CITY UNIVERSITY

LONDON

Department of Mechanical Engineering & Aeronautics

OPTIMAL DESIGN OF WATER DISTRIBUTION NETWORKS WITH RELIABILITY CONSIDERATIONS

by

Driss KHOMSI

A Thesis submitted to City University

for the Degree of Doctor of Philosophy

in Mechanical Engineering

JUNE
1994

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter VIII

Chapter IX

ACKNOWLEDGEMENT

I wish to thank all those who have helped and encouraged me during all the stages of this project. In particular I would like to express my sincerest gratitude to my thesis supervisors:

- Prof. A. R. D. Thorley, Head of the Department of Mechanical Engineering and Aeronautics, for his skilful, valuable and sympathetic guidance during the course of this research. His encouragement and constant support will always be remembered.

- Dr. G. A. Walters, Senior Lecturer in the School of Engineering at the University of Exeter, for his original and expertise ideas, tactful guidance and for his frequent valuable advice.

- Prof. D. Ouazar, Prof. in Civil Engineering Department at the Ecole Mohammadia d'Ingénieurs, Rabat, Morocco, for his overall help, fruitful discussions and useful suggestions.

I wish also to acknowledge the financial support and the efficient help provided by the British Council under a Link agreement between Ecole Mohammadia d'Ingénieurs, Rabat, Morocco, the Department of Mechanical Engineering and Aeronautics at City University, London, and the School of Engineering at the University of Exeter, Exeter.

The Ecole Mohammadia d'Ingénieurs and particularly the Civil Engineering Department for their support.

I also remember my daughter, Asmae, who has brought me immeasurable happiness and a smile at each day, and also my sister Saloua.

Finally I would like to express my deepest respect to my friend Prof. Abdelmajid Semlali Hassani, my thanks to my mother, Aicha, for her encouragement and my sincerest love to my wife, Btissam Mamou, to them all this work is dedicated. Their love, understanding and reassurance have made possible the completion of this work.

SYNOPSIS

The overall aim of this research has been to develop new algorithms and computer software that may be used to assess the reliability of water distribution systems. Such a tool can be used by design engineers to create systems which are both economical in total cost commensurate with meeting targets for a specified level of reliability.

The introduction describes how water supply and distribution systems are normally designed, what they comprise and problems associated with failure or lack of availability of an adequate supply to the end user. This is followed by a résumé of current methods and algorithms for the analysis of networks and a detailed examination of the previous work on network optimisation and reliability.

Three main algorithms exist for the analysis of water networks. These are the Hardy-Cross methods, the Newton-Raphson methods and the Linear method. A computer program based on the Linear method, which is known to be the most reliable, is proposed for the hydraulic analysis part of the present work.

With respect to reliability, a full discussion of the topic, including all the various factors which influence it such as the stochastic nature of customer demands, the apparently random occurrence of pipe breakages and the concept of repair time, is presented. A reliability analysis model, that incorporates simultaneously the three reliability factors mentioned, for the assessment of nodal and system availabilities, is proposed, from which an efficient computer program has been developed and tested.

Two models for the design of optimal water distribution systems, based on reliability criteria, have been developed, programmed and tested. The first model makes use of the entropy principle for producing 'reliable' distributions of flow and the Linear Programming technique is used for computation of the least cost design. In the second model, however, a Genetic Algorithm procedure, that incorporates the new reliability analysis model and which is superior to other models has been formulated.

The thesis concludes with a comparison between the two methods formulated as a result of this research and applied to realistic practical systems, plus suggestions for further work to improve the optimisation of water distribution networks.

NOTATION

A_d	= pipe section, m^2
A_{Fit}	= average fitness of a population of strings
A_{Net}	= network availability
$Anode_i$	= availability of node i
a_p	= coefficient of pump curve
As_T	= target system availability
α, β	= pre-specified values of probability
b_k	= net head loss in path k, $k = 1,...,(N_F - 1)$, m
b_p	= coefficient of pump curve
C, C_i	= Hazen-Williams coefficient (for link i)
C_{ij}	= Hazen-Williams Coefficient of pipe of diameter j in link i
$Check_i$	= 0/1 variable used in the Path_Q computer program
C_f	= large constant
Cl_i	= cost of link i
$Cost_{NET}$	= total network cost (£)
$Cost_{St}$	= total network cost for string St including penalty cost (£)
$Cost_{NET}(St)$	= network cost for string St without penalty cost (£)
c_p	= coefficient of pump curve
D	= pipe diameter, m
D_{eq}	= equivalent diameter, m

D_{ij}	=	diameter of pipe j in link i, m
D_{min}, D_{max}	=	minimum and maximum diameter, m
D_t	=	theoretical diameter, m
Dt_i	=	theoretical diameter of link i, m
d_p	=	coefficient of pump curve
ΔE	=	difference in total head between two fixed head nodes, m
$\delta(h)$	=	defining length of schema h
ΔH_i	=	head loss in link i, m
$\Delta H_i(j)$	=	head loss in link i corresponding to loading condition j, m
$E[f_j]$	=	known value of expectation function f_j
e^p	=	coefficient of pump curve
E_p	=	pump energy
ϵ	=	precision of computation for determining L_s
F	=	algebraic function
F^*	=	optimal value of F
FailTime	=	time that supply is in failure state
F_i	=	number of failures per unit time of pipe i
F_{ij}	=	probability of insufficient supply at node i of the jth violation of minimum pressure
Fit	=	raw fitness
Fit(h)	=	schema average fitness
Fit_{St}	=	raw fitness of string St
Fit'_{St}	=	scaled fitness of string St
$f_{ji}(y)$	=	expectation function f_j of the ith probability of the discrete random variable y
$Fnode_i$	=	probability of hydraulic failure of node i

f_Q	=	function of flow Q for a given pipe
f_P	=	function of flow Q for a pump
f'_Q	=	derivative of f_Q
f'_P	=	derivative of f_P
g	=	acceleration due to gravity (9.81 m/s^2)
G_i	=	gradient of function f_Q evaluated at $Q = Q_i$
γ	=	constant depending on the units used in the Hazen-Williams head loss equation
h	=	schema
H_0	=	original head at the source, m
Hc	=	cutoff head of the pump, m
H_k^{min}	=	minimum required head at node k, m
H_k^{min} (j)	=	minimum required head at node k for loading condition j, m
H_i	=	pressure at node i, m
h_L	=	total head loss in a pipe including minor loss, m
h_{LM}	=	minor loss, m
h_{LP}	=	head loss, m
H_{pump}	=	pump head, m
J_d	=	hydraulic gradient at peak flow
J_{ij}	=	hydraulic gradient of pipe of diameter j in link i
J_{min}, J_{max}	=	lowest and highest hydraulic gradients
k	=	index for counting violations of minimum pressure
K_1, K_2	=	constants equal to $4/\pi$ and 10.70 respectively
K_B, K_S, K	=	arbitrary positive constants
KLoad	=	ratio of the actual demand to the time averaged demand
$KLoad_i$	=	demand factor for zone i

K_m = minor loss constant

K_{max} = number of times the pressure test was violated

K_p = pipe constant which is function of length, diameter and roughness

K_{Scal} = constant used in the Linear Normalization

L = pipe length, m

L_b = length of a string St

L_i = length of link i, m

L_s = length of a sub-string corresponding to a decision variable

m = minor loss coefficient

M = number of expectation functions

m(h,t) = number of copies of schema h at time t

n = coefficient

N = number of subsystems

$Ncopies_{St}$ = number of copies of string St

N_D = number of decision variables

N_d = number of candidate decision variables

Nd(i) = set of outflows, including any demand, from node i

N_F = number of fixed head nodes

n(i) = number of permissible diameters in link i

N_J = number of nodes

N_L = number of loops

NLoad = total number of demand conditions

N_p = number of pipes

N_{pop} = population size

N_x = number of segments connected in series in a link

$o(h)$	=	order of schema h
p_0	=	probability of generating initial GA population
p_{0i}	=	proportion of the total supply to the network provided by source i
$p_{crossover}$	=	probability of crossover
pD_i	=	probability of failure per day of pipe i of diameter D_i
pD_{eq}	=	probability of failure of a link of the equivalent diameter D_{eq}
$Pf(k)$	=	set of links in path k, $k = 1,\ldots,(N_F - 1)$
p_i	=	probability of occurrence of subsystem i
P_i	=	probability of flow arriving at node i, $i = 1,\ldots,N_J$
$PKLoad_i$	=	probability of demand condition i
$Pl(i)$	=	set of links in loop i
$p_{mutation}$	=	probability of mutation
$Pn(i)$	=	set of links in the path from the source to node i
PNetwork	=	probability of no pipes being out of action
$Ppipe_i$	=	probability of failure of pipe i
q_{0j}	=	external inflow at node j, m^3/s
Q	=	flow, m^3/s
Q^{avg}	=	overall average demand, m^3/s
Q_i^{avg}	=	time averaged demand at node i, m^3/s
Q_i	=	flow at link i, m^3/s
Q_{ij}	=	flow of link i for loading condition j, m^3/s
q_{in}	=	flow into the node, m^3/s
Q_{min}, Q_{max}	=	minimum and maximum design flow, m^3/s
q_{out}	=	flow away from the node, m^3/s
Q_i^{ext}	=	external demand or supply at node i, m^3/s

Q_{ij}^{ext} = external demand for loading condition i at node j, m^3/s

Q_p = peak flow, m^3/s

$Rank_{St}$ = rank of solution St

ρ = constant cost coefficient

rD_i = breakage rate per year and per km of pipe i of diameter D

Rs = system reliability

Rs_T = target system reliability

S = system entropy

S_0 = entropy of sources;

S_B = Boltzman's entropy

Scheck = control variable used in the PATH_Q computer program

S_i = entropy of node i

S_N = nodal entropies

S_S = Shannon's entropy

St = string of binary/integer values

St_i = string i of binary/integer values

St_i' = new string i resulting from the application of genetic operators to string St_i

T_0 = total supply or demand, m^3/s

Θ = exponent greater than 1

T_i = total outflow (including any demand from node i), m^3/s

TotalTime = time interval considered, day

T_{rep} = repair time during which a pipe remains in a failed state, day

V_d = design velocity at peak flow, m/s

V_{min}, V_{max} = minimum and maximum design velocities, m/s

x_i = decision variable i, $i=1,...,N_d$

x_i^* = optimal value of x_i

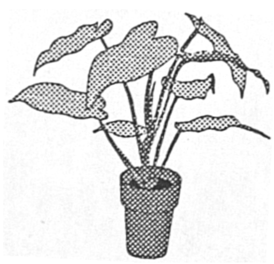
x_{min}, x_{max} = minimum and maximum values of decision variables x

X_{ij} = segment length of pipe of diameter j in link i, m

y, y_i = discrete random variables

Chapter I

WATER DISTRIBUTION SYSTEMS



Chapter I

WATER DISTRIBUTION SYSTEMS

1.1 INTRODUCTION

Water is one of the essential fluids upon which life is based. Without it death and extinction would follow. The purpose of a water supply system therefore is to provide an adequate supply of suitable water at points of need. The system provides a link between points of availability, ie the sources of water, and the users, be they agricultural, domestic or industrial.

If the sources of supply are some distance from the users, and also if the users are large in number and spread over a wide geographical area, relatively long distance trunk supply mains will be required. The more general distribution to the various consumers will be through distribution networks. The trunk mains will often be of large diameter, eg 1-4 metres, whilst the pipes in the distribution networks will be in general in the range of 150-500 mm or less.

The design of a water supply system is influenced by many factors. However, it may be postulated that there are, say, three key stages through which the engineering design process evolves. These are - the initial feasibility studies, the functional design and the risk assessment or fault analysis.

During the feasibility studies the overall economic and engineering issues are reviewed to assess the viability and worthwhileness of the scheme, leading to a specification for a functional design if it is agreed that the scheme should go ahead.

The functional design is when details of the system are worked out, pipeline routes, sizes and materials chosen, the location of pumping stations, selection of pumps, valves and control gear made, operating strategies evolved, and so on. This is also where feedback loops start to appear in the design process - not only in relation to the risk assessments and the impact that these may have on the design, but also in relation to optimisation, scope for developments in future demand or, perhaps, even potential changes in available sources.

The major concern in this project is with "optimisation". Historically, this process has been interpreted in many different ways by engineers and researchers. Early interpretations tended to focus on the optimisation of capital cost, eg the cheapest layout of the pipe network and, perhaps, pumping stations. Other interpretations look, for example, at optimising operational costs, saving energy, and so on. The full literature will be discussed presently, However, it is only recently that optimisation from the perspective of the consumer has begun to be addressed - and particularly so in this study.

To lay the foundation for the development of these ideas various features of all water distribution systems are now discussed in more detail, following which various techniques for the functional analysis are reviewed. A detailed review of the literature on optimisation will then be given in Chapter 3.

1.2 SYSTEM DESIGN

The principal driving force in pipe network design is the demand. The demand is created by domestic consumers and by industrial, commercial and agricultural needs. Some allowance must also be made for losses from the system and for "free" usage, ie public parks and leisure use, fountains, firefighting, etc.

Globally, agriculture accounts for about 55 per cent of water demand, industry 35 per cent and domestic usage about 10 per cent (Park, 1986), but without doubt the level and character of demand will both vary from country to country, and through time within each country. In the United Kingdom during the present century, agriculture demand for irrigation water has been limited and industrial and domestic demand have been dominant (Park, 1982).

Over 99 per cent of England and Wales is now served by the privatised water supply. Most domestic users have unmetered supply, and are charged in proportion to property rateable values, not on the basis of water used, which would require water metering. Domestic consumption currently averages about 125 l/person/day (Park, 1986) and by the end of the century, it is estimated to become around 140-160 l/person/day (I.W.E.S, 1984).

1.2.1 Domestic usage

It is estimated that around 33 per cent of domestic demand is used in WC flushing, a further 17 per cent is used in bathing and showering, 12 per cent is used in clothes washing machines and the remaining 38 per cent is used in various ways which include hand washing, drinking, cooking, cleaning, outside use such as garden watering and car washing and luxury appliances such as dishwashers etc.

Domestic demands vary within the same country and from country to country. Among the causes of different domestic consumption there are:

- variation in habit, as regards frequency of use of appliances and fittings;
- volume per use;
- levels of appliance ownership;
- climatic conditions.

Compared to the UK domestic consumption, it was found in the US that there is a very high level of use of WC flushing because the average volume/flush is 5 gallons, twice the UK level. Moreover flow rate for showers is much higher in the US than in the UK (5 US gallons/minute in US, 1 gallon/minute in UK) and also the frequency of showering in US is greater than that in UK (about 1 per person per day against an average 1 1/2 baths or showers per person per week in the UK) (I.W.E.S, 1984).

Consumption of water due to garden watering is weather dependent and occurs over a short period. In the UK it constitutes a small part of the average annual household, consumption accounting for less than 5 per cent (I.W.E.S, 1984) though in other countries, due to climate conditions garden watering can exceed the total of the remaining part of domestic demand (Power et al., 1981).

Tourist demand can be added to consumer requirements. This appears within the spring/summer months in general. The amount of water needed is variable, depending upon the area served and the type of accommodation used (Hooper, 1981). Tourist demand may in some circumstances create serious problems of supply. For example, from a survey by South West Water Authority (1978), it was found that at the height of the tourist season the global demand could increase by 20 per cent over the annual average.

1.2.2 Industrial, Commercial and Agricultural usage

The second category of water consumption is represented by industrial, commercial and agricultural usage. Water consumption in small engineering workshops and high industry may range up to 3 m^3 per 1000 m^2 of factory floor per day (I.W.E.S, 1983).

Unlike the domestic demand, industrial consumption can be influenced by controlling consumption, ie by means of reduction of wastage of water, recycling water, or even by changing the production process to another one which uses less water.

Consumption in small factories where water is not used in the manufacturing process, commercial consumption in restaurants, shops and offices, and institutional consumption in schools, hospitals and government offices can be assessed from the water authorities.

The use of public water supply for agricultural purposes in the UK is relatively small in terms of annual average demand (about 500 Ml/d, less than 5 per cent of the total). About 40 per cent of all agricultural water requirements are met by direct abstraction (I.W.E.S, 1984).

1.2.3 'Free' usage

Firefighting, park irrigation, sewer jetting, flushing and the maintenance of mains are considered as 'free' usage. In general, water used for firefighting is not recorded and measurement can be difficult to make since it is not known in advance when and where water will be needed. The total quantity used is likely to be small, and is usually ignored in estimating future demands (I.W.E.S, 1983).

1.2.4 Losses

Losses represent water wastage through leakage, bursts and illegal use. Losses start at the consumer's premises and may result from a broken service pipe downstream of the meter, defective ball valves and WC cisterns, and dripping taps. In the distribution system, water leaks through openings due to main bursts and failures of the various network fittings such as joints, hydrants and meter boxes. Illegal use results from illicit connections to the distribution mains.

All water distribution systems leak. Estimated losses from water networks vary widely depending on factors such as pressure, age and maintenance. It is extremely rare to find a rate of leakage less than 15 per cent. Such a figure would be associated with a well maintained system. In the UK it is taken as 25-50 per cent. Losses of 50 per cent and more are not uncommon. For instance, in the city of Kathmandu in Nepal, the rate of losses was 75 per cent in 1973 (I.W.E.S, 1983). The knowledge of leakage rate is an important factor in assessing future demands.

1.2.5 Variation/Uncertainty of demands

Demand profiles vary from day to day, season to season and year to year. Among the features causing this variation are the size of the area, its type in terms of industrial, commercial, tourist, domestic area or a combination of these, the weather, the temperature and the time of the year.

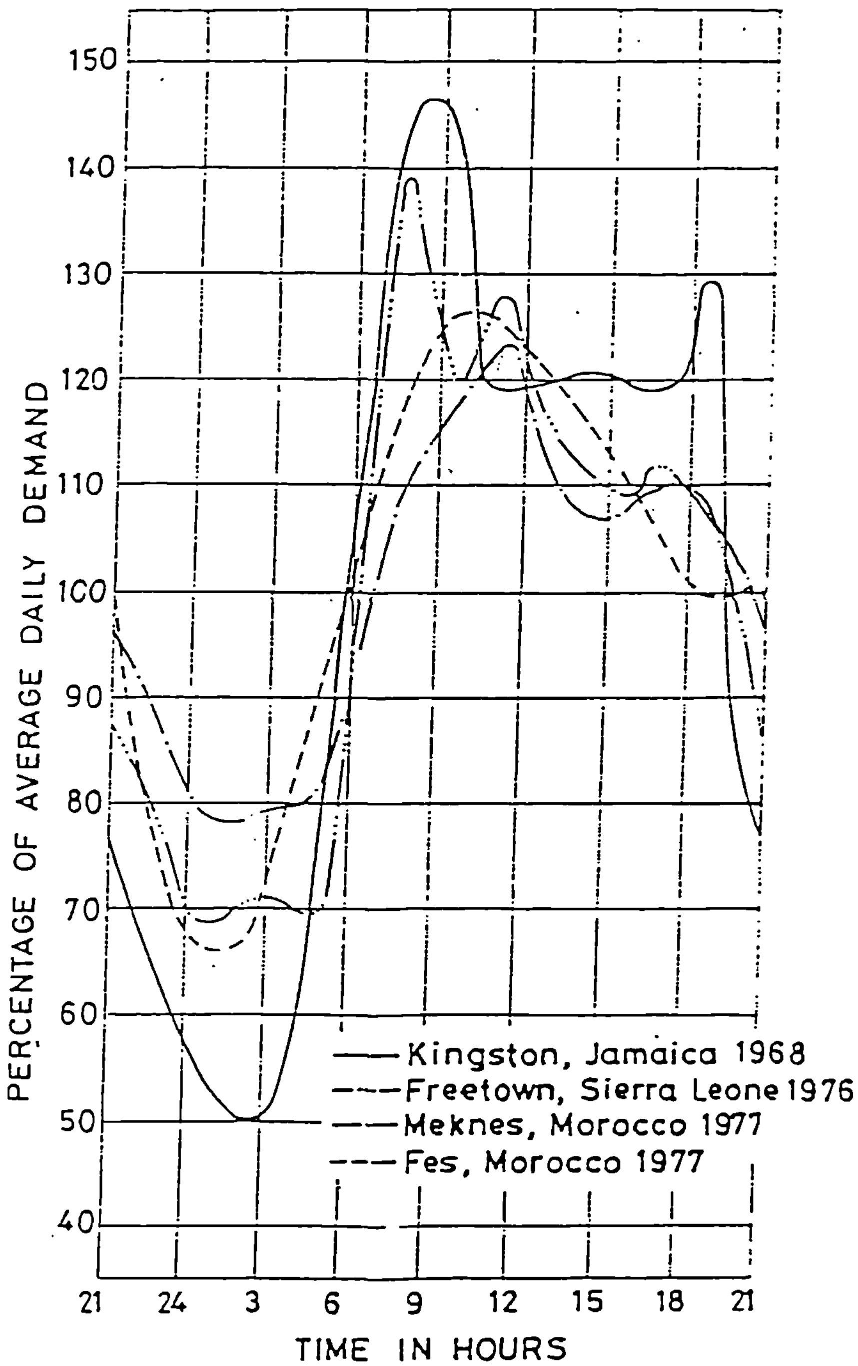


Figure 1.1 Daily Variation in the Rate of Water Consumption

(I.W.E.S, 1983)

Typical changes in the rate of usage that occurs during the day within a distribution system is given in Fig. 1.1.

Consumption peaks can be observed in both winter or summer. Winter peaks do not correspond to a real use of water but are caused by loss of water due to leaks and bursts in system links. On the other hand summer peaks may be explained by a rise in consumption resulting from two main effects, the "tourist effect" and the "garden watering effect". However the proportion of summer peak demand due to garden watering is not known (I.W.E.S., 1984).

Some peaking coefficients have been produced in the literature, ranging (for local distribution mains) from three times average demand to a higher factor of up to six times winter demand. It has also been observed that the ratios of peak week, peak day and peak hour to average (or winter) demands have not changed as per capita unmetered demand has risen (I.W.E.S., 1984).

In designing water networks, demands have been generally taken as deterministic values. However in reality they are not. In order to ensure a design is adequate all, or at least a number of the expected critical, conditions must be considered. In recent literature, the ability of a network to operate under various demand patterns has been equated with a reliable network (Templeman, 1982).

1.3 THE ELEMENTS OF PIPE NETWORKS

To meet the needs of the various consumers, as introduced in the preceding sections, the network that evolves comprises pipes for conveyance, pumps for providing the energy to drive the fluid, reservoirs and storage tanks to provide a buffer between the rate of supply and rate of demand, and valves and meters to control and monitor the demand respectively.

1.3.1 Pipes

The pipework of a distribution system comprises a number of sections with different purposes. Mains intended to convey water in bulk from one part of the network to another are called **trunk distribution mains.** These may contain some branch connections but generally no consumer connections. **Secondary mains** provide the basis of the system and their use is to link the service mains with the service reservoirs and/or with the trunk distribution mains. Some direct connections to these pipes may be allowed especially for large individual demands.

Service mains play the role of supplying water from the secondary mains to the smaller consumers. A pipe size of more than 100 mm is generally adopted by water authorities since there are numberous connections to individual households and for reasons of firefighting in residual areas. About 70 per cent of the UK distribution system is 150 mm or smaller in diameter and made from cast iron (Lackington, 1983).

Service pipes are the small diameter pipes conveying a supply of water from the service main into the customers' property. For domestic supplies a diameter less than 25 mm is generally used. The majority of service pipes are 13 mm nominal diameter, but other consumers may require larger sizes. Service pipes are frequently further subdivided in the **communication pipe** which links the service main to the boundary of the premises being supplied, and the **supply pipe** which is the portion within the boundary.

1.3.2 Pipe Materials

Cast iron, spun iron, ductile iron, spun ductile iron, asbestos cement, uPVC (unplasticized polyvinyl chloride) and MDPE (medium density polyethylene, mainly used by British Gas for many years) are the most common pipe materials used for manufacturing and fabrication of water mains. Steel pipes are frequently used for trunk distribution mains. Copper and galvanized mild steel are used for service pipes. Cast or spun iron pipes and lead pipe is for practical purposes no longer used, though both materials were widely adopted in the past (I.W.E.S, 1983). The water industry has been somewhat conservative in its selection and use of pipe materials since the provision and maintenance of mains is an expensive and long-term work. About 85 per cent of distribution pipework in the UK is cast iron; Asbestos cement and uPVC make up a large proportion of the remaining 15 per cent.

Spun ductile iron pipe which includes the advantages of ductile iron pipe over gray iron pipe (eg greater strength, 30 per cent lighter, more resistance to fracture under beam loading conditions) was first manufactured in the UK on a commercial scale in 1961 (Water Authority Association Advisory Committee, 1983).

Asbestos cement pipe which is fabricated from asbestos fibres and ordinary Portland cement have been used in the UK for more than 50 years (Water Authority Association Advisory Committee, 1982), while uPVC pressure pipe has been available to the water industry only for about 20 years (Ibid, 1982).

Recently, however significant technological advances have been seen in the world of pipe manufacturers to the extent that designers are more willing to vary their choice, with confidence in the whole-life performance of the selected material (Latham, 1990).

1.3.3 Pumps and Pumping Stations

Pumps are used to raise a fluid from one level to a higher level. Many different types of pumps can be used. However, centrifugal pumps are most frequently used within water distribution systems. Details of these pumps and others may be obtained from textbooks (eg Anon, 1968, 1982).

Basic pumping plants in use for application in the water industry are lift stations and booster stations. Lift stations are necessary where water cannot be supplied by gravity from one point to another. Such pumping may be either to service reservoirs or water towers. Booster stations, on the other hand may be applied to increase the carrying capacity of a main especially at period of peak demands, so avoiding or postponing replacement or duplication. They may also be applied to overcome local deficiencies in pressure or to maintain or increase output of a pumping station with increasing friction-head or demand. Guidelines on the selection of these pumping plants, their sites, operation, control and maintenance are fully described in I.W.E.S (1984).

1.3.4 Valves

The movement of water within the distribution network needs at all time to be controlled. Examples include maintaining system pressure within min/max operating limits, controlling leakage, maintenance and repair or adding new connections. **Valves** are an essential and integral part of the system: the more there are the more flexibility there is. The term 'valve' is widely used for both controlling devices, eg pressure reducing valve, and for throttling devices, eg partially closed valve.

Block valves are intended to fully open or fully shut and should not be used for flow regulation. The common sluice or sliding-gate valve is usually used for this purpose, as it is comparatively cheap and virtually water tight when shut.

Pressure throughout the system needs to be great enough to satisfy nodal demands. However, excessive pressures are neither desirable nor economic and also increase leakage and wear and tear on fittings including customer's own water-using apparatus. Pressure reduction in such situations may be achieved through the use of **pressure-reducing valves** or **pressure-control valves**, which are designed to maintain a preset pressure or flow in the downstream side of the valve for all flows which are lower than the upstream pressure.

Pressure-sustaining valves are similar to pressure-reducing valves but are designed to protect upstream rather than downstream pressure.

1.3.5 Reservoirs and Storage

Other major components which may appear within water distribution systems are **service reservoirs** and **water towers**. The former is a receiving tank for treated water, situated generally on high ground near a centre of population. The latter is a form of service reservoir, but elevated artificially above ground to create the necessary pressure throughout the area served which is usually of flat topography. Both tanks fulfil the same purpose of:

(1) Dampening hourly consumer demand peaks;

(2) Providing contingency storage;

(3) Compensating for variation in water quality.

Service reservoirs and water towers are many and varied ranging from small individual steel tanks through massive masonry or brick structures to those of large reinforced concrete construction built recently.

1.4 DESIGN PRACTICE

Design methods for water distribution networks vary considerably between the different industries which use pipe networks. To date the methods used can be included in four approaches:

(1) Feasible solution;

(2) Trial-and-error "optimisation";

(3) Rules-of-thumb;

(4) Mathematical optimisation.

Approaches (1), (2) and (3) can be performed manually. However, the mathematical optimisation approach requires a computer facility, especially for medium and large networks.

The first approach refers to any feasible set of diameters found by the designer that will function hydraulically for a given network without concern for the least cost solution.

The second approach, trial-and-error "optimisation", does not use methods of theoretical optimisation such as Linear, Non-Linear Programming, Geometric Programming or Genetic Algorithms or others, for the choice of pipe sizes. Rather, it only compares alternative feasible link sizes to find a least-cost by trial and error.

In the third approach some rules-of-thumb are combined by the designer to reach a practical solution. The selection of pipe diameters is performed on a criterion such as a peak flow velocity of 1.8 m/s, and a check to ensure that the solution obtained is hydraulically feasible.

For the last approach, mathematical optimisation techniques are used to find the optimal pipe sizes. A résumé of the techniques used is presented in Chapter 3.

1.4.1 Design parameters

Usually the parameters taken into account in the design of water networks are some criteria on Flow, Velocity, Pressure, Hydraulic Gradient and Diameters. The number and the type of criteria change between design teams.

1.4.1.1 Flow

This parameter is the most important variable in pipe size selection. Though most of the time water networks are operating under normal conditions, ie average flow pattern, peak flows are used for the design. Usually, peak flows are linked to the average flows via peaking factors, eg peak hourly flow is obtained by average flow times peak hourly factor.

1.4.1.2 Velocity

Limits on velocity are different from country to country. In France pipe velocities have to be between 0.5 and 1.0 m/s (Didier, 1980). According to Didier, the minimum limit has to be respected in order to avoid a deposit settling out, while the maximum limit increases energy consumption and the risk of deterioration of hydraulic components (eg valves and joints).

In the US bounds on velocities are considered in conjunction with the importance of the flow. Walski (1985) pointed out that velocities of the order of 0.6 m/s (2ft /s) at average flow and less than 2.4 m/s (8 ft/s) at peak flow are good practice.

1.4.1.3 Pressure

Constraints on pressure throughout the system are essential in the design of water mains though the UK Water Acts (Section 39, 1945) did not prescribe a fixed pressure for supplies to users. Instead, the Water Acts refer to constancy of supply. Weak pressures are indicative of problems of supply while excessive pressures are not desirable and not economic. In effect, excessive pressures cause deterioration of hydraulic fittings, moreover water use increases and leakage shoots up. In a recent work it was shown that leakage varies exponentially with pressure (Patton and Horsley, 1980).

Once again bounds on pressure are not fixed for most countries. In the UK, most main designers use a minimum distribution pressure of between 15m to 20m head for most consumers, whereas in France this limit is dropped to 10m (Didier, 1980) while it is around 14m (20 psi) during fire conditions in the US (Walski, 1985).

Maximum distribution pressures on the other hand vary also between countries. Generally they range from 40m to 60m or even more (Didier, 1980, I.W.E.S., 1984 and Walski, 1985).

Desirable pressures are about 30m to 50m which are frequently achieved through dividing a distribution system into zones, each valved such that it is separated from the others on the system or linked by pressure-control valves.

1.4.1.4 Diameters

Bounds on pipe sizes are, in general, only put on the minimum size since the maximum size depends on the importance of the area served. In the UK, mains in the network range from as small as 50 mm (2 in) to 450 mm (18 in) and above. The vast majority of small size are 75 mm (3 in), 100 mm (4 in) and 150 mm (6 in). Whilst current design practice has standardized on 100-150 mm for most new mains (Latham, 1990), there is no statutory duty to provide any particular capacity for firefighting (I.W.E.S., 1984). However, in the US some rules specify a minimum diameter of 150 mm for systems providing fire protection and at least 50 mm for systems where fire protection is absent (Walski, 1985). In France on the other hand, the minimum diameter is in the interval 40-60 mm in the situation of a system with no fire protection and at least 100 mm for fire protection systems (Didier, 1980).

1.4.1.5 Hydraulic gradient

Preferred velocity and preferred hydraulic gradient through the system are used generally as a basis for choosing pipe diameters. A method for achieving this is described below.

1.4.2. The sizing of water mains

When distribution flows are known, pipe diameters can be determined on sizing criteria often stated in terms of a preferred velocity at peak flow or average flow (eg 1.5 m/s at average flow to 2.5 m/s for fire situation) or in terms of a uniform hydraulic gradient (eg 4 m/1000 m at peak flow). Given a fixed velocity, the use of the continuity equation allows the determination of the theoretical pipe size as follows:

$$D_t = (K_1 * Q_p/V_d)^{1/2} \tag{1.1}$$

Where

D_t = theoretical diameter, m
Q_p = peak flow, m^3/s
V_d = design velocity at peak flow, m/s
K_1 = $4/\pi$.

If the hydraulic gradient is selected with the use of the Hazen-Williams formula, the size of the pipe can be obtained by:

$$D_t = (K_2/J_d)^{0.21} * (Q_p/C)^{0.38} \tag{1.2}$$

Where

D_t = theoretical diameter, m
J_d = h/L = hydraulic gradient at peak flow, 4/1000
Q_p = peak flow, m^3/s
C = Hazen-Williams C coefficient
K_2 = 10.7 for D_t, L, h in m and Q in m^3/s.

The size D_t found is a theoretical value which is usually rounded up or down to the closest available nominal size.

While these approaches have the advantage of being simple and rapid they are on the other hand less than ideal methods since they ignore the cost factor and hence the optimal pipe size.

1.5 SYSTEM FAILURE AND BREAKDOWN

The concept of failure has usually been associated with events such as breakage of pipes and, perhaps in some cases, loss of pressure due to electrical power failure. In this study the concept is taken to a greater depth, but some preliminary comments are appropriate in this opening chapter.

1.5.1 Overall System Failure

Massive failures such as these should occur only rarely, or never at all. The commonest cause of a temporary, but total, failure is probably loss of electrical power to the motors driving the pumps. They are fairly rare in most developed countries but in some parts of the world will be quite frequent, ie one or more times per week. Far less likely, but not to be completely ignored, are dam breaks and droughts. A more insidious failure is contamination due to the ingress of contaminated water when the system operating pressure is sub-atmospheric. Much more common, however, are failures of sub-sections of a system due to failures of individual components.

1.5.2 Pipeline Failures

Pipe breaks occur despite precautionary measures such as material selection and proper mainlaying. Basically, the variety of factors reported to affect the failure characteristics of water mains are:

(1) defects in manufacture (observed generally immediately after installation);

(2) poor backfill and bedding (because loads on the pipe are not evenly distributed as was assumed during design);

(3) internal/external corrosion of the fabric of the main;

(4) excessive load from traffic;

(5) pressure surging (eg unexpected transient pressures not catered for during design, or due to poor maintenance and possible corrosion/fatigue effects);

(6) ground movement, through expansion of soils, subsidence or, freezing/thawing conditions;

(7) joint deterioration (caused by waterhammer for example), may erode the bedding under a pipe which is then no longer supported adequately and more likely to burst under external loads;

(8) accidental damage caused by operations of other utility (eg new construction);

(9) pipe diameter (the rate of bursts increases as pipe diameter decreases, (Robert and Regan, 1974)).

In particular each type of material is different with respect to the common cause of failures, the type of failures, and the rate of failures.

For cast iron pipes, the chief causes of mechanical failures are ground movements, corrosion and uneven loading. The three common types of failure encountered are circumferential fractures, longitudinal fractures and holes.

Holes are caused by local corrosion whilst longitudinal fractures are generally associated with preferential corrosion damage along a longitudinal section of the pipe and ground loading.

Circumferential fractures are commonly observed with small pipe sizes and generally due to beam loading (Robert and Regan, 1974).

Asbestos cement pipe is made up of asbestos fibres and ordinary Portland cement as outlined above. Deterioration of these pipes is commonly due to external attack in aggressive soils (eg those with low pH and/or high sulphate content) and to conveyance of a negative Langelier Index water, which may leach the lime from the cement matrix leading to a weakened pipe.

Failure of uPVC pipe can result from fatigue caused by rapid pressure oscillation or under normal gravity head conditions. The pipe may fail at a high stress point. These are areas within the pipe caused by:

(1) presence of scratches on the inside surface;

(2) deep notches on the outer surface;

(3) foreign bodies within the pipe wall;

(4) External point loads.

These causes are discussed in detail elsewhere (Stephens and Gill, 1982 and Kirby, 1981).

Some historical data on pipe breakage rates have been published in some studies in the literature (eg O'Day, 1982; Walski and Pellicia, 1982; Weiss et al., 1985; Kettler and Goulter, 1985). The typical approach is to develop a regression equation for the break rates, as a function of pipe diameter of water mains, such as a power regression (Mays, 1989). Robert and Regan (1974) and Kettler and Goulter (1985) have found that the rate of breaks increases as pipe diameter decreases. However for the work of Kettler and Goulter, data collected corresponds to the city of Winnipeg, Canada, which has extreme temperature variation and hence these data should be applied elsewhere only with caution (Walters and Knezevic, 1988). Cullinane (1987) and Su et al. (1987) reported breakage rates for Philadelphia and St. Louis cities respectively. Male, Walski and Slutsky (1990) presented an inventory of water mains (6 in - 24 in) by borough and diameters for five boroughs (Bronx, Brooklyn, Mahattan, Queens and Staten Island) in New York city. Mays (1989) reported rates of breakage for San Diego and St. Louis. It was observed that for the city of San Diego, the cast iron pipes had higher break rates than the asbestos-concrete pipes, by almost a factor of ten for 6 in diameter.

Several authors (Walski and Wade, 1986; Walski, Wade and Sharp, 1987; Weiss et al., 1985) have demonstrated the seasonality of pipe break rates. Typically, break rates increase significantly during a period of cold weather (I.W.E.S, 1984).

