critical node (8)	Next critical node (9)
2.170	3	6	3	9	3	6	8	5
2.500	7	8	7	9	7	8	6	5
2.750	3, 7	9	3, 7	9	3, 7	9	6, 8	5
2.775	9	6, 8	9	6, 8	9	3, 7	6, 8	5
2.800	9	6, 8	9	6, 8	9	6, 8	6, 8	5



(a)



(b)

Figure 7.1. Networks under normal operating condition with all flows in litre per second.

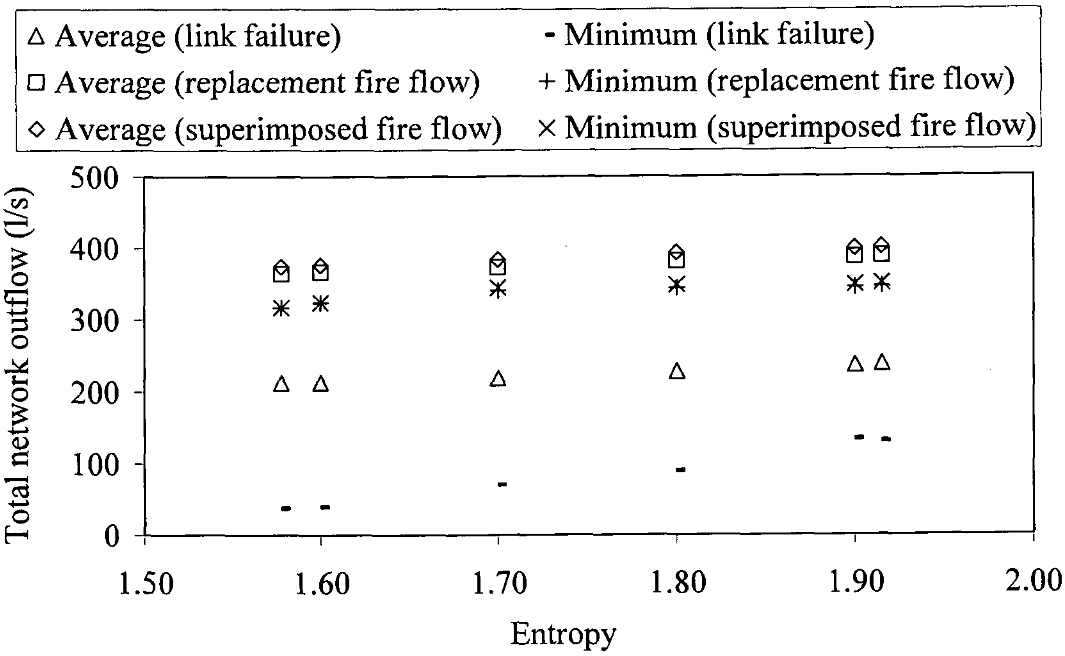


Figure 7.2. Flow supplied vs. entropy for the two-loop network under critical operating conditions.

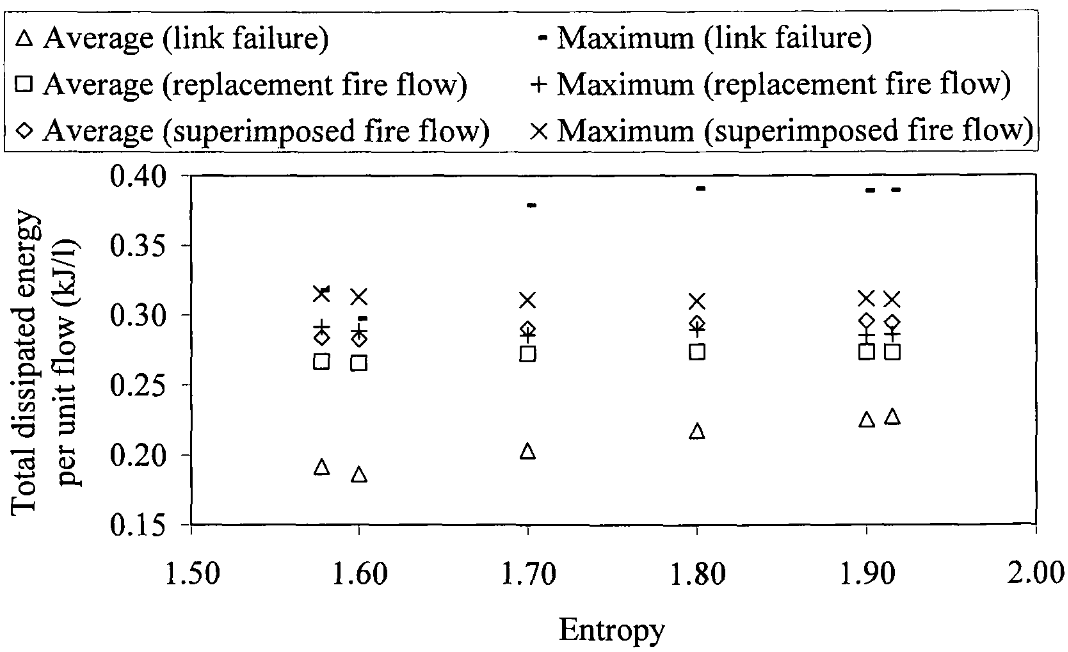


Figure 7.3. Total dissipated energy per unit flow vs. entropy for the two-loop network under critical operating conditions.

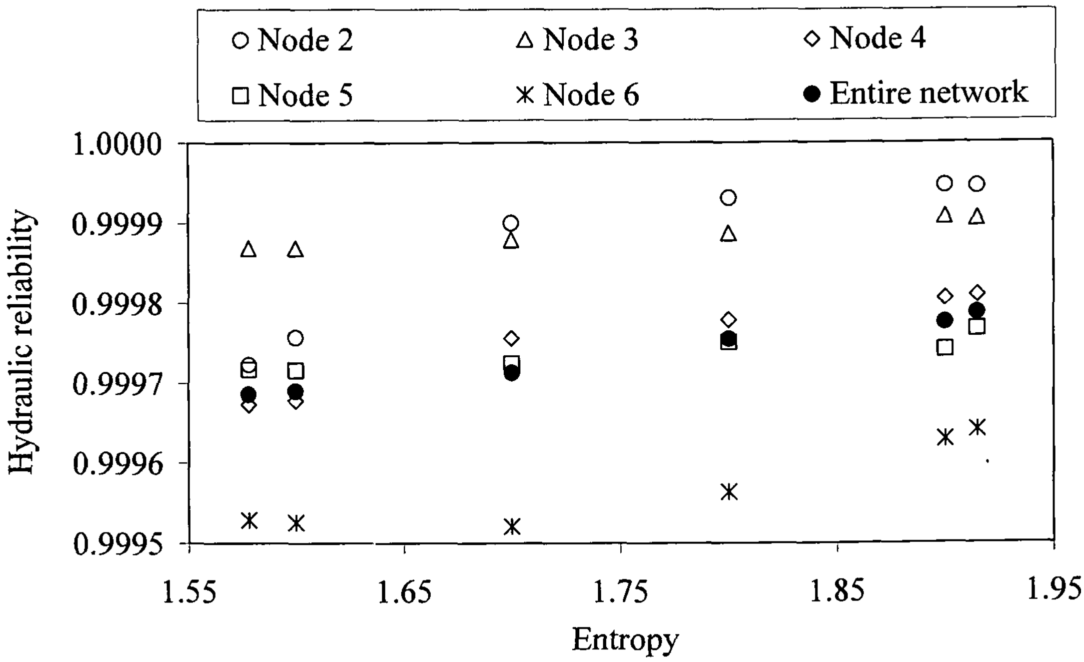


Figure 7.4. Hydraulic reliability vs. entropy for the two-loop network.

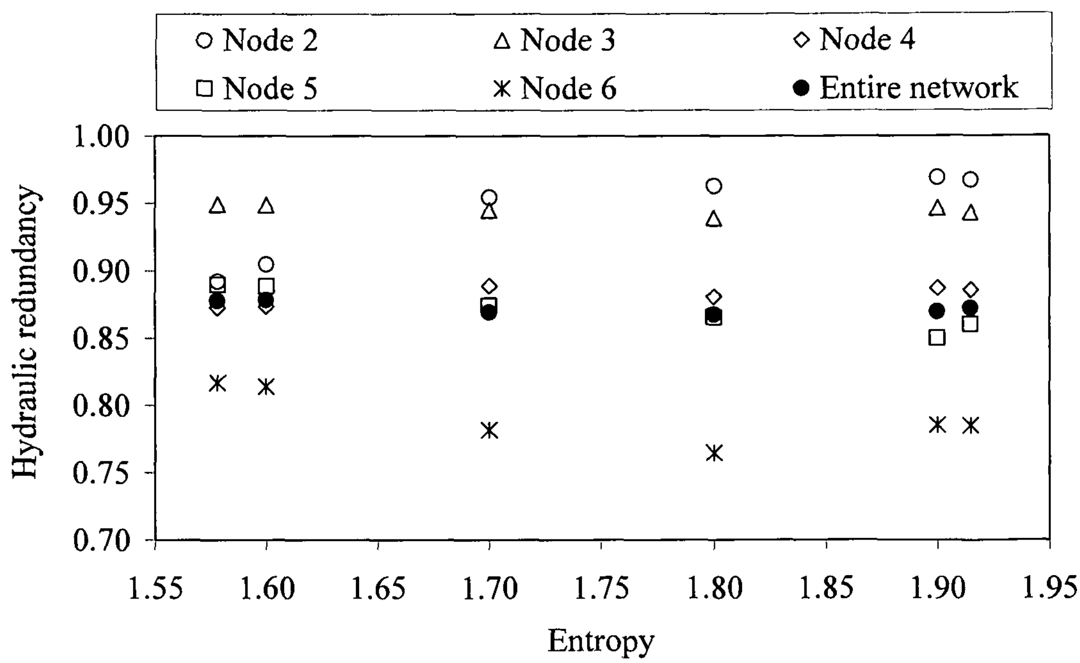


Figure 7.5. Hydraulic redundancy vs. entropy for the two-loop network.

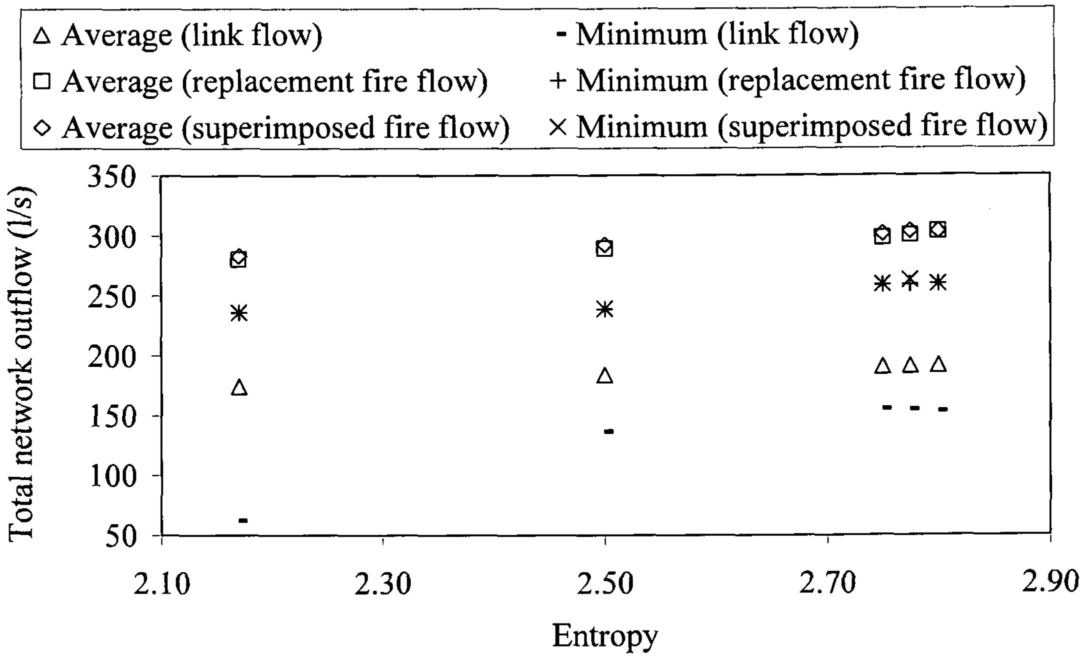


Figure 7.6. Flow supplied vs. entropy for the four-loop network under critical operating conditions.

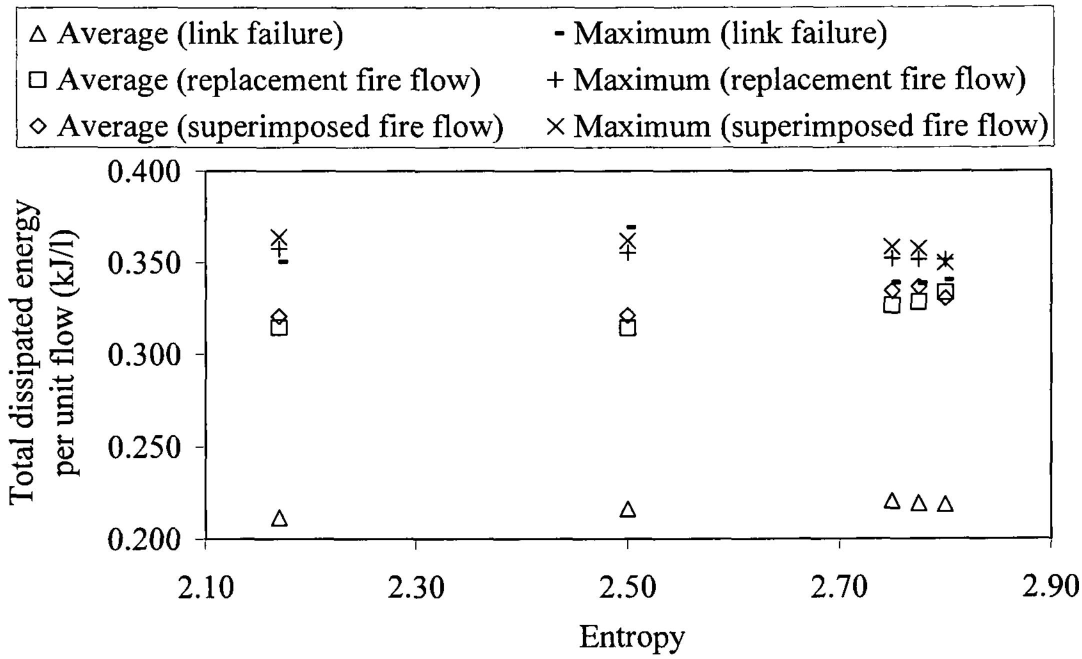


Figure 7.7. Total dissipated energy per unit flow vs. entropy for the four-loop network under critical operating conditions.

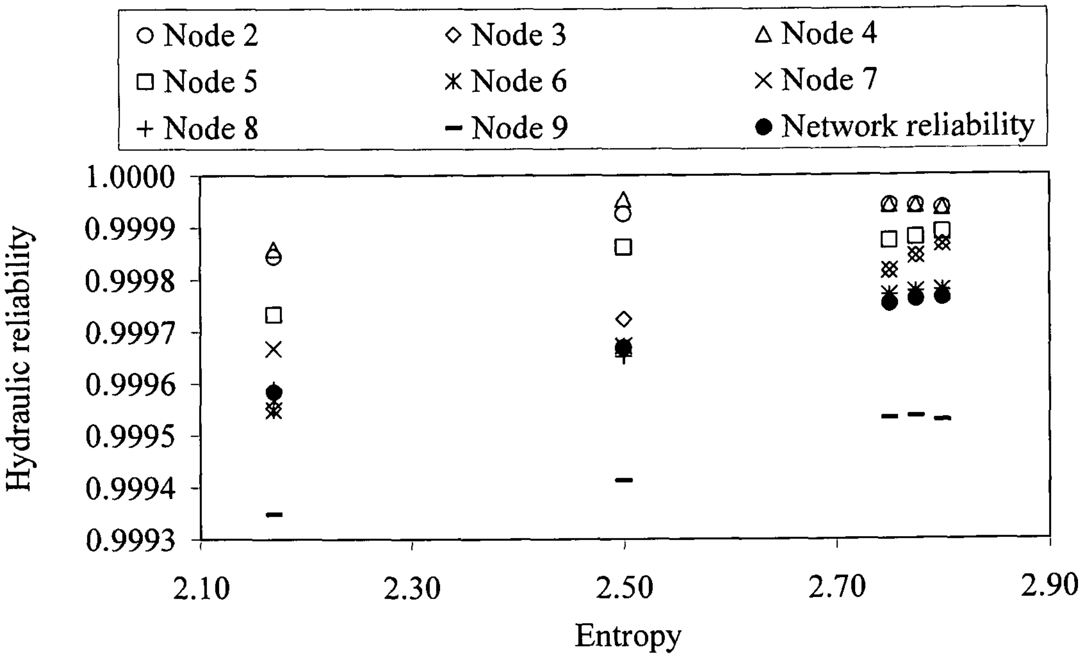


Figure 7.8. Hydraulic reliability vs. entropy for the four-loop network.

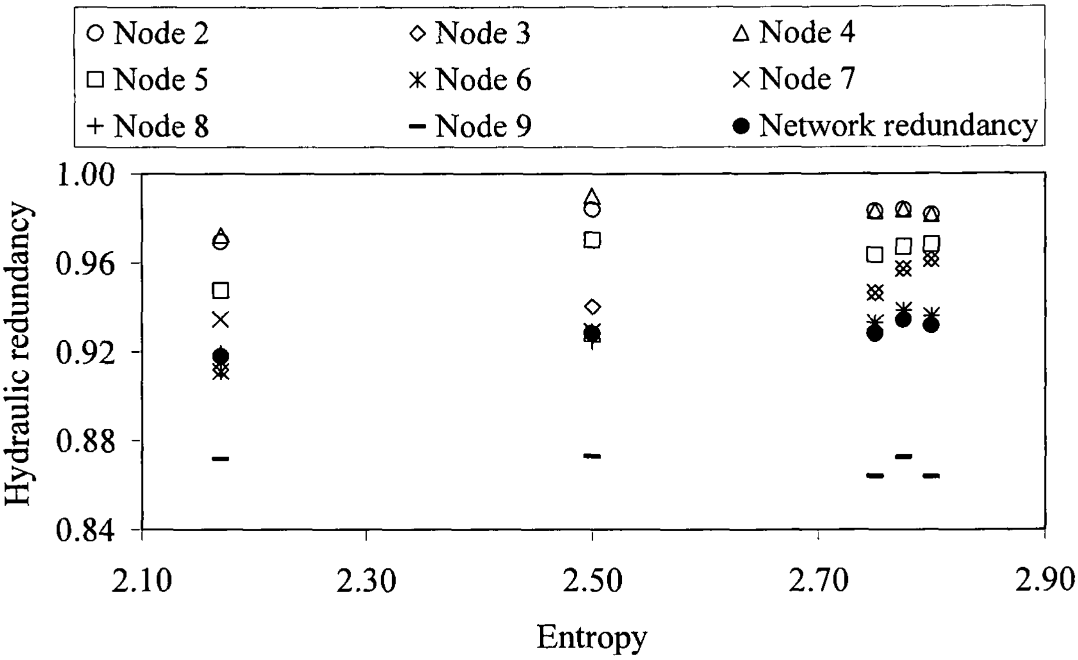


Figure 7.9. Hydraulic redundancy vs. entropy for the four-loop network.

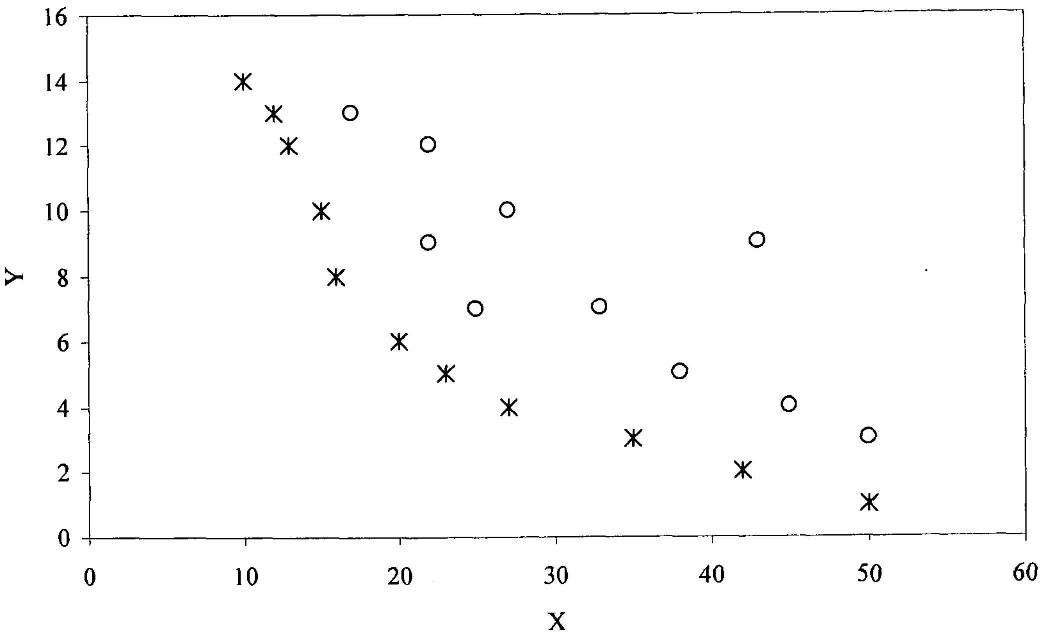


Figure 7.10. Example of Pareto Optimal solution.

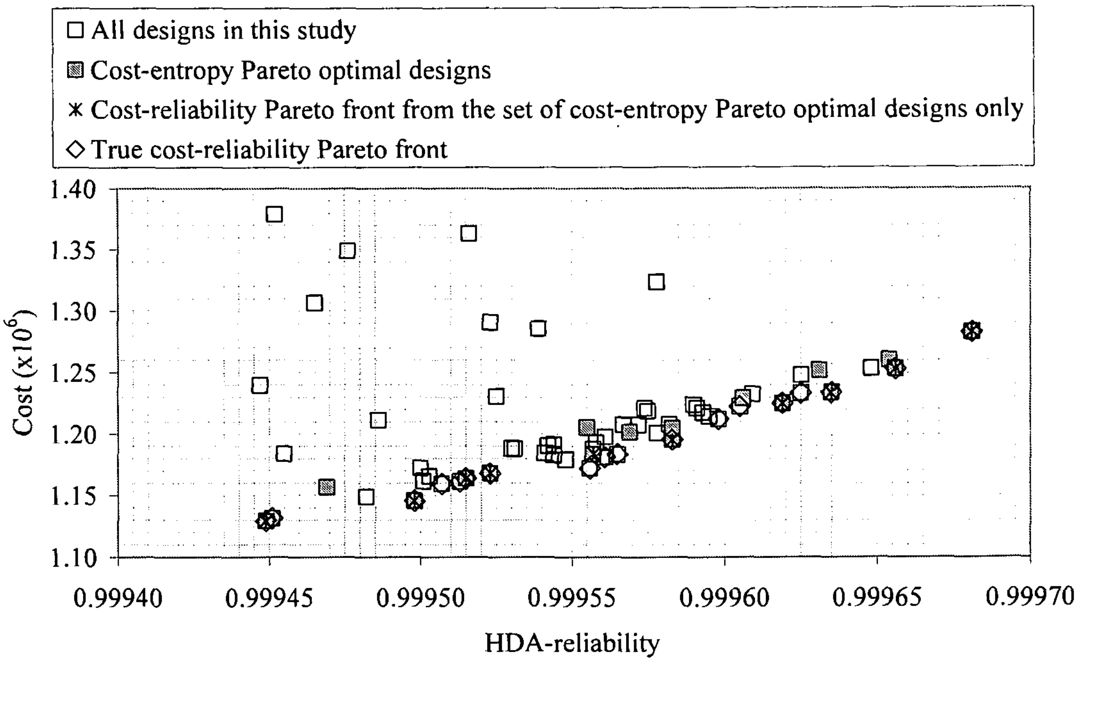


Figure 7.11. Plots of cost against HDA-reliability showing the cost-entropy and the cost-reliability Pareto optimal layouts.

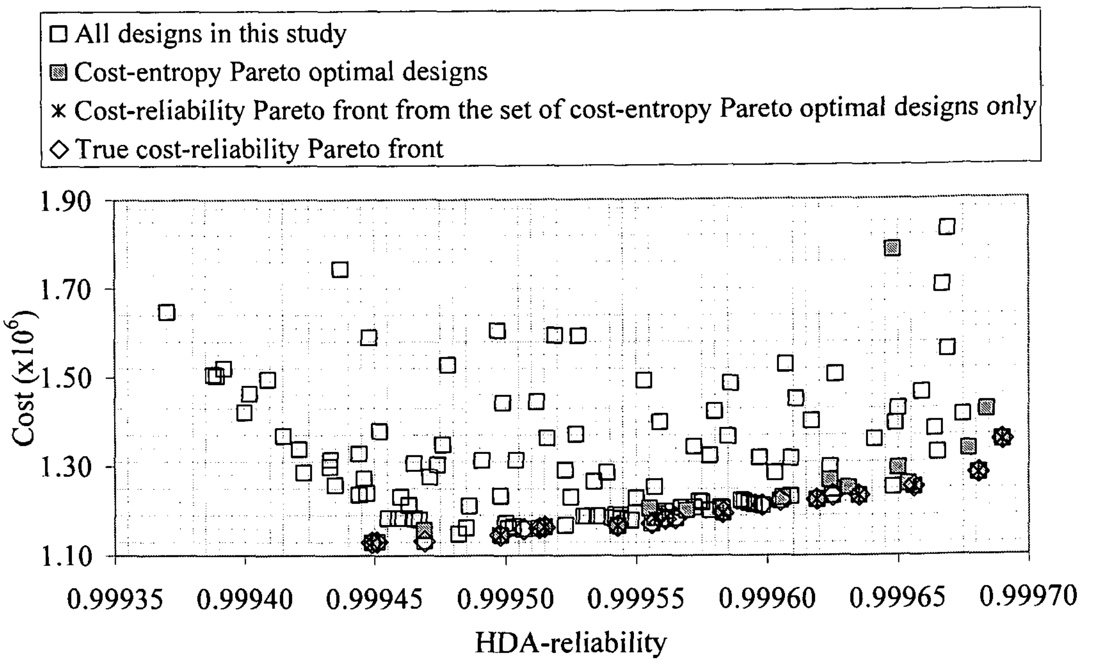


Figure 7.12. Plots of cost against HDA-reliability showing the cost-entropy and the cost-reliability Pareto optimal layouts considering alternative flow directions.

CHAPTER 8 THE MAXIMUM ENTROPY APPROACH TO THE OPTIMUM DESIGN OF WATER DISTRIBUTION NETWORKS USING GENETIC ALGORITHMS AND DISCRETE PIPE DIAMETERS

8.1 INTRODUCTION

The problem of designing optimum water distribution networks has been reviewed in Chapter 5. The design optimization is complex especially when the aspect of reliability is considered in the design process. Several design optimization methods proposed by several researchers, which attempt to simplify the process of the optimization, have also been discussed briefly in the same chapter. The strengths and weaknesses of theses methods were highlighted. The following study looks into the application of the entropy constrained approach to the design optimization of water distribution networks using a stochastic search method. As in the previous chapters, the investigation is limited to gravity networks only. However, each link in the resulting designs in the present study would have a single discrete, instead of continuous, pipe diameter size. Genetic Algorithms were chosen among other stochastic optimization methods available. These algorithms have been used quite extensively by several researchers in the past few decades in obtaining new designs of water distribution networks and extension to the existing networks, e.g. Simpson et al. (1994), Halhal et al. (1997), Savic and Walters (1997), Vairavamoorthy and Ali (2000). The results of their investigations show that GA can perform satisfactorily despite the complexity of the problem due to the discrete nature of the decision variables and the non-linearity of the constraint functions. Also, Walski et al. (2003) have stated that the superiority of the Genetic Algorithms in comparison to other design optimization methods for water distribution networks has been acknowledged. The process of adding the maximum entropy constraint to the design optimization procedures has been explained in Chapter 5. The entropy-constrained design problem for cases with continuous pipe diameters has also been summarised in that chapter. For the case in which discrete pipe diameters are used, the optimization problem can be formulated as follows.

Problem 8

$$\underset{\forall D_{ij} \in D_D}{Minimize} \quad C = \sum_{ij \in IJ} \gamma_d L_{ij}^d \tag{5.18}$$

Subject to:

$$h_{ij} = \alpha L_{ij} \left(\frac{q_{ij}}{C_{ij}} \right)^{1.852} \frac{1}{D_{ij}^{4.87}} \ , \quad \forall ij \in IJ \tag{4.4}$$

$$\sum_{j \in NU_n} q_{jn} - \sum_{k \in ND_n} q_{nk} = q_n \ , \quad n = 1, \ldots, NN - 1 \tag{4.1}$$

$$\sum_{ij \in IJ_l} h_{ij} = 0 \ , \quad l = 1, \ldots, NLP \tag{4.7}$$

$$\sum_{ij \in IJ_p} h_{ij} = h_p \ , \quad p = 1, \ldots NP \tag{4.9}$$

$$\frac{\pi v_{\min}}{4} \le \frac{q_{ij}}{D_{ij}^2} \le \frac{\pi v_{\max}}{4} \ , \quad \forall ij \in IJ \tag{5.21}$$

$$H_s - H_{\max,n} \le \sum_{ij \in IJ_n} h_{ij} \le H_s - H_{\min,n} \ , \quad \forall n \tag{5.23}$$

$$D_{ij} \in D_D \tag{5.25}$$

$$S \ge S_{\min} \tag{5.34}$$

All symbols in the above problem have been previously defined and those definitions are unchanged. No links may be eliminated from the network by the above procedures since the link sizes must be selected from the set of the available discrete pipe diameters, D_D. Unlike Problem 5, the requirement for the non-negativity of flow in the above problem formulation is no longer required. The flow directions in the network are allowed to change in the resulting designs and the value of $S_{\min}$ is calculated for each set of flow directions using Yassin-Kassab (1998) algorithm as explained later. Next, the GA-based design optimization program for water distribution networks is described. A sensitivity study on the GA parameters is then carried out to determine their appropriate values for solving the above problem. The strengths and weaknesses of the GA-based procedures in the design optimization of water distribution networks with the maximum entropy constraint are examined.

8.2 OVERVIEW OF THE GA-BASED DESIGN OPTIMIZATION PROGRAM

The steps involved in the design optimization of water distribution networks using Genetic Algorithms are summarised in Figure 8.1. A design optimization program was developed and written in FORTRAN95. It consists of several routines including the GA and the maximum entropy routines. The standard GA routine used in the program was developed by Anderson (1995). Binary alphabet was used to represent the decision variables and modification to the binary strings was carried out using three basic GA operators, i.e. tournament selection, cross-over and mutation. Knuth subtractive method for pseudo random numbers generation (WH Press et al., 1992, pp. 273-274) was used in the program. Other methods for random number generation may be used if considered more appropriate, for example Simpson et al. (1994) used a linear congruential generator (Gentle, 2003). The maximum entropy routine, on the other hand, was developed by Yassin-Kassab (1998). It is capable of calculating the maximum entropy flows for water distribution networks with single and multiple sources. At each generation, hydraulic analyses were carried out as many times as the number of population using Newton-Raphson method based on the head system of equations. More details of the program components are described next following the brief description of the scope of the program.

8.2.1 SCOPE OF THE PROGRAM

The strong points of the program and the search procedures can be summarised as follows:

1. Given that a high entropy value of a water distribution network corresponds to a high level of hydraulic reliability of the network, as has been shown in the earlier chapters, the GA-based design program is capable of producing inexpensive yet

highly reliable designs since it incorporates the maximum entropy constraint in the optimization procedures. The advantage of quantifying the reliability level of the network by means of the entropy measure is attained without adding unnecessary burden in the computations since the network reliability value is not actually calculated at this stage.

2. The program can handle discrete values of design variables, i.e. discrete pipe diameters. This feature is favoured by practicing engineers since the design model resembles the real network more closely and there is no need for rounding off the diameters to find the final solution.
3. The program requires a set of initial flow directions as input. However, depending on the sizes of the chosen diameters in the trial solutions, the outcomes of the Newton-Raphson analysis may give a new set of different flow directions for each of the trial solutions. Considering that networks whose link-flows travel through the shortest paths from the source to each demand node should intuitively have low network costs, the GA optimization procedures therefore would guide the search towards optimum flow directions, i.e. the shortest paths, while the maximum entropy constraint would ensure that high level of network reliability is maintained regardless of the flow directions in the network. This facility is particularly useful when designing multi-looped networks that have more than one source of supply.
4. The entropy of a water distribution network is a function of the pipe flow rates. Its value changes with the change of the flow directions in the network. Therefore, in the optimization procedure (Problem 8), the value of the maximum entropy constraint may vary and must be calculated for every new set of flow directions.
5. The potential for finding the optimum set of flow directions is an advantage that stochastic optimization approaches have over deterministic methods. In deterministic search procedures a starting point has to be specified, which includes specifying the flow directions in the network. The procedures then search towards a minimum cost solution based on this set of flow directions. However, the problem of specifying optimum flow directions on which the design is based is not a straight forward task, especially for large networks with many loops and/or when multiple flow patterns are considered.
6. The search process in the GA optimization procedures is carried out over a population of trial solutions and spreading throughout the solution space. This

technique increases the chances of finding the global optimum solution significantly. Also, the ability of the GA search method in finding the optimum set of flow directions contributes greatly towards the identification of the optimum solution as will be seen later in this chapter.

Apart from the above advantages, there are also some limitations to the program and the GA method. These shortcomings are as follows:

1. The program can only design gravity networks. Pumps, valves and service reservoirs are not considered in the present study. The optimization procedures also take into account the cost and the level of network hydraulic reliability only. Other aspects of design, such as water quality and the sensitive customers in the network, which should be considered in a wider context of design, are not included in this study.
2. Only one flow pattern can be handled by the program. However, multiple patterns can be incorporated by specifying a different set of constraints for every flow pattern considered.
3. In Genetic Algorithms, constraint functions cannot be handled directly in the process of updating the trial solutions. In the design optimization of a water distribution network, once all the new trial solutions have been generated, a hydraulic analysis is required to check if any constraints were violated by each trial solution and to obtain the corresponding amount of the violations. This contributes to the high computing time of the GA search procedures, in particular, and the stochastic optimization methods in general. However, this slight limitation is compensated by the apparent ability of the stochastic methods in finding the global optimum solution.

8.2.2 INPUT AND OUTPUT OF THE PROGRAM

It has been mentioned earlier that the GA-based design optimization program requires an initial set of flow directions to start with. Other required data are the network topography, available source heads, minimum head requirement and candidate diameters and their corresponding costs per unit length. In the present study, the binary codes that represent the candidate diameters are also required in the input file. However, this process can be easily automated by including a routine to generate binary numbers within the program.

The output of the program consists of an optimum and several near optimum solutions. Provided that all the constraints are satisfied, the near optimum solutions may have higher costs due to different sets of flow directions, which correspond to different maximum entropy values. If any of these values were found to be higher than the entropy of the optimum solution, reliability calculations can be carried out on these candidate designs to obtain a more definite comparison of their reliability levels. Finally, if the high level of reliability is verified, the design may be selected as the final design. For every candidate solution, the data provided in the output file are the nodal heads, nodal outflows, link flows and their directions, pipe diameters and their total cost, constraint violations, maximum entropy and the actual entropy value of the resulting design.

8.2.3 PENALTY FUNCTIONS

The way in which violations of constraints are accounted for in Genetic Algorithms is by means of cost penalty functions outside the search procedures. A cost penalty is assigned to each of the violated constraints in proportion to the magnitude of the violations. The role of the hydraulic solver in the GA-based optimization of water distribution networks in identifying these violations is therefore crucial.

In the present study, different cost penalty functions are used for different sets of constraints depending upon their accuracy requirements. However, all the functions take the following form

$$Cost\ Penalty = Unit\ of\ constraint\ violated \times PC \tag{8.1}$$

in which PC is the penalty cost multiplier. For loop and path constraints, i.e. Equations (4.7) and (4.9), the value of PC was set to 10,000 unit cost/metre of head. Nevertheless, the iteration in the Newton-Raphson analysis would converge towards a solution that satisfies these two constraints. In the case where convergence in the Newton-Raphson iteration was not achieved, a very high value of the cost penalty (equal to 10^9 unit cost in the present study) was assigned to the trial solution to ensure its exclusion from the subsequent tournament selection. For the nodal head and the entropy constraints, a higher constraint violation corresponds to higher cost penalty. This was intended to increase the chances of the trial solutions with higher constraint violations being excluded from the selection process, hence increasing the rate of convergence. Three values of *PC* for the nodal head constraints were specified, i.e. 100,000 unit cost/metre of head when the head at a demand node is violated by 0.01 metres or more, 90,000 unit cost/metre of head when there is a violation of equal to or greater than 0.005 metres and 80,000 unit cost/metre of head when the violation is less than 0.005 metres. Meanwhile, the values of *PC* for the entropy constraint were 10^6 unit cost when 1.0^{-5} or higher units of entropy were violated, 500,000 unit cost when the violation is between 1.0^{-6} to 1.0^{-5} and 10,000 unit cost if the entropy violation is less than 1.0^{-6}. The values of *PC* for the entropy constraint are higher than the *PC* values for the other constraints since slight difference in the entropy value of a network may lead to a difference in the reliability level of the design.

The above procedure is a crude way of guiding the search towards solutions that satisfy all the constraints. Heuristic analysis shows that the value of *PC* affects the accuracy of the resulting designs in terms of satisfying the constraint functions and the above method appears to be quite effective. For simplicity, the flow velocity constraints were excluded from the present study. Another set of cost penalty functions are required when these constraints are included in the optimization procedures. Once all the cost penalties have been obtained, the total cost of each trial solution is the sum of the design cost of the network and its cost penalties. The solution fitness is then obtained as the inverse of the total cost, i.e.

$$Total\ Cost = Network\ Cost + \sum_{NC} Cost\ Penalty \tag{8.2}$$

in which *NC* is the number of constraints, and

$$Fitness = \frac{1}{Total\ Cost} \tag{8.3}$$

8.2.4 TERMINATION CRITERIA

The termination requirement in GA iteration is usually the completion of a certain number of generations (Simpson et al., 1994; Savic and Walters, 1997; Vairavamoorthy and Ali, 2000). In the present study, three termination criteria are used in the program (Anderson, 1995) including the specified maximum number of generations of 500. The second stopping criterion refers to the condition in which there are no trial solutions in the current generation better than the best solutions from the previous several generations. Several best and near best results are saved at the end of each generation. This collection is then updated at each successive generation. When several generations (15 generations in this study) have elapsed with no updates to the collection, the optimum solution is considered to have been reached and the search process is terminated. The third condition is when the standard deviation of the fitness of the population in the current generation is very small (i.e. less or equal to a pre-specified value; 10^{-15} was used in the present study to account for the double precision values used in the program). This condition is usually an indication that the optimum solution is either found or almost found since an improvement in the population fitness in the next generation would be very small or insignificant. A minimum number of generations, e.g. 100 in this study, is specified before the second and third stopping criteria are applied. This is to ensure a wide enough search space has been explored before the program may be terminated.

8.3 PARAMETER SENSITIVITY AND SEARCH EFFICIENCY ANALYSES

Three layouts were selected from the previous chapters to be used in the parameter sensitivity and search efficiency studies. These layouts are the 2-loop 6-node and 4-loop 9-node layouts in Figure 7.1, and the 3-loop 12-node layout in Figure 6.5b. It may be noted that maximum entropy designs based on these layouts have been generated in the previous chapters using continuous pipe diameters. In this chapter, the maximum entropy designs based on these layouts were re-generated using Genetic Algorithms and discrete pipe diameters. The same set of constraints as that for the previously generated designs were used and the resulting discrete pipe diameter designs were compared to the previous designs. Since all of the previous designs were generated using continuous pipe diameters, for comparison to be valid, the discrete candidate diameters used in the GA-based design optimization must include the set of diameters obtained previously for the maximum entropy designs based on continuous pipe diameters. The GA-based program was then run with the aim to obtain the same combination of pipe diameters. For example, 8 discrete candidate diameters were specified for designing the 2-loop 6-node layout using Genetic Algorithms. 6 of these candidate diameters were obtained from the previous maximum entropy design with continuous pipe diameters (Chapter 7, Table 7.1) and the other two were chosen from the continuous diameters of the design with the smallest entropy value. The flow directions of the designs from which the diameters were selected were the same as those in Figure 7.1a, which represent the shortest paths from the source to each demand node. The maximum entropy value for this set of flow directions is 1.915 and the corresponding cost of the design is £0.263 million, which is based on the γ and e values of 900 and 2.4, respectively (Equation 5.17). Since the layout comprises of 7 links, with 8 candidate diameters for each link, there are 8^7 = 2,097,152 possible pipe combinations for feasible and infeasible solutions. The GA-based design procedures were able to obtain the same combinations of pipe diameters and flow directions as the previous maximum entropy design in a relatively short period of time, i.e. 38 seconds of CPU time at the most on a Pentium 4, 1.4GHz PC with 256 MB RAM.

The above 2-loop 6-node layout is small and simple. Also, it may be argued that the 6 candidate diameters, which were selected from the previous maximum entropy design with continuous pipe diameters, helped the search in the GA procedures towards finding the same set of pipe diameters. To disprove this notion, 16 candidate diameters were specified for designing the symmetrical 4-loop 9-node layout. This set of candidate diameters comprises the diameters from the previous maximum as well as minimum entropy designs whose flow directions are as indicated in Figure 7.1b. The GA search procedures may therefore produce a design with the smallest or the highest entropy value based on the specified sets of candidate diameters and flow directions or a completely different design should the flow directions in the resulting design change. The maximum entropy value for the shortest-path flow directions is 2.799 and this time the search space has $16^{12} = 2.815 \times 10^{14}$ of possible pipe combinations. The GA search method was capable of finding the maximum entropy solution with the same sets of diameters and flow directions as the previously generated design without any difficulty.

The third layout comprises of 14 links. 16 candidate diameters were specified for this layout in the same way as that for the previous two layouts, i.e. the candidate diameters include the diameters of the maximum entropy design from the continuous diameter case whose flow directions represent the shortest paths (Figure A4.2 option 1). This layout was a little more complicated to design since the search space is larger, i.e. 7.206×10^{16} pipe diameter combinations. However, the GA-based design program managed to obtain the maximum entropy design with the shortest-path flow directions. The maximum entropy value for the shortest-path flow directions is 2.671. However, the actual entropy value of the optimum solution found using GA was only 2.669. This discrepancy is small and may be attributed to the set of discrete candidate diameters used as well as the cost penalty function for the entropy constraint as follows. Firstly, the maximum entropy value of the network may not be achievable exactly with discrete pipe sizes. Secondly, although the specified candidate diameters are discrete, some of their values are almost the same, e.g. 388mm and 390mm, 194mm and 197mm, 230mm and 234mm. Therefore, the selection of certain combination of pipe diameters results in a lower total cost (design cost + cost penalties) at the expense of the accuracy of the entropy constraint. Comparison between the GA-based discrete-diameter designs and the previously generated continuous-diameter designs for all three layouts can be observed in Table 8.1.

For the sensitivity study, three basic GA parameters, i.e. population size and cross-over and mutation probabilities, were examined and the effects of different combinations of population size and cross-over probability were analysed. For each combination, the GA-based design program was run five times with different starting points for the random number generator, which would give five different sequences of pseudo random numbers and five different initial populations. This was done to increase the chances of finding the global optimum solution. Population size and cross-over probability are the dominant parameters that drive the search in GA towards the optimum solution. The probability of mutation, on the other hand, is very low. It contributes towards ensuring that the search process is not trapped in a local optimum point. Many researchers have proposed several different combinations of the GA parameters (see e.g. Mitchell, 1999, pp. 175-177, for the list of combinations). However, no conclusive results on what is the best combination of the parameters have been achieved to date. The typical value of the population size lies between 50-1000 populations (Mitchell, 1999). It is logical to expect the required number of population to increase with the increase of the solution search space in order to maintain the breadth of the exploration, hence maintaining the chances of finding the global optimum solution. The recommended range for the cross-over probability is between 0.6-1.0 (Goldberg, 1989) and the probability of mutation is usually between 0.01-0.1 (Savic and Walters, 1997). In the present study, the mutation probability is function of the number of bits in the chromosome string, i.e. the number of bits that represent one trial solution. For example, there are 7 links and 8 candidate diameters in the 2-loop 6-node layout design problem. Therefore one candidate diameter can be represented by a three-bit binary number and one trial solution comprises of $3 \times 7 = 21$ binary bits. The probability of mutation used in the GA operation for this layout was $1/21 \approx 0.05$.

In the above analysis, the flow directions of the resulting GA-based maximum entropy designs may be different than the previously generated designs from which the candidate diameters were selected. This may cause the comparison between those designs to be invalid. However, for each set of flow directions, there is only one combination of link flows, hence pipe diameters, which corresponds to the maximum entropy value. Therefore, although the resulting designs may not be comparable to the previous designs, the addition of the maximum entropy constraint in the GA-based design optimization of water distribution networks provides an unambiguous condition of optimality and a definite target to achieve regardless of the resulting flow directions in the network. Also, for the first two layouts, the GA-based design program was able to reach convergence in less than 100 generations for some combinations of population size and cross-over probability. In order to see exactly at which generation the convergence was achieved, the minimum required generation was eliminated and all three stopping criteria were applied from the very first generation. The results of the analysis are shown in Tables 8.2 – 8.4.

From the tables, it is clear that as the size of the problem increases, the size of population required in the GA search procedures also increases. The increase in the CPU time as the number of population becomes larger seems to be compensated by the reduction in the number of generations required by the Genetic Algorithms to reach the optimum solution. Also, the efficiency of the GA-based design optimization of water distribution networks does not seem to be affected by the introduction of the maximum entropy constraint. However, it is rather difficult to obtain a definite comparison since different researchers used different combinations of population size, cross-over and mutation probabilities as well as different cost penalty functions. For example, the GA-based design procedures of Savic and Walters (1997), which does not include the entropy constraint, required more than 130 generations to obtain the optimum solution to a problem with 1.48 x 10^9 possible pipe combinations based on a population size of 50. Their cost penalty equations are non-linear functions of the generation number, whose values gradually increase with the increase of the generation number. Meanwhile in the present study, less than 130 generations were required to obtain the optimum solution to a problem with 2.815 x 10^{14} possible pipe combinations based on a population size of 100. It should also be noted that the efficiency of the hydraulic solver used in the search procedures contributes towards the overall efficiency of the method.

8.4 FLOW DIRECTION ANALYSIS

The potential of the GA-based design procedures in finding the optimum design of water distribution networks with optimum flow directions is now examined. The designs based on the three layouts used in the sensitivity analysis were re-generated using various initial flow directions. The aim is to obtain the same combinations of pipe diameters as the previous GA-based designs and the flow directions that represent the shortest paths from the source to each demand node. Seven sets of initial flow directions shown in Figure 8.2 were used for designing the 2-loop 6-node layout and the GA-based design program was run using a combination of population size of 500 individuals and a cross-over probability of 0.9. These values were obtained from the sensitivity analysis. For the 4-loop 9-node layout, nine sets of initial flow directions shown in Figure 8.3 were specified and the values of the population size and cross-over probability used were 500 and 1.0, respectively. Finally, 23 different sets of initial flow directions as shown in Figure A4.2 (options 2-24) were specified for designing the 3-loop 12-node layout. A population size of 1000 and a cross-over probability of 0.7 were used for this layout.

For each set of the initial flow directions, the GA-based design optimization program was run five times with five different sequences of pseudo random numbers. The typical plots of the cost of the best solution in each generation for the three layouts studied in this chapter are shown in Figures 8.4, 8.5 and 8.6, which corresponds to the 2-loop 6-node, 4-loop 9-node and 3-loop 12-node layout, respectively. In each run, the optimum solutions shown in the above figures, which correspond to the shortest-path flow directions and the set of pipe diameters shown in Table 8.1 were always identified. This result is very encouraging considering the huge numbers of possible pipe combinations and flow directions especially for the larger 3-loop 12-node layout. The ability of the GA-based design optimization in identifying the optimum set of flow directions may have a significant contribution towards its ability in obtaining the global optimum solution to the design problem of water distribution networks. Although in the above analysis only networks with one source supply were considered, the facility should, at least in theory, be transferable to multiple source networks in which identification of the optimum flow directions is a lot more complicated especially when multiple loops are present.

8.5 SUMMARY AND CONCLUSIONS

Problem formulation for the maximum-entropy design optimization of water distribution networks with discrete pipe diameters has been presented in this chapter. This problem has never been attempted before due to the limitation of the available optimization methods in dealing with discrete non-linear problems. The application of stochastic optimization method seems to hold the key to this issue. Genetic

Algorithms were used in this chapter and a design optimization program utilising the algorithms and the maximum entropy constraint was developed. Brief descriptions of the strengths and weaknesses of the program as well as the search procedures in obtaining the optimum solution have also been presented in this chapter.
Several hypothetical layouts whose maximum entropy designs have been identified were used to validate the design program and to analyse sensitivity of the GA parameters as well as the efficiency of the search procedures. Only the dominant parameters, i.e. the population size and cross-over probability, were analysed in this chapter and it was found that these parameters are quite sensitive to the type and size of the problem and they affect the efficiency of the optimization procedures. On top of that, the form of the cost penalty function has a noticeable influence on the accuracy of the outcome, hence on the effectiveness of the method. This is a slight drawback to the Genetic Algorithms since the appropriate combinations of the parameters are problem specific and they are not known in advance. More studies are also required to determine the suitable cost penalty function.