A study of the failure records of uPVC and ductile iron in similar conditions carried out by the Water Research Centre (Critchley and Habershon,1981) indicated that the recent failure record of uPVC is as good as, or even better than, ductile iron and this is shown by Fig. 1.2. Previously, there had been fatigue failures. Typical rates of break are presented in Fig. 1.3.

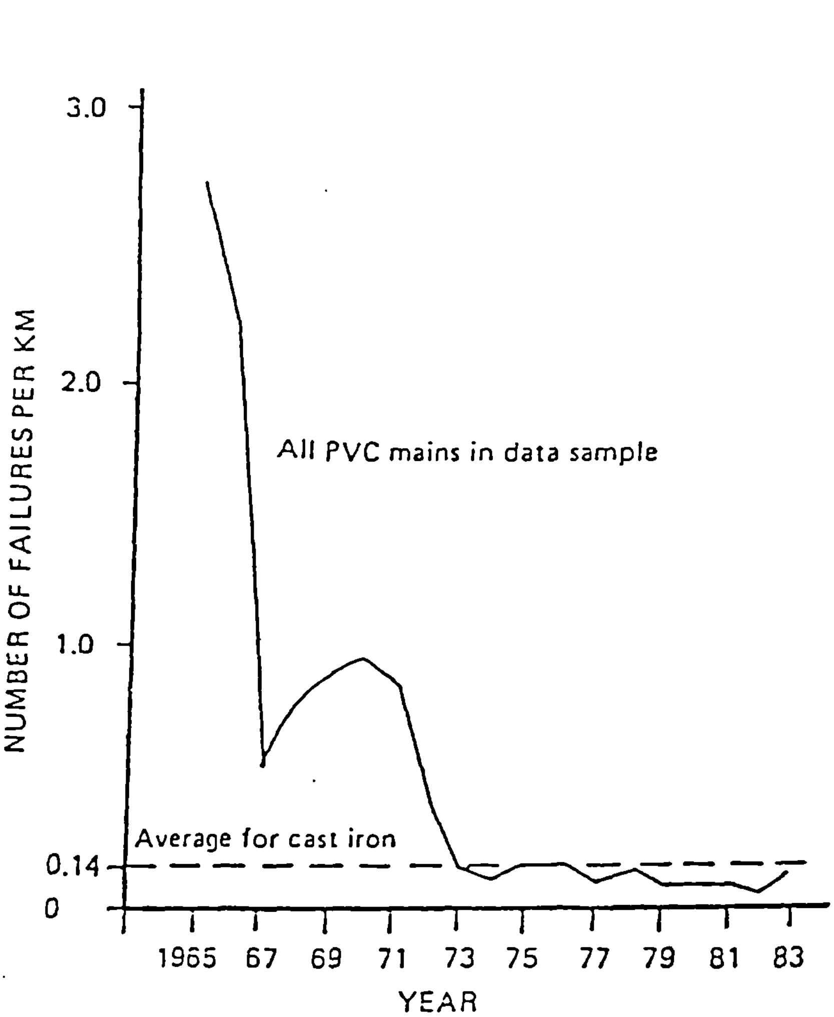


Figure 1.2 ***Annual Failure Rates of uPVC Water Mains***

(I.W.E.S, 1984)

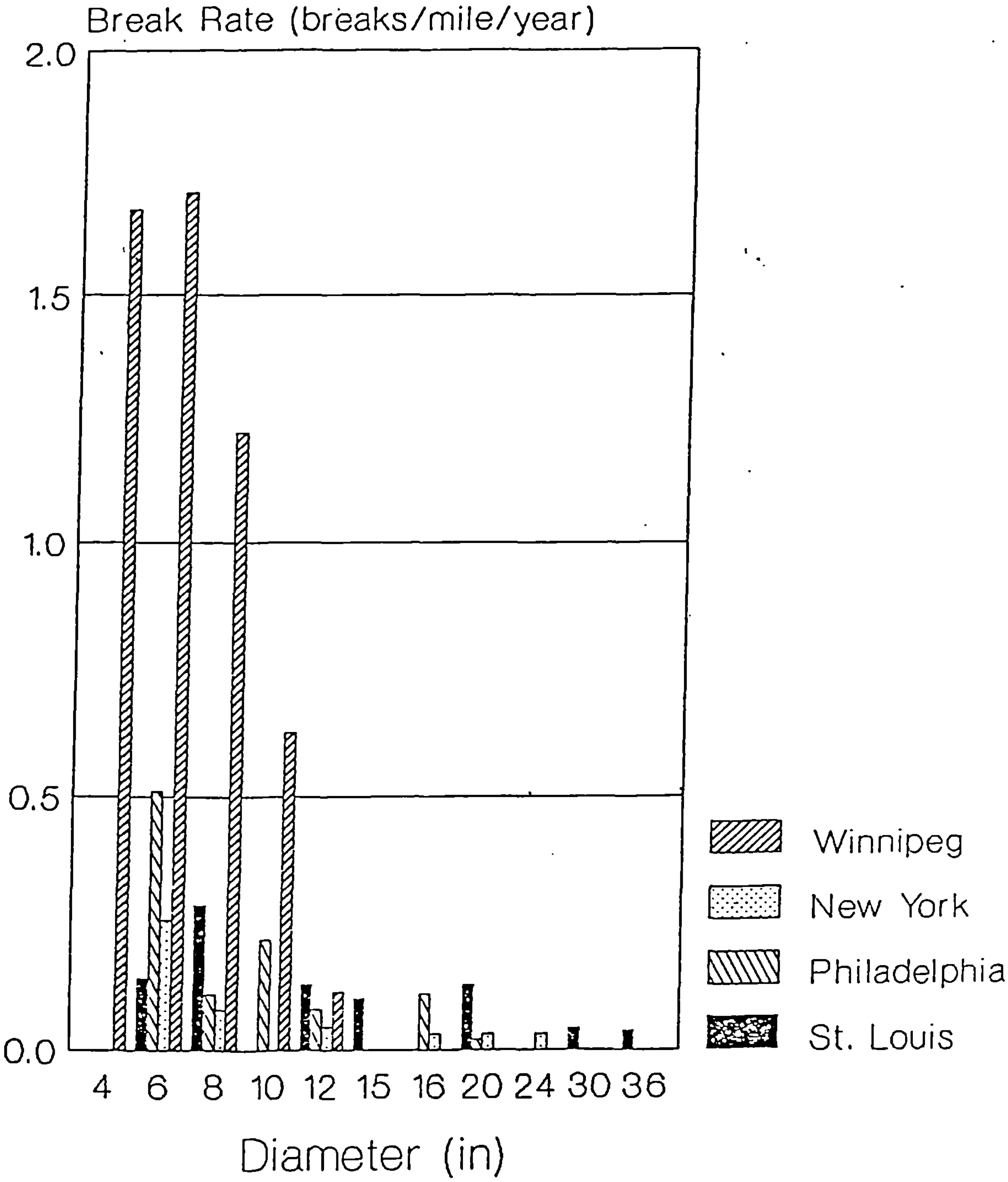


Figure 1.3 Break Rate as Function of Diameter

1.5.3 Component Failure

Design criteria for networks are usually based on the relevant code of practice for maximum permissible pressures in the pipes. However, this will prove useless unless the permissible pressures in the casings and bodies of the various components can also withstand the same pressures, and also impact and shock loads due to the motion of moving parts within these components. This problem which is important is beyond the scope of this work. Thorley (1991) discussed these aspects.

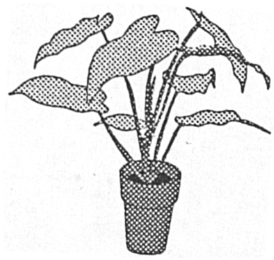
1.5.4 Failure in terms of Lack of an Adequate Supply

All of the above are expressed in the language of the designer, engineer and supplier. Failure, as perceived by the consumer, is when water is not available in sufficient quantity at an adequate pressure. Failure is now perceived as 'lack of availability' of water and a 'reduction in the reliability' of supply. This approach is introducing more subtle aspects of the breakdown of the system.

In the next Chapter, various approaches to the classical design and analysis of water supply networks are developed to give further background. This is necessary since the more refined techniques will still ultimately contribute to the assessment of optimisation with respect to availability and reliability. The full Literature Review will then follow.

Chapter II

WATER DISTRIBUTION NETWORK ANALYSIS



Chapter II

WATER DISTRIBUTION

NETWORK ANALYSIS

2.1 INTRODUCTION

Steady state analysis of flow and pressure in water distribution systems has been and still is a major task for many water engineers. The governing equations are non-linear and cannot be solved directly. Most of the techniques applied for solving these equations involve gradient methods to deal with the non-linear terms. Consequently, convergence problems are always a possibility particularly if ill-conditioned data such as, for example poor pump descriptions or other components are employed. Therefore, there is no absolute guarantee of convergence.

Historically flow and pressure distributions were most often calculated using a loop method ie equations were expressed in terms of the unknown flowrates in the pipes. The node method is another formulation of the problem. That is, the equations are expressed via the unknown heads at junctions within the system. Several techniques have been suggested and tested. A summary of the major techniques available for use on microcomputers has been provided by Thorley and Wood (1986). The most popular techniques are:

(1) **Hardy-Cross** Methods (*flow formulation* and *node formulation*);

(2) **Newton-Raphson** Methods (*flow formulation* and *node formulation*);

(3) **Linear** Method.

Details of these methods can be found in Wood and Rayes (1981). Less commonly used models are based on the Finite Element Method (Collins and Johnson, 1975), the Graph Theoretical Approach (Kesavan and Chandrashekar, 1972) and the Non-linear Optimisation Techniques (Collins et al., 1976).

2.2 BASIC EQUATIONS

A general description of water distribution networks has been presented in the previous Chapter. When links and nodes within a system constitute a closed path, they form loops. Moreover, when junction nodes, fixed head nodes (eg water towers and service reservoirs) and primary loops are identified the following relationship holds:

$$N_P = N_J + N_L + N_F - 1 \tag{2.1}$$

Where

N_P = number of pipes;

N_J = number of junction nodes;

N_L = number of loops;

N_F = number of fixed head nodes.

It turns out that this identity is directly related to the basic hydraulic equations which govern steady state flow in water distribution systems.

The distribution of flows through the network under a certain loading pattern must satisfy two laws: **Conservation of Mass** (Continuity equations) and **Conservation of Energy**. For an incompressible fluid the conservation of mass is replaced by conservation of flow at each junction node:

$$\sum (q_{in}) - \sum (q_{out}) = Q^{ext} \qquad (2.2)$$

Where

q_{in} = flow into the node;

q_{out} = flow away from the node;

Q^{ext} = external demand or supply at the node.

There are N_J continuity equations. For each primary loop, ie an independent closed path which does not contain secondary loops within it, conservation of energy must hold, that is, the sum of head losses in the loop must be equal to zero. If h_L is the head loss in a pipe and H_{pump} is the head of a pump contained in the loop, the following equation holds:

$$\sum (h_L) - \sum (H_{pump}) = 0 \qquad (2.3)$$

Strictly, the total head loss in a pipe must incorporate the kinetic term, ie $V^2/2g$, where V is the flow velocity and g is the acceleration due to gravity (9.81 m/s^2). Since velocities in water networks are in general in the range of 0.5 to 2.5 m/s, the kinetic term is insignificant (0.22 m for V = 1.5 m/s) compared to the other terms in Eq. 2.3 and thus usually ignored.

There exist N_L independent primary loops in the network. Finally, the difference in total head (ΔE) between two fixed head nodes (ground level plus pressure), where the head is constant for the simulation period, must be conserved. If there are N_F such nodes, then there are $(N_F - 1)$ independent equations of the form:

$$\Delta E = \sum (h_L) - \sum (H_{pump}) \qquad (2.4)$$

It should be noted that Eq. 2.4 is more general and Eq. 2.3 is a special case of it, ΔE is equal to zero for a path forming a closed circuit. As a result, the conservation law is expressed by $(N_L + N_F - 1)$ energy equations, and the total number of equations $(N_J+N_L+N_F - 1)$, as outlined in Eq. 2.1 constitutes a set of N_P non-linear algebraic equations.

2.3 ANALYSIS ALGORITHMS

Principally, three iterative approaches have been applied to solving the set of equations described in the previous section. These are the Linear theory, the Newton-Raphson and the Hardy-Cross techniques.

In 1963, Martin and Peters published an algorithm using the Newton-Raphson method to solve the non-linear network equations. Shamir and Howard (1968) showed that pumps and valves could be incorporated as well as being able to solve for unknowns besides the nodal heads.

The Linear method was first proposed by Wood and Charles in 1972 for simple networks and, later extended to include pumps (Tavallaee, 1974).

The Hardy-Cross method (Cross, 1936), originally developed in 1936, is attractive for hand calculation and easily coded. For large complex networks, however, it has been found that it is not reliable since it converges slowly if at all (Jeppson, 1983).

2.4 CHOICE OF ANALYSIS ALGORITHM

From the previous section it becomes clear that the selection of an efficient algorithm for the analysis of water networks is between the Newton-Raphson approach and the Linear method. The former method is reported to converge more quickly and require less computer storage than the latter. A disadvantage of the Newton-Raphson method is that it requires an accurate initial guess to ensure convergence.

The linear method has the following advantages:

(1) Rapid convergence, though somewhat less rapid than the Newton-Raphson;

(2) Does not require an initial starting point;

(3) Has more flexibility in the representation of pumps, and the capability to analyse all components.

In a very recent study, Wood and Funk (1993) studied extensively the reliability of the three main procedures. The study was performed on an extensive data base. These data were provided by consulting engineers and water distribution engineers and represent actual and proposed distribution systems in the US. Most of the data were sent to one of the authors (Wood) because analysis difficulties had been encountered. Different samples of water networks were analysed by the three approaches and the conclusions were:

(1) The Hardy-Cross method has exhibited significant convergence problems especially with larger systems.

(2) The most promising algorithms were the Newton-Raphson and the Linear Methods. The convergence of both of them is virtually assured if reasonable data are employed. This conclusion was also reached and refined by Altman and Boulos (1992).

(3) Of the two, the Linear Method has slightly better convergence characteristics since the Newton-Raphson method showed one failure among the cases tested. No failure, however, was encountered with the Linear Method.

For the above reasons, the Linear method is chosen for the analysis of water networks herein. The author has developed a computer program, **LMANLS**, coded in Pascal. LMANLS will be used as a tool for solving network equations in the computer programs being developed in this research. For reasons of simplicity and rapidity of computation, the Hazen-Williams formulation is retained as a head loss equation in the LMANLS procedure. The Linear method algorithm is presented in appendix A.

Chapter III

LITERATURE REVIEW

Chapter III

LITERATURE REVIEW

3.1 INTRODUCTION

With the availability of computing power, the application of Operations Research methodologies was accelerated such as optimisation and by the mid of 1960s, optimisation techniques for the design of water distribution networks were beginning to appear in the literature (Labye, 1966; Karmeli et al., 1968). Most of these early models were, however, restricted in application to relatively simple branched networks and small looped networks (Jacoby, 1968 and Watanatada, 1973). For these early works, Jacoby who has used a non-linear formulation has noticed that his approach, which will be reviewed later, tended to reduce some links to practically zero and the demand can be satisfied without the loop-forming links. However, in the absence of a minimum size limit of the decision variables (diameters) and if a single pattern demand is considered, the cost optimisation process will automatically remove the redundant links from the design. Consequently, the optimal design will degenerate into a tree configuration and the loops will then be lost (Templeman, 1982).

Among the reasons given by the engineers for adopting loops in the distribution networks rather than trees is that, loops increase system security, flexibility and reliability. In case of failure of one or more links, the demand nodes may be supplied by another path which does not contain the failed pipes. The question of how best to incorporate the reliability issues in the design of water networks has not yet satisfactorily resolved (Fujiwara and Silva, 1990).

Most of the work presented in this Chapter up to 1991 was reviewed by the author (Khomsi, 1991). The review started from 1966 and recent papers (1993) were also examined. This Chapter presents a survey of the literature on Optimal Design and Reliability Analysis of Water Distribution Systems. The Review basically concentrates on three major parts, namely: optimisation of water distribution networks, reliability of water distribution networks and reliability-based optimal design.

3.2 OPTIMISATION OF WATER DISTRIBUTION NETWORKS

This section contains a review of the published approaches to the optimal distribution design *without reliability considerations*. The focus is on looped networks, which is the most common case for urban systems. A great deal of effort related to hydraulic system optimisation had been spent to assist the water engineer. Walski (1985) identified over 60 optimisation problems (branched and looped networks). Table 3.1 summaries most of the work done in the optimisation of looped networks area. Given a specific set of links in the network layout, the optimisation models determine pipe diameters, pump capacities, heights of water towers, and other design parameters subject to the governing hydraulic equations satisfying steady flow conditions and various constraints on pipe diameters, flows, and nodal heads. The problem is even more complicated if constraints are loading conditions (eg pipe bursts and stochastic demands).

The objective function of the models focuses exclusively on monetary costs including acquisition, operation and maintenance costs. Important capabilities of the models include the type of system analysed (looped or looped and branched) the number of sources allowed (single or multiple) and the number of loading design conditions handled. Solution tools range from linear programming (LP) to fully non-linear (NL) optimisation techniques.

To interpret Table 3.1, some comments are required. After the first and second columns where the date and the author(s) are presented respectively, the third column is related to the decision variables selected. In general, three types of unknown are used:

(1) The length of the link. In this case, the pipe ie the link, is assumed to be composed of segments of different diameters. The problem is linear with respect to the link length (Labye, 1966; Karmeli et al., 1968; Gupta, 1969; and Alperovits and Shamir,1977).

(2) The diameter of the link. Two formulations were suggested in the literature. The first one considers diameters as continuous pipe sizes where the cost of the pipe is non-linear in a continuous concave function in the range of commercially available sizes (eg Jacoby, 1968; Watanatada, 1973), while the second one treats diameters as discrete pipe sizes (eg Gessler, 1982).

(3) The length of pipe segments as in the first type. However the objective is not the optimisation of the pipe cost but to minimise the change of cost by replacing one size with another (Kally, 1972 and Morgan and Goulter, 1985).

The fourth column distinguishes between the type of system which can be handled by the model. In fact if the system is fed by a source which is located above it, this system is classified as supplied by gravity. On the other hand, if the source head is insufficient and water is pumped to the system, the system is classified as pumped.

For the pump, if either the head or the flow is known, the cost for pump energy is linear with respect to the other. Some models take this relationship and knowing the flow put the energy cost of the pump into the objective function as a linear term and solve for the head (Robinson and Austin, 1976). Others use non-linear functions and iterative approaches to select the optimum head (Alperovits and Shamir, 1977).

Besides the pipe and the pump energy costs, a number of formulations incorporate the installation cost, called capital cost, and the maintenance cost into the objective function. These formulations consider the pump cost as a non-linear function of horsepower or a term composed of a product of pumped flow and head each raised to different exponents (Watanatada, 1973; Shamir, 1974; Deb, 1978; Rowell and Barnes, 1982).

Year	Author	Decision Variable	Type of System	Algorithm and Comments
1968	Jacoby	D	Pumped	NLP, unconstrained by incorporating flow constraints into a Merit function. Single loading condition
1969	Lai and Schaake	D	Pumped	LP with assumed pressure surface Multiple loading condition.
1972	Kally	X_{ij}	Pumped	LP, unknowns are the change of lengths of given diameter in a link. Single loading condition
1973	Watanatada	D	Pumped	NL, unconstrained by Lagrangian of flow constraints. Single loading condition
1974	Shamir	D	Pumped	NLP, same formulation as Watanatada but takes into a count other components. Multiple loading condition
1977	Alperovits Shamir	X_{ij}	Pumped	LPG, Two-level hierarchical scheme LP with gradient correction based on change in head. Multiple loading condition
1981	Quindry Brill Liebman	D	Pumped	Two-level hierarchical scheme similar to Alperovits and Shamir model but gradient with respect to change in Flow. Multiple loading condition
1982	Rowell Barnes	D^*	Pumped	Two-level scheme, NLP for layout using single load and design assumption then Integer Programming for adding redundancy. Multiple loading condition

D^* = Discrete Diameter
D = Continuous Diameter
X_{ij} = Segment of pipe of diameter i in link j
LP = Linear Programming
NLP = Non-Linear Programming
LPG = Linear Programming Gradient method of Alperovits and Shamir (1977)

Table 3.1 Water Distribution System Optimisation Models (continued)

Year	Author	Decision Variable	Gravity Pumped	Algorithm and Comments
1982	Gessler	D^*	Pumped	Enumeration Multiple loading condition
1985	Morgan Goulter	X_{ij}	Gravity	LP similar to Kally's formulation, optimal layout is found using heuristic. Multiple loading condition
1987	Fujiwara Jenchaimahakoon Edirisinghe	X_{ij}	Gravity	Modified LPG method of Alperovits and Shamir. Single loading condition
1987	Lansey Mays	D	Pumped	NLP using Augmented Lagrangian function to deal with head bounds. Multiple loading condition
1990	Fujiwara Khang	Q H	Pumped	Two-phase Decomposition approach. In the first a NLPG method similar to that of Alperovits and Shamir but the gradient is computed using the optimal Lagrange multipliers. In the second phase a NL program is solved where the head losses are fixed along the links. Multiple loading condition
1991	Kessler Shamir	Q D	Pumped	Decomposition scheme that uses a Minimum Cost Flow Algorithm to determine flow and similar approach to that of Shamir and Alperovits to find diameters. Multiple loading condition

D^* = Discrete Diameter
D = Continuous Diameter
H = Head
X_{ij} = Segment of pipe of diameter i in link j
LP = Linear Programming
NLP = Non-Linear Programming
LPG = Linear Programming Gradient method of Alperovits and Shamir (1977)
NLPG = Non-Linear Programming Gradient method of Fujiwara et al. (1990).

Table 3.1 Water Distribution System Optimisation Models

The final column gives a very brief comment on the optimisation technique used and the capability of the model to handle multiple loads or not.

Good reviews of this work have been outlined by Shamir (1974, 1979) who has summarized the approaches developed during the 1970s, Walski (1985) and Walters (1988). The most important models are discussed below.

In 1968, Jacoby suggested the use of a numerical gradient technique to minimise the looped hydraulic network cost. The procedure starts with an arbitrary set of diameters and flows or head losses and moves in a random direction, seeking for a local optimum. The objective function to be optimised, called the Merit function, was the combined cost of pumps and pipes and penalties for violation of the continuity and the energy equations. The procedure amounts to a Lagrangian function. No guidance, however, is given on how to change the penalty weights to move them to correct values. In addition, the author concluded that the procedure outlined would require a good deal of engineering judgement.

Following the principal idea of using non-linear programming techniques for optimally designing hydraulic distribution systems, Watanatada (1973) and Shamir (1974) developed optimisation procedures incorporating a formal Lagrangian function and adding the equality constraints to the objective functions. Watanatada's approach used a combination of Box (1966) and Haarhoff and Buys (1970) and the variable metric techniques (Davidon, 1959) for transforming the constraints into the objective function to solve the non-linear problem, while Shamir's procedure applied the reduced gradient technique. The initial cost of pumps and multiple loads could be considered in the formulation.

A second promising line of development for the optimisation of looped systems was the use of a hierarchical scheme, which fixed either the flows or heads, solved a linear programming problem for the optimum pipe diameter then, using the dual variables from the linear program, adjusted the flow or the pressure distribution. In effect, the formulation of the problem as a linear programming one was first proposed by Labye (1966) for optimising branched networks. A link is considered as being made up of segments of different sizes and the sum of segments of different diameters equals the length of the link. The optimal design problem is then a linear programming problem with respect to the length of different diameters. The advantage of this linear programming modelling is the use of discrete pipe diameters that are commercially available, rather than the direct use of diameters as decision variables which are considered as continuous during the optimisation process, and which, once the optimal solution is reached, have to be rounded up/down to commercially available sizes.

In their 1977 paper, Alperovits and Shamir took advantage of the easily accessible and computationally efficient linear programming algorithms and extended Labye's model by adding loop equations to their formulation (the sum of head loss around each loop is equal to zero). The first level algorithm is to start with a known flowrate distribution satisfying the continuity equations and to set up the linear programming model with the only unknown being the pipe segment lengths. In the second level, the authors made use of the dual variables from the sub-optimal linear program to find a cost descent direction, ie a gradient search with which the linear problem is modified. They adopted a fixed step size strategy in which a fixed step size is tried, and if the cost of the linear program decreases, a new LP problem is set up with the newly obtained flow rates. If the cost of the linear program for this fixed step size increases, reduce the step size and try again. If the cost of the LP for a very small step size increases, stop. In the later part of their article, Alperovits and Shamir extended this technique which they have called Linear Programming Gradient (LPG) method to the case of multiple loading conditions and some other cases.

Apparently, the descent direction found by the authors is actually the Steepest Descent direction (Fujiwara et al., 1987). In non-linear programming theory, however, it is well known that this approach is the most intuitive but also the slowest descent direction. Indeed, in practice it requires hundreds of iterations to make very little progress toward the solution. Moreover, Quindry et al. (1979) published a comment with the correction of the gradient terms in the original method which improved the solution in the example problem.

Fujiwara et al. (1987) solved the drawback mentioned above, which is related to the proper search direction in the Alperovits and Shamir model. In fact they have used the BFGS method (Broyden-Fletcher-Goldfarb-Schanno) to determine the search line, and the backtracking line search (Dennis and

Schnabel, 1983) to determine the step size. A detailed theoretical analysis of the LPG method has been presented later by Kessler and Shamir (1989).

In their 1981 paper, Quindry et al. published a procedure where the pressure was considered as fixed in the sub-program and updated by a hierarchical scheme. Based on the work of Lai and Schaake (1969), the sub-program can determine the optimal continuous diameter for a link. A gradient term was derived from the node equations and directed the change in the pressure distribution for the next iteration. However, this derivation was incorrect as stated by Templeman (1982). Templeman pointed out that the linearization of the original cost which is originally not linear was a source of difficulty and error. Indeed, the global solution resulting from the linearized problem may not be even a minimum at all for the original problem. In addition, as the optimisation problem was formulated, other hydraulic components cannot be designed.

Gessler (1982, 1985) proposed a heuristic model based upon an enumeration procedure to seek the best solution. With the availability of a set of possible pipe sizes, the algorithm performs cost and hydraulic analysis tests on all potential combinations to determine a list of alternative optimal solutions. Logic is incorporated in the model to perform size and cost tests which discard some combinations without performing the analysis of the network. He then noticed that his enumeration technique is NP-hard, ie the computer time required to solve the problem increases exponentially with the number of unknowns. However, he contends that by limiting the number of diameters and grouping pipes through engineering judgement, the problem can be reduced to a size which can be handled within a reasonable time. No indication was given on the network size that can be handled by the model. Furthermore the proposed scheme is not an optimisation technique that would produce the global optimum design.

Morgan and Goulter (1985) published a heuristic algorithm to analyse hydraulic systems under multiple loads and determine the optimum layout and design of looped water networks. Two standard techniques were used for this purpose. The linear programming algorithm and a Hardy-Cross network solver. This model uses heuristics to define the critical nodes under different loading conditions, and their pressure constraints acting as the active constraints in a linear program. The objective considered is not the cost of the pipe but to minimise the change of cost by replacing the size of a pipe with another (Kally, 1972). A change of this nature in a given link thus required constraints in the model to ensure that the sum of segments of length of pipe equals the link length. A second heuristic was developed to remove redundant links on the basis of a weighting procedure. The global algorithm works as follows. First, an initial solution of the layout and pipe diameters is given and analysed by the Hardy-Cross solver. Second, a heuristic procedure is performed to remove unnecessary links on the basis of weighting each link, ie small link weight indicates that the link does not fully contribute to the distribution of flows and pressures. Next, given the new layout, the new flow distribution and weighting, the linear programming model gives the least cost pipe diameters. Finally, the new pipe sizes are passed back to the network simulator to compute their corresponding set of flows and the iterative procedure continues until the value of the objective function is zero, indicating that the linear programming model does not choose to replace any part of the network and the maximum weighting for any link in the system is greater than a previously specified value.

Lansey and Mays (1987,1989) developed a non-linear model for designing water systems under multiple loads. Besides the pipes, the model is able to size other components such as pumps and reservoirs. In the conclusion of the papers, the authors stated that their model, like the previous ones, should be considered as a guidance tool, that cannot fully solve the problem. In addition, as mentioned above the tendency towards a branched configuration by applying the model to an initially looped system is unavoidable.

In recent years the optimisation of water systems has been performed by using some decomposition techniques so as to reduce the size of the original problem with the objective of finding better local minima.

Fujiwara and Khang (1990) proposed a non-linear two-phase decomposition method for the design of networks. In each step of the optimisation process, the first and the second phases are applied. With reference to the first phase, a new method which the authors called the Non-Linear Programming Gradient (NLPG) method was suggested. The NLPG is basically similar to that of Alperovits and Shamir (1977) and Quindry et al. (1981). Indeed, in these previous works the gradient search was derived from the dual variables whilst in this phase the gradient is computed using the optimal Lagrange multipliers of a non-linear solution. As the decision variables were link flows and pump head, the gradient is computed to modify these variables so as to find a reduction in the system cost. In the second phase however, the authors made use of the solution obtained from the first phase to formulate another non-linear programming problem by fixing the head losses along the links and the resulting program is solved for the flow distribution and pump heads. The new solution is used again as the starting point for the first phase and the second phase is processed. These two phases are repeated until no further improvement can be achieved.

The overall algorithm was applied to the extension of the New York water supply system and it was found that the optimal solution is better than those published previously (eg Quindry et al., 1981 and Morgan and Goulter, 1985). The model was extended to deal with new expansions of existing networks, multiple sources and multiple loading problems and it was stated that the model cannot guarantee the global optimality of the obtained solution.

In 1991, Kessler and Shamir suggested a decomposition scheme for the optimisation of water networks. They proposed a global program in which the objective function is the minimisation of the total system cost (pipes, pump and water tower costs) and the constraints were the continuity equations, the head constraints (energy equations) and bounds on nodal heads (minimum and maximum heads) and link flows (minimum flow required). This program was decomposed into two sub-models: the Fixed Head sub-model and the Fixed Flow sub-model. These will interchange information (head, flow) until no better solution can be obtained. The reduction in the fixed head model was achieved through:

(1) considering heads as fixed values satisfying a feasible solution;

(2) considering only the continuity equations with minimum limits on link flows;

(3) expressing the objective function elements in terms of link flow: for pipe costs this can be obtained by eliminating pipe diameter from the Hazen-Williams equation.

This formulation is similar to some of the published researches (eg Lai and Schaake, 1969 and Quindry et al., 1981), except for the decision variable: link flow is considered in this model whereas pipe sizes were selected in the others. As formulated, the authors used the efficient minimum cost flow algorithms (Kennington and Helgason, 1980) to solve for the flows.

For the fixed flow sub-model, it is basically the Alperovits and Shamir (1977) model, except that the head constraints are defined by the incidence matrix and another formulation for the pump. It was noted that usually the solution is obtained after two iterations but the problem of more than one loading conditions was not successfully solved.

3.2.1 Conclusion

In summary, there is an obvious tendency to use the two level algorithm along with decomposition techniques in the field of hydraulic distribution systems in recent years. They are creative and heuristic in nature. However they lack the ability to design a complete water distribution network. Their limitations are due to the size of the network, the number of loading conditions and the type of hydraulic elements designed.

A common characteristic of all the models discussed is that they can only guarantee a local minimum with a sub-branched configuration if the system is forced to be looping. Otherwise, all redundancy is removed by the optimisation process resulting in a tree network. This is obvious since it is always cheaper to supply a fixed quantity of water to a node by one pipe rather than by two or more. The cost optimisation process will reduce the redundant links down to zero and remove them from the design.

3.3 RELIABILITY OF WATER DISTRIBUTION SYSTEMS

Compared to reliability analysis literature in other fields such as power supply systems, computer and chemical engineering etc, the literature on reliability analysis for water networks is scant. This section is divided into two parts:

(1) Water supply reliability;

(2) Water distribution network reliability.

The early work concentrates on reliability computations of the source or pumps, assuming the conveying pipeline and the distribution network are perfectly reliable. Later work recognizes the importance of the distribution system reliability and recommends some methodologies.

Since the report is aimed at water distribution networks, water supply reliability will be briefly described. Excellent and comprehensive references will be given.

3.3.1 Water Supply Reliability

Basically the methods that have been used for the assessment of water supply reliability are:

(1) Fault Tree analysis;

(2) Frequency and Duration analysis;

(3) Markov Chains methods.

For a general study of reliability, the system may be viewed as a single supply area connected to a single demand area (Fig. 3.1).

Reliability for such configurations has been developed by Endrenyi (1978), Billinton and Allan (1984), Shamir and Howard (1981, 1985) and Hobbs (1985b). More references on these reliability methods and their applications can be found in Billinton (1972) and the IEEE subcommittee on the Application of Probability Methods (1978).

For a rather more detailed analysis the system's supply may be modelled with more components (Fig. 3.2). Each component of the system is characterized by the probability function of time to failure and time to repair. When these distributions are assumed to be exponential, the Mean Time Between Failures (MTBF) and the Mean Time To Repair (MTTR) are the only parameters required to entirely characterize the system (Wagner et al., 1988a). These models have been applied by Tangena and Koster (1983), Shamir and Howard (1985) and Hobbs (1985a).

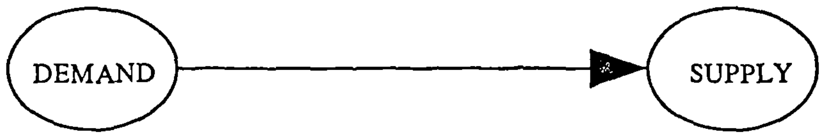


Figure 3.1 System with Lumped Supply - Lumped Demand

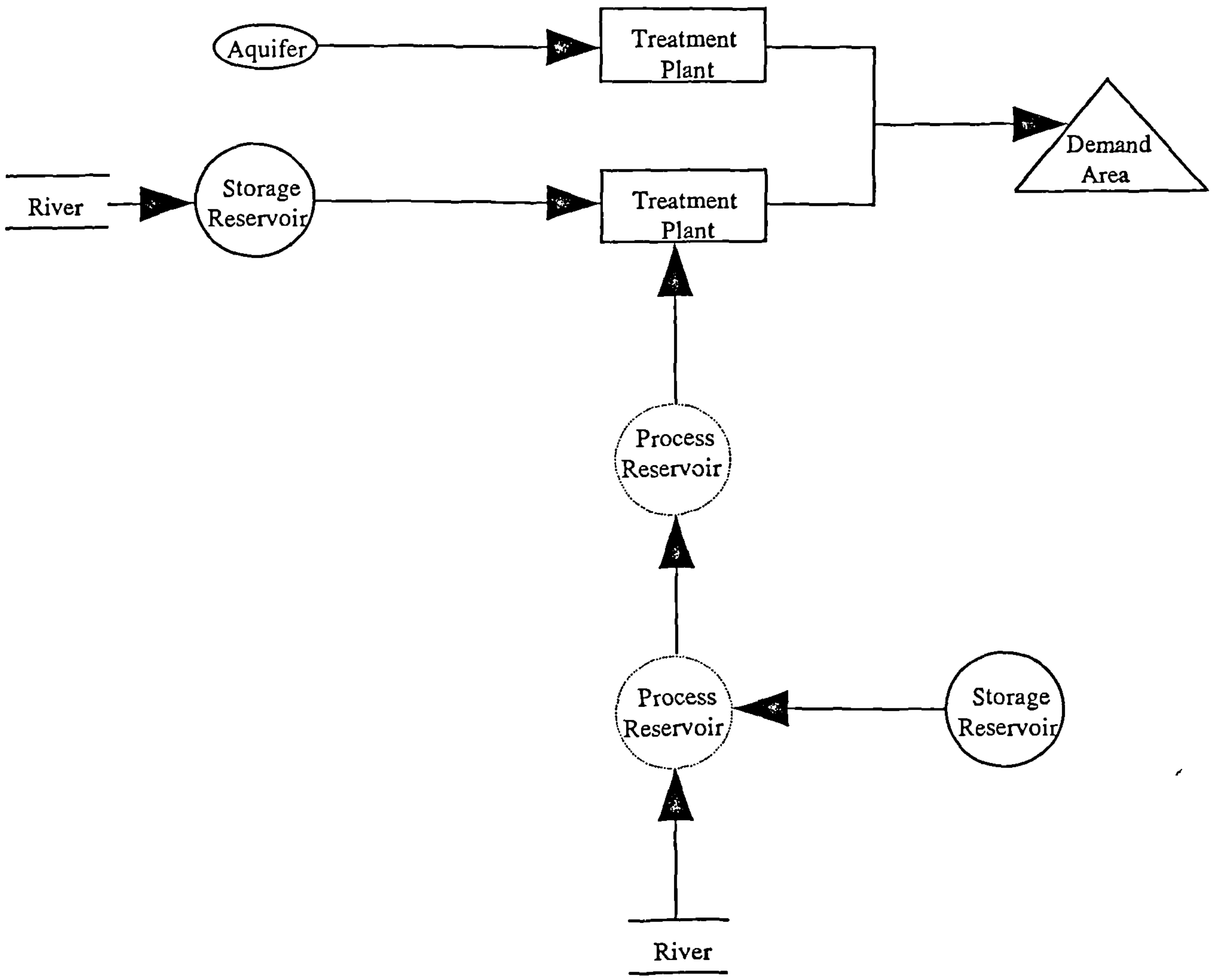


Figure 3.2 Bulk Transmission System (Tangena & Koster, 1983)

Fault Tree analysis is also often used to analyse water supply systems. Fault tree analysis considers the different ways in which component failures lead to supply shortfall and computes the associated probabilities. Henley and Kumamoto (1981) give useful introductions to this field. Willie (1978) and De Jong et al. (1983) provide an application of these methods to water supply systems.