On the other hand, the above slight limitation of the GA search method is overcome by its apparent ability to obtain the global optimum point. In the optimization of the hypothetical layouts used in the sensitivity analysis, the maximum entropy designs were always identified. The GA search procedures were also capable of finding the optimum flow directions in the network regardless of the initial flow directions specified. This facility seems to contribute towards the ability of the method in identifying the global optimum solution.

Table 8.1. Comparison of diameters (mm) between the GA-based designs and the previously generated designs with continuous diameters

2-loop 6-node layout			4-loop 9-node layout			3-loop 12-node layout		
GA Candidate Diam.	Previous Design*	GA-based Design	GA Candidate Diam.	Previous Design*	GA-based Design	GA Candidate Diam.	Previous Design**	GA-based Design
100	261	261	100	294	294	100	253	253
165	367	367	151	201	201	109	388	390
185	235	234	156	294	294	158	234	230
234	294	294	164	221	221	176	197	197
235	234	234	201	164	164	187	230	230
261	234	234	207	201	201	194	310	310
294	185	185	216	207	207	197	194	194
367			221	216	216	211	158	158
			271	221	221	221	221	221
			286	207	207	230	176	176
			294	164	164	234	194	194
			317	216	216	253	187	194
			349			310	211	211
			610			388	109	100
			615			390		
			620			400		
Costs (£10^6)								
	0.263	0.263		0.292	0.292		1.183	1.182

* Tanyimboh (1993)

** Tanyimboh and Sheahan (2002)

Table 8.2. The number of optimum solutions achieved in every 5 runs of the GA-based design program.

Population size	Cross-over probability	Number of convergence per 5 runs		
		2-loop 6-node (Layout 1)	4-loop 9-node (Layout 2)	3-loop 12-node (Layout 3)
100	0.7	4	0	0
	0.9	2	1	0
	1.0	2	1	0
500	0.7	5	5	2
	0.9	5	5	3
	1.0	5	5	2
1000	0.7	5	5	5
	0.9	5	5	4
	1.0	5	5	4

Note: For every layout, the GA-based design optimization program was run 5 times with different starting points. The numbers in the above table indicate how many times the optimum solution was identified from different starting points.

Table 8.3. The average number of generations required by the GA-based design program to reach the optimum solution (out of 5 runs).

Population size	Cross-over probability	Average number of generations		
		2-loop 6-node (Layout 1)	4-loop 9-node (Layout 2)	3-loop 12-node (Layout 3)
100	0.7	67	n/a	n/a
	0.9	63	110	n/a
	1.0	71	183	n/a
500	0.7	40	83	158
	0.9	38	88	160
	1.0	39	78	144
1000	0.7	34	72	130
	0.9	34	69	135
	1.0	31	73	120

Table 8.4. The average CPU time required by the GA-based design program to reach the optimum solution (out of 5 runs).

Population size	Cross-over probability	Average CPU time (seconds)		
		2-loop 6-node (Layout 1)	4-loop 9-node (Layout 2)	3-loop 12-node (Layout 3)
100	0.7	5.2	n/a	n/a
	0.9	5.5	87.9	n/a
	1.0	6.2	102.0	n/a
500	0.7	19.9	326.0	896.5
	0.9	19.7	348.8	487.3
	1.0	20.3	335.8	460.0
1000	0.7	37.1	639.6	860.2
	0.9	38.3	626.6	908.0
	1.0	36.6	683.8	845.0

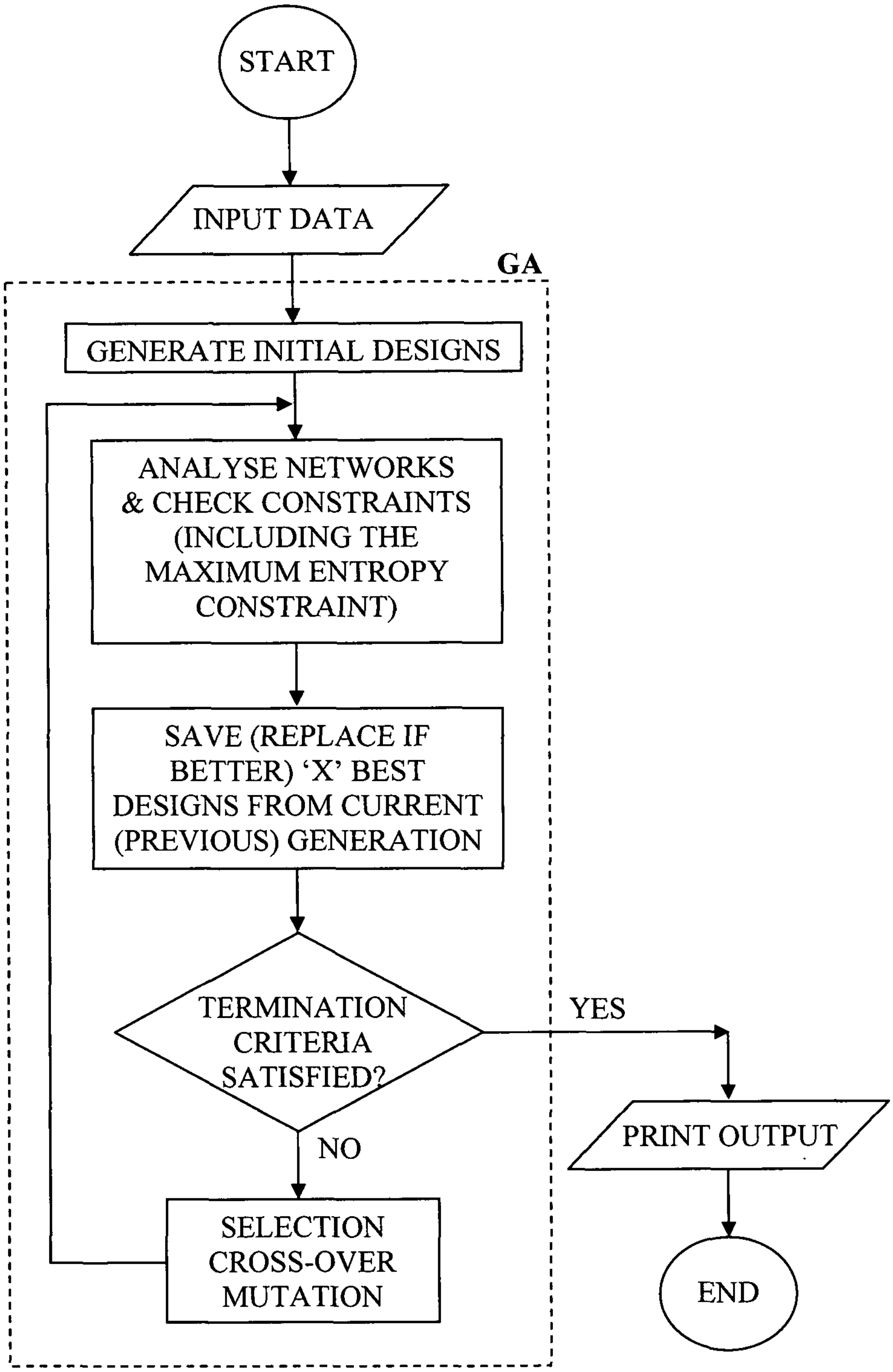


Figure 8.1. Schematic of the GA-based design optimization procedures.

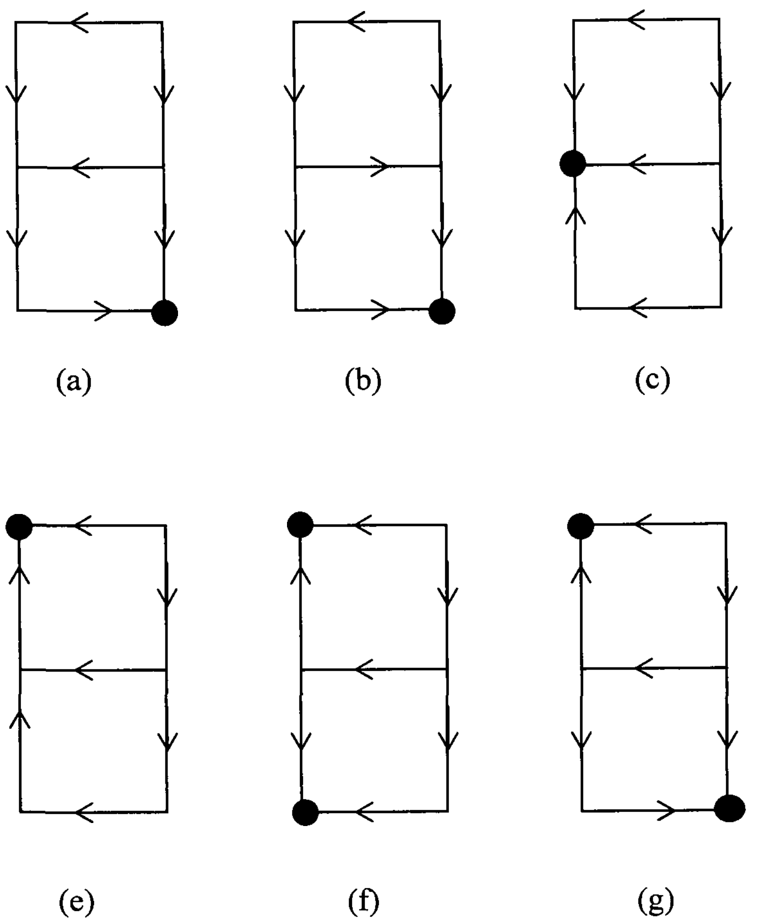


Figure 8.2. Different sets of initial flow directions for the GA-based design of the 2-loop layout.

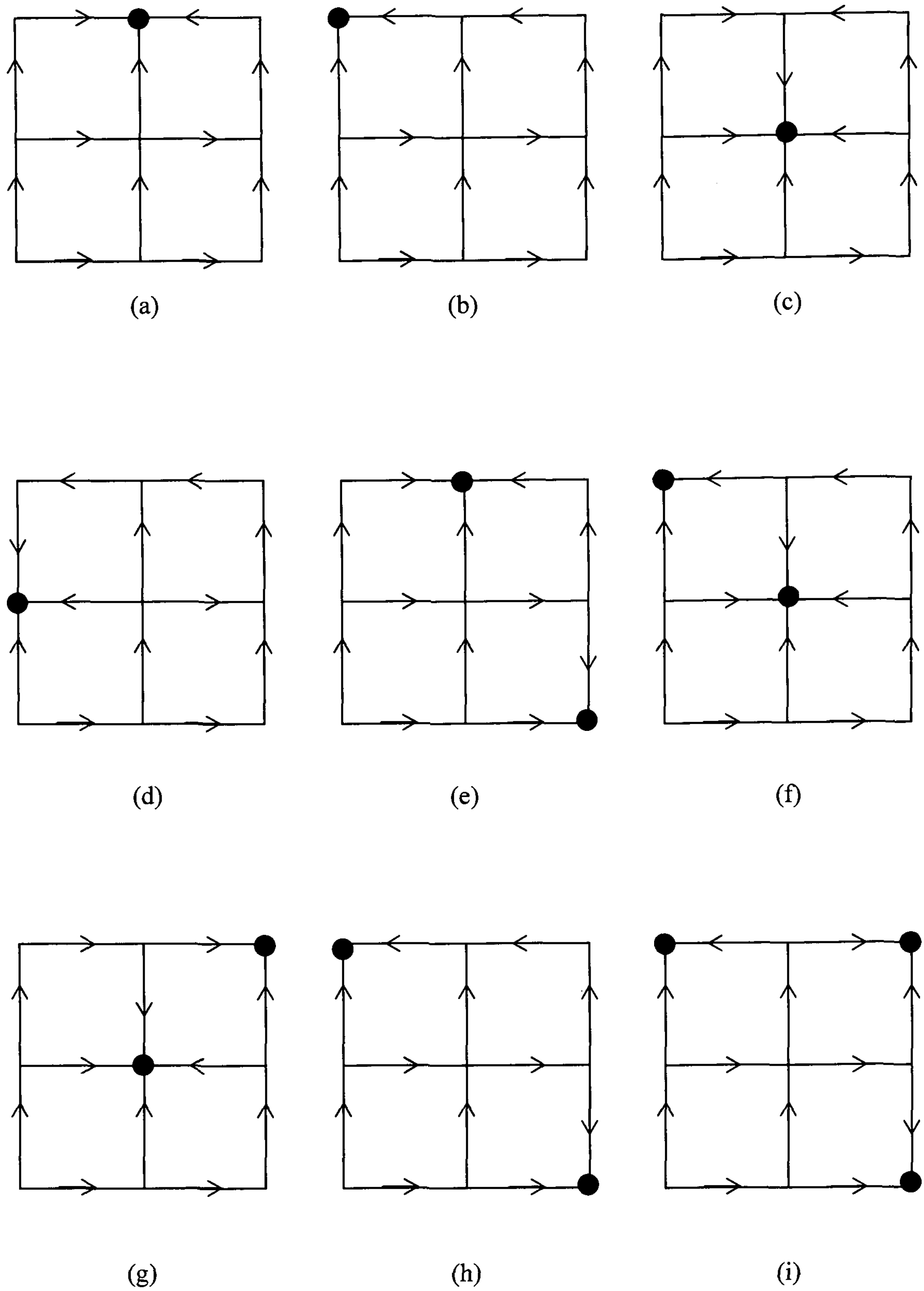


Figure 8.3. Different sets of initial flow directions for the GA-based design of the 4-loop layout.

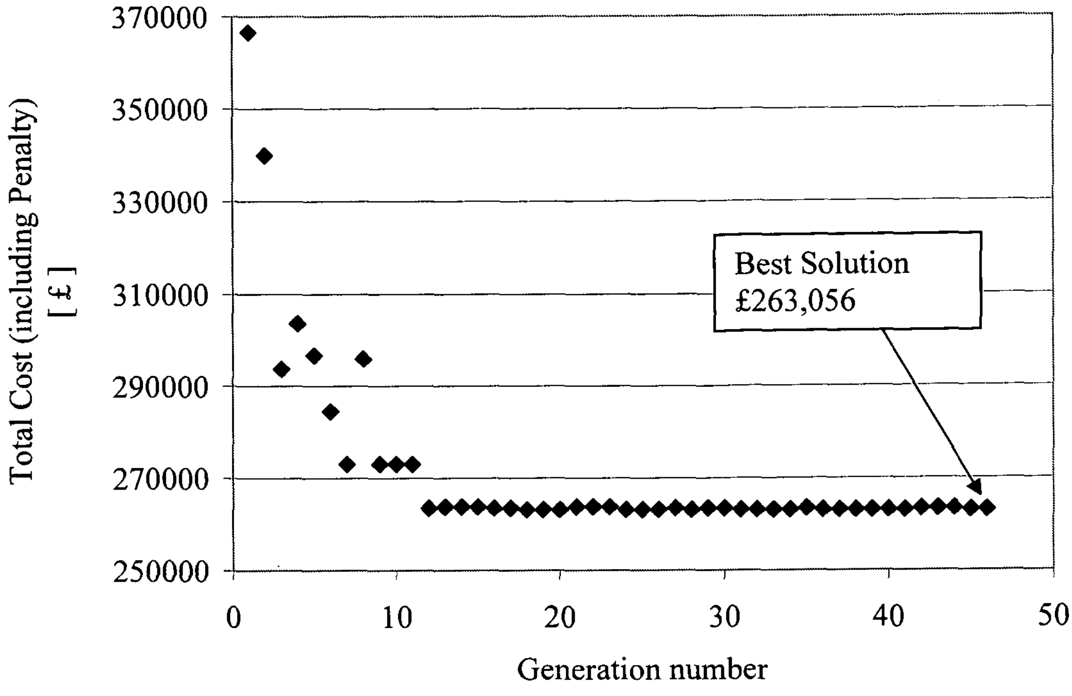


Figure 8.4. Typical progress of the optimization of the 2-loop 6-node layout.

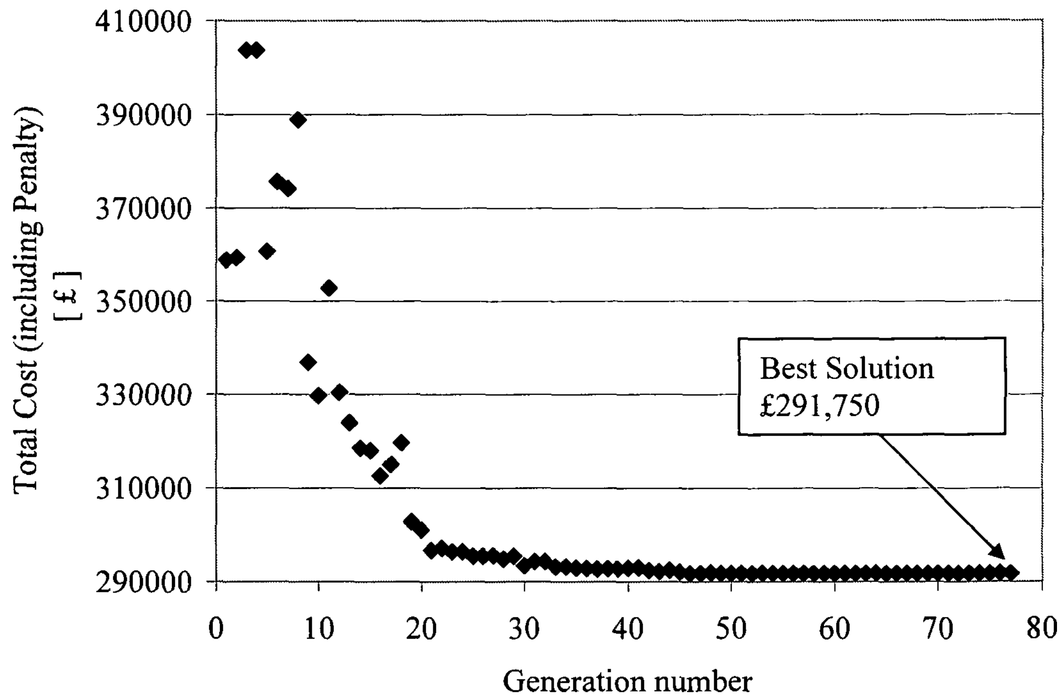


Figure 8.5. Typical progress of the optimization of the 4-loop 9-node layout.

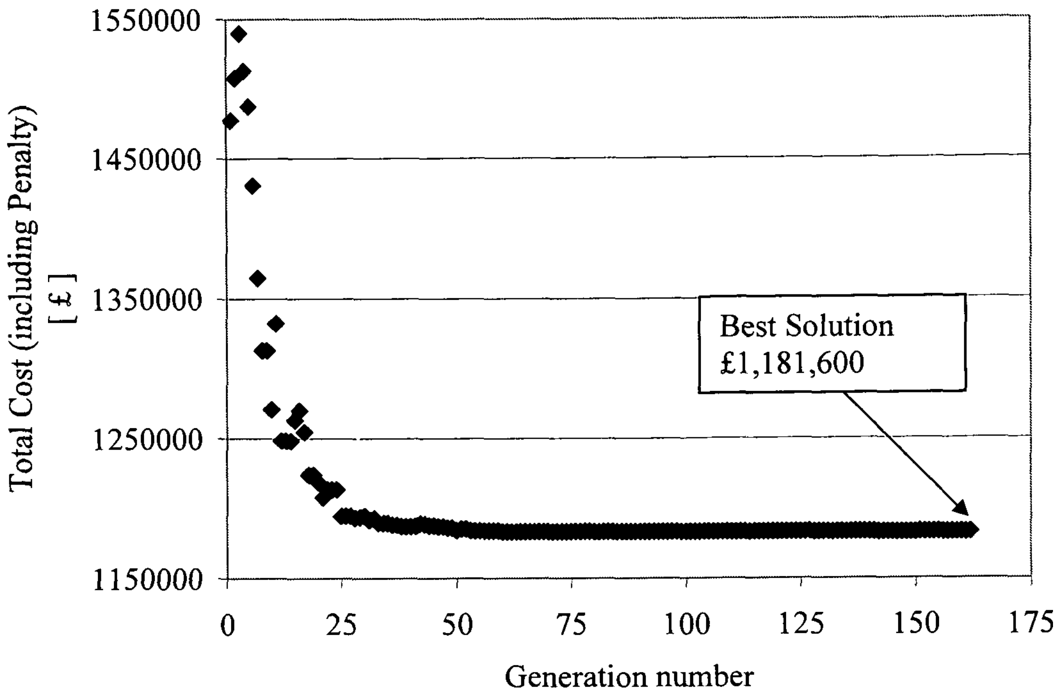


Figure 8.6. Typical progress of the optimization of the 3-loop 12-node layout.

CHAPTER 9 SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

9.1 INTRODUCTION

The problem of obtaining optimum design of reliable water distribution networks is extremely difficult to solve. Research into this issue has been going on for several decades and many different methods have been proposed by many researchers, each with its own advantages and disadvantages. Most of the traditional design methods minimize the capital cost of the network without putting much emphasis on its performance. Templeman (1982) has pointed out that these methods would inevitably try to eliminate all the redundancy in the network and result in a tree-type branch network with each demand node supplied by a single path.

To increase flexibility and reliability, some methods suggested that the existence of loops in the network should be maintained by means of a minimum pipe diameter constraint so that no links may be eliminated in the optimization process. However, the flexibility of the resulting designs of these methods is questionable. Although no links are eliminated from the network, the optimization process would try to reduce the diameter of the loop completing links into the minimum size possible. The usefulness of these links as alternative flow paths in the event of failure of larger pipes is very low. Wagner et al. (1988a) have pointed out that the performance of a water distribution network cannot be determined based on connectivity alone.

Other methods tried to incorporate some kind of reliability measure directly into the optimization procedures. Two main problems are immediately apparent from this approach. Firstly, there is no comprehensive and widely applicable method for quantifying the reliability of water distribution networks. Many researchers have proposed many different methods, each having its own strengths and weaknesses. Also, the definition of the reliability is different from one researcher to another. Secondly, in most of the reliability models, the process of obtaining the reliability value involves simulation of failure modes in the network to gauge the network performance under such situations. This process can be tedious and impractical for real water networks since there will be so many failure conditions and their combinations that have to be considered. Also, there are uncertainties associated with the failure rates and the duration of failure in the network. Furthermore, factors like the age of the network and soil conditions, which will affect the failure rates in the network, need to be accounted for in a wider context of reliability. All of these issues add to the difficulties in obtaining an accurate value of reliability even for a relatively small network.

Some researchers therefore suggested the use of entropy as a surrogate measure of the reliability of water distribution networks. This approach was pioneered by Awumah et al. (1989) who proposed the first entropy function for water networks based on Shannon's informational entropy (Shannon, 1948). However, their entropy function violates two of the fundamental properties of Shannon's entropy. Tanyimboh and Templeman (1993a) were the first who proposed a rigorous entropy function for water distribution networks and their entropy has been used in the present study. The advantages of entropy as a surrogate measure for water network reliability stems from its ease of computation and incorporation into the design optimization procedures. Also, it has been shown that the reliability of a distribution network improves as the entropy value of the network increases (Tanyimboh and Templeman, 1993b). This idea forms the basis for the investigations carried out in this research.

9.2 SUMMARY AND CONCLUSIONS OF THE PRESENT RESEARCH

Entropy is a measure of uncertainty. Therefore, in water distribution systems it is mainly applicable to looped water networks in which there is uncertainty associated with the distribution of flows in the network. In this research, the investigations were carried out on simple gravity-driven looped networks. Head dependent analysis method was used in the present study and the results were compared to the previous demand driven analysis results. Previous studies have shown that the relationship between entropy and reliability of water distribution networks appears to be quite strong. However, there are many aspects of the design that can affect the reliability level as well as the entropy value of the network. A comprehensive study into the possible influence of some of these aspects on the relationship between entropy and reliability has been carried out in the present research. The issues investigated were the possible influence of layouts, flow directions, cost functions and modelling errors.

It has been demonstrated that the influence of the above issues on the entropy-reliability relationship is negligible. The results of the investigations also support the previous conviction that higher value of entropy corresponds to a better network performance. An important aspect of the entropy-based design is the increase of the average size and uniformity of the pipe diameters in the network as the value of the entropy increases; with the maximum entropy designs having the largest average size and the most uniform pipe diameters. These characteristics lead to the high reliability level of the maximum entropy design since larger diameters generally generate lower head losses and uniform pipe diameters produces equal flow paths. Failure of a pipe in the network would not cause too much stress to the network since alternative paths are able to carry the redirected flows. It has also been shown that the increase in the network cost due to the increase in the entropy value seems to be quite modest in comparison to the improved performance level. Another important issue is the fact that the level of similarity in the performance of designs with equal maximum entropy values is much higher compared to other designs with different entropy. This quality is important for any surrogate performance measure in order to be able to distinguish between networks with different levels of performance.

The above analyses provide more evidence of the effectiveness of the maximum entropy approach in generating optimum designs of water distribution networks. It leads to a further study of the characteristics of the maximum entropy designs. Hydraulic predictability in terms of the locations of the critical links and nodes in the network was investigated. It was found that in maximum entropy designs these locations were more intuitively predictable than in other designs. For example, in the maximum entropy designs, the critical nodes could be found either at the terminal nodes or at the nodes with high demands while the critical links were often located near the source or connected to a node with large demand. The selection of these intuitive locations as the critical points in the network can assist designers in the analysis of the network performance during the design process by reducing the amount of calculations required since analysis can be concentrated on the critical areas. By contrast, the locations of the critical links and nodes in other designs were not easily predictable and could be anywhere in the network.

The maximum entropy approach to the layout optimization of water distribution networks has also been demonstrated in this thesis. The method is efficient since the entropy was used as a preliminary filter to identify the optimum layouts and the reliability values were calculated for these candidate layouts only. The method is also robust in the sense that it is capable of identifying the layouts that belong to the true Pareto optimal set, which is the set that represents the trade-off between cost and reliability of the network.

In all the above studies, the use of the HDA method in the analysis generally produced better results than the DDA method. This can be seen from the correlation between entropy and reliability, which appeared stronger when the HDA method was used. The locations of the critical links and nodes in the maximum entropy designs were also more consistent with one's intuitive expectation when the analysis was carried out using the HDA. Finally, the set of true Pareto optimal layouts could be identified more accurately using the HDA method.

The present research has also studied the application of the maximum entropy approach to the optimum designs of water distribution networks with discrete pipe diameters. Genetic Algorithms were employed in the optimization procedures. The study showed that different combinations of the GA parameters influence the efficiency of the search procedures quite considerably. Sensitivity study was required to determine the appropriate combinations of these parameters for solving different design optimization problems. The form of the cost penalty function was also found to have an effect on the efficiency of the search method. Despite all the above limitations, the method was capable of identifying the maximum entropy designs with optimum set of flow directions. Also, the introduction of the maximum entropy constraint does not affect the efficiency of the GA method.

The general conclusions drawn from the present research are as follows:

1. Entropy is a suitable performance measure for water distribution networks. The maximum entropy designs are not only highly reliable but also hydro-spatially predictable. As such, designing a network to carry maximum entropy flows could help designers in the process of obtaining the optimum solution.
2. The maximum entropy approach to the layout optimization of water networks is robust and the use of the HDA method complements the approach considerably.
3. Genetic Algorithms are capable of producing designs with optimum set of flow directions. This faculty contributes towards the ability of GA in finding the global optimum solution to the design problem.
4. The introduction of the maximum entropy constraint does not seem to reduce the efficiency of the GA search method. On the contrary, it provides a definite objective for the optimization procedures, which may improve the efficiency of the method in finding the global optimum solution. On average, the optimum solution was found in 4 out of 5 runs in this study provided that the right combination of parameters has been identified.
5. GAs can produce minimum-cost maximum-entropy designs with discrete pipe diameters

9.3 SUGGESTIONS FOR FUTURE RESEARCH

The advantages of the maximum entropy approach to the design optimization of water distribution networks have been highlighted in the present study. The approach has also been successfully applied to hypothetical networks to obtain the optimum designs with discrete pipe diameters. However, it is by no means the absolute decision making tool in the design optimization of water networks. Many issues still need to be explored to further substantiate the applicability of the approach to a more general water distribution system design problem. Some of these issues are discussed next.

It has been said earlier that the maximum entropy design approach applies to looped water distribution networks only. Many looped water networks, however, are prone to water quality problems, e.g. water age. This is triggered by the uneven distribution of flow and velocity in the network, which causes some parts of the network experiencing flow congestion while in other parts water can flow freely to supply the demands. It is often difficult to pin point in advance the areas in the network sensitive to this problem. Designers have to rely on simulations to asses the ability of the preliminary designs to cope with this issue. The maximum entropy approach, on the other hand, ensures that the distribution of flow in the network is as uniform as possible. It follows that the velocity distribution in the network is also highly uniform. In theory, the uniform velocity distribution should reduce the chances of the flow in the network from being congested, hence reducing the water quality problem. However, investigations are needed to clarify this issue since, in practice, the distribution of flow in the network is often highly non uniform due to the uneven distribution of demands. Also, the velocity constraints, with minimum velocity as the limiting constraint, have to be explicitly considered in tackling the above issue.

Another possible area of research is the consideration of the sensitive customers in the network in the design optimization process. Some customers like hospitals, schools, factories, etc. may need constant supplies of water. The nodes at which these sensitive customers are located need to be more reliable than other nodes in the network. This issue may be approached using the nodal entropy values in addition to the network entropy. However, the framework in which these values are used in the design optimization procedures remains to be addressed. Also, the correlation between the nodal entropy and reliability of water distribution networks has never been previously analysed in an explicit way.

A robust method for layout and cost optimization of pipe networks based on the maximum entropy approach has been presented in Chapter 7. The problem of determining the optimum flow directions for the design is no longer an issue when Genetic Algorithms are used in the optimization process. However, a more efficient method in which a direct search towards the optimum solution without having to obtain the preliminary designs of all the possible layout configurations is desired. This method can then be extended to include other aspects of designs such as multiple demand patterns, location of sensitive customers and other network components like pumps valves and storage tanks. Finally, studies into the application of the maximum entropy approach to obtain the optimum designs - in terms of layout, pipe sizes and the network performance - of water distribution networks need to be extended further to include large real-life networks. These studies are very important since many aspects of design encountered in real water networks may not be accounted for in the study of hypothetical networks. Hence, the context in which entropy is used in this more general design problem needs to be explored.

REFERENCES

Ackley, J. R. L., Tanyimboh, T. T., Tahar, B. and Templeman, A. B. (2001) Head-driven analysis of water distribution systems. **Water Software Systems: Theory and Applications**, Ulanicki, B., Coulbeck, B. and Rance, J. (Eds.), Volume 1, 183-192.

Afshar, M. H., Akbari, M. and Marino, M. A. (2005) Simultaneous layout and size optimization of water distribution networks: engineering approach. *ASCE J. Infrastructure Systems*, **11**, (4), 221-230.

Alperovits, E. and Shamir, U. (1977) Design of optimal water distribution systems. *Water Resources Research*, **13**, (6), 885-900.

Ang, W. and Jowitt, P. W. (2003) Some observation on energy loss and network entropy in water distribution networks. *Engineering Optimization*, **35**, (4), 375-389.

Awumah, K., Goulter, I. and Bhatt, S. K. (1989) Entropy approach to redundancy and reliability assessment of water distribution networks. *Water Resources Research report*, **12**, Department of Civil Engineering, University of Manitoba, Winnipeg, Canada.

Awumah, K., Goulter, I. and Bhatt, S. K. (1990) Assessment of reliability in water distribution network using entropy based measures. *Stochastic Hydrology and Hydraulics*, **4**, (4), 309-320.

Awumah, K., Goulter, I. and Bhatt, S. K. (1991) Entropy-based redundancy measures in water distribution networks. *ASCE J. Hydraulic Engineering*, **117**, (5), 595-614.

Awumah, K. and Goulter, I. (1992) Maximizing entropy-defined reliability of water distribution networks. *Engineering Optimization*, **20**, (1), 57-80.

Bao, Y. and Mays, L. W. (1990) Model for water distribution system reliability. *ASCE J. Hydraulic Engineering*, **116**, (9), 1119-1137.

Basu, P. C. and Templeman, A. B. (1984) An efficient algorithm to generate maximum entropy distribution. *International J. Numerical Methods in Engineering*, **20**, 1039-1055.

Basu, P. C. and Templeman, A. B. (1985) Structural reliability and its sensitivity. *Civil Engineering Systems*, **2**, (1), 3-11.

Bell, M. G. H. (1983) The estimation of an origin-destination matrix from traffic counts. *Transportation Science*, **17**, (2), 198-217.

Bhave, P. R. (1981) Node flow analysis of water distribution systems. *ASCE J. Transportation Engineering*, **107**, (4), 457-467.

Bhave, P. R. (1991) **Analysis of Flow in Water Distribution Networks**, Technomic Publishing Co., Lancaster, Pa.

Carey, M. and Hendrickson, C. (1984) Bounds on expected performance of networks with links subject to failure. *Networks*, **14**, (3), 439-456.

Chandapillai, J. (1991) Realistic simulation of water distribution system. *ASCE J. Transportation Engineering*, **117**, (2), 258-263.

Chiu, C. L. (1987) Entropy and probability concepts in hydraulics. *ASCE J. Hydraulic Engineering*, **113**, (5), 583-600.

Chiu, C. L. (1988) Entropy and 2-D velocity distribution in open channels. *ASCE J. Hydraulic Engineering*, **114**, (7), 738-756.

Chiu, C. L. (1989) Velocity distribution in open channel flow. *ASCE J. Hydraulic Engineering*, **115**, (5), 576-594.

Chiu, C. L. (1991) Application of entropy concept in open channel flow study. *ASCE J. Hydraulic Engineering*, **117**, (5), 615-628.

Cover, T. M. and Thomas, J. A. (1991) **Elements of Information Theory**, John Wiley & Son Inc.

Cormen, T. H., Leiserson, C. E. and Rivest, R. L. (2001) **Introduction to Algorithms**, MIT Press, Cambridge, Mass.

Cullinane, M. J., Lansey, K. E. and Mays, L. W. (1992) Optimization-availability-based design of water distribution networks. *ASCE J. Hydraulic Engineering*, **118**, (3), 420-441.

Dorigo, M., Maniezzo, V. and Colorni, A. (1996) The ant system: optimization by a colony of cooperating ants. *IEEE Transactions on Systems, Man and Cybernetics*, **26**, 29-42.

Erlander, S. (1977) Accessibility, entropy and the distribution and assignment of traffic. *Transportation Research*, **11**, 149-153.

Fujiwara, O. and De Silva, A. U. (1990) Algorithm for reliability based optimal design of water networks. *ASCE J. Environmental Engineering*, **116**, (3), 575-587.

Fujiwara, O. and Ganesharajah, T. (1993) Reliability assessment of water supply systems with storage and distribution networks. *Water Resources Research*, **29**, (8), 2917-2924.

Fujiwara, O. and Kang, D. (1990) A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resources Research*, **26**, (4), 539-549.

Fujiwara, O. and Tung, H. D. (1991) Reliability improvement for water distribution networks through increasing pipe size. *Water Resources research*, **27**, (7), 1395-1402.

Gargano, R. and Pianese, D. (2000) Reliability as tool for hydraulic network planning. *ASCE J. Hydraulic Engineering*, **126**, (5), 354-364.

Gentle, J. E. (2003) **Random Number Generation and Monte Carlo Methods**, Springer, New York, USA.

Gouldberg, D. E. (1989) **Genetic Algorithms in Search, Optimization and Machine Learning**, Addison-Wesley Publishing Co., Reading, Mass.

Goldberg, D. E. and Kuo, C. H. (1987) Genetic algorithms in pipeline optimization. *J. Computing in Civil Engineering*, **1**, (2), 128-141.

Goulter, I. (1995) Analytical and Simulation models for Reliability Analysis in Water Distribution Systems. **Improving Efficiency and Reliability in Water Distribution Systems**, Cabrera, E. and Vela, A. F. (Eds.), Volume 14, 235-266.

Guiasu, S. (1977) **Information Theory with Applications**, McGraw-Hill Inc.

Gupta, R. and Bhave, P. R. (1996) Comparison of methods for predicting deficient network performance. *ASCE J. Water Resources Planning and Management*, **122**, (3), 214-217.

Halhal, D., Walters, G. A., Ouazar, D. and Savic, D. A. (1997) Water network rehabilitation with structured messy genetic algorithms. *ASCE J. Water Resources Planning and Management*, **123**, (3), 137-146.

Holland, J. H. (1975) **Adaptation in Natural and Artificial Systems**, MIT Press, Cambridge, Mass.

Isaacs, L. T. and Mills, K. G. (1980) Linear theory methods for pipe network analysis. *ASCE J. Hydraulics Division*, **106**, (7), 1191-1201.

Jaynes, E. T. (1957) Information theory and statistical mechanics. *Phys. Rev.*, **106**, 620-630 and **108**, 171-190.

Jeppson, R. W. (1976) **Analysis of Flow in Pipe Networks**, Ann Arbor Science, Ann Arbor (Mich.).

Jones, D. S. (1979) **Elementary Information Theory**, Clarendon Press, Oxford.

Kalungi, P. (2003) **A Holistic Approach to the Optimal Long-term Upgrading of Water Distribution Networks**. PhD Thesis, University of Liverpool, UK.

Kalungi, P. and Tanyimboh, T. T. (2003) Redundancy model for water distribution systems. *Reliability Engineering and System Safety*, **82**, (3), 275-286.

Kapur, J. N. (1989) **Maximum Entropy Models in Science and Engineering**, John Wiley and Sons Inc.

Kapur, J. N. and Kesavan, H. K. (1987) **The Generalized Maximum Entropy Principle (with Application)**, Sandford Educational Press, Waterloo (Ont.).

Kessler, A. and Shamir, U. (1989) Analysis of the linear programming gradient method for optimal design of water supply networks. *Water Resources Research*, **25**, (7), 1469-1480.

Khinchin, A. I. (1953) The entropy concept in probability theory. *Uspekhi Matematicheskikh Nauk*, **8**, (3), 3-20. Translation in A. I. Khinchin (1957) **Mathematical Foundations of Information Theory**, Dover, New York, 1-28.

Kumar, V. (1987) Entropic measure of manufacturing flexibility. *International J. Production Research*, **25**, 957-966.

Levine, R. D. and Tribus, M. (1979) **The Maximum Entropy Formalism**, M. I. T. Press.

Li, X. (1987) **Entropy and Optimization**, Ph.D. Thesis, Department of Civil Engineering, University of Liverpool, U.K.

Li, X. and Templeman, A. B. (1988) Entropy-based optimum sizing of trusses. *Civil Engineering Systems*, **5**, 121-128.

Loganathan, G., Greene, J. and Ahn, T. (1995) Design heuristic for globally minimum cost water distribution systems. *ASCE J. Water Resources Planning and Management*, **121**, (2), 182-192.

Loganathan, G. V., Sherali, H. D. and Shah, M. P. (1990) A two-phase network design heuristic for minimum cost water distribution systems under a reliability constraint. *Engineering Optimization*, **15**, (4), 311-336.

Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seah, H. Y. and Tan, C. L. (2003) Ant colony optimization for design of water distribution systems. *ASCE J. Water Resources Planning and Management*, **129**, (3), 200-209.

Marti, K. (2005) **Stochastic Optimization Methods**, Springer, Berlin, Germany.

Martin, D. W. and Peters, G. (1963) The application of Newton's method to network analysis by digital computer. *J. Institution of Water Engineers*, **17**, 115-129.

Mays, L. W. (1989) **Reliability Analysis of Water Distribution Systems**, ASCE.

Mays, L. W. (2000) **Water Distribution Systems Handbook**, McGraw-Hill, New York, USA.

Mitchell, M. (1999) **An Introduction to Genetic Algorithms**, MIT Press, Cambridge, Mass.

Morgan, D. R. and Goulter, I. C. (1985) Optimal urban water distribution design. *Water Resources Research*, **21**, (5), 642-652.

Mountain, L. and Steele, D. (1983a) Prior information and the accuracy of turning flow estimates. *Traffic Engineering and Control*, **24**, (12), 582-588.

Mountain, L. J. and Westwell, P. M. (1983b) The accuracy of estimation of turning flows from traffic counts. *Traffic Engineering and Control*, **24**, (1), 3-7.

Mountain, L., Maher, M. and Maher, S. (1986a) The estimation of turning flows from traffic counts. 1. At four-arm intersections. *Traffic Engineering and Control*, **27**, (10), 501-507.

Mountain, L., Maher, M. and Maher, S. (1986b) The estimation of turning flows from traffic counts. 2. At five-arm intersections. *Traffic Engineering and Control*, **27**, (11), 566-569.

Munro, J. and Jowitt, P. W. (1978) Decision analysis in the ready-mixed concrete industry. *Proceedings of The Institution of Civil Engineers, Part 2*, **65**, 41-52.

Provan, J. S. and Ball, M. O. (1983) The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM J. Computing*, **12**, (4), 777-788.

NAG Ltd. (1995) **NAG Fortran Library Manual**. Mark 18, Vol. 4. The Numerical Algorithms Group Ltd , Oxford, UK.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992) **Numerical Recipes in FORTRAN**, 2nd edn, Cambridge University Press.

Quindry, G. E., Brill, E. D. and Liebman, J. C. (1981) Optimization of looped water distribution systems. *ASCE J. Environmental Engineering Division*, **107**, (4), 665-679.

Rossman, L. A. (2000) **EPANET 2 Users Manual**. U.S. Environmental Protection Agency, Cincinnati, USA.

Rowell, W. F. and Barnes, J. (1982) Obtaining layout of water distribution systems. *ASCE J. Hydraulics Division*, **108**, (1), 137-148.

Savic, D. A. and Walters, G. A. (1997) Genetic algorithms for least-cost design of water distribution networks. *ASCE J. Water Resources Planning and Management*, **123**, (2), 67-77.

Shamir, U. and Howard, C. D. D. (1968) Water distribution systems analysis. *ASCE J. Hydraulics Division*, **94**, (1), 219-234.

Shannon, C. E. (1948) A Mathematical theory of communication. *Bell System Technical J.*, **27**, (3), 379-428.

Simpson, A. R., Dandy, G. C. and Murphy, L. J. (1994) Genetic algorithms compared to other techniques for pipe optimization. *ASCE J. Water Resources Planning and Management*, **120**, (4), 423-443.

Stützle, T. and Hoos, H. H. (2000) MAX-MIN ant system. *Future Generation Computer Systems*, **16**, 889-914.

Su, Y., Mays, L. W., Duan, N. and Lansey, K. E. (1987) Reliability-based optimization model for water distribution systems. *ASCE J. Hydraulic Engineering*, **142**, (12), 1539-1556.

Surendran, S., Tanyimboh, T. T. and Tabesh, M. (2005) Peaking demand factor-based reliability analysis of water distribution systems. *Advances in Engineering Software*, **36**, 789-796.

Tabesh, M. (1998) **Implications of the Pressure Dependency of Outflows on Data Management, Mathematical Modelling and Reliability Assessment of Water Distribution Systems**. PhD Thesis, University of Liverpool, UK.

Tahar, B., Tanyimboh, T. T. and Templeman, A. B. (2002) Pressure-dependent modelling of water distribution systems. In *Proceedings of 3rd International*

Conference on Decision Making in Urban and Civil Engineering, London. Khosrowshahi F (Ed.). CD-ROM.

Tanyimboh, T. T. (1993) **An Entropy-based Approach to the Optimum Design of Reliable Water Distribution Networks**. PhD Thesis, University of Liverpool, UK.

Tanyimboh, T. T. (2003) Reliability analysis of water distribution systems. In *Urban and Rural Water Systems for Sustainable Development: Proceedings of the 30th IAHR Congress,* Thessaloniki. Ganoulis, J., Maksimovic, C. and Kaleris, V. (Eds.), ISBN 960-243-594-1, 321-328.

Tanyimboh, T. T., Setiadi, Y., Mavrokoukoulaki, A. and Storer, C. (2002) The suitability of information-theoretic entropy as a surrogate performance measure for water distribution systems. In *Proceedings of 3rd International Conference on Decision Making in Urban and Civil Engineering*, London. Khosrowshahi F (Ed.), CD-ROM.

Tanyimboh, T. T. and Sheahan, C. (2002) A maximum entropy based approach to the layout optimization of water distribution systems. *J. Civil Engineering and Environmental Systems*, **19**, (3), 223-253.

Tanyimboh, T. T., Tabesh, M. and Burrows, R. (2001) Appraisal of source head methods for calculating reliability of water distribution networks. *ASCE J. Water Resources Planning and Management*, **127**, (4), 206-213.

Tanyimboh, T., Tahar, B. and Templeman, A. (2003) Pressure-driven modelling of water distribution systems. *Water Science and* Technology, **3**, (1-2), 255-262.

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) Optimum design of flexible water distribution networks. *Civil Engineering Systems*, **10**, (3), 243-258.

Tanyimboh, T. T. and Templeman, A. B. (1993c) Maximum entropy flows for single-source networks. *Engineering Optimization*, **22**, (1), 49-63.

Tanyimboh, T. T. and Templeman, A. B. (1998) Calculating the reliability of single-source networks by the source head method. *Advances in Engineering 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.

Tanyimboh, T. T. and Templeman, A. B. (2004) A new nodal outflow function for water distribution networks. In *Proceedings of 4th International Conference on Engineering Computational Technology*, Lisbon. BHV Topping and CA Mota Soares (Eds.). Paper 64.

Templeman, A. B. (1982) Discussion of "Optimization of looped water distribution systems". *ASCE J. Environmental Engineering Division*, **108**, (3), 599-602.

Templeman, A. B. and Li, X. (1985) Entropy Duals. *Engineering Optimization*, **9**, (2), 107-119.

Templeman, A. B. and Li, X. (1987) A maximum entropy approach to constrained non-linear programming. *Engineering Optimization*, **12**, (2), 191-205.

Templeman, A. B. and Li, X. (1989) Maximum entropy and constrained optimisation. **Maximum Entropy and Bayesian Methods**, J. Skilling (Ed), Kluwer Academic Publishers, 447-454.

Templeman, A. B. and Yassin-Kassab, A. (2002) Robust Calibration of Computer Models of Engineering Networks with Limited Data. **Optimization in Industry**, I. C. Parmee and P. Hajela, (Eds.), Springer-Verlag, London, 71-81.

Tolson, B. A., Maier, H. R., Simpson, A. R. and Lence, B. J. (2004) Genetic algorithms for reliability-based optimization of water distribution systems. *ASCE J. Water Resources Planning and Management*, **130**, (1), 63-72.

Vairavamoorthy, K. and Ali, M. (2000) Optimal design of water distribution systems using genetic algorithms. *Computer-aided Civil and Infrastructure Engineering*, **15**, 374-382.

Van Zuylen, H. J. and Willumsen, L. G. (1980) The most likely trip matrix estimated from traffic counts. *Transportation Research*, **14B**, (3), 281-293.

Wagner, J. M., Shamir, U. and Marks, D. H. (1988a) Water distribution reliability: analytical methods. *ASCE J. Water Resources Planning and Management*, **114**, (3), 253-275.

Wagner, J. M., Shamir, U. and Marks, D. H. (1988b) Water distribution reliability: simulation methods. *ASCE J. Water Resources Planning and Management*, **114**, (3), 276-294.

Walski, T. M., Chase, D. V., Savic, D. A., Grayman, W., Beckwith, S. and Koelle, E. (2003) **Haestad Methods. Advanced Water Distribution Modeling and Management**, Haestad Press, Waterbury, CT, USA.

Walski, T. M. and Pelliccia, A. (1982) Economic analysis of water main breaks. *J. American Water Works Association*, **74**, (3), 140-147.

Walters, G. A. (1995) Discussion on: Maximum entropy flows in single-source networks. *Engineering Optimization*, **25**, (2), 155-163.

Wong, S. C., Tong, C. O., Wong, K. I., Lam, W. H. K., Lo, H. K., Yang, H. and Lo, H. P. (2005) Estimation of multiclass origin-destination matrices from traffic counts. *ASCE J. Urban Planning and Development*, **131**, (1), 19-29.

Wood, D. J. and Charles, O. A. (1972) Hydraulic network analysis using linear theory. *ASCE J. Hydraulics Division*, **98**, (7), 1157-1170.

Xu, C. and Goulter, I. C. (1999) Reliability-based optimal design of water distribution networks. *ASCE J. Water Resources Planning and Management*, **125**, (6), 352-362.

Yao, D. D. (1985) Material flow and information flow in flexible manufacturing systems. *Material Flow*, **2**, 143-149.

Yassin-Kassab, A. (1998) **Entropy-based Inference and Calibration Methods for Civil Engineering System Models under Uncertainty.** PhD Thesis, University of Liverpool, UK.

Yassin-Kassab, A., Templeman, A. B. and Tanyimboh, T. T. (1999) Calculating Maximum Entropy Flows in Multi-source, Multi-demand Networks. *Engineering Optimization*, **31**, (6), 695-729.

Yates, D. F., Templeman, A. B. and Boffey, T. B. (1984) The computational complexity of determining least capital cost designs for water supply networks. *Engineering Optimization*, **7**, (2), 143-155.

APPENDIX A. INFORMATION REGARDING ENTROPY, RELIABILITY AND THE SUMMARY OF THE DEMAND DRIVEN ANALYSIS RESULTS

APPENDIX A1 – THE PATH-BASED APPROACH FOR CALCULATING THE MAXIMUM ENTROPY FLOWS IN SINGLE SOURCE NETWORK

To demonstrate the path-based approach for calculating the maximum entropy flows in single source networks, Tanyimboh and Templeman (1993c) used a single source network presented in Figure 3.3 and re-presented below as Figure A1.1. The equal path flows from the source to each demand node are shown in Figure A1.2 below. For example, node 4 is served by three paths 1-2-3-4, 1-3-4 and 1-4, each of which must carry 5 unit of flow, which is one-third of the demand at node 4. Also, node 5 is served by three paths 1-2-5, 1-2-3-5 and 1-3-5 and each path must carry 8 unit of flow, which is the demand at node 5 divided equally amongst the path supplying the node. Finally, the maximum entropy flow for each link is obtained by adding the flow for all paths through that link. The resulting maximum entropy link flows are identical as those shown in Figure 3.5.

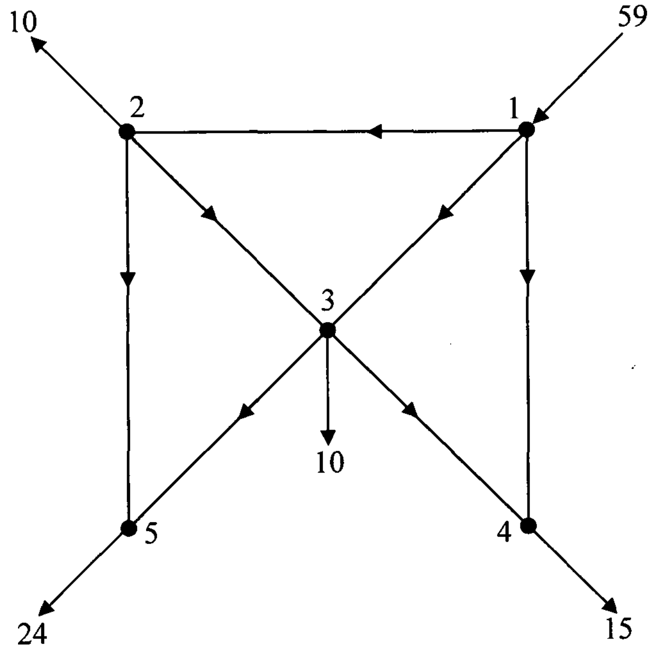


Figure A1.1. Single source network (Tanyimboh and Templeman, 1993c).

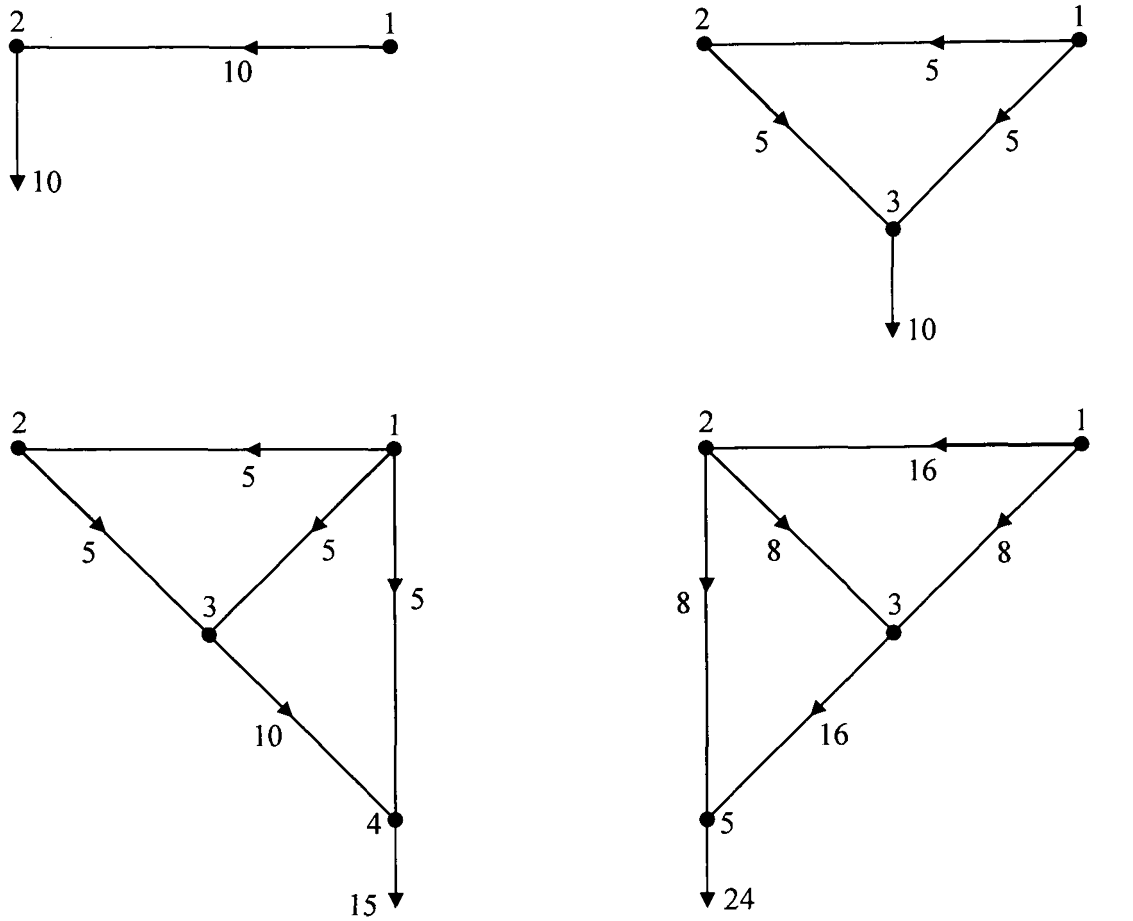


Figure A1.2. Equal path flows from the source to each demand node for network shown in Figure A1.1 (Tanyimboh and Templeman, 1993c).

APPENDIX A2 – DERIVATION OF TANYIMBOH'S RELIABILITY

It was mentioned in Chapter 5 that the first part of the Equation (5.12) is the basic definition of hydraulic reliability, which is defined as the time-averaged value of the ratio of the flow delivered to the flow required, i.e.

$$\overline{R} = \frac{1}{T}\left(p(0)T(0) + \sum_{m=1}^{M} p(m)T(m) + \sum_{m=1}^{M-1}\sum_{n=m+1}^{M} p(m,n)T(m,n) + \cdots \right) \tag{A2.1}$$

in which $\overline{R}$ is the lower bound to the reliability since terms corresponding to the network configuration with more than two pipes simultaneously unavailable are missing. The motivation for Equation (5.12) originated from an observation (Tanyimboh et al., 2001) that if, instead of the flow delivered, the shortfall in supply is used in Equation (A2.1) then the resulting expression gives the system unreliability $\overline{U}$ as

$$\overline{U} = \frac{1}{T}\left(p(0)[T - T(0)] + \sum_{m=1}^{M} p(m)[T - T(m)] + \sum_{m=1}^{M-1}\sum_{n=m+1}^{M} p(m,n)[T - T(m,n)] + \cdots \right)$$

$$= \left(p(0) + \sum_{m=1}^{M} p(m) + \sum_{m=1}^{M-1}\sum_{n=m+1}^{M} p(m,n) + \cdots \right) - \overline{R} \tag{A2.2}$$

Equation (A2.2), like (A2.1), gives a lower bound because, in practice, it is unlikely to include all possible combinations of unavailable links. Therefore, the complement, $1-\overline{U}$, is an upper bound to the reliability of the system. It has been demonstrated that by averaging the lower and upper bound estimates of reliability, an improved estimate of the reliability can be obtained (Tanyimboh et al., 2001). Hence

$$R = \frac{\overline{R} + \left(1 - \overline{U}\right)}{2} \Rightarrow 2R - 1 = \overline{R} - \overline{U} \tag{A2.3}$$

Substituting for $\overline{U}$ from equation (A2.2) and rearranging gives

$$R = \overline{R} + \frac{1}{2}\left(1 - p(0) - \sum_{m=1}^{M} p(m) - \sum_{m=1}^{M-1}\sum_{n=m+1}^{M} p(m,n) - \cdots \right) \tag{A2.4}$$

And after substituting for $\overline{R}$, Equation (A2.4) is identical to Equation (5.12).

APPENDIX A3 – SUMMARY OF THE PREVIOUS DDA-BASED STUDY ON THE POSSIBLE INFLUENCE OF LAYOUT ON THE RELATIONSHIP BETWEEN ENTROPY AND RELIABILITY (TANYIMBOH AND SHEAHAN, 2002)

For the analysis, Tanyimboh and Sheahan (2002) generated 65 different layout configurations based on the network in Figure 6.1. These layouts are shown in Figure A3.1. The results of the DDA analysis are shown in Figure A3.2. The plots of entropy against reliability show that there is an increase in the reliability value as the entropy of the network increases. There is also a gradual increase in cost as the reliability value increases, which is expected, but this increase seems to be outweighed by the increase in the performance of the network. Figure A3.2 also shows that there is some scatter in the plot of the entropy against reliability, which gives an R^2 value of 0.5. Although the relationship does not seem very strong, the fact that there is a correlation between the increase of entropy and reliability seems to indicate that the influence of different layouts chosen in the design process is quite small. Therefore, the choice of layout in a design process can be determined by other factors, for example distribution of demands or topological requirements. So long as the maximum entropy value of the network is high, the reliability of the network can also be expected to be high.

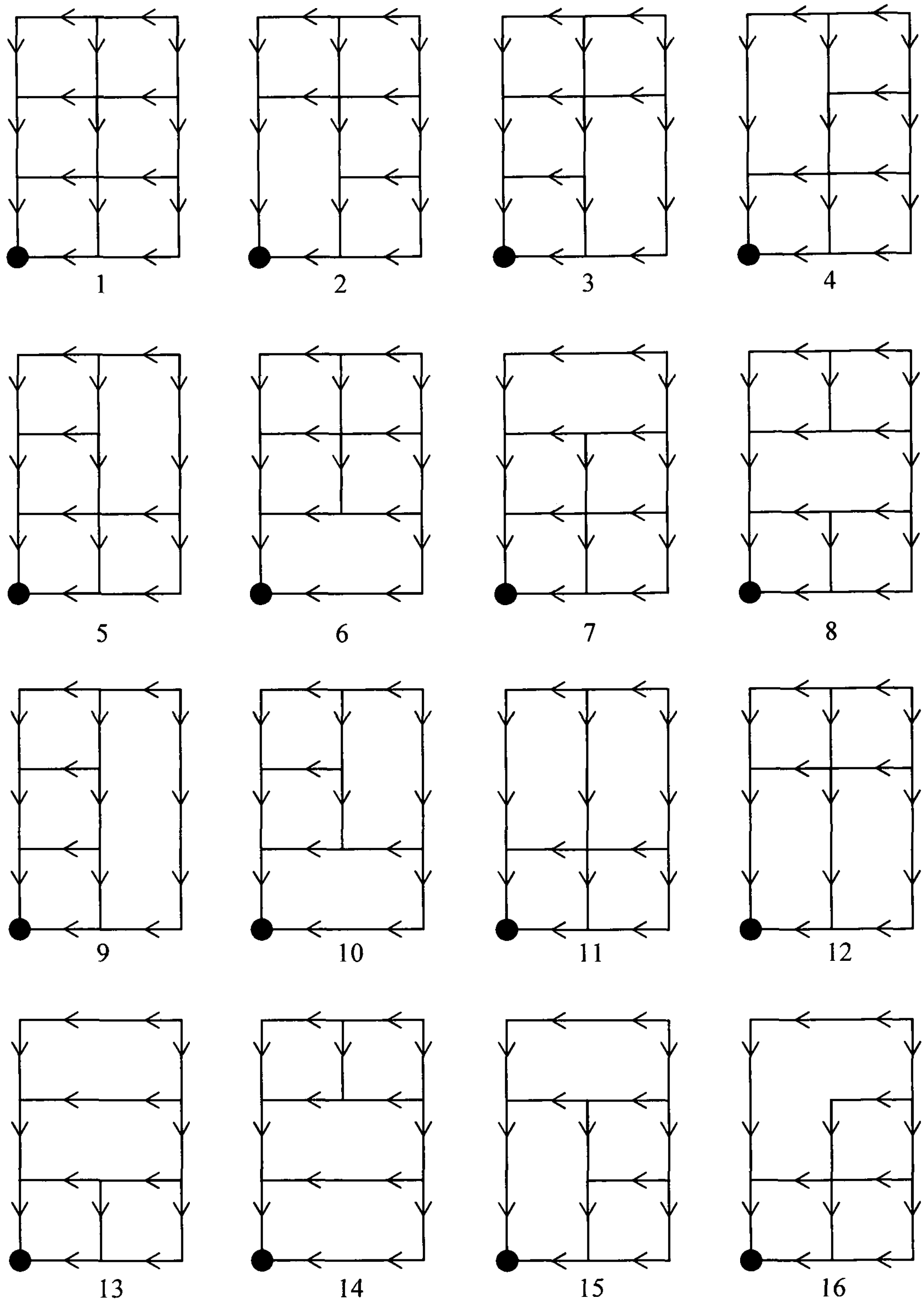


Figure A3.1. Alternate layouts based on the network of Figure 6.1 (Tanyimboh and Sheahan, 2002).

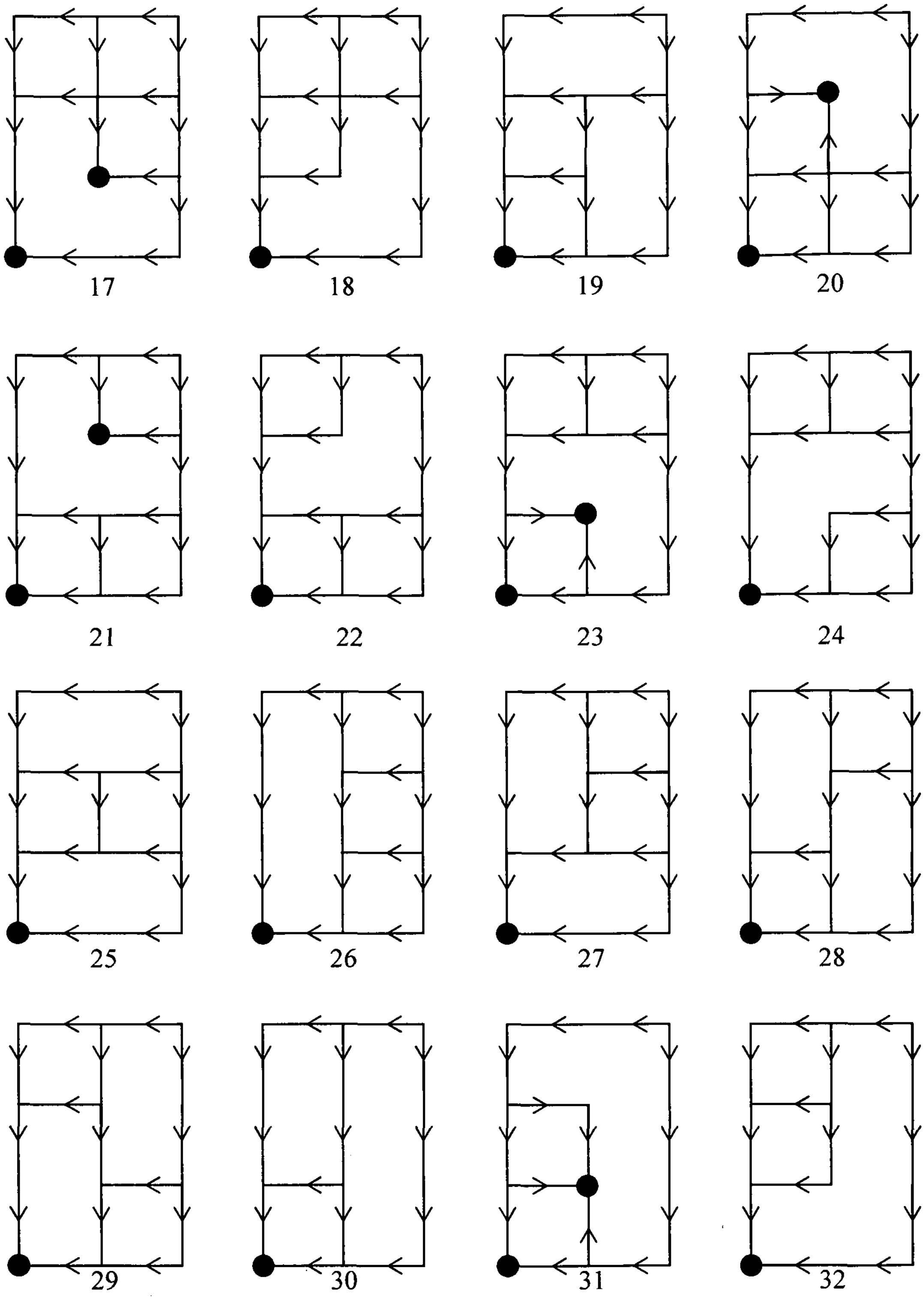


Figure A3.1. (Continued).

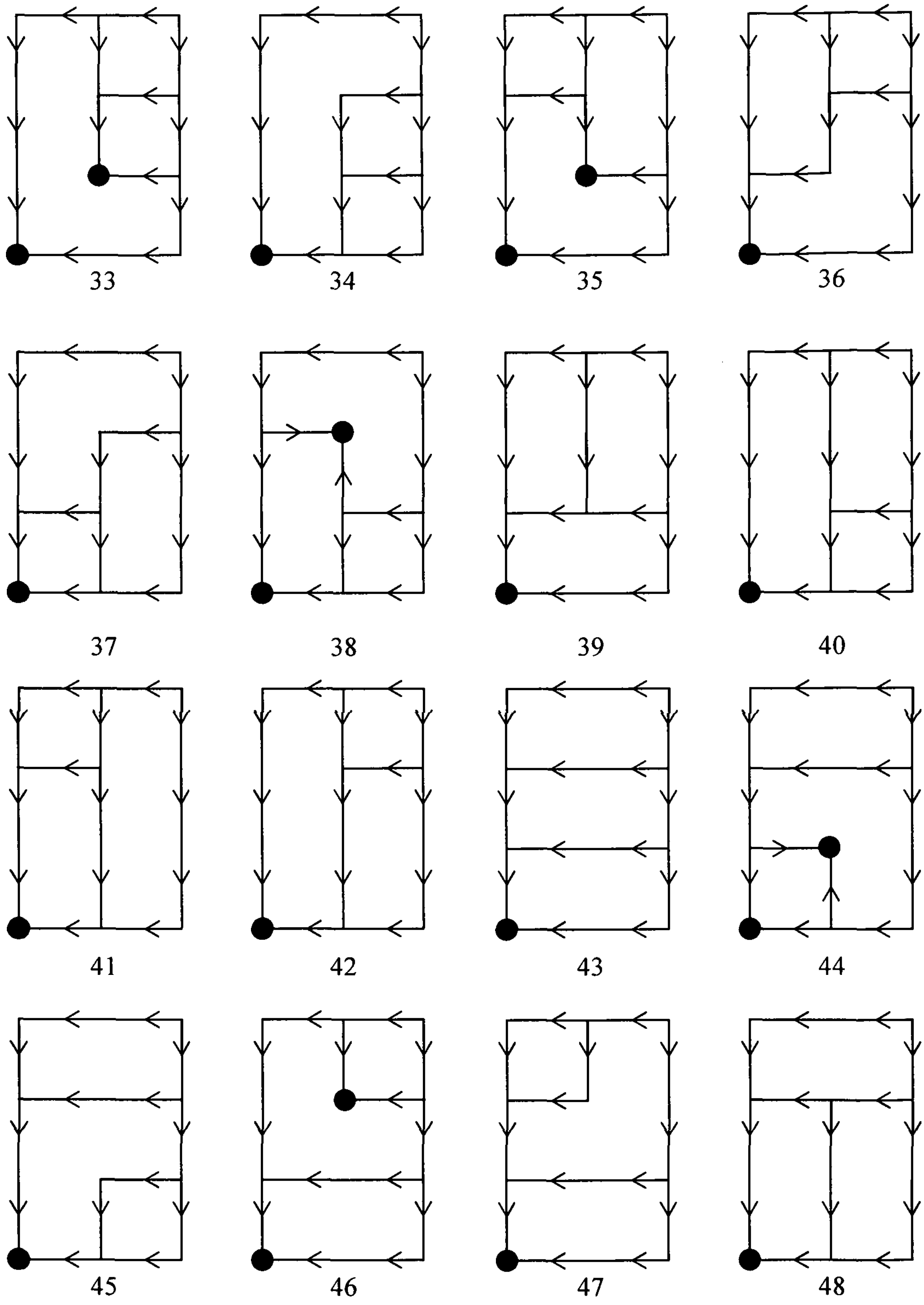


Figure A3.1. (Continued).

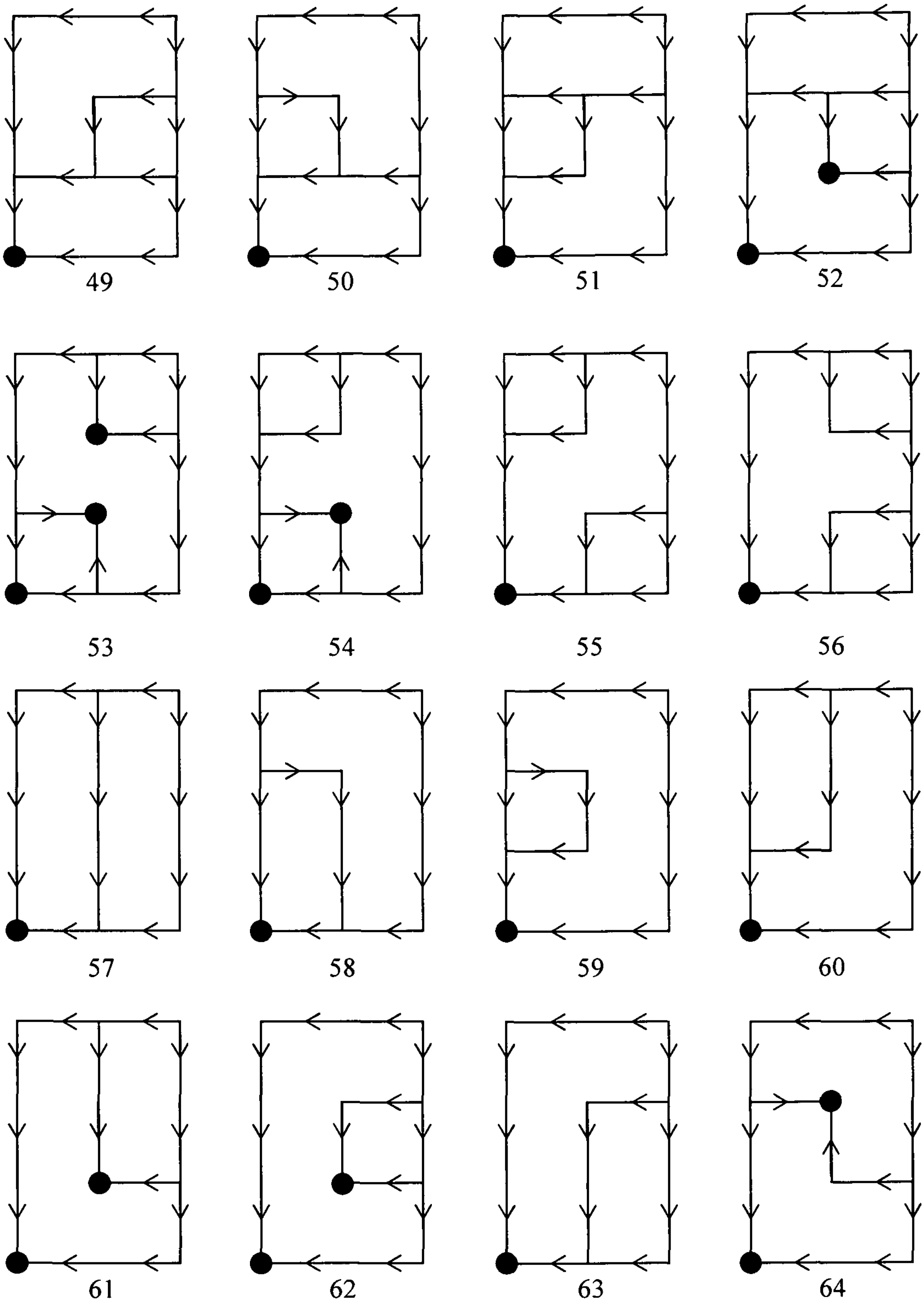


Figure A3.1. (Continued).

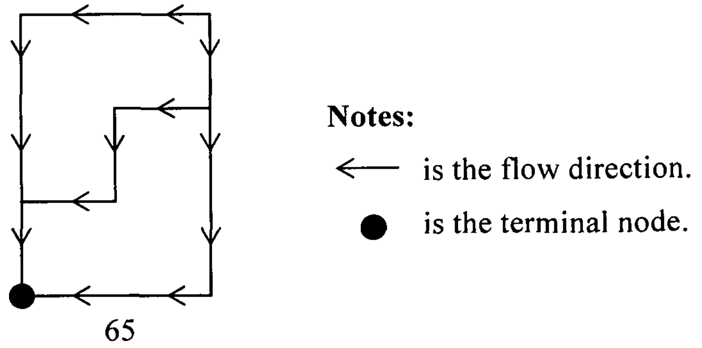


Figure A3.1. (Continued).

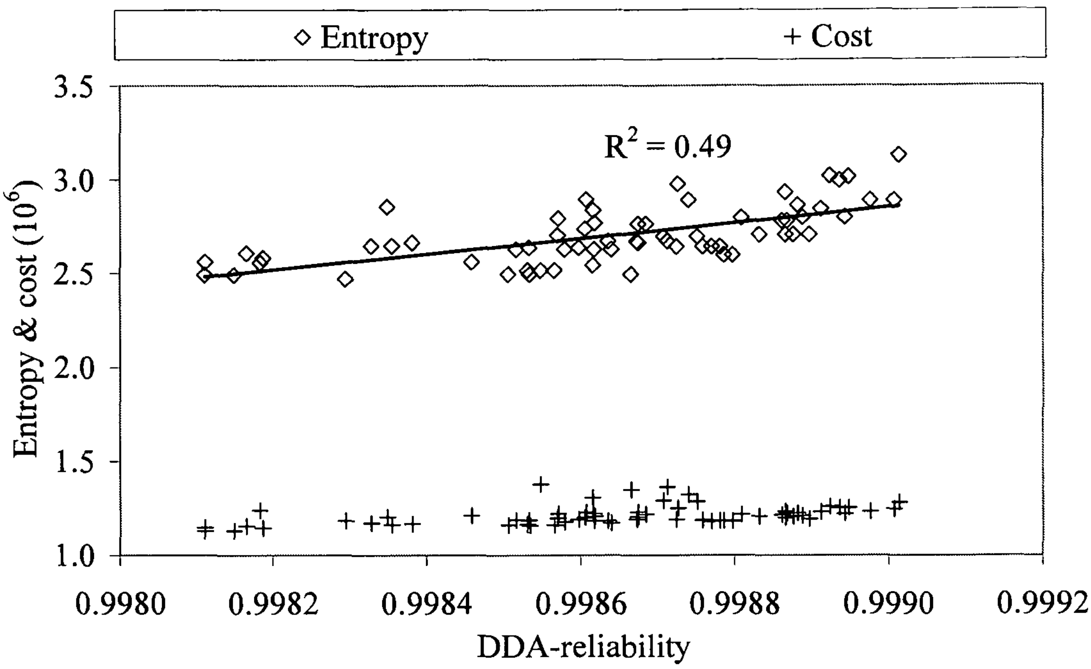


Figure A3.2. Effect of layout on the entropy-reliability relationship analysed by the DDA method.

APPENDIX A4 – RESULTS OF THE DDA-BASED STUDY OF THE POSSIBLE INFLUENCE OF FLOW DIRECTIONS ON THE RELATIONSHIP BETWEEN ENTROPY AND RELIABILITY

Figures A4.1, A4.2 and A4.3 show the complete sets of flow directions for the respective 2-loop, 3-loop and 6-loop layouts used in the present study. The optimum designs generated based on these flow directions are presented in Tables A4.1 to A4.3. Meanwhile, the results of the DDA analysis are shown in Figures A4.4 to A4.7. The same conclusions can be drawn from the DDA-based study to those obtained from the HDA method in which the narrow ranges in the entropy and reliability values lead to inconclusive results when the three layouts were analyzed separately. A more apparent relationship was obtained when the results were combined together. However, the DDA results seem to show a weaker relationship between entropy and reliability compared to the HDA results.

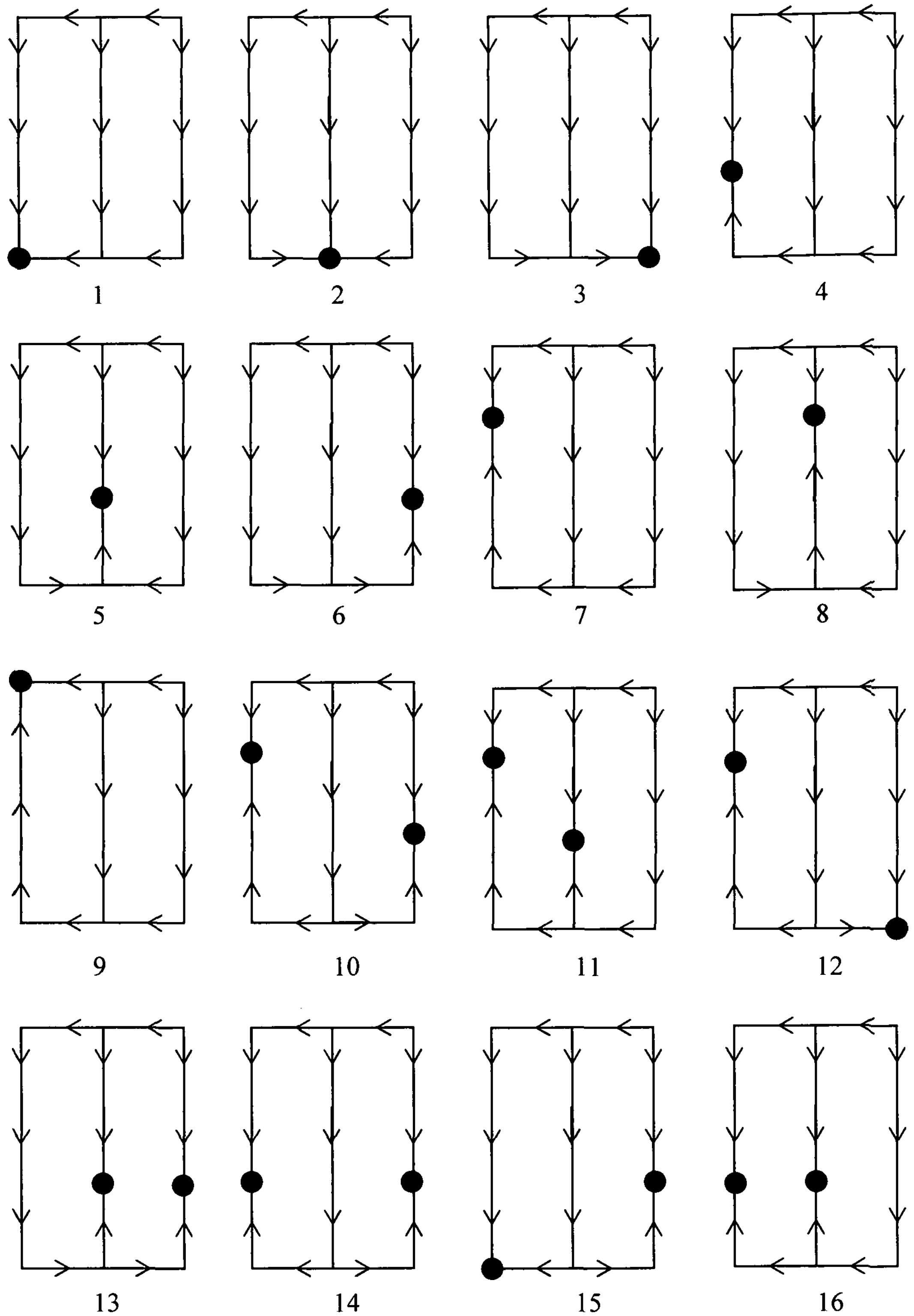


Figure A4.1. Alternate flow directions based on the layout of Figure 6.3a.

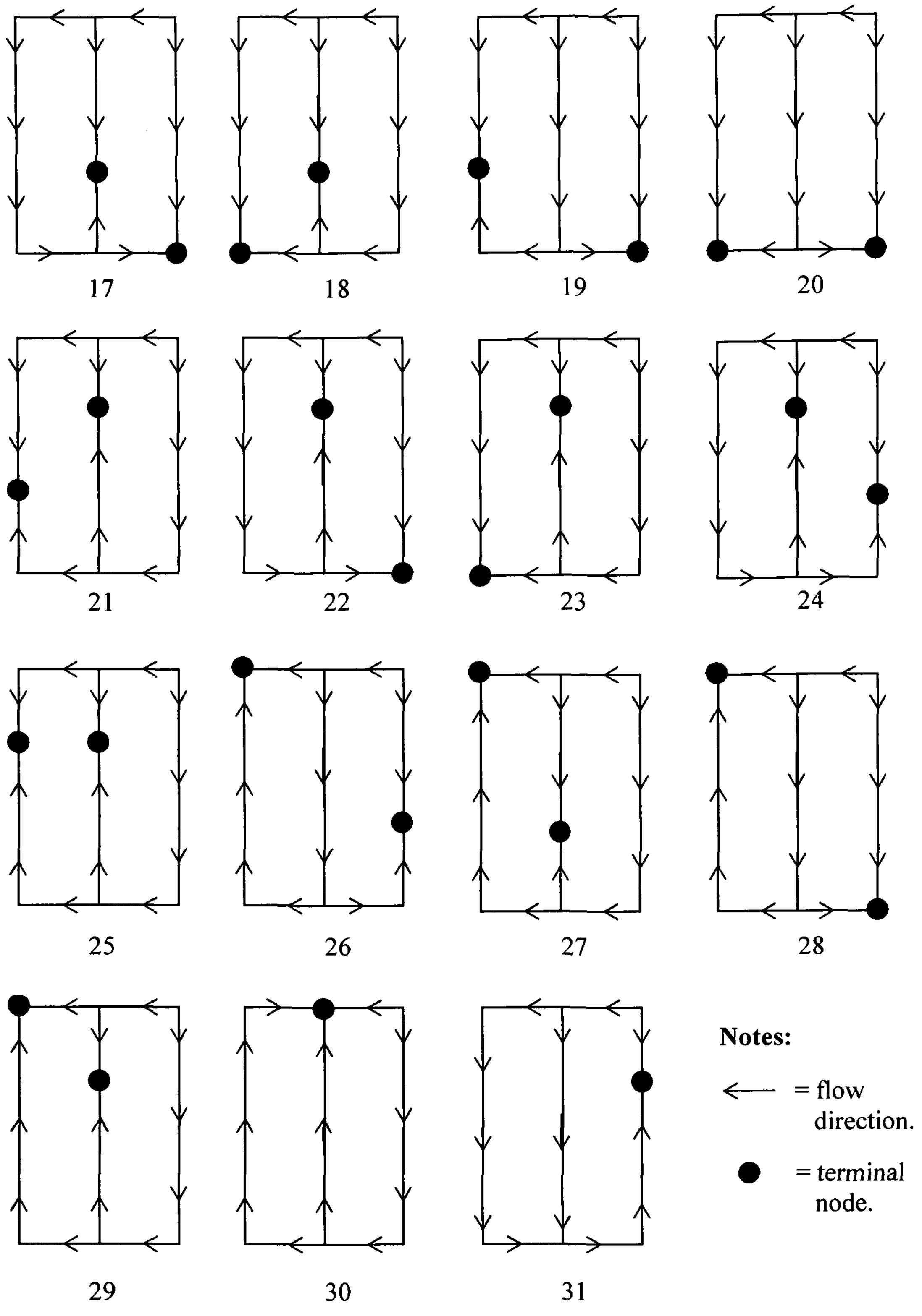


Figure A4.1. (Continued).

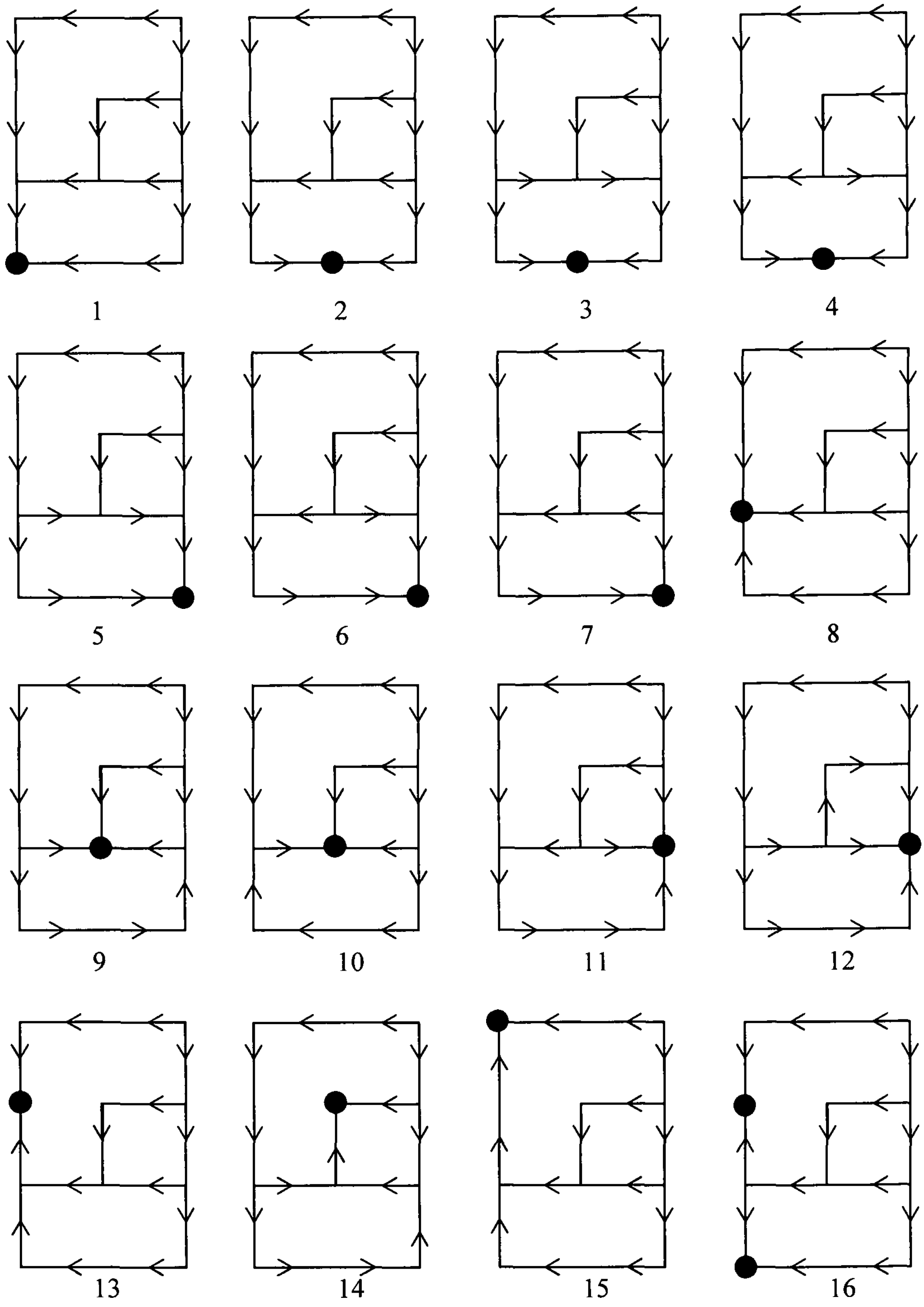
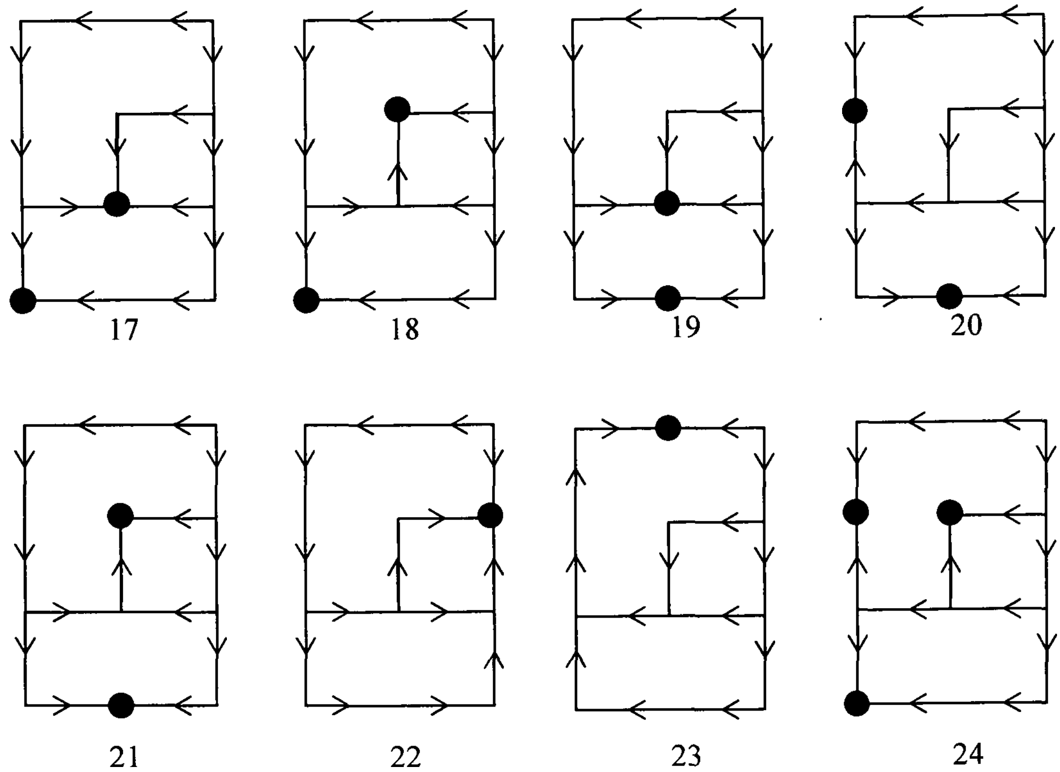


Figure A4.2. Alternate flow directions based on the layout of Figure 6.3b.



Notes:

← = flow direction.

● = terminal node.

Figure A4.2. (Continued).

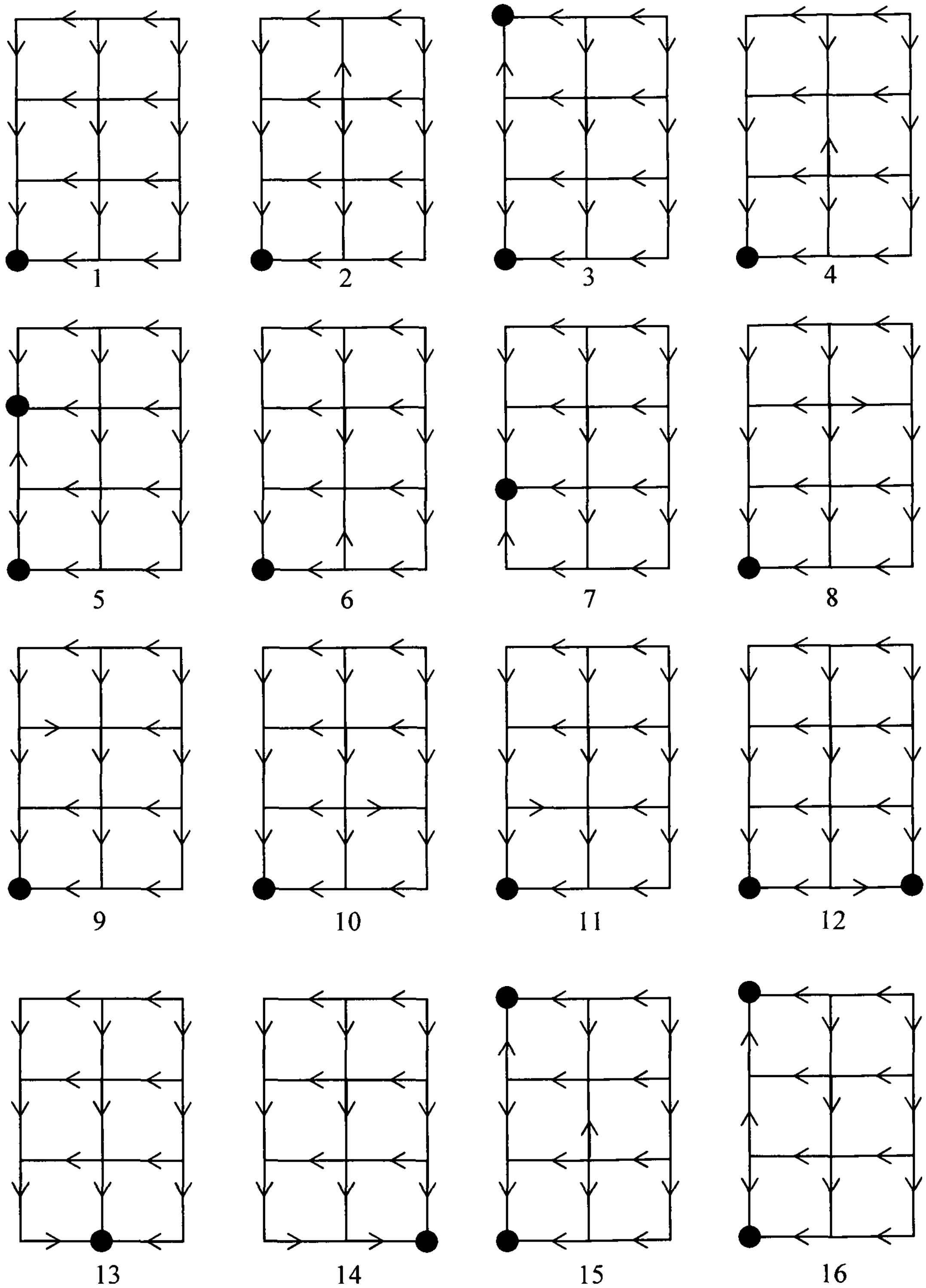


Figure A4.3. Alternate flow directions based on the layout of Figure 6.1.

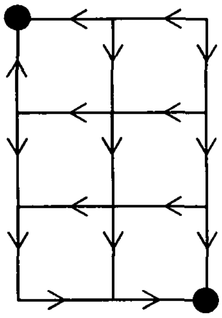
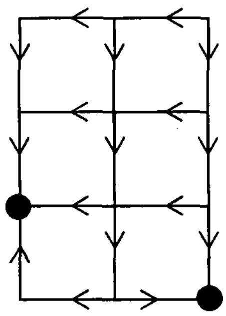
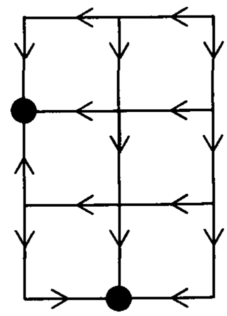
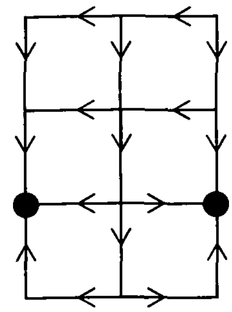
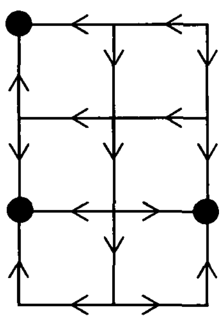
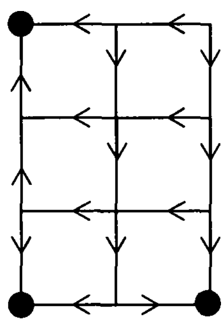
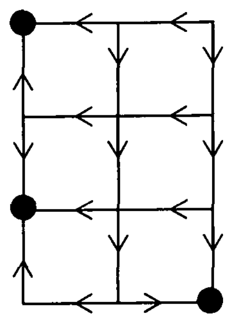
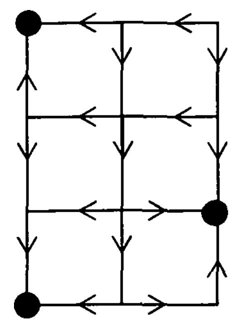
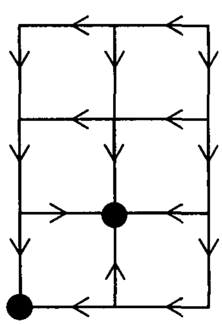
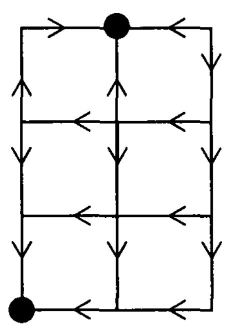
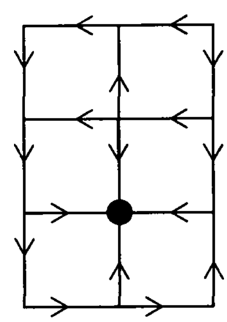
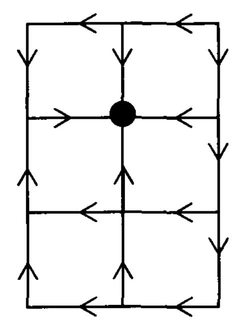
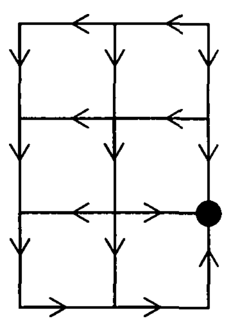
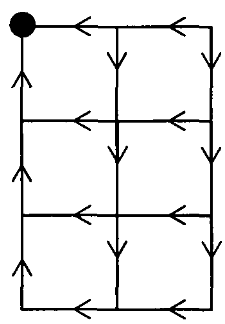
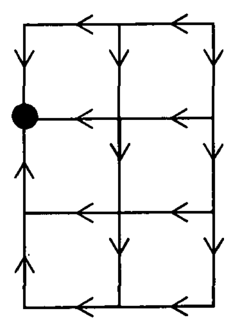
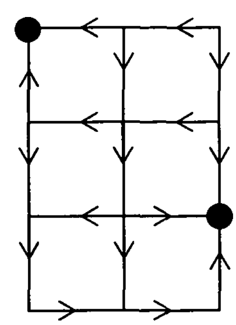

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

Figure A4.3. (Continued).