However, Shamir and Howard (1985) pointed out that the fault tree method is used to calculate availability at demand junctions by means of the cut set procedure. They added that this method is workable only as long as the system is not too complex, since identifying all the cut sets becomes complicated and carrying out the computations is expensive.

Markov chain methods were also used for evaluation of water supply systems (Beim and Hobbs, 1988).

The concept of Frequency and Duration analysis was also used (Hobbs, 1985b; Hobbs and Beim, 1986; Duan and Mays, 1987,1990). This method indicates how frequently shortfalls of a given severity occur, and how long they last, and is one which is used in many electric utility planning studies (Billinton and Allan, 1983). Duan and Mays (1990) provided applications of the method by treating five numerical examples.

3.3.2 Water Distribution Network Reliability

This section surveys the methods for computing the reliability of systems which have the classical series-parallel structure and a general topology. It also presents the different ways in which reliability was assessed.

Reliability is one of the system objectives. According to British Standards, the definition of reliability of a non-repairable manufactured item or component in a system (B.S. 3811, 1974) is "The ability of an item to perform a required function under stated conditions for a stated period of time. This may be expressed as a probability". A similar US definition (Mays et al. 1986) is "The probability that a system performs its mission within specified limits for a given period of time in a specified environment".

At present (1994) researchers do not concur on a universally accepted definition of reliability for water distribution networks even for existing systems.

Network reliability analysis models have been successfully developed and applied for electrical, chemical and mechanical engineering processes, and the electronics industry but do not exist in water distribution system analysis.

One excellent and comprehensive literature review on the topic of reliability analyses has been presented by Mays and Cullinane (1986) and was later up dated by one of the authors (Mays, 1989).

Basically, water network reliability was assessed by means of:

(1) Analytical Methods (eg Connectivity, Reachability, Path Enumeration (path sets, cut sets etc));

(2) Mechanical and Hydraulic Availabilities

(3) Stochastic Simulation Methods (eg Monte Carlo)

3.3.2.1 Analytical Methods

Analytical methods can provide a good assessment of the reliability of networks, but are generally computationally intensive and time consuming. In general, these methods relate reliability to connectivity and reachability. Connectivity refers to the connection of every node to at least one source while reachability corresponds to a situation where a specified node is connected to a source (Wagner et al., 1988a).

From the reliability theory (Billinton and Allan, 1983), calculation of mathematical system reliability requires knowledge of the precise reliability of the basic subsystems or components and the set of all possible subsystem (component) failures. If a system has a series-parallel structure, the assessment of its reliability is a straightforward task. However for complex systems, additional modelling and evaluation techniques are necessary in order to determine the reliability of such systems. Those techniques include the conditional probability approach, cut, path and tie set analyses, tree diagram (event tree), logic diagram (Fault tree) and connection matrix techniques. Most of these techniques which are of similar form but derived from different lines of consideration (Tung, 1985) transform the logic operation of the system or topology of the system into a structure that consists only of series and parallel components, paths or branches.

The methodology presented by Mays et al. (1986) for computing the minimal cut sets was the first to correctly apply the definition of minimal cut sets in stochastic process theory to the water distribution systems that account for the hydraulics of the system. This methodology simulates the performance of a pipe system under various failure modes of pipes using a network solver and detects the minimal cut sets by comparing the computed hydraulic heads with the required hydraulic heads. It was implemented on a numerical example with seven links, two loops and four nodes. Nodal and system reliabilities were computed using reliability theory.

Tung (1985) states that there is very little work done on attempting to quantify the system reliability of water distribution networks. Six techniques were briefly described for evaluating reliability of a system with complex configuration in his paper. These techniques are:

(1) conditional probability approach;
(2) connection matrix method;
(3) cut set analysis;
(4) path set analysis;
(5) event tree analysis;
(6) fault tree analysis.

Reliability was defined as the probability that flow can reach all the demand points in the network. Assuming that all pipes within the numerical example given have the same failure probability (5 per cent), the author found that five methods out of the six outlined (path set is not included) yield practically the same system reliability. But from the computational view point the cut set method with a first order approximation (ie one pipe is cut at a time) is the most efficient. However this conclusion is based on analysing only a very simple network (9-pipe and 7-node network).

Quimpo and Shamsi (1987) used the concept of terminal-pair reliability which is defined as the probability that a specified vertex (node) can communicate with another specified vertex in the stochastic network. A stochastic network consists of arcs and vertices that fail independently. In water distribution networks this is illustrated by the probability that a demand point receives water from the source on the basis of the connection of the demand point and the source as stated by the authors. They then choose arbitrarily a minimal path and cut set analysis for calculating terminal-pair reliabilities, which represent the reliabilities from the source to different nodes called point reliabilities. A path set is a set of arcs which forms a connection between the source and demand point, and is minimal when it fails if any of its arcs fail. On the other hand, a cut set is a set of arcs whose removal disconnects the source from the demand points. It is minimal if it contains no cut set which is also a cut set (Billinton and Allan, 1983). As pointed out in the summary of the paper, point reliability values can be used to construct a reliability surface for the whole network which may help water authorities to identify sectors of the system which need maintenance or rehabilitation.

In their second paper (Shamsi and Quimpo, 1988) a preventive maintenance strategy based upon the use of network reliability using the same techniques mentioned above to determine point reliability, was presented. These nodal reliabilities are compared to a threshold level specified by management. Whenever the computed reliabilities are less than the preset limit, their locations have to be found and the corresponding pipes/blocks of hydraulic elements have to be repaired or replaced according to the cost of each operation ie if a pipe cannot be repaired at a cost less than that of replacement it should be replaced. The technique outlined was applied to a numerical example. Reliabilities of individual components or strings of components connected in series with each other, such as a pipe, pump and valve, were evaluated using the exponential distribution (Billinton and Allan, 1983) for constant failure rate of individual components. It should be noted that no hydraulic simulation was performed to find the path sets in the example problem. They were found using graph theory.

Wagner et al. (1988a) presented three analytical approaches for evaluating system reliability: reachability, connectivity (as defined above), and a third measure expressed in terms of the probability that a system can meet a specified level of flow at each junction. Analytical solutions to the reachability and connectivity problem were presented based on the algorithms of Satyanarayana and Wood (1982) and Rosenthal (1977). Since a reachable node may still not receive sufficient supply under a preset pressure, they then defined the third issue of reliability as the probability that a given node receives its demand under random component failures. The three indices were computed for determining the reliability of two numerical examples and it was found that the probability of the system being able to deliver the required supply is less than the probability of it simply being connected. However, for the third index of reliability, the procedure assumed that demand was constant throughout the day.

3.3.2.2 Mechanical and Hydraulic Availabilities

In this class of approaches hydraulic availability is defined as the percentage of time that the demand can be delivered at or above the required residual pressure, while mechanical reliability is measured in terms of the probability that the component is operable at time t given that it was as good as new at time zero (Cullinane, 1986).

Cullinane (1986) presented concepts for evaluating water distribution system reliability using both mechanical and hydraulic availability concepts for assessment of nodal and system performance. The two concepts were combined to give an expected value of nodal reliability. The author stated that reliability at critical nodes may be much more important than the overall system reliability and then he defined system reliability as the average of the total nodal reliabilities. However, since all nodes did not have the same importance, the procedure for evaluating system reliability should take into account this characteristic by, for instance, weighting each node by its demand. The author combined the mechanical and hydraulic availabilities using expected value analysis. The proposed procedure calculated the hydraulic availability with and without specified links operational. The expected value of the availability was computed as the sum of the product of the link mechanical availability and associated hydraulic availability plus the product of link mechanical unavailability and associated hydraulic availability. Link failure was assumed to be independent, and a heuristic was proposed to identify candidates for failure evaluation.

In his second paper, Cullinane (1987) applied the procedure to a small numerical example (2 loops, 6 links and 5 nodes). Hydraulic availability was computed using an existing extended time period network solver WADISO, (Headquarters Department of the Army, 1987), while the mechanical availability, of each link was calculated on the basis of MTBF and MTTR for a repairable component, drawn from mechanical and electrical systems reliability (Billinton and Allan, 1983). Multiple link failures were evaluated; however, it was concluded that in general, multiple link failures could be neglected because of the high availability of system components. This is in agreement with the conclusions of other researches (eg Wagner, Shamir and Marks, 1986). Moreover, the approach is useful, compared to previous methods, since it considers flow performance factors as well as component failure. However, only steady flow patterns were used.

3.3.2.3 Stochastic Simulation Methods

More realistic measures of reliability can be offered by stochastic simulation techniques but the major problem with these techniques is linked to the large number of calls for network analyses to evaluate availabilities. Simulation methods allow the analyst to obtain a variety of reliability indices such as the number, location, duration and impact of failures; furthermore, they allow greater flexibility in the choice of the type of system component to be analysed.

The simulation approach proposed by Wagner et al. (1988b) consists of two parts:

(1) the simulation section which generates failure and repair states according to pre-established probability distributions; and

(2) the hydraulic network solution section which solves the non-linear hydraulic equations for the heads and flows for the full or reduced network.

The simulation proceeds by randomly generating failure times of pipes and pumps according to specified probability functions. When a link is out of use, it is removed from the network. The corresponding heads and flows to the reduced system are obtained using a network solver. The new heads at the demand nodes are used to predict the behaviour of the system. Once a link has failed, a random repair time is generated. The system is then assumed to function in the partially complete configuration until the repair time is reached, another pipe fails, the reservoir empties or the pump fails. The simulation program records the total duration of normal, and reduced service, the failure mode at each node, the nodal shortfall and other measures of reliability such as total duration of failure time for each pipe, or the total number of breaks within the system. It was noted in the paper that the simulation method is time consuming but no ideas were given about the period of time spent in analysing the two numerical examples presented and, simulation runs are hard to optimise. The authors of this paper have shown the flexibility and the usefulness of the simulation methods. However, like many previous works the stochastic nature of the demand was not considered.

Bao and Mays (1990) work presented a method for the assessment of both nodal and system reliabilities. The method used random demands, pressure heads and pipe roughnesses generated using Monte Carlo techniques. The random variables generated are then tested by a network solver (KYPIPE, Wood, 1980) and the nodal reliability, which was defined as the probability that a demand node is satisfied (the probability that the pressure head at the given node is greater than or equal to the minimum bound required), can be computed. For the system reliability, three measures were suggested: (1) the minimum nodal reliability within the system, (2) the mean of all nodal reliabilities and (3) the mean of all nodal reliabilities as in (2) but weighted by their demands.

The number of data sets of demands, pressure heads and/or roughness coefficients, the type of probability distribution and the the parameters for the random variables must be entered into the computer program written for this purpose. It was found that about 500 iterations were required for the number generation and hydraulic simulation to compute nodal and system reliabilities. Moreover mechanical failure could be added to the program, requiring a random variable to indicate when the component fails. However, as the computed reliabilities were unrealistically low (eg 0.915 at node 4), far more simulation would be required to predict higher reliability values. Monte Carlo simulations are suitable techniques for evaluation of the reliability of even complex systems; however, they are expensive to run and cannot provide precise estimates of reliability without long run times. This limits its practical application especially within an optimisation scheme for water networks.

3.3.3. Conclusions

It appears from the hydraulic network reliability review that:

(1) there is no standard universal definition of reliability of water systems;

(2) there is also no accepted single technique for evaluating reliability;

(3) reliability indices reported are scarce compared to other fields such as power supply systems, computer and chemical engineering;

(4) the best approaches are those which define reliability in terms of supplying a node sufficiently, not simply in terms of its connection to a source. Indeed, in hydraulic systems, a node already connected to the system may receive insufficient supply or no supply at all if the pressure is below the required levels.

(5) stochastic simulations may help as techniques for the assessment of water network reliability.

3.4 RELIABILITY-BASED OPTIMAL DESIGN OF WATER DISTRIBUTION NETWORKS

In his 1968 paper, Jacoby noticed that his approach tended to automatically reduce some redundant network element sizes to practically zero and the prescribed node flow requirements can be satisfied without the links which form loops.

Watanatada (1973) raised the problem of reliability for the first time when he tried to optimise a simple water distribution network composed of four nodes, five links and one pump. He showed that an important characteristic common to the networks tested was the trend of the network to be branching by the redundancy of some pipes and their tendency to disappear. This was examined by studying effects of the minimum permissible diameter (D_{min}). When D_{min} is equal to zero the cost is the cheapest one and the network has a tree-like layout. As D_{min} increases the network gradually changes from a branched configuration to fairly uniform sized looping arrangement at $D_{min} = 18$ in. The transition from $D_{min} = 0$ to 24 in was associated with a rise in total cost of about 40%, which, as he has asserted may be loosely regarded as the price to be paid for a more reliable system. At the end of his paper, he suggested that future work should be on the development of a more complete network model that explicitly incorporates measures of reliability.

Alperovits and Shamir (1977) noticed that when a network is designed for a single loading, the optimal design will have a branched configuration unless a minimum permissible diameter is specified (sub-branched network). They suggested that more work should define the network reliability, not in terms of forcing the network to have a fully looped configuration but, in terms of a performance criterion for specified emergency situations.

The traditional approach to reliability in a distribution network is to provide loops throughout the system (redundancy). Bhave (1978) suggested a relatively simple technique which determines the least cost minimal spanning tree ie branched network, and simply joins the ends of the branches to create the required redundancy loops. Rowell and Barnes (1982) tried to optimise the hydraulic system by means of optimising the layout. The reliability was provided by ensuring that each node has an alternative means of supply. Two programs were used. The first one is a NL programming model which determines an economical tree-like configuration for the major pipe links, whilst the second program, which incorporates an integer programming model, selects the loop-forming links to add to the former program in order to minimise the cost of providing a specified level of reliability ie any node is connected to at least two links. A common disadvantage to these methods is the fact that they provide a sub-branched network. In addition, they do not guarantee the maintenance of hydraulic consistency, as Goulter (1987) pointed out.

Goulter and Morgan (1985) have developed two linked linear programming models for both sizing and determining the layout of water distribution networks. The reliability was expressed in terms of redundancy. Looping constraint sets were incorporated in the layout program which will guarantee that each node is connected to at least two links. The model maintained hydraulic consistency but it was very optimistic since it assumed, rather than checked, that all connections to a junction were independently capable of supplying the required demand.

Kettler and Goulter (1983) reported an attempt to use the probability of component failure in major supply paths. In this approach, the reliability issue was incorporated directly in a linear programming model through constraints restricting the average number of breaks per year permitted in each pipe. A Poisson distribution was used to compute the relevant probabilities of interest on the basis of the average failure rate in each link. The technique, which is iterative, uses the dual variables of previous linear programming solutions to find "improved" solutions. However, one of the authors pointed out (Goulter, 1987), that the authors were not able to solve explicitly for reliability according to their own definition.

Coals and Goulter (1985) and Goulter and Coals (1986) presented three approaches by which the probability of failure of individual pipes can be related to a measure of the overall system reliability in linear programming minimum cost design procedures. The first approach considers the probability of the failure of a path supplying a junction. The second approach addresses the probability of node isolation ie the probability of the simultaneous failure of all links supplying a node. The last approach assumes that all links connected to a junction should be able to supply the demand of that junction on an individual basis. The second approach is more reliable than the first one, as was claimed by the authors, however, a node may be isolated without the failures of the links directly connected to it. Indeed, other pipes not directly connected to a node may fail and then isolate the node and possibly other nodes. This approach (node isolation) can be improved by examining the probability of such events in the same way as the calculation of the probability of node isolation for those pipes directly linked to the node. In such cases, a new problem will arise associated with the number of all possibilities of node isolation which increases the computational effort for a real network. In addition the probability of node isolation possesses a theoretical weakness as was asserted by Su et al. (1987) since a node may not be adequately supplied even if there is one link connecting it to the rest of the system. Su et al. pointed out also that the assumption made for the third approach (all the pipes connecting to a node have similar diameters) is not applicable in real pipe networks.

Tung (1986) approached the problem by assuming that the required head, pipe roughness coefficients and demand node are random variables with known probability distributions using a chance-constrained model (Charnes and Sterdy, 1966) for least cost design. Knowing the probability distribution of demands, required heads and roughness coefficients which were assumed normal, a series of chance constraints were formulated as follows:

(1) the probability that the network flow into each node is greater than the actual demand was greater than a pre-specified value, α;

(2) the probability that the head is greater than the required minimum was also greater than some pre-specified value, β.

These two constraints were transformed into their deterministic equivalent (see Charnes and Cooper, 1963) before running the chance-constrained model. The reliability consideration was expressed in terms of the uncertainties in the design procedure, and may result in a more 'reliable' design than would be determined on the basis of average conditions. No methodology was suggested as to how to choose α and β and how to achieve the optimal design at least cost.

Using a similar procedure to that of incorporating chance constraints in the formulation of optimal water network design, Goulter and Bouchart (1987) have developed an approach based upon both mechanical reliability (component failure) and flow probability based on reliability issues (chance-constraints). In this combined approach, chance constraints on pipe breakage were formulated to restrict the probability that the number of breaks in a given link would be greater than some predetermined value. No numerical example was given. However, the authors noticed in the summary of their article that the improvement of the system reliability is further complicated after the first trial of the method. In effect, this difficulty is linked to the question of what steps should be taken to improve, at least cost, the reliability of the junctions which have been detected as possessing an unsatisfactory level of reliability. In addition, one of the authors (Goulter, 1987) pointed out that the problem of how to combine breakage and demand accedence probabilities into the same criterion for measuring reliability is, as yet, some way off.

Su et al. (1987) have developed a combined approach for water distribution network optimisation under reliability considerations. This approach combined three models. The first one is a non-linear programming model using the generalized reduced-gradient model by Lasdon et al. (1984). The second model is a network solver developed by Wood (KYPIPE, 1980) based on the linear method for the analysis of hydraulic systems. The last model incorporates the reliability analysis for determining the nodal and system reliability using the minimum cut set method. The three models are linked in the following way: For each iteration of the optimisation model, the reliability model computes the values of the nodal and system reliabilities using the same approach as was proposed by Mays et al. (1986). Minimum cut sets are determined by closing a pipe or combination of pipes in the system before running the network simulator which gives both pressure heads at nodes and link flows. If the pressure head of the node under examination is not satisfied this pipe or combination of pipes is a minimum cut set of the system, as well as of the demand point whose pressure head is below the required bound. It should be noted that if any one of the pipes in this cut set is already a minimum cut set of the demand junction by definition, this cut set is not a minimum set of the system (Billinton and Allan, 1983). This procedure is repeated until all the combinations of pipes have been considered. The reliability model then calculates the value of system and nodal reliability and returns to the optimisation model.

This model is among the few which exist up to now that incorporate a reliability measure, but on the other hand it did not completely solve the problem. Walters and Knezevic (1989) pointed out in their discussion of the paper that the use of the cut set method, which is suitable for the calculation of the reliability of general systems subject to any one or combination of component failures, is not really justified in the evaluation of water systems. In fact, it is only under extreme conditions or situations that all pipes in a minimum cut set (of two or more pipes) will be in the failure state at any one time in real network. Since failure of pipes is usually repaired within one or two days, this will give a greater probability than that of considering all pipes being simultaneously out of use.

Jacobs and Goulter (1989) suggested the use of graph theoretic principles to measure reliability of water networks. For a given network, knowing the number of links and nodes, the authors found that the optimal reliable network is a 'regular' graph, ie, a network in which each node has an equal number of links incident to it. Such a requirement is with no doubt of limited practical application in water distribution systems since nodes in the central areas of a network are usually more connected than those located at the periphery, as was pointed out by one of the authors in another publication (Goulter, 1992).

Kessler et al. (1990) proposed the concept of an invulnerable network for designing reliable networks. A topologically invulnerable network is a system which is able to sustain a single failure by means of adopting two alternative paths to every node. Their approach, which is composed of three stages, deals with single-source systems. The model works as follows:

(1) Two disjointed paths are allocated between the source and every consumer, selected by the designer using graph theory and engineering judgement. First, the given network is numbered using Depth First Search (Even, 1979) and a numbering scheme based on the work of Even and Tarjan (1976). Second, another algorithm is used to find two distinct spanning trees both of which are rooted in the source (Itai and Rodeh, 1984).

(2) The network diameters are sized using the linear programming formulation of Labye for branched networks. The head constraints are included for the two paths from each node.

(3) The solution must be tested for a set of loading conditions since hydraulic consistency is not incorporated into the optimisation parts.

However, just an application of the first stage of the algorithm is presented. Moreover, as pointed out in the paper, there is no way to determine the best pair of trees prior to a full hydraulic evaluation of each pair. The proposed guidance was the evaluation of the shortest distance tree by means of the Dijkstra algorithm (Itai and Rodeh, 1984) for tree number 1. Similarly, tree number 2 may be evaluated on the same basis, except that the links already chosen by tree number 1 may be assigned to a zero cost. Such an assignment will force the second tree toward a maximum overlapping between the trees.

Fujiwara and Silva (1990) proposed a heuristic procedure incorporating two models. The first one uses the linear programming gradient model (Alperovits and Shamir, 1977) modified by Quindry et al. (1979) and Fujiwara et al.

(1987), to size the network and to determine link flows. The second is a reliability model where the reliability index is measured as the complement of the ratio of the expected minimum total shortfall in flow to total demands. The minimum shortfall in the system flow is computed using the maximum flow algorithm of the capacitated network (Bazaraa et al., 1990). An iterative approach is subsequently used to modify the given link flows in order to have a significant increase in system reliability. In effect, the improvement of system reliability is obtained via the increase of flow capacity along a longest path (Bazaraa et al., 1990) where a "length" for each link is assigned in such a way that the selected path gives a significant increase in reliability at a small increase in system cost. The new link flows are supplied back to the linear programming gradient model and new pipe sizes are obtained. This process continues until a satisfactory reliability value is met.

After illustrating the method through a small network, the authors underlined the fact that further refinements of their work are required since flow capacity defined in the maximum flow model does not give a clear physical meaning and the system reliability estimated does not take into account hydraulic consistency.

Loganathan et al. (1990) used also a two-phase heuristic method to minimise network costs including reliability issues. In the first phase, the authors used an iterative procedure based on LP programming techniques to find a tree system, whilst, in the second phase, a heuristic algorithm is suggested to ensure that any node within the system is supplied by at least two paths from the sources. The redundant loop-forming links are designed to have pipe sizes as small as possible and if it is necessary, the tree found in the first phase is modified to force a feasible solution. This method is like two others previously published in the literature (Bhave, 1978 and Rowell and Barns, 1981) which defined reliability in terms of forcing the network to have loops even with minimum diameters for the redundant links. However, It was reported that minimum pipe sizes are more prone to failure than larger diameters (Robert and Regan, 1974). Moreover, the solution obtained with this model is not really a reliable network since for instance, in the event of mechanical failure of larger pipes, the system will suffer from inadequate pressures.

Another line of development for the reliability of water distribution systems is Entropy. Templeman (1989) was the first to use the principal of Maximum Entropy for the system to assign most likely flows to alternative paths through a looped network.

Awumah et al. (1991) proposed a method that minimises the cost of a network while imposing, in addition to the usual hydraulic constraints, a set of nodal reliabilities based on the concept of entropy. The entropy constraints are obtained by enforcing a minimum permissible redundancy at each junction. Awumah et al. (1991) used the Quindry et al. (1981) formulation since flows were used as decision variables and are also the variables of the redundancy. However, minimum bounds on diameters were not stipulated, to allow deletion of links if cheaper to do so while still maintaining the desired level of redundancy. The method was applied to a previously published example of Morgan and Goulter (1985) and it was found that the result obtained was close to that obtained by Morgan and Goulter (1985) which used an intensive iterative approach that incorporates 37 different loads.

In their second paper, Awumah and Goulter (1992) presented an alternative approach to that proposed above. In this approach, the objective function is to maximise the overall network reliability measured in terms of the concept of entropy. The constraints are the necessary hydraulic constraints (head loss, continuity and loop energy equations for single and multiple loads), minimum level of nodal pressures (for single and multiple loads) and a constraint on the network cost (budget constraint). The overall procedure works as follows. First, the model is run without the budget constraint to obtain a more reliable network corresponding to a local maximum network cost. Second, the budget constraint is incorporated into the model and reduced (in value) regularly in the next runs. The authors made use of the available graph theoretic parameter (Nodal Pair Reliability; Kim et al., 1972) to compare the change in the overall system entropy with the mean system reliability calculated on the basis of a nodal pair reliability index. As a consequence, it was found, for the example given, that a remarkable similarity can be seen between the shapes of two curves: The first one refers to the network cost saving versus reliability and the second one relates cost saving to system entropy. However Tanyimboh and Templeman (1993b) pointed out that if this holds for water distribution networks generally, it could be interpreted as evidence of a close relationship between entropy and mechanical reliability. One difficulty that remains with the entropy approach, pointed out by one of the authors (Goulter, 1992) is what a particular value of entropy measure means in absolute reliability or redundancy terms.

More recently, Tanyimboh and Templeman (1993c) presented a method for designing flexible water distribution networks. Flexibility was achieved through maximisation of the entropy of link flows. The method is based on a non-linear minimisation problem in which the system cost is optimised subject to continuity equations, energy equations and bounds (minimum and maximum) on pressure heads, flow velocities, diameters and non-negativity of flows. Reliability of the system was measured in terms of entropy. This model differs from the previous ones in terms of incorporating nodal entropic constraints into the non-linear program, which necessitates the non-negativity of flows due to the function logarithm used in entropy. Illustration of the method was performed using the Alperovits and Shamir (1977) example. It was found however that the concept of looping is sensitive to the value of entropy used: When the entropy is zero the system will degenerate into a tree shape. As the value of entropy increases, the network becomes less and less implicitly branched. Therefore, to keep loops, the authors suggested an entropy value considerably higher than zero. Obviously, looped systems are more reliable and more flexible than branched systems and the addition of entropy constraints in the optimisation scheme can reduce the tendency towards tree networks. However, the relationship between entropy and reliability has yet to be established.

3.4.1 Conclusion

Basically reliability can be modelled in different ways. These include:

(a) Provision of loops;

(b) Provision of uncertainty in design demand and pressure heads (chance constraints);

(c) Graph theory: cut sets, regular graph, invulnerable network;

(d) Heuristic approaches: path failure, node isolation, simultaneous failure of all links supplying a node, the complement of ratio of the expected minimum total shortfall in flow to total demand;

(e) Entropy: redundancy, flexible networks.

In short, all the methods provide useful insights into network design but all have their disadvantages. None of the models incorporating the listed measures of reliability gives a complete answer to the problem. The most promising index seems to be related to the randomness of demands, the randomness of pipe bursts and those using the entropy concept.

3.5 SUMMARY

To date, the optimisation of water distribution networks remains largely unsolved due to many limitations such as the non-linearity of the governing equations, the size of the networks and the complexity of the systems containing miscellaneous components. Yates, Templeman and Boffey (1984) concluded that discrete pipe size optimisation for distribution networks is NP-hard. A common characteristic of all the published models is that they guarantee only local minima, with sub-branched configurations.

With respect to reliability, all past work shows that uncertainties in the full definition and quantification of reliability still remain. The best indices seem to be the analytical methods, where the reliability is defined on the basis of the probability of providing the flows demanded at the required pressure heads rather than being simply determined upon connectivity or a similar index, but they are computationally intensive. The stochastic simulation is also a good indicator but the problem of its incorporation into optimisation is not yet solved. Entropy seems to provide good insights in this area. However, more work is needed to establish the relationship between entropy and water distribution reliability.

The literature review shows also that for models which include reliability specifications, no universally acceptable procedure for their definition is yet available.

Chapter IV

RELIABILITY BASED OPTIMAL DESIGN

Chapter IV

RELIABILITY BASED OPTIMAL DESIGN

4.1 INTRODUCTION

The previous work on optimisation of water networks summarised in chapter three has clearly shown that the published methods for optimally designing reliable networks do not give a complete answer to the problem.

The causes of the shortcomings of the past studies are various and mainly related to the complexity of the analysis of water distribution systems, the optimisation techniques used and more specifically the absence of both a standard definition of reliability and a well-defined conceptual framework for the overall approach to the problem of water system reliability evaluation.

The complexity of water networks is due to the non-linear nature of the hydraulic equations involved in the analysis in addition to the various hydraulic components that a normal system comprises.

Most of the optimisation procedures proposed so far are essentially gradient search techniques. Such algorithms can only guarantee local minima associated with tree-like shapes when the consideration of reliability issues is not taken into account. This is obvious, since it is always cheaper to convey a fixed quantity of water by one pipe rather than by two or more.

When non-linear optimisation techniques are used, the usual formulation of the problem is to consider pipe sizes as continuous decision variables which are rounded up/down to commercially available diameters, once the algorithms meet convergence criteria. The transition from a continuous solution to a discrete solution was reported to pose some difficulties linked to the feasibility and the globality of the discrete solution (Gessler and Walski, 1985).

Ideally, a water system should allow all consumers to draw the desired quantity and quality of water at the desired time at an adequate pressure. However this objective cannot be achieved throughout the entire life of a given network.

At this time (1994), there is no universal definition of reliability for water networks. As outlined in the previous chapter, the definition of reliability of a non-repairable manufactured item or component in a system according to British Standards (B.S. 3811, 1974) is "The ability of an item to perform a required function under stated conditions for a stated period of time. This may be expressed as a probability". When talking about water distribution networks, the definition of reliability quoted is inadequate for two reasons. First, because faults and failures do not cause complete breakdown of the system. Second, since system components are repairable items, the system's life is not determined by failure of an individual component. Therefore the concept of **Availability** has greater relevance to the evaluation of water distribution systems than pure reliability concepts.

Most aspects of water network reliability have been principally reported in the literature, but have not been simultaneously taken into account in the assessment of reliability. These include connectivity, link capacity, the concept of repair time, the stochastic nature of the demands and the randomness of pipe breaks.

The connectivity index, which refers to the probability that each node is connected to a source, has been reported as useful for identifying systems and/or nodes with serious problems due to insufficient redundancy. However, connection of nodes to a supply source is only a necessary, but not at all sufficient condition since the capacity of links is involved.

The concept of repair time, which is considered as the length of time required for a failed pipe to regain its operational state, is of fundamental importance when discussing reliability aspects of water network design as was stressed by Walters and Knezevic (1988). For instance, if, following a pipe rupture, it takes on average two days to return a pipe into service, then an annual breakage rate of 0.1 per km means that on average a 1 km long pipeline will be out of service for 0.2 days per year and will consequently be available for use 99.945 % of the time.

Pipe and system component breakages and hydraulic failures due, for example, to exceptionally high demands or inadequate pipe sizes are common causes of unreliability in water networks. These failures may cause a reduction in availability which could violate the supplier's obligation to its individual customers. They may also imply loss of revenue through a shortfall in water supplied as well as an increase in repair costs. A more complete model would be one that includes both the probabilistic nature of pipe bursts and the demands. Monte Carlo techniques can fill this gap. However, their heavy computational requirement limits their practical application especially within an optimisation scheme for looped systems.

4.2 POSSIBLE APPROACHES CONSIDERED

The above discussions further justify the need for the development of new methods for optimising water distribution networks that incorporate a sensible definition of reliability and, that may be used to yield systems which are both economical in total cost and meet reliability specifications.

More specifically, the objectives are:

(i) For the **reliability** aspect:

(1) To develop a new index for the definition of reliability of water networks;

(2) To develop a method that can be used for rapid assessment within an optimisation procedure;

(3) To incorporate the concept of repair time;

(4) To incorporate the probabilistic nature of both the demands and pipe bursts;

(5) To evaluate both nodal and system reliabilities.

(ii) For the **optimisation** aspect:

(1) To consider explicitly the discrete nature of the pipe sizes;

(2) To apply a new technique to the optimisation of water networks that is not influenced by the linearity or non-linearity of the objective function and the constraints involved, and that can find a near global optimal solution for the reliability based optimal design.

These objectives can be achieved by the development and testing of the following possible methods:

(a) **Flow Assignment plus Linear Programming Optimisation**

In this method the required reliability is imposed on the network by assigning flow capacity to the network links in a manner based upon reliability criteria. Then, the Linear Programming technique will be performed in order to find the optimal network cost and link diameters.

(b) **Genetic Algorithms with Reliability Tester**

This method applies a Genetic Algorithm that is a global search technique to find the least cost design of water networks subject to the technical constraints and reliability specifications. The Reliability Tester which is incorporated into the Genetic Algorithm, allows a rapid assessment of both nodal and system reliabilities.

These two methods are developed for use on PCs, implemented and tested in the following Chapters.

Attention is limited to the optimal design of single source networks. Excluded are the aspects of network layouts and the optimisation of the water distribution system components such as pump capacities, and heights and volumes of of water towers etc.

Chapter V

FLOW ASSIGNMENT MODULE

Chapter V

FLOW ASSIGNMENT MODULE

5.1 INTRODUCTION

In this method, the required reliability is imposed on the network by assigning flow capacity to the network links in a fashion based on reliability criteria. The determination of 'reliable' flow may be achieved through using the **Entropy Principle**. In other words, before optimising water networks, distribution of link flows will be obtained via the technique of maximum entropy. It should be noted that recently, entropy has been used as a new measure for assessing the reliability of water networks. Since the redundancy in a network is related to the reliability of that network, works by Awumah et al. (1991) and Awumah and Goulter (1992) are related to optimising network layout on the basis of entropic redundancy as reviewed in Chapter three. The method used for computing link flows on an entropy basis is that proposed by Tanyimboh and Templeman (1993). The historical development of the notion of entropy will be reviewed and its character as a means of measuring reliability will be discussed. Then the work of Tanyimboh and Templeman (1993) will be reviewed and applied to a small water distribution network.

Although first associated with classical thermodynamics, entropy has become known as an important and powerful concept in a variety of fields (Kapur, 1983). The notion of entropy has a long history in statistical mechanics. It is a central idea in information theory (Shannon, 1948). Its maximisation has been used as a model building principle: Kullback (1959) has built a theory for statistical inference. Jaynes (1957) has studied physical systems and Wilson (1970) has derived urban and regional models.

Entropy has been used as a measure of concentration, decentralization or variation (Erlander, 1977, 1980; Theil, 1967; Tribus, 1969), and also in probability concepts in hydraulics (Chiu, 1987, 1988).

In the optimisation area, entropy has also been used in connection with linear programming as a measure of accessibility and efficiency in the solution of the distribution problem in transportation planning (Erlander, 1977) and, in non-linear constrained optimisation it was used as a surrogate solution technique (Templeman and Xingsi, 1987).

In recent years, entropy has been applied as a measure of flexibility within manufacturing systems (Yao, 1985 and Kumar, 1987). In these works, flexibility within manufacturing systems refers to the ability of the whole produced system to overcome, without significant change in production capacity, the failure of one of the units in the system.

In 1989, Templeman published a paper about entropy and optimisation in civil engineering. He had shown how the entropy principle can help in assessing the most likely link flows in a looped pipe network from incomplete data.

Originally, the classical thermodynamic entropy as defined by Clausius was concerned only with macroscopic states of matter such as temperature, pressure, volume, etc. Whilst later on, Boltzmann examined microscopic states for thermodynamic systems. For a given system having N subsystems, if p_i refers to the probability of occurrence of subsystem i, Boltzmann has defined the entropy of the system as:

$$S_B = -K_B \sum_{i=1}^{N} p_i \, Ln(p_i) \qquad (5.1)$$

Where

S_B = entropy;

K_B = Boltzmann (positive) constant.

In his turn, Shannon (1948) has provided an essential advance in the use of entropy in new areas, for quantifying the disorder of systems other than thermodynamic ones. He has shown that, as in thermodynamic systems, entropy can also measure the amount of uncertainty in any probabilistic distribution. Shannon's entropy is the same as Boltzmann's entropy:

$$S_S = - K_S \sum_{i=1}^{N} p_i \, Ln(p_i) \qquad (5.2)$$

Where

S_S = Shannon's entropy;
p_i = the probability of occurrence of event i;
K_S = arbitrary positive constant.

Shannon's function S_S has several properties:

(1) S_S is a continuous and symmetric concave function for any fixed N with respect to all its arguments;

(2) S_S has a global maximum Ln(N) in case of equi-probabilities (p_i = 1/N), which increases in magnitude with the increase of the number of subsystems.

(3) S_S should be zero if one of the N probabilities is equal to one;

(4) S_S possesses the normality condition (It is axiomatic that all p_i are collectively exhaustive and mutually exclusive):

$$\sum_{i=1}^{N} p_i = 1 \qquad (5.3)$$

An interesting contribution of Jaynes (1957) is the use of Shannon's entropy in a reverse sense, ie, he has extended Shannon's entropy to generate a probability distribution which would have maximum entropy and which must contain minimum bias. He has shown theoretically how to avoid the introduction of bias in solving problems with incomplete information. This method is known as the **Maximum Entropy Principle.**

Suppose a random process can be described by a discrete random variable y which may take various discrete values of y_i, i = 1,...,N. If p_i refers to the probability that y has the value y_i, maximisation of Shannon's entropy subject to Eq. 5.3 only is a simple calculation which yields uniform probabilities, p_i, and an entropy ratio:

$$S_S/K_S = \mathrm{Ln}(N) \tag{5.4}$$

This result is derived using no information in addition to the axiomatic normality condition. If one assumes that some information about this stochastic process is available in terms of M expectation functions (see Templeman, 1989) of the form:

$$\sum_{i=1}^{N} p_i\, f_{ji}(y) = E[f_j] \qquad j = 1,\ldots,M \tag{5.5}$$

Where, $f_{ji}(y)$ and $E[f_j]$ are known. It is assumed that $M < N - 1$. If not, Eqs 5.3 and 5.5 suffice to determine uniquely the unknown probabilities. These probabilities p_i can be determined by solving the following constrained non-linear program:

Problem PO

Maximise $\quad S_S/K_S = -\sum_{i=1}^{N} p_i\, \mathrm{Ln}(p_i)$

Subject to:

1) $\sum_{i=1}^{N} p_i = 1$

2) $\sum_{i=1}^{N} p_i\, f_{ji}(y) = E[f_j] \qquad j = 1,\ldots,M$

In problem P0, it is axiomatic that p_i is not less than zero, $i = 1,...,N$. Templeman and Xingsi (1985) showed that the above optimisation problem P0 can be transformed into its dual form and then easily solved using standard techniques for unconstrained NL programming.