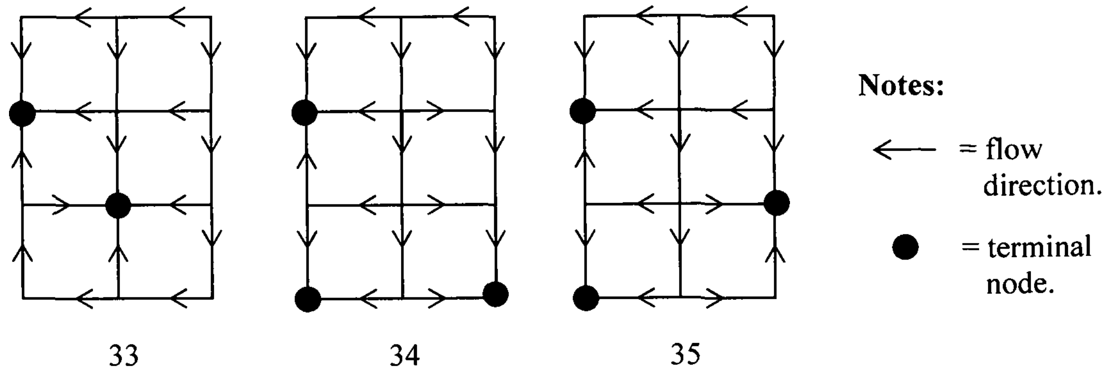


Figure A4.3. (Continued).

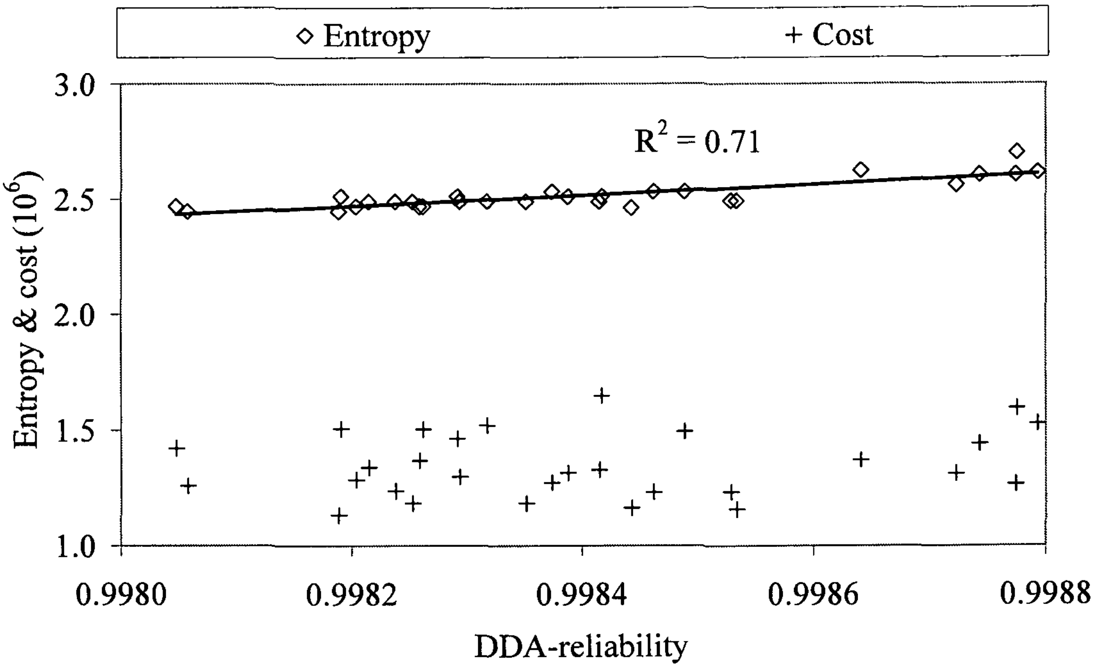


Figure A4.4. Effect of flow direction on the relationship between entropy and DDA-reliability for the 2-loop designs.

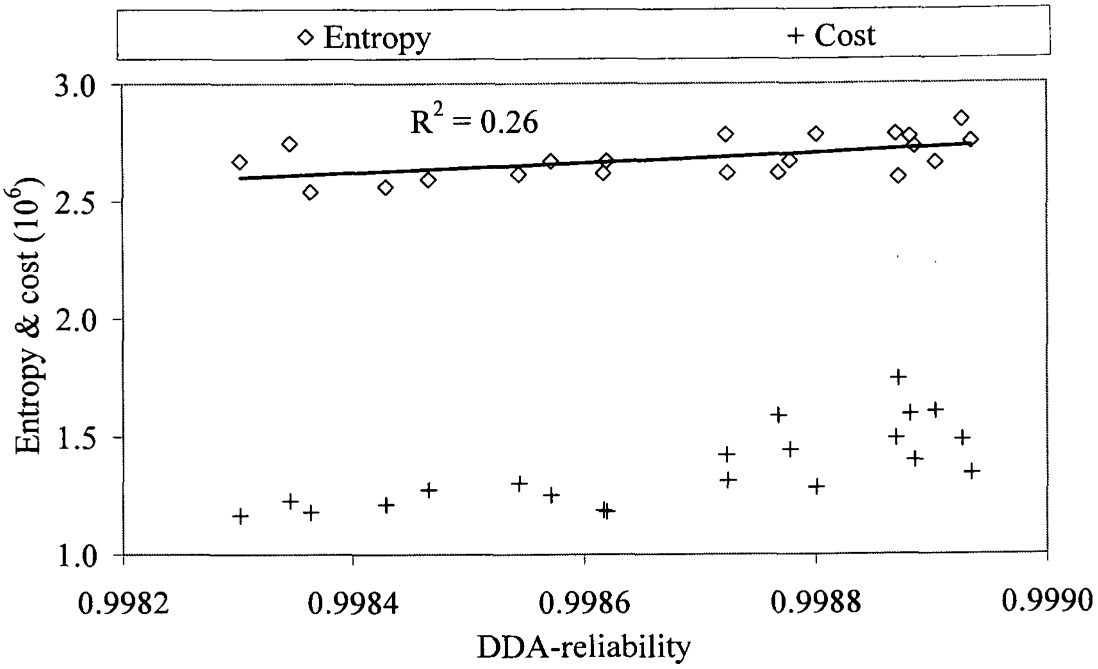


Figure A4.5. Effect of flow direction on the relationship between entropy and DDA-reliability for the 3-loop designs.

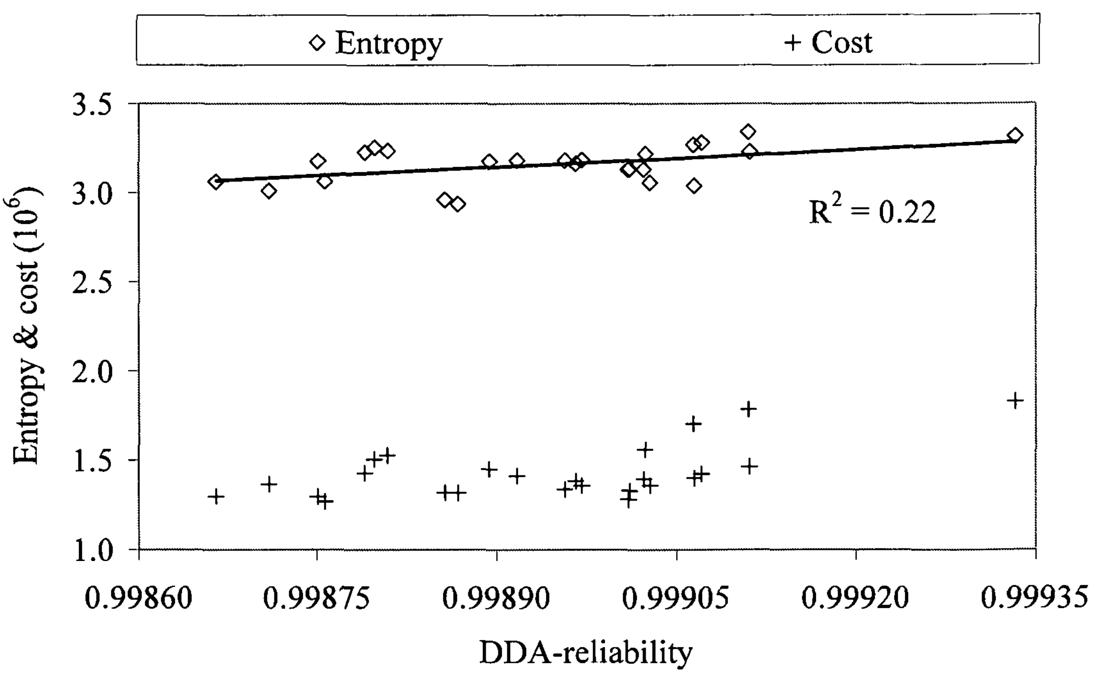


Figure A4.6. Flow direction effect on the relationship between entropy and DDA-reliability for the 6-loop designs.

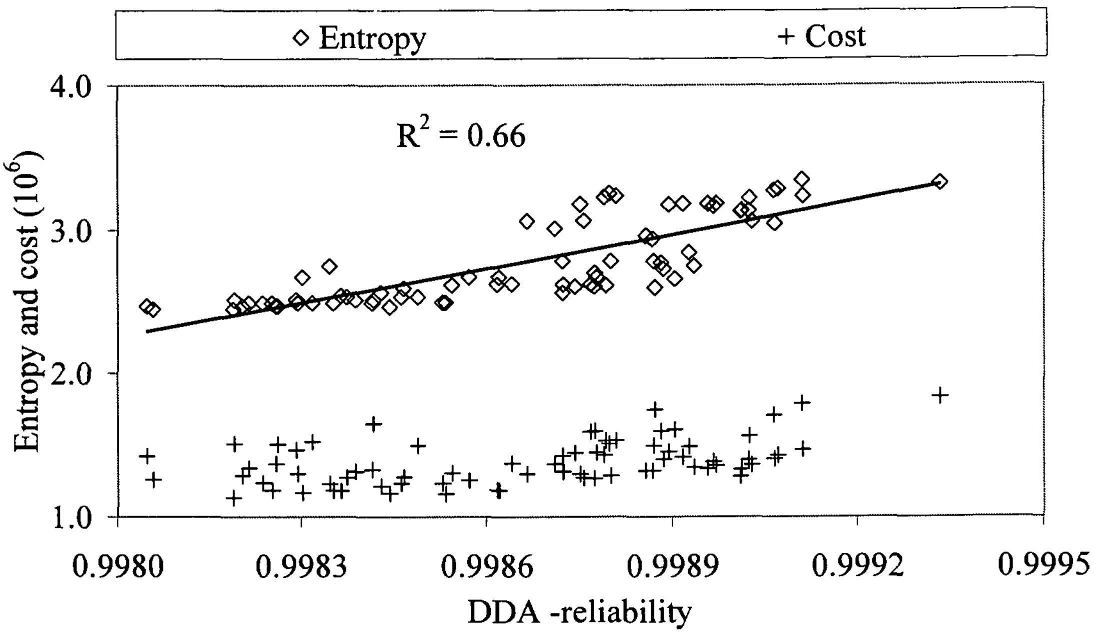


Figure A4.7. Analysis of the possible effect of flow direction on the relationship between entropy and cost against DDA-reliability for the 2-, 3- and 6-loop designs.

Table A4.1. Pipe diameters for the two-loop designs in the flow direction analysis.

Pipe	Pipe diameters (mm) for the design number indicated									
	1	2	3	4	5	6	7	8	9	10
1 - 2	373	383	407	362	377	439	354	332	317	430
2 - 3	270	302	320	222	328	343	165	367	100	170
1 - 4	274	258	222	300	284	170	319	304	336	180
2 - 5	262	239	257	287	178	280	308	100	343	393
3 - 6	242	278	296	186	305	321	105	345	208	116
4 - 7	250	233	192	277	261	124	298	282	315	140
5 - 8	234	209	229	261	130	255	284	208	318	375
6 - 9	218	258	277	146	287	304	174	327	258	147
7 - 10	204	183	113	238	220	218	261	245	282	163
8 - 11	176	134	170	212	214	208	240	291	277	348
9 - 12	124	199	226	208	239	260	268	284	318	242
10 - 11	166	134	171	209	189	259	237	220	260	202
11 - 12	166	140	185	247	203	228	294	255	340	268

Table A4.1. (Continued).

Pipe	Pipe diameters (mm) for the design number indicated									
	11	12	13	14	15	16	17	18	19	20
1 - 2	270	406	437	426	423	308	410	338	402	398
2 - 3	179	171	394	228	272	236	362	276	231	273
1 - 4	384	226	181	181	181	353	227	320	226	227
2 - 5	195	367	186	363	332	193	188	193	332	295
3 - 6	123	117	375	194	245	202	342	249	197	246
4 - 7	367	197	141	141	141	334	198	299	197	199
5 - 8	151	347	144	344	311	150	147	150	311	271
6 - 9	148	149	362	160	222	167	327	226	164	223
7 - 10	341	124	187	166	173	304	127	266	125	128
8 - 11	142	317	160	313	275	147	175	158	275	228
9 - 12	246	247	330	170	138	174	289	140	174	139
10 - 11	326	110	233	207	216	287	135	244	114	122
11 - 12	272	273	310	210	118	216	265	129	217	131

Table A4.1. (Continued).

Pipe	Pipe diameters (mm) for the design number indicated								
	21	22	23	24	25	26	27	28	29
1 - 2	272	420	307	445	227	439	238	416	186
2 - 3	236	398	277	425	181	106	113	109	118
1 - 4	385	226	358	180	413	178	412	225	438
2 - 5	114	109	113	108	116	420	197	394	117
3 - 6	201	380	249	408	123	180	181	177	179
4 - 7	369	197	339	140	397	137	396	197	423
5 - 8	154	175	161	168	151	404	151	376	150
6 - 9	165	367	226	396	148	231	233	231	231
7 - 10	343	125	311	181	374	163	372	123	401
8 - 11	232	262	242	250	225	379	138	349	220
9 - 12	171	336	141	368	243	289	291	290	289
10 - 11	328	124	294	225	361	199	358	106	389
11 - 12	211	317	118	353	268	310	312	311	310

Table A4.2. Pipe diameters for the three-loop designs in the flow direction analysis.

Pipe	Pipe diameters (mm) for the design number indicated									
	1	2	3	4	5	6	7	8	9	10
1 - 2	253	270	374	283	387	301	277	239	410	268
1 - 4	388	385	302	380	281	371	356	397	252	387
2 - 3	234	252	362	266	375	285	260	219	398	252
3 - 6	197	218	341	235	356	257	229	179	380	220
4 - 5	230	247	249	332	234	332	252	212	180	170
4 - 7	310	292	146	167	123	139	244	331	152	339
5 - 8	194	214	216	310	198	310	222	174	121	114
6 - 9	158	185	325	207	342	232	201	130	366	189
7 - 10	221	162	252	211	176	136	100	266	223	299
9 - 12	176	246	173	220	242	272	325	140	322	208
10 - 11	194	114	209	166	118	155	207	244	262	280
7 - 8	187	212	266	204	223	162	296	161	203	100
8 - 9	211	248	239	196	208	232	325	169	104	196
11 - 12	109	204	119	170	212	241	291	199	301	243

Table A4.2. (Continued).

Pipe	Pipe diameters (mm) for the design number indicated										
	11	12	13	14	15	16	17	18	19	20	21
1 - 2	322	463	195	412	141	199	308	330	336	198	356
1 - 4	346	130	420	247	447	414	348	322	323	418	295
2 - 3	307	454	171	400	100	175	294	316	323	175	344
3 - 6	281	439	100	382	210	109	268	293	300	108	323
4 - 5	344	197	228	113	244	253	178	100	180	273	100
4 - 7	100	192	351	200	374	327	290	312	258	317	283
5 - 8	321	251	193	195	212	222	127	179	129	244	176
6 - 9	259	428	181	368	254	144	248	274	282	133	307
7 - 10	203	149	282	187	299	222	236	239	191	169	190
9 - 12	323	281	175	309	204	149	130	119	207	223	203
10 - 11	253	210	261	239	279	195	209	213	152	127	149
7 - 8	103	114	178	239	197	216	115	162	119	243	170
8 - 9	269	311	206	144	239	261	139	196	126	299	187
11 - 12	300	258	220	286	241	116	139	142	158	180	158

Table A4.3. Pipe diameters for the six-loop designs in the flow direction analysis.

Pipe	Pipe diameters (mm) for the design number indicated									
	1	2	3	4	5	6	7	8	9	10
1 - 2	302	172	273	257	275	282	296	414	342	344
1 - 4	361	435	374	395	363	366	370	232	332	334
2 - 3	192	252	100	195	152	191	179	197	289	195
2 - 5	228	214	262	153	221	199	227	363	172	281
4 - 7	275	290	278	358	289	303	290	345	270	150
3 - 6	138	209	144	143	100	138	121	151	259	143
5 - 8	239	190	244	180	298	176	251	161	291	334
6 - 9	182	208	159	239	135	182	140	161	105	190
7 - 10	169	175	171	184	174	248	189	198	169	233
8 - 11	184	169	184	137	196	231	236	177	202	133
9 - 12	162	167	153	176	131	233	173	153	151	147
4 - 5	226	324	245	146	211	195	222	284	177	284
5 - 6	175	127	234	235	100	176	146	135	212	175
7 - 8	179	200	182	287	198	128	183	255	171	244
8 - 9	178	157	188	114	257	279	125	193	221	145
10 - 11	119	132	123	147	127	225	148	155	116	193
11 - 12	135	131	140	115	156	100	225	152	151	154

Table A4.3. (Continued).

Pipe	Pipe diameters (mm) for the design number indicated									
	11	12	13	14	15	16	17	18	19	20
1 - 2	331	308	313	310	207	226	268	303	282	337
1 - 4	346	342	357	326	414	372	342	347	363	289
2 - 3	234	194	204	205	100	100	100	181	152	184
2 - 5	231	235	232	229	198	340	278	238	228	280
4 - 7	243	236	263	208	372	296	212	242	284	100
3 - 6	194	141	155	159	160	244	162	124	100	129
5 - 8	155	254	231	223	210	368	232	274	299	388
6 - 9	294	187	215	229	227	246	198	144	126	157
7 - 10	164	100	143	100	185	175	100	100	148	176
8 - 11	238	257	112	201	137	194	199	319	127	325
9 - 12	100	169	238	388	168	116	364	163	218	147
4 - 5	236	238	232	237	166	217	258	239	219	335
5 - 6	259	178	197	209	293	158	271	148	100	156
7 - 8	123	238	180	347	305	209	339	259	209	133
8 - 9	226	285	217	373	120	338	371	127	299	114
10 - 11	108	167	100	287	149	128	271	180	100	226
11 - 12	181	114	189	336	119	155	312	213	165	198

Table A4.3. (Continued).

Pipe	Pipe diameters (mm) for the design number indicated				
	21	22	23	24	25
1 - 2	306	228	273	304	308
1 - 4	297	345	358	294	344
2 - 3	100	100	100	100	223
2 - 5	317	366	275	314	207
4 - 7	100	253	243	100	259
3 - 6	165	256	160	163	182
5 - 8	404	401	279	389	106
6 - 9	126	258	114	178	273
7 - 10	188	100	100	189	259
8 - 11	338	295	325	299	113
9 - 12	153	130	169	158	156
4 - 5	363	230	256	365	213
5 - 6	231	165	221	259	242
7 - 8	140	297	265	141	100
8 - 9	120	356	133	169	158
10 - 11	239	196	184	241	230
11 - 12	200	133	217	100	112

APPENDIX A5 – SUMMARY OF THE DDA-BASED RESULTS FROM THE INVESTIGATION OF THE POSSIBLE INFLUENCE OF COST FUNCTION ON THE ENTROPY-RELIABILITY RELATIONSHIP

Figures A5.1 and A5.2 show the DDA analysis results of the possible influence of cost function on the relationship between entropy and reliability of water distribution network. The results seem to indicate that the entropy-reliability relationship remains strong despite the difference in the cost function used. Hence the influence of the cost function is negligible. Figure A5.2 also shows that more uniform pipe diameters leads to higher value of network reliability, which is as expected.

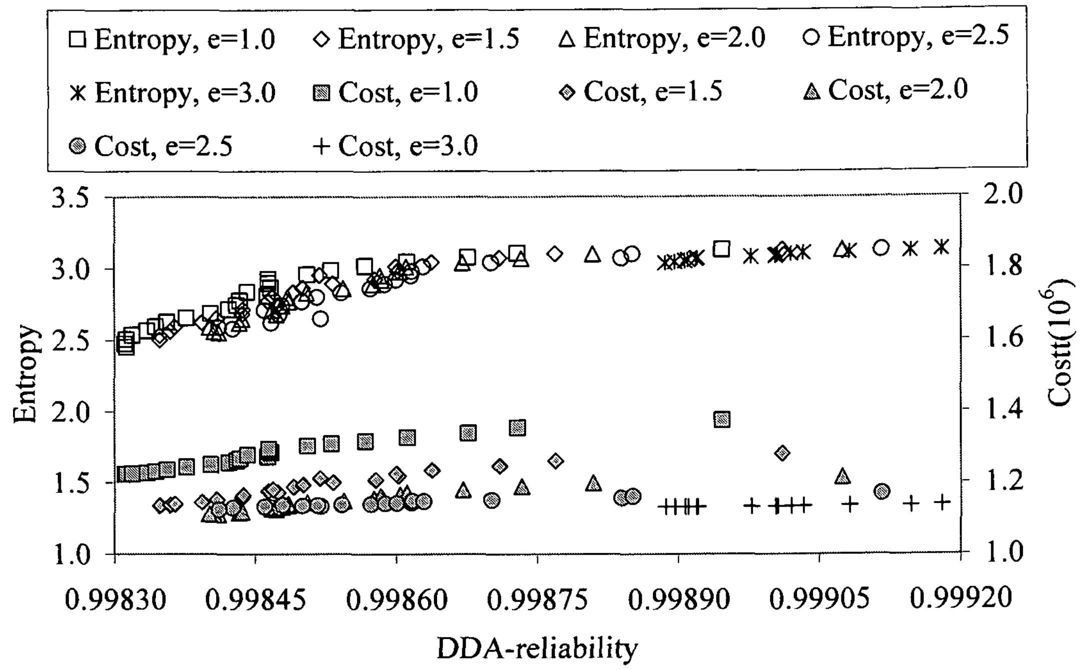


Figure A5.1. Influence of cost function on the correlation between entropy and DDA-reliability.

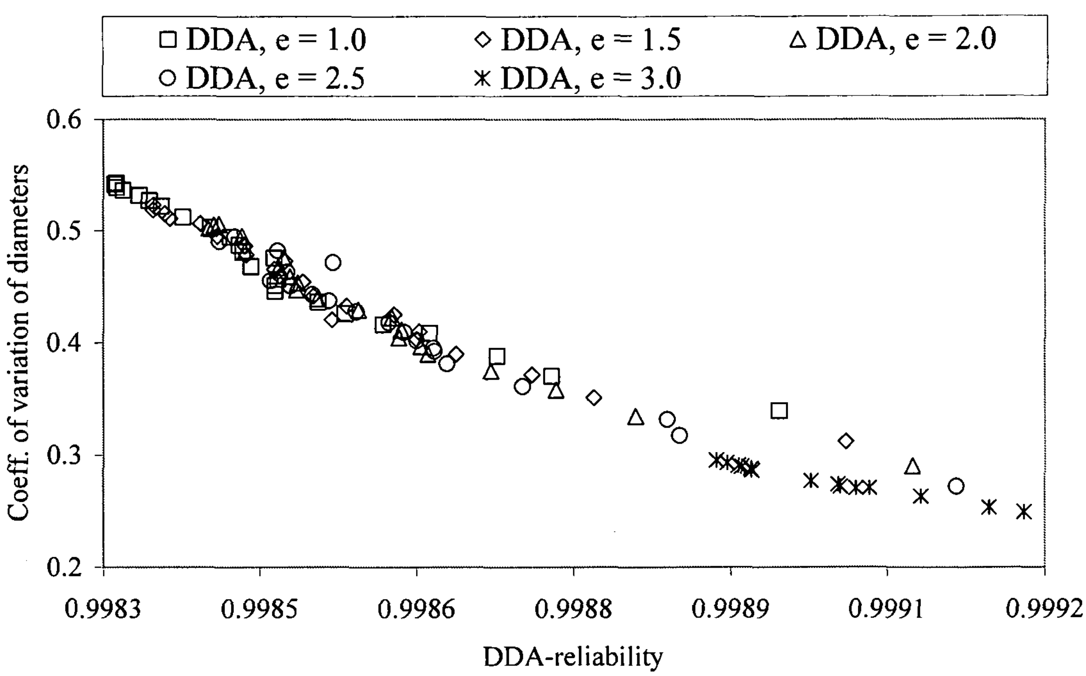


Figure A5.2. Relationship between coefficient of variation of diameters and DDA-reliability for designs generated using different cost functions.

APPENDIX A6 – SUMMARY OF THE DDA-BASED RESULTS FROM THE STUDY OF THE POSSIBLE INFLUENCE OF MODELING ERRORS ON THE ENTROPY-RELIABILITY RELATIONSHIP

To investigate the possible influence of modeling errors on the entropy-reliability relationship using the DDA method, the designs were first validated by checking the surplus or deficit in head at the critical node(s). Figure A6.1 shows the distribution of the surplus heads, which seems to follow the normal distribution with the heads at the critical nodes for most of the designs being equal or very close to zero. This suggests that the accuracy of the resulting designs is acceptable. To check the effect of the small differences in head at the critical nodes on the reliability values, the performance of designs with an excess or shortfall in capacity was compared to slightly adjusted designs, which satisfied the demands exactly. To achieve this, the head at the source, for designs with a surplus or deficit in head at the critical node was artificially altered so that, at the critical node, the head was precisely equal to the desired service head of 30 m. These designs were then re-analysed and their reliability values obtained. The results of this analysis are shown in Figure A6.2 in which the reliability values before and after the head modification for all the designs in this study are plotted against each other. It shows that all the designs have virtually identical pairs of reliability values, which indicates that the small discrepancies in head are insignificant. Figure A6.3 shows that there is no correlation between the DDA-reliability and surplus head, which further suggests that the influence of the slight modeling errors on the entropy-reliability relationship is insignificant.

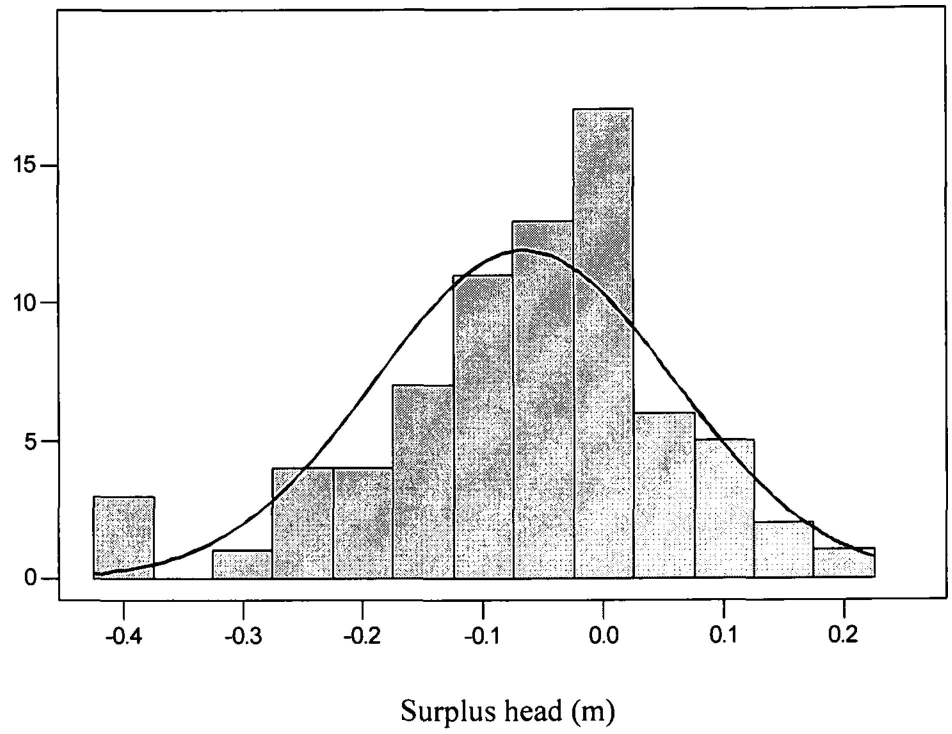


Figure A6.1. Distribution of surplus heads.

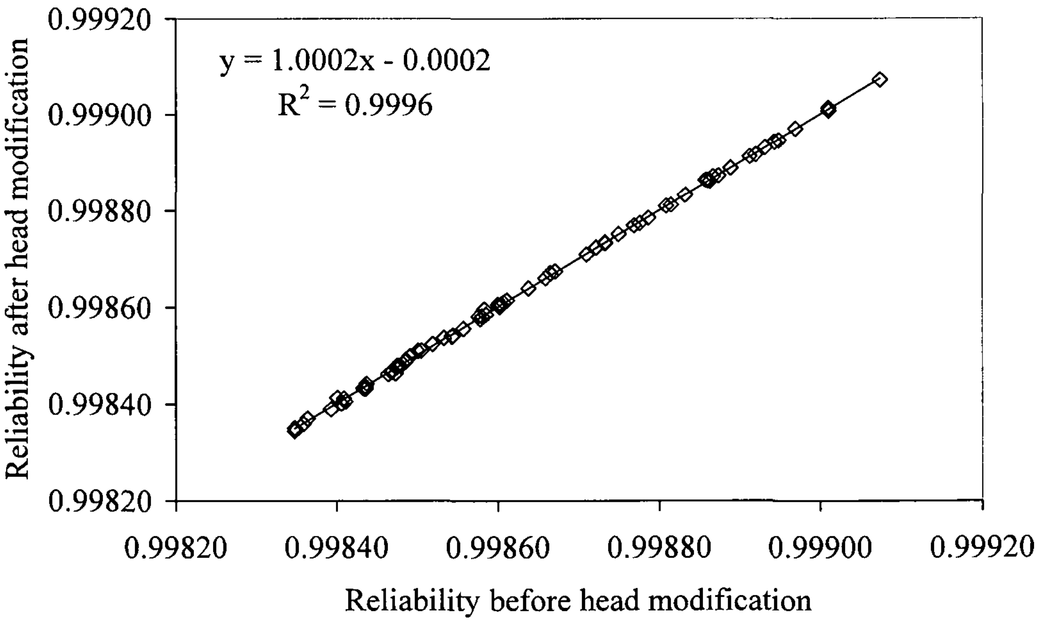


Figure A6.2. DDA-reliability verification.

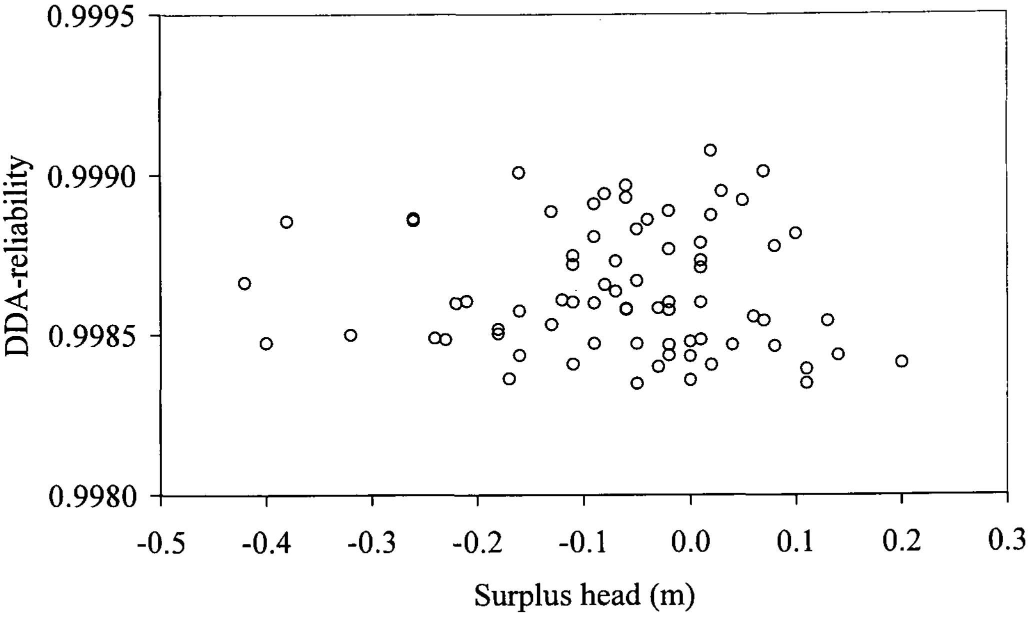


Figure A6.3. DDA-reliability against surplus head.

APPENDIX A7 – PERFORMANCE OF DESIGNS WITH EQUAL MAXIMUM ENTROPY VALUES – SUMMARY OF THE DDA-BASED STUDY

Figure A7.1 shows the coefficient of variation of the DDA-reliability (CVR) for the designs within the Equal Maximum Entropy Groups (EMEGs) compared to four possible comparators (Tanyimboh and Sheahan, 2002) – see Chapter 6. A total of 137 designs were analysed - 65 designs were taken from Tanyimboh and Sheahan (2002) and the rest were generated in the present study and have been previously used in the flow directions analysis in Chapter 6. There are 29 EMEGs in total as shown in Figure A7.1. The CVR values of the designs within the EMEGs seem to be lower than other designs. The weighted average of the CVRs of all the EMEGs is also much lower than those of the four comparators. This indicates that the similarity in the reliability level of the designs with equal maximum entropy value is very high.

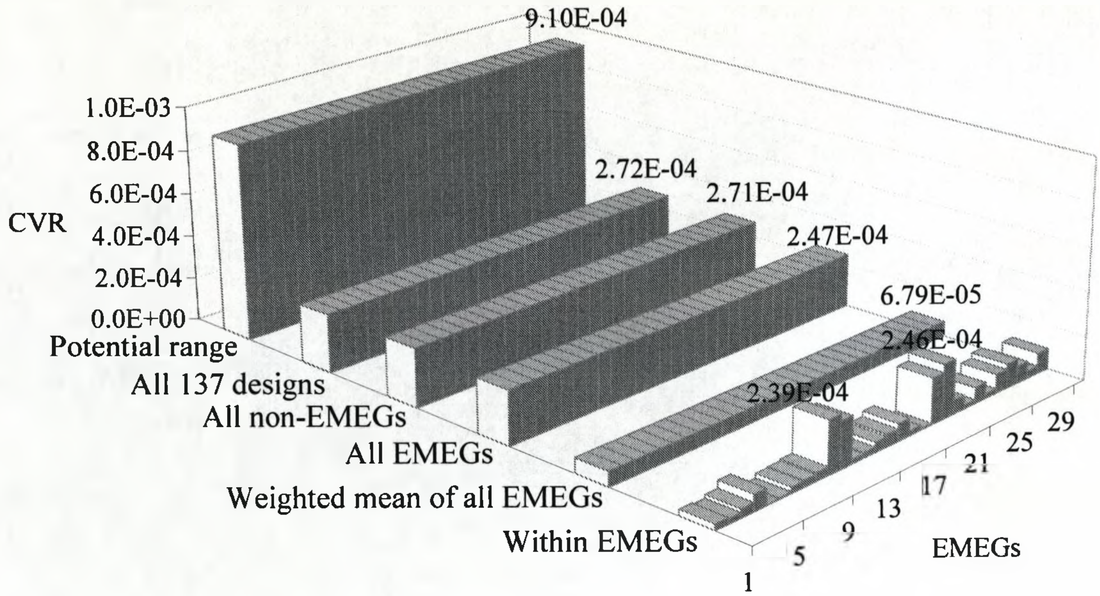


Figure A7.1. Coefficients of variation of DDA-reliabilities.

APPENDIX A8 – SUMMARY OF THE DDA RESULTS OF THE HYDRAULIC PREDICTABILITY ANALYSIS OF ENTROPY CONSTRAINED DESIGNS OF WATER DISTRIBUTION NETWORKS

In the DDA analysis, the *notional* usable source head was used to measure the network performance (Tanyimboh, 1993), i.e. the head at the source that was required to satisfy the minimum head requirements at all demand nodes, especially the critical ones. It is termed the notional source head since the value is sometimes extremely large and not feasible in practice. Another performance measure used in DDA analysis was the total dissipated energy by the pipe network given in Equation (5.15).

In pipe failure analysis based on the DDA method, the critical pipe is the pipe which, when fails, will give the highest useable source head or cause the network to dissipate the highest total energy and the corresponding critical node is the node which has the lowest pressure head in the case of the failure of the critical pipe (Tanyimboh, 1993). In fire fighting situation, the critical node is the node at which the occurance of a fire fighting demand would lead to the highest usable source head or total dissipated energy.

For the two-loop network in Figure 7.1a, the results shown in Tables A8.1 to A8.3 are more or less as expected. Pipe 1-3 is the critical pipe in the network since it is connected to the source and it lies in the only path supplying the largest demand in the network at node 5. At the failure of the critical pipe, it is quite obvious that node 5 is the critical node due to its large demand. In fire fighting situation, node 6 is the critical node in the network. Under normal condition, node 6 is the terminal node in the network; hence it has the lowest residual head. Large increase in demand therefore leads to a huge reduction in the pressure head causing the network to suffer greatly. For the network in Figure 7.1b, the critical pipe in the network can be either pipe 1-2 or 1-4 due to the network symmetry. The results in Table A8.4 support this contention. Also, when the critical pipe in the network fails, the critical node is located at node 9, which is as expected. In a fire-fighting situation, it seems that the critical node should be at node 9, i.e. the terminal node. However, for a replacement fire flow, the increase from the normal demand of 62.5 l/s to the fire-fighting demand of 250 l/s at node 9 is less compared to the increase from the normal to the fire-fighting demand at any other nodes, which is from 20.8 l/s to 250 l/s. This seems to be the reason that the results of the DDA analysis in Table A8.5 show that nodes 3 and 7 are more critical than node 9 (Tanyimboh, 1993). The selection of nodes 3 and 7 as the critical nodes is due to the fact that, under normal operating condition, both nodes are the most downstream nodes supplied only by a single path. Therefore, the large increase in demand at those nodes increases the head losses in the network significantly. When the fire demand is superimposed onto the design demand (Table A8.6), the results yielded by the DDA analysis are in accordance with the expectation that the critical node is located at node 9. The results in all the Tables show that the maximum entropy designs are in general more predictable that other designs. Figures A5.1 to A5.6 also show the improvement in the network performance of the two networks analysed in this study as their entropy values increase.

Table A8.1. Critical links and nodes for single-link failures for the two-loop network.

Network entropy (1)	Assessment criteria: Total usable source head – Critical link (2)	Total usable source head – Next critical link (3)	Total usable source head – Critical node (4)	Total dissipated energy – Critical link (5)	Total dissipated energy – Next critical link (6)
1.578	1 - 3	3 - 5	6	1 - 3	3 - 5
1.600	1 - 3	3 - 5	6	1 - 3	3 - 5
1.700	3 - 5	1 - 3	5	3 - 5	1 - 3
1.800	1 - 3	3 - 5	6	1 - 3	3 - 5
1.900	1 - 3	3 - 5	5	1 - 3	3 - 5
1.915	1 - 3	3 - 5	5	1 - 3	3 - 5

From Tanyimboh (1993).

Table A8.2. Critical nodes for nodal fire-fighting demands replacing design demands for the two-loop network.

Network entropy (1)	Assessment criteria: Total usable source head – Critical node (2)	Total usable source head – Next critical node (3)	Total dissipated energy – Critical node (4)	Total dissipated energy – Next critical node (5)
1.578	2	6	2	6
1.600	2	6	2	6
1.700	2	6	2	6
1.800	6	2	6	2
1.900	6	5	6	2
1.915	6	4	6	2

From Tanyimboh (1993).

Table A8.3. Critical nodes for superimposed nodal fire-fighting demands for the two-loop network.

Network entropy (1)	Assessment criteria			
	Total usable source head		Total dissipated energy	
	Critical node (2)	Next critical node (3)	Critical node (4)	Next critical node (5)
1.578	2	4	2	4
1.600	2	4	2	4
1.700	4	6	4	6
1.800	6	4	6	4
1.900	6	5	6	5
1.915	6	5	6	5

Table A8.4. Critical links and nodes for single-link failures for the four-loop network.

Network entropy (1)	Assessment criteria				
	Total usable source head			Total dissipated energy	
	Critical link (2)	Next critical link (3)	Critical node (4)	Critical link (5)	Next critical link (6)
2.170	1 - 4	8 - 9	7	1 - 4	4 - 5
2.500	6 - 9	5 - 6	9	6 - 9	5 - 6
2.750	1 - 2, 1 - 4	2 - 3, 4 - 7	9	1 - 2, 1 - 4	2 - 5, 4 - 5
2.775	1 - 2, 1 - 4	2 - 3, 4 - 7	9	1 - 2, 1 - 4	2 - 5, 4 - 5
2.800	1 - 2, 1 - 4	6 - 9, 8 - 9	9	1 - 2, 1 - 4	2 - 3, 4 - 7

From Tanyimboh (1993).

Table A8.5. Critical nodes for nodal fire-fighting demands replacing design demands for the four-loop network.

Network entropy (1)	Assessment criteria			
	Total usable source head		Total dissipated energy	
	Critical node (2)	Next critical node (3)	Critical node (4)	Next critical node (5)
2.170	3	7	3	7
2.500	7	3	7	3
2.750	3, 7	9	3, 7	6, 8
2.775	3, 7	9	3, 7	6, 8
2.800	3, 7	9	3, 7	6, 8

From Tanyimboh (1993).

Table A8.6. Critical nodes for superimposed nodal fire-fighting demands for the four-loop network.

Network entropy (1)	Assessment criteria			
	Total usable source head		Total dissipated energy	
	Critical node (2)	Next critical node (3)	Critical node (4)	Next critical node (5)
2.170	3	7	3	7
2.500	7	3	7	3
2.750	3, 7	9	3, 7	9
2.775	3, 7	9	3, 7	9
2.800	9	3, 7	9	3, 7

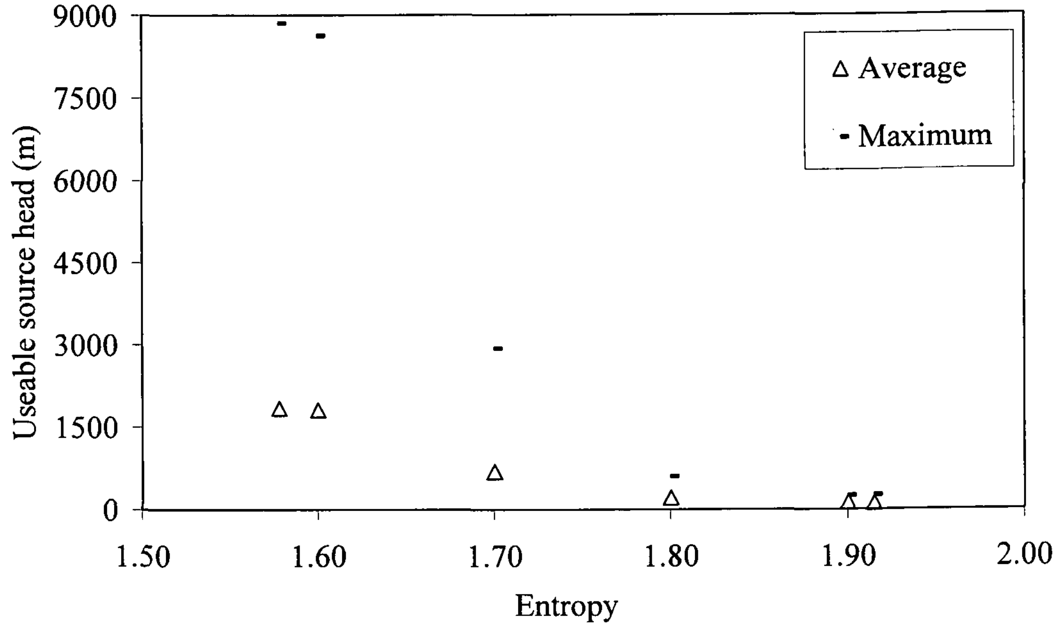


Figure A5.1. Total usable source head vs. entropy from the link failure analysis of the two-loop network.

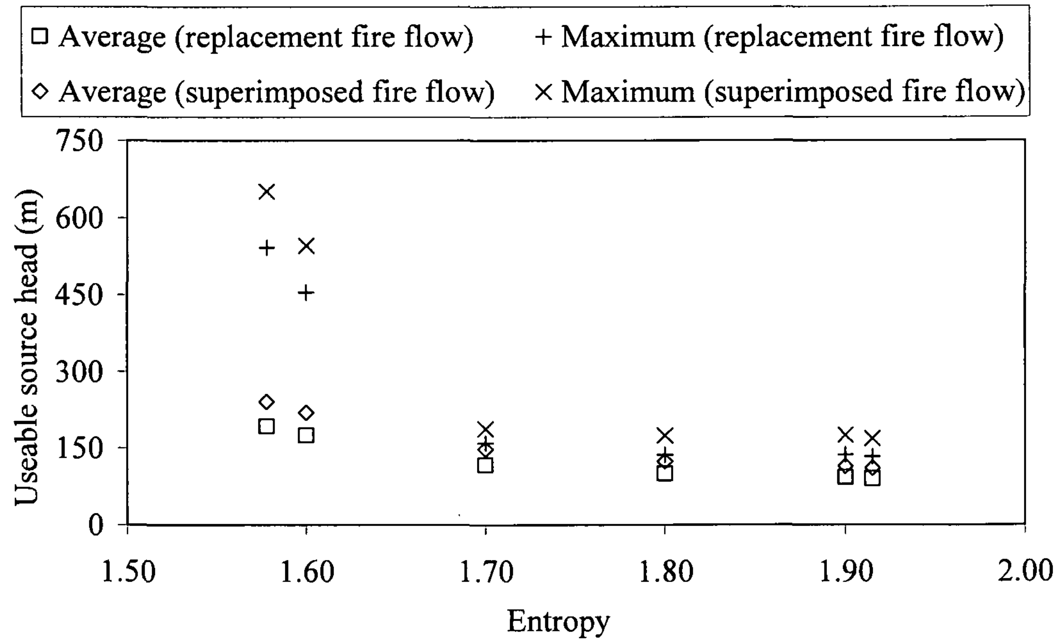


Figure A5.2. Total usable source head vs. entropy from the analysis of the two-loop network under fire-fighting conditions.

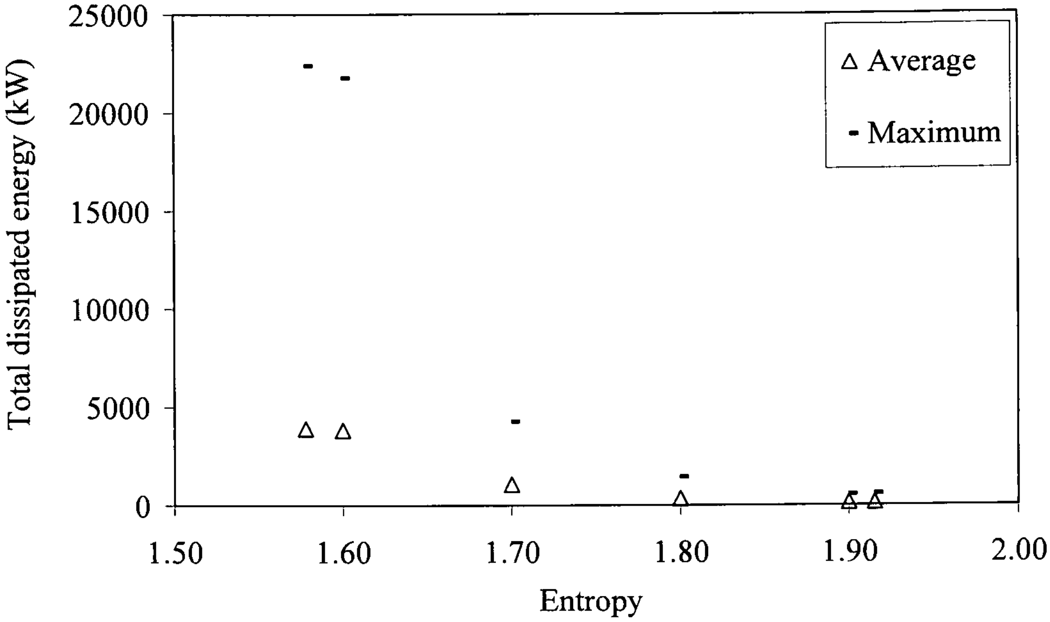


Figure A5.3. Total dissipated energy vs. entropy from the link failure analysis of the two-loop network.

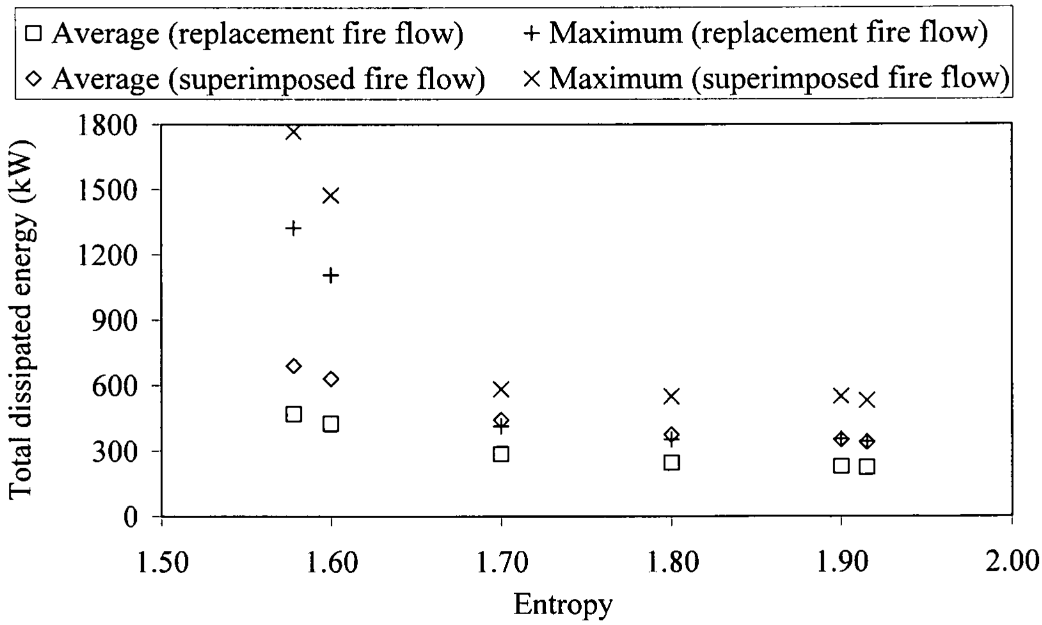


Figure A5.4. Total dissipated energy vs. entropy from the analysis of the two-loop network under fire-fighting conditions.

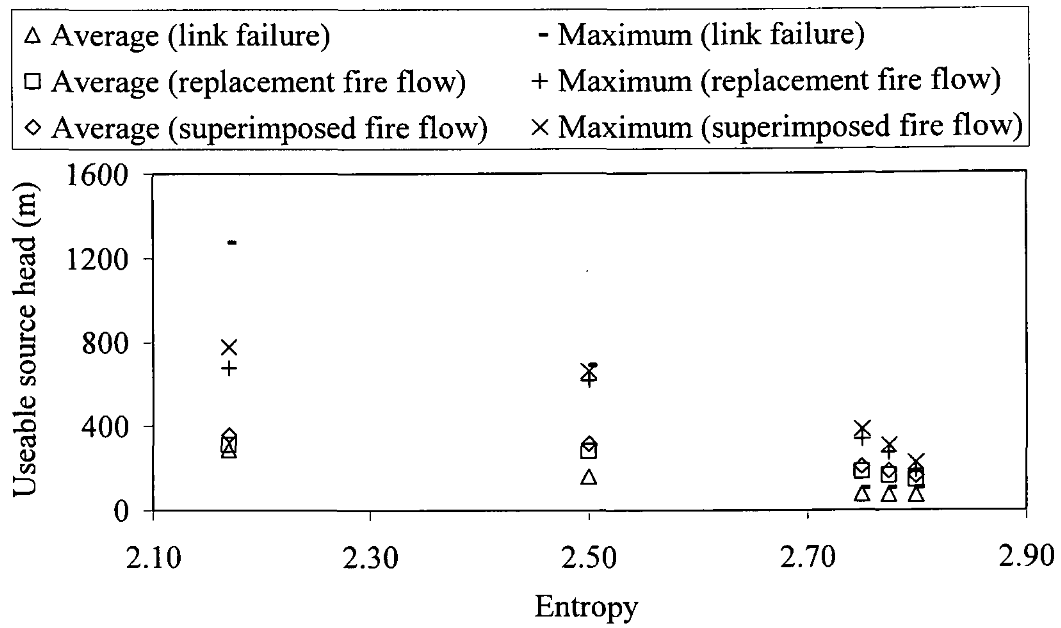


Figure A5.5. Total usable source head vs. entropy from the analysis of the four-loop network under critical operating conditions.

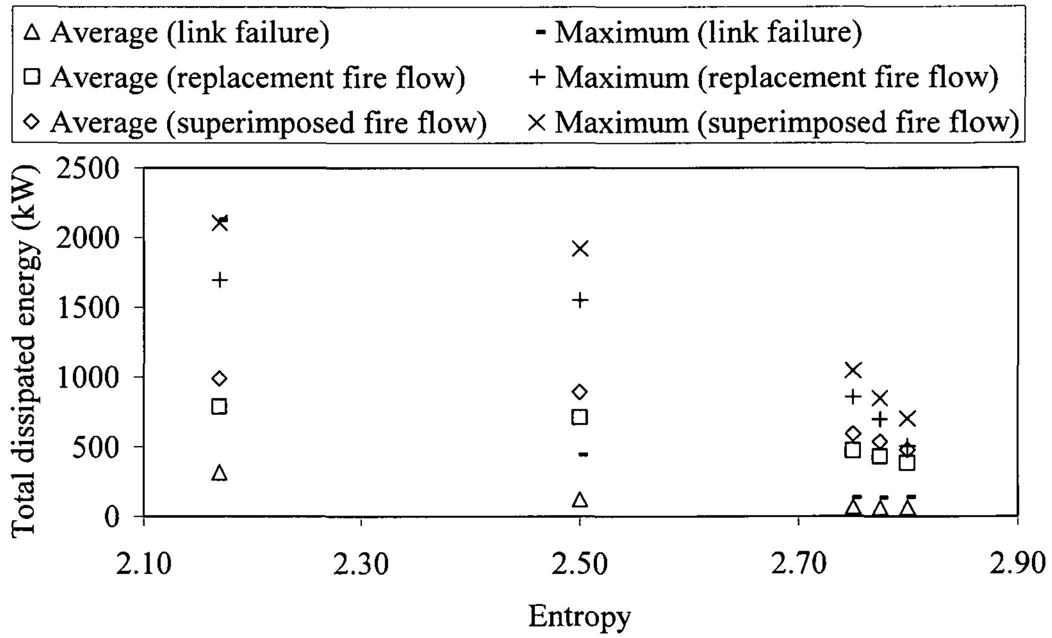


Figure A5.6. Total dissipated energy vs. entropy from the analysis of the four-loop network under critical operating conditions.

APPENDIX A9 – SUMMARY OF THE DDA-BASED RESULTS OF THE MAXIMUM ENTROPY APPROACH TO THE LAYOUT OPTIMIZATION OF WATER DISTRIBUTION NETWORKS

Figures A91 and A9.2 show that the maximum entropy approach manage to identify the designs close to the true Pareto optimal set. Although the message depicted by the DDA method is not as strong as that by the HDA, the potential of this method in layout optimization study is noticeable, which was the basis of the investigations carried out in Chapter 7.

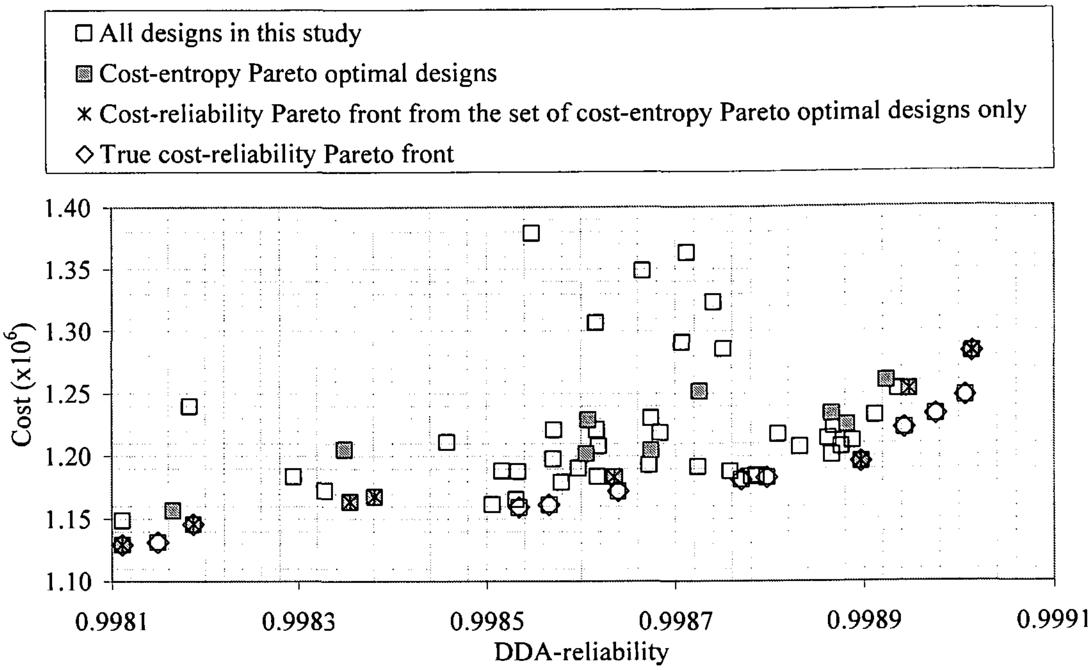


Figure A9.1. Plots of cost against DDA-reliability showing the cost-entropy and the cost-reliability Pareto optimal layouts (Tanyimboh and Sheahan, 2002).

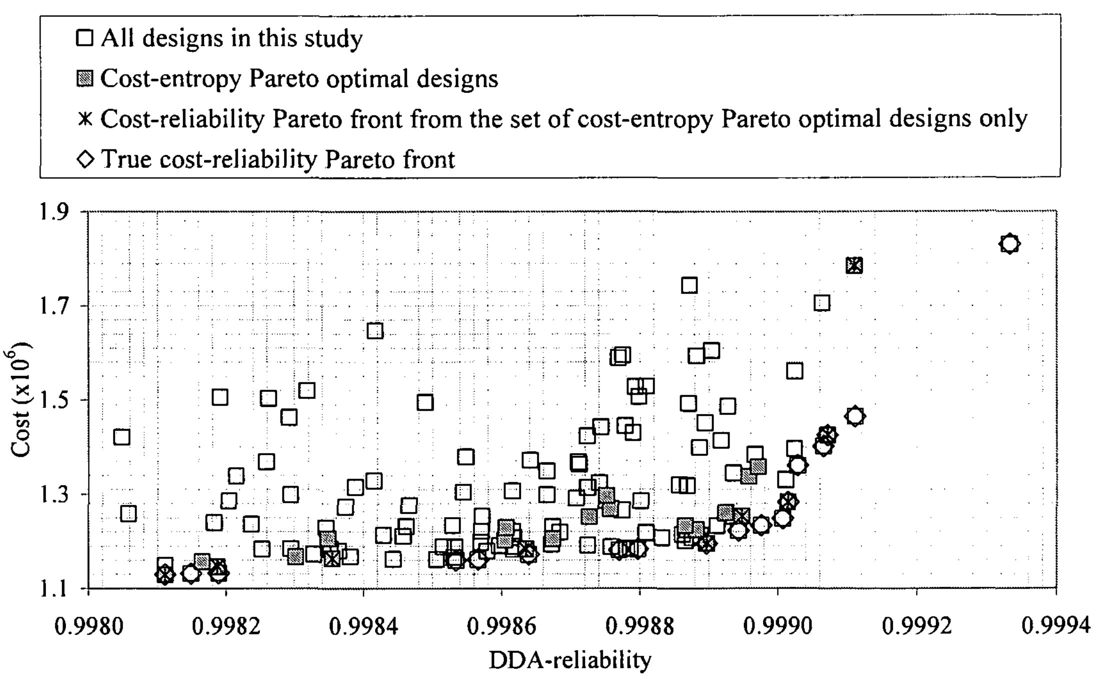


Figure A9.2. Plots of cost against DDA-reliability showing the cost-entropy and the cost-reliability Pareto optimal layouts considering alternative flow directions.

APPENDIX B. PUBLICATIONS

- "The Suitability of Information-theoretic Entropy as A Surrogate Performance Measure for Water Distribution Systems". In *Proceedings of 3rd International Conference on Decision Making in Urban and Civil Engineering*, London, 2002.
- "Water System Entropy: A Study of Redundancy as a Possible Lurking Variable". In *Proceedings of 9th International Conference on Civil and Structural Engineering Computing*, Egmond Aan Zee, the Netherlands, 2003.
- "Hydraulically Predictable Water Distribution Networks". In *Proceedings of 4th International Conference on Engineering Computational Technology*, Lisbon, Portugal, 2004.
- "Modelling Errors, Entropy and the Hydraulic Reliability of Water Distribution Systems". *Advances in Engineering Software* (2005), 36(11-12), 780-788.
- "Joint Layout, Pipe Size and Hydraulic Reliability Optimization of Water Distribution Systems". Submitted for possible publication in *ASCE Journal of Water Resources, Planning and Management*, 2006.

THE SUITABILITY OF INFORMATION-THEORETIC ENTROPY AS A SURROGATE PERFORMANCE MEASURE FOR WATER DISTRIBUTION SYSTEMS

Tiku Tanyimboh, Lecturer, Department of Civil Engineering, University of Liverpool, Brownlow Street, Liverpool L69 3GQ, England. Email: ttanyi@liv.ac.uk
Yohan Setiadi, PhD student, Department of Civil Engineering, University of Liverpool, Brownlow Street, Liverpool L69 3GQ, England
Athina Mavrokoukoulaki, Masters student, Department of Civil Engineering, University of Liverpool, Brownlow Street, Liverpool L69 3GQ, England
Catherine Storer, Masters student, Department of Civil Engineering, University of Liverpool, Brownlow Street, Liverpool L69 3GQ, England

ABSTRACT

Reliability-based optimal design of water distribution systems is computationally a very difficult problem to solve. However, it has been known for some time that entropy-based design optimisation approaches offer several computational advantages. There is a body of evidence which suggests that entropy is a possible general reliability surrogate. The aim of this study was to assess the strength of the relationship between entropy and reliability under more general and tightly defined conditions including the influence of pipe flow directions. Minimum cost designs were obtained using non-linear programming for a progression of entropy values and the designs assessed by calculating their reliabilities. The results would appear to reinforce the idea that the association between entropy and reliability is strong.

Keywords: Information-theoretic entropy, reliability, water distribution systems, statistical correlation, design optimisation

INTRODUCTION

Urban water distribution networks are designed to supply water every day and the effects of any interruptions to the supply should be minimised. Reliability based performance assessment provides a means of checking that the network will have sufficient capacity to at least satisfy the minimum standards following hydraulic and/or mechanical failure events or other instances involving the unavailability of components. Obviously, the need to safeguard the reliability of supply has to be balanced against the advantages of using available funds economically.

Thus optimal designs of water distribution systems are generally obtained by minimising a cost objective function whilst satisfying a range of constraints including the constitutive equations and nodal service pressure constraints. Other issues considered may include bounds on velocities and diameters. It is well known that the conventional optimum design process inherently reduces the redundancy of the system to the extent that the reliability is probably compromised (Templeman, 1982). Several techniques have been used in an attempt to reconcile the conflicting goals of reliability/redundancy and cost minimisation including the specification of minimum pipe flow rates or diameters (e.g. Alperovits and Shamir, 1977).

Unfortunately, accurate and meaningful reliability measures are difficult to calculate (Wagner et al., 1988). The difficulties associated with the determination of reliability have led researchers to investigate the possibility of using entropy as a surrogate measure for reliability (Awumah et al., 1991; Coelho, 1997). A recent approach that would appear to be very promising consists of the addition of an entropy constraint to the basic constraints set, which therefore produces entropy-constrained minimum-cost designs (Tanyimboh, 1993). Furthermore, it has recently been suggested that minimum-cost maximum-entropy designs of water distribution systems can be used to identify good layouts in that designs based on these layouts have the potential to achieve a reasonable compromise between reliability and cost (Tanyimboh and Sheahan, in press). Recent studies have addressed issues related to potential difficulties arising from the absence of a one-to-one mapping between entropy and reliability due to the invariance of the entropy function (Tanyimboh and Sheahan, in press). The reason for this is that, for a given probability scheme, the entropy function can distinguish between combinations, but not permutations, of the probabilities (Tanyimboh and Sheahan, in press).

Given pipe flow rates the entropy of a water distribution system can be easily calculated (Tanyimboh and Templeman, 1993a; Yassin-Kassab et al., 1998). There is a body of evidence which suggests that, for water distribution networks, the association between entropy and reliability is strong (Tanyimboh and Templeman, 2000). It should, however, be borne in mind that for a given network configuration the entropy value along with the reliability and cost depend on the set of pipe flow directions assumed for the design of the network. Due to the multiplicity of feasible sets of flow directions associated with any non-dendritic layout, the potential influence of flow directions on the relationship between entropy and reliability cannot be ignored. All previous studies involving entropy have, however, used a single set of pipe flow directions based implicitly on the shortest path concept. Secondly, much of the work on the reliability-related properties of entropy-constrained designs has relied on a range of intuitively sensible considerations (Tanyimboh, 1993; Tanyimboh and Templeman, 1993b) to demonstrate that reliability generally increases as entropy increases. However, the influence of the entropy level on the reliability of designs derived from a single layout has not been investigated using quantified reliability measures.

The aim of this paper is to demonstrate that, for a specified layout of a water distribution network, the correlation between entropy and reliability would appear to be extremely strong. The effects of the choice of flow directions have also been investigated and the results would appear to suggest that the correlation seems fairly strong irrespective of the choice of flow directions.

INFORMATIONAL ENTROPY CALCULATION

Shannon's entropy function (Shannon, 1948), which is a measure of the amount of uncertainty in a finite probability distribution, is

$$S/K = -\sum_{i=1}^{I} p_i \ln p_i \tag{1}$$

in which S is the entropy; K is an arbitrary positive constant often taken as 1; p_i is the probability associated with the ith outcome, i = 1, 2, 3, ..., I; I represents the number of outcomes. For a single probability space the normality condition is satisfied automatically, i.e.

$$\sum_{i=1}^{I} p_i = 1 \tag{2}$$

The entropy function for water distribution networks is (Tanyimboh, 1993)

$$\frac{S}{K} = -\sum_{j \notin IN} (Q_j/T)\ln(Q_j/T) - \frac{1}{T}\sum_{j=1}^{J} T_j \left((Q_j/T_j)\ln(Q_j/T_j) + \sum_{i \in N_j} (q_{ij}/T_j)\ln(q_{ij}/T_j) \right) \tag{3}$$

in which the subscript j represents all nodes including source nodes while J denotes the number of nodes and IN the set consisting of source or input nodes. Also, T is the total supply; T_j is the total flow reaching node j, including any external inflow; q_{ij} is the flow rate in pipe ij while N_j represents all nodes immediately upstream of and connected to node j; Q_j is the demand at demand nodes or supply at source nodes. The first term in the large parentheses is zero for demand nodes because, for those nodes, the external inflow is zero. Because desired nodal demands and supplies are usually specified, network entropy is basically a function of internal pipe flow rates only and can be quickly calculated from Eq. (3).

OPTIMAL DESIGN OF WATER DISTRIBUTION SYSTEMS USING ENTROPY

This can be summarised in general terms as follows.
Minimise cost:

$$Cost = \gamma \sum_{ij} L_{ij} D_{ij}^{e} \tag{4}$$

Subject to:

$$h_{ij} = \alpha L_{ij} (q_{ij} / C_{ij})^{1.852} / D_{ij}^{4.87} \qquad \forall ij \tag{5}$$

$$\sum_{ij \in l} h_{ij} = 0 \qquad \forall l \tag{6}$$

$$\sum_{ij \in p} h_{ij} = h_p \qquad \forall p \tag{7}$$

$$\sum_{ij \in t} h_{ij} \le H_s - H_{\min,t} \qquad \forall t \tag{8}$$

$$\sum_{ij \in IJ_j} q_{ij} = Q_j \qquad \forall j \tag{9}$$

$$D_{\max} \ge D_{ij} \ge D_{\min} \qquad \forall ij \tag{10}$$

$$S / K \ge S_m \tag{11}$$

In the above problem e = constant. This user-specified cost exponent is thought to have a range of 1.0 to 2.5 (Fujiwara and Khang, 1990). The cost coefficient γ, whose value depends on a range of factors, is specified by the user. The parameters C_{ij}, D_{ij}, h_{ij}, L_{ij} and q_{ij} are the roughness coefficient, diameter, headloss, length and flow rate, respectively, for pipe ij; Q_j = inflow or outflow at node j; IJ_j represents pipes incident on node j; p (p = 1, 2, ...) represents the pth path having a known value of headloss, h_p; l (l = 1, 2, ...) represents the lth loop; t (t = 1, 2, ...) represents a path from a specified source to a terminal node t, i.e. a node with no other nodes downstream of it; H_s = head at a specified source; $H_{\min,t}$ = minimum allowable head at terminal node t; K = arbitrary constant; S = entropy; S_m = desired entropy value; $D_{\min}$ = minimum allowable pipe diameter; $D_{\max}$ = maximum pipe diameter. Equations (5-7, 9) are the constitutive equations while Eq. (8) guarantees the nodal minimum service pressures. Equation (11) ensures the entropy of the network is equal to the specified value, S_m.

RELIABILITY CALCULATION

Assuming a constant demand value, the reliability of a water distribution system can be taken as

$$R = \frac{1}{T}\left(p(0)T(0) + \sum_{m=1}^{M} p(m)T(m) + \sum_{\substack{m=1 \\ \forall n \succ m}}^{M-1} p(m,n)T(m,n) + \cdots \right) + \frac{1}{2}\left(1 - p(0) - \sum_{m=1}^{M} p(m) - \sum_{\substack{m=1 \\ \forall n \succ m}}^{M-1} p(m,n) - \cdots \right) \tag{12}$$

where R is the system reliability; p(0) is the probability that no pipe is unavailable; p(m) is the probability that only pipe m is unavailable; p(m, n) is the probability that only pipes m and n are unavailable. Similarly, T(0), T(m) and T(m, n) are the respective total flows supplied with no pipes unavailable, only pipe m unavailable, and only pipes m and n unavailable. The range of the summations involving two simultaneously unavailable pipes emphasises that all permutations of any given combination of pipes represent a single operating condition. Finally, M stands for the number of pipes while T represents the total demand. A derivation of the foregoing reliability formula, which corresponds to the time-averaged value of the ratio of the flow delivered to the flow required, is contained in Tanyimboh and Sheahan (in press).

STRENGTH OF THE ENTROPY-RELIABILITY RELATIONSHIP

This aspect was studied using the network of Figure 1. The water level at the source is 100m while demand nodes have elevations of 0m and required service heads of 30m. All pipes are 1000m long with a Hazen-Williams coefficient of 130. Pipe breakage rates were calculated using a formula from Cullinane et al. (1992). The performance of the WDS with the broken pipes isolated was simulated using EPANET as described in Tanyimboh et al. (2001). Figure 2 shows that the relationship is very strong. The influence of the cost function exponent e is also shown in Figure 2. The γ and e values used were 800 and 1.5; 1600 and 2.0, respectively. For e = 1.5 the rank correlation coefficient is 0.992 and for e = 2.0 the value is 0.991. The diameters of the individual designs were compared as shown in Figure 3 using the standard deviation. This is considered reasonable for the present study because the pipes have the same length. It is clear from Figure 3 that the pipe diameters become more uniform as the entropy value increases. It can also be seen that higher values of the cost function exponent e are associated with more uniform pipe sizes.

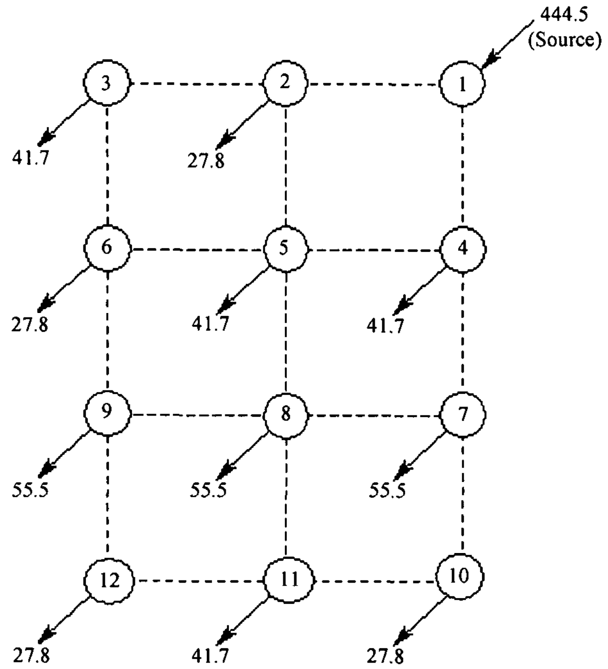


Figure 1. Network of supply and demand nodes with all demands in litres per second

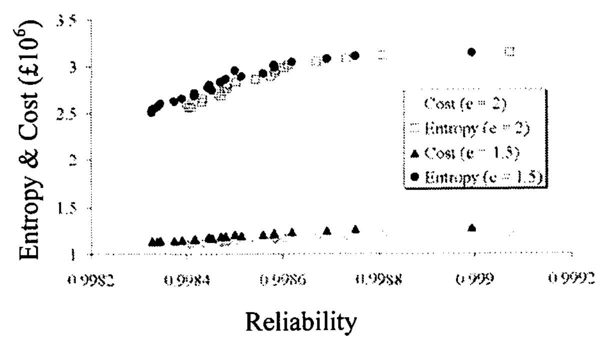


Figure 2. Plots of entropy & cost *vs* reliability

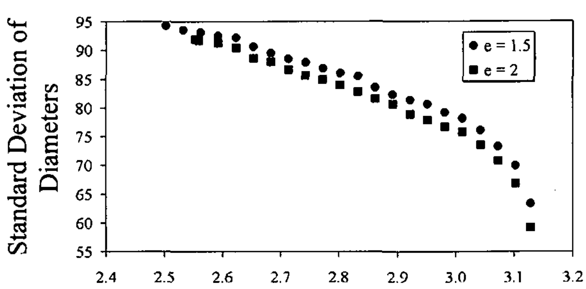


Figure 3. Plots of standard deviation of diameters *vs* entropy

INFLUENCE OF FLOW DIRECTIONS

The entropy of a WDS depends on the pipe flow rates which, in turn, obviously depend on the flow directions. To assess the influence of flow directions on the relationship between entropy and reliability, maximum entropy designs were generated for the flow directions shown in Figure 4 using a cost function exponent e of 1.5 (as in Tanyimboh and Sheahan, in press). The results were assessed by combining the designs obtained with maximum entropy designs for the full range of layouts for the network of Figure 1 from Tanyimboh and Sheahan (in press). The flow directions in Tanyimboh and Sheahan (in press) were based on the shortest path concept whereas those in Figure 4 are more arbitrary. Figure 5 shows that the new designs generally lie in the expected region of the graph. Tanyimboh and Sheahan (in press) obtained a rank correlation coefficient of 0.7. Adding the new designs (based on Figure 4)

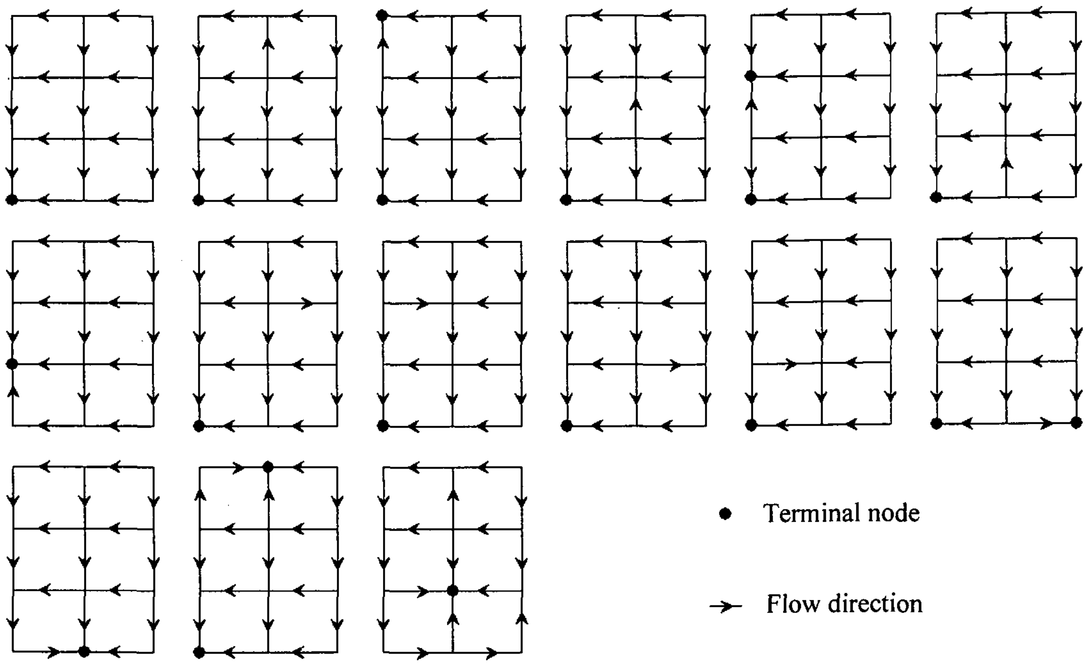


Figure 4. Diagram of design flow directions for the six-loop layout

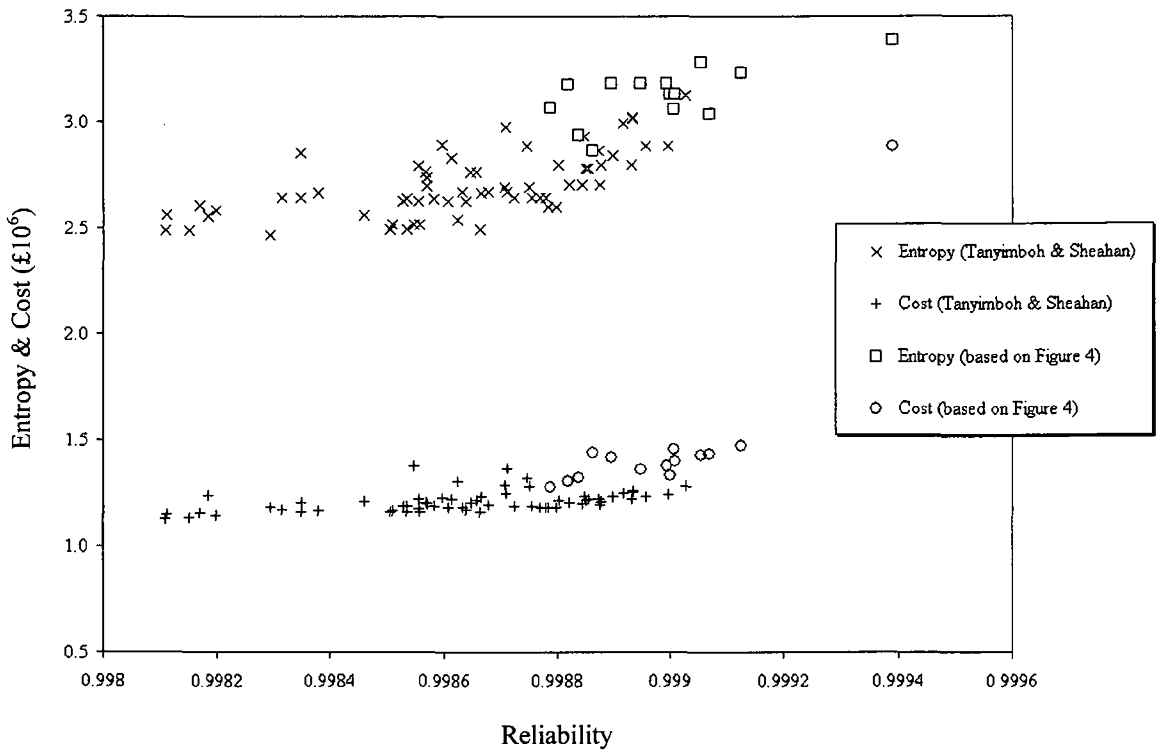


Figure 5. Plots of entropy & cost *vs* reliability to assess the influence of flow directions

leads to a correlation coefficient of 0.791. Therefore, it would appear that pipe flow directions do not impact the relationship between entropy and reliability in a dramatic way. This is a particularly reassuring result. It is self-evident from Figure 5 that the most reliable design is relatively very expensive. Referring to Figure 4, this design has a single terminal node at Node 8 and the least direct flow directions. It has much larger pipe diameters than the rest of the designs, which would account for the large increase in cost. Taking all the costs together, finally, there appears to be a strong diminishing returns effect (cost vs reliability) associated with the most reliable and costly design.

CONCLUSIONS

On a realistic scale, the calculation of reliability remains formidable for water distribution systems. By contrast, flow entropy is easy to compute. While the entropy function is not a direct substitute for reliability, it seems somehow to capture its essential properties. In this paper, examples have been used to test the strength of the relationship between entropy and reliability for a range of conditions. For the network in this study, the main conclusions are as follows: 1. With an entropy constraint, the higher the marginal cost of larger diameter pipes, the more uniform the pipe diameters. 2. The higher the entropy value, the more uniform the pipe diameters. 3. In general, reliability increases as entropy increases, the correlation being generally strong.

REFERENCES

Alperovits E and Shamir U (1977) Design of optimal water distribution systems. *Wat. Resour. Res.*, **13** (6), 885-900.

Awumah K, Goulter I and Bhatt SK (1991) Entropy-based redundancy measures in water distribution network design. *J. Hydraulic Eng.,* ASCE, **117** (5), 595-614.

Coelho ST (1997) *Performance in Water Distribution.* Research Studies Press, Taunton, England.

Cullinane MJ, Lansey KE and Mays LW (1992) Optimization-availability-based design of water distribution networks. *J. Hydraulic Eng.,* ASCE, **118** (3), 420-441.

Fujiwara O and Khang D (1990) A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resources Res.*, **26** (4), 539-549.

Shannon C (1948) A math. theory of communication, *Bell Syst. Tech. J.,* **27** (3), 379-428.

Tanyimboh TT (1993) *An Entropy-based Approach to the Optimum Design of Reliable Water Distribution Networks.* PhD thesis, Department of Civil Engineering, University of Liverpool, England.

Tanyimboh TT and Sheahan C (in press) A maximum entropy based approach to the layout optimization of water distribution systems. *Civil Engineering and Environmental Systems*

Tanyimboh TT, Tabesh M and Burrows R (2001) Appraisal of source head methods for calculating reliability of water distribution networks. *J. Water Res. Plan. and Management,* ASCE, **127** (4), 206-213.

Tanyimboh TT and Templeman AB (1993) Calculating maximum entropy flows in networks. *J. Operational Res. Society,* **44** (4), 383-396.

Tanyimboh TT and Templeman AB (1993) Optimum design of flexible water distribution networks. *Civil Engineering Systems,* **10** (3), 243-258.

Tanyimboh T and Templeman A (1993) A quantified assessment of the relationship between the reliability and entropy of water distribution systems. *Eng. Opt.*, **33** (2), 179-199.

Templeman A (1982) Discussion of "Optimization of looped water distribution systems." *J. Env. Eng. Div.,* ASCE, **108** (3), 599-602.

Wagner JM, Shamir U and Marks DH (1988) Water distribution reliability: analytical methods. *J. Water Res. Planning and Management,* ASCE, **114** (3), 253-275.

Yassin-Kassab A, Templeman A and Tanyimboh T (1999) Calculating maximum entropy flows in multi-source, multi-demand networks. *Eng. Opt.,* **31** (6), 695-729.

Paper 122

Proceedings of the Ninth International Conference
on Civil and Structural Engineering Computing,
B.H.V. Topping (Editor),
Civil-Comp Press, Stirling, Scotland.

Water System Entropy: A Study of Redundancy as a Possible Lurking Variable

Y. Setiadi, T.T. Tanyimboh, A.B. Templeman and B. Tahar
Department of Civil Engineering
University of Liverpool, United Kingdom

Abstract

This paper concerns an investigation of the possible influence of apparent network redundancy or insufficient capacity on the relationship between the entropy and reliability of water distribution systems. Pressure-driven simulation was used to analyse entropy-constrained minimum-cost designs and the results are compared with those obtained previously using demand-driven analysis. The study shows that the entropy-reliability relationship is much stronger when pressure-dependent analysis is used. Apparent network redundancy also appears not to have a significant effect on the entropy-reliability relationship. The redundancy investigated consists of a slight over- or under-capacity due to minor differences between the design optimisation model and subsequent simulation models.

Keywords: entropy, reliability, water distribution systems, pressure-dependent analysis, demand-driven analysis.

1 Introduction

Entropy as a surrogate measure for the reliability of water distribution systems (WDS) has been investigated for some time. It has the computational advantages of being easy to calculate, minimal data requirements and ease of incorporation into optimisation procedures. Reliability, on the other hand, is very computationally demanding. Based on this argument, the possible use of entropy as an indicator of reliability is very desirable.

To analyse the hydraulic behaviour of water distribution systems, most of the previous studies used demand-driven simulation models. The simulation assumes that demands in the network are fully satisfied regardless of the pressure in the system. The models give acceptable results when the systems are subject to normal operating conditions. However, WDSs are subject to component failures or very large demands, which may result in a reduction of the pressure in the system. When this happens, demand-driven analysis often gives results that indicate that the system is still supplying the full demand at lower, and sometimes, negative pressures. The validity of such results is obviously questionable. This study, however, uses head-dependent analysis (HDA) as well as demand-driven analysis (DDA). The pressure-dependent simulation approach has been suggested to provide more realistic results when WDSs operate under subnormal pressure conditions [1].

In addition to the use of pressure-dependent modelling, this is the biggest study of the relationship between entropy and reliability. The analyses involved the possible influence of redundancy or insufficient capacity, in the form of a small surplus or deficit in head at the critical node, on the relationship between entropy and reliability. This was done by comparing the performance of network designs with a small excess or shortfall in capacity with the same designs adjusted to satisfy the demands exactly. The possible effect of the cost objective function was also investigated. Different cost functions will produce different designs. These differences may have an effect on the entropy-reliability relationship. Another aspect was the influence of layouts on the relationship between entropy and reliability. Since the entropy of a WDS is a function of pipe flow rates, different layouts of the distribution system may have a significant impact on the relationship between entropy and reliability.

Overall, the results from the present study appear to strengthen the notion that the relationship between entropy and reliability is strong, with HDA generally yielding much better correlation than DDA.

2 Informational Entropy

Entropy, in the context of information theory, was first introduced by Shannon [2]. He developed a way of measuring the levels of information or uncertainty in different probability distributions. His entropy function can be written as

$$S/K = -\sum_{i} p_i \ln p_i \tag{1}$$

in which S is the entropy, K is an arbitrary positive constant often taken as 1, p_i is the probability associated with the *i*th outcome.

The values of entropy in this paper were calculated using the entropy function for WDS developed by Tanyimboh [3]

$$\frac{S}{K} = -\sum_{j\in IN}(Q_j/T)\ln(Q_j/T) - \frac{1}{T}\sum_{j=1}^{J} T_j\left((Q_j/T_j)\ln(Q_j/T_j) + \sum_{i\in N_j}(q_{ij}/T_j)\ln(q_{ij}/T_j)\right) \tag{2}$$

in which the subscript j represents all nodes including source nodes while J denotes the number of nodes and IN the set consisting of source or input nodes. Also, T is the total supply, T_j is the total flow reaching node j, including any external inflow while N_j represents all nodes immediately upstream of and connected to node j. Q_j is the demand at demand nodes or supply at source nodes and q_{ij} is the flow rate in pipe ij. The first term in the large parentheses is zero for demand nodes since there is no external inflow at these nodes. The entropy of a WDS is basically a function of internal pipe flow rates only since the desired nodal demands and supplies are usually specified.

3 Optimal Design of Water Distribution Systems

The design optimisation was carried out using a Fortran program [3] with the cost as the objective function to be minimized. The cost function used in this study is

$$Cost = \gamma \sum_{ij} L_{ij} D_{ij}^{e} \tag{3}$$

in which e is the cost exponent that is specified by the user and thought to have a range of 1.0 to 2.5 [4]. The cost coefficient γ, whose value depends on a range of factors, is also specified by the user. L_{ij} and D_{ij} are the length and diameter of pipe ij, respectively.

In the optimisation process, the program minimized the cost subject to several sets of constraints. These constraints consist of pipe head loss, continuity, conservation of energy and entropy equations. Lower and upper limits of the pipe diameter and pipe flow non-negativity constraints were also used.

3.1 Pipe Head Loss Equation

When water flows through pipes it experiences loss of energy due to friction at the pipe walls. This loss of energy is often called head loss, which is the energy loss per unit weight. Head loss is also caused by pipe bends, fittings and changes in cross-sectional area. The pipe head loss in this study was calculated using the Hazen-Williams empirical formula given below.

$$h_{ij} = \alpha L_{ij} (q_{ij} / C_{ij})^{1.852} / D_{ij}^{4.87} \quad \forall ij \tag{4}$$

For the above equation h_{ij} is the head loss along pipe ij and α is a constant, equal to 10.67 in S.I. units. The parameters L_{ij}, q_{ij}, C_{ij} and D_{ij} are the length, flow rate, roughness coefficient and diameter, respectively, for pipe ij. Minor losses of energy, for example due to fittings, were considered negligible in this study.

3.2 Continuity Equation

Continuity states that all the flows coming into a point must be equal to all the flows going out of that point. Therefore, the continuity equation can be written as

$$\sum_{ij \in IJ_j} q_{ij} = Q_j \qquad \forall j \tag{5}$$

in which the summation includes all the flows upstream and downstream of the node considered, indicated by IJ_j, and Q_j represents the external inflow or outflow at node j.

3.3 Conservation of Energy Equations

Two sets of equations for conservation of energy were considered in this study. They are the loop and path equations. The loop equation requires the net head loss around each loop in a pipe network to be zero. Therefore, the equation can be written as

$$\sum_{ij \in l} h_{ij} = 0 \qquad \forall l \tag{6}$$

in which l (l = 1, 2, ...) represents the *l*th loop.

The path equation ensures that the total head loss along any path is equal to the difference in head between the end points of that path. The equation may therefore be written as

$$\sum_{ij \in p} h_{ij} = h_p \qquad \forall p \tag{7}$$

where p (p = 1, 2, ...) represents the *p*th path having a known value of head loss, h_p.

Another equation is used to ensure that nodal service pressures are satisfactory. This equation is given below.

$$\sum_{ij \in t} h_{ij} \le H_s - H_{\min,t} \qquad \forall t \tag{8}$$

In this equation t (t = 1, 2, ...) represents a path from a specified source to a terminal node t, i.e. a node with no other nodes downstream of it, H_s is the head at a specified source and $H_{min,t}$ is the minimum allowable head at terminal node t.

3.4 Entropy Constraint

As mentioned earlier, one of the advantages of entropy is that it can be incorporated into the optimisation procedure as one of the constraints. This entropy constraint takes the form of

$$S / K \ge S_{\min} \tag{9}$$

in which K is an arbitrary constant, S is the entropy and S_{min} is the minimum desired entropy value. This equation will ensure that the entropy of the network does not fall below the specified value.

Apart from the constraints mentioned above, the diameter of the pipes should not lie outside the predetermined range. Within this range the value of the diameter is considered continuous to simplify the optimisation. This diameter restriction takes the following form

$$D_{\max} \geq D_{ij} \geq D_{\min} \qquad \forall ij \tag{10}$$

in which D_{ij} is the diameter of pipe ij while D_{max} and D_{min} are the maximum and minimum allowable pipe diameters, respectively.

4 Network Analysis

Two methods of network analysis were used in this study for comparison purposes, the first being Demand Driven Analysis (DDA) and the other Head Dependent Analysis (HDA).

4.1 Demand Driven Analysis (DDA)

This method of network analysis has been widely used in the water industry for many years. Unfortunately, there are a few disadvantages arising from the use of this method. DDA does not take into consideration the relationship between the nodal outflows and the pressure within the system. It assumes that the demands of the system are fully satisfied regardless of the pressure in the system. In consequence, when the pressure drops below the required level, network analysts would have no information on how much water would be delivered by the system under the available pressure regime. In this situation some customers would receive reduced supplies and, in the worst scenario, they might not receive any supply at all.

The drop in pressure in the distribution system can be triggered by many factors. Excessive abstraction at one demand node, for example in a fire fighting situation, may cause the pressure in the neighbouring abstraction points to drop below the required level. The analysis using this method was carried out using a computer software called EPANET, which is freely available on the World Wide Web and is provided by the American Environmental Protection Agency.

4.2 Head Dependent Analysis (HDA)

Pressure dependent analysis has long been suggested to surpass demand driven analysis, particularly for networks under subnormal operating conditions. It is well known that outflows from a WDS are dependent upon the pressure within that system and therefore the DDA assumption that demands are always satisfied regardless of the pressure in the system is flawed. HDA takes into consideration the pressure dependency of outflows, and in consequence, the results are more realistic. Nevertheless, this method is not yet commonly used in the water industry since more research and verification of the true relationship between network pressure and nodal outflows are still necessary.

Some researchers, however, have proposed several assumed head-outflow relationships. For example, Wagner et al. [5] suggested the following function

$$\frac{Q_j^{avl}}{Q_j^{req}} = 0 \quad ; \qquad H_j < H_j^{\min} \tag{11a}$$

$$\frac{Q_j^{avl}}{Q_j^{req}} = \left(\frac{H_j - H_j^{\min}}{H_j^{des} - H_j^{\min}} \right)^{\frac{1}{n_j}} \quad ; \qquad H_j^{\min} \le H_j < H_j^{des} \tag{11b}$$

$$\frac{Q_j^{avl}}{Q_j^{req}} = 1 \quad ; \qquad H_j \ge H_j^{des} \tag{11c}$$

in which Q_j^{avl} and Q_j^{req} are the actual outflow that can be delivered by the system and the required outflow or demand, respectively. H_j is the actual head at node j, H_j^{min} is the minimum nodal head at node j, below which there would be no outflow and H_j^{des} is the desired head at node j, above which the outflow would be equal to the demand. Values of the exponent parameter, n_j, are thought to lie between 1.5 and 2 [6]. With regard to Equation (11b), Q_j^{avl} is set to zero if H_j is less than H_j^{min} or equal to Q_j^{req} if H_j reaches H_j^{des}, as shown in Equations (11a) and (11c), respectively.

The HDA in this study was carried out using a computer program called PRAAWDS, which stands for Program for the Realistic Analysis of the Availability of Water in Distribution Systems [1]. This program calculates the actual flow delivered under normal and subnormal pressure conditions.

5 Reliability

In this paper the reliability is defined as the time-averaged value of the ratio of the flow delivered to the flow required [3]. By assuming a constant demand value, the reliability of a water distribution system can be written as

$$R = \frac{1}{T}\left(p(0)T(0) + \sum_{m=1}^{M} p(m)T(m) + \sum_{\substack{m=1 \\ \forall n \succ m}}^{M-1} p(m,n)T(m,n) + \cdots \right)$$
$$+ \frac{1}{2}\left(1 - p(0) - \sum_{m=1}^{M} p(m) - \sum_{\substack{m=1 \\ \forall n \succ m}}^{M-1} p(m,n) - \cdots \right) \tag{12}$$

in which R is the system reliability, p(0) is the probability that no pipe is unavailable, p(m) is the probability that only pipe m is unavailable and p(m, n) is the probability that only pipes m and n are unavailable. Similarly, T(0), T(m) and T(m, n) are the respective total flows supplied with no pipes unavailable, only pipe m unavailable, and only pipes m and n unavailable. Finally, M stands for the number of pipes while T represents the total demand. A derivation of the above reliability formula is contained in Tanyimboh and Sheahan [7]. For DDA, T(m) and T(m,n) were calculated using the source head method [8].

6 Results and Discussion

The investigation was carried out using the network of Figure 1. The source has piezometric level of 100m while demand nodes have elevations of 0m. The desired nodal service head for fully satisfactory performance is 30m and the nodal head corresponding to zero nodal outflow is 0m. All pipes are 1000m long with a Hazen-Williams coefficient of 130. The value of n_j used in Equation (11) is 2.