The use of the maximum entropy formalism is a very important idea in the sense that when designing water distribution networks, the available information is, in general, the source of supply and the nodal demands. Therefore the maximum entropy principle can be used for assigning least biased values to flows. Knowing the set of the most likely link flows, any optimisation technique which requires an initial estimate of flows can be applied to solve for the optimal system cost and diameters.

Another point which is more important than the determination of the least biased flows is that the reliability of a water network can be expressed in terms of entropy. Entropy may be viewed as a measure of the spread of the distribution of water molecules within the pipes of the system and the accessibility of water molecules to all the demand nodes. The higher the value of the entropy S_S, the more even the distribution. High entropy means that the system has maximum disorder, ie, the maximum choice for water molecules to move to any node via any pipe. However, in water systems, due to the restrictions expressed in terms of continuity equations water molecules have less freedom in choosing their routes. So a reduction in the value of S_S occurs. On the other hand, low entropy may be interpreted by low accessibility of flow elements to the system nodes: there will be a concentration of flow into some pipes as in branched systems.

We have introduced enough concepts to be able to formulate an optimisation problem that permits the design of flexible and "reliable" networks based on Jaynes's maximum entropy formalism.

5.2 REVIEW OF THE WORK OF TANYIMBOH AND TEMPLEMAN

Consider a general network with N_J nodes, N_P pipes, N_L loops and N_F fixed head nodes. The available data are the demands Q_i^{ext}, $i = 1,...,N_J$, the N_F

sources and all flow directions specified. Tanyimboh and Templeman (1993b) defined system entropy in terms of the sum of the sources' entropy and nodal entropies as:

$$S/K = S_0 + S_{NJ} \tag{5.6}$$

Where

S = Shannon's entropy;

K = positive constant;

S_0 = entropy of sources;

S_{NJ} = nodal entropies.

5.2.1 Sources' entropy

In the above equation (5.6), S_0 refers to the entropy of the external inflows defined as:

$$S_0 = -\sum_{j \in N_F} p_{0j} \, Ln(p_{0j}) \tag{5.7}$$

p_{0j} is the proportion of the total supply to the network that is provided by source j. Its value is given by:

$$p_{0j} = q_{0j} / \sum_{j \in N_F} q_{0j} = q_{0j}/T_0, \qquad \forall j \in N_F \tag{5.8}$$

Where

q_{0j} = external inflow at node j;

T_0 = total supply or demand.

5.2.2 Nodal entropies

Nodal entropies, S_{NJ}, was defined by the following equation:

$$S_{NJ} = -\sum_{i=1}^{N_J} P_i\, S_i \tag{5.9}$$

Where

P_i = probability of flow arriving at node i, i = 1,....,N_J;

S_i = entropy of node i.

For any node belonging to N_J, the resulting entropy S_i for the outflows including any demand is:

$$S_i = -\sum_{ik \in Nd(i)} p_{ik}\, Ln(p_{ik}), \qquad \forall\, i \tag{5.10}$$

The set Nd(i), i = 1,....,N_J, consists of all the outflows, including any demand, from node i, i = 1,...., N_J. In Eq. 5.10, the probability p_{ik} which refers to the fraction of T_i carried by link i–k, where T_i, i = 1,....,N_J, is the total outflow (including any demand from node i), is given by:

$$p_{ik} = q_{ik} / \sum_{ik \in Nd(i)} q_{ik} = q_{ik}/T_i, \qquad \forall\, i, \forall\, ik \in Nd(i) \tag{5.11}$$

The symbol q_{ik} is used for both internal and external inflows and outflows. For an external inflow, the first subscript will be zero and the second, the source node number. Also, the second subscript for a demand will be zero while the first will be the number of the node where the demand occurs. Otherwise, q_{ik} is the pipe flow from node i to node k. Tanyimboh and Templeman (1993b) reported that the probability P_i, i = 1,....,N_J was found by adding the probability of flow arriving at the node by each path. They stated that a convenient formula for P_i is given by:

$$P_i = T_i/T_0, \qquad \forall\, i \tag{5.12}$$

Having S_0, S_i and P_i for any node i of the network, the entropy of the overall system can be obtained. *Maximisation of the system entropy S/K* for a general network having more than one source, subject to the continuity equations, constitutes a constrained non-linear problem that can be solved by any available and efficient non-linear programming code. Detail of the derivation of all the equations involved and an illustration of the model can be found in Tanyimboh and Templeman (1993b). However, for a single-source network a non-optimisation approach based on the fact that all the paths supplying a node should distribute equally the required demand was presented in Tanyimboh and Templeman (1993b). This accords with two principles: Laplace's principle of insufficient reason and Jayne's maximum entropy formalism (Jaynes, 1957). Once the distribution of flow within a network is known, the LP technique can be applied for the least cost design. This will be dealt with in Chapter 8.

5.2.3 Summary

It should be noted that, for the formulation proposed by Tanyimboh and Templeman:

(1) The derivation of the final entropy of a system is preferred to and more justified than that suggested by Awumah and Goulter (1991) since, as stated by Khinchin (1953), the entropy of a system cannot be a simple sum of the entropy at each node;

(2) S_0 is equal to zero for a single source network;

(3) S_{NJ} may be interpreted as the sum of nodal entropies weighted by their total outflows, including nodal demands (see Eqs. 5.9 and 5.12).

5.3 APPLICATION OF THE MODEL OF TANYIMBOH AND TEMPLEMAN

It is clear from the published works that entropy may be used as an index for the assessment of water network reliability. The method used herein for the allocation of flows to a water network is that proposed by Tanyimboh and Templeman (1993) for single source networks.

For the purpose of illustration, a single source network (Network A) that is shown by Fig. 5.1 is selected. The topological data are:

$N_F = 1;$

$N_J = 6;$

$N_P = 8;$

$N_L = 2.$

External demands are given in Table 5.1.

Node	Demand (l/s)
source	- 80
1	00
2	10
3	10
4	30
5	20
6	10

Table 5.1 Nodal demands of Network A

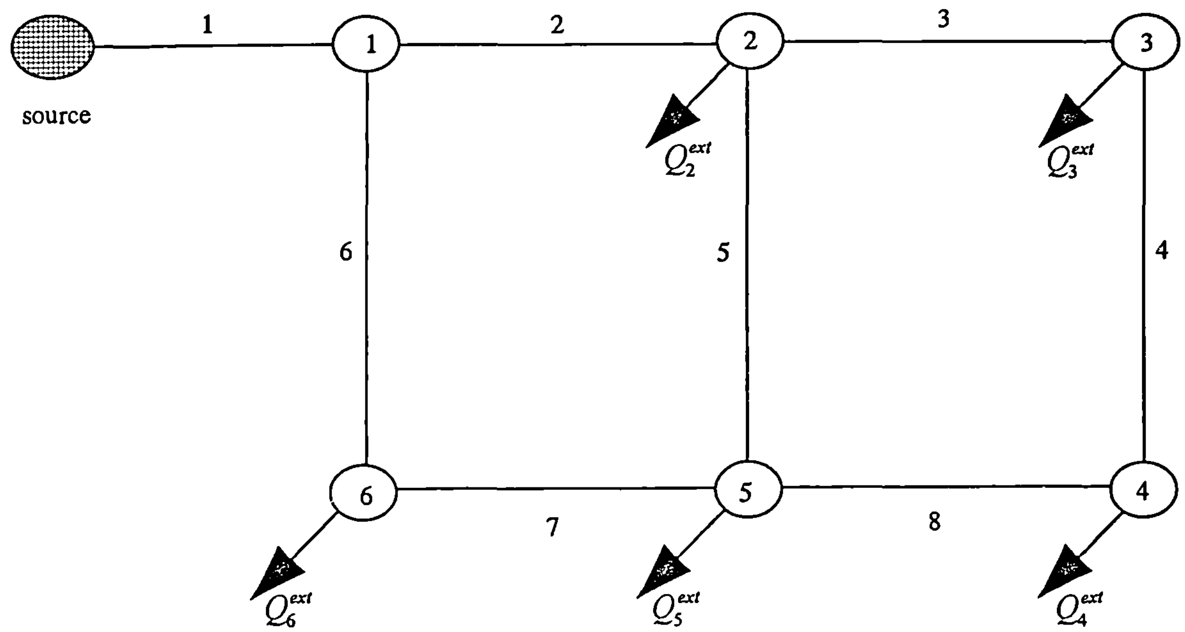


Figure 5.1 Network A

The available information consists of the external demands and the flow directions. These are essential for identifying nodal paths. The method used is based on the identification of nodal paths. Then the demands are supplied equally to nodes by all their paths. Paths and link flows for network A were determined as stated above and are summarised in Tables 5.2 and 5.3 respectively.

Node	Demand (l/s)	No of Paths	Path Links	Flow/Path (l/s)
1	00	1	1	00
2	10	1	1 2	10
3	10	1	1 2 3	10
4	30	3	1 2 3 4 1 2 5 8 1 6 7 8	10
5	20	2	1 2 5 1 6 7	10
6	10	1	1 6	10

Table 5.2 Paths for Network A

Link	Flow by Tanyimboh & Templeman Model (l/s)
1	80
2	50
3	20
4	10
5	20
6	30
7	20
8	20

Table 5.3 Link Flows for Network A

5.4 SOFTWARE

For small network sizes, like network A, the approach can be applied manually. However for somewhat larger and more general networks, an efficient algorithm is required.

PATH_Q was developed to implement the Tanyimboh and Templeman approach for single-source networks. The algorithm starts by identifying all paths serving any node (path enumeration) and then computes the flow distribution of the system ie, each node within the network is taken in turn and its demand is divided equally amongst all its supplying paths. Nodes that are supplied by only one path are first processed. However, for a node that is served by more than one path, it can be shown that its total number of paths is the sum of paths to all nodes upstream of, and directly supplying this node. **PATH_Q** makes use of two global types of variables: $Check_i$, $i = 1,....,N_J$ and Scheck. $Check_i$, $i = 1,....,N_J$, which initially takes the value of zero is a control variable for any node that indicates if the node is processed ($Check_i$ = 1) or not ($Check_i$ = 0). Scheck on the other hand is incremented by one each time a node has been processed, and is used as a stopping criterion for the algorithm. ie, when Scheck becomes equal to the total number of system junctions (N_J), the algorithm stops, indicating that all nodes have been processed, and then, all link flows are known. Basically, the algorithm is as follows:

Program **Path_Q**

```
Start_Path
Scheck = 0
Check_i = 0
First_Search
Repeat
      Second_Search
Until  Scheck  =  N_J
Result_Path
```

Start_Path refers to the routine that allows the entry of the required topological data for the network (N_F, N_J, N_P and the connecting links to any node) and the external demands. The global variables Scheck, and $Check_i$ are initialised to zero before the execution of the procedure **First_Search.**

The First_Search procedure identifies and determines nodes which are supplied by one path only. Each of these nodes is processed such that its demand is supplied by its path. As mentioned above, for the nodes processed, $Check_i = 1$ (i refers to a processed node) and Scheck will be equal to the number of processed nodes.

The main part of the algorithm contains a repeat-until loop that incorporates the **Second_Search** procedure which is repeated as after as it is required until the convergence criterion is met. Second_Search deals with any node that is served by more than one path. When the algorithm reaches its end, the link flows will be output by the procedure **Result_Path.**

Other algorithms based on supplying equally each node by its paths (the path model) are discussed in Tanyimboh and Templeman (1993b).

Running Path_Q with Network A gives exactly the result provided by Table 5.3. Before closing this Chapter it is worth noting that it would be interesting to make a comparison between a design produced on the basis of the entropy principle and an arbitrary distribution of flow. This comparison should consider the reliability aspects. That is, one should:

(1) Select samples of networks with different sizes and topologies;

(2) Give an arbitrary distribution of flow to a sample of networks and apply Path_Q to the same network for the distributions of flow;

(3) Apply an optimisation technique to the resulting distributions of flow of the same network, for the optimal cost;

(4) Assess the reliability of the optimal designs corresponding to the two distributions of flow for the same network.

These points will be dealt with in the three following Chapters. Chapter 6 introduces the application of the Linear Programming technique to the design of water networks while the reliability aspect is discussed in depth in Chapter 7. Examples of the measure of reliability in terms of entropy will be discussed in Chapter 8. The final part of the comparison mentioned above is performed in Chapter 8.

Chapter VI

LINEAR PROGRAMMING OPTIMISATION MODULE

Chapter VI

LINEAR PROGRAMMING OPTIMISATION MODULE

6.1 INTRODUCTION

This chapter is related to the previous chapter which allows the determination of system link flows on the basis of the maximisation of Shannon's entropy. Having obtained the "reliable" link flow distribution within the network, the optimisation of the total cost of the network can be performed. At first sight, the selection of pipe sizes seems to be very non-linear. However, this problem can easily be transformed into a linear form. The clue is, instead of choosing pipe diameters as decision variables, consider a link as being made up of a set of "candidate diameters", and the decision variables being the lengths of segments of constant diameter within the link as first described by Labye (1966). In such a case one can take advantage of the very powerful and easily accessible **Linear Programming** algorithm and make use of it to find the optimal network cost and the pipe diameters. Moreover, the advantage of this LP modelling is the direct use of discrete pipe diameters that are commercially available. In the rest of this chapter, the formulation and the solution of the optimisation of water networks using the LP technique is discussed. A computer program which has been developed for this purpose is also presented and tested.

6.2 STATEMENT OF THE PROBLEM

The optimisation of a simple hydraulic distribution network can be stated as follows:

GIVEN:

1) A set of demand nodes;

2) A set of links;

3) A set of normal loading (demand) conditions;

4) A fixed link flow pattern;

5) A set of commercially available pipe diameters and costs;

6) A set of minimum performance levels for normal loading conditions.

FIND:

1) Link diameters;

2) The minimum total cost of the network.

SUBJECT TO:

1) Satisfying steady state flow conditions,

2) Satisfying minimum performance levels under normal loading conditions.

The mathematical relationships between head loss, diameter and length of pipe are easily formulated and are very amenable for use in LP models.

Indeed, consider the water system (Network C) given in Fig. 6.1. Network C is taken from Coals and Goulter (1985) where a single link of 1000m (link number 1) has been added between the first node and the source for distinguishing between the source and the nodes. For this network, the topological data are: $N_F = 1$, $N_J = 9$, $N_P = 13$, and $N_L = 4$.

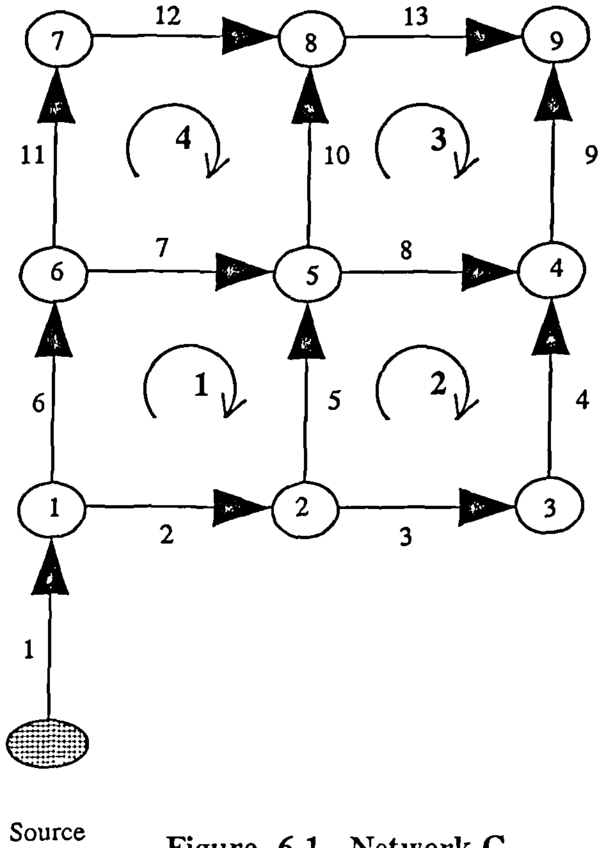


Figure 6.1 Network C

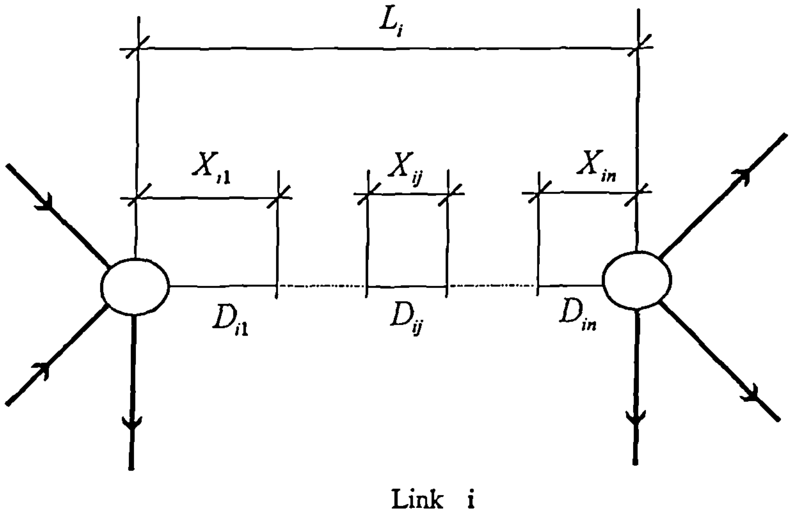


Figure 6.2 Representation of Segment X_{ij} in a Link of Length L_i

Assume that any link i, $i = 1,\ldots,N_P$ of network C may, in fact, be made up of one or more spans of pipes of different discrete diameters (D_{ij}, $i = 1,\ldots,N_P$, $j = 1,\ldots,n(i)$ where n(i) is the total number of candidate diameters of link i connected in series as shown in Fig. 6.2. When more than one pipe is used, the total length of pipes (X_{ij}, $i = 1,\ldots,N_P$, $j = 1,\ldots, n(i)$) employed should be equal to the length of the link (L_i):

$$\sum_{j=1}^{n(i)} X_{ij} = L_i \qquad \forall\ i \in N_P \tag{6.1}$$

Eq. 6.1 holds if the pipe diameters D_{ij}, $i = 1,\ldots,N_P$, $j = 1,\ldots, n(i)$ of the segment of length X_{ij} are only available in discrete commercial sizes.

For network C, if the flow Q_i along any link i, $i = 1,\ldots,N_P$, is known (evaluated either on the basis of an initial flow pattern satisfying continuity equations or determined from the application of the maximum entropy principle suggested in chapter 5), the hydraulic gradient J_{ij}, $i = 1,\ldots,N_P$, $j = 1,\ldots, n(i)$, can be computed in advance for each candidate diameter D_{ij} from, for instance, the Hazen-Williams equation as:

$$J_{ij} = \gamma\,(Q_i/C_{ij})^{1.852}\, D_{ij}^{-4.87} \tag{6.2}$$

Where γ is a constant depending on the unit used and C_{ij} is the Hazen-Williams coefficient of diameter j in link i.

Therefore, the head loss along link i ($i = 1,\ldots,N_P$) composed of n(i) segment X_{ij} of candidate diameter D_{ij} is given by:

$$\Delta H_i = \sum_{j=1}^{n(i)} J_{ij}\, X_{ij} \tag{6.3}$$

Head losses ΔH_i, $i = 1,...,N_P$, are required for evaluating the following constraints (Node, Path and Loop) applied to water networks:

(1) **Node:** nodal heads must be equal or greater than the required minimum bound:

$$\sum_{i \in Pn(k)} \Delta H_i <= H_0 - H_k^{min} \qquad \forall\ k \in N_J \tag{6.4}$$

Pn(k) is the set of links in the path from the source to node k; H_0 is the original head at the source and H_k^{min} is the minimum required head at node k.

(2) **Path:** the total head loss along any path between two fixed nodes must equal the difference in head between those nodes:

$$\sum_{i \in Pf(k)} \Delta H_i = b_k \qquad \forall\ k \in (N_F - 1) \tag{6.5}$$

Where Pf(k) is the set of links in path k, $k = 1,...,(N_F - 1)$ associated with known net head loss b_k.

(3) **Loop:** the total head loss around a loop must equal zero, ie, a loop is a particular path whose $b_k = 0$.

$$\sum_{i \in Pl(k)} \Delta H_i = 0 \qquad \forall\ k \in N_L \tag{6.6}$$

Pl(k) is the set of links in loop k, $k = 1,...,N_L$.

Finally, information on the cost $CostD_{ij}$ of each pipe of diameter D_{ij} ($i = 1,...,N_P$, $j = 1,..., n(i)$) is required. The total cost of the supply of pipes and fittings, trench excavation and laying, expressed as a cost per unit length, can however, include an allowance for future maintenance, repair and replacement.

The cost of a link Cl_i is then:

$$Cl_i = \sum_{j=1}^{n(i)} CostD_{ij} X_{ij} \qquad \forall\ i \in N_p \tag{6.7}$$

The total cost for the whole network is as follows:

$$Cost_{NET} = \sum_{i=l}^{N_P} Cl_i \tag{6.8}$$

Having defined the objective function to be minimised and the technical constraints, the design of network C can be expressed mathematically for a fixed flow pattern (single load) in terms of the LP formulation, as:

Problem L1

<u>OBJECTIVE FUNCTION:</u>

$$\text{Minimise } Cost_{NET} = \sum_{i=1}^{N_P} \sum_{j=1}^{n(i)} CostD_{ij} X_{ij} \tag{6.9}$$

<u>CONSTRAINTS:</u>

1) Length: length of segments in each link must be equal to the length of that link:

$$\sum_{j=1}^{n(i)} X_{ij} = L_i \qquad \forall\ i \in N_p \tag{6.10}$$

2) Node: minimum permissible head at each node must be satisfied.

$$\sum_{i \in Pn(k)} \sum_{j=1}^{n(i)} J_{ij} X_{ij} \leq H_0 - H_k^{min} \qquad \forall\ k \in N_J \tag{6.11}$$

3) Path: head loss around any path between two nodes must equal the difference in head between those two nodes.

$$\sum_{i \in Pf(k)} \sum_{j=1}^{n(i)} J_{ij} X_{ij} = b_k \qquad \forall\ k \in (N_F - 1) \tag{6.12}$$

4) Loop: the total head loss round a loop must equal zero.

$$\sum_{i \in Pl(k)} \sum_{j=1}^{n(i)} J_{ij} X_{ij} = 0 \qquad \forall\ k \in N_L \tag{6.13}$$

5) Non-negativity:

$$X_{ij} \geq 0 \tag{6.14}$$

Having set up both the objective function and the constraints, it becomes clear that the design problem of water systems can be formulated in linear terms and hence can be handled by any efficient Simplex-based linear package such as LINDO (Schrage, 1987) for the least cost of the network and the optimal pipe diameters. Furthermore, solution of problem L1 also gives the nodal heads for the system.

It should be noted that the formulation outlined deals only with a single loading condition and does not design water towers or pumps.

6.3 SELECTING LINK CANDIDATE DIAMETERS

To solve problem L1, (i) the network data (eg external demands, link lengths, source head and minimum nodal heads, unit costs of pipe sizes etc.) have to be specified, (ii) a distribution of link flows is required, and (iii) the candidate diameters per link and their number have to be determined. Having the distribution of link flows, the sets of candidate diameters for each link can be selected by applying one of two approaches:

(1) Selection on the basis of minimum and maximum velocities (eg 0.5 and 2.5 m/s).

(2) Selection on the basis of lowest and highest allowable hydraulic gradients (eg 0.0005 and 0.05 were used by Alperovits and Shamir, 1977);

In the first approach, the theoretical pipe sizes Dt_i (in m) for a given individual link can be found from the following equation (Orth, 1986):

$$\left(K_1 \cdot Q_{max} / V_{max}\right)^{1/2} \leq Dt_i \leq \left(K_1 \cdot Q_{min} / V_{min}\right)^{1/2} \qquad (6.15)$$

Where $K_1 = 4/\pi$; Q_{min} and Q_{max} are minimum and maximum design flows, m^3/s; and V_{min} and V_{max} are minimum and maximum design velocities, m/s. Minimum and maximum flows may be different if multiple loading conditions are considered. For a single load, however, the two values are equal.

With reference to the second approach, if the Hazen-Williams equation is used to represent the head loss, the theoretical pipe diameters for an individual link are given by:

$$(K_2/J_{max})^{0.21} \cdot (Q_{max}/C)^{0.38} \leq Dt_i \leq (K_2/J_{min})^{0.21} \cdot (Q_{min}/C)^{0.38} \qquad (6.16)$$

Where K_2 = 10.70 (for Dt_i, in m; and Q_{min}, Q_{max} in m^3 /s); C is the Hazen-Williams coefficient; J_{min} and J_{max} are the lowest and highest hydraulic gradients. As above, for a single load, minimum and maximum flows are equal.

The first approach was chosen for simplicity. Given bounds on velocities (V_{min} and V_{max}), link flows Q_i, i = 1,...,N_p, and the discrete commercially available pipe diameters, the set of candidate diameters and their number for each link can be obtained by running a simple procedure called ***SizeD***.

6.4 SOFTWARE

LP packages such as LINDO (Schrage, 1987) are expensive and their use for solving problem L1 is not a simple task. In effect, application of LINDO to solve problems like problem L1 requires that the objective function and all the constraints involved are written explicitly in terms of the decision variables before running the program. This is not a simple task, especially with larger networks where the numbers of both the decision variables and the constraints increase rapidly. Therefore, a specific computer program for the design of water networks by LP algorithm is desirable. **LNOPTNET**, is a linear optimisation model that was developed during this research to tackle programs like problem L1. However only single-source networks (Problem L2) can be solved. Problems L1 and L2 are similar except that equation 6.12 which deals with the head losses between two fixed nodes is not required for problem L2. LNOPTNET is a computer program that uses only the network data, solves directly the linear program and gives the least cost of the network and the optimal pipe sizes.

LNOPTNET, which runs on PCs, comprises eight procedures:

Program LNOPTNET

```
Start_LP
    Objective_Function_Coefficients
    RightHandSide_Coefficients
    Length_Coefficients
    Node_Coefficients
    Loop_Coefficients
    Simplex
Result_LP
```

Before examining any of the procedures belonging to the above algorithm, it should be noted that the routine used in LNOPTNET for solving problem L2 is the Press et al. (1988) routine published in 'Numerical Recipes' book. This routine (SIMPLX) was selected: (i) because it was recommended by several academic members of staff. (ii) Because it is available in the Pascal Language, which is the programming language used in this project. The SIMPLX routine was slightly modified to give as output only the required results (the objective function and the decision variable values).

The SIMPLX routine is a powerful one that requires (1) the identification of the type of constraints ('$\leq$' constraints, '$\geq$' constraints and '$=$' constraints) related to the problems considered and their total number, (2) the coefficients of the LP matrix.

With reference to problem L2, the '$\geq$' constraints are absent, and for the rest of the types of constraints, there are $N_P + N_L$ '$=$' constraints (N_P Length constraints and N_L Loop constraints) and N_J '$\leq$' constraints (Node constraints). Thus the total number of constraints is $N_J + N_P + N_L$.

The LP matrix coefficients are:

(1) The coefficients of the objective function (unit costs of diameters);

(2) The right-hand side coefficients of the constraints (total link lengths, zeros for the loop constraints and the differences between the head at the source and the nodal minimum heads);

(3) The left-hand side coefficients of the constraints. These are (i) the 0-1 coefficients corresponding to the Length constraints, and (ii) the known hydraulic gradients J_{ij} (Node and Loop constraints) for each D_{ij}, Q_i combination.

The procedure *Start_LP* is related to the input of the required network data (eg N_J, N_P, N_L, the source head H_0, bounds on nodal heads, link lengths, external demands, Hazen-Williams coefficients, number of candidate diameters per link n(i), costs of unit length of diameters, etc.).

Objective_Function_Coefficients procedure is called to prepare the coefficients of the objective function. These coefficients are the unit costs of the pipe diameters.

The *RightHandSide_Coefficients* routine allows the determination of the coefficients of the right-hand sides of all constraints. These are zeros for the loop constraints, the lengths for the length constraints and the head differences (H_0 - H_k^{min}, k = 1,...,N_J) between the original source head and the minimum heads of nodes.

The procedure *Length_Coefficients* prepares the 0-1 coefficients related to the N_P Length constraints.

Node_Coefficients and *Loop_Coefficients* sub-programs allow the preparation of the known coefficients, and relate to the N_J Node constraints and the N_L Loop constraints. These coefficients are either zeros for decision variables not involved in the constraints or the constant hydraulic gradients (J_{ij}) of the pipes involved in the node and loop restrictions.

Finally, the *Simplex* routine uses all the coefficients prepared by the above procedures to solve the linear problem while *Result1* organises the output of the solution with the least network cost and optimal pipe sizes.

6.5 EXAMPLE

The numerical example chosen to run with LNOPTNET is Network C displayed in Fig. 6.1 below. For consistency, the link and the node numbers used in the work of Coals and Goulter (1985) have been changed. Table 6.1 gives the correspondence between the two sets of numbers used. The Demand of nodes is given by Table 6.2. The length and the Hazen-Williams coefficient are equal to 1000m and 130 for all pipes respectively. The head at the source is assumed to be 50m and the minimum permissible pressure head is set to 25m. The set of candidate diameters with their Costs which are provided by Table 6.3 are those utilised by Coals and Goulter (1985).

In addition to the above data, the flow directions have to be specified and an initial flow distribution is required. These are take from Coals and Goulter (1985) and shown in Fig. 6.3.

Link Numbers used	Link Numbers of Coals and Goulter (1985)
Nodes	
1	1
2	2
3	3
4	6
5	5
6	4
7	7
8	8
9	9
Links	
1	-
2	1
3	2
4	5
5	4
6	3
7	6
8	7
9	10
10	9
11	8
12	11
13	12

Table 6.1 Correspondence between Links and Nodes used here and those used in Coals and Goulter (1985)

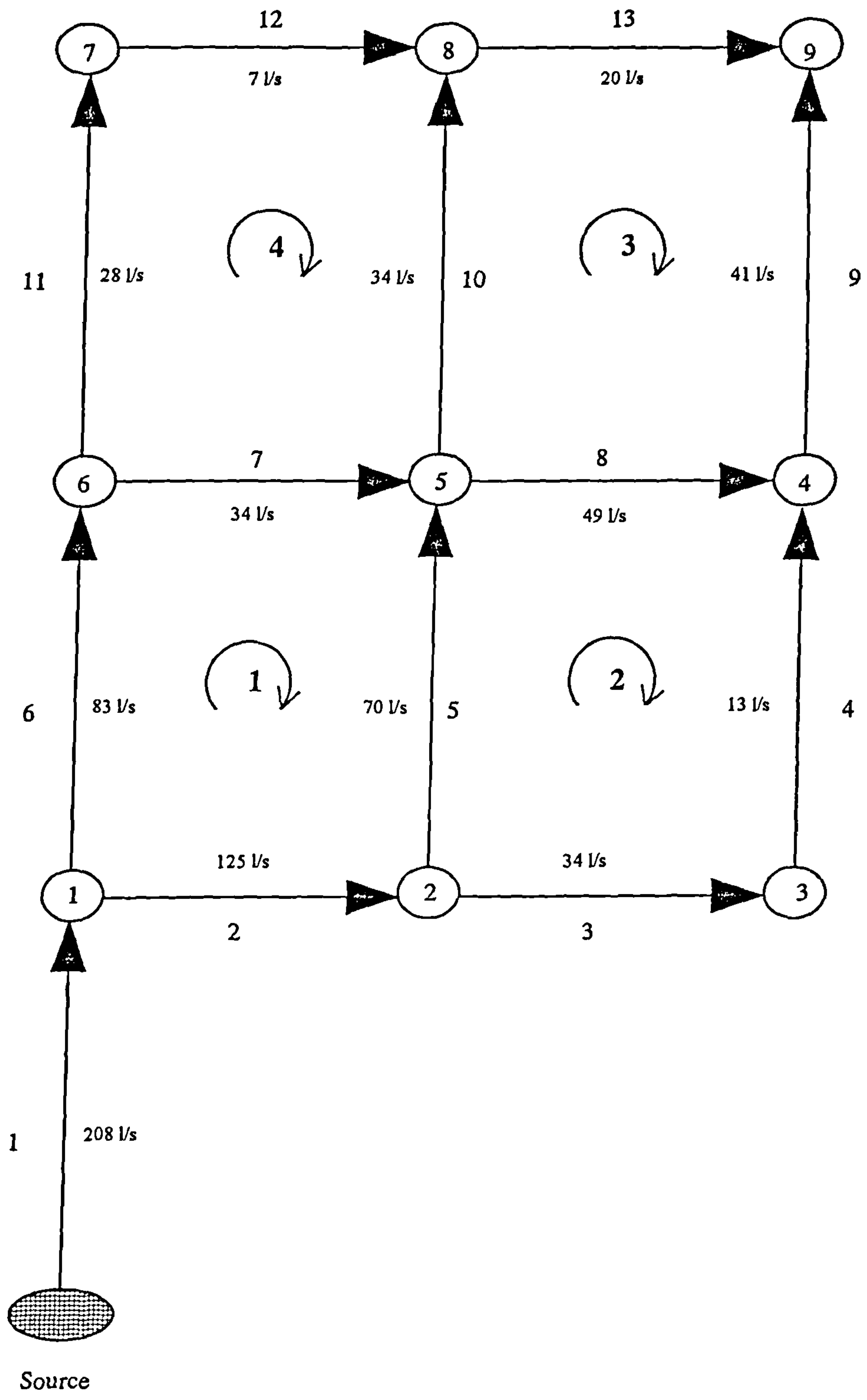


Figure 6.3 Assumed Flow Directions of Network C

For simplicity all the six pipe sizes in Table 6.3 are taken as candidate diameters for all links. However, before running LNOPTNET with network C, strings of pipes for pressure and loop constraints have to be specified. These are given in Table 6.4. Execution of LNOPTNET with Network C gives an optimal network cost of $639303. The optimal diameters with their corresponding length are summarised in Table 6.5.

Node	Demand (l/s)	
1	-208	(Source)
2	21	
3	21	
4	21	
5	21	
6	21	
7	21	
8	21	
9	61	

Table 6.2 Demand of Nodes in Network C

Diameter (mm)	Cost ($/m)
100	14.3
150	16.9
200	24.1
250	43.2
300	69.2
350	98.6

Table 6.3 Cost Data for Network C

Path	Number of Links in Path	Link Number				
		Node Equation				
1	1	1				
2	2	1	2			
3	3	1	2	3		
4	4	1	2	3	4	
5	3	1	2	5		
6	2	1	6			
7	3	1	6	11		
8	4	1	6	11	12	
9	5	1	6	11	12	13
		Loop Equation				
10	4	7	-5	-2	6	
11	4	8	-4	-3	5	
12	4	13	-9	-8	10	
13	4	12	-10	-7	11	

Table 6.4 Strings of Pipes for Pressure and Loop Constraints

Link	Length (m)	Diameter (mm)
1	1000	350
2	1000	350
3	1000	200
4	1000	200
5	1000	300
6	1000	300
7	297	200
	703	250
8	939	250
	61	300
9	13	250
	987	300
10	193	200
	807	250
11	40	150
	960	200
12	1000	150
13	1000	200

Table 6.5 Optimal Solution for Network C

6.6 COMMENT ON THE LP SOLUTION

Solution of problem L2 by LNOPTNET is characterised by:

(1) Most of the elements, X_{ij} are zero in the optimal solution. This is exemplified by the result summarised in Table 6.5. For instance, LNOPTNET has selected for pipe 1, a segment length, $X_{15} = 1000$m, $X_{1j} = 0$ for $j = 1,...,4$.

(2) When a single diameter for a link has not been retained from those available, this link will consist of at most two sections with different diameters. The sum of their individual lengths is equal to the length of this link. Links 7, 8, 9, 10 and 11 are examples.

(3) Links with two segments should be made up of two adjacent pipe sizes. (eg 200 and 250 mm in link 7) This adjacency property holds if and only if pipe costs are a strictly convex function of a power of pipe diameters. A full proof of this condition can be found in the paper of Fujiwara and Dey (1987).

(4) On the practical side, the result of the design by the application of LNOPTNET should not be seen as a finished work. This result may be adjusted where appropriate. Indeed, segments that are too small to be of practical significance in two-section links should be eliminated after the optimisation process to obtain a more satisfactory engineering solution. This may be achieved, for example, by substituting the smaller diameter by the larger one in the same link. In this event, the head loss in the network would be reduced resulting in a relative increase in both the minimum nodal pressures and the system cost. For example, if a segment in a link of less than, say 50 m (13m of 250 mm in link 9 and 40m of 150 mm in link 11), is considered too small in practice, it can be substituted by the larger pipe size in the same link (13m of 300 mm in link 9 and 40m of 200 mm in link 11). This modification is accompanied with an insignificant increase in the system cost (0.1 %) as mentioned above.