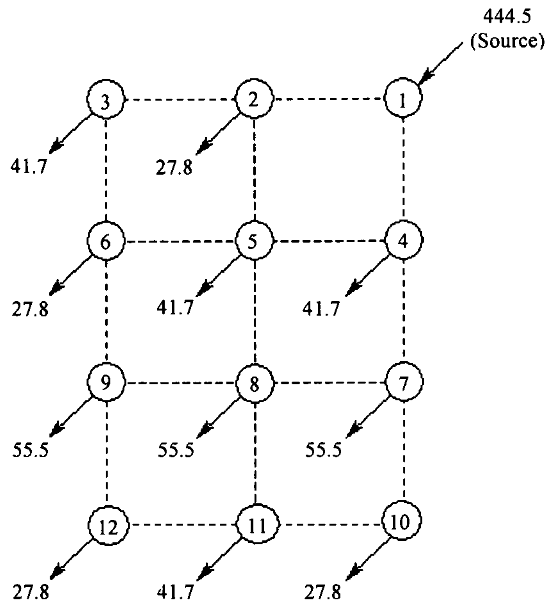


Figure 1: Network of supply and demand nodes with all demands in litre per second

Different programs were used in the optimisation procedure in the design process and in the subsequent hydraulic simulations of the designs. Round-off errors, especially in the diameters, produced small surpluses or deficits in head at the critical nodes. Hayuti [11] suggested that redundancy or insufficient capacity in the WDS, in the form of the surplus or deficit in head at the critical node, might have a bearing on the observed relationship between entropy and reliability given the very small differences between the reliability values.

In this study, the performance of designs with an excess or shortfall in capacity was compared to slightly adjusted designs which satisfied the demands exactly. The head at the source, for all designs with a surplus or deficit in head at the critical node was altered so that, at the critical node, the head was precisely equal to the desired service head. These designs were then re-analysed and their reliability values obtained. The results of this analysis are shown in Figure 2 in which the reliability values before and after the head modification for all the designs in this study are plotted against each other. It shows that all the designs have virtually identical pairs of reliability values.

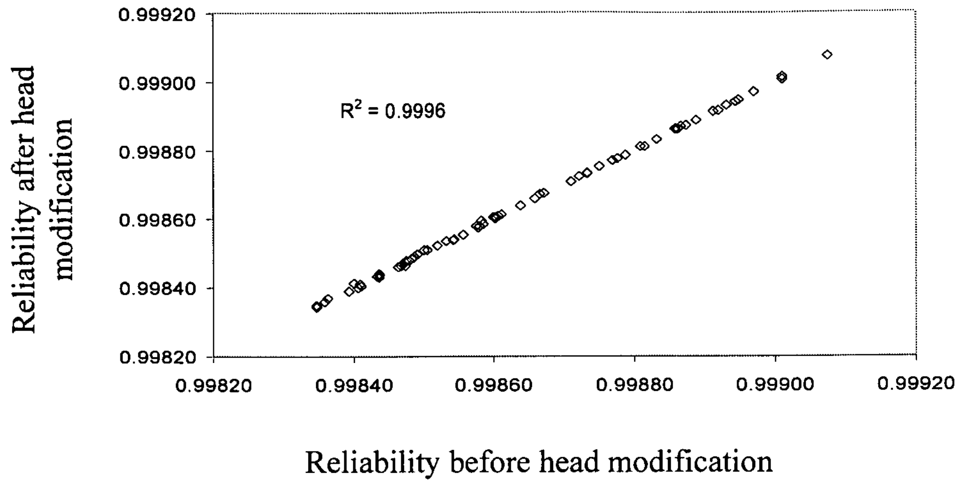


Figure 2: DDA-reliability verification for all the designs in the present study

Figures 3 and 4 show plots of reliability against surplus head for DDA and HDA, respectively, for all the designs. They confirm that the correlation between reliability and the surplus head is insignificant. Figure 5 also shows that any correlation between entropy and the surplus head is negligible. The results in Figures 2 to 5 suggest that the influence of the small surplus or deficit in head at the critical node upon the entropy-reliability relationship is negligible.

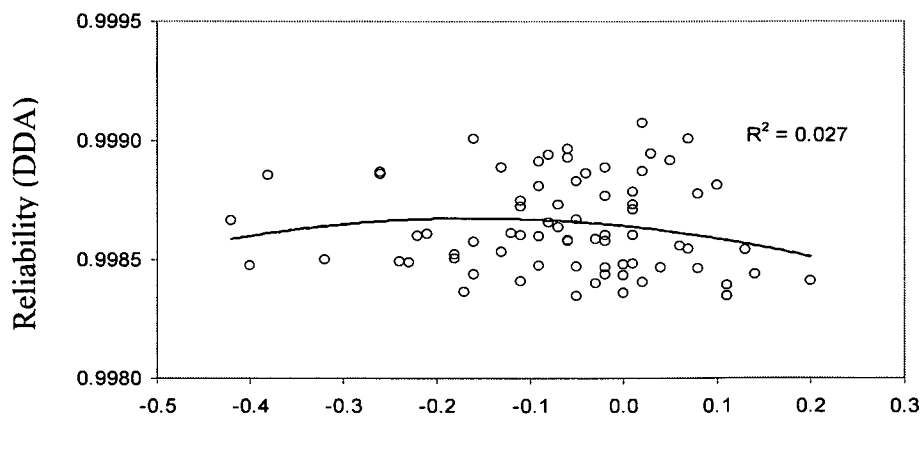


Figure 3: DDA-reliability *vs* surplus head at the critical node

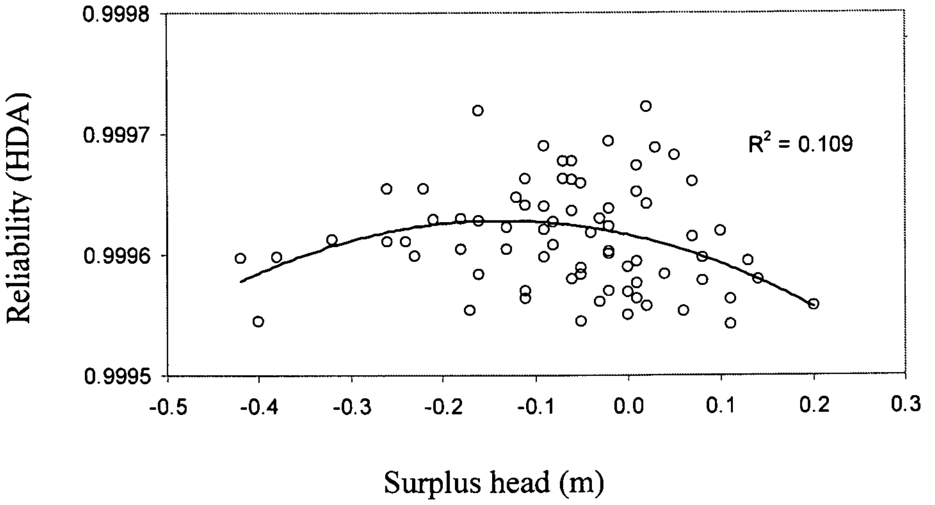


Figure 4: HDA-reliability *vs* surplus head at the critical node

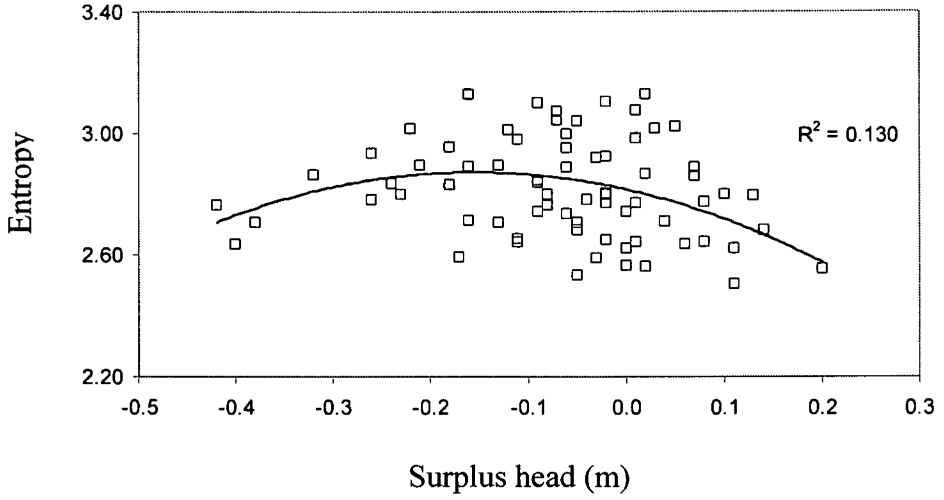


Figure 5: Entropy *vs* surplus head at the critical node

The effects of different layouts have also been investigated in this study. The entropy of a WDS is a function of the pipe flow rates and the pipe flow rates are affected by the layout of the network. While the effects of layouts on the relationship between entropy and reliability have been investigated using DDA [7, 9], the present study investigated this issue using pressure dependent analysis. Figure 6 shows plots of entropy and cost against reliability for the maximum-entropy minimum-cost designs from Reference [9]. It appears that the relationship between entropy and reliability is a great deal stronger when analysed by the HDA method.

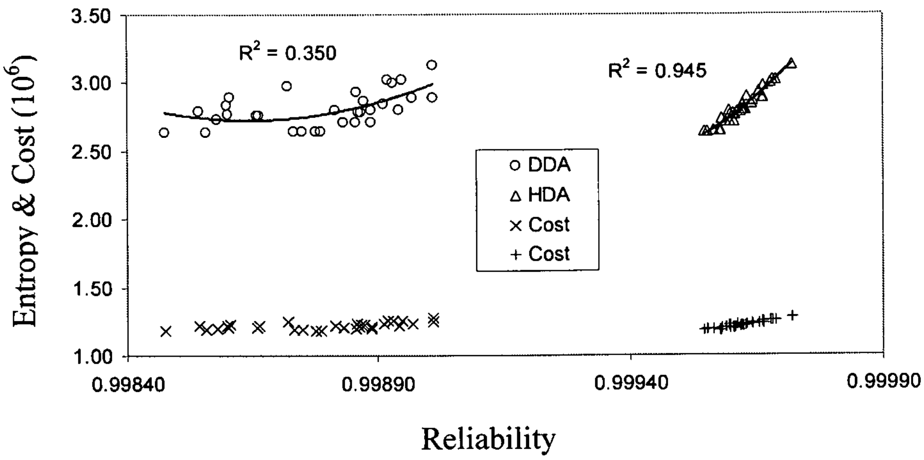


Figure 6: The effect of layout on the relationship between entropy and reliability with values of γ = 800 and e = 1.5

Another issue investigated in this study concerns the influence of the cost objective function of Equation (3) on entropy-reliability relationship. The 43 entropy-constrained minimum-cost designs were taken from Reference [10]. These designs, all of which are based on the six-loop 17-pipe layout in Figure 1, range from the smallest to the largest possible entropy values. Based on Equation (3), the first twenty-two designs were generated using values of γ and e of 800 and 1.5, respectively. The results of these designs can be seen in Figure 7 below. The remaining designs, the results of which are shown in Figure 8, were generated using γ = 1600 and e = 2.0.

Looking at the two graphs of entropy against reliability, it is evident that the relationship between entropy and reliability remains strong despite the differences in the values of γ and e. The two plots of entropy against reliability follow a very similar pattern and, again, HDA shows a stronger relationship than DDA. It should be noted that the costs have been plotted twice in Figures 6 to 9, even though the two cost data sets are identical in each case, for completeness and ease of reference.

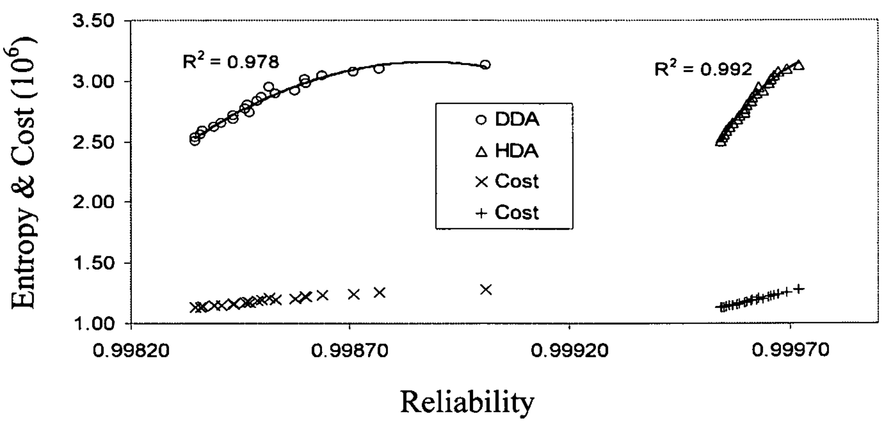


Figure 7: Plots of entropy *vs* reliability for γ = 800 and e = 1.5

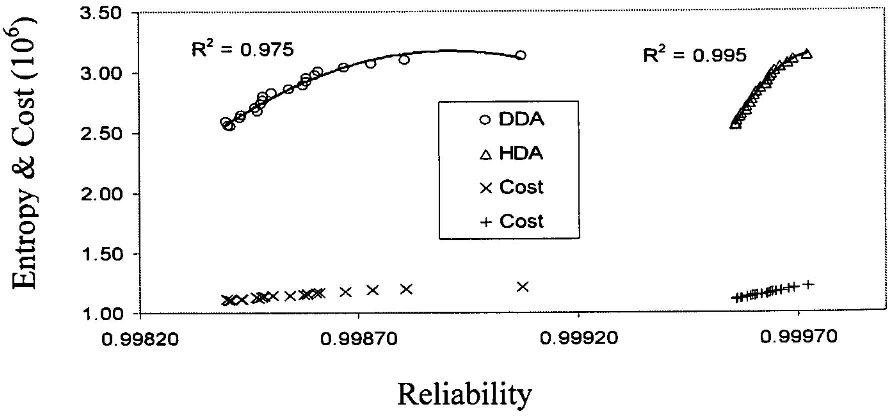


Figure 8: Plots of entropy *vs* reliability for $\gamma = 1600$ and e = 2.0

Overall, the comparison between HDA and DDA can be clearly observed in Figure 9, in which the results for all the designs are shown. It confirms that the correlation between entropy and reliability is much stronger when the analysis is done using HDA.

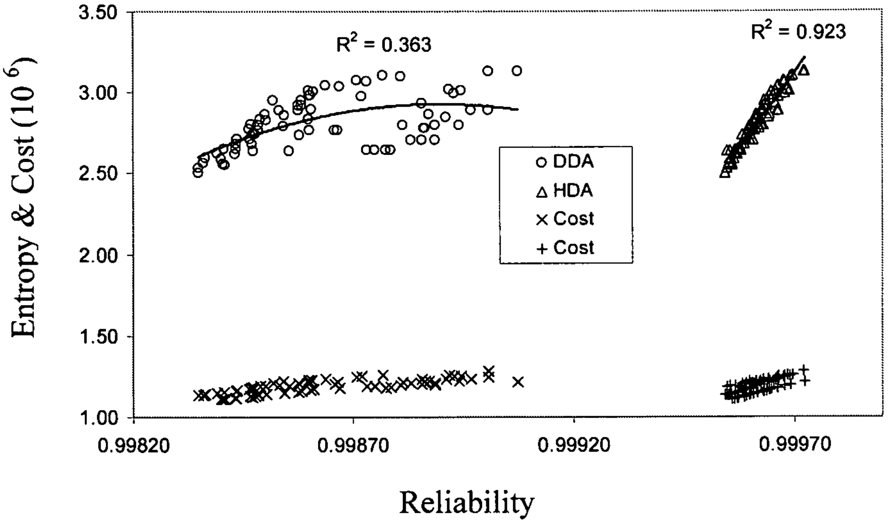


Figure 9: Plots of entropy *vs* reliability for all the entropy-constrained minimum-cost designs in the present study

7 Conclusions

For the network investigated, there seems to be a strong relationship between entropy and reliability. This relationship is stronger when the analysis is carried out using HDA in comparison to DDA. Redundancy or insufficient capacity in the form of small surpluses or deficits in head at the critical nodes have no real influence on the entropy-reliability relationship.

Acknowledgment

The authors are grateful to the UK Engineering and Physical Science Research Council for its financial support for the PRAAWDS component of this research under Grant GR/NO2672. MEng. Student Claire Sheahan also contributed to this research as part of her project.

References

[1] Tahar, B., Tanyimboh, T.T., Templeman, A.B., "Pressure-dependent Modelling of Water Distribution Systems", in "Proceedings of the Third International Conference on Decision Making in Urban and Civil Engineering", Khosrowshahi, F., (Editor), London, 2002.

[2] Shannon, C., "A Mathematical Theory of Communication", Bell System Technical Journal, 27(3), 379-428, 1948.

[3] Tanyimboh, T.T., "An Entropy-based Approach to the Optimum Design of Reliable Water Distribution Networks", PhD Thesis, Department of Civil Engineering, University of Liverpool, UK, 1993.

[4] Fujiwara, O., Kang, D., "A Two-phase Decomposition Method for Optimal Design of Looped Water Distribution Networks", Water Resources Research, 26(4), 539-549, 1990.

[5] Wagner, J.M., Shamir, U., Marks, D.H., "Water Distribution Reliability: Simulation Methods", Journal of Water Resources Planning and Management, ASCE, 114(3), 276-294, 1988.

[6] Gupta, R., Bhave, P.R., "Comparison of Methods for Predicting Deficient Network Performance", Journal of Water Resources Planning and Management, ASCE, 122(3), 214-217, 1996.

[7] Tanyimboh, T.T., Sheahan, C., "A Maximum Entropy Based Approach to the Layout Optimization of Water Distribution Systems", Journal of Civil Engineering and Environmental Systems, 19(3), 223-253, 2002.

[8] Tanyimboh, T.T., Tabesh, M., Burrows, R., "Appraisal of Source Head Methods for Calculating Reliability of Water Distribution Networks", Journal of Water Resources Planning and Management, ASCE, 127(4), 206-213, 2001.

[9] Tanyimboh, T.T., Templeman, A.B., "A Quantified Assessment of the Relationship Between the Reliability and Entropy of Water Distribution Systems", Engineering Optimisation, 33(2), 179-199, 2000.

[10] Tanyimboh, T.T., Setiadi, Y., Mavrokoukoulaki, A., Storer, C., "The Suitability of Informational-theoretic Entropy as a Surrogate Performance Measure for Water Distribution Systems", in "Proceedings of the Third International Conference on Decision Making in Urban and Civil Engineering", Khosrowshahi, F., (Editor), London, 2002.

[11] Hayuti, M.H., "Performance Assessment of Water Distribution Systems", MSc.(Eng.) Dissertation, Department of Civil Engineering, University of Liverpool, UK, 2002.

Proceedings of the Fourth International Conference
on Engineering Computational Technology,
B.H.V. Topping and C.A. Mota Soares (Editors),
Civil-Comp Press, Stirling, Scotland.

Paper 66

Hydraulically Predictable Water Distribution Networks

T.T. Tanyimboh and Y. Setiadi
Department of Civil Engineering
University of Liverpool, United Kingdom

Abstract

It would appear that the entropy-constrained approach to the design optimization of water distribution systems has the potential to generate designs which have a high degree of hydraulic predictability in the sense that, in general, the higher the entropy value, the greater the likelihood that the hydraulic properties of the design achieved will be intuitively obvious. This point is illustrated in this paper using results for two networks. Using pressure-dependent modelling to simulate critical operating conditions including fire-fighting and pipe failure, the key nodes and distribution mains are identified. The criteria used include the nodal hydraulic reliability and pressure-dependent rate of flow delivery; and the deterioration in the hydraulic performance of the network. A new hydraulic performance indicator has been introduced, based on the rate at which a water distribution network dissipates energy relative to the amount of water supplied.

Keywords: water distribution networks, uncertainty, informational entropy, design optimisation, hydraulic reliability, fire-fighting, pressure-dependent modelling, energy dissipation.

1 Introduction

There is a considerable amount of uncertainty associated with the design and operation of water distribution systems. These include: long-term projections of the growth in demand; the spatial distribution of the nodal demands coupled with diurnal and seasonal consumption patterns; variations in electricity/energy tariffs; bursts and component failures; possible changes in pipe diameters and roughness with age. In view of the operational uncertainties, evidence is provided herein which shows that informational entropy, which is a measure of uncertainty, can be used to help control the location of critical pipes and nodes in a water distribution system. Following a pipe failure/removal or large localised increase in demand, flows in pipe networks are rerouted in complex ways which are generally difficult to predict prior to a full simulation of the network concerned.

Results are presented in this paper which suggest that if the entropy is sufficiently large, the hydraulic characteristics of the network with regard to flow rerouting due to link failures or increased network flows may be less unpredictable. It appears, also, that entropy can be used to force the critical nodes for fire fighting flows to be near the design terminal nodes of a network or other intuitively more obvious locations and, in consequence, to reduce the need to explicitly consider fire fighting flows at all nodes of a network at the design stage. The ability to predetermine the most likely critical nodes and links of a water distribution system may find applications in certain reliability analysis and optimal design formulations [1]. The determination of the critical links and nodes is based on several measures including the actual amount of water supplied under critical operating conditions, hydraulic reliability and energy dissipation rate. The hydraulic simulations of the sample networks used were carried out using pressure-dependent modelling.

2 Performance Assessment

2.1 Energy Dissipation

Based on a suggestion that the efficiency of a pipe can be gauged from the rate at which the pipe dissipates energy [2], Tanyimboh and Templeman [3] compared alternative designs for a pipe network using the respective energy dissipation rates of the designs. Tanyimboh and Templeman [3] stressed the need for caution when assessing water distribution systems using the rate of energy dissipation. Their approach was based on demand-driven simulation with a stipulation that all demands be fully satisfied.

Herein, consistency in the comparisons is achieved using a new measure, the rate of energy dissipation per unit flow delivered by the network under stressed conditions. It is worth recalling that in a head-dependent modelling environment [4, 5], alternative designs which deliver the same amount of water under normal operating conditions may supply different quantities of water under pressure-deficient conditions. The energy dissipation rate per unit flow rate of the water, taken over the entire water distribution network, would appear to represent a measure of hydraulic performance. Thus, for the same rate of water supply, the amount of energy dissipated would increase as the stress on the network increases. The equation for the rate of energy dissipation, E, is

$$E = \rho g \sum_{ij \in IJ} Q_{ij} h_{ij} \tag{1}$$

in which ρ is the density of water; g is the acceleration due to gravity; Q_{ij} and h_{ij} are flow rate and head loss in pipe ij, respectively; IJ represents the links of the network.

2.2 Reliability and Pressure Dependent Analysis

The hydraulic performance of water distribution networks is assessed herein based on the amount of water supplied or the demand satisfaction ratio for the network or node as appropriate. The amount of water that the network can actually supply as opposed to the demand was obtained by modelling the relevant operating conditions using the head-dependent analysis approach [4, 5]. Essentially, the system of equations for the network can be set up using the flow continuity equations for each node, *i*, as follows

$$F_i(H_i, H_j) = \sum_{j \in N_i} Q_{ij} - Q_i = 0 \tag{2}$$

where Q_{ij} is the flow in link *ij* (i.e. pipe, pump or valve), Q_i is the head-dependent outflow at node *i* and N_i represents all the nodes connected to node *i*. H_i and H_j are the piezometric heads at nodes *i* and *j*, respectively. Thus Q_i is the actual amount of water that the network can supply at node *i*, which may be less than or equal to the demand. Herein the pressure-dependent nodal outflow function proposed by Wagner et al. [6] was used. The head-dependent network analysis was carried out using the PRAAWDS computer program [4, 5].

The reliability measure used in this research is the time-averaged value of the ratio of the flow supplied to the flow required. This reliability measure requires head-dependent modelling of the distribution system as described above and pipe failure/availability rates for the network. The details of the procedure and relevant equations can be found in Tanyimboh [7] and the references therein.

2.3 Informational Entropy

Entropy is a measure of the amount of uncertainty associated with a probability scheme. The values of entropy in this paper were calculated as follows [8].

$$\frac{S}{K} = -\sum_{j \in IN} (Q_j/T)\ln(Q_j/T) - \frac{1}{T}\sum_{j=1}^{J} T_j \left((Q_j/T_j)\ln(Q_j/T_j) + \sum_{i \in N_j} (q_{ij}/T_j)\ln(q_{ij}/T_j) \right) \tag{3}$$

in which the subscript *j* represents all nodes including source nodes while *J* denotes the number of nodes and *IN* the set consisting of source or input nodes. Also, *T* is the total supply, T_j is the total flow reaching node *j*, including any external inflow while N_j represents all nodes immediately upstream of and connected to node *j*. Q_j is the outflow at demand nodes or supply at source nodes and Q_{ij} is the flow rate in pipe *ij*. The first term in the large parentheses is zero for demand nodes since there is no external inflow at these nodes.

3 Examples, Results and Discussion

3.1 Sample Networks

The demonstration of some of the predictable properties of maximum entropy designs of water distribution networks is based on the two sample networks considered below. The assessments herein are based on designs which were obtained using an entropy-constrained cost minimisation approach which uses the usual design constraints for water distribution networks along with a specified minimum entropy value for each design. The entropy value of each design is thus equal to the specified minimum value. Further details including the pipe diameters can be found in Tanyimboh [9]. All pipes are 1000 m long with a Hazen-Williams coefficient of 130.

PRAAWDS [4, 5] was used to determine the hydraulic performance of the designs and the cases corresponding to the unavailability of individual pipes, based on pressure-dependent modelling as summarised in Subsection 2.2. All the demand nodes of the two samples networks have an elevation of zero. The nodal piezometric head below which outflow is zero was taken as zero, this being the elevation of the nodes, while the demand is satisfied in full when the residual head reaches 30 m. The value of 30 m for the desired residual head was chosen to ensure that the results of the present study, which uses pressure-dependent modelling, are comparable to Tanyimboh [9]. The comparison, however, is not included in this paper. The reliability and other performance indicators used were obtained as described in Section 2.

3.2 Critical Operating Conditions

The three critical operating conditions described shortly were used to assess the predictability of the hydraulic performance of the designs, on the basis of the locations of the critical pipes and nodes. The critical pipe is the pipe the removal of which causes the greatest deterioration in the performance of the distribution network, as determined by the increase in the energy dissipated (i.e. the greatest increase) or the reduction in the amount of water supplied (i.e. the greatest reduction). The determination of the critical node is somewhat context dependent. For example, for the pipe failure/removal simulations, the critical node is the node with the smallest post-failure demand satisfaction ratio. For the fire-fighting simulations, the critical node is the node which is least able to cope with the fire-fighting demand.

The critical operating conditions include:

1. The unavailability of each pipe in turn.

2. A large demand of 0.25 m^3/s at each node in turn, in place of the normal demand at that node, with all other nodal demands at their normal design values. This

situation would be akin to a fire occurring under a locally ideal (i.e. zero) background demand condition.

3. A large demand of 0.25 m^3/s at each node in turn, in addition to the normal demand at that node, with all other nodal demands at their normal design values. This situation would be akin to a fire occurring under a more adverse background demand condition.

3.3 Network 1

The layout of the network, with nodal demands, is shown in Figure 1. The piezometric head at the source was taken as 50 m to achieve a residual head of 30 m at node 6 under normal operating conditions which, therefore, means that the demand at each node is fully satisfied. Performance indices for the respective designs have been summarised graphically in Figures 2 and 3. The difference in cost between the least and most expensive designs is 5.2% [9]. The critical nodes and links are shown in Tables 1 to 3, with each table corresponding to one of the three critical operating conditions mentioned earlier. The performance of the designs under the two simulated fire-fighting situations is shown in Figures 4 and 5.

3.3.1 Discussion

Previous research has established that, in general, the hydraulic performance or reliability of water distribution systems increases as the entropy increases (e.g. [10]). Tanyimboh and Templeman [3] observed that, as the entropy increased, overall system-wide improvements in performance did not in general occur at the expense of deteriorations at the micro level. Figure 3 illustrates this property very clearly, thanks to the ability to calculate the actual amount of water delivered at each node using head-dependent modelling. This paper, however, is primarily concerned with the hydraulic predictability of maximum-entropy designs. Figure 3 shows that the reliability of Node 2 is low, compared to Node 5, for the design with the smallest entropy value. This would appear to be counterintuitive given that Node 2 is close to the source. On the other hand, the reliability of Node 2 for the design with the maximum entropy value is higher than then rest of the nodes. This would appear to suggest that maximum entropy designs are hydraulically more predictable.

The system-wide performance indicators in Figure 2a (average value and worst case) are consistent with Figure 3 in that they also show an overall improvement in performance. Figure 2a suggests that the marginal increase in the amount of energy dissipated for the extra flow delivered is not disproportionate. Turning once more to hydraulic predictability and focussing on individual nodes and pipes, the reasonableness of the results in Tables 1 to 3 can be explained in general in terms of the following considerations.

The criteria for critical nodes include:

1. The distance of the node from the source;
2. The number of paths supplying the node;
3. The normal demand at the node;

4. The increase in demand at the node during a simulated fire-fighting situation.

The criteria for critical pipes include:

1. The location of the pipe relative to the source;
2. The location of the pipe relative to the major demand nodes;
3. The number of alternative supply pipes incident on the major demand nodes. For example, a single pipe supplying a major demand node is likely to be more critical than two pipes supplying another node with a similar demand.

It is worth observing that the above-mentioned criteria for the pipes and nodes can apply individually or in concert, often in complex ways. In a similar vein, although

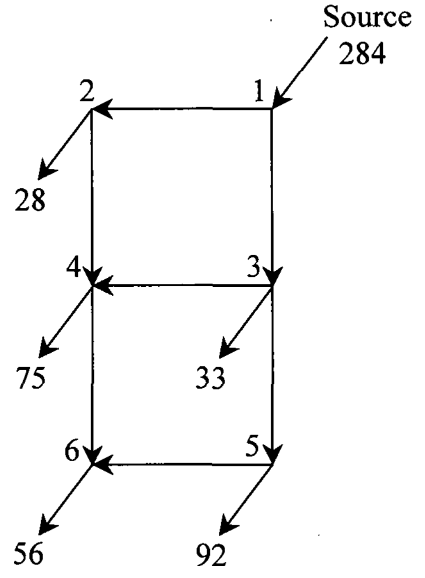


Figure 1: Network 1 under normal operating conditions with demands in l/s

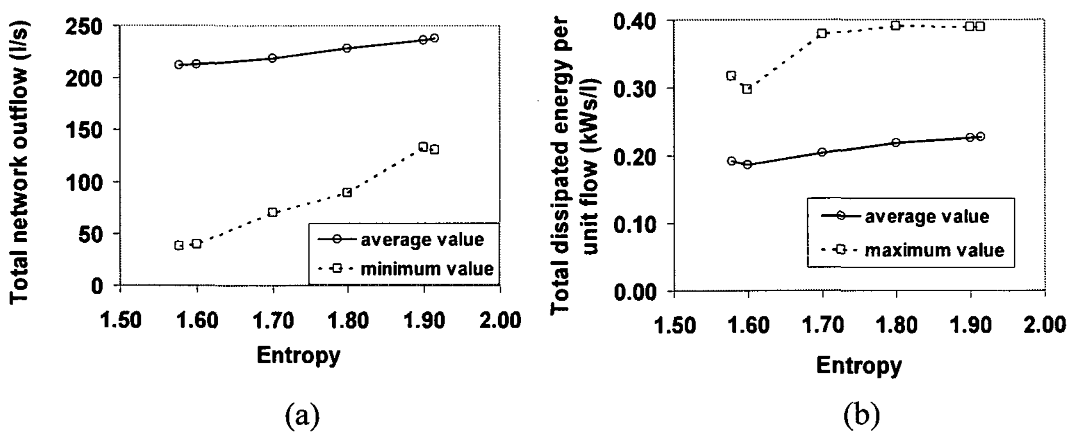


Figure 2: Total flow supplied and total dissipated energy per unit flow supplied for single-pipe failures for Network 1

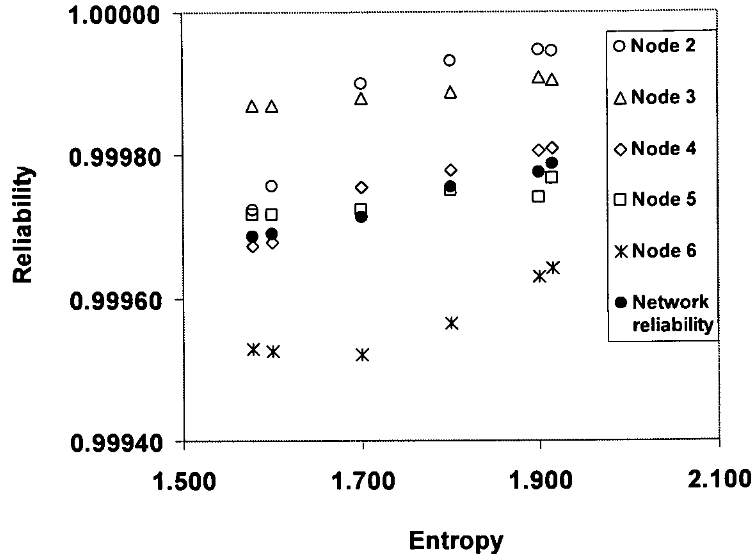


Figure 3: Hydraulic reliability for Network 1

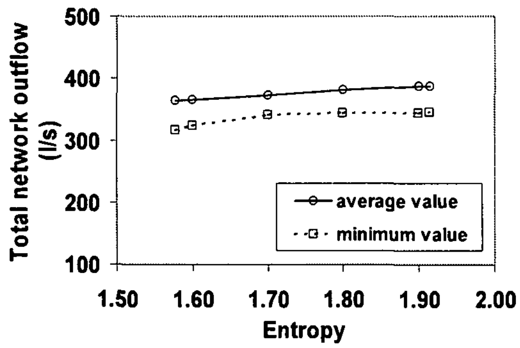
Assessment Criteria	Network Elements	Network Entropy					
		1.578	1.600	1.700	1.800	1.900	1.915
Total flow supplied	Critical link	1 - 3	1 - 3	1 - 3	1 - 3	1 - 3	1 - 3
	Next critical link	3 - 5	3 - 5	3 - 5	3 - 5	3 - 5	3 - 5
	Critical node	6	6	6	6	5	5
Total energy dissipated per unit flow supplied	Critical link	1 - 3	1 - 3	1 - 3	1 - 3	1 - 3	1 - 3
	Next critical link	1 - 2	1 - 2	1 - 2	1 - 2	1 - 2	1 - 2
	Critical node	6	6	6	6	5	5

Table 1: Critical links and nodes for Network 1 for single-link failures

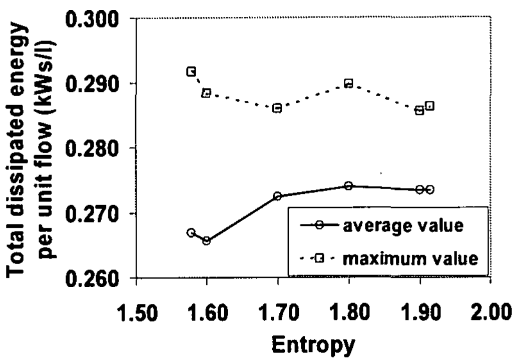
the 'flow delivered' and 'energy dissipated' parameters are not unrelated, the properties they measure differ in a subtle way. For example, Table 1 shows that there is broad agreement between the two parameters. Where they lead to different results, each outcome is in general justifiable. Taking an example from Table 1 and looking at Figure 1, it can be seen that Node 5 has the largest demand and is supplied by a single path. Node 6, on the other hand is supplied by three paths. Thus, bearing in mind the respective distances from the source, it seems reasonable to expect that Node 5 would be a critical node. It can be seen in Table 1 that the

Assessment Criteria	Network Elements	Network Entropy					
		1.578	1.600	1.700	1.800	1.900	1.915
Total flow supplied	Critical node	2	2	6	6	6	6
	Next critical node	6	6	2	2	5	4
Total energy dissipated per unit flow supplied	Critical node	6	6	6	2	6	6
	Next critical node	4	4	2	6	2	2

Table 2: Critical nodes for Network 1 for nodal fire-fighting demands replacing normal demands



(a)

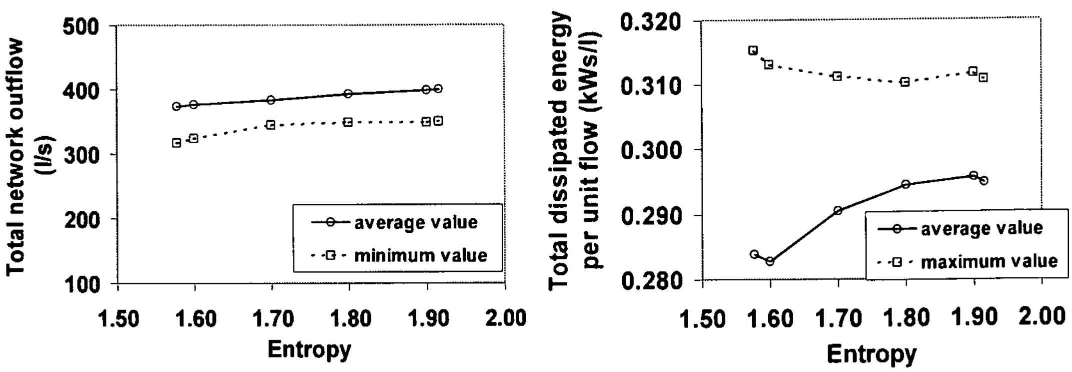


(b)

Figure 4: Total flow supplied and total dissipated energy per unit flow supplied for nodal fire-fighting demands replacing normal demands for Network 1

Assessment Criteria	Network Elements	Network Entropy					
		1.578	1.600	1.700	1.800	1.900	1.915
Total flow supplied	Critical node	2	2	6	6	6	6
	Next critical node	6	6	4	4	5	5
Total energy dissipated per unit flow supplied	Critical node	5	5	5	5	5	4
	Next critical node	6	6	6	4	4	5

Table 3: Critical nodes for superimposed nodal fire-fighting demands for Network 1



(a) (b)

Figure 5: Total flow supplied and total dissipated energy per unit flow supplied for superimposed fire-fighting demands for Network 1

designs with the highest entropy values have Node 5 as the critical node while the rest of the designs do not. This would appear to suggest that maximum entropy designs are somewhat more predictable. Unfortunately, for reasons of brevity, a comprehensive discussion of each of the tables and graphs of the results is not included herein. However, in general they follow a similar pattern and have been included for completeness.

3.4 Network 2

The layout of the network is shown in Figure 6. The piezometric head at the source was taken as 53.5 m to achieve a residual head of 30 m at node 9 under normal operating conditions which, therefore, means that the demand at each node is fully

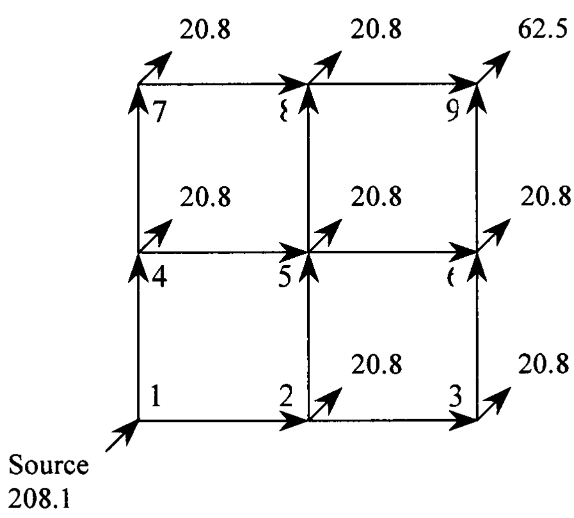
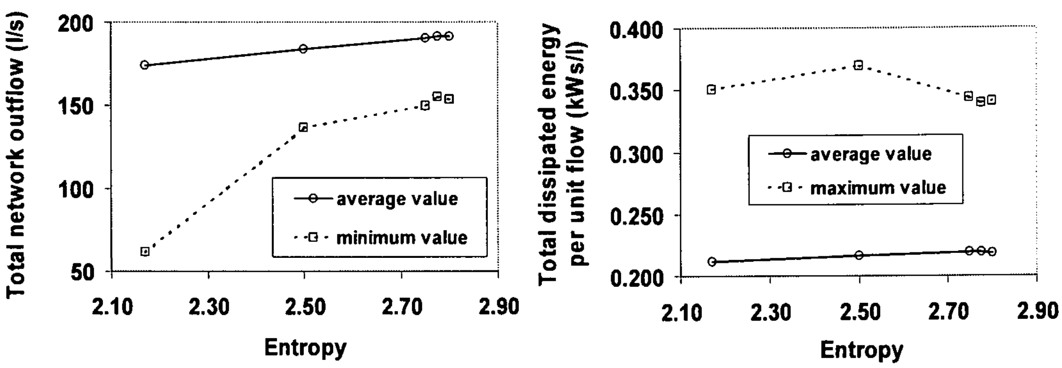


Figure 6: Network 2 under normal operating conditions with demands in l/s

satisfied. Performance indices for the respective designs have been summarised graphically in Figures 7 and 8. The difference in cost between the least and most expensive designs is 5.4% [9]. The critical nodes and links are shown in Tables 4 to 6, with each table corresponding to one of the three critical operating conditions mentioned previously in Subsection 3.2. The performance of the designs under the two simulated fire-fighting situations is shown in Figures 9 and 10. Whereas the designs for Network 1 were generated in a manner such that a more or less uniform coverage of the entropy range was achieved, the entropy values used for the designs of Network 2 were selected to achieve a greater coverage of the upper end of the entropy range so as to provide a comparison of near-maximum entropy designs [9].



(a) (b)

Figure 7: Total flow supplied and total dissipated energy per unit flow supplied for single-link failures for Network 2

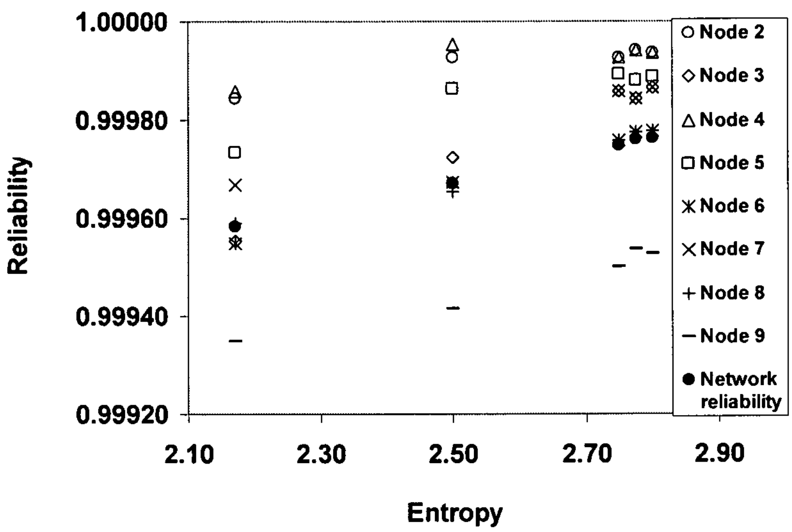


Figure 8: Hydraulic reliability for Network 2

Assessment Criteria	Network Elements	Network Entropy				
		2.170	2.500	2.750	2.775	2.800
Total flow supplied	Critical link	1 - 4	1 - 2	1 - 2 & 1 - 4	1 - 2 & 1 - 4	1 - 2 & 1 - 4
	Next critical link	4 - 5	5 - 6	2 - 3 & 4 - 7	2 - 5 & 4 - 5	2 - 3 & 4 - 7
	Critical node	7	9	9	9	9
Total energy dissipated per unit flow supplied	Critical link	1 - 4	1 - 2	1 - 2 & 1 - 4	1 - 2 & 1 - 4	1 - 2 & 1 - 4
	Next critical link	4 - 5	1 - 4	2 - 3 & 4 - 7	2 - 5 & 4 - 5	2 - 3 & 4 - 7
	Critical node	7	9	9	9	9

Table 4: Critical links and nodes for single-link failures for Network 2

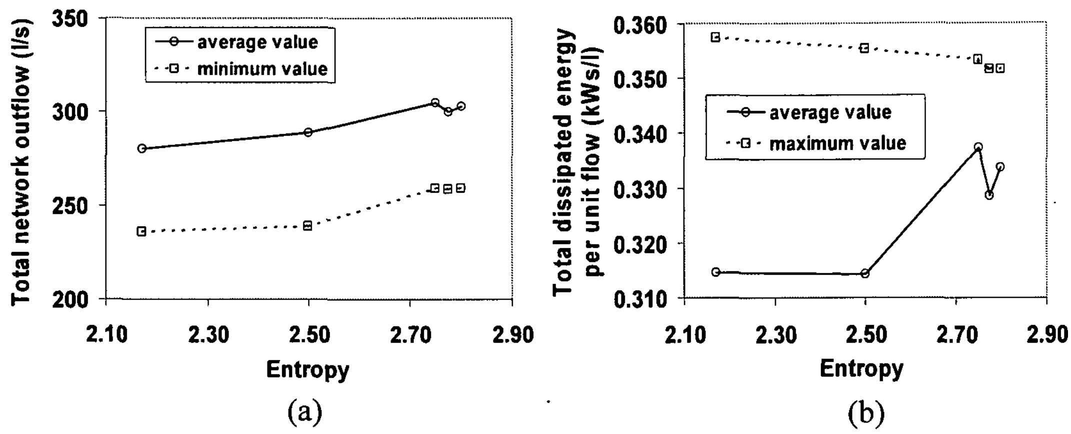


Figure 9: Total flow supplied and total dissipated energy per unit flow supplied for fire-fighting demands replacing normal demands for Network 2

3.4.1 Discussion

Network 2 has been included herein to show that the hydraulic predictability properties observed in Network 1 are not the result of some peculiarities of Network 1. This discussion of Network 2 is short as the findings are generally similar to Network1. For example, with reference to Figure 8, it can be seen that the reliability of Node 3 for the design with the smallest entropy value is low relative to its location with respect to the source. On the other hand, for the maximum entropy design, the rank order of the nodal reliabilities (largest to smallest) is as follows: Nodes 2 and 4; Node 5; Nodes 7 and 3; Nodes 6 and 8; Node 9. This order mirrors the respective distances of the nodes from the source exactly, except that Node 5 has twice as many supply paths as Nodes 3 and 7 and so Node 5 has a higher reliability

Assessment Criteria	Network Element	Network Entropy				
		2.170	2.500	2.750	2.775	2.800
Total flow supplied	Critical node	3	7	6, 8	3, 7	6, 8
	Next critical node	6	8	9	6, 8	3, 7
Total energy dissipated per unit flow supplied	Critical node	8	6	5	6, 8	6, 8
	Next critical node	5	5	6, 8	5	5

Table 5: Critical nodes for fire-fighting demands replacing normal demands for Network 2

Assessment Criteria	Network Element	Network Entropy				
		2.170	2.500	2.750	2.775	2.800
Total flow supplied	Critical node	3	7	9	9	9
	Next critical node	6	8	6, 8	3, 7	6, 8
Total energy dissipated per unit flow supplied	Critical node	8	6	5	6, 8	6, 8
	Next critical node	5	5	6, 8	5	5

Table 6: Critical nodes for superimposed fire-fighting demands for Network 2

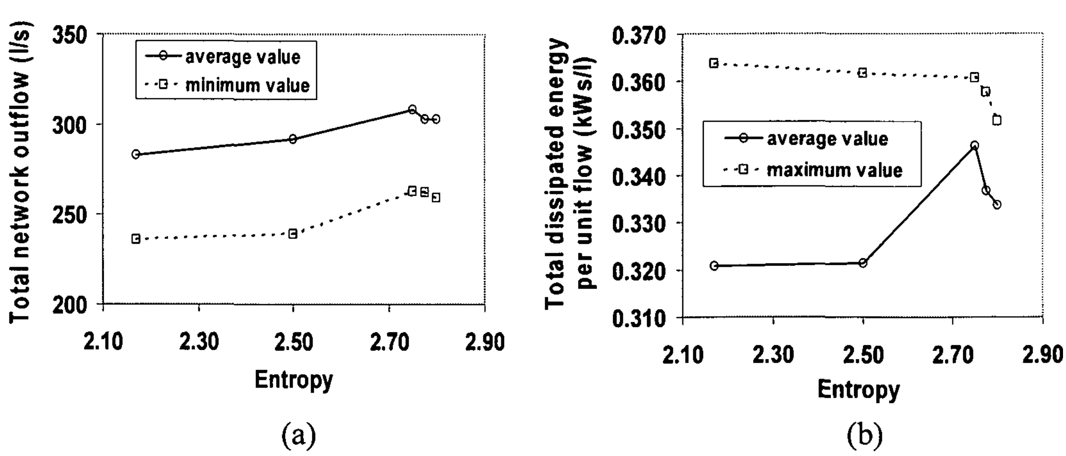


Figure 10: Total flow supplied and total dissipated energy per unit flow supplied for superimposed fire-fighting demands for Network 2

than Nodes 3 and 7. It may be noted that Node 9 has twice as many supply paths as Nodes 6 and 8. However, in addition to the greater distance of Node 9 from the source, its demand is three times larger. Once again these results, i.e. Figure 8, would appear to suggest that maximum entropy designs are hydraulically more predictable.

4 Conclusions

It has been shown that the entropy-constrained approach to the design optimization of water distribution systems has the potential to generate designs which have a high degree of hydraulic predictability. In general, the higher the entropy value, the greater the likelihood that the hydraulic properties of the design achieved will be intuitively obvious.

Results have been presented for two simple networks. Using pressure-dependent modelling and three simulated critical operating conditions including fire-fighting and pipe failure, the key nodes and distribution mains were identified. The criteria used included the nodal hydraulic reliability and pressure-dependent rate of flow delivery, and the deterioration in the hydraulic performance of the network as determined by the energy dissipation rate. A new hydraulic performance indicator has been introduced, based on the rate at which a water distribution network dissipates energy relative to the amount of water supplied.

The results presented would appear to be encouraging enough to warrant verification on more complex networks.

Acknowledgment

The authors are grateful to the UK Engineering and Physical Sciences Research Council for its financial support for the PRAAWDS component of this research under Grant GR/NO2672. Professor Andrew Templeman's contribution to the development of PRAAWDS is gratefully acknowledged.

References

[1] Xu, C., Goulter, I., "Reliability-based optimal design of water distribution networks", J. Water Resources Planning and Management, ASCE, 125(6), 352-362, 1999.

[2] Rowell, W.F., Barnes, J., "Obtaining layout of water distribution systems", J. Hydraulics Division, ASCE, 108(1), 137-148, 1982.

[3] Tanyimboh, T.T., Templeman, A.B., "Optimum design of flexible water distribution networks", Civil Engineering Systems, 10(3), 243-258, 1993.

[4] Tanyimboh, T.T., Tahar, B., Templeman, A.B., "Pressure-driven modelling of water distribution systems", Water Science and Technology - Water Supply, 3(1-2), 255-262, 2003.

[5] Tanyimboh, T.T., "Availability of water in distribution systems", in "Decision Support in the Water Industry under Conditions of Uncertainty", Walters, G.,

Savic, D., Khu, S-T., King, R. (Editors), Centre for Water Systems, Exeter, 159-164, 2004.
[6] Wagner, J.M., Shamir U., Marks D.H., "Water distribution reliability: simulation methods", J. Water Resources Planning and Management, ASCE, 114(3), 276-294, 1988.
[7] Tanyimboh, T.T., "Reliability analysis of water distribution systems", in "Urban and Rural Water Systems for Sustainable Development: Proceedings of the 30th IAHR Congress", Ganoulis J., Macsimovic, C., Kaleris, V. (Editors), ISBN 960-243-594-1, Thessaloniki, 321-328, 2003.
[8] Tanyimboh, T.T., Templeman, A.B., "Calculating maximum entropy flows in networks", J. Operational Res. Society, 44(4), 383-396, 1993.
[9] Tanyimboh, T.T., "An Entropy-based Approach to the Optimum Design of Reliable Water Distribution Networks", PhD thesis, University of Liverpool, 1993.
[10] Tanyimboh, T., Templeman, A., "A quantified assessment of the relationship between the reliability and entropy of water distribution systems", Engineering Optimization, 33(2), 179-199, 2000.

ELSEVIER

Advances in Engineering Software 36 (2005) 780–788

ADVANCES IN ENGINEERING SOFTWARE

www.elsevier.com/locate/advengsoft

Modelling errors, entropy and the hydraulic reliability of water distribution systems

Y. Setiadi, T.T. Tanyimboh*, A.B. Templeman

Department of Civil Engineering, University of Liverpool, Brownlow Street, Liverpool L69 3GQ, UK

Received 25 November 2003; accepted 17 March 2005
Available online 29 August 2005

Abstract

This paper reports on an investigation of the possible influence of modelling errors on the relationship between the entropy and hydraulic reliability of water distribution systems. The errors are due to minor differences between the design optimisation and subsequent simulation models, which lead to small discrepancies between the capacity of the network and the required supply. Pressure-dependent analysis was used for the hydraulic simulations. It is shown that any correlation between the redundancy or undercapacity due to the modelling errors and the hydraulic reliability is insignificant. The results, therefore, provide yet more evidence that the entropy-reliability relationship is strong.

Keywords: Entropy; Hydraulic reliability; Water distribution systems; Pressure-dependent analysis; Demand-driven analysis; Network analysis; Design optimisation

1. Introduction

Reliability analysis is an important component in the design, operation and maintenance of water distribution systems (WDS). Many researchers have tried to incorporate reliability in the design of WDS as one of the objectives to be optimised without unduly increasing the cost of the system. It has been shown, however, that the problem of calculating reliability exactly for a WDS is extremely difficult to solve [1]. As an answer to this problem, several researchers have proposed an alternative approach using a surrogate measure, e.g. entropy [2–5]. Entropy as a surrogate measure for the reliability of WDS has the computational advantages of being easy to calculate, minimal data requirements and ease of incorporation into optimisation procedures [6].

To analyse the hydraulic behaviour of water distribution systems, most previous studies used demand-driven simulation models, which assume that demands in the network are fully satisfied regardless of the pressure in the system. The models give acceptable results when the systems are subject to normal operating conditions. However, WDSs are subject to component failures or very large demands, which may result in a reduction of the pressure in the system. When this happens, demand-driven analysis (DDA) often gives results that indicate that the system is still supplying the full demand at lower, and sometimes, negative pressures. The validity of such results is obviously questionable. This study uses the head-dependent analysis (HDA) approach, which has been suggested to provide more realistic results when WDSs operate under subnormal pressure conditions [7–9]. However, due to the relative unfamiliarity of HDA in industry and the research community, DDA results have also been included for comparison and cross-checking purposes.

The main objective of this study was to assess critically the possible influence of modelling errors, namely redundancy or insufficient capacity, in the form of a small surplus or deficit in head at the critical node(s), on the relationship between the entropy and hydraulic reliability of water distribution systems. The small surpluses or shortfalls in the heads at the critical nodes are actually modelling errors, which are brought about by several factors as explained subsequently in this paper. The errors are

* Corresponding author. Tel.: +44 151 794 5229; fax: +44 151 794 5218.

E-mail address: ttanyi@liv.ac.uk (T.T. Tanyimboh).

0965-9978/$ - see front matter
doi:10.1016/j.advengsoft.2005.03.028

seemingly small and at first glance would appear to be insignificant. However, any impact they may have on the relationship between the entropy and hydraulic reliability has never been studied. The importance of this issue lies in the fact that hydraulic reliability values for alternative designs of water distribution systems are generally high which means that the differences in the reliability values tend to be very small. The very small range of variation in hydraulic reliability leaves open the possibility that seemingly small design and modelling errors may have a significant or disproportionate impact on the calculated relationship between the entropy and reliability. Due to the overall computational demands, a small network was used as the basis of the study. A range of minimum-cost designs with different entropy values were generated using non-linear programming. The hydraulic reliability and other properties of the designs were then evaluated to yield the data that were used subsequently for the various analyses herein.

A secondary objective of this paper is to strengthen the evidence in the literature that a relationship exists between the entropy and hydraulic reliability of water distribution systems using a combination of sensitivity analysis and head-dependent modelling as described next. The possible effects of different cost objective functions are being studied also in the present research, which is continuing. Different cost objective functions will produce different designs. These differences may have an effect on the entropy-reliability relationship and early results from this study based on the DDA approach have been reported [10]. The above-mentioned results on the cost function effects are verified herein using the more realistic HDA modelling approach, so as to validate and strengthen the background evidence that a strong relationship exists between the entropy and hydraulic reliability of water distribution systems.

Tanyimboh and Templeman [11] investigated the influence of the layouts of WDS on the relationship between entropy and reliability. Since the entropy of a WDS is a function of the pipe flow rates, different arrangements of the pipes in the distribution system may have a significant impact on the relationship between entropy and reliability. The HDA method has been used in this study to verify their results on the influence of the layout of the system on the relationship between entropy and hydraulic reliability [11], thereby producing yet more background evidence of a strong entropy-reliability relationship.

Thus, this paper has two inter-related aims. First, the results of an analysis of the possible influence of the modelling errors mentioned above on the correlation between entropy and hydraulic reliability are presented. The correlation between the hydraulic reliability and apparent over-/under-capacity and between the entropy and over-/under-capacity were assessed along with the relationship between entropy and hydraulic reliability. Second, background evidence of a strong relationship between entropy and hydraulic reliability is provided based on a re-evaluation of the layout and cost objective function effects using HDA. The main conclusion from this study is that the results appear to strengthen the notion that the relationship between entropy and reliability is strong, with HDA generally yielding much better correlation than DDA.

2. Literature review

Reliability is a good measure to assess the merit of a WDS. There is, however, no general agreement about the definition of reliability. Bao and Mays [12], for example, define reliability as a function of the probabilities of the heads at demand nodes being above the minimum required level. Fujiwara and De Silva [13], on the other hand, used the shortfalls in the flow delivered to measure WDS reliability. In this paper, reliability is defined as the time-averaged value of the ratio of the flow delivered to the flow required [6,14].

To evaluate the exact value of reliability, all possible failure scenarios have to be considered. This can be very complicated and tedious for large networks. Researchers have proposed the use of Monte Carlo simulation [12] and minimum cut set techniques [15] to help reduce the complexity of the calculations. The difficulties faced in calculating the reliability motivated researchers to find surrogate measures of reliability for WDS. One of the surrogate measures is entropy, which is based on Shannon's informational entropy [16]. Awumah et al. [2,3], who proposed several WDS entropy functions, first suggested the use of Shannon's entropy in WDS analysis as a surrogate measure for the reliability. Awumah and Goulter [4] also proposed that entropy could be incorporated in the optimisation formulation as a means of quantifying the reliability in the design of WDSs. Further work by Tanyimboh and Templeman [5,6,17] led to the entropy function used herein, which is well established. Tanyimboh and Templeman [5] suggested that higher flexibility of a distribution network could be achieved by maximising the entropy of the network. They also showed that the methodology could produce resilient designs without a substantial increase in cost.

Methods for calculating the most likely values of link flows in a WDS for which the available data are insufficient for a full hydraulic analysis have also been developed by Tanyimboh and Templeman [17,18]. Their algorithm for maximizing the entropy of single-source networks was generalised for multi-source multi-demand networks by Yassin-Kassab et al. [19]. The above works led to further investigation of the apparent relationship between the entropy and hydraulic reliability of WDS [11], which showed that high network reliability can be expected when the network carries maximum entropy flows. The possible use of entropy to find the optimum layout of a WDS has also been investigated by Tanyimboh and Sheahan [14]. Templeman and Yassin-Kassab [20] have suggested that entropy could be used in calibrating WDSs to find the most likely pipe characteristics in WDSs in which the roughness coefficients of the pipes have been lost or have changed with time. Recently, Ang and Jowitt [21] have explored the relationship between WDS energy loss and entropy to help gain a deeper understanding of the properties of entropy. As in an earlier study by Tanyimboh and Templeman [5], they concluded that the importance of a pipe in a water distribution network can be related to the amount of energy that the network dissipates following the removal or closure of that pipe.

3. Informational entropy

Shannon's entropy function [16] enables the levels of information or uncertainty of different probability distributions to be compared quantitatively and can be written as

$$S/K = -\sum_{i} p_i \ln p_i \quad (1)$$

in which S is the entropy, K is an arbitrary positive constant often taken as 1, p_i is the probability associated with the ith outcome.

The values of entropy in this paper were calculated using the entropy function for WDS [5,17], i.e.

$$\frac{S}{K} = -\sum_{j\in IN}(Q_j/T)\ln(Q_j/T) - \frac{1}{T}\sum_{j=1}^{J} T_j[(Q_j/T_j)\ln(Q_j/T_j) + \sum_{i\in N_j}(q_{ij}/T_j)\ln(q_{ij}/T_j)] \quad (2)$$

in which the subscript j represents all nodes including source nodes, while J denotes the number of nodes and IN the set consisting of source or input nodes. Also, T is the total supply, T_j is the total flow reaching node j, including any external inflow while N_j represents all nodes immediately upstream of and connected to node j. Q_j is the demand at demand nodes or supply at source nodes and q_{ij} is the flow rate in pipe ij. Thus the entropy of a WDS is basically a function of internal pipe flow rates only since the desired nodal demands and supplies are usually specified.

4. Optimal design of water distribution systems

The design optimisation was carried out using a Fortran program called PEDOWDS (Program for Entropy-based Design Optimization of Water Distribution Systems) [6], with the cost as the objective function to be minimized. The program is based on the NAG library routine E04UCF [22], which is a routine for constrained non-linear programming. E04UCF uses sequential quadratic programming and requires gradients of the objective and constraint functions. For every design, the design optimization program PEDOWDS was run using six different starting points. This was to ensure that the same optimum design was achieved several times, from different starting points in order to verify its optimality. The design that has the lowest cost with no violation of the constraints was then selected. A hydraulic analysis of the network was then carried out to further validate the result of the optimisation by verifying that no constraints were violated.

The cost function used in this study is

$$\text{Cost} = \gamma \sum_{ij} L_{ij} D_{ij}^{e} \quad (3)$$

in which e is the cost exponent that is specified by the user and thought to have a range of 1.0–2.5 [23]. The cost coefficient γ, whose value depends on a range of factors [24], is also specified by the user. L_{ij} and D_{ij} are the length and diameter of pipe ij, respectively. For the present study, it may be assumed that the costs are calculated in pounds sterling.

The constraints used in the optimisation consist of the pipe head loss, continuity, conservation of energy and entropy equations. Lower and upper limits of the pipe diameter and pipe flow non-negativity constraints were also used. The pipe head loss due to friction was calculated using the Hazen–Williams empirical formula, given as

$$h_{ij} = \alpha L_{ij}(q_{ij}/C_{ij})^{1.852}/D_{ij}^{4.87} \quad \forall\, ij \quad (4)$$

For the above equation, h_{ij} is the head loss along pipe ij and α is a constant, equal to 10.67 in SI units. The parameters L_{ij}, q_{ij}, C_{ij} and D_{ij} are the length, flow rate, roughness coefficient and diameter, respectively, for pipe ij. The continuity equation can be written as

$$\sum_{ij\in IJ_j} q_{ij} = Q_j \quad \forall\, j \quad (5)$$

in which the summation includes all the flows upstream and downstream of the node considered, indicated by IJ_j, and Q_j represents the external inflow or outflow at node j. Two sets of equations for conservation of energy were used, namely the loop and path equations. The loop equation requires the net head loss around each loop in a pipe network to be zero, i.e.

$$\sum_{ij\in l} h_{ij} = 0 \quad \forall\, l \quad (6)$$

in which l (l=1,2,...) represents the lth loop. The path equation ensures that the total head loss along any path is equal to the difference in head between the end points of that path. The equation may be written as

$$\sum_{ij\in p} h_{ij} = h_p \quad \forall\, p \quad (7)$$

where p ($p=1,2,\ldots$) represents the pth path having a known value of head loss, h_p.

Similarly, the equation used to ensure that nodal service pressures are satisfactory is

$$\sum_{ij\in t} h_{ij} \le H_s - H_t^{\mathrm{des}} \quad \forall\, t \tag{8}$$

where t ($t=1,2,\ldots$) represents a path from a specified source to a terminal node t, i.e. a node with no other nodes downstream of it, H_s is the head at a specified source and H_t^{des} is the desired head at terminal node t. The value of H_t^{des} may vary from one region to another. OFWAT (Office of Water Services for England and Wales), for example, specifies the value of H_t^{des} as 10 m on the customer's side of the main stop tap at the property boundary for a flow of 9 l/min. In practice, however, the pressure is difficult to measure at this point and water companies usually adopt a surrogate pressure head of 15 m in the adjacent water main serving the property [25]. OFWAT also stipulates that if the pressure drops below 7 m for more than once in a 28-day period and each occasion lasts for more than 1 h, then the customers affected are entitled to compensation.

As mentioned earlier, one of the advantages of entropy is that it can be incorporated into the optimisation procedure as one of the constraints. This entropy constraint may take the form

$$S/K \ge S_{\min} \tag{9}$$

in which K is an arbitrary constant which is set to unity in this work, S is the entropy and $S_{\min}$ is the minimum desired entropy value. This equation will ensure that the entropy of the network does not fall below the specified value of $S_{\min}$. $S_{\min}$ can take any value between zero and the maximum entropy value for the network under consideration, which depends on the chosen layout. The maximum entropy value is thus calculated in advance using special-purpose algorithms for calculating maximum entropy flows in networks [19]. These algorithms enable the maximum entropy value to be evaluated quickly without the need for any mathematical programming.

Apart from the constraints mentioned above, the diameters of the pipes should not lie outside the predetermined range. Within this range the value of the diameter is considered continuous to simplify the optimisation. However, an equivalent two-phase approach which yields segmental pipes based on commercial pipe sizes could be used instead [14]. The above-mentioned diameter restriction takes the following form

$$D_{\max} \ge D_{ij} \ge D_{\min} \quad \forall\, ij \tag{10}$$

in which D_{ij} is the diameter of pipe ij, while $D_{\max}$ and $D_{\min}$ are the maximum and minimum allowable pipe diameters, respectively.

5. Network analysis

Two methods of network analysis were used in this study for verification and cross-correlation purposes, the first being Demand Driven Analysis (DDA) and the other Head Dependent Analysis (HDA).

5.1. Demand driven analysis

This method of network analysis has been widely used in the water industry for many years. Unfortunately, there are a few disadvantages arising from the use of this method. DDA does not take into consideration the relationship between the nodal outflows and the pressure within the system. It assumes that the demands of the system are fully satisfied regardless of the pressure in the system. In consequence, when the pressure drops below the required level, network analysts would have no information on how much water would be delivered by the system under the available pressure regime. In this situation some customers would receive reduced supplies and, in the worst scenario, they might not receive any supply at all [7,8].

The drop in pressure in the distribution system can be triggered by many factors. Excessive abstraction at one demand node, for example, in a fire fighting situation, may cause the pressure in the neighbouring abstraction points to drop below the required level. The analysis using this method was carried out using a computer software called EPANET [26], which is freely available on the Internet and is provided by the US Environmental Protection Agency.

5.2. Head dependent analysis

Pressure dependent analysis has long been suggested to surpass demand driven analysis, particularly for networks under subnormal operating conditions. It is well known that outflows from a WDS are dependent upon the pressure within that system and, therefore, the DDA assumption that demands are always satisfied regardless of the pressure in the system is often inappropriate. HDA takes into consideration the pressure dependency of nodal outflows, and in consequence, the results are more realistic. Nevertheless, this method is not yet commonly used in the water industry since more research and verification of the true relationship between network pressure and nodal outflows are still necessary.

Some researchers have proposed several assumed head–outflow relationships. For example, Wagner et al. [27] suggested the following functions

$$\frac{Q_j^{\mathrm{avl}}}{Q_j^{\mathrm{req}}} = 0; \quad H_j < H_j^{\min} \tag{11a}$$

$$\frac{Q_j^{\mathrm{avl}}}{Q_j^{\mathrm{req}}} = \left(\frac{H_j - H_j^{\min}}{H_j^{\mathrm{des}} - H_j^{\min}}\right)^{1/n_j}; \quad H_j^{\min} \le H_j < H_j^{\mathrm{des}} \tag{11b}$$

$$\frac{Q_j^{avl}}{Q_j^{req}} = 1; \quad H_j \geq H_j^{des} \tag{11c}$$

in which Q_j^{avl} and Q_j^{req} are the actual outflow that can be delivered by the system and the required outflow or demand, respectively. H_j is the actual head at node j, H_j^{min} is the nodal head at node j below which there would be no outflow and H_j^{des} is the desired head at node j, above which the outflow would be equal to the demand. Values of the exponent parameter, n_j, are thought to lie between 1.5 and 2 [28]. Another example of the head–outflow relationship is the following function proposed by Fujiwara and Ganesharajah [29]

$$\frac{Q_j^{avl}}{Q_j^{req}} = \frac{\int_{H_j^{min}}^{H_j} (H_j - H_j^{min})(H_j^{des} - H_j^{min})\mathrm{d}H}{\int_{H_j^{min}}^{H_j^{des}} (H_j - H_j^{min})(H_j^{des} - H_j^{min})\mathrm{d}H}; \tag{12}$$

$$H_j^{min} \leq H_j < H_j^{des}$$

The HDA in this study was carried out using a similar head–outflow curve [30] in a FORTRAN computer program called PRAAWDS (Program for the Realistic Analysis of the Availability of Water in Distribution Systems) [8,9], which calculates the actual flow delivered under normal and subnormal pressure conditions.

6. Reliability calculation

In this paper, the reliability of a water distribution system is calculated by assuming a constant demand value. The reliability function can be written as [14]

$$R = \frac{1}{T}\left[p(0)T(0) + \sum_{m=1}^{M} p(m)T(m) + \sum_{m=1_{\forall n>m}}^{M-1} p(m,n)T(m,n) + \cdots\right] + \frac{1}{2}\left[1 - p(0) - \sum_{m=1}^{M} p(m) - \sum_{m=1_{\forall n>m}}^{M-1} p(m,n) - \cdots\right] \tag{13}$$

in which R is the system reliability, $p(0)$ is the probability that no pipe is unavailable, $p(m)$ is the probability that only pipe m is unavailable and $p(m,n)$ is the probability that only pipes m and n are unavailable. Similarly, $T(0)$, $T(m)$ and $T(m,n)$ are the respective total flows supplied with no pipes unavailable, only pipe m unavailable, and only pipes m and n unavailable. Finally, M is the number of pipes, while T represents the total demand. For DDA, $T(m)$ and $T(m,n)$ were calculated using the basic source head method [6,31,32]. Eq. (13) measures the reliability as the time-averaged value of the ratio of the flow delivered to the flow required [6]. This equation has two main components, shown using two pairs of square brackets. The first part of the equation corresponds to the basic definition of hydraulic reliability as stated above. The second part is a correction function whose value approaches zero as more and more multiple-component failure simulations are included in the first part [14].

7. Case study, results and discussion

The study was based on the network of Fig. 1. The source has a piezometric level of 100 m, while demand nodes have elevations of 0 m. The desired nodal service head for fully satisfactory performance is 30 m and the nodal head corresponding to zero nodal outflow is 0 m, this being the elevation of the nodes. All pipes are 1000 m long with a Hazen–Williams coefficient of 130. The above-mentioned desired head value H^{des} of 30 m was chosen to facilitate comparisons with previous DDA-based studies [11]. It is common to specify the value of the head corresponding to zero outflow as zero or the elevation of the node. However, any appropriate site-specific values may be used.

The maximum entropy value of the network was calculated using the special-purpose algorithms mentioned in Section 4. The optimisation program PEDOWDS was then used to generate the maximum entropy design as explained in Section 4. A traditional minimum-cost design, obtained using PEDOWDS without the entropy constraint,

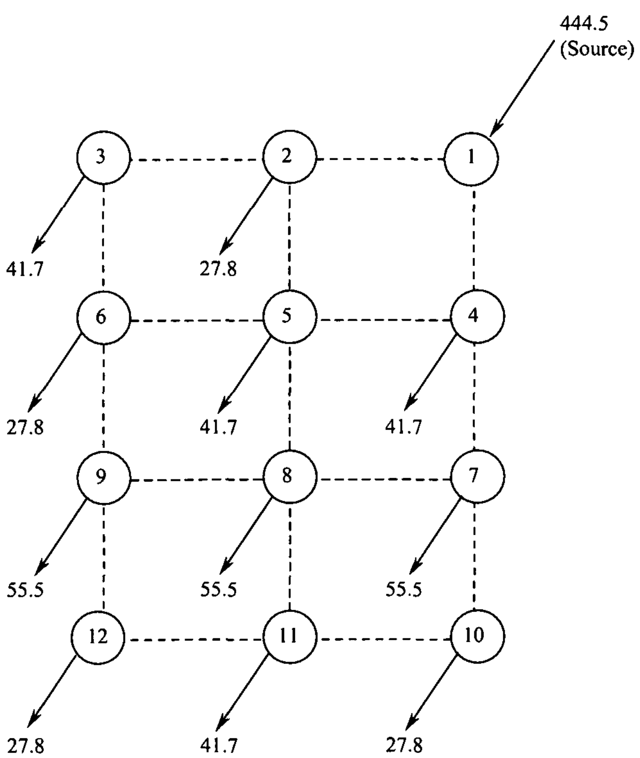


Fig. 1. Network of supply and demand nodes with demands in litres per second.

yielded the minimum entropy design. A range of designs with different entropy values between the maximum and minimum values were then produced. A total of 43 designs were generated in this way [10]. For each design, three to six different starting points were used, with a stipulation that the same cheapest design be obtained several times, to increase the chances of finding a global minimum. The rest of the designs in this study, taken from Refs. [11,14], were generated in a slightly different way. Instead of using the full set of pipes shown in Fig. 1, different combinations of those pipes were chosen. The maximum entropy design for each of the layouts obtained in this way was produced as in Section 4.Thus, a total of 31 maximum entropy designs were obtained from Refs. [11,14].

The programs used in the design optimisation process (PEDOWDS) and subsequent hydraulic simulations (EPANET and PRAAWDS) have slight differences in some of their coefficients and units. These differences, including the rounding off of the pipe diameters, which are continuous herein, produced 'small' surpluses or deficits in head at the critical nodes with the largest being −0.4 m. The critical node is the node with the lowest pressure-head in the system, whose value is specified as 30 m in this study for networks under normal operating conditions. The distribution of the surplus heads, which seems to follow the normal distribution, is illustrated in Fig. 2a. It shows that the heads at the critical nodes for most of the designs are equal or very close to zero. Fig. 2b and c, obtained from the HDA analysis, show the total outflows delivered by the full network and the distribution of the ratio of the actual to the required nodal outflow (i.e. demand satisfaction ratio), respectively. From the latter two graphs, it seems that, under normal operating conditions, the total outflow delivered by most of the designs is approximately equal to the total demand required. This would appear to suggest that the accuracy of the results is acceptable and that the small surpluses or deficits in heads and outflows are negligible.

EPANET and PRAAWDS were also used to carry out pipe failure simulations on each of the designs in order to obtain its hydraulic reliability value, as explained in Section 6. The possibility that redundancy or insufficient capacity in the WDS, in the form of the surplus or deficit in head at the critical node, might have a bearing on the observed relationship between entropy and reliability given the very small differences between the reliability values has not been addressed in previous studies.

First of all, however, to check the effect of the small differences in head at the critical nodes on the reliability values, the performance of designs with an excess or shortfall in capacity was compared to slightly adjusted designs, which satisfied the demands exactly. To achieve this, the head at the source, for all designs with a surplus or deficit in head at the critical node was artificially altered so that, at the critical node, the head was precisely equal to the desired service head of 30 m. These designs were then re-analysed and their reliability values obtained. The results

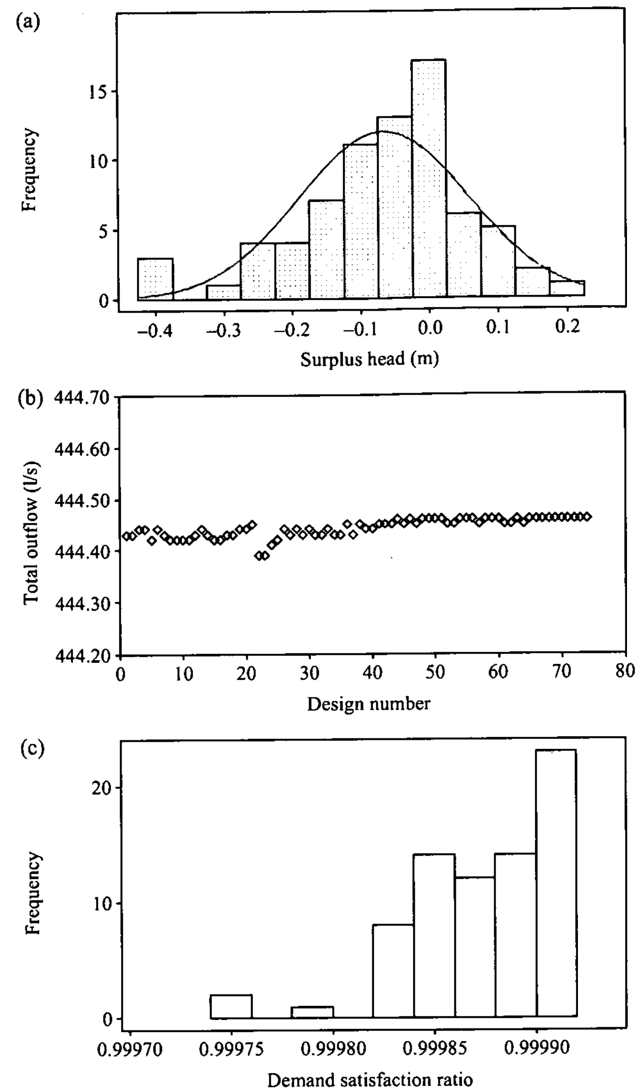


Fig. 2. (a) Distribution of surplus heads. (b) Total outflows delivered by the networks with all pipes available. (c) Distribution of demand satisfaction ratios.

of this analysis are shown in Fig. 3 in which the reliability values before and after the head modification for all the designs in this study are plotted against each other. It shows that all the designs have virtually identical pairs of

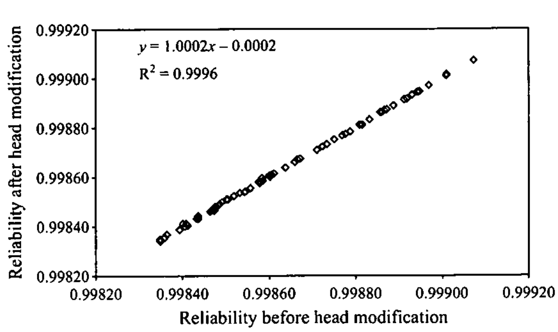


Fig. 3. DDA-reliability verification.

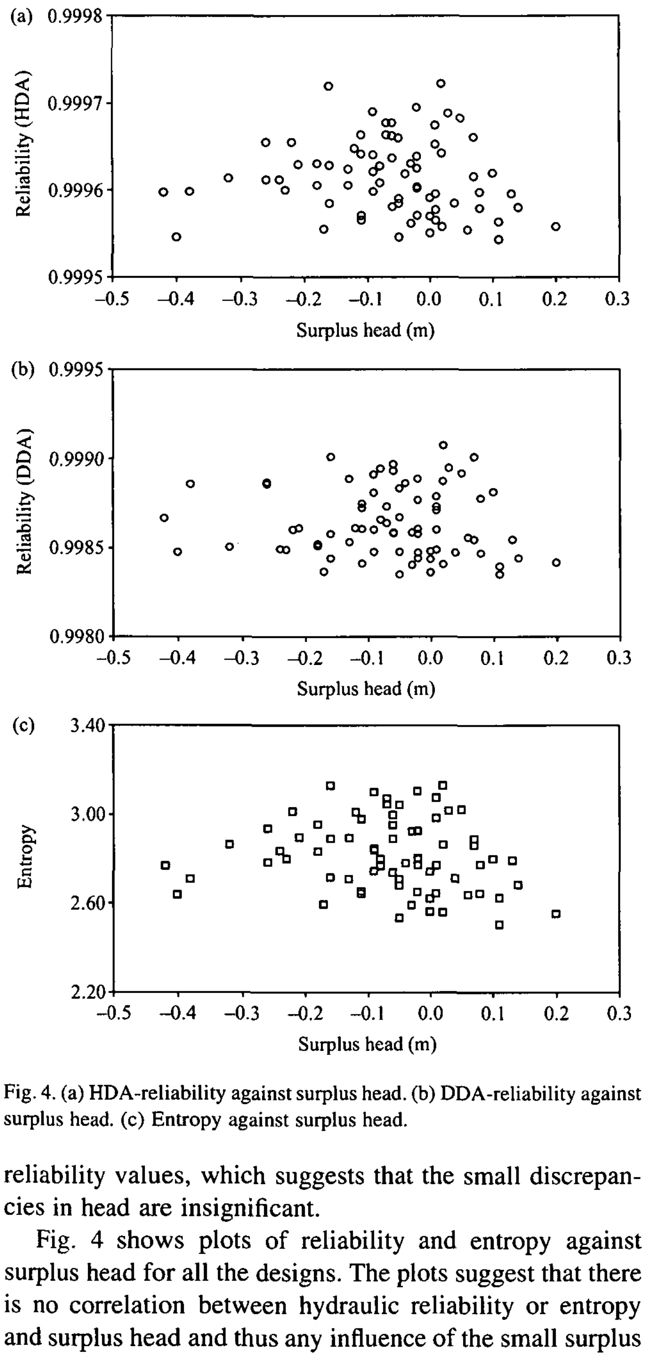


Fig. 4. (a) HDA-reliability against surplus head. (b) DDA-reliability against surplus head. (c) Entropy against surplus head.

reliability values, which suggests that the small discrepancies in head are insignificant.

Fig. 4 shows plots of reliability and entropy against surplus head for all the designs. The plots suggest that there is no correlation between hydraulic reliability or entropy and surplus head and thus any influence of the small surplus or deficit in head at the critical node upon the entropy-reliability relationship is insignificant. It may nevertheless be noted that the difference between the most reliable design and the least reliable design is only 0.000726 and 0.000180, for DDA and HDA, respectively.

Having demonstrated that the perceived relationship between entropy and hydraulic reliability is not attributable to modelling errors or WDS redundancy/overcapacity, the remainder of this section compares the entropy–reliability relationship as assessed on one hand by DDA and on the other hand by HDA. One reason for this is that it is vital to demonstrate that the conclusions reached are not purely the outcome of the HDA modelling approach.

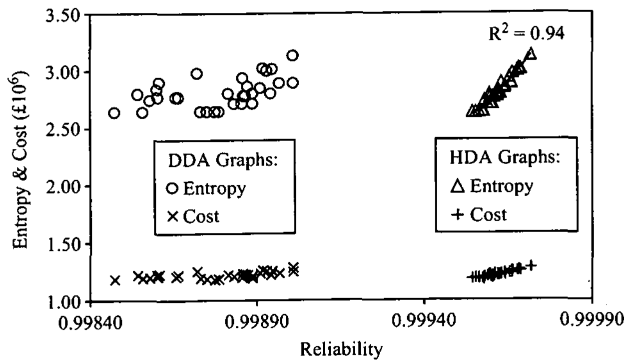


Fig. 5. Effect of layout on the entropy–reliability relationship for $\gamma=800$ and $e=1.5$.

The entropy of a WDS is a function of its pipe flow rates and the pipe flow rates are affected by the layout of the network, i.e. the configuration of the pipes. The effects of layouts on the relationship between entropy and reliability have been investigated previously using only DDA [11,14]. Pressure dependent analysis has been used herein to verify those results so as to provide more background evidence of the entropy–reliability relationship. More importantly, it is essential to demonstrate that the present assessments based on HDA give results which are consistent with previous investigations. Indeed it turns out, based on the evidence herein, that HDA not only confirms but also strengthens previous conclusions relating to the relationship between entropy and hydraulic reliability. Fig. 5 shows plots of entropy and cost against reliability for the maximum-entropy minimum-cost designs from Refs. [11,14].

To further ascertain that the conclusions reached herein are not merely the outcome of the HDA modelling approach, the 43 17-pipe 6-loop designs based on the full set of links shown in Fig. 1 [10] were re-analysed using HDA. Based on Eq. (3), 22 of the designs were generated using values of γ and e of 800 and 1.5, respectively. The entropy–reliability results for these designs can be seen in Fig. 6. The remaining 21 designs, for which the results are shown in Fig. 7, were generated using $\gamma=1600$ and $e=2.0$. It is worth emphasizing that the preliminary results reported in Ref. [10] did not involve HDA. Looking at the graphs of

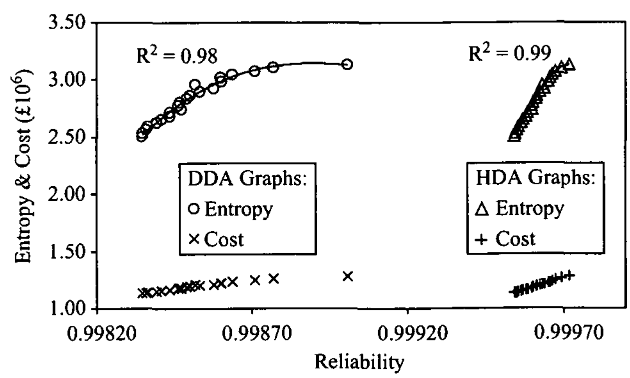


Fig. 6. Plots of entropy vs. reliability for $\gamma=800$ and $e=1.5$.

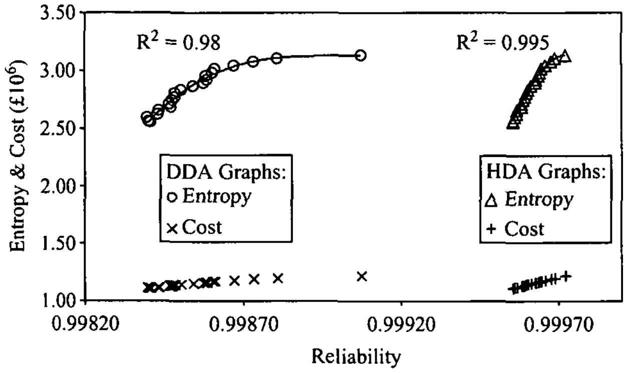


Fig. 7. Plots of entropy vs. reliability for $\gamma = 1600$ and $e = 2.0$.

entropy against reliability in Figs. 6 and 7, it is evident that the relationship between entropy and reliability remains strong despite the differences in the values of γ and e. The two plots of entropy against reliability follow a very similar pattern and, again, HDA shows a stronger relationship than DDA. This would appear to suggest that the observed entropy–reliability relationship is not unduly sensitive to the coefficient, e, of the cost objective function used herein, Eq. (3). It should be noted that the costs have been plotted twice in Figs. 5–7, even though the two cost data sets are identical, for completeness and ease of reference.

Overall, the relationship between entropy and reliability and the difference between HDA and DDA can be observed clearly in Fig. 8, in which the results for all the 74 designs studied herein are shown. It demonstrates that the correlation between entropy and reliability is much stronger when the analysis is done using HDA. In particular, the correlation coefficient value of $r^2 = 0.94$ for HDA is remarkable considering that 32 different layouts (i.e. 31 from Refs. [11,14] plus the six-loop layout of Fig. 1) and exponent, e, values of 1.5 and 2.0 were involved.

The rather simple sample network used herein is considered appropriate for this study which requires numerous calculations of the respective maximum entropy values, design optimization runs (each of which requires three to six runs of the optimization program, PEDOWDS) and evaluation of the reliability values. Furthermore, the reliability of each design was calculated twice, i.e. one value based on DDA and the other based on HDA. Each reliability value entails the hydraulic simulation of all the operational conditions required in Eq. (13). It could be pointed out also that, additionally, the DDA reliability values were calculated twice (Fig. 3). Therefore, perhaps it can be appreciated that the range of results in this study would be significantly more difficult to reproduce using a much larger network. To put this in perspective, the results reported in this paper would perhaps correspond to a laboratory-level study ahead of a full-scale field trial or pilot-type study. An advantage of the laboratory-level investigation is that it enables a large number of issues to be addressed quickly at minimal cost whilst tackling the research questions, for example, whether a concept is feasible, at least in principle. Many questions about the relationship between entropy and reliability have been tackled in this way, to the point where real-world demonstrations or applications now appear close. In this study, the network of Fig. 1 enabled useful results to be obtained while keeping the computations manageable.

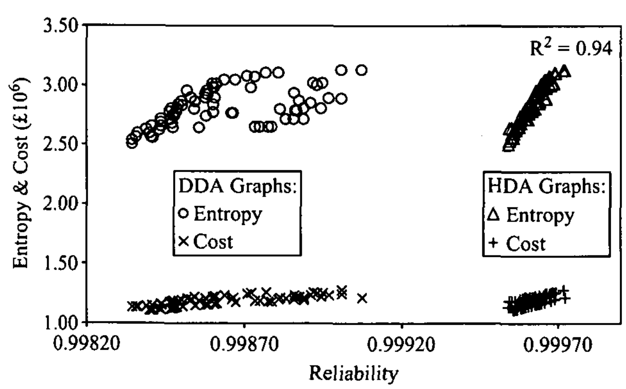


Fig. 8. Plots of entropy vs. reliability for all the designs.

8. Summary and conclusions

Recognising the importance of the reliability of WDS and the difficulties in obtaining its value, researchers strive to find alternative approaches to quantify the reliability. Entropy has been suggested to provide good representation of the reliability value for WDS. Its strengths stem from the ease of computation and, in the design of WDS, it can be easily incorporated into the optimisation procedure. In this study, the effects of modelling errors on the relationship between entropy and reliability have been investigated. The errors are small surpluses or deficits in the head at the critical nodes. Other aspects were examined, also, to provide further background evidence of a strong relationship between entropy and reliability using designs based on different layouts and cost objective functions. Results for a sample network were presented and discussed.

For the sample network, there seems to be a strong relationship between entropy and reliability. This relationship is stronger when the analysis is carried out using HDA in comparison to DDA. Small, unavoidable modelling errors, in the form of small surpluses or deficits in the pressure heads at the critical nodes, would appear to have no real influence on the entropy–reliability relationship. In particular, the new HDA results herein are very encouraging. They strongly support and reinforce previous DDA-based conclusions and would appear to suggest that there is some potential in the water distribution entropy research. While this study is based on a simple network, the results appear to warrant research on larger, more complicated networks, which may include components other than pipes.

Acknowledgements

The authors are most grateful to the UK Engineering and Physical Science Research Council for its financial support for the PRAAWDS component of this research under Grant GR/NO2672. The financial support from the University of Liverpool and Universities UK under the Overseas Research Students Awards scheme for Yohan Setiadi is also acknowledged. We would like to thank the referees whose comments helped to improve the presentation of the paper.

References

[1] Provan JS, Ball MO. The complexity of counting cuts and of computing the probability that a graph is connected. SIAM J Comput 1983;12(4):777–88.

[2] Awumah K, Goulter IC, Bhatt SK. Assessment of reliability in water distribution networks using entropy-based measures. Stoch Hydrol Hydraul 1990;4(4):325–36.

[3] Awumah K, Goulter IC, Bhatt SK. Entropy-based redundancy measures in water-distribution networks. ASCE J Hydraul Eng 1991;117(5):595–614.

[4] Awumah K, Goulter I. Maximizing entropy defined reliability of water distribution networks. Eng Optimiz 1992;20(1):57–80.

[5] Tanyimboh TT, Templeman AB. Optimum design of flexible water distribution networks. Civil Eng Syst 1993;10(1):243–58.

[6] Tanyimboh TT. An entropy-based approach to the optimum design of reliable water distribution networks. PhD Thesis, University of Liverpool, UK; 1993.

[7] Ackley JRL, Tanyimboh TT, Tahar B, Templeman AB. Head-driven analysis of water distribution systems. In: Ulanicki B, Coulbeck B, Rance J, editors. Water software systems: theory and applications, vol. 1, 2001. p. 183–92.

[8] Tanyimboh T, Tahar B, Templeman A. Pressure-driven modelling of water distribution systems. Water Sci Technol 2003;3(1–2):255–62.

[9] Tahar B, Tanyimboh TT, Templeman AB. Pressure-dependent modelling of water distribution systems. In: Khosrowshahi F, editor. Proceedings of third international conference on decision making in urban and civil engineering, London, 2002.

[10] Tanyimboh TT, Setiadi Y, Mavrokoukoulaki A, Storer C. The suitability of information-theoretic entropy as a surrogate performance measure for water distribution systems. In: Khosrowshahi F, editor. Proceedings of third international conference on decision making in urban and civil engineering, London, 2002.

[11] Tanyimboh TT, Templeman AB. A quantified assessment of the relationship between the reliability and entropy of water distribution systems. Eng Optimiz 2000;33(2):179–99.

[12] Bao Y, Mays LW. Model for water distribution system reliability. ASCE J Hydraul Eng 1990;116(9):1119–37.

[13] Fujiwara O, De Silva AU. Algorithm for reliability-based optimal design of water networks. ASCE J Environ Eng 1990;116(3):575–87.

[14] Tanyimboh TT, Sheahan C. A maximum entropy based approach to the layout optimization of water distribution systems. J Civil Eng Environ Syst 2002;19(3):223–53.

[15] Su Y, Mays LW, Duan N, Lansey KE. Reliability-based optimization model for water distribution systems. ASCE J Hydraul Eng 1987; 142(12):1539–56.

[16] Shannon C. A mathematical theory of communication. Bell Syst Tech J 1948;27(3):379–428.

[17] Tanyimboh TT, Templeman AB. Calculating maximum entropy flows in networks. J Oper Res Soc 1993;44(4):383–96.

[18] Tanyimboh TT, Templeman AB. Maximum entropy flows for single-source networks. Eng Optimiz 1993;22(1):49–63.

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

[20] Templeman AB, Yassin-Kassab A. Robust calibration of computer models of engineering networks with limited data. In: Parmee IC, Hajela P, editors. Optimization in industry. London: Springer; 2002. p. 71–81.

[21] Ang W, Jowitt PW. Some observation on energy loss and network entropy in water distribution networks. Eng Optimiz 2003;35(4): 375–89.

[22] NAG Ltd. NAG Fortran library manual. Mark 17. vol. 4. Oxford, UK: The Numerical Algorithms Group Ltd; 1995.

[23] Fujiwara O, Kang D. A two-phase decomposition method for optimal design of looped water distribution networks. Water Resour Res 1990; 26(4):539–49.

[24] Lansey K, Duan N, Mays L, Tung Y-K. Water distribution system design under uncertainties. ASCE J Water Resour Plann Manage 1989;115(5):630–45.

[25] Office of water services. Levels of service for the water industry in england and wales. 2002–2003 Report.

[26] Rossman LA. EPANET 2 users manual. Cincinnati: US Environmental Protection Agency; 2000.

[27] Wagner JM, Shamir U, Marks DH. Water distribution reliability: simulation methods. ASCE J Water Resour Plann Manage 1988; 114(3):276–94.

[28] Gupta R, Bhave PR. Comparison of methods for predicting deficient network performance. ASCE J Water Resour Plann Manage 1996; 122(3):214–7.

[29] Fujiwara O, Ganesharajah T. Reliability assessment of water supply systems with storage and distribution networks. Water Resour Res 1993;29(8):2917–24.

[30] Tanyimboh TT, Templeman AB. A new nodal outflow function for water distribution networks. In: Topping BHV, Mota Soares CA, editors. Proceedings of fourth international conference on engineering computational technology, paper 64, Lisbon, 2004.

[31] Tanyimboh TT, Templeman AB. Calculating the reliability of single-source networks by the source head method. Adv Eng Softw 1998; 29(7–9):499–505.

[32] Tanyimboh TT, Tabesh M, Burrows R. Appraisal of source head methods for calculating reliability of water distribution networks. ASCE J Water Resour Plann Manage 2001;127(4):206–13.

Joint Layout, Pipe Size and Hydraulic Reliability Optimization of Water Distribution Systems

Tiku T. Tanyimboh[1] and Yohan Setiadi[2]

Abstract

A multicriteria maximum entropy approach to the joint layout, pipe size and reliability optimization of water distribution systems is presented in this paper. The capital cost of the system is taken as the principal objective, and so the trade-offs between cost, entropy, reliability and redundancy are examined sequentially in a large population of optimal solutions. The novelty of the method stems from the use of the maximum entropy value as a preliminary filter, which screens out a large proportion of the candidate layouts at an early stage of the process before the designs and their reliability values are actually obtained. This technique, which is based on the notion that the entropy is potentially a robust hydraulic reliability measure, contributes greatly to the efficiency of the proposed method. In addition, maximum entropy designs help reveal the optimum performance of the candidate layouts and so promote a like-with-like comparison of the layouts. The use of head dependent modeling for simulating pipe failure conditions in the reliability calculations also complements the method in locating the true Pareto-optimal front. The computational efficiency, robustness, accuracy and other advantages of the proposed method are demonstrated by application to a sample network.

Keywords: entropy, head-dependent modeling, multicriteria design optimization, reliability, water distribution system

[1] Senior Lecturer, Dept of Civil Eng., University of Strathclyde, Glasgow G4 0NG, UK. (Corresponding author). Tel: + 44 (0)141 548 4366. Fax: + 44 (0)141 553 2066. Email: tiku.tanyimboh@strath.ac.uk.
[2] Ph.D. Student, Dept of Eng., University of Liverpool, Liverpool L69 3GQ, UK. Email: y.setiadi@liv.ac.uk.

INTRODUCTION

It is widely accepted that the supply of water in urban areas should be available on demand. However, a water distribution system (WDS) is a network with many components that are subject to random failures. To help increase the reliability of supply, WDSs usually have loops to reduce the possibility of some of the demand points being separated from the rest of the network if a pipe is not in service. A WDS can suffer from mechanical and hydraulic failure. Mechanical reliability measures the probability that the component or system being considered is operational at any time while hydraulic reliability is a measure of the probability that the system can supply enough water at the right pressure. Although hydraulic reliability depends on mechanical reliability, it is largely governed by the hydraulic performance of the network. In turn, the performance depends on the layout and capacities of the pipes; the locations and capacities of storage facilities and pumps; the spatial and temporal variations in supply and demands; and the locations of valves and other appurtenances. The above characterization of reliability may appear straightforward. In practice, however, the issue is so complicated that there is no universally adopted practical definition of reliability in the context of WDSs.

Because the reliability of a WDS is inherently linked to its layout, reliability considerations significantly increase the complexity of simultaneously optimizing the layout and components of a WDS. The literature is replete with models for the design of WDSs, but they generally do not optimize the hydraulic reliability explicitly. Joint layout and pipe-size optimization with regard to reliability is extremely complex and there is no completely satisfactory model for solving this problem. Furthermore, virtually no layout optimization models incorporate hydraulic reliability in a formalized way. Therefore, this gives added impetus to the search for a quantified surrogate for reliability and, fortunately, extensive research based on head-dependent modeling has demonstrated the potential of entropy as a robust hydraulic reliability measure (Setiadi et al., 2005).