With reference to the head loss in network C, this alteration of pipes 9 and 11 has been accompanied with a relative increase in nodal pressures as expected. This may be demonstrated by running LMANLS with network C for the two sets of diameters. Results of these runs are shown in Table 6.6 where all nodal pressures have been increased except node 1, which is not concerned with the redistribution of flow and node 6.

Node	Pressure (m) Before Adjustment of Diameters	Pressure (m) After Adjustment of Diameters
1	38.20	38.20
2	33.61	33.63
3	27.32	27.36
4	26.26	26.31
5	30.28	30.32
6	33.64	33.62
7	28.72	29.05
8	27.36	27.46
9	25.00	25.08

Table 6.6 Nodal Pressures of Network C for the optimal and the adjusted solutions

Other suggestions for the adjustment of the LP solutions can be found in the work of Orth (1986). Later in this work (Chapter 8), the practical solution adopted is that being made only of one section per link.

(5) The number of candidate diameters per link is related to the specified limits on velocity. For a link, the minimum velocity constrained the maximum candidate diameter and the maximum velocity constrained the minimum size (see Eq. 6.15). This number of candidate diameters may be zero. For instance, suppose that for Network C the maximum velocity has been set to 1.5 m/s. From Eq. 6.15 (Q_{max} = 0.208 m^3/s and V_{max} = 1.5 m/s), the minimum size, Dt_1, corresponding to link 1, is greater than 400 mm which is not available in the set of the candidate diameters considered in this example.

(6) Finally, it should be noted that the Linear Programming technique applied to the optimisation of water networks mail fail to produce feasible solutions if the constraints are not satisfied. A non-feasible solution can be met with Network C if, for example, the head at the source is changed to 40m instead of 50m. Indeed for the problem of non-feasibility, Alperovits and Shamir (1977) have used different hydraulic gradients for different pipe flows (eg for Q = 33.33l/s, $0.0004 < \Delta h/L < 0.0228$; for Q = 180.56 l/s, $0.0025 < \Delta h/L < 0.043$). Moreover, this is the reason for the introduction of **Dummy Valves** in the LP formulation in the same work of Alperovits and Shamir.

Chapter VII

RELIABILITY TESTER

Chapter VII

RELIABILITY TESTER

7.1 INTRODUCTION

Reliability of drinking water distribution systems has recently become a focus of attention in the water industry. Its significance to network design was recognised when optimisation techniques were applied to looped distribution systems (Jacoby 1968 and Watanatada 1973). It was found that when looped networks were optimised under a single pattern of demands and on a cost basis, the cost optimisation process reduced the redundant lines down to zero and the optimal design lost the loops leaving only a branched configuration.

One of the reasons water engineers incorporate loops in a network rather than having trees, is that loops increase system security, flexibility and reliability. Cost optimisation tends to reduce system reliability, and hence constraints are needed within the optimisation process to retain a required reliability standard.

At present researchers do not concur on a universally accepted definition of reliability, even for analysing existing systems. There is also no well-established methodology for evaluating water system reliability. Since system components are repairable items, it was seen in Chapter 4 that the concept of availability is more appropriate to the performance of water distribution systems than pure reliability concepts. Consequently, reliability issues may be usefully quantified on the basis of the availability of adequate supply to customers rather than the connectivity of the customers to the supply point. Indeed a connection of any node to a source is a necessary but not sufficient condition since, in a hydraulic network, a node linked to a source may or may not receive any water, depending on the pressures in the system.

As stressed in Chapter 4 also, the other important factor in the assessment of water system reliability is the concept of Repair Time. A reasonable evaluation of the network reliability should incorporate this concept.

In Chapter 3, the important approaches published in the literature for the assessment of the reliability of water networks have been outlined. These include: (a) provision of loops, (b) uncertainty in demand and pressure heads (chance constraints), (c) graph theory (eg cut sets), (d) heuristic approaches (eg path failure, node isolation), and (e) entropy.

It was also mentioned in Chapter 3 that the use of the cut sets method for considering the reliability of networks subject to component failures is perhaps not appropriate for many reasons. Among these (1) it is only under unrealistically extreme conditions that all pipes in a minimum cut set (of two or more) would be in a failure state at any one time in a real network, since a failed pipe is usually repaired within one or two days. (2) Identification of all minimum cut sets requires an extensive computation. (3) the supply to a node could fail completely without being entirely isolated by broken pipes and finally, (4) basic cut set methods do not allow the variability in the demands to be taken into account (Walters and Cembrowicz, 1993).

More realistic measures of reliability can be offered by stochastic simulation techniques, but the major problem with these methods is linked to the very large number of network analyses required to evaluate availabilities. Wagner et al. (1988b) demonstrated the usefulness and the flexibility built into such techniques in a scheme incorporating the uncertain nature of failure events and repair times. However the stochastic nature of the demand was not incorporated.

Bao and Mays (1990) assessed hydraulic reliability on the basis of providing adequately the nodal supplies under specified pressure bounds where the demands were determined by a Monte-Carlo technique. Monte-Carlo techniques are robust and suitable for systems of any complexity; however, their heavy computational requirement limits their practical application especially within an optimisation scheme for looped systems.

More recently, entropy was used as a surrogate reliability measure. Awumah et al. (1991) used this principle as a means of incorporating redundancy in the optimisation of water distribution systems. Tanyimboh and Templeman (1993) suggested that flexible networks can be achieved through maximising the entropy of flows. The addition of such constraints in the non-linear optimisation problem of looped systems can reduce the risk of losing the loops. However the relationship between entropy and reliability has yet to be properly established.

7.2 SPECIFICATION FOR A NEW MODEL

The previous work and all the above mentioned approaches provide useful insights into how reliability can be incorporated into water network design, but all have their limitations.

A new method is sought with the following objectives:

1- To develop a method that can be used for rapid numerical calculation of reliability within an optimisation procedure. Such a method must avoid the need for the large number of network analyses required by Monte Carlo type models.

2- To incorporate simultaneously the probabilistic nature of both the demands and pipe failures.

3- To incorporate the concept of repair time.

4- To evaluate nodal and system availabilities on the basis of availability of supply at the demand nodes.

All these objectives will be taken into account in the following stochastic approach which the writer calls the "**Reliability Tester**". The description of the Reliability Tester given in this Chapter is reproduced in Khomsi et al. (1994). Now an appropriate measure of reliability must be specified before describing the approach.

7.2.1 Reliability Measure Adopted

For this work, reliability may best be expressed in terms of the probability that a given node receives its supply within specified limits or constraints. This reflects consumer expectations from a reliable system. In other words, as far as the consumer is concerned an availability can be defined as the proportion of time that a satisfactory supply is available:

$$A = 1 - \text{FailTime}/\text{TotalTime} \qquad (7.1)$$

Where

A	= Availability
FailTime	= Time that supply is in failure state
TotalTime	= Time interval considered

System availability can be defined as the average of nodal availabilities weighted by the demands to take into account the distribution of the demands. This index is a reasonable assessment of a system's general performance. It is suggested that systems should be designed for a specified system availability.

7.3 THE RELIABILITY TESTER

The reliability tester developed herein can be used either for the analysis and expansion of existing networks or for the design of new ones. Due to its rapid assessment of availability, it can also be incorporated into an optimisation framework.

The reliability tester comprises five models or procedures which are:

- Demand Model;

- Pipe Failure Model;

- Topological Model;

- Analysis Model;

- Overall Probabilistic Model.

7.3.1 Demand Model

Demands that must be provided at minimum specified pressures have in the past generally been taken as deterministic values in designing water networks. However, in reality, they are not. They change during the day, the week and the seasons, and develop over a period of years. They will also follow some form of probability distribution.

To ensure reliable delivery of water to the users, a water system must be designed to accommodate a range of expected demand patterns incorporating exceptionally high demands over part or all the system, these being the chief causes of hydraulic failure. The greater the demands in the network, the larger the pressure drops. Designing networks on this basis models operating conditions more realistically than adopting only one peak loading condition.

One simple way of representing the variable nature of expected loading conditions is to define a demand factor, KLoad, as the ratio of the actual demand to the time averaged demand. For simplicity, this factor is taken to apply uniformly over the network. ie

$$\text{KLoad} = \text{Demand / Time Averaged Demand} \tag{7.2}$$

There will be a relationship between demand factor and probability which can be determined from previously recorded data for customer demands. This can be presented in the form of a probability density function for demands, as in Fig. 7.1. The demand can be split into a finite number of zones usually equally spaced along the x-axis as shown in the same figure. For each zone, the mean demand factor can be determined with its corresponding probability such that:

$$\sum_{i=1} \text{PKLoad}_i = 1 \tag{7.3}$$

Where:

NLoad = Total No. of demand conditions,
PKLoad_i = Probability of demand condition i.

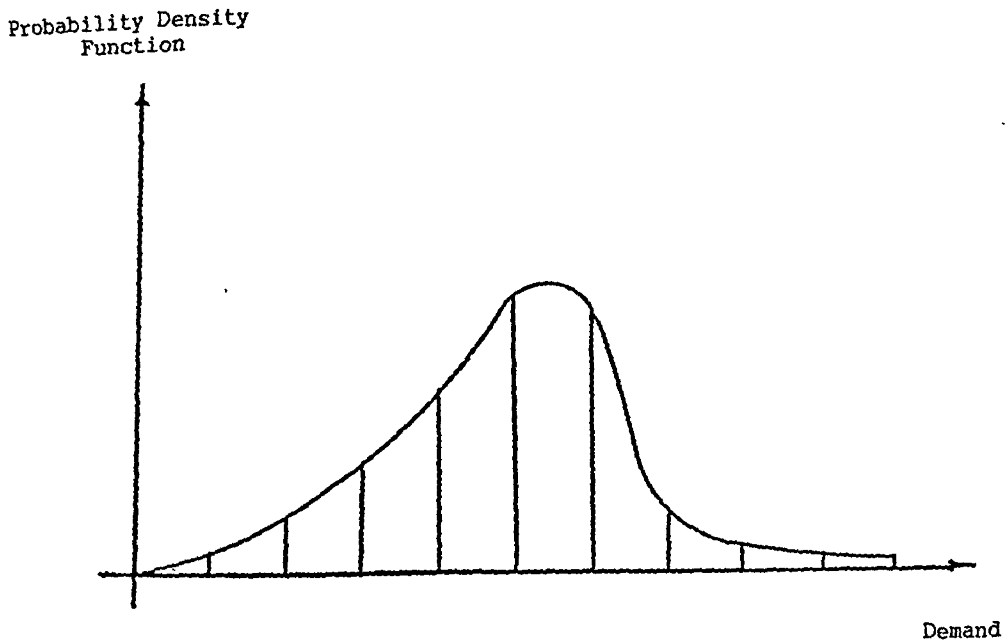


Fig. 7.1 Arbitrary Probability Density Function of Demands

(Bouchart & Goulter, 1991)

7.3.2 Pipe Failure Model

The second type of failure encountered in water networks is mechanical failure due to pipe breakage events.

It is assumed in this model that:

(1) The number of failures per unit time, F_j, $j = 1,..., N_p$ (number of pipes in the network) for each pipe j is known;

(2) Only one pipe fails at a time;

(3) Breakage events for pipes are independent;

(4) A pipe remains in a failed state for a fixed time interval, T_{rep}.

The second assumption was addressed when the cut sets method was discussed earlier. In addition, Su et al. (1987) concluded, based on the high availability of pipelines, that multiple pipe failure was generally so unlikely as to make analysis of such events unnecessary. An exception to this generalisation would be failures associated with fluid transient etc following power failure or similar events.

The assumption of independence of break events means that failure of one link does not imply an increased or decreased probability that other links will break. This has been questioned by some authors (Walski and Pelliccia 1982). However no data are available that can be used to develop a more complex dependent failure model. Due to the lack of such data this assumption is a reasonable simplification of the problem.

Once a pipe has failed, the model assumes it is isolated from the network which then functions in a reduced state for a fixed time interval until a repair is effected. The probability of pipe j failing, $Ppipe_j$, and the network consequently being in a reduced state is therefore $T_{rep}.F_j$, where T_{rep} and F_j have the same time base. Pipe failure probabilities are used later on for the assessment of nodal and system availabilities.

7.3.3 Topological Model

This model arranges all the data needed for the analysis of both the full network, where all the pipes are in an operational state, and the reduced network, where one of the pipes has failed and is removed from the network.

For a full network it is clear that when a pipe is removed, the data describing the network structure and connectivity will change. The objective of the topological model is to find and reorganise from the full network the data necessary for the reduced system before performing the analysis.

If for a full network there are N_P links, N_J nodes and N_L loops, in the reduced network there will generally be N_P - 1 links, N_J nodes and N_L- 1 loops. The exception to this is when the failed pipe is not part of a loop, in which case the network becomes disconnected, and water cannot be supplied to one or more nodes.

7.3.4 Analysis Model

According to Chapter 2, the technique selected for solving the non-linear algebraic equations for flow and pressure is the Linear Method. This method was chosen because of its high convergence characteristics, as outlined previously, and also because, unlike other approaches, it does not need an initial guess for link flows, which would further complicate the computation for the reduced networks.

The energy head loss equation used is the Hazen-Williams formula since it is widely adopted by practising engineers in the analysis of water distribution networks.

It should be noted that for the full network, once the analysis is performed, the resulting flows and pressures cannot be used as a basis for computing flows and pressures of a reduced network. This is obvious since flows will be rerouted in complex ways. However, when a pipe is out of action, the number of continuity equations will not change (N_J equations) but, continuity equations at the upstream and downstream ends of the failed pipe are affected. Moreover, if the failed pipe is among any loop pipes, this loop will be opened and then the number of energy equations will be reduction by 1 leaving only N_P -1 equations to be solved.

Therefore, it was necessary to develop an analysis model which is capable of dealing with both full and reduced networks. In the Reliability Tester, the topological model prepares the data and the analysis model solves the equations.

7.3.5 Overall Probabilistic Model

Since neither system component breakages alone nor hydraulic failures alone give an appropriate assessment of system reliability, these two factors are considered concurrently in this model.

The probabilistic model is connected directly to the demand and pipe failure models. These provide the necessary probabilities of loading conditions and of pipe failures respectively. It is also connected to the analysis model which supplies the calculated flows and pressures. The overall probabilistic model enables the nodal and system availabilities to be determined in the following fashion.

7.3.6 Nodal Availability

The nodal availability was defined earlier as the probability that a given node receives a sufficient supply at or above a minimum pressure.

The computation of the nodal availability requires the analysis of all topological states of the network (full network and reduced networks) under all the loading conditions. This computation can be summarized as follows:

1) Consider the Full Network

For each demand factor $KLoad_i$ with its corresponding probability $PKLoad_i$, i = 1,...NLoad:

- Analyse the system to obtain pressures H_j, where j refers to the node number, j = 1,... N_J;

- If $H_j < H_j^{min}$ then the probability of an insufficient supply at node j with a full network and this demand factor is:

$$F_{jk} = PKLoad_i \cdot PNetwork \tag{7.4}$$

Where PNetwork is the probability of no pipes being out of action and k is an index for counting violations of minimum pressure.

Strictly,

$$PNetwork = \prod_{ii=1}^{N_p} (1 - Ppipe_{ii}) \tag{7.5}$$

Where

$Ppipe_{ii}$ = probability of pipe ii, ii = 1,...,N_p, being in a failed state.

However, as only single pipe failure events are considered below, PNetwork is taken to be:

$$PNetwork = 1 - \sum_{ii=1}^{N_p} Ppipe_{ii} \tag{7.6}$$

2) Consider Reduced Networks

For each pipe Lk with its corresponding probability of failure $Ppipe_{Lk}$, Lk = 1,...,N_p and for each loading condition $KLoad_i$ with its corresponding probability $PKLoad_i$, i = 1,... NLoad:

- Assume pipe Lk is out of service,

- Analyse resulting reduced network for H_j, j = 1,..., N_J

- If $H_j < H_j^{min}$ then $F_{jk} = Ppipe_{Lk} \cdot PKLoad_i$ (7.7)

In this way all combinations of failure mode and loading conditions are considered.

3) Sum Nodal Pressure Violations

For each node j, $j = 1,..,N_J$, the probability of hydraulic failure is:

$$Fnode_j = \sum_{k=1}^{K_{max}} F_{jk} \tag{7.8}$$

Where

K_{max} = No of times the pressure test was violated

4) Calculate Nodal Availabilities

For each node j, $j = 1,...., N_J$ the availability is:

$$Anode_j = 1 - Fnode_j \tag{7.9}$$

7.3.7 System Availability

System availability can be expressed as a mean of all nodal availabilities weighted by their demands.

$$A_{Net} = \sum_{j=1}^{N_J} (Anode_j \cdot Q_j^{ext}) / \sum_{j=1}^{N_J} Q_j^{ext} \tag{7.10}$$

Where

Q_j^{ext} = Time Averaged demand at node j,

A_{Net} = Network availability.

7.4 SOFTWARE

The algorithm for determining nodal and system availabilities can be summarized as follows:

<u>Algorithm</u>

```
Pipe = 0
Repeat
    Topological Model
    Load = 1
    Repeat
        Demand Model
        Analysis Model
        Pipe Failure Model
        Overall Probabilistic Model
        Load = Load + 1
    Until Load > Nload
    Pipe = Pipe + 1
Until Pipe > P
```

A computer program for evaluating nodal and system availabilities was developed in TURBO PASCAL on a PC. The input data needed are those which describe the full network topology, the design demands with the KLoad demand factors and their corresponding probabilities, and the probabilities and constant duration of pipe failures (T_{rep}).

7.5 EXAMPLE

For illustration and clarification of the Reliability Tester a hypothetical three-loop, eight-pipe and six-node network (Network D), is shown in Fig. 7.2, and used for demonstration. Mean demands and ground levels for each node are listed in Table 7.1. The network is supplied by one source at node one.

Pipe diameters have been deliberately chosen to be insufficient to represent an overloaded network in need of reinforcement. It should be noted that in the design of a new network, demands considerably greater than the mean demands would be used for sizing the pipes. This would then give much higher levels of system and nodal availabilities than the values used for this example.

Link data (pipe length, diameter) are summarised in Table 7.2. Minimum permissible pressure and Hazen-Williams coefficients are assumed to be 20m and 130 respectively for all the system.

As outlined earlier the model requires:

(1) probabilities for pipe failures, $Ppipe_i$, $i = 1,...,N_P$

(2) duration of pipe failures, T_{rep}

(3) the probabilistic distribution of the demands, $PKLoad_i$,

$i = 1,...NLoad$.

With respect to the first point, the probability of pipe failures can be determined on the basis of rate of breakage of pipes. Some historical data on pipe breakage rates that have been published in a few studies in the literature were presented in Chapter 1.

For this application the average of pipe breakage rates of all the cities mentioned in Chapter 1, excluding the typical Winnipeg data, has been used.

Node Number	Mean Demand (L/s)	Ground Level (m)
1	-150	200 (Source)
2	20	158
3	30	158
4	40	148
5	30	155
6	30	173

Table 7.1 Characteristics of Demand Nodes for Network D

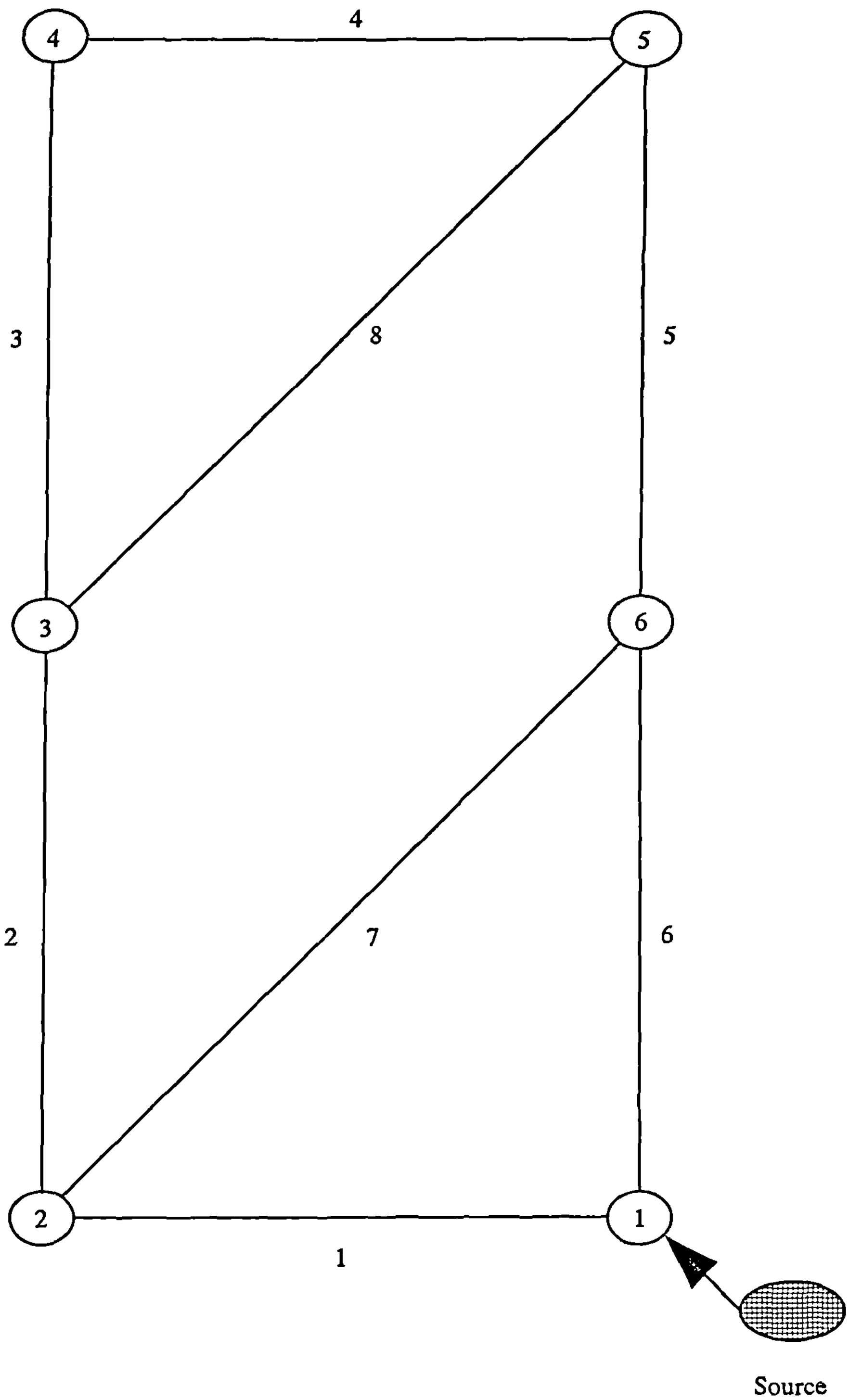


Figure 7.3 Network D

Link Number	Length (m)	Diameter (mm)
1	1000	250
2	1000	250
3	1000	100
4	1000	200
5	1000	200
6	1000	300
7	1000	150
8	1000	150

Table 7.2 Characteristics for Network D

Knowing rate of breakage, rD_i, expressed in breaks/km/year of pipe i, $i = 1,...N_p$, of length L_i and diameter D_i, the mean probability of failure of this pipe in a day can be expressed by:

$$pD_i = rD_i \cdot L_i / 365 \tag{7.11}$$

Where

pD_i = Mean probability of failure per day of pipe i of diameter D.

L_i = Length of pipe i in km.

rD_i = Breakage rate per year of pipe i of diameter D.

Diameters and their corresponding probabilities of failure per kilometre are presented in Table 7.3.

The assumed values of mean pipe probabilities of failure (*) correspond to the probabilities of failure of diameters that have not been available in the literature.

In this example, the duration of the mechanical failure of a pipe T_{rep}, is taken as 1 day for each breakage.

Diameter (mm)	Mean Probability of Failure ($km^{-1}day^{-1}$)
50	0.002000
75	0.001500*
80	0.001000*
100	0.000901
125	0.000680*
150	0.000468
175	0.000300*
200	0.000192
225	0.000170*
250	0.000150*
275	0.000130*
300	0.000107
350	0.000107
375	0.000090*
400	0.000071
450	0.000071*
500	0.000071*
525	0.000071*
600	0.000060*

Table 7.3 Mean Pipe Failure Probabilities by Diameter

* Assumed values.

With reference to point (3) which deals with the probabilistic distribution of the demands, the daily demand data used in this example corresponds to a region in South West England for the period of 1976 to 1989 inclusive (South West Water, 1993). The data was first processed so that a distribution function relating probability to demand was obtained based on the information contained in this sample of 14 years.

The selection of the most suitable probability function was based on a sensitivity analysis. Four types of distribution were considered for this application. These were the Normal, the Log-Normal, the Weibul and the Exponential. Using the Kolmogorov-Smirnov Test, the normal distribution was found to be the closest theoretical distribution function for this daily water consumption sample. The distribution function and the graphical representation of the data on normal distribution probability paper are presented in Figures 7.3 and 7.4.

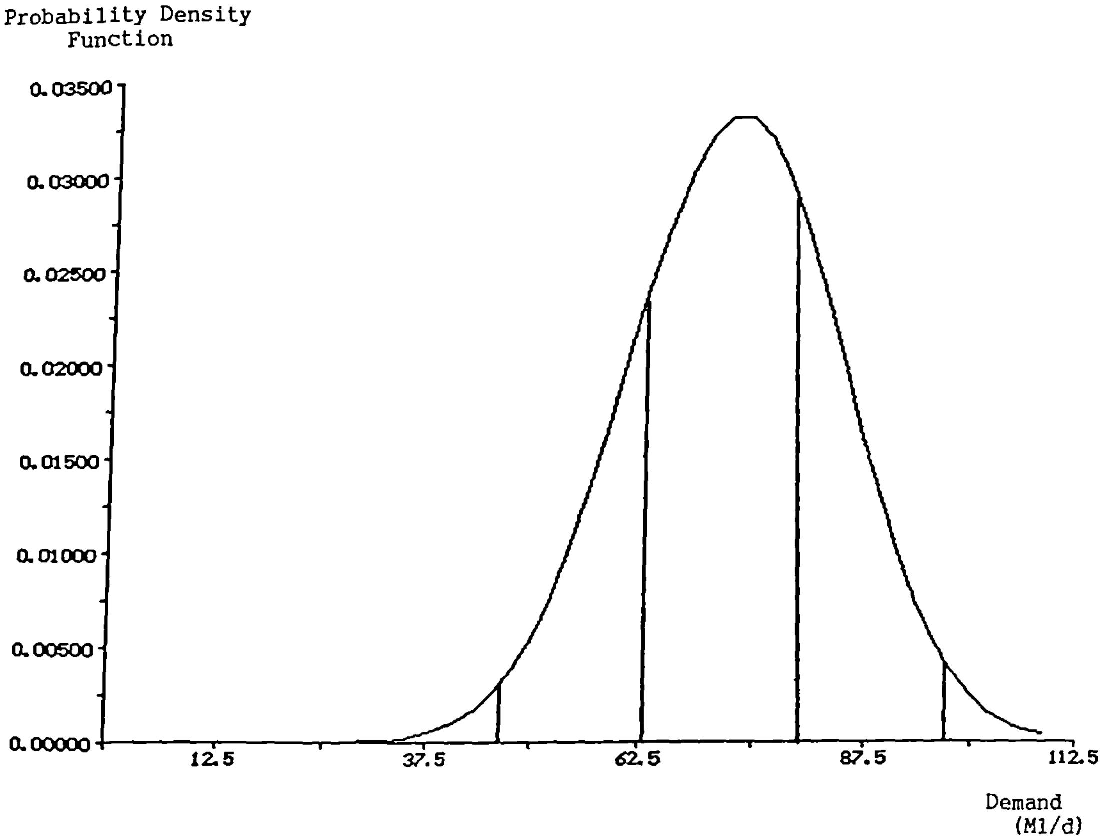


Fig. 7.3 Distribution of Daily Demands (1976-1989)

for a region in South West England

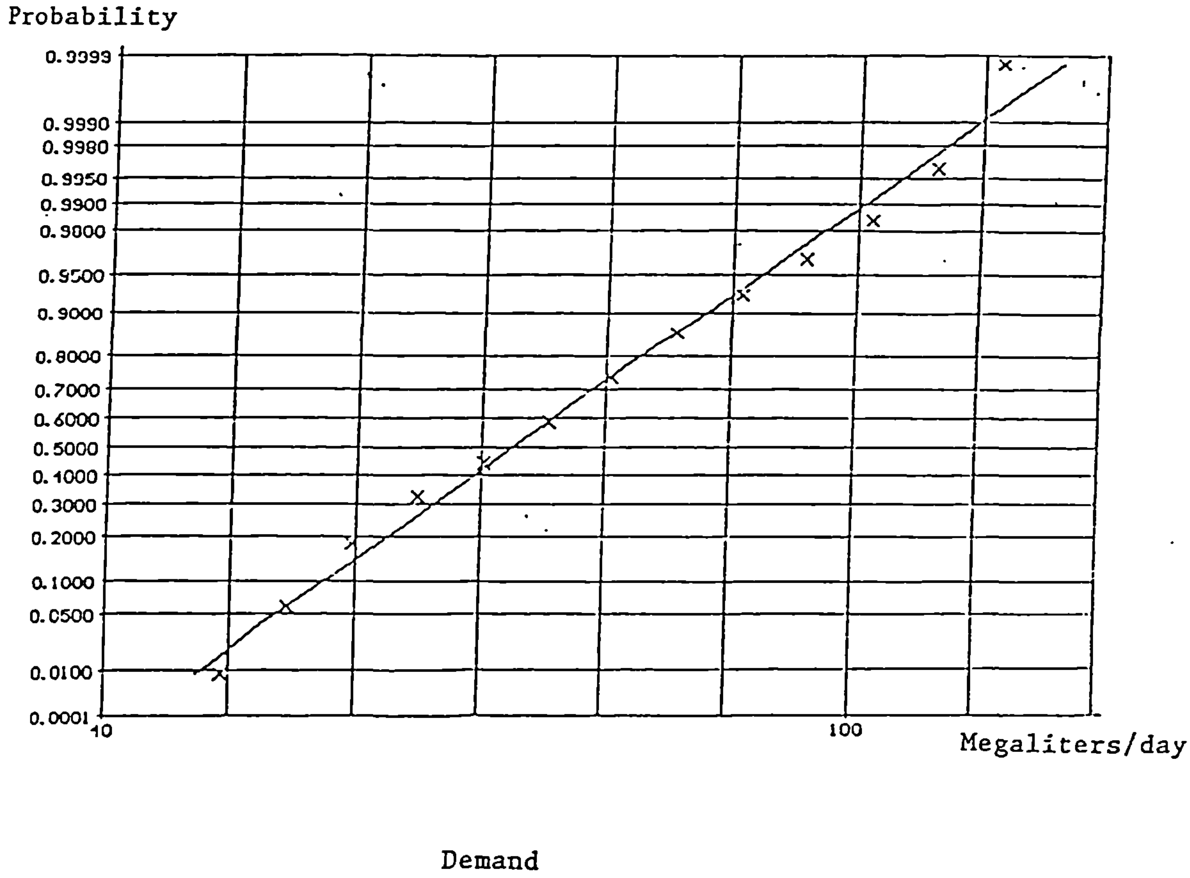


Figure 7.5 Graphical Representation of Daily Demands on Normal Distribution probability paper

After splitting the area under the curve of Fig. 7.3 into 5 equal zones (NLoad = 5) along the x-axis as shown in the same figure, the demand factor Kload can be determined as:

$$KLoad_i = Q_i^{ext} / Q^{avg}, \qquad \forall\, i \in NLoad \tag{7.12}$$

where

$KLoad_i$ = Demand factor for zone i;

Q^{avg} = Overall Average Demand ;

Q_i^{ext} = Time Averaged demand at zone i.

The probabilities corresponding to $KLoad_i$, i = 1,...,NLoad, are computed directly from the normal distribution, the results being shown in Table 7.4. The demand patterns for all nodes of Network D are determined using Eq. 7.12 such that:

$$Q_{ij}^{ext} = KLoad_i * Q_j^{avg} \tag{7.13}$$

Where

Q_{ij}^{ext} = External Demand for loading condition i, i = 1,...,NLoad, at node j, j = 1,...,N_J;

$KLoad_i$ = Demand factor i, i = 1,...,NLoad;

Q_j^{avg} = Time averaged demand at node j, j = 1,...N_J.

Kload	0.56	0.77	0.99	1.21	1.43
Probability	0.0209	0.2127	0.4900	0.2545	0.0219

Table 7.4 Demand Factors and their Probabilities

Running the reliability tester with Network D (Run #1) gives an artificially low system availability of 91.26%. Table 7.5 shows details of nodes at which pressure is below minimum for full and reduced states. Nodal availabilities are presented in Table 7.6. Examination of Table 7.6 shows also that node 1 has an availability of 100 %, corresponding to sufficient supply of the source as assumed. The least reliable node is the one furthest from the source (node 4), while the highest availability corresponds to the node nearest to the source (node 2). Nodal availabilities for the rest of the junctions (3, 5 and 6) are close to each other.

The Reliability Tester is a powerful tool that can help in detecting nodes with insufficient supply and can also help in determining the major sources of unreliability. In the absence of pipe breakage, failures are due to infrequent sets of high demands. With the pipe diameters and loadings specified in this example, it can be seen from Table 7.5 that no such failure was produced for node 2. Node 4 failed with a demand factor of 1.21, with the remaining nodes failed only when a demand factor of 1.43 was imposed.

When pipe breakages are considered, the supply pressure resulting from the reduced network and probabilistic demands is in all cases insufficient somewhere in the system, except when pipe 1 fails with a demand factor of 0.56 or when pipes 3, 7 or 8 fail with demand factor less than or equal to 0.99. (see Table 7.5). Not surprisingly, pipes 1 and 6 are the greatest cause of unreliability in this system since they are the links connected directly to the source. Almost 41% of the number of failures are due to these pipes. Breakage of pipe 2 causes 16% of all failures, with insufficient pressure at node 3 for all demand factors, and at nodes 4 and 5 for all but the lowest demands. Node 6 fails only when the demand factor is 0.99 and over. Failure of pipe 7 is the least critical for the system.

Kload		0.56	0.77	0.99	1.21	1.43
Probability		0.0209	0.2127	0.4900	0.2545	0.0219
State of the Network	Probability					
Full	0.996932	0	0	0	4	3, 4, 5, 6
Reduced						
Pipe 1	0.000150	0	2, 3, 5, 6	2, 3, 4, 5 6	2, 3, 4, 5 6	2, 3, 4, 5 6
Pipe 2	0.000150	3	3, 4, 5	3, 4, 5, 6	3, 4, 5, 6	3, 4, 5, 6
Pipe 3	0.000901	0	0	0	4, 5, 6	4, 5, 6
Pipe 4	0.000192	4	4	4	4	3, 4, 6
Pipe 5	0.000192	5	4, 5	3, 4, 5	3, 4, 5	2, 3, 4, 5
Pipe 6	0.000107	5, 6	3, 4, 5, 6	2, 3, 4, 5 6	2, 3, 4, 5 6	2, 3, 4, 5 6
Pipe 7	0.000468	0	0	0	4	3, 4, 5, 6
Pipe 8	0.000468	0	0	0	4, 5, 6	4, 5, 6

TABLE 7.5 Nodes at which Pressure is below minimum

Besides having the possibility of diagnosing water networks, the Reliability Tester can help planners and designers of water distribution networks to improve system and nodal reliabilities. In fact, any alternative improvement options proposed by the designer can be simulated by the model, which has the ability to detect better alternatives and highlight poorer schemes.

Node Number	Initial Availability (%) Run #1	Availability (%) Run #2	Availability (%) Run #3	Availability (%) Run #4
1	100.0000	100.0000	100.0000	100.0000
2	99.9767	99.9767	99.9893	99.9893
3	97.7594	97.7585	99.9516	99.9519
4	72.3072	97.7300	99.9255	99.9255
5	97.7176	97.7293	97.7572	99.9458
6	97.7396	97.7506	97.7800	99.9860

Table 7.6 Nodal Availabilities (%) for Network D

For instance, the effect of increasing the minimum pipe size on the availability can be studied by the model. The minimum pipe size (100 mm) for Network D was increased to 150 mm (Run #2). With this alteration, the system availability becomes 98.04 % (previously 91.26%), demonstrating a significant increase in overall availability, and a very large increase in the availability of node 4 (see Table 7.6). This occurs for two reasons. First, the greater capacity of larger pipes leads to a reduction in hydraulic failures for the complete network under high demands (node 4 now fails only with Kload = 1.43 for the complete network). Second, in the pipe failure model larger pipes fail less frequently. This latter result is in line with the comment by Fujiwara and Tung (1991) that improvement in the reliability of water systems may be achieved through increasing pipe sizes. The Reliability Tester does however assume that there are no technical problems in adopting larger minimum pipe sizes, such as the violation of minimum flow velocities.

If increased nodal availabilities for junctions 3, 4 and 5 are required (initial availability around 97.72%), this may be achieved by increasing pipe sizes along the paths supplying these nodes. For example, increasing the diameter of link 1 from 250 mm to 300 mm (Run #3) increases the system availability to 99.08%, with nodal availabilities of greater than 99.90% for junctions 3 and 4. (See table 7.6). Similar improvements for nodes 5 and 6 can be achieved by increasing the diameter of pipe 6 from 300 to 350 mm (Run #4). In this way system and nodal availabilities greater than 99.92% are achieved.

Calculation of nodal and system availabilities for Network D requires very little computer time. Approximately 2 seconds on a 486 IBM compatible PC were required.

7.6 NETWORK DESIGN

As shown above, the Reliability Tester has the ability of detecting nodes with insufficient supply in addition to the major causes of unreliability. Furthermore, it was seen earlier that the Reliability Tester can help designers and planners to improve nodal and system availabilities. However this possibility is only offered when dealing with small size networks. If for a medium or large size network, the information on availability (for the system or the individual nodes) computed by the model, is considered not satisfactory, an algorithm for improving the reliability can be applied such as that of Fujiwara and Tung (1991) or Bouchart and Goulter (1991).

Ideally, improvement of nodal and system availabilities would be achieved through an optimisation model. In such a model, the objective could be the maximisation of the overall availability of the network subject to a cost limitation and technical constraints. Technical constraints would include bounds on pipe velocities, especially on the permissible minimum velocity since the optimisation procedure will tend to select large diameter pipes which are less prone to mechanical failure and reduce hydraulic failures.

An alternative approach is to optimise the network on the basis of minimum cost, with constraints on the reliability of the system and individual nodes.

In either model, the decision variables would be pipe diameters. The reliability tester would be incorporated into both optimisation models to calculate nodal and system availabilities.

7.7 CONCLUSIONS

The reliability tester developed herein is a new model based on a stochastic simulation for fast assessment of nodal and system availabilities in a water distribution network.

The model can be used to analyse an existing water system, identifying critical nodes with serious supply problems, and the major causes of unreliability. This approach is therefore useful in the detailed investigation of promising options for the improvement of reliability in existing systems.

The model can also be used in the design of new systems and the expansion of existing networks where the focus is on provision of a system of specified reliability or where the reliability is to be maximised.

Owing to its simplicity and fast computation, the model can also be efficiently and successfully incorporated into an optimisation procedure such as that presented in Chapter 9, for determining least cost design of water distribution networks under reliability constraints.

Chapter VIII

FLOW ASSIGNMENT PLUS LINEAR PROGRAMMING METHOD

Chapter VIII

FLOW ASSIGNMENT PLUS LINEAR PROGRAMMING METHOD

8.1 INTRODUCTION

This chapter implements the first approach to the optimal design of water networks with reliability issues. As mentioned in Chapter 5, reliability of water systems is, in this approach, assessed in terms of the assignment of link flows on the basis of entropy. With the distribution of flow within the network thus defined, the LP technique outlined in Chapter 6 is applied for finding both the optimal set of pipe diameters and the network cost.

In the remainder of this Chapter, the detail of the Flow assignment plus LP method is presented. Applications of the model to the design of three water networks are presented. Network C presented previously in Chapter 6 and Networks E and F defined later. Finally, comparisons of the entropy-based flow assignment method and arbitrary distributions of flow ends this Chapter.

8.2 DETAILS OF THE MODEL STRUCTURE

This model makes use of the two computer programs developed: PATH_Q, for the assignment of flow within a given network, and LNOPTNET for the optimal design of this network. It also uses the routine SizeD for determining candidate diameters related to each link.

Given a network having topological data and external demands, PATH_Q is first applied to obtain the link flows (Q_i, $i = 1,...,N_p$). However, before calling LNOPTNET, candidate diameters for each link must be specified. These may be obtained by running SizeD. The data required are:

(1) The link flows, Q_i, $i = 1,...,N_p$ that are supplied by PATH_Q;

(2) The minimum and maximum bounds on velocity V_{min} and V_{max}; and

(3) The set of commercially available pipe diameters: D_i, $i = 1,...,N_d$, where N_d is the maximum number of available sizes.

The outputs of SizeD are:

(1) Number of candidate diameters n(i), $i = 1,...,N_p$, for link i; and

(2) Candidate diameters for each link i: D_{ij}, $i = 1,...,N_p$, $j = 1,..., n(i)$.

In addition to the candidate pipe sizes for each link and their numbers, LNOPTNET requires the unit costs of each standard pipe size, $CostD_i$, $i = 1,..., N_d$ and the minimum nodal pressures.

The execution of LNOPTNET gives both the optimal network cost and pipe segments (X_{ij}, $i = 1,..,N_p$, $j = 1,..,n(i)$) with their corresponding diameters.

In summary, with respect to this approach, the design of water networks with reliability specifications, measured in terms of entropy, is represented by the following three-step algorithm:

Algorithm

PATH_Q;
SizeD;
LNOPTNET.

8.3 EXAMPLES

Under this heading, LNOPTNET is run with three example problems: Networks C, E and F. Network C (see Fig. 6.1) is a four-loop network whilst Networks E and F, which are shown in Figures 8.1 and 8.2 below, are larger and more realistic networks. For simplicity all three water systems are assumed to be on a level plane although this restriction is not necessary since the pressure required may be different from node to node. Networks C, E and F are designed on the basis of the stochastic demand factors proposed in Chapter 7. The demand factor selected is Kload = 1.43 giving design demands equal to 1.43 * Mean demands. With this value, the Reliability Tester will cause no hydraulic failures for the complete network and so in the absence of pipe failures the most pressure-critical nodes will always receive their demands at adequate pressure. All design demands for the three networks presented later are likewise assumed to be weighted by the same demand factor, Kload = 1.43.

The cost of pipes is commonly considered as a function of the pipe diameter D and the length L (Water Research Center, 1977; U.S. Army Corps of Engineers, 1980). The unit cost, CostD (eg \$/m, £/m) is usually taken as:

$$\text{CostD} = \rho D^{\Theta} \tag{8.1}$$

Where ρ = constant cost coefficient and Θ = exponent greater than 1. For this research, the cost coefficients considered (ρ and Θ) are those used by Tanyimboh and Templeman (1993). These are: ρ = £900 and $\Theta = 2.4$ when D is in metres.

For the purpose of comparison, Network C will be designed on the basis of two distributions of flow: (1) for the flow distribution of Coals and Goulter (1985), (2) for the flow distribution resulting from the application of the path method.

8.3.1 Network C

Results of running PATH_Q with Network C and the flow distribution proposed by Coals and Goulter are tabulated in Table 8.1.

To apply SizeD to the two distributions of flow for the candidate diameters, the minimum and maximum limits on velocity should be specified. These, which are also considered for the design of networks E and F, are 0.5 m/s and 3.0 m/s respectively. Such bounds may cause the numbers of candidate diameters for links to vary. For reasons of reducing the number of decision variables, at most 5 candidate diameters, starting from the minimum pipe size for each link, are considered if they are available under the velocity constraints. This strategy will be followed in the design of the others networks (E and F). Table 8.2 summarised the candidate diameters for each distribution of flow.

Link Number	Flow by Path_Q (l/s)	Coals & Goulter Flow Distribution (l/s)
1	208.0	208.0
2	104.0	125.0
3	38.2	34.0
4	17.2	13.0
5	44.8	70.0
6	104.0	83.0
7	44.8	34.0
8	34.3	49.0
9	30.5	41.0
10	34.3	34.0
11	38.2	28.0
12	17.2	7.0
13	30.5	20.0

Table 8.1 The Two Flow Distributions for Network C

Link	Candidate Diameters (mm) PATH_Q	Coals & Goulter (1985)
1	300 350 375 400 450	300 350 375 400 450
2	225 250 275 300 350	250 275 300 350 375
3	175 200 225 250 275	125 150 175 200 225
4	100 125 150 175 200	75 80 100 125 150
5	175 200 225 250 275	175 200 225 250 275
6	225 250 275 300 350	200 225 250 275 300
7	150 175 200 225 250	125 150 175 200 225
8	125 150 175 200 225	150 175 200 225 250
9	125 150 175 200 225	150 175 200 225 250
10	125 150 175 200 225	125 150 175 200 225
11	150 175 200 225 250	125 150 175 200 225
12	100 125 150 175 200	75 80 100 125
13	125 150 175 200 225	100 125 150 175 200

Table 8.2 Candidate Diameters for the Two Flow Distributions for Network C

Path	Number of Links in Path	Link Number
	Node Equation	
1	1	1
2	2	1 2
3	3	1 2 3
4	4	1 2 3 4
5	3	1 2 5
6	2	1 6
7	3	1 6 11
8	4	1 6 11 12
9	5	1 6 11 12 13
	Loop Equation	
10	4	7 -5 -2 6
11	4	8 -4 -3 5
12	4	13 -9 -8 10
13	4	12 -10 -7 11

Table 8.3 Strings of Pipes for Pressure and Loop Constraints

Before executing LNOPTNET with Network C for the two different flow distributions, strings of pipes for pressure and loop constraints have to be specified. These are given in Table 8.3.

The results of running LNOPTNET are given in Table 8.4. It follows from Chapter 6 that the solutions given should not be considered as finished. These solutions must be adjusted where appropriate (Links with two segments) for the reasons discussed in Chapter 6 (Section 6.6). Moreover, for two-section

Link	PATH_Q		Coals & Goulter (1985)	
	X (m)	D (mm)	X (m)	D (mm)
1	1000.00	400	1000.00	400
2	1000.00	300	127.28	300
			872.72	350
3	361.33	200	1000.00	200
	638.67	225		
4	1000.00	175	1000.00	150
5	788.50	225	445.76	250
	211.50	250	554.24	275
6	1000.00	300	1000.00	275
7	788.50	225	421.53	200
	211.50	250	578.47	225
8	1000.00	225	1000.00	250
9	1000.00	225	1000.00	225
10	1000.00	225	465.21	200
			534.79	225
11	361.33	200	450.05	175
	638.67	225	549.95	200
12	1000.00	175	1000.00	125
13	1000.00	225	1000.00	200
Cost:	£426648		£426242	

Table 8.4 Optimal Solutions of Network C for the two flow Distributions

Link	L (m)	D (mm)
1	1000	400
2	1000	300
3	1000	225
4	1000	175
5	1000	225
6	1000	300
7	1000	225
8	1000	225
9	1000	225
10	1000	225
11	1000	225
12	1000	175
13	1000	225

Table 8.5 Practical Solution of Network C for the Path Approach

links, lengths of segments should be rounded up/down in order to obtain a more satisfactory engineering solution. Therefore, one of the practical solutions corresponding to the design on the basis of the path approach (PATH_Q) is presented in Table 8.5.

The practical solution presented in Table 8.5 is composed of one single diameter per link. 225 mm are proposed for Links 3, 5, 7 and 11. This alteration corresponds to an insignificant increase (0.033 %) in the network cost compared to that given in Table 8.4. The solution shown in Table 8.5 has been analysed using **LMANLS** to see the impact of this alteration of segment lengths on the system pressures. Table 8.6 gives the results of the analysis of network C where it can be seen that all nodal pressures are greater than the specified minimum bound (25m).

Node	Pressure (m)
1	43.84
2	36.92
3	32.16
4	28.12
5	31.41
6	36.92
7	32.16
8	28.12
9	25.22

Table 8.6 Nodal Pressures of Network C

With respect to the network costs of the two solutions, it may be seen from Table 8.4 that the path approach gives a solution cost (£426648) a little bit greater than that produced on the basis of the flow distribution of Coals and Goulter (£426242).

Comparison of the reliabilities of the two solutions was performed using the Reliability Tester. This will be dealt with later in subsection 8.4, after the optimisation of networks E and F.

8.3.2 Network E

The second numerical example selected is Network E (Awumah, Bhatt and Goulter, 1989). Network E (N_F = 1, N_J = 12, N_p = 18 and N_L = 6) is shown with an assumed flow direction in Fig. 8.1. Table 8.7 lists the network demands. The source head and the minimum pressure for all nodes are set to 55m and 30m respectively. Hazen-Williams coefficients are equal to 130. All network links are 1000m long.

Node	Demand (l/s)
Source	-531
1	83
2	28
3	42
4	42
5	42
6	42
7	56
8	56
9	56
10	28
11	28
12	28

Table 8.7 Demands of Network E

Flows resulting from the PATH_Q program and an arbitrary flow distribution are detailed in Table 8.8. The sets of candidate diameters for network E, resulting from running SizeD with this network for the two flow distributions, and the strings of pipes for pressure and loop constraints are listed in Appendix B (Tables B1 and B2).

The last step in the design of Network E is performed by running LNOPTNET. The optimal costs and the optimal diameters are tabulated in Table 8.9.

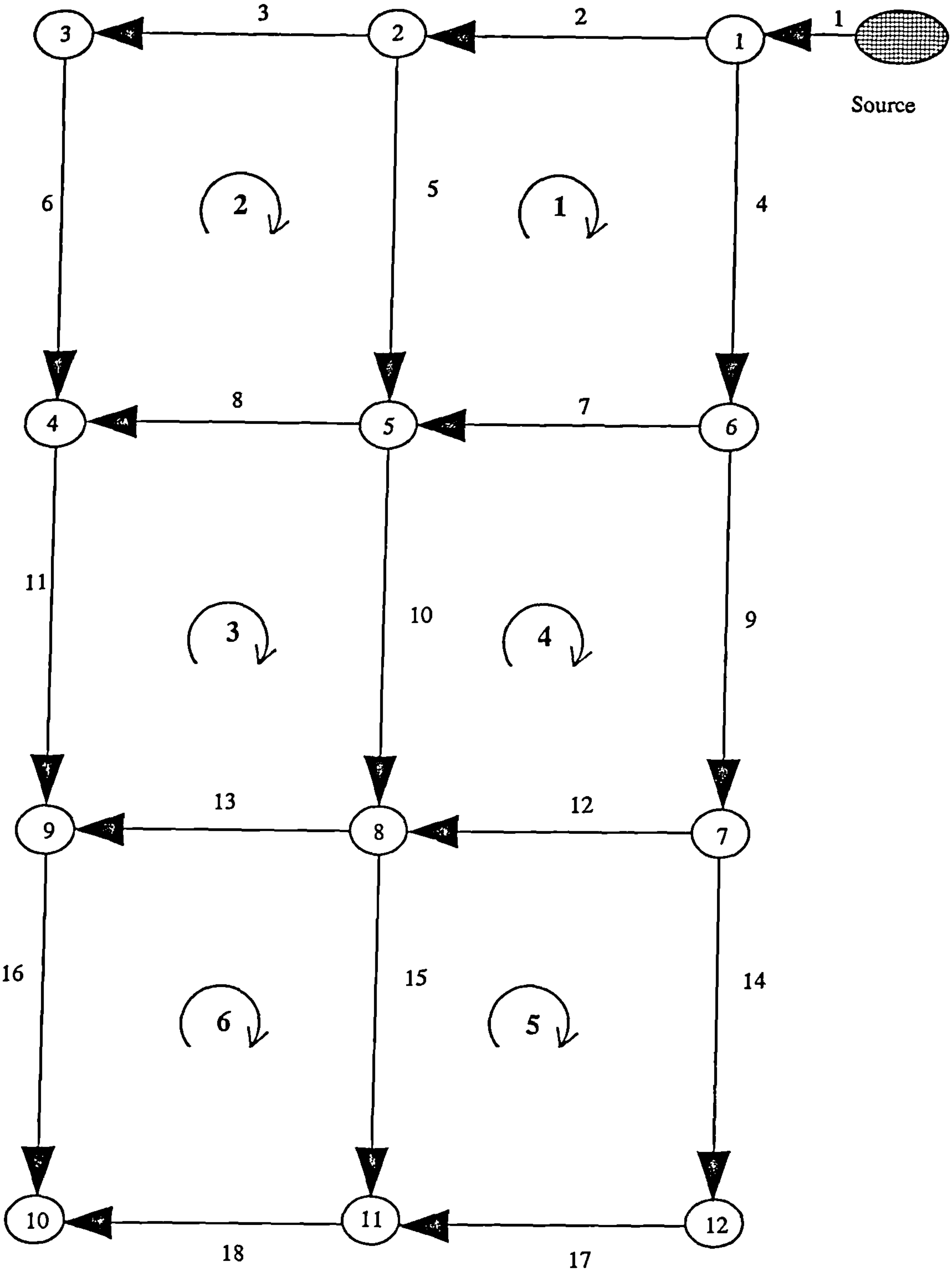


Figure 8.1 Network E

Link	Flow by PATH_Q (l/s)	Arbitrary Flow Distribution (l/s)
1	531.00	531.00
2	183.87	248.00
3	68.13	92.00
4	264.13	200.00
5	87.73	128.00
6	26.13	50.00
7	87.73	16.00
8	52.27	32.00
9	134.40	142.00
10	81.20	70.00
11	36.40	40.00
12	40.60	44.00
13	36.40	28.00
14	37.80	42.00
15	29.40	30.00
16	16.80	12.00
17	9.80	14.00
18	11.20	16.00

Table 8.8 The Two Flow Distributions for Network E

Examination of Table 8.9 shows that for Network E, the maximum network cost (£1002756) corresponds to the arbitrary flow distribution. For the entropy basis approach, the path approach gives a network cost (£988845) less than that corresponding to the arbitrary distribution (£1002756). It should be noted that this result is the opposite to that seen with network C.

Link	PATH_Q		Arbitrary Flow Distribution	
	X (m)	D (mm)	X (m)	D (mm)
1	1000.00	600	1000.00	600
2	891.18	400	438.54	450
	108.82	450	561.46	500
3	372.33	250	1000.00	300
	627.67	275		
4	661.81	450	1000.00	400
	338.19	500		
5	1000.00	300	67.55	300
			932.45	350
6	1000.00	200	1000.00	250
7	1000.00	300	52.79	150
			947.21	175
8	1000.00	250	619.34	200
			380.66	225
9	1000.00	350	399.20	350
			600.80	375
10	34.73	275	1000.00	275
	965.27	300		
11	1000.00	225	1000.00	250
12	662.43	225	1000.00	250
	337.57	250		
13	79.33	200	1000.00	225
	920.67	225		
14	30.22	200	1000.00	225
	969.78	225		
15	1000.00	225	834.01	200
			165.99	225
16	1000.00	200	1000.00	150
17	1000.00	150	182.36	150
			817.64	175
18	1000.00	150	1000.00	200
Cost:	£988845		£1002756	

Table 8.9 Optimal Solutions of Network E for the two flow Distributions

8.3.3 Network F

Network F ($N_F = 1$, $N_J = 7$, $N_p = 13$ and $N_L = 6$), which is shown in Fig. 8.2, is the final water system selected for comparison between the flow distributions approaches. The nodal demands are presented in Table 8.10. All network links are 1000m long and have the same Hazen-Williams coefficient (130). The source head is assumed to be 90m while the minimum pressure for all nodes is 30m.

Node	Demand (l/s)
Source	-248.82
1	00.00
2	42.90
3	42.90
4	42.90
5	34.32
6	42.90
7	42.90

Table 8.10 Demands of Network F

Table 8.11 summarises the results of the application of the PATH_Q software along with an arbitrary distribution of flow.

Appendix B (Tables B3 and B4) contains also the sets of candidate diameters, resulting from running the SizeD program, and the strings of pipes for pressure and loop constraints for Network F.

The execution of LNOPTNET with network F for the two flow distributions gives the results summarised in Table 8.12.

Link	Flow by PATH_Q (l/s)	Arbitrary Flow Distribution (l/s)
1	248.82	248.82
2	80.44	120.60
3	33.60	50.00
4	50.41	3.80
5	41.11	10.90
6	24.67	34.00
7	22.88	2.00
8	8.58	2.00
9	2.86	30.32
10	42.18	32.00
11	84.37	50.70
12	80.44	20.00
13	168.38	128.22

Table 8.11 The Two Flow Distributions for Network F

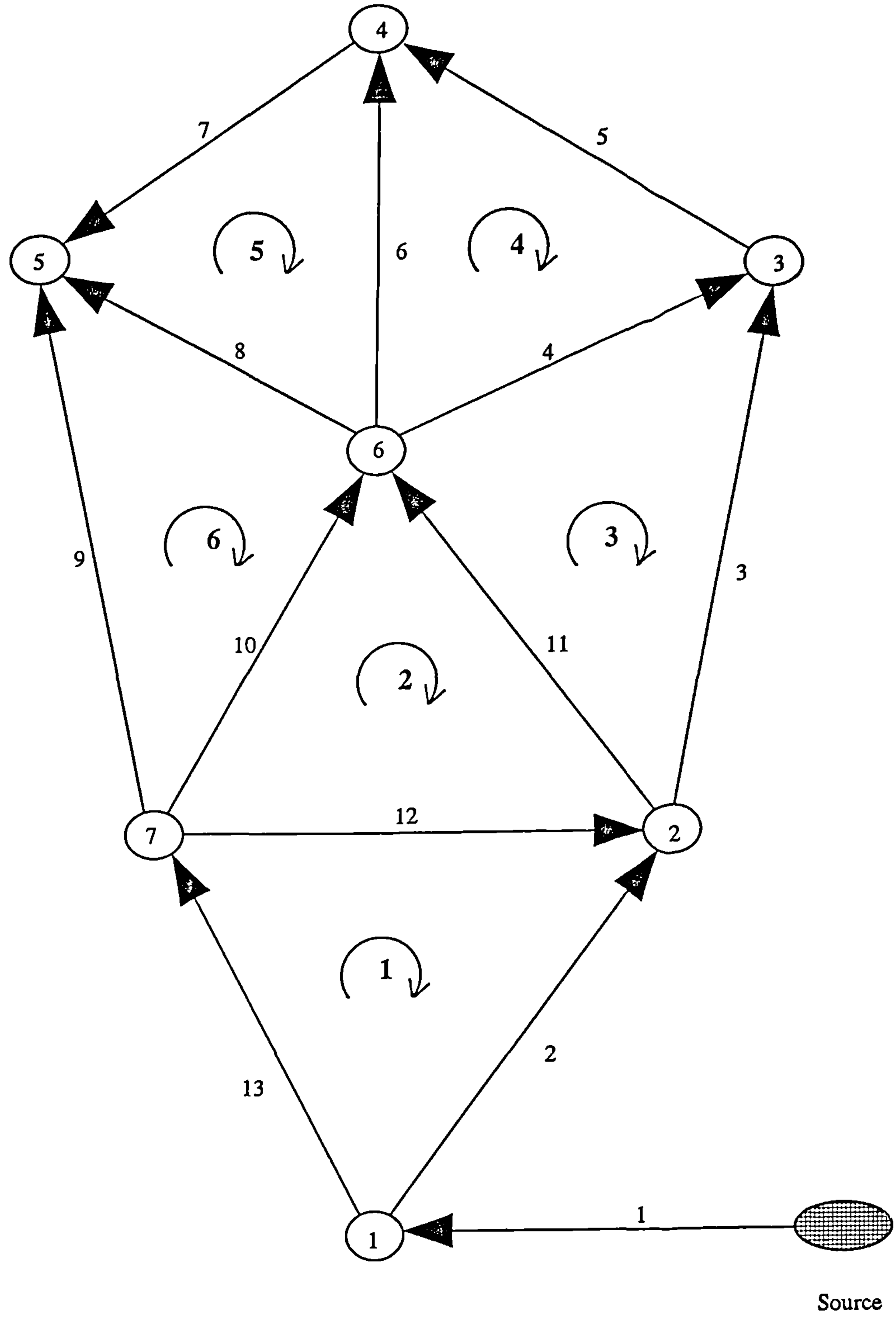


Figure 8.2 Network F

Link	PATH_Q		Arbitrary Flow Distribution	
	X (m)	D (mm)	X (m)	D (mm)
1	1000.00	375	1000.00	350
2	605.92	200	700.92	250
	394.08	225	299.08	275
3	207.69	150	339.92	175
	792.31	175	660.08	200
4	1000.00	225	1000.00	80
5	1000.00	225	1000.00	150
6	770.53	150	35.14	150
	229.47	175	964.86	175
7	1000.00	200	1000.00	75
8	37.14	80	513.84	50
	962.86	100	486.16	75
9	501.93	50	262.18	125
	498.07	75	737.82	150
10	764.43	175	96.89	150
	235.57	200	903.11	175
11	1000.00	275	1000.00	225
12	564.21	250	1000.00	175
	435.79	275		
13	1000.00	300	921.59	275
			78.41	300
Cost:	£345513		£254636	

Table 8.12 Optimal Solutions of Network F for the

two flow Distributions

For network F, examination of Table 8.12 shows that the greater optimal cost (£345513) of the two distributions of the flow is related to the design obtained by the application of the path approach. The arbitrary distribution of flow produces a lower cost (£254636).

8.4. COMPARISON OF FLOW ASSIGNMENT METHODS

8.4.1 Cost

The results of the application of LNOPTNET to the sample networks show that the design on the basis of entropy gives cheaper solution (£988845) than that produced on an arbitrary basis (£1002756) for network E while for networks C and F, the opposite has been found. As far as the cost is concerned, the optimal cost of a network depends on the distribution of flow as was exemplified by the Linear Programming Gradient Method of Alperovits and Shamir (1977).

To finish the comparison between the design of the three networks, the reliability aspect must be considered. This point is addressed in the following subsection.

8.4.2 Reliability

It was seen in Chapter 5 that designing water networks on the basis of distributing link flows using entropy may result in producing reliable networks.

It would be interesting to verify this point. This can be investigated by applying the Reliability Tester to the solutions produced by the arbitrary flow distributions and the entropy based distributions.

It should be noted that for convenience, the comparison should be directly made on the solutions resulting from the application of LNOPTNET before any alteration to the optimal lengths.

Examination of the solutions, after running LNOPTNET, shows that some links are made of two segments. This means an increase in the number of nodes, with no external demands, of the original system. To maintain the same parameters of the original network and reduce the run time of the reliability tester, an equivalent diameter with its corresponding probability of failure may be sought as follows.

For a link of total Length L consisting of N_x segments connected in series of Length X_i and Diameter D_i, the equivalent diameter D_{eq} that will produce the same head loss resulting from the application of the Hazen-Williams equation to segments of a link in series is given as:

$$D_{eq} = \left[L / \sum_{i=1}^{N_x} X_i / D_i^{4.87}\right]^{1/4.87} \tag{8.2}$$

Likewise, for a link consisting of N_x segments connected in series of Length X_i, of Diameter D_i, and Probability of failure pD_i related to Length X_i, the probability of failure pD_{eq} may be obtained from:

$$pD_{eq} = \sum_{i=1}^{N_x} pD_i \tag{8.3}$$

Strictly, the product of the probabilities of failures should be subtracted from Eq. 8.3. However, since these probabilities are very low (eg 0.000192 for 200 mm), Eq. 8.3 is a reasonable approximation.

The set of commercially available diameters with their probabilities of failure used in this report are presented earlier in Table 7.3.

In addition to the probabilities of pipe failures, the probabilistic distribution of the demands and the duration of pipe failures are required for running the reliability tester. These are the same as those considered for computing the system availability related to Network D. ie, $T_{rep} = 1$ day and the demand rates and their probabilities are given in Chapter 7 by Table 7.4.

Running the reliability tester with networks C, E and F gives the artificial system availabilities presented in Table 8.13. Table 8.14 gives also the optimal system costs for comparison.

		PATH_Q Model	Arbitrary Flow Distribution
Network C	Cost	£426648	£426242
	A_{Net} (%)	99.98	99.97
Network E	Cost	£988845	£1002756
	A_{Net} (%)	99.99	99.98
Network F	Cost	£345513	£254636
	A_{Net} (%)	99.98	99.96

Table 8.13 Optimal Costs and Availabilities of Networks C, E and F

Examination of Table 8.13 shows that designing water networks on the basis of the maximisation of entropy results in more reliable networks than those designed on an arbitrary flow distribution. This point is clearly exemplified by the relatively low values of the availabilities corresponding to the arbitrary distributions of flow, especially for network F (99.96%) compared to the availabilities resulting from the entropic flow distribution (99.98%).

8.5 CONCLUSIONS

From the above discussion, it becomes clear that the optimal design of water networks on the basis of the entropy principle produces more reliable networks. This result has become possible thanks to the reliability tester. The maximum artificial system availability reached by the application of the entropy principle to the numerical examples is 99.99%. This maximum availability obtained by running the Reliability Tester with the sample networks corresponds to a design produced on the basis of the stochastic demand factor of 1.43 that does not cause a hydraulic failure for the full network. This is the reason for the relatively high availabilities obtained for the three numerical examples.

Chapter IX

GENETIC ALGORITHM MODEL

Chapter IX

GENETIC ALGORITHM MODEL

9.1 INTRODUCTION

This Chapter presents an alternative approach to the optimisation of water networks with reliability considerations. Optimisation will be performed via the technique of Genetic Algorithms (GAs). GAs have been promoted as a class of general purpose search strategies that manoeuvre through complex spaces in a near optimal way. GAs have also been shown to be robust search methods which use concepts borrowed from the natural world. The Genetic Algorithms discussed in the present chapter incorporates the Reliability Tester model, previously presented in Chapter 7, to assess nodal and system availabilities.

In the rest of this Chapter, GAs are explored, and examined as an engineering optimisation tool for the minimisation of the cost of water networks under reliability specifications.

9.2 DESCRIPTION OF GENETIC ALGORITHMS

To get an insight into genetic algorithms, a look at what they are and where they come from is undertaken. In doing so, the mechanics of the algorithm are presented, with an attempt to gain some idea of why they work. Then a more rigorous explanation of the underlying search processes is presented. Finally, implementations of the algorithm to numerical applications are presented.

9.2.1 What are Genetic Algorithms?

GAs are a class of stochastic improvement algorithm; they were invented to mimic some of the processes observed in nature: these algorithms solve problems of finding good artificial chromosomes by manipulating the material in the chromosomes blindly. They know nothing about the problems they are solving; the only information they are given is an evaluation of each string they produce, and their only use of that evaluation is to bias the selection of artificial chromosomes so that those with the best evaluations tend to reproduce more often than those with a poor evaluation. In a sense, GAs enforce the survival of the fittest among a population of artificial chromosomes (strings). The algorithms are genetic because the string manipulations employed resemble the mechanics of natural genetics. Every generation, a new set of artificial chromosomes is generated using components (sub-strings) of the fittest of the old generation; an occasional new part is tried for good measure.

Yet, one should not assume that genetic algorithms are a simple random walk through some parameter space; these methods are not coin flipping by a fancy name. GAs can be viewed as parallel search algorithms that efficiently exploit old information to seek in a huge space trial points with above average performance. Indeed, by considering many strings as potential candidate solutions, the risk of getting trapped in a local optimum is greatly reduced.

GAs have been introduced and developed by John Holland (1975) and his students in the Computer and Communications Sciences Department at the University of Michigan. The main goals of their research have been twofold:

1) to abstract and understand, mathematically, the adaptive processes of natural systems,

2) to design software for artificial systems that retain the important mechanisms of natural systems.

This approach has led to important discoveries in both natural and artificial systems science.

The power of these algorithms is derived from a very simple heuristic assumption: that the best solution will be found in regions of the search space containing relatively high performance of good solutions; and that these regions can be identified by judicious and robust sampling of the space. Holland (1975) showed how simple mathematical models of population genetics can efficiently and implicitly make use of this heuristic. GAs implement these models by iteratively manipulating a population of strings using genetic operators(eg **Selection, Crossover** and **Mutation).**

While established as a valid approach to optimisation problems requiring efficient and effective searches, GAs are computationally simple, and powerful. This is so, because they place a minimum of requirements and restrictions on the user prior to engaging the search procedure. The user simply codes the problem as a finite length string, characterizes the objective or objectives (biologists call this objective the fitness function) as a black box, and turn the GA crank. The genetic search then takes over, seeking near-optima through the combined action of its operators.

9.2.2 Overview of the Theory

The fundamental theorem of GAs published by Holland (1975) will be briefly reviewed. More details on the explanation of the theorem can be found elsewhere (Goldberg, 1989). As it is discussed further in the following Chapter, GAs work on populations of strings. The theory is based on the concept of schemata (schema in singular). A schema is a similarity template describing a subset of strings with similarities over certain string positions.

The basic structure processed by GAs is the string. Assume that we have a finite binary string of length L_b (number of bits in the string), and we wish to describe a particular similarity. For instance, consider the two strings St_1 and St_2, each with a length $L_b = 5$:

$$St_1 = 10111$$

$$St_2 = 11100$$

We can see that both of the strings have 1's in the first position. Such similarity can be described by for example introducing a star $*$ in all positions where we are disinterested in the particular bit value. As a consequence, the similarities in the first and the third positions can be described as follows:

$$1_{****}. \quad \text{and} \quad {}_{**}1_{**}.$$

Note that the combined similarity can be described by the string 1_*1_{**}, having 1 at the first and third positions respectively.

These schemata, or similarity templates, apply to not only strings St_1 and St_2 but also describe the subset of strings in each schema. For instance, the schema 1_{****} describes a subset of $2^4 = 16$ strings, each with a 1 in the first position. The particular schema (1_*1_{**}) contains a subset of $2^3 = 8$ strings, each with 1 in the first and the third position.

In general, not all schemata are generated equally. Some are more specific than others. Some have defining positions that span a greater or lesser proportion of the string. The specificity of schema h (its number of fixed positions), is called the order of schema $o(h)$. For example, $o(h = 1_*1_{**}) = 2$ and $o(h = 1_{****}) = 1$.

Another factor used in measuring the quality of a schema is its **defining length** $\delta(h)$, defined as the distance between the outermost defining positions of the schema.

For example, the defining length of any one-bit schema is 0:

$$\delta(h = 1_{****}) = \delta(h = {}_{**}1_{**}) = 0.$$

For the two-order schema, the $\delta(h = 1_*1_{**})$ can be computed by substruction the position indices of the outermost defining positions as:

$$\delta(h = 1_*1_{**}) = 3 - 1 = 2.$$

Using the concepts of order and defining length, the fundamental theorem of GAs, otherwise known as the schema theorem can be written as follows (Goldberg, 1989):

$$m(h,t+1) \geq m(h,t) \cdot \frac{Fit(h)}{A_{Fit}} \left[1 - p_{crossover} \cdot \frac{\delta(h)}{L_b} - p_{mutation} \cdot o(h) \right] \quad (9.1)$$

Where

$m(h,t)$ = number of copies of schema h at time t;

$m(h,t+1)$ = number of copies of schema h at time t + 1;

A_{Fit} = Average fitness of the population;

$p_{crossover}$ = Probability of crossover;

$\delta(h)$ = Defining length of schema h;

L_b = length of string St_i

$p_{mutation}$ = Probability of mutation;

$o(h)$ = Order of schema h;

$Fit(h)$ = Schema average fitness, defined by:

$$Fit(h) = \frac{\sum_{St_i \in h} Fit(St_i)}{m(h,t)} \quad (9.2)$$

The schema average fitness Fit(h) is the average of the fitness values of all strings St_i which currently include the schema h.

Schemata are an interesting notational device for discussing similarities in strings. More than this, they provide the basic means for analysing the performance of GAs.

In Eq. 9.1, $p_{crossover}$ and $p_{mutation}$ refer to the probabilities of applying the genetic operators crossover and mutation respectively. These will be more thoroughly discussed in the following sections.

The factor multiplying the m(h,t) may be thought of as a growth factor. If it is larger than one, the expected number of schemata h, will continue to grow; otherwise, it can do no more than remain constant in number. It is worth pointing out that Eq. 9.1 holds for all schemata contained in the population. In other words, a simple GA processes all schemata in this manner. Highly fit schemata tend to survive because of the factor $Fit(h)/A_{Fit}$. Short definition lengths are also preferred with a high crossover probability, $p_{crossover}$ (in general close to 1). Moreover, due to the fact that mutation probability, $p_{mutation}$ is often quite small; this has a little effect except on schemata of very high order.

In short the schema theorem says that a schema h, is expected to grow in subsequent generations if:

(1) It has above average fitness;

(2) It is relatively short; and

(3) It is of low order.

When all three conditions are met, the schema in question is termed a *building block*. These building blocks are combined and recombined by GAs to seek the best solution.

9.2.3 Genetic Algorithm Essentials

This section investigates a simple genetic algorithm, both its mechanics and why it works. The mechanics of the process are surprisingly simple. The algorithm does nothing more complex than string copying and partial string swapping. The explanation of why it works is much more subtle and powerful. The simplicity of operation and the power of effect are among the main attractions of the GA approach.

Genetic algorithms are derived from a simple model of population genetics based on the following assumptions:

(1) Artificial chromosomes, which can undergo genetic transformations are fixed length strings having a finite number of position values (eg 0/1) at each position;

(2) A population contains a finite number of artificial chromosomes; and

(3) Each population individual has a fitness, or relative ability to survive and produce offspring.

Before going into details, it may help to give a brief overview of how GAs work. During each iteration of the algorithm (a generation), the fitness of each individual in the population is determined and strings are stochastically selected to produce offspring according to their relative fitness. Pairs of successful offspring are chosen to mate and produce the offspring of the next generation. Variation is introduced by the use of the genetic operators: Crossover and Mutation. By application of crossover, each offspring draws part of its genetic material from one parent and part from another. Moreover, new genetic material is occasionally introduced through mutation. The artificial chromosomes which survive will, over time, be those which have proved to be the most fit. In other words, the search is directed towards regions containing strings with above average fitness.

To be a simple GA which gives good practical results in the sense of Goldberg (1989), a procedure must contain the following types of operators:

1 - *Selection*;

2 - *Crossover*;

3 - *Mutation*.

In order to produce a new population, strings from the current population have to follow a certain procedure inspired from the natural world: First, artificial chromosomes are selected from the current population. Second, they are split up, and recombined and finally 'mutated' to form new chromosomes for the following generation.

9.2.3.1 Selection

The first key step in executing a GA is selection. The purpose of this step is to lead the genetic search in a specified direction: regions of high observed average fitness. This concept causes the best chromosomes to proliferate in the future generations and the least fit members to be ruled out.