The aim of this paper is to demonstrate the effectiveness of the maximum entropy (ME) approach to the joint layout, pipe-size and reliability optimization of WDSs. Some of the properties worth highlighting include simplicity, computational efficiency, robustness, diversity among the non-dominated solutions and the ability to locate the true Pareto-optimal (PO) front. A key feature of the technique is that it operates on a small portion of the solution space and, as such, its efficiency is not overly affected by the size of the WDS. The main difference from an earlier demand-driven analysis (DDA) study (Tanyimboh and Sheahan, 2002) is that the present work used head-dependent analysis (HDA) and a vastly expanded population of solutions. The HDA results herein show that the non-dominated solutions found are virtually identical to the PO front. The rest of this paper includes the following sections: literature review; description of proposed approach; overviews of entropy, reliability, HDA and ME pipe-size optimization; demonstration and discussion of proposed approach.

LITERATURE REVIEW

Layout optimization models include those that begin with a spanning tree and then add loop-completing links in an attempt to meet some reliability criteria while minimizing the cost of the network. Other models start with the candidate links of the network and then remove some of these links in an attempt to reduce the cost while satisfying some reliability criteria. Rowell and Barnes (1982) used two main steps. In Step 1, an optimal spanning tree layout was identified and its pipe sizes determined by solving a non-linear minimum cost flow problem. In Step 2, loop-completing pipes were added in order to provide an alternative supply path to each demand node. However, the designs were not hydraulically consistent (Goulter and Morgan, 1984). Loganathan et al. (1990) adopted a conceptually similar approach and assured hydraulic consistency by redesigning the WDS following the addition of the loop-completing links. The loops were completed using minimum-diameter pipes whose usefulness is questionable (Tanyimboh and Templeman, 1993a).

Kessler et al. (1990) used the concept of two trees to design a WDS that would be invulnerable to any single-pipe failure. Two spanning trees were selected with graph theory algorithms such that they overlapped and ensured the existence of an alternative path to each demand node following any single-link failure. Invulnerability was provided by designing each tree so that, on its own, it could supply all the demands of the WDS. However, the method was only applicable to single-source networks and there was no means of determining the best pair of trees prior to a full design and evaluation of all the possible pairs of trees.

The models just described first select a core tree, then add loop-completing links while the models considered next start with the potential links, and then eliminate those that are less cost effective. Awumah et al. (1989) developed a heuristic integer-programming model. The main constraint, from a reliability point of view, was that each node had to be connected by at least two links. However, the reliability was not quantified. Morgan and Goulter (1985) developed an LP-based heuristic with the advantage that, for a given network, the number of variables and constraints remained constant no matter the number of demand patterns considered explicitly in the design. The method enhanced the resilience of the WDS by using a multiplicity of flow patterns. However, the formulation did not have an in-built capability for removing unwanted pipes.

In Afshar et al. (2005), a minimum-cost design having all the candidate pipes was obtained using minimum-diameter constraints. Then, heuristics were used to identify the pipes to be considered for removal. The minimum-diameter constraints for these pipes were then relaxed and the WDS redesigned to allow the pipe-size optimization program to remove the chosen pipes as appropriate. However, the method did not consider the hydraulic reliability of the network explicitly. Awumah and Goulter (1992) presented a non-liner programming (NLP) model based on an entropy-type function. The optimization process started with an upper bound upon the capital cost so large that it would be inactive at the solution. Then the model was re-run for successively lower allowable costs until no more cost reductions were possible. However, the method seemed unable adequately to preserve loops (Tanyimboh, 1993). Jacobs and Goulter (1989) developed a layout-only model by applying graph theory principles. However, costs and hydraulic considerations were not addressed.

The literature is replete with models for optimizing the design of WDSs. The overwhelming majority do not optimize the hydraulic reliability explicitly and those that consider the hydraulic reliability often do not optimize the layout. The present review does not include "components sizing only" models due to the limitations of space. Usually, the layout is pre-specified based on a number of considerations including practicability and cost. The model in Su et al. (1987) was an NLP formulation in which the conservation of mass and energy equations were satisfied using a WDS simulation model. The reliability was calculated by the minimum cut set method and defined as the probability of having sufficient flow and pressure. However, the use of simulation as part of the optimization process required a large amount of computer time. In addition, the gradients of the reliability function were calculated using finite differences. The computational expense of calculating these gradients can be considerable, as the evaluation of the reliability requires a large number of pipe-failure simulations. The model of Cullinane et al. (1992) was an NLP formulation similar to Su et al. (1987) at the conceptual level. In Fujiwara and De Silva (1987), reliability was optimized using an LP-based heuristic. However, the conservation of energy equations were ignored in the reliability calculations.

DESCRIPTION OF LAYOUT, PIPE SIZE AND RELIABILITY OPTIMIZATION METHOD

There is strong evidence which suggests that, for a fixed layout, the hydraulic reliability increases as the informational entropy of the pipe flow rates increases (Setiadi et al., 2005). Tanyimboh and Templeman (1993a) recommended that WDSs be designed to carry ME flows and the reasons are that: (a) Maximum-entropy WDSs would be more reliable than traditional minimum-cost designs. (b) ME designs would appear not to be unduly expensive. (c) ME designs are computationally easier to produce because the flows are calculated first and then the pipes sized (Tanyimboh and Sheahan, 2002). ME flows can be calculated readily using an efficient algorithm that does not require formal optimization methods (Yassin-Kassab et. al., 1999). The algorithm is non-iterative and only requires the solution of a system of non-linear equations. Pipe sizing for a WDS is much simpler if the pipe flow rates are fixed in advance (Alperovits and Shamir, 1977); *inter alia*, the number of decision variables is more or less halved and the nodal flow continuity equations removed.

The unique advantage that ME designs have in layout optimization is that the ME value can be calculated using only the nodal demands and network connectivity. Pipe lengths, diameters and roughness coefficients are not required. This permits screening of the layouts using the ME values only. Also, ME designs help reveal the optimum performance of the candidate layouts and so promote a like-with-like comparison of the layouts. The design of a WDS is a multicriteria optimization problem, and a natural solution strategy is to generate a diverse population of non-dominated designs. By taking the capital cost of the WDS as the principal objective, the trade-offs between cost, entropy, reliability and redundancy are examined sequentially. If each solution in a set of solutions is better than the rest of the solutions in at least one objective, then the solutions are mutually non-dominated. The non-dominated set of solutions is PO if it includes all the non-dominated solutions of the problem. The overall approach can be summarized as follows:

1. Generate the candidate layouts. In theory, this is generally a large combinatorial problem. In practice, however, the options are often limited by practical considerations. This aspect might be addressed using any suitable technique including graph theory. However, it is not addressed herein as there is not enough space.
2. (a) Calculate the ME flows and ME value for each layout.

   (b) Using the ME values -- while recognizing the importance of diversity among the ultimate non-dominated solutions -- screen out a portion of the layouts as appropriate.
3. For the layouts not discarded in 2b, size the pipes to carry the ME flows found in 2a. This can be done by LP or other methods e.g. genetic algorithms.
4. Using only the cost and entropy, identify the cost-entropy non-dominated (CEND) designs. Discard the rest of the designs.
5. Calculate the hydraulic reliability and redundancy of the CEND designs.
6. Using only the cost and hydraulic reliability, identify the cost-reliability non-dominated (CRND) designs. Discard the rest of the designs.
7. Using only the cost and hydraulic redundancy, identify the cost-redundancy non-dominated designs. Discard the rest of the designs.

INFORMATIONAL ENTROPY FUNCTION FOR WATER DISTRIBUTION SYSTEMS

Shannon (1948) derived the informational entropy function as a quantitative measure of the amount of uncertainty that a probability distribution represents. For an exhaustive probability scheme with mutually exclusive events (i.e. $p_1 + p_2 + ... + p_n = 1$; n = number of events), Shannon's entropy function is

$$S = -\sum_{i=1}^{n} p_i \ln p_i \tag{1}$$

in which S is the entropy and p_i is the probability of the ith event.

Several applications of entropy in WDSs have been reported (Setiadi et al., 2005). Tanyimboh and Templeman (1993a, 1993b) developed the WDS entropy function in a framework that enabled pipe flow rates to be interpreted as probabilities. For example, the particles of water in a WDS follow different paths through the system and the probability that a particle travels through a particular path depends on the flow rates in the pipes in the path. For a network in which the pipe flow rates and directions are known, the entropy function is

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

where S = WDS entropy; S_0 = entropy of source supplies; S_i = entropy of node i; $P_i = T_i/T$ = fraction of the total flow through the network that reaches node i; T_i = total flow reaching node i; T = sum of the nodal demands; N = number of nodes in the network;

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

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

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

where Q_{ij} = outflow ij from node i; ND_i represents the outflows, including any demand, at node i. For any given WDS, the entropy value depends solely on the pipe flow rates since the nodal demands are usually specified. When computerized, Eqs. (2-4) are very easy and rapid to evaluate; the calculations are easily done by hand for small networks (Tanyimboh and Templeman, 1993b, c). This simplicity and computational speed are in part the reason for the research into the various applications of entropy in WDSs. The WDS entropy (Eq. 2) is a measure of the uniformity of the pipe flow rates (Tanyimboh, 1993). For example, for any node, S_i (Eq. 4) attains its maximum value if the outflows, including any demand, Q_{ij} are identical. Similarly, the maximum value of Eq. (1) is $S = ln(n)$, which corresponds to the uniform probability distribution, i.e. $p_1 = p_2 = \ldots = p_n = 1/n$.

HYDRAULIC RELIABILITY CALCULATION

There is no universally agreed definition for the hydraulic reliability of WDSs. Herein, the definition used is a measure of the system's ability to satisfy the nodal demands and is taken as the time-averaged value of the ratio of the flow delivered to the flow required (Tanyimboh et al., 2001). By assuming a constant demand value, this can be written as (Tanyimboh and Sheahan, 2002)

$$R = \frac{1}{T}\left(p(0)T(0) + \sum_{m=1}^{M} p(m)T(m) + \sum_{m=1}^{M-1}\sum_{n=m+1}^{M} p(m,n)T(m,n) + \cdots \right)$$
$$+\frac{1}{2}\left(1 - p(0) - \sum_{m=1}^{M} p(m) - \sum_{m=1}^{M-1}\sum_{n=m+1}^{M} p(m,n) - \cdots \right) \qquad (5)$$

in which R = hydraulic reliability; M = number of links (i.e. pipes, pumps and valves); $p(0) = a_1a_2a_3...a_M$ = 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; $u_m = 1 - a_m$; $p(m, n) = p(0)(u_m/a_m)(u_n/a_n)$ = 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.

Eq. (5) has two main parts, i.e. the terms associated with the two pairs of large parentheses. The first part corresponds to the proportion of the total demand that the system satisfies on average. However, it is often impracticable to simulate all the configurations of the WDS with multiple components out of service when calculating the hydraulic reliability. Consequently, the calculation of the first part of Eq. (5) generally underestimates the reliability in practice. The second part of Eq. (5) is an estimate of the amount by which the first part underestimates the reliability. Eq. (5) has been verified numerically (Tanyimboh and Tabesh, 1997) while a full derivation can be found in Tanyimboh and Sheahan (2002). Herein, the nodal demands were taken as constants. In practice, however, water consumption varies with time. The incorporation of variations in demands is currently an area of active research (Surendran et al., 2005) and, as yet, no entirely satisfactory and/or easy-to-use model has been developed. The hydraulic reliability values for the sample network described later were calculated using Eq. (5) with assumed pipe availability (a_m) values based on a formula in Cullinane et al. (1992).

Another measure of reliability is the hydraulic redundancy, which is the time-averaged value of the fraction of the total demand satisfied when one or more components are out of service. It is thus a measure of resilience or invulnerability. Given R and $p(0)$, the evaluation of the hydraulic redundancy, FT, is simple, i.e.

$$FT = \frac{R - p(0)T(0)/T}{1 - p(0)} \qquad (6)$$

The derivation and further characterization of Eq. (6) can be found in Tanyimboh and Templeman (1998) while Kalungi and Tanyimboh (2003) have shown the importance of assessing WDSs using FT in addition to R.

HEAD DEPENDENT MODELING

A WDS may not be able to satisfy all of the nodal demands in full if there is insufficient pressure, for example, if the demands exceed the capacity of the network or some of the components are not in service. Unfortunately, the conventional DDA approach is, in general, unable to cope with such situations and the results can be misleading (Tanyimboh et al., 2003). Recognizing the relationship between the nodal outflows and heads under pressure-deficient conditions, an alternative approach, often referred to as *head-dependent* analysis (HDA), aims to determine the actual nodal outflows that the system can provide. Therefore, using the flow continuity equations, nodal head-outflow relationships can be incorporated in the system of equations for the network as follows:

$$F_i(H_i,H_j)=\sum_{j\in N_i}Q_{ij}-Q_i(H_i)=0; \qquad i=1, \ \ldots, \ N-1 \tag{7}$$

where H_i and H_j = piezometric heads at nodes i and j, respectively; N_i represents all the nodes connected to node i; Q_{ij} = flow in link ij (i.e. pipe, pump or valve) expressed in terms of H_i and H_j; $Q_i(H_i)$ = head-dependent outflow at node i. A commonly used relationship proposed by Wagner et al. (1988) is

$$\frac{Q_i(H_i)}{Q_i^{req}}=\left(\frac{H_i-H_i^{\min}}{H_i^{des}-H_i^{\min}}\right)^{0.5}; \qquad H_i^{\min}<H_i<H_i^{des} \tag{8}$$

where Q_i^{req} = demand at node i; H_i^{min} and H_i^{des} = piezometric heads at node i below which there would be no outflow and above which the outflow would be equal to the demand, respectively. $Q_i(H_i) = 0$ if $H_i \le H_i^{min}$; $Q_i(H_i) = Q_i^{req}$ if $H_i \ge H_i^{des}$. HDA was used to simulate the pipe closures (unavailabilities) for the reliability calculations herein (Eq. 5). This was essential because the minimum-cost designs (described in the next section) did not have any spare capacity and any pipe closures resulted in pressure-deficient conditions. A prototype Fortran computer program called PRAAWDS (Tanyimboh et al., 2003) was used for the post-design hydraulic simulations.

MAXIMUM ENTROPY PIPE SIZE OPTIMIZATION

The approach used to generate minimum-cost maximum-entropy (MCME) designs has two steps (Tanyimboh and Sheahan, 2002). Step 1 involves the calculation of the ME pipe flows rates -- the decision variables -- for the given nodal demands, layout and flow directions, along with the ME value. This step can be summarized as

Problem 1 - Maximize entropy:

$$\underset{\forall Q_{ij}}{Maximize} \ S=S_0+\sum_{i=1}^{N}P_iS_i \tag{2}$$

Subject to:

$$\sum_{j\in N_i}Q_{ij}=Q_i; \quad \forall i \tag{9}$$

$$Q_{ij}\ge 0; \quad \forall_{ij} \tag{10}$$

where the Q_i are the inflows (for source nodes) or demands (for demand nodes). The entropy (Eq. 2) is maximized subject to nodal flow continuity (Eqs. 9). Problem 1 is convex because the objective function is a summation of concave functions and the constraints are linear. Consequently, there is a unique set of ME flows for any given network and specified flow directions (Tanyimboh, 1993; Tanyimboh and Sheahan, 2002). As explained earlier, Yassin-Kassab et al. (1999) have proposed an efficient algorithm for solving Problem 1.

In Step 2, the pipes are sized to carry the ME flows Q^*_{ij} obtained in Step 1. This step can be summarized in general terms as follows.

Problem 2 - Minimize cost:

$$\underset{\forall D_{ij}}{Minimize} \quad Cost = \gamma \sum_{ij} L_{ij} D_{ij}^{e} \tag{11}$$

Subject to:

$$h_{ij} = \alpha L_{ij} (Q_{ij}^{*} / C_{ij})^{1.852} / D_{ij}^{4.87} \qquad \forall ij \tag{12}$$

$$\sum_{ij \in l} h_{ij} = 0 \qquad \forall l \tag{13}$$

$$\sum_{ij \in p} h_{ij} = h_p \qquad \forall p \tag{14}$$

$$\sum_{ij \in t} h_{ij} \le H_s - H_t^{des} \qquad \forall t \tag{15}$$

$$D_{\max} \ge D_{ij} \ge D_{\min} \qquad \forall ij \tag{16}$$

In the above problem e = constant (taken as 1.5 herein; Fujiwara and Khang, 1990). The cost coefficient γ was taken as 800. It may be regarded as a scaling parameter and has no real effect on the results. α = dimensionless conversion factor (10.67 in S.I. units); C_{ij}, D_{ij}, h_{ij}, L_{ij} and Q_{ij}^* = Hazen-Williams roughness coefficient, diameter (the decision variable), headloss, length and ME flow rate, respectively, for pipe *ij; p (p = 1, 2, ...)* represents the *p*th path having a known value of headloss, h_p; *l (l = 1, 2, ...)* represents the *l*th loop; *t (t = 1, 2, ...)* represents a path from a specified source to a critical or terminal node *t*; H_s = head at a specified source; H_t^{des} = minimum allowable head at critical or terminal node *t*; D_{min} = minimum allowable pipe diameter (100 mm); D_{max} = maximum pipe diameter (600 mm). The objective function (Eq. 11) is the cost of the system; Eq. (12) is the Hazen-Williams pipe headloss formula; Eqs. (13-14) are for the conservation of energy; Eq. (15) ensures the heads at demand nodes are high enough. The solution to Problem 2 yields the cost and pipe diameters D_{ij} of a particular design. It must be emphasized that, as presented, Problems 1 and 2 are general-purpose descriptions and, as such, do not conform to the requirements of any particular approaches for their computational solutions. Problem 2 was solved using a prototype Fortran computer program called PEDOWDS, which uses sequential quadratic programming (Tanyimboh, 1993; Setiadi et al., 2005).

SAMPLE NETWORK AND RESULTS

The network chosen to illustrate the proposed technique is shown in Figure 1a. This grid-type network inherently lends itself to layout optimization and entropy maximization studies. Previous related studies on this network include Setiadi et al. (2005); Tanyimboh and Sheahan (2002); Tanyimboh and Templeman (2000); and Awumah et al. (1991). All pipes have a length of 1000 m and Hazen-Williams roughness coefficient of 130; nodal elevations are all zero. The nodal piezometric heads above which demands are satisfied in full are H_i^{des} = 30 m and below which outflows cease are H_i^{min} = 0 m; the head at the source node is H_s = 100 m. The values of the other relevant parameters are as stated at the end of the previous section.

For the full set of links in Figure 1a, there are 65 fully looped layouts. For each fully looped layout, an MCME design was produced as described in Problems 1 and 2 based on an assumed set of flow directions. The layouts, flow directions and other details can be found in Tanyimboh and Sheahan (2002). 72 additional designs were generated herein with alternative flow directions. Flow directions in a WDS are generally not predictable without a hydraulic simulation of the WDS. However, many design methods, including the one used in this study, require the flow directions in the WDS to be specified *a priori*. These pre-specified flow directions would affect the resulting design in terms of cost, performance and ME value. The identification of the optimum designs under a wide range of layouts and flow directions is therefore highly desirable. The additional flow directions were generated based on three layouts only, these being Figures 1a, 1b and 1c. Based on the flow directions in Tanyimboh and Sheahan (2002), Figure 1a had the largest ME value (S = 3.12900), Figure 1b the median value (S = 2.67122) and Figure 1c the next smallest (S = 2.49454). These layouts were chosen based on an assumption that it would be impracticable to identify, design and calculate the hydraulic reliability values for all the feasible flow directions of the 65 layouts. The additional flow directions were chosen somewhat arbitrarily provided they were feasible. For example, all demand nodes should be reachable.

After calculating the ME flows (Problem 1), the computational solution of the pipe sizing problem (Problem 2) was obtained using PEDOWDS as explained in the preceding section. As in Tanyimboh and Sheahan (2002), the NLP and continuous diameter approach was used to reduce the overall computational effort considering both the design phase and subsequent reliability calculations. Several measures were used to increase the confidence in the optimality and hydraulic feasibility of the solutions found, e.g. multiple starting points in the design phase (Problem 2) and subsequent hydraulic simulations to reveal any surplus heads at the terminal nodes (for suboptimal designs) or insufficient nodal heads (for infeasible designs).

The reliability and redundancy values were then calculated as described in Eqs. (5) and (6). For each design, HDA analyses were performed for the WDS with all pipes in service and for each degraded configuration with one pipe out of service. Due to the small size of the network and by virtue of the error correction (i.e. second) term of Eq. (5), cases with two or more pipes out of service were not considered as in previous studies. Figure 2 is a plot of the costs against the hydraulic reliability values, which shows the non-dominated sets and true PO front. For comparison purposes, Figure 3 shows the updated (HDA) reliability values for the Tanyimboh and Sheahan (2002) flow directions. Figure 4 takes the analysis a step further and shows the costs versus the hydraulic redundancy values. Only 9 (81.8%) of the 11 CRND designs are cost-redundancy non-dominated. This demonstrates the importance of assessing the redundancy explicitly, in addition to the reliability (Kalungi and Tanyimboh, 2003). For comparison purposes, the values for all the designs including the non-dominated sets are shown. The layouts and flow directions corresponding to the 9 'best' designs obtained are shown in the appendix in Figure A1. The pipe diameters and additional details can be found in Setiadi (2006).

The proposed approach was assessed using the "generational distance" (GD) which measures the average distance between the non-dominated solutions and the true PO front (Deb, 2001), i.e.

$$GD = \frac{1}{NS}\left(\sum_{i=1}^{NS} d_i^2\right)^{1/2} \tag{17}$$

where NS = number of non-dominated solutions; d_i = Euclidean distance, in the space of the objective functions, between the ith non-dominated solution and the nearest solution in the true PO front;

$$d_i = Min\left\langle\left[\sum_{m=1}^{2}\left(f_m^{(i)} - f_m^{*(k)}\right)^2\right]^{1/2} \quad ; \quad k = 1, ..., NP\right\rangle \tag{18}$$

$f_m^{(i)}$ = value of the mth objective function for the ith non-dominated solution; $f_m^{*(k)}$ = value of the mth objective function for the kth PO solution; NP = number of PO solutions. The GD values for the CRND sets in Figures 2 and 3 are 6.4×10^{-6} and 7.0×10^{-6}, respectively. For comparison purposes, the DDA version of Figure 3 in Tanyimboh and Sheahan (2002) yields a GD value of 9.2×10^{-4} for the CRND set. Finally, for the sample network, the typical CPU times for a Pentium 4 personal computer (256 MB RAM, 1400 MHz processor) were as follows. Calculation of the ME flows, i.e. Problem 1: 0.2 seconds; solving the MCME design, i.e. Problem 2: 0.5 seconds; calculating the hydraulic reliability and redundancy including the HDA simulations of the full and degraded configurations, i.e. Eqs. (5) and (6): 0.32 seconds.

DISCUSSION

The method proposed is very quick and effective and Figure 2 demonstrates a key feature of the technique. The bulk of the designs do not belong to the CEND set, and the performance (hydraulic or otherwise) of designs that are not members of the CEND set would not be evaluated when using the proposed approach. As observed by Tanyimboh and Sheahan (2002), this filtering contributes to the efficiency of the method. The fraction of designs that are CEND is expected to decrease as the total number of layouts increases, which assures the practicability of the technique. Thus, the proportion of CEND designs in Figure 3 with 65 designs is 26.2 % (i.e. 17/65) while, in Figure 2 with 137 designs, only 16.8 % (i.e. 23/137) of the designs are CEND. This is an *extremely important* property: it shows that the efficiency of the method is not affected in an adverse way by the size of the WDS in terms of the number of pipes and thus the number of alternative layouts.

Furthermore, as there are only 23 CEND designs in Figure 2, at most 23 reliability evaluations (Eq. 5) would be required when using the proposed approach. This figure is really very small for a joint layout, pipe size and reliability optimization model. For example, Tanyimboh and Sheahan (2002) observed that the reliability is often evaluated in each iteration of components sizing optimization routines (Cullinane et al., 1992; Fujiwara and Tung, 1991; Su et al., 1987). To illustrate this point, the numbers of iterations for the MCME designs (Problem 2) herein were of the order of about 50 to 100. It is worth repeating that most models for optimizing the design of WDSs do not consider both layout selection and a quantified measure of hydraulic reliability.

Figure 2 shows that the CRND set locates the true CRPO front excellently, with GD = 6.4×10^{-6}. To put this in perspective, for the solutions in Figure 2, the maximum GD value would be about 0.55. Also, the CRND designs are fairly uniformly spread out along the PO front. This is a desirable feature as it provides a range of alternative solutions and highlights the trade-off between cost and reliability. Additional performance indicators could be used to help determine the 'best' designs. Herein, the redundancy (Eq. 6) was used as an example (Figure 4). By contrast, other layout and pipe size optimization techniques generally yield a single solution.

It is worth observing that the CRND set in Figure 2 is merely a subset of the CRPO front; the proposed approach would not necessarily identify all the CRPO designs. However this, in reality, is another extremely important property that contributes to the computational efficiency of the present approach, so long as the CRPO front is located accurately. In practice, this merely requires a more or less uniformly distributed set of CRND designs, as in Figure 2 for example. Indeed, as mentioned in Step 2b of the proposed procedure, for networks with large numbers of candidate layouts, a fraction of the layouts could be discarded purely on the basis of their entropy values. Compared to the effort involved in generating a complete design and evaluating its hydraulic reliability, the calculation of the ME value is a relatively simple exercise.

Finally, Figure 3 has fewer data points than Figure 2 as Figure 2 also includes designs based on other (often, but not always, less appropriate) sets of flow directions in addition to those of Figure 3. Figure 2 suggests the proposed approach and the relationship between entropy and hydraulic reliability are robust enough as virtually all of the CEND designs are close to the CRPO front. Thus, even though different layouts and flow directions yield different local optimum designs, there seems to be a strong correlation between entropy and hydraulic reliability. Indeed, Tanyimboh and Sheahan (2002) have shown that, in cases where different layouts or flow directions have identical ME values (due to the invariance property of the entropy function) the corresponding MCME designs (Problem 2) usually possess essentially identical hydraulic reliability properties.

SUMMARY AND CONCLUSIONS

The maximum entropy approach to the joint layout, pipe size and hydraulic reliability optimization of WDSs has been demonstrated. The approach has been shown to be dominant in several aspects in comparison to its predecessors (e.g. Afshar et al., 2005; Cullinane et al., 1992; Fujiwara and Tung, 1991; Kessler et al., 1990; Su et al., 1987). Firstly, the proposed method proves more effective since it concentrates on a small proportion of the available layouts and designs and hence the size of the WDS does not affect the method in an adverse way. Secondly, the method quantifies the hydraulic reliability explicitly in the search for the optimum layout configuration. Thirdly, unlike other methods which generally produce a single solution, the outcome of the procedure in this paper is a range of non-dominated solutions from which a design can be selected based on the trade-offs and other relevant considerations. Finally, the use of HDA in the reliability calculations adds to the accuracy of the ME method in identifying the true cost-reliability PO front.

More research is necessary to improve the applicability of the proposed method to real-world WDSs. For example, the entropy approach has yet to be applied in the optimization of WDSs involving components other than pipes, i.e. pumps, valves and service reservoirs. Also, the uncertainties and fluctuations in demands, which will affect the values of the network entropy and reliability and thus their correlation, were not considered herein. The results presented would appear to suggest that the time is ripe to use actual, discrete pipe sizes for MCME designs. In conclusion, the ME approach presented in this paper provides an excellent foundation for the search towards the ultimate design optimization tool for WDSs. The method is quick, robust and accurate, as demonstrated by its capacity to locate the true Pareto-optimal front.

Acknowledgement

The authors are grateful for the financial support of the UK EPSRC for the PRAAWDS research (GR/NO2672) and ORS scheme and the University of Liverpool for the second author.

APPENDIX. REFERENCES

[1] Afshar, M. H., Akbari, M., and Marino, M. A. (2005). "Simultaneous layout and size optimization of water distribution networks: engineering approach." *J. Infrastructure Systems,* 11(4), 221-230.

[2] Alperovits, E., and Shamir, U. (1977). "Design of optimal water distribution systems." *Water Resources Research,* 13(6), 885-900.

[3] Awumah, K., Bhatt, S. K., and Goulter, I. C. (1989). "An integer programming model for layout design of water distribution networks." *Engineering Optimization,* 15(1), 57-70.

[4] Awumah, K., Goulter, I., and Bhatt, S. K. (1991). "Entropy-based redundancy measures in water distribution network design." *J. Hydraul. Eng.,* 117(5), 595-614.

[5] Awumah, K., and Goulter, I. (1992). "Maximizing entropy-defined reliability of water distribution networks." *Eng. Opt.,* 20(1), 57-80.

[6] Cullinane, M. J., Lansey, K. E., and Mays, L. W. (1992). "Optimization-availability-based design of water distribution networks." *J. Hydraul. Eng.,* 118(3), 420-441.

[7] Deb, K. (2001). Multi-Objective Optimization using Evolutionary Algorithms, Wiley.

[8] Fujiwara, O. and de Silva, A. U. (1987). "Algorithm for reliability-based optimal design of water distribution networks." *J. Env. Eng.,* 116(3), 575-587.

[9] Fujiwara, O., and Khang, D. (1990). "A two-phase decomposition method for optimal design of looped water distribution networks." *Water Resour. Res.,* 26(4), 539-549.

[10] Fujiwara, O., and Tung, H. D. (1991). "Reliability improvement for water distribution network through improving pipe size" *Water Resources Research,* 27(7), 1395-1402.

[11] Goulter, I. C., and Morgan, D. R. (1984). Discussion of "Obtaining the layout of water distribution systems." *J. Hydraulics Div.* 109(1), 67-68.

[12] Jacobs, P., and Goulter, I. C. (1989). "Optimization of redundancy in water distribution networks using graph theoretic principles." *Eng. Opt.* 15(1), 71-82.

[13] Kalungi, P., and Tanyimboh, T. T. (2003). "Redundancy model for water distribution systems." *Reliability Engineering and System Safety,* 18(3), 275-286.

[14] Kessler, A., Ormsby, L., and Shamir, U. (1990). "A methodology for least-cost design of water distribution networks." *Civ. Eng. Systems,* 7(1), 20-28.

[15] Loganathan, G., Sherali, H., and Shah, M. (1990). "A two-phase network design heuristic for minimum cost water distribution system under a reliability constraint." *Eng. Opt.,* 5(4), 311-336.

[16] Morgan, D. R., and Goulter, I. C. (1985). "Optimal urban water distribution design." *Water Resour. Research,* 21(5), 642-652.

[17] Rowell, W. F., and Barnes, J. (1982). "Obtaining layout of water distribution systems." *J. Hydraulics Div.,* 108(1), 137-148.

[18] Setiadi, Y. (2006). Entropy-based Design Optimization of Water Distribution Networks. PhD thesis, University of Liverpool, UK, submitted.

[19] Setiadi, Y., Tanyimboh, T. T., and Templeman, A. B. (2005) "Modeling errors, entropy and the reliability of water distribution systems." *Advances in Engineering Software*, 36(11-12), 780-788.

[20] Shannon, C. (1948). "A math. theory of communication." *Bell Syst. Tech. J.,* 27(3), 379-428.

[21] Su, Y., Mays, L. W., Duan, N., and Lansey, K. E. (1987). "Reliability-based optimization model for water distribution systems." *J. Hydraulic Eng.,* 114(12), 1539-1556.

[22] Surendran, S., Tanyimboh, T. T., and Tabesh, M. (2005). "Peaking factor based reliability analysis of water distribution systems." *Advances in Engineering Software,* 36(11-12), 789-796.

[23] Tanyimboh, T. T. (1993). An Entropy-based Approach to the Optimum Design of Reliable Water Distribution Networks, PhD thesis, University of Liverpool, UK.

[24] Tanyimboh, T. T., Tabesh, M., and Burrows, R. (2001). "An appraisal of source head methods for calculating the reliability of water distribution networks." *J. Water Res. Plng and Mngt,* 127(4), 206-213.

[25] 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.

[26] Tanyimboh, T. T., and Tabesh, M. (1997). "The basis of the source head method of calculating distribution network reliability." *3rd Internat. Conf. on Water Pipeline Systems,* The Hague, 211-220.

[27] Tanyimboh, T. T., Tahar, B., and Templeman, A. B. (2003). "Pressure-driven modelling of water distribution systems." *Water Sc. and Tech. - Water Supply,* 3(1-2), 255-262.

[28] Tanyimboh, T. T., and Templeman, A. B. (1993a). "Optimum design of flexible water distribution networks." *Civil Engineering Systems,* 10(3), 243-258.

[29] Tanyimboh, T. T., and Templeman, A. B. (1993b). "Calculating maximum entropy flows in networks." *J. Operational Research Society,* 44(4), 383-396.

[30] Tanyimboh, T. T., and Templeman, A. B. (1993c). "Maximum entropy flows for single-source networks." *Engineering Optimization,* 22(1), 49-63.

[31] Tanyimboh, T. T., and Templeman, A. B. (1998). "Calculating the reliability of single-source networks

by the source head method." *Advances in Engineering Software*, 29(7-9), 499-505.

[32] 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.

[33] Wagner, J. M., Shamir, U., and Marks, D. H. (1988). "Water distribution reliability: simulation methods." *J. Water Resour. Plng and Mgmt*, 114(3), 276-294.

[34] Yassin-Kassab, A., Templeman, A., and Tanyimboh, T. (1999). "Calculating maximum entropy flows in multi-source, multi-demand networks." *Eng. Opt.*, 31(6), 695-729.

APPENDIX. NOTATION

The following symbols are used in this paper:

a_m	= probability that link *m* is in service at any given moment;
C_{ij}, D_{ij}	= Hazen-Williams roughness coefficient and diameter, respectively, for pipe *ij*;
D_{min}, D_{max}	= minimum allowable and maximum available pipe diameters, respectively;
d_i	= Euclidean distance, in the space of the objective functions, between the *i*th non-dominated solution and the nearest solution in the Pareto-optimal front;
e	= exponent in cost function;
FT	= hydraulic redundancy;
$f_m^{(i)}$	= value of the *m*th objective function for the *i*th non-dominated solution;
$f_m^{*(k)}$	= value of the *m*th objective function for the *k*th Pareto-optimal solution;
GD	= generational distance;
H_i, H_s	= piezometric head at node *i* and head at source node *s*, respectively;
H_i^{des}	= piezometric head at node *i* above which outflow equals demand;
H_i^{min}	= piezometric head at node *i* below which outflow equals zero;
h_{ij}, h_p	= headloss in pipe *ij* and path *p*, respectively;
I	= the set of all source nodes;
L_{ij}	= length of pipe *ij;*
M, N	= number of links (i.e. pipes, pumps and valves) and nodes, respectively;
N_i	= the set of all the nodes connected to node *i*;
ND_i	= the set of outflows, including any demand, at node *i;*
NP, NS	= number of Pareto-optimal and non-dominated solutions, respectively;
P_i	= fraction of total flow through network that reaches node *i;*

p_i = probability of ith event;

$p(m), p(m, n)$ = respective probabilities that only link m and only links m and n are not in service;

$p(0)$ = probability that all links are in service;

Q_i = inflow (for sources) or demand (for demand nodes) at node i;

Q_i^{req} = demand at node i;

Q_{0i} = inflow at source node i;

$Q_i(H_i)$ = head-dependent outflow at node i;

Q_{ij}, Q^*_{ij} = flow and maximum entropy flow in link ij, respectively;

R = hydraulic reliability;

S = informational entropy;

S_i, S_0 = entropy of node i and entropy of source supplies, respectively;

T = sum of nodal demands;

T_i = total flow reaching node i;

$T(m), T(m, n)$ = respective total flows supplied with only link m and only links m and n out of service;

$T(0)$ = total flow supplied with all links in service;

u_m = probability that link m is not in service;

α = dimensionless conversion factor (10.67 in S.I. units); and

γ = coefficient in cost function.

FIGURE CAPTIONS

Fig. 1 Alternative layouts for the sample network (with demands in liters per second)

Fig. 2 Cost versus hydraulic reliability for all designs

Fig. 3 Cost versus hydraulic reliability for the Tanyimboh and Sheahan flow directions

Fig. 4 Cost versus hydraulic redundancy for all designs

Fig. A1 Optimal layouts and flow directions

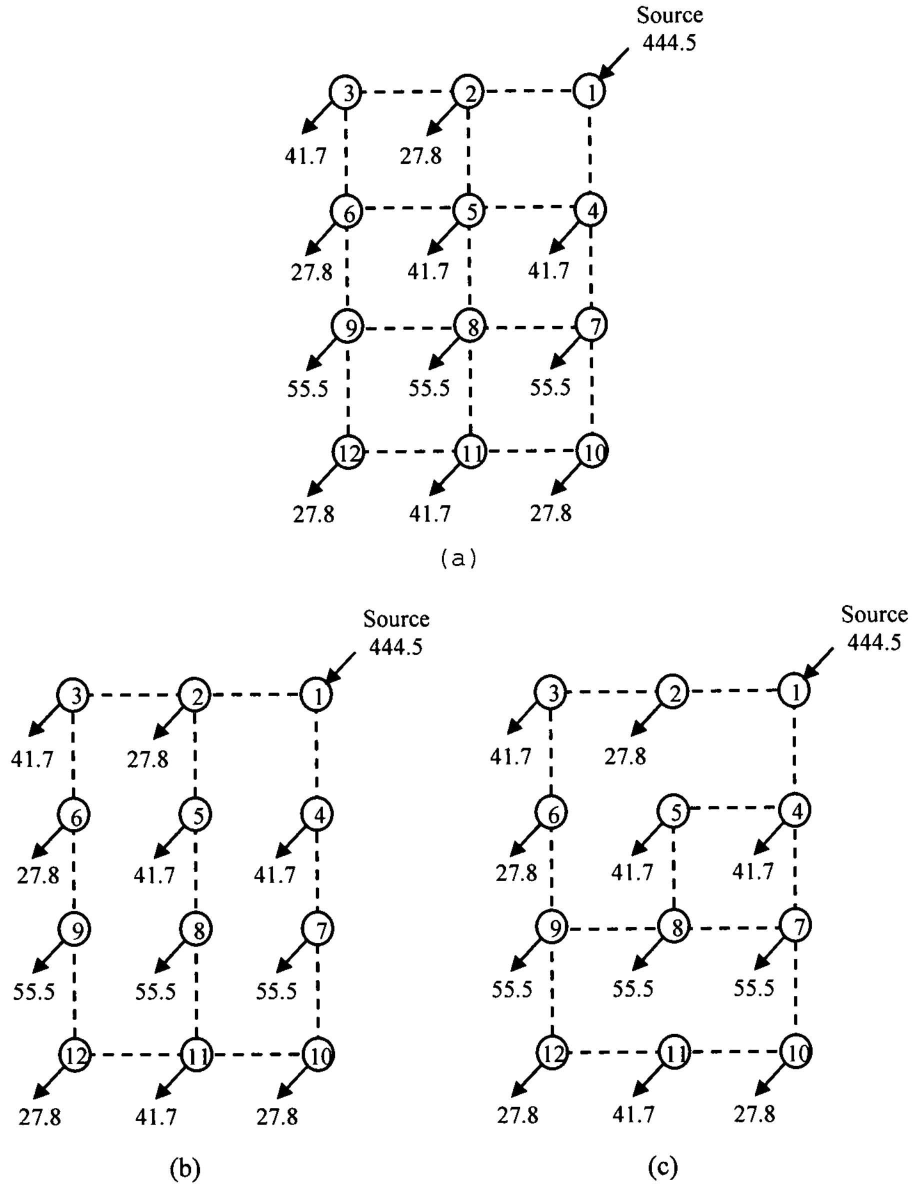


Figure 1 Alternative layouts for the sample network (with demands in liters per second)

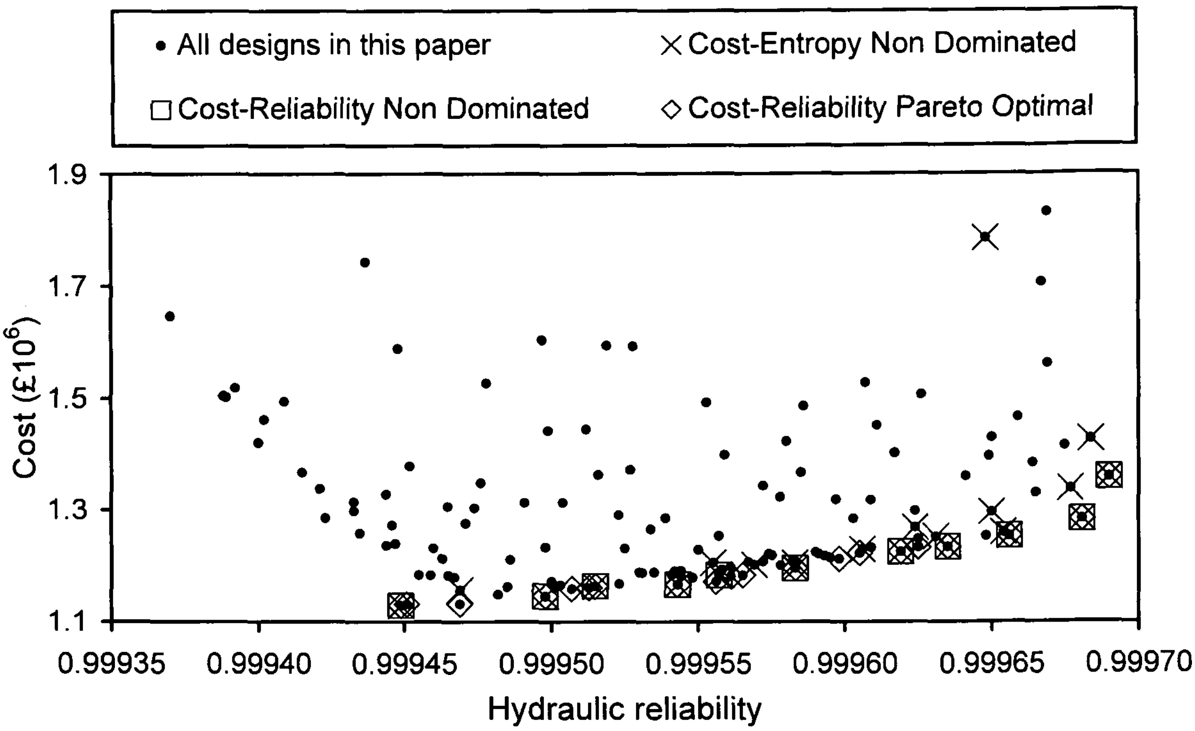


Figure 2 Cost versus hydraulic reliability for all designs

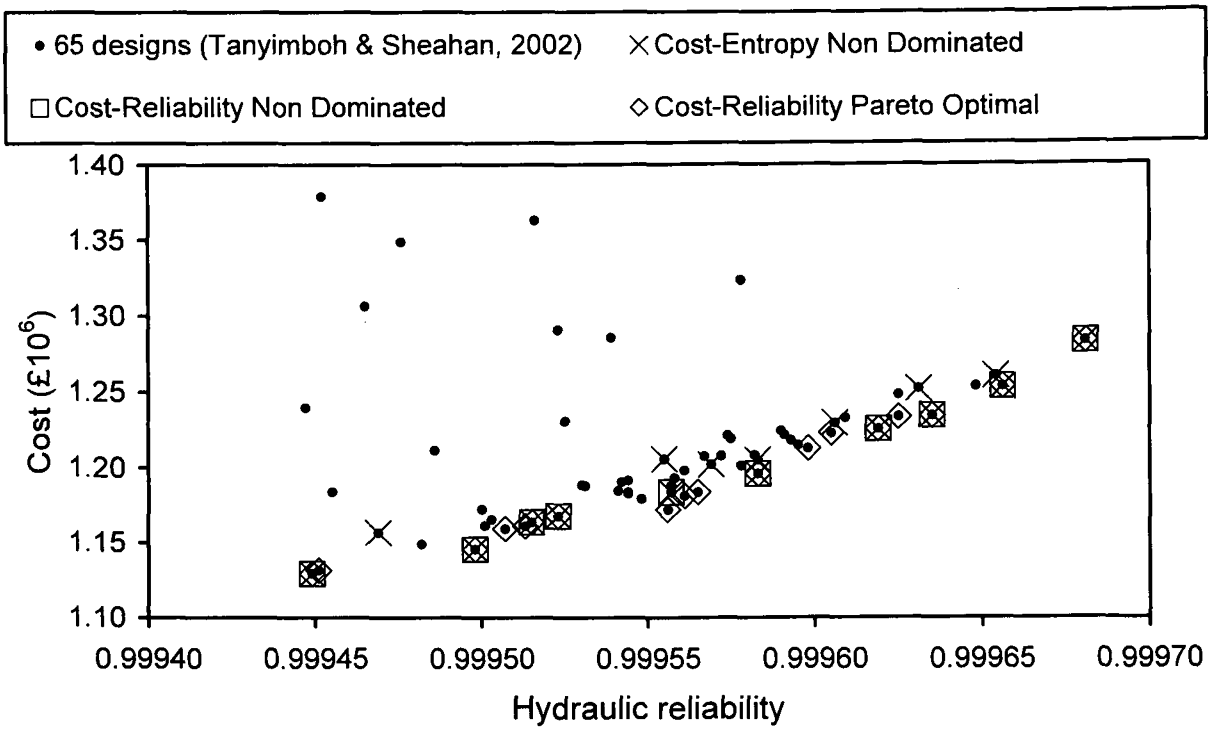


Figure 3 Cost versus hydraulic reliability for the Tanyimboh and Sheahan flow directions

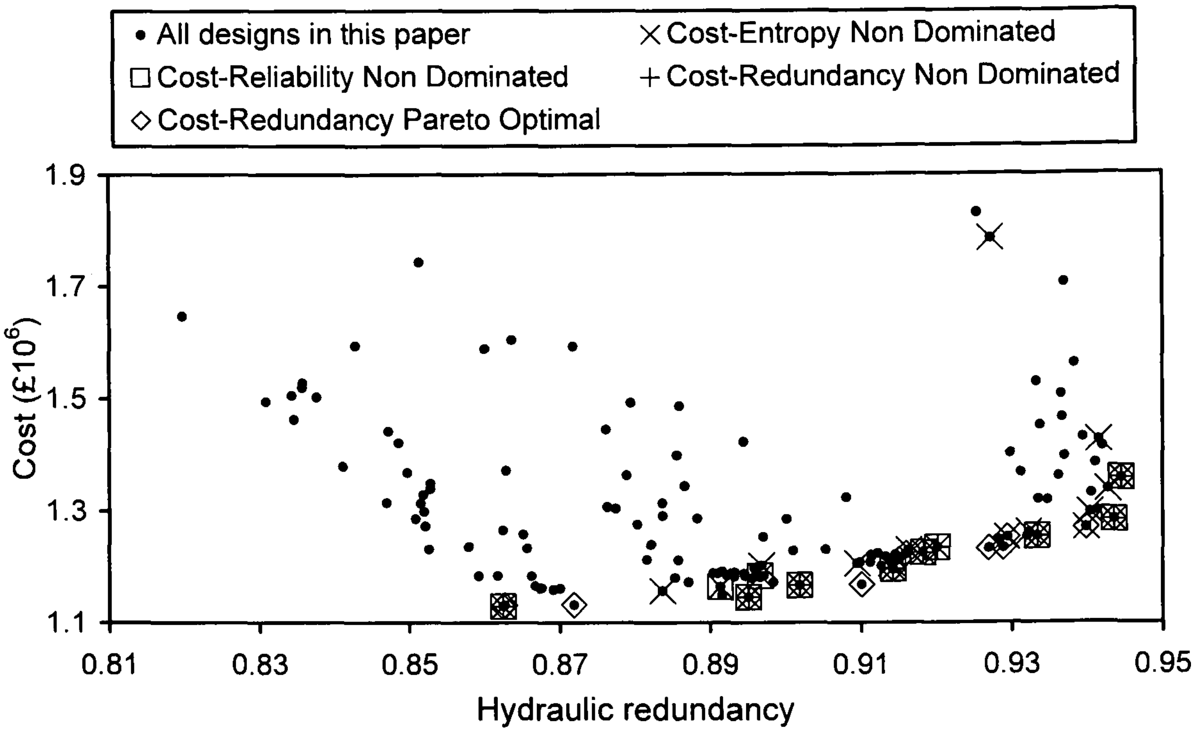


Figure 4 Cost versus hydraulic redundancy for all designs

APPENDIX. OPTIMAL LAYOUTS AND FLOW DIRECTIONS

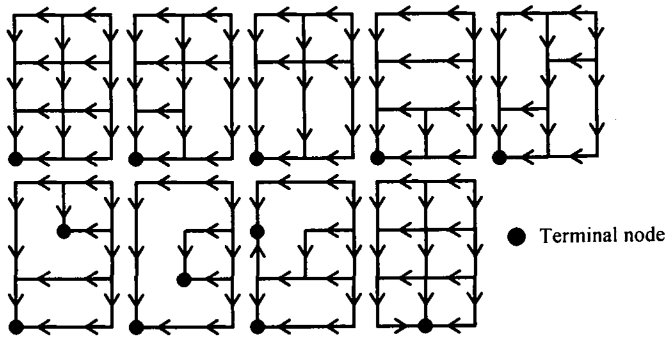


Figure A1 Optimal layouts and flow directions