There are many ways to perform selection effectively. One commonly-used and perhaps the easiest technique is **roulette wheel selection.** The choice of a string in the current population can be obtained by the following procedure:

(a) Compute the total sum of fitnesses of the population strings; call the result population fitness;

(b) Generate j, a random number between 0 and population fitness;

(c) Return the first population member whose fitness, added to the fitnesses of the preceding population strings, is greater than or equal to j.

This procedure is referred to as roulette wheel selection because it can be viewed as allocating pie-shaped slices or segments, on a roulette wheel to population strings, with each slice proportional to the string's fitness. Selection of a particular string from the current population to be a reproduction candidate can be viewed as a spin of the wheel, with the winning string being the one in whose slice the roulette spinner stops. It is worth noting that when using this technique the string fitness values should be positive numbers as they are proportional to the probability of selection.

9.2.3.2 Crossover

While selection represents an elitist process which retains only the most fit strings of a population for mating, it does not in any way improve the quality of any single string in the population. It is the crossover operator that allows the characteristics of the population strings to be altered. Many GA practitioners believe that crossover is the genetic workhorse, a high performance search technique that acts rapidly to combine what is good in the initial population, and that continues to spread good schemata throughout the population as the GA runs. In fact crossover, which causes long jumps in the search space, is the only operator that is thought to distinguish GAs from all others optimisation algorithms.

Now let us examine how the crossover transform is applied. Again several different ways of carrying out this operation are possible. The conventional approach is described for illustration. First two strings are selected and a crossing site, called the Crossover-Point (fixed for both of the strings), is generated randomly. Then, position values are swapped between the two strings following the crossover-point, so that two new offspring arise.

For simplicity in the following example, each of the two strings used has identical elements:

	Parent 1	[0000000]	Parent 2	[1111111]
Separate:		[000 ....]		[... 1111]
		[... 0000]		[111 ....]
Recombine:		[000 ....]		[... 0000]
		[... 1111]		[111 ....]
Result:	Child 1	[0001111]	Child 2	[1110000]

It should be noted that this method is usually referred to as one point crossover.

9.2.3.3 Mutation

If selection and crossover provide much of the innovation of the genetic search, what then is the role of the diversity-generating or mutation operator? Mutation is a necessary component of GAs: in the beginning, mutation safeguards the genetic search process from an early loss of valuable genetic material and after substantial convergence it refines solutions after selection and crossover have narrowed the search.

Usually mutation is performed with a low probability rate, for instance, 0.001: thus, when mutation is applied to a binary string during a run of a standard GA, each bit in the string will have a one in one thousand chance of being randomly replaced. If the mutation rate is too high, information dissolves and the process degenerates into a random search (Eigen, 1987). Once again, there are many mutation operators. For example, in a binary-coded GA, one commonly used operator replaces a 1 with a 0 or vice versa if a probability test is passed.

9.2.3.4 The Algorithm

The structure of an algorithm that can be applied to a wide range of problem domains is shown in the following diagram:

Initialise Population: Randomly generate an initial Population of size N_{pop}
While Not (terminate condition) DO
    Compute the Fitness function of each member of the population;
    For i = 1 To (N_{pop}/2) Do

        Selection: Pick two parents on a roulette wheel basis;
        Crossover: Crossover the parents based on crossover probability to produce two new offspring;
        Mutation: Mutate each offspring based on mutation probability;

    EndFor
END.

9.2.4 Advantages of Genetic Algorithms

In summary the principle attractions of GAs are:

(a) Globality: the main advantage of this stochastic search is its ability to achieve a near global optimum while most search techniques seek for a local optimum solution. This is due basically to: (1) a parallel search; seeking from a population of points instead of a simple point and (2) the fact that a diverse population of solutions is maintained from generation to generation. Sets of solutions will tend to converge on each local optimum, but these will eventually be left as the overall search identifies new solutions in more profitable regions.

(b) Decision Variable independence: GAs require any continuous valued decision variable to be discretised, for the process of mapping onto a binary string. Yet, they handle integer and discrete valued variables efficiently. Most engineering design problems tend to involve discrete choices such as pipe sizes, beam section, which are very suited to binary representation.

(c) Domain independence: the algorithm works on the coding of a problem, ie each decision variable of the problem is represented by, for instance, a sub-string of 0's and 1's, so that it is easy to write a general computer program for solving many different optimisation problems.

(d) Non-Linearity: Many conventional optimisation techniques are based on unrealistic assumptions of linearity, convexity, differentiability etc. None of these are needed by GAs. The only requirement is the ability to compute some quality function (fitness function) which may be highly complicated and non-linear.

(e) Flexibility: GAs do not require that the constraints should be expressed explicitly in terms of design variables.

(f) Robustness: As a consequence of the previous advantages, GAs are inherently robust, they can cope with a large spectrum of problems, they can work with highly non-linear problems and they do it in a very efficient manner (Goldberg et al., 1989).

(g) Parallel Nature: Not only are GAs inherently parallel search techniques but also due to the independence of processing every individual solution in the population, computation can be performed in parallel. This implicit parallelism of GAs makes them the most suitable for design optimisation in a parallel computing environment. Attempts have been made for such implementation (Pettey et al., 1987).

9.2.5 Application of a simple Genetic Algorithm

In the previous sections a simple GA was investigated. Now an application of the algorithm to the minimisation of a simple algebraic function is examined. This example is taken from Bunday and Garside (1987). The problem is to minimise the following function:

Problem GA1:

$$\text{Minimise} \qquad F = 3x_1^2 + 4x_1x_2 + 5x_2^2 \tag{9.3}$$

$$\text{Subject to} \qquad \begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \\ x_1 + x_2 &\geq 4 \end{aligned} \tag{9.4}$$

The anticipated solution is $F^* = 44$ at $x_1^* = 3$ and $x_2^* = 1$. As reported in the literature, to solve an optimisation problem by GAs one should specify 5 components:

1) A string representation of the solution to the problem;

2) A way of generating an initial population of solutions;

3) A fitness function measuring the quality of solutions;

4) Genetic operators that improve solutions during the run of the GAs; and

5) Values of the parameter utilised by the genetic search (population size; probabilities of applying operators etc).

9.2.5.1 String representation of the solution

Since standard GAs work on coded variables, a chromosomal representation of the solution is required. Owing to its simplicity, binary coding is adopted. The problem we face is the determination of the length of a sub-string (No. of bits),

L_s, per decision variable. L_s relies on the type of decision variable involved: continuous, integer or discrete. In general, given a continuous design variable, x_i, which can take any value between a minimum value x_{min} and a maximum value x_{max}, the sub-string length L_s required to a precision of ϵ, may be estimated from the following expression (Goldberg, 1989):

$$2^{L_s} \geq \left[(x_{max} - x_{min})/\epsilon + 1\right] \tag{9.5}$$

For problem GA1, if we suppose that the decision variable x_i, can take any value between 0.0 and 15.0 (inclusive) to a precision of 1.0, then $2^{L_s} \geq 16$. This gives a value of L_s of 4, with the following 16 4-digit combinations of 0 and 1:

1 :	0000
2 :	0001
3 :	0010
4 :	0011
5 :	0100
6 :	0101
7 :	0110
8 :	0111
9 :	1000
10 :	1001
11 :	1010
12 :	1011
13 :	1100
14 :	1101
15 :	1110
16 :	1111

The decoding of the strings will produce the corresponding decimal digits which will then represent real values of the variables. This 'parameter' mapping is within the control of the user. The procedures *extract_parm* which removes a sub-string from a full string, and *map_parm* that maps the unsigned integers to the range $[x_{min}, x_{max}]$ presented in Goldberg (1989) are used in this example.

9.2.5.2 Initial Population

Having coded the decision variables x_1 and x_2 as finite-length strings, the initial population of solutions can be set up. It is a common practice when beginning a genetic search to initialise a population of N_{pop} strings by randomly generating bits with equal probability, p_0 (eg 0.5), for zero and one. Indeed, a random number is generated between 0 and 1 and compared to p_0. If the generated number is greater than or equal to p_0 then the bit value is 1 otherwise 0. If there are N_D decision variables (x_i, i=1,...,N_D), this process is repeated $N_D.L_s.N_{pop}$ times. Table 9.1 shows a population created in this way.

9.2.5.3 Fitness Function

Fitness values of solutions are the only information that GAs exploit to move to high performance space regions. To be more precise, a fitness value is used to guide the selection component to choose the most fit strings for crossover and mutation. A fitness value, which must be positive as required by computation of selection probability, expresses the quality or fitness of a solution.

Originally, GAs were designed to deal with maximisation problems. The common practice used to transform a minimisation problem to maximisation problem is to maximise the negative objective function. This approach is not feasible as mentioned earlier. Thus minimisation problems can be solved for example by using a simple device:

$$Fit = C_f - F \qquad (9.6)$$

Where Fit is the fitness function and C_f is a constant large enough to prevent negative values of fitness. For problem GA1, C_f can be estimated by putting both of the variables to their maximum values, 15, in Eq. 9.3. This gives C_f = 2700.

Another convenient and efficient approach to sort out this problem will be explained, when the optimisation of water networks is addressed.

				Generation 0				
St_i	x_1		x_2		F	Fit		Selection Probability Fit/ΣFit
1	1111	15	1000	8	1475	1225		0.085
2	1010	10	0111	7	825	1875		0.129
3	0000	0	0010	2	20*	0*		0.000
4	1000	8	1110	14	1620	1080		0.075
5	1100	12	0000	0	432	2268		0.157
6	0000	0	1000	8	320	2380		0.164
7	1110	14	0001	1	649	2051		0.142
8	1100	12	1011	11	1565	1135		0.078
9	1101	13	1110	14	2215	485		0.033
10	0100	4	1010	10	708	1992		0.137
	Σ					14491	1	

Table 9.1 Processing Generation 0
*Constraints violation

9.2.5.4 Genetic Operators

The genetic operators used herein to solve problem GA1 are:

(1) The simple roulette wheel selection;

(2) One point crossover; and

(3) Bit mutation, ie replace a 1 with 0 or vice versa.

9.2.5.5 Parameter Values

Hidden behind the conceptual simplicity of GAs, there are a variety of parameters such as population size and probabilities of mutation and crossover. Effective values of these parameters for bit string representation have been intensively studied (De Jong, 1975; Grefenstette, 1986 and Schaffer et al., 1989). Again this problem will be discussed further when the optimisation of water systems is examined. Note that the object of this section is to present a simple illustration of the way in which GAs deal with optimisation problems. For this example the genetic parameters for problem GA1 are:

1. Population size: $N_{pop} = 10$;
2. Probability of Crossover: $p_{crossover} = 0.8$;
3. Probability of mutation: $p_{mutation} = 0.03$;
4. Number of generations = 50;

9.2.5.6 Implementation and Results

The work now to be described was carried out using a PASCAL program, developed by the author from Goldberg's "Simple Genetic Algorithm" (1989).

Table 9.1 shows also the binary sub-strings of x_1 and x_2 and their numerical values, the values of the function F and its corresponding fitness (columns 6 and 7), and finally the probabilities of selection. Note that the maximum fitness (2380) corresponding to the minimum value of the function F of generation 0 (320) has the highest probability of selection (16.4% for St_6) and the worst strings has the lowest probability of selection (0.0% for St_3). A zero in the fitness column indicates that the particular combination of the variables does not satisfy the imposed constraint (Eq. 9.4).

Application of the three genetic operators to the members of generation 0 results in producing new solutions comprising generation 1. The new strings generated and their corresponding function values F and fitnesses are summarised in Table 9.2.

St_i	binary coding	x_1	x_2	F	Fitness
1	01010100	5	4	235	2465
2	10111001	11	9	1164	1536
3	01100100	6	4	284	2416
4	00111101	3	13	1028	1672
5	01010010	5	2	135	2565
6	11000100	12	4	704	1996
7	01011100	5	12	1035	1665
8	11100001	14	1	649	2051
9	01110100	7	4	339	2361
10	01010011	5	3	180	2520

Table 9.2 Results of Generation 1

A complete analysis of the problem by computer is listed in Appendix C. It will be seen that the optimum solution of problem GA1 ($F^* = 44$ at $x_1^* = 3$ and $x_2^* = 1$) has been found at generation 10. It should be noted that only improved solutions are listed in Appendix C.

9.2.6 Summary

The detailed mechanics of a simple genetic algorithm have been presented. GAs operate upon populations of strings. The strings are coded to represent the underlying parameter set. Selection, crossover, and mutation are applied to successive string populations to generate new string populations. The operations performed are simple string copies and partial string swaps, yet the effect is extremely powerful. A simple genetic algorithm has been introduced to deal with optimisation of a simple algebraic function with the aim of illustrating both the detail and power of the method.

Having introduced the basic concepts of GAs in section 9.2, we now move to consider how these might be applied in the context of water network design with reliability issues.

9.3 OVERALL MODEL

In the previous section, it was shown how GAs can be applied to a simple optimisation problem. Now their application to the design of water networks with reliability specifications is considered. Earlier a simple GA incorporating standard forms of genetic operators was presented. In this section, however, improved forms of genetic operators, associated with a new version of GAs are implemented to speed up the convergence of the search.

9.3.1 The problem

Design of water networks under reliability constraints and stochastic demands can be stated verbally as:

GIVEN:

1) A set of demand nodes;
2) A set of links;
3) A set of normal loading (demand) conditions;
4) A set of stochastic demands;
5) A set of commercially available pipe diameters and costs;
6) A set of minimum performance levels for normal loading conditions and stochastic loads.

FIND:

1) Link diameters;
2) The minimum total cost of the network.

SUBJECT TO:

1) Satisfying steady state flow conditions;
2) Satisfying minimum performance levels under normal and stochastic loading conditions;
3) Satisfying a minimum system reliability.

This problem can be viewed as finding the optimal network cost and diameters subject to the technical constraints (steady state flow conditions and minimum nodal pressures) and the constraint on reliability.

Therefore, for stochastic demands, the optimal design of a water network having N_F fixed head nodes, N_J nodes, N_P pipes and N_L loops can now be stated as:

Problem GA2

$$\text{Minimise } \; Cost_{NET} = \sum_{i=1}^{N_P} CostD_i \, L_i \tag{9.7}$$

Subject to

$$\sum_{i \in Pn(k)} \Delta H_i(j) \leq H_0 - H_k^{min}(j) \qquad \forall \; k \in N_J, \forall \; j \in NLoad \tag{9.8a}$$

$$\sum_{i \in Pf(k)} \Delta H_i(j) = b_k \qquad \forall \; k \in (N_F - 1), \forall \; j \in NLoad \tag{9.8b}$$

$$\sum_{i \in Pl(k)} \Delta H_i(j) = 0 \qquad \forall \; k \in N_L, \forall \; j \in NLoad \tag{9.8c}$$

$$Rs \geq Rs_T \tag{9.8d}$$

Where

$Cost_{NET}$ = total network cost;

$CostD_i$ = unit cost of diameter D_i in link i;

L_i = length of link i;

Pn(k) = set of links in the path from the source to node k;

H_0 = original head at the source;

$H_k^{min}(j)$ = minimum required head at node k for loading condition j;

Pf(k) = set of links in path k, k = 1,...,(N_F - 1) associated with known net head loss b_k;

$Pl(k)$ = set of links in loop k, $k = 1,...,N_L$;

Rs = system reliability;

Rs_T = target reliability;

$\Delta H_i(j)$ = head loss in link i corresponding to loading condition j, may be defined using the Hazen-Williams equation by:

$$\Delta H_i(j) = \gamma L_i (Q_{ij}/C_i)^{1.852} D_i^{-4.87} \tag{9.9}$$

Where γ is a constant depending on the units used, Q_{ij} is the flow of link i for loading condition j, C_i is the Hazen-Williams coefficient of diameter D_i in link i.

Eq. 9.8a. expresses the node constraints: nodal heads must be equal to or greater than the required minimum bound. Eqs. 9.8b. and 9.8c. refer to the path constraints: the total head loss along any path between two fixed nodes must equal the difference in head between those nodes. However, Eq. 9.8d. states that the reliability of the system must be greater than a pre-defined value.

We have seen in Chapter 7 that water networks reliability can be efficiently assessed by the simple and computationally fast Reliability Tester. In addition, the reliability tester can be successfully incorporated into an optimisation scheme.

It is worth noting that Eqs. 9.8 can be solved by the reliability tester. The problem then is reduced to finding the least cost design of water networks and the optimal set of diameters given a pre-specified reliability value subject to the technical and reliability constraints. These can by efficiently assessed by the reliability tester.

It is worth noting that problem GA2 does not contain constraints on velocity and therefore the solution of GA2 should be tested for acceptable velocities.

9.3.2 The GA in the Design of Water Networks

It should be noted that the standard Genetic Algorithm, which is referred to by some GA researchers as the **Generational Replacement** (GR) scheme, is the algorithm that is already exemplified through problem GA1. The basis of the GR is to generate all the next population from the current population. In other words, at each generation of the GA, all members of the population are generated. The **Steady State Reproduction** (SSR) is an alternative search strategy. The main difference between the GR and the SSR is that within each generation, only a few members of the population are changed (usually one or two). The standard genetic algorithm (GR) and the alternative search strategy (SSR) are developed, programmed, tested, and compared. The algorithm with the best performance will be adopted for the design of water networks.

So far, discussions of GAs have focused on searching unconstrained objective functions. Typical engineering problems often involve a set of constraints which must be fulfilled. At first, it would appear that constraints pose no particular problem. Indeed, GAs generate a sequence of decision variables to be tested. One simply runs the model, evaluates the objective function, and checks to see if any constraints are violated. If not, the fitness will be the objective function evaluation. If, on the other hand, constraints are violated the solution is infeasible and thus has no fitness. This is the approach adopted in problem GA1. This procedure is fine except that many problems are highly constrained; finding a feasible solution is almost as difficult as finding the best. Furthermore, the foundations of GA theory, however, say that GAs optimise by combining partial information from *all* the population. Therefore, the infeasible solutions should contribute by providing information and not just be thrown away. Consequently, we might want to rate infeasible solutions as well, perhaps degrading their fitness ranking in relation to the degree of constraint violation. This is what is done in penalty methods.

In a penalty method, a constrained problem is transformed to an unconstrained problem by associating a cost or penalty with constraint violations. This cost is included in the objective function evaluation. Interior penalty function approaches (Golberg, 1987) and exterior penalty function approaches (Lin and Hajela, 1992) have been described in the GA literature.

Furthermore, it was outlined in the previous section that the application of GAs to optimisation problems requires 5 components to be specified. These are: (1) a string representation of the solution; (2) a way of seeding the initial population; (3) a fitness function; (4) genetic operators and finally (5) values of the genetic parameters. In the design of water networks, this strategy will be followed.

9.3.2.1 String representation

Pipe sizes may be taken as decision variables in the design of water networks. In such a case, these decision variables are not continuous and must be chosen from the set of the discrete available sizes. With conventional optimisation techniques, if the problem is formulated as non-linear, one common practice is to consider pipe sizes as continuous, and, when the optimal solution is found, diameters are rounded up/down to the nearest commercially available sizes. Most of the time two situations can arise: either the practical solution is infeasible or it is not optimal. However GAs can handle discrete problems efficiently. Pipe sizes may remain discrete but coded in finite length strings. *Hence the final solution obtained after running the GA will be at the same* time optimal and feasible.

9.3.2.1.1 Binary Coding

Usually in engineering problems, the range of variation of the decision variables x_i, $i = 1,...,N_d$, is known. For example, in the design of water networks, the minimum and the maximum values of the diameters can be known. For a single source network, usually the source flow is known and by considering constraints on velocities, the maximum pipe size can be determined. The minimum pipe size can be taken as the minimum available size or the minimum practical size for firefighting (eg 100 mm). Once the minimum and maximum diameters (D_{min} and D_{max}) are known, the number of candidate decision variables, N_d, also becomes known. Each of these N_d discrete variables can be represented in the form of a sub-string. One simple way of doing this is the use of binary coding. In the binary coding representation, the length L_s of the binary sub-string has to be determined. With reference to the first example (problem GA1), the variables were continuous and L_s was determined using Eq. 9.5. In the design of water networks, where the problem is discrete, the length L_s can also be determined by Eq. 9.5 with an accuracy ϵ equal to 1 only (Lin and Hajela, 1992).

Under these conditions, one may proceed by mapping an integer variable, i, between 0 and N_d-1 using binary coding, and then assign each i, i = 0,...,N_d-1, to its corresponding pipe size D_{i+1}. Therefore, the sub-string length can be determined such that:

$$L_s \geq Ln(N_d)/Ln(2) \tag{9.10}$$

For example, suppose that N_d = 32 discrete diameters, from Eq. 9.10, L_s will be equal to 5. Table 9.3 shows the number of individual binary sub-strings for a practical range of sub-string length.

Sub-string length L_s	No. of possibilities
3	8
4	16
5	32
6	64
7	128
8	256
9	512

Table 9.3 String capacity for binary mapping of discrete variables

If the number of decision variables, N_d, were equal to the number of possibilities (ie L_s = $Ln(N_d)/Ln(2)$), a one-to-one correspondence between pipe sizes and the sub-string chromosomes could be readily established. However, in most applications this is not possible, and the excessive binary strings must be assigned in an appropriate manner. Consequently, this problem which may increase running times of GAs, has to be solved in applications where binary coding is involved. Approaches for handling this situation are described in Lin and Hajela, 1992.

9.3.2.1.2 Integer Coding

Another way for representing efficiently pipe diameters within a GA search is the use of **Integer Coding.** In this representation scheme, decision variables are first mapped to an equivalent number of integer variables. Then a one-to-one correspondence between pipe sizes and the integer numbers can be performed. For instance, assume that for a given network, the number of pipes $N_p = 5$ and the number of candidate diameters (in mm) is $N_d = 8$. These are sorted from the minimum size to the maximum size such that:

[50, 75, 80, 100, 125, 150, 175, 200]

The mapping of these pipe sizes into an integer coding can be performed in the following way:

[1, 2, 3, 4, 5, 6, 7, 8]

Examples of solutions encoded in the integer coding are:

$St_1 = 1\ 5\ 7\ 8\ 2$

$St_2 = 3\ 8\ 3\ 5\ 4$

In these strings, the corresponding sets of diameters are [50, 125, 175, 200, 75] and [80, 200, 80, 125, 100] respectively.

It is worth noting that for the integer coding, there are no excessive strings such as those encountered with binary coding. Also one-to-one correspondence gives no bias, whereas binary representation will give bias towards some values. Therefore, the run time may be reduced.

9.3.2.2 Initial Population

For a population of size N_{pop}, the initial set of strings may be generated at random. For binary coding, this may be achieved by tossing a fair coin as described in section 9.2.5.2. For the integer coding, however, the initial population may be created by choosing randomly integer numbers between 1 and N_d. In water distribution design, the decoding procedure which determines the pipe diameters that correspond to the coded strings will depend on the representation scheme used (binary coding or integer coding).

9.3.2.3 Fitness Function

It was seen earlier that the fitness function returns a measure of how good any encoding solution is. In the previous sections, a device of dealing with minimisation problems was also introduced since GAs were originally designed to tackle maximisation problems (see Eq. 9.6). With the GR algorithm, the same approach will be adopted.

Before addressing completely this problem it is worth noting that GAs may be struck by a malady called *Premature Convergence* if **Fitness Scaling** is not considered (Goldberg, 1989). Premature convergence is typically encountered in a population of a small size with a few extraordinary individuals dominating among mediocre individuals during the first stages of a GA. In fact, if left to the standard selection rule ($p^i_{selection}=Fit_i/\Sigma Fit_i$), the best members would take over a significant proportion of the finite population in a single generation, leading to an undesirable premature convergence. Late, during the process of the GA a critical situation can rise when the population average fitness becomes close to the population best fitness: average members will get the same numbers of copies in future generations. In this situation, fitness scaling techniques may also help by giving reproduction chances to the members with a slight edge that are far in excess of the amount by which they are superior (Davis, 1991).

Genetic algorithms practitioners have solved this and related problems by transforming the evaluations of the strings in various ways. The technique used here is the Linear Scaling scheme proposed by Goldberg (1989). This technique requires a linear relationship between the raw fitness and the scaled fitness. The coefficients of this linear relationship are selected such that the average scaled fitness is equal to the raw average fitness and for the raw maximum fitness, the maximum scaled fitness is equal to two times the raw average fitness.

In the beginning of section 9.3.2 it was mentioned that the design of water networks is a constrained problem. However, it should be noted that the topic of constrained function optimisation is of recent interest in GA research (Richardson et al., 1989). As mentioned earlier, it was reported in the GA literature that the ways utilised to transform a constrained problem into an unconstrained one were the use of penalty function approaches; but according to Rajeev and Krishnamoorthy (1992), the standard penalty approaches are ideally suited for sequential searches but may not be appropriate for GAs which process in parallel using population of points in the search space. Richardson et al. (1989) stressed that penalties, which are functions of the distance from feasibility are more successful. A formulation based on the violation of the network availability is proposed.

Consider a solution represented by a coded string St. The decoding of St gives the candidate diameters D_j, $j = 1,...,N_p$, extracted from the N_d available diameters. Having these diameters, the network cost, $Cost_{NET}(St)$, can be computed by Eq. 9.7.

Application of the reliability tester to the solution obtained results in the assessment of both nodal and system availabilities. Given a target (specified) system availability, which is a reasonable index of a system's general performance, that the system is desired to have, two situations can arise:

(1) the computed system availability is greater than or equal to the target availability:

$$A_{Net} \geq As_T \tag{9.11}$$

In this case, the solution may be kept, since the conditions on availability are fulfilled.

(2) the computed system availability is less than the target availability:

$$A_{Net} < As_T \tag{9.12}$$

In this condition, the quality of this solution depends on how far the computed availability is from the target availability.

Therefore, the total cost of solution St, $Cost_{St}$, may be calculated as the sum of the network cost, $Cost_{NET}(St)$, and the penalty cost, $Cost_{PNLT}(St)$, for violation of the reliability requirement such that:

$$Cost_{St} = Cost_{NET}(St) + Cost_{PNLT}(St) \tag{9.13}$$

9.3.2.3.1 Network Cost Function

With reference to Eq. 9.11 the reliability requirement is fulfilled, that is, there is no penalty cost, and, the solution cost is only the network cost that can be calculated using Eq. 9.7:

$$Cost_{St} = Cost_{NET}(St) = \sum_{i=1}^{N_P} CostD_i \, L_i \tag{9.14}$$

9.3.2.3.2 Penalty Cost Function

If there is a violation of the reliability constraint, a penalty cost has to be added to the network cost to represent this violation. Various functions were tried, with the following being that adopted:

$$Cost_{PNLT}(St) = Cost_{NET}(St) \cdot r \cdot \left(1 - \frac{A_{Net}}{As_T}\right) \quad (for\ A_{Net} \leq As_T) \tag{9.15}$$

Where

A_{Net} = computed system availability;

As_T = target availability;

r = penalty coefficient.

Eq. 9.15 states that for any value of the penalty coefficient r, the penalty cost becomes practically zero when A_{Net} tends to As_T. On the other hand, as A_{Net} decreases, ie for low system availabilities, the penalty cost, $Cost_{PNLT}$, and the total system cost $Cost_{St}$ increase. Having determined the total cost of a solution St, the corresponding fitness has now to be evaluated.

9.3.2.3.3 Raw Fitness

Since the design of water networks is a minimisation problem, the raw fitness should be subtracted from a constant (C_f), large enough, in order to prevent negative values of fitness. For a solution St, the raw fitness, Fit_{St}, is given by:

$$Fit_{St} = C_f - Cost_{St} \tag{9.16}$$

C_f may be taken as the network cost (given by Eq. 9.14) of a network which is assumed to be entirely consisting of the maximum pipe diameter allowed (D_{max}).

9.3.2.3.4 Scaled Fitness

According to the linear scaling technique (Goldberg, 1989), the scaled fitness, Fit'_{St}, for a solution St may be assessed as:

$$Fit'_{St} = a_f \cdot Fit_{St} + b_f \tag{9.17}$$

Where a_f and b_f are the linear coefficients linking the raw and scaled fitnesses that may be obtained as explained earlier in sub-section 9.3.2.3.

9.3.2.4 Improved Operators For Standard GAs

The genetic operators used in the optimal design of water networks are the same as those presented earlier to solve problem GA1. However improved forms of these genetic operators are selected that have been reported to speed up the convergence of the standard genetic search.

9.3.2.4.1 Selection

Different ways are possible to implement the selection process. Earlier, selection was performed via the simple roulette wheel approach with the purpose of providing bias in the population to provide more fit members and to rid the population of less fit members. In the design of water networks, this objective can be reached by an alternative approach for standard GAs, that has been reported to outperform the conventional roulette wheel selection method (Goldberg, 1989): the **Remainder Stochastic Sampling Without Replacement** method. In her dissertation, Brindle (1981) has confirmed the inferiority of roulette wheel selection observed earlier by De Jong (1975) compared to the remainder stochastic sampling without replacement. Therefore, the remainder stochastic sampling without replacement scheme will be adopted for the selection process within the GR algorithm.

Like the roulette wheel selection, the remainder stochastic sampling procedure starts by computing the number of copies ($Ncopies_{St} = N_{Pop} \cdot Fit_{St}/\Sigma Fit_{St}$) that a string St may have according to its probability of selection. In the roulette wheel selection the number of copies, being a real number, is rounded to the nearest integer value. However, for the remainder stochastic method, the fractional parts of the expected number of copies are treated as probabilities and if the expected number is greater than one, such a string will receive for certain one copy and another with a probability equal to the fractional part of the expected number. For instance, assume that the expected number of copies of a string St turns to have a value of 1.6. This string will definitely receive a single copy and another with probability 0.6.

The Pascal code that implements the remainder stochastic sampling without replacement selection used here is that published by Goldberg (1989).

9.3.2.4.2 Crossover

The second modification suggested for the design of water networks is implemented through the use of **Two Crossover Points.** This has been reported to improve the performance of the standard genetic search (Booker, 1987; Syswerda, 1989). In 1982, investigations by Booker demonstrated and showed that a two crossover points operator can indeed improve performance of GAs. This result was confirmed by optimising 10 times a set of 5 test functions (Booker, 1987). Therefore the two crossover points system is adopted.

To effect crossover, a set of crossover parameters are generated randomly. These consist of a pair of strings and two crossing sites. The crossover is carried out by swapping the set of bits between the two cut points. A Pascal routine called *Two_Point_Crossover* that implements this process has being developed.

9.3.2.4.3 Mutation

If binary coding is used, the same mutation operator utilised in the illustrative example (problem GA1) may be adopted for the optimisation of water networks. However, with integer coding, the bit mutation scheme cannot be applied. A new mutation operator is sought.

For binary coding, mutation was performed by replacing a 1 with a 0 or vice versa. In the same manner, for integer coding, mutation of an integer in a given string may be achieved by choosing at random any integer among the N_d integers allowed except that integer. For example, assume that for a given network, the number of pipes $N_p = 5$ and the number of candidate diameters is $N_d = 8$ (1, 2, 3, 4, 5, 6, 7, 8). For a string St_1 (St_1 = 1 5 7 8 2), assume that the probability test, to apply mutation, for the third position (7) has passed, a mutation of the integer 7 may take any value from 1 to 8 except 7 (eg 3). The result is a new string:

$St_1' = 1\ 5\ 3\ 8\ 2$

Algorithm

```
Given a string St
FOR any integer I  in St  DO
      IF (the probability test of mutation is passed) THEN
            REPEAT
               Generate an integer J  among the Nd integers
            UNTIL J <> I
      ENDIF
END.
```

9.3.2.5 Parameter Values

With the standard GAs, it has long been acknowledged that the parameters controlling GAs can have a significant impact on their performance. Moreover, GAs theory gives little guidance for their proper choice. Basically three works have been published in this area. De Jong (1975) performed several computational experiments, on five minimisation functions, to try to gain some insight into the influence of population size, the probability rates of crossover and mutation operators, and a number of other parameters, on the efficacy of genetic search. Two performance measures were designed for this purpose: the on-line performance and the off-line performance. The on-line performance is simply the mean of all function evaluations up to a given number of trials while the off-line performance is the average best fitness of a population up to the given number of trials. Empirical results produced a set of numerical parameters that was generally found to yield the best on-line and off-line performances: N_{pop} = 50-100, $p_{crossover} = 0.60$ and $p_{mutation} = 0.001$.

Grefenstette (1986) has suggested a more robust approach for the optimal selection of these parameters. Indeed, he has developed a **Meta-Genetic Algorithm** that takes not only the design variables as chromosomal representation but also values of the desired parameters (N_{pop}, $p_{crossover}$ and $p_{mutation}$). Applications of these algorithms to the minimisation functions used in the above study by De Jong, to generate the best on-line performance, give the following recommended parameters: $N_{pop} = 30$, $p_{crossover} = 0.95$ and $p_{mutation} = 0.01$.

Grefenstette's combination of parameter values, which recommended a smaller population size and much higher rates of applying the genetic operators than did De Jong, have been proven useful across a variety of problem domains (Davis, 1989).

More recently Grefenstette's results were reinforced by the work of Schaffer et al. (1989) that has consumed more than 12 months of CPU time (1.5 CPU years on Sun 3 and VAX machines). They have used the De Jong test functions and some additional problems (five other functions) which were more complicated and multimodal. They were able to show that robust parameter settings found by their search indicated that good on-line performance can be expected with: $N_{pop} = 20\text{-}30$, $p_{crossover} = 0.75\text{-}0.95$ and $p_{mutation} = 0.005\text{-}0.01$.

It becomes clear that for standard GAs, good parameter values may be taken from the above fairly recent works in this area, since there is some supporting evidence that the reported parameter sets are function independent. For these reasons, the values of the parameters applied for designing water networks are:

$N_{pop} = 30;$

$p_{crossover} = 0.95;$

$p_{mutation} = 0.01.$

By adopting this selection, one can avoid any exploratory experiments which would inevitably increase the overheads of the implementation.

9.3.3 The Standard Genetic Algorithm Adopted

For water network design, the algorithm used in problem GA1 can also be applied. However, the Linear fitness scaling technique and the improved genetic operators, for both binary and integer codings, discussed earlier are added to the algorithm. Therefore, solution of problem GA2 may be reached by the following scheme:

<u>Algorithm</u>

Initialise Population: Randomly generate an initial Population of size N_{pop}. Each member of the population consists of Np candidate pipe sizes.

WHILE NOT (terminate condition) DO

Compute the Fitness function of each member of the population based on running the Reliability Tester which gives the system availability. Having the system availability, the fitness function can be determined on the basis of network cost ($Cost_{NET}$) only if there is no violation of the availability specification or on both network cost ($Cost_{NET}$) and penalty cost ($Cost_{PNLT}$) for violation of the availability specification.

Compute the Scaled Fitness of each member of the population by applying Linear Scaling;

FOR i = 1 To (N_{pop}/2) DO

Selection by Stochastic Remainder Method: Pick two parents

Crossover (One-point/Two-point Crossover): Crossover the parents based on crossover probability to produce two new offspring;

Mutation (Bit/Integer Mutation): Mutate each offspring based on mutation probability;

ENDFOR

ENDWHILE.

This algorithm which the author calls **OPTNET1** has been developed and written in Pascal. OPTNET1 calls the reliability tester for computing the system availability of a candidate solution, which is essential for the determination of the raw fitness of the candidate solution. If the reliability test is passed, the raw fitness is reduced to the network cost given by Eq. 9.14. Otherwise, the raw fitness corresponds to the total cost (see Eq. 1.15) including the penalty cost. OPTNET1 is implemented and tested in the following section.

Since OPTNET1 includes the reliability tester model, the probabilistic demand factors, the repair time and the probabilities for pipe failures should be specified. For the optimisation of water networks in this Chapter, the same probabilistic demand factors (see Table 7.4) and repair time (T_{rep} = 1 day) used in Chapter 7 will also be utilised. For unit costs of pipe diameters, the cost equation taken from the literature (Tanyimboh and Templeman, 1993) is also used here.

9.3.4 Example: Network G

Under this heading, the standard genetic algorithm is applied to the optimisation of a water system: Network G. This system has been deliberately selected to be a very small size (one loop) to illustrate rapidly the application of the GA search to the design of water networks.

The layout of Network G ($N_F = 1$, $N_J = 4$, $N_P = 5$, and $N_L = 1$) is shown in Fig. 9.1. Fig 9.1 shows also the network demands. Ground levels are equal to 140m for all nodes, which are supplied by one source with a pressure head at 200m. Link lengths are set to 1000m for all pipes. Minimum acceptable pressure and Hazen-Williams coefficients are assumed to be 20m and 130 respectively for all the system.

Application of OPTNET1 to Network G requires the determination of the N_d candidate diameters which, for simplicity, may be held constant for all links. As mentioned in subsection 9.3.2.1, N_d may be determined on the basis of selecting the minimum and maximum pipe sizes for Network G. The minimum diameter may be taken as D_{min} = 100 mm for reasons of firefighting. Since

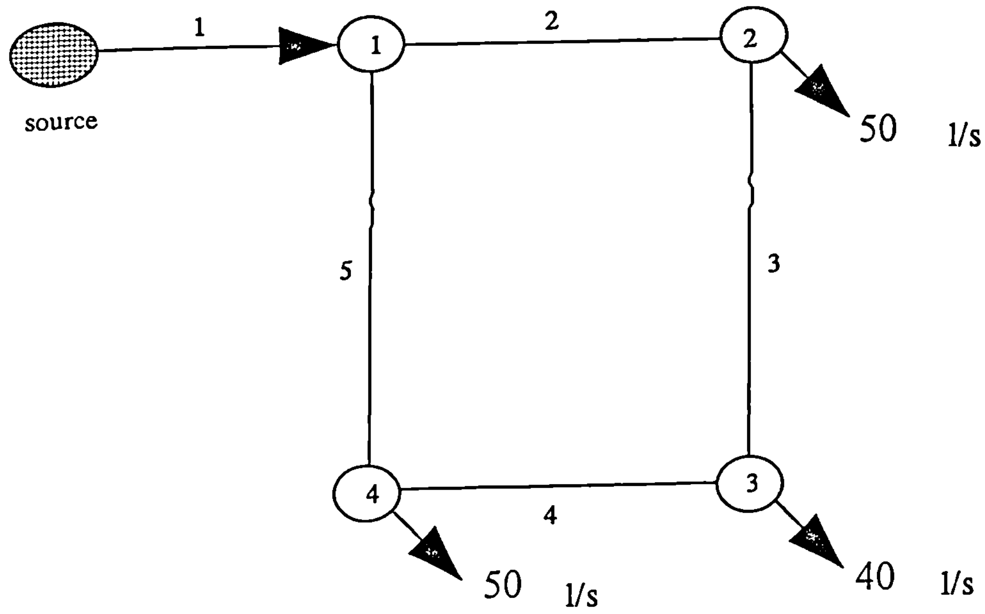


Figure 9.1 Network G

Network G is a single-source system, the total network demand which is known (140 l/s) and a minimum acceptable velocity (say $V_{min} = 0.5$ m/s) may be used to determine the maximum pipe size. According to Eq. 1.1, D_{max} is given by:

$$D_{max} = (4/\pi * 0.14/0.5)^{1/2}$$

$$D_{max} \cong 600\text{mm}.$$

Therefore, the number of candidate diameters for Network G is $N_d = 16$ (see Table 7.3). Note that N_d is a large value for this small problem. However, the objective of this first example is to show rapidly how to apply the standard genetic search to the design of water systems.

For binary coding, with the availability of N_d, the length L_s of the binary sub-string representing each diameter may be determined by Eq. 9.10. This gives:

$$L_s \geq \text{Ln}(16)/\text{Ln}(2)$$

$$L_s \geq 4.$$

Consequently, each member of the population comprises $L_s * N_p = 4 * 5 = 20$ bits. However, with integer coding, as seen earlier, each member of the population will only have N_p (5) integer numbers. For the purpose of comparison between binary and integer codings, the maximum number of generations has been set to 30 generations which gives for network G: $30 * (N_{pop} + 1) = 930$ fitness evaluations.

Another important parameter that must be specified before executing OPTNET1 is the *target availability* value that is supplied by the user. It was mentioned in Chapter 7 that the system availability defined as the average of nodal availabilities weighted by the demands, to take into account the

St_i	Diameters (mm)	Total Cost of Strings (£)
1 :	125 600 500 450 375	17109471.6633*
2 :	125 450 525 600 275	16493302.4650*
3 :	200 350 150 450 450	9308296.7941*
4 :	175 450 100 100 500	8401445.1355*
5 :	125 525 175 300 300	8092579.8076*
6 :	225 300 500 350 350	7867905.7778*
7 :	150 525 100 125 300	6769642.9146*
8 :	175 500 250 225 125	6433307.4651*
9 :	225 200 525 275 250	6216649.7761*
10 :	600 100 250 150 250	5709367.7727*
11 :	100 500 125 175 175	5387841.5871*
12 :	225 125 125 450 250	5139235.1378*
13 :	250 100 275 450 275	5023229.5686*
14 :	200 125 300 225 350	4392263.4949*
15 :	450 100 150 100 400	4238000.7686*
16 :	450 100 100 150 375	3959091.4877*
17 :	225 275 100 175 400	3679647.2474*
18 :	175 225 125 375 125	3543686.0471*
19 :	225 100 350 150 175	3226371.2627*
20 :	525 375 150 450 500	589610.6002
21 :	600 500 300 375 175	583898.6687
22 :	500 250 350 600 175	553118.0673
23 :	375 400 400 100 600	552821.6868
24 :	275 275 150 525 350	544614.9679*
25 :	300 200 500 300 275	506723.6312*
26 :	300 500 150 300 525	471784.2155
27 :	500 175 175 450 350	402833.7197
28 :	400 250 100 275 525	368014.9126
29 :	275 400 175 275 250	304127.9184*
30 :	275 350 100 200 350	237902.1182*

Table 9.4 Processing the Initial Population - Binary Coding

(* Violated solutions)

distribution of the demands, is a reasonable assessment of a system's general performance. The problem then is to specify a target availability (As_T), that the designers wish to reach, and apply OPTNET1 to search for economical solutions that do not violate this availability issue. For Networks considered herein, this value, which depends on the designers and planners of water networks, has been fixed at 99.95 %.

OPTNET1 was first run three times with Network G using **binary coding**. As mentioned earlier, just 30 generations have been considered for each trial. For bigger networks however, larger numbers of generations may be required.

In Fig. 9.2 results of the three independent trials, using different starting points, are displayed. Some details of the best run are summarised in Tables 9.4 and 9.5. Table 9.4 gives (1) the member numbers, (2) the diameters of the initial population and (3) the total network cost of each population member. It should be noted that OPTNET1 includes a routine that sorts the population costs in a decreasing order and hence, the minimum population cost is that having the biggest rank (N_{pop}). Table 9.5 presents the result of the last generation.

It should be noted that for Network G, many values for the penalty coefficient r (1, 5, 10, 15, 20, 25, 30, 35 and 40) have been tried. The penalty coefficient r = 30, which has been held constant throughout the runs, has been found suitable and allowed the search to include some infeasible solutions during the genetic process as previously outlined (see Section 9.3.2). Reduction of the penalty factor makes the infeasible solutions more prominent in the search. Increasing the penalty factor will exclude infeasible solutions from the search.

Fig. 9.2 shows the best-of-generation network costs of each generation as the solution proceeds for the three independent runs. At first, performance is poor, but through the action of the genetic operators, better and better strings are formed. The best solution has been found in the 3rd trial in which the total network cost corresponds to the network cost only with no violation of the availability constraint. This optimal solution, which has been found at generation 22 of the best run, is as follows:

<u>Diameters</u>: (mm) **350, 275, 200, 100** and **225**.

<u>Cost</u>: **£160638.**

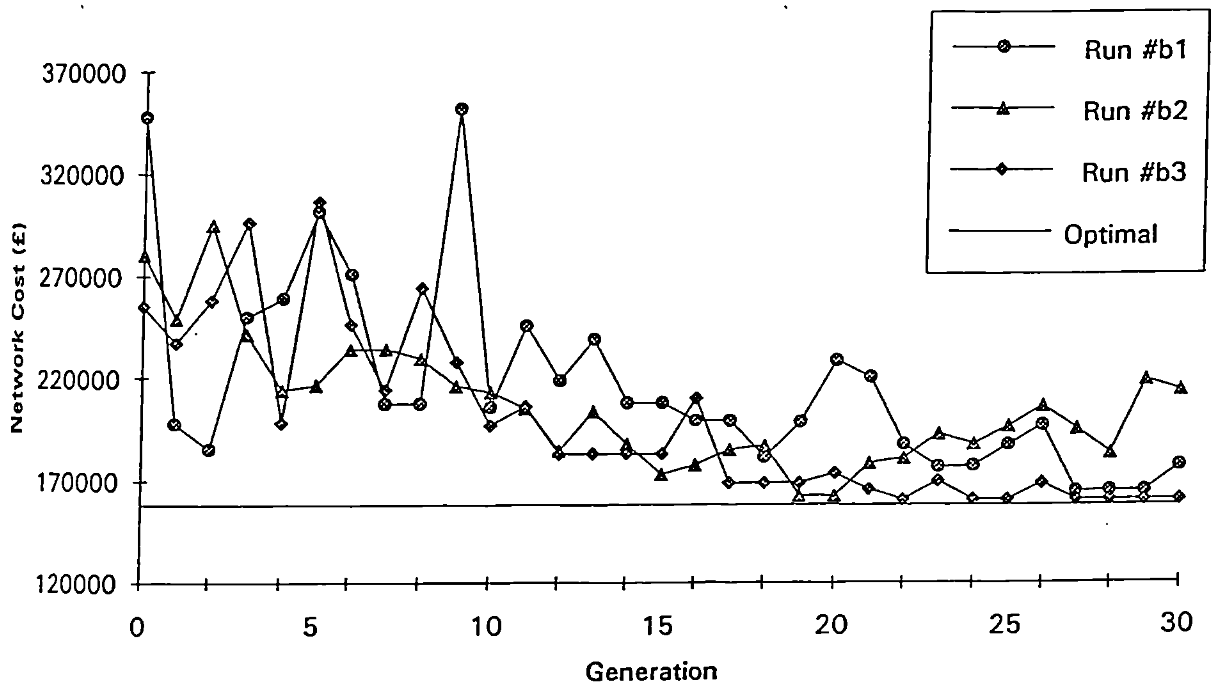


Figure 9.2 Best-of-generation Network Costs for the hree Runs

Application of OPTNET1 to Network G - Binary coding

St_i	Diameters (mm)	Total Cost of Strings (£)
1 :	150 250 125 275 275	3352382.4265*
2 :	350 275 100 125 375	1344067.9019*
3 :	300 275 125 125 350	518906.2021*
4 :	500 275 100 450 400	446944.3649
5 :	450 275 150 500 250	385336.8454
6 :	450 275 150 500 250	385336.8454
7 :	350 225 200 500 375	372452.9583
8 :	500 150 225 225 250	351548.0446*
9 :	350 250 150 175 500	298476.6610
10 :	350 275 450 225 225	295652.6721
11 :	350 375 275 275 275	279767.2461
12 :	375 275 200 225 400	269913.8826
13 :	350 275 275 225 375	264245.3378
14 :	350 275 250 225 375	255941.7323
15 :	350 275 225 225 375	248723.4294
16 :	400 275 125 175 375	245760.7813
17 :	450 225 200 225 275	242119.1061
18 :	350 225 250 125 275	236469.4520*
19 :	350 175 250 225 350	216010.7905
20 :	350 175 250 225 350	216010.7905
21 :	400 225 200 225 275	209512.2809
22 :	375 275 225 225 225	201367.7016
23 :	350 275 125 275 275	200397.8648
24 :	350 175 250 275 275	199698.8794
25 :	350 175 250 275 275	199698.8794
26 :	350 275 200 225 275	197665.8980
27 :	350 275 100 275 250	189556.1653
28 :	300 250 225 275 275	188659.3418
29 :	350 275 225 125 275	184875.9565
30 :	300 275 200 225 275	175263.2008

Table 9.5 Processing the Final Population - Binary Coding

(* Violated solutions)

The first diameter (350) corresponds to the first link in the network, the second diameter (275) to link number 2 and so on.

OPTNET1 was executed a further three times with Network G, but this time using **integer coding**. Fig. 9.3 shows the results related to the three runs.

From Fig. 9.3, it can be seen that the best optimal solution (£159553) has been found in the third run at generation 27. Fig. 9.3 shows also that solution costs obtained using integer coding are slightly better than those corresponding to the use of binary coding. Indeed, with respect to the best solutions found using both binary and integer codings, the best solution has been found using integer coding. This solution has the following set of optimal diameters and cost:

Diameters: (mm) 350, 200, 100, 250 and 250.

Cost: £159553.

	Run Number	Best Cost (£)	Difference from Global (%)
Binary coding			
	1	164603	4.03
	2	162253	2.54
	3	160638	1.52
Integer coding			
	1	163176	3.12
	2	161351	1.97
	3	159553	0.83

Table 9.6 Run Comparisons

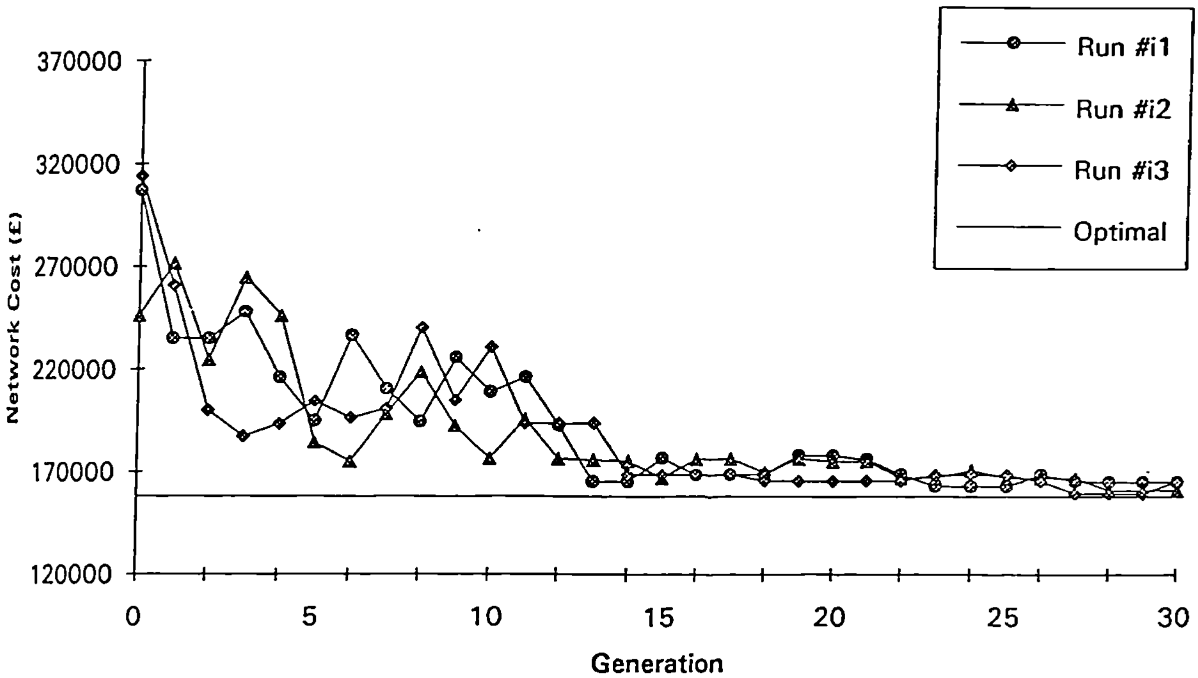


Figure 9.3 Best-of-generation Network Costs for the three Runs

Application of OPTNET1 to Network G - Integer coding

Examination of figures 9.2 and 9.3 shows that with the use of integer coding, a 'real' convergence is observed for network G from generation 25 to generation 30 for the three runs. Moreover, for Network G the best solution found using integer coding is less economical (0.7 %) than that obtained with binary coding. This result has been confirmed by running OPTNET1 (for binary and integer codings) with network G many times.

Each run for network G has required approximately 5 min on a 486 IBM compatible PC.

9.3.5 Enumeration

For the purpose of highlighting the performance of OPTNET1, network G was sized using a complete enumeration scheme. A simple program called GLOBAL, which incorporates the Reliability Tester and a routine for computing the total network cost (penalty cost is included), was written for this purpose. GLOBAL performs all combinations of the 16 diameters allowed for each link of network G. For each combination, the Reliability Tester is applied, and its corresponding total cost is calculated. GLOBAL retains the best solution found so far. It should be noted that for network G the total number of solutions examined by enumeration is 16^5 (1 048 576 solutions). GLOBAL was run using the same computer machine (486 IBM compatible PC) and has required an extensive run time of approximately 4 days. The global solution found is:

<u>Diameters</u>: (mm) **350, 225, 150, 200** and **250.**

<u>Cost</u>: **£158233.**

9.3.6 Conclusions

As seen above, the global solution for network G has a cost of £158233. Examining the results presented in Table 9.6 shows that OPTNET1 finds solutions near to the global solution (increase in cost of 0.8% - 1.5% for the best runs). The difference between the global solution and the best solutions related to each trial may be further improved by increasing the number of generations allowed. This number has been set deliberately to 30 for the purpose of comparison between the standard genetic algorithm (OPTNET1) and the steady state genetic search (OPTNET2) which is presented in the following section.

Irrespective of whichever coding is used, one finishes with the conclusion that genetic algorithms are very effective techniques in finding near-optimal solutions, as solutions within 0.8-1.5% of global optimum were found by examining only a small proportion (0.09%) of the possible designs for an insignificant run time (5 min) compared to that (4 days) required by complete enumeration.

Having applied the standard genetic algorithm to the design of water network G, we move now to the application of the steady state genetic algorithm to the design of the same network and network F which was presented previously in chapter 8.

9.4 THE STEADY STATE GENETIC ALGORITHM

The performance of the standard genetic algorithm may again be improved by applying the *Elitist* strategy. The main idea of the elitist strategy is to copy the best member of each generation into succeeding generations. It will be seen later that this strategy is built into the steady state genetic algorithm.

In this section, the focus will be on the difference between the standard genetic algorithms and the Steady State Genetic Algorithms. Specifically, these include overlapping generations, partial replacement and independent genetic operators. Then, the SSGAs will be programmed in Pascal, implemented and applied to Networks C and F.

The steady state genetic search is similar to the GA procedures normally used in classifier systems. Holland (1975) was the first to describe an algorithm with most of the features of SSGAs and later these were used in other works (Goldberg, 1983; Whitley and Kauth, 1988; Syswerda, 1989). The main difference between SSGAs and the standard GAs is the use of overlapping generations, ie, instead of replacing the current population by a new one, only a few members (typically one or two) of the population are changed. Among the advantages of this algorithm, there are:

(1) Good members of the population are not destroyed and float to the top of the list where their genetic material is preserved.

(2) Poor members sink to the bottom where they are more likely to be deleted, but can still be parents if they are lucky. These characteristics provide an automatic elitism to protect good members of the population.

It is worth noting that the technique that will be used for the design of water networks, which is called *Steady State without Duplicates*, has been used with great success by a number of genetic algorithm practitioners (Whitley and Kauth, 1988; Syswerda, 1989; Davis, 1991). The purpose of the non-duplication is to discard from the population offspring that are duplicates of current solutions. Consequently, each string within the population will be unique. Moreover, the steady state without duplication algorithm may be further improved, as found herein, by insertion of generated members that are better than the worst member in the current population.

The same strategy that was followed in section 9.2 for the application of GAs to the optimisation of water networks will be adopted. The string representation and the way of seeding the initial population will not change. However, the technique of fitness function, the way the genetic operators are applied and finally the values of the genetic parameters will be discussed.

9.4.1 The Scaled Fitness

Recently, Davis (1991) suggested the Linear Normalization technique as a surrogate fitness scaling scheme which requires a sorting of the raw fitnesses, followed by generation of scaled fitnesses that begin with a constant value and decrease linearly. Parameters of this technique are the constant value and the rate of decrement.

To test the performance of each of the fitness scaling techniques outlined, (Goldberg's Linear scaling and Davis Linear Normalization) critical fitness values presented in Table 9.7 are used. These are taken from Davis (1991) for two reasons. First, because they exemplify a quick convergence that may occur during the execution of GAs if scaling techniques are not used. Indeed, the existence of a super member (200) will probably eliminate all its competitors and dominate the population in one or two generations. Second, they reflect a situation, especially encountered at the end of GAs runs, when the

Raw Fitness	200.00	9.00	8.00	8.00	4.00	1.00
Linear Scaling	76.67	31.38	31.38	31.14	30.19	29.48
Linear Normalization	100.00	80.00	60.00	40.00	20.00	1.00

TABLE 9.7 Examples of Fitness Scaling Techniques

fitnesses of several members become close (eg 9, 8, 8). If this happens, it is desirable to increase the pressure of the selection to allow the fittest solutions to have more copies than the others. Examination of the result of the two techniques shows that the linear scaling does no better than having the raw fitnesses: the super member and the closeness of the other members still persist. However, with the Linear Normalization, the super member will be selected, but not so much that it will quickly dominate the population. Moreover, the Linear Normalization heightens the competition in the close race. Therefore, for the reasons outlined above, the Linear Normalization will be used in the steady state Genetic search.

In the design of water networks, if the raw fitness corresponds to the total cost of a solution St (see Eq. 9.13), sorting the raw fitnesses into an ascending numerical order permits a correspondence between the minimum raw fitness and the best (cheapest) solution. On the other hand, the maximum raw fitness will be associated with the worst (most expensive) solution. When these raw fitnesses are sorted, the scaled fitness may be assessed as:

$$\mathrm{Fit'}_{St} = K_{Scal} \cdot (N_{pop} + 1 - \mathrm{Rank}_{St}) \tag{9.18}$$

Where

N_{pop} = size of population;

$\mathrm{Fit'}_{St}$ = scaled fitness;

K_{Scal} = constant used in the Linear Normalization, eg 10;

Rank_{St} = rank of the solution St; $\mathrm{Rank}_{St} = 1, \ldots, N_{pop}$.

Eq. 9.18 states that the first solution (rank = 1) will receive the highest scaled fitness and the last solution (rank = N_{pop}) will have the lowest scaled fitness. Values of the scaled fitness vary linearly between these bounds. Thus, the transformation of a minimisation problem to a maximisation problem required by GAs has been achieved without requiring a large constant as suggested earlier.

9.4.2 Genetic Operators

Basically, the three genetic operators used in the generational replacement search may also be applied to the steady state genetic algorithms. However, the operators that have been reported to enhance the performance of the STGAs will be adopted. Since STGAs worked individual by individual there is no need for any sophisticated selection operator and the simple and quick roulette wheel process may fill this gap.

Syswerda (1989) has made a comparison of three crossover operators using the STGAs. These are one-point crossover, two-point crossover and uniform crossover. By treating six different optimisation problems, he has found that two-point crossover was consistently better than one point crossover and the uniform crossover, generally works better than either. Therefore, the uniform crossover will be used.

In brief, to effect the uniform crossover, two parents (St_1 and St_2) are selected and a crossover mask, which is a string having the same length (eg $L_s = 8$) as the population member, has to be generated randomly. If the crossover mask is made up of '0' and '1', this means that the first offspring will take the same genes of the first parent having at their positions '0' and the same genes from the second parent having at their positions '1', and, vice versa for the second offspring. An example of the uniform crossover applied to the integer coding is shown as follow:

St_1	1 5 4 7 9 2 1 8	
St_2	9 3 8 2 3 4 5 7	
Mask	1 0 0 1 1 0 1 0	Yields:
St_1'	9 5 4 2 3 2 5 8	
St_2'	1 3 8 7 9 4 1 7	

Having selected the crossover operator, the bit mutation or the integer mutation used in the above section are suitable and may be also used.

It was mentioned that in the standard GAs, selection is performed first and then, the selected members undergo crossover and mutation. It is common to apply a combined operator in which mutation is embedded with crossover (Goldberg, 1989). With the aim of improving the performance the STGAs, Davis (1991) proposed to separate crossover and mutation, and, one or the other of these operators will be applied during any iteration of the genetic search. He proposed that the choice of any operator may be performed by the roulette wheel selection. These modifications are adopted in the present work.

9.4.3 Parameter Values

Most genetic implementations keep the crossover and the mutation rates fixed over the course of a run. Recent work by Booker (1987) suggests that varying the crossover rate may be beneficial. Davis (1991) has confirmed this result by his experimental work. Indeed, Davis proposed to alter both crossover and mutation rates as the run is in progress in order to gain improvements in the genetic search. This was achieved by specifying a minimum and maximum values of each operator. At the beginning of a run these rates are set to their minimum and then will be changed gradually over the run and finally set to their maximum when the run is completed. A full description of how to produce these values efficiently and effectively is given in Davis (1991). These values that are also used here are:

$$p_{crossover} = 0.70 - 0.50;$$

$$p_{mutation} = 0.30 - 0.50.$$

For the size of the population, the same population size used in the generational replacement algorithm, applied to water network design, has been adopted:

$$N_{pop} = 30;$$

9.4.4 The Steady State Genetic Algorithm Adopted

Algorithm

Initialise Population: Randomly generate an initial Population of size $N_{pop.}$
Each member of the population consists of Np candidate pipe sizes.
Compute the Fitness function of each member of the population based
on running the Reliability Tester which gives the system availability.
Having the system availability, the fitness function can be determined
on the basis of network cost ($Cost_{NET}$) only if there is no violation
of the availability specification or on both network cost ($Cost_{NET}$) and
penalty cost ($Cost_{PNLT}$) for violation of the availability specification.
Compute the Scaled Fitness of each member of the population by
applying the Linear Normalization Scaling technique;
WHILE NOT (Terminate Condition) DO
FOR I = 1 To 2 DO
REPEAT
Selection by Roulette wheel technique: Pick two parents
Selection by Roulette wheel technique: Pick one operator
IF Operator = Crossover then Uniform Crossover: Crossover
the parents based on the crossover probability
to produce two new offspring;
IF Operator = Mutation Mutate each offspring based on
mutation probability;
Duplication Condition: the first offspring is taken and its
genes (diameters) are compared to the other genes
of each population member. In the event of being
different to all the members it will be retained.
Otherwise, the same procedure is applied to the second
offspring. If each offspring has an identical copy within
the population the process restarts and continues until
the continues until the condition of non-duplication
is satisfied;
UNTIL (No Duplication)
Compute the Fitness of the member retained;
Compute the Scaled Fitness: insert the member retained into the
population, sort the population and apply the Linear
Normalization technique to compute the scaled fitness values;
Insertion Test: the retained offspring is compared to the worst member
within the population. It is only inserted on the condition
that is better than the worst;
ENDFOR
ENDWHILE.

It is worth noting that the repeat-until loop within the above algorithm tests the duplication condition. The test on duplication is very quick since only comparisons of diameters are involved. Note that the resulting population may be composed of members having the same cost but different diameters.

OPTNET2, which implements the above algorithm has been developed during this research. Like OPTNET1, OPTNET2 used the same sets of data. Tests of OPTNET2 are provided in the following subsection.

9.4.5 Examples

Under this heading, OPTNET2 is applied to two numerical examples: Networks G and F. The coding used is the integer representation for the reasons observed in the previous section.

9.4.5.1 Network G

One of the convergence criteria used in the standard genetic algorithm is the number of generations allowed. In the steady state genetic algorithms, however, the concept of generation is not used. Instead, the convergence criterion applied is a desired number of function evaluations (Davis, 1991). Previously network G was optimised for a total number of function evaluations of 930. Here again, this number of trials is also considered.

As in subsection 9.3.4, OPTNET2 was run three times with network G for 930 function evaluations. Table 9.8 shows the initial population of strings, generated at random for the third run, while Table 9.9 displays the final population.

St_i	Diameters (mm)					Total Cost of Strings (£)
1 :	225	525	275	150	525	9235717.9390*
2 :	225	500	450	400	150	8807991.1564*
3 :	100	350	175	525	300	8606600.6451*
4 :	150	125	500	400	125	7579059.6420*
5 :	100	375	100	300	450	7138852.5675*
6 :	100	125	450	350	275	6622622.1158*
7 :	100	275	300	100	450	5974069.8874*
8 :	100	150	225	350	375	5089788.7120*
9 :	150	300	375	175	200	4613773.3192*
10 :	225	350	125	150	225	2991762.0094*
11 :	600	150	150	175	450	1549328.9547*
12 :	350	500	350	150	100	1134061.0626*
13 :	500	525	350	100	500	608764.9609
14 :	375	225	200	600	500	564130.9597
15 :	275	375	525	225	150	540812.6497*
16 :	500	175	250	225	600	505762.3394
17 :	525	525	175	175	375	496344.2026
18 :	525	300	200	400	450	492886.6118
19 :	375	450	400	375	350	475657.9860
20 :	275	525	250	375	300	457740.2507*
21 :	375	250	450	350	450	455081.5125
22 :	275	175	200	500	450	445023.8877*
23 :	350	175	400	275	525	418295.0516
24 :	275	400	400	350	400	412493.8851
25 :	350	125	450	375	400	396288.6047
26 :	450	200	400	275	400	391567.1443
27 :	300	225	400	275	500	386072.4250
28 :	450	225	225	100	400	327119.0924*
29 :	300	275	525	175	225	321168.3353
30 :	350	225	250	125	275	236469.4520*

Table 9.8 Processing the Initial Population - Network G

(* Violated solutions)

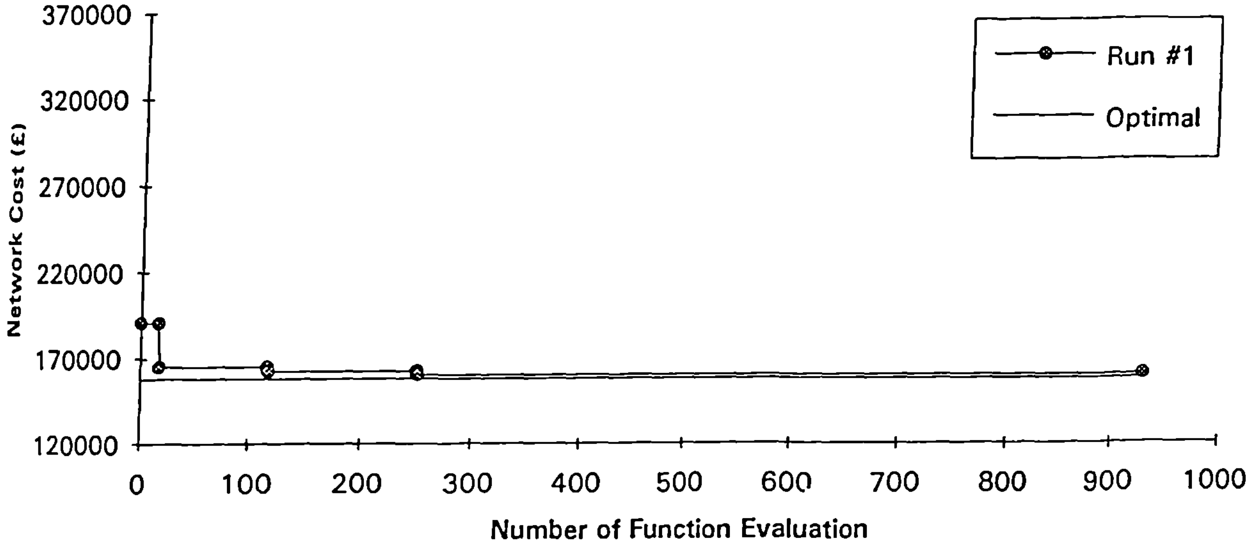


Figure 9.4.1 Best Network Costs for Run #1

- Application of OPTNET2 to Network G

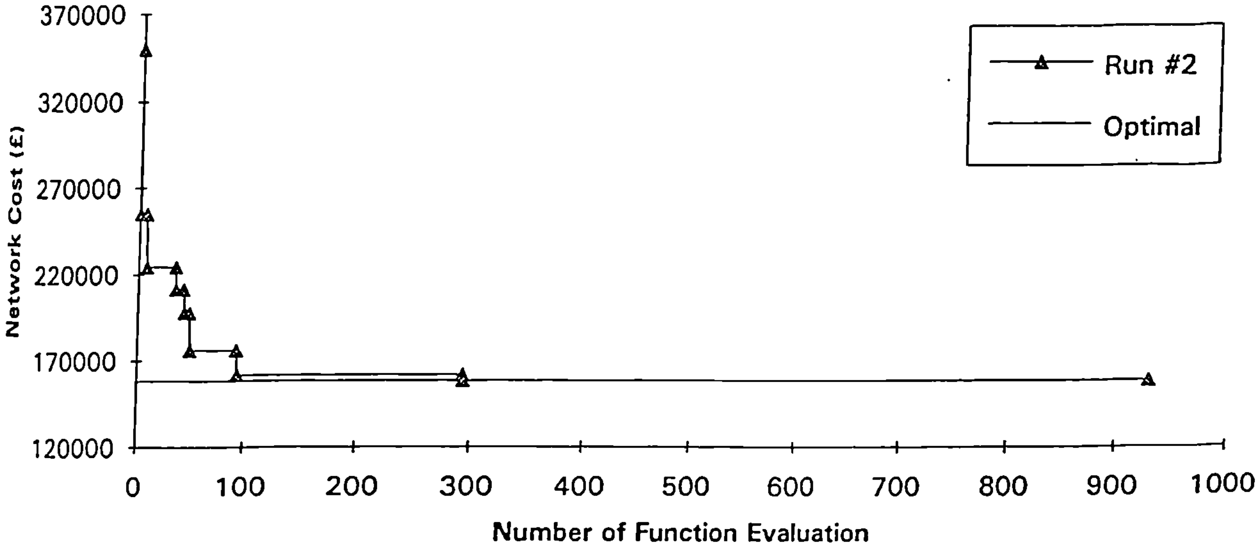


Figure 9.4.2 Best Network Costs for Run #2

- Application of OPTNET2 to Network G

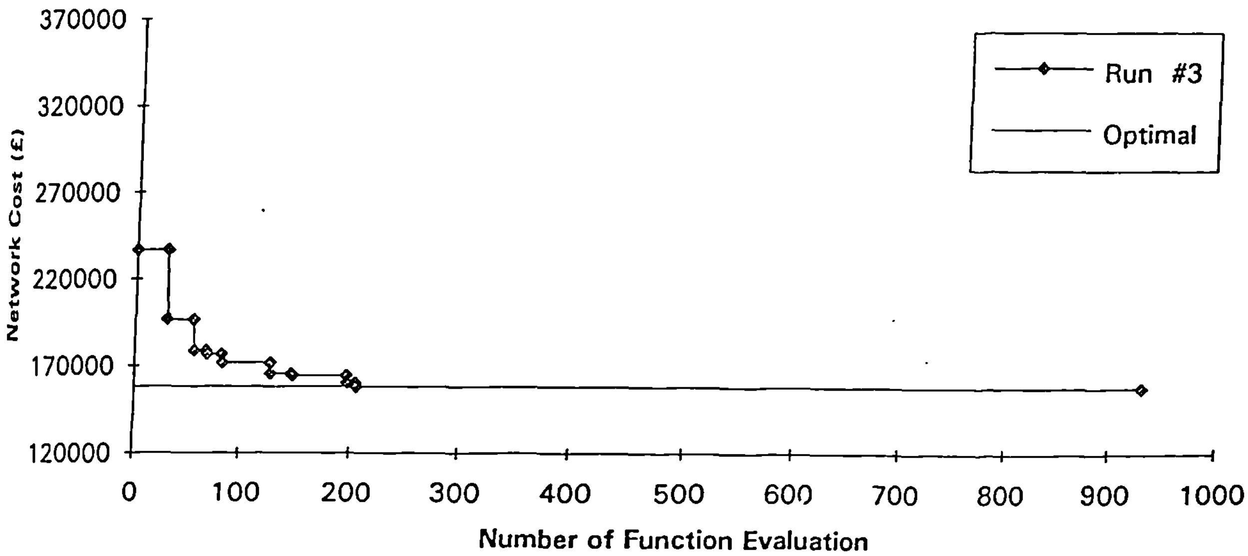


Figure 9.4.3 Best Network Costs for Run #3

- Application of OPTNET2 to Network G

St_i	Diameters (mm)	Total Cost of Strings (£)
1 :	350 225 225 200 250	173840.3842
2 :	350 225 200 225 250	173840.3842
3 :	350 275 150 200 250	173754.6155
4 :	350 275 200 150 250	173754.6155
5 :	350 250 225 150 250	171628.8482
6 :	350 225 150 250 250	171628.8482
7 :	350 250 150 225 250	171628.8482
8 :	350 225 125 175 250	171213.5202
9 :	350 225 150 150 250	170203.6543
10 :	350 250 175 200 250	169695.5274
11 :	350 250 200 175 250	169695.5274
12 :	350 225 225 175 250	168655.0627
13 :	350 225 175 225 250	168655.0627
14 :	350 275 175 150 250	168569.2939
15 :	350 275 150 175 250	168569.2939
16 :	350 225 100 175 250	168305.5358
17 :	350 225 125 250 250	168268.7456
18 :	350 275 200 100 250	167856.4188
19 :	350 225 200 200 250	167662.5461
20 :	350 200 150 250 250	165451.0100
21 :	350 250 150 200 250	165451.0100
22 :	350 250 200 150 250	165451.0100
23 :	350 250 175 175 250	164510.2059
24 :	350 225 150 225 250	164410.5452
25 :	350 225 175 200 250	162477.2245
26 :	350 225 200 175 250	162477.2245
27 :	350 200 125 250 250	162090.9075
28 :	350 250 150 175 250	160265.6885
29 :	350 250 175 150 250	160265.6885
30 :	350 225 150 200 250	158232.7071

Table 9.9 Processing the Final Population - Network G

Some of the features of the last population need to be highlighted. First, the solution costs are between £173840 and £158233, giving a region within which best costs are located. This offers the possibility of taking as an alternative to the optimum one of these solutions which may actually be preferred to the optimum, based on other non-quantifiable measures. This is a major benefit of the steady state genetic algorithm. In contrast, this was not observed with the generational replacement scheme where results of the last generation vary between £3352382 which is infeasible and £175263 (see Table 9.5). Second, as shown by Table 9.5, in the Generational Replacement scheme the best solution (£159553) was lost during the course of the further runs. For the steady state genetic search, however, the best string is preserved automatically. Third, the technique of non-duplication has been applied successfully ie, each member of the final population is unique. Since the test of duplication is performed on pipe sizes only, different solutions with the same cost may be members of the final solution. Examples are strings 1, 2; 3, 4; and 5, 6, 7. For the standard genetic algorithm, duplicate solutions which reduce the diversity of the population may be present in the final result as shown in Table 9.5 (eg strings 5, 6 and 24, 25).

Finally, the most important feature related to the application of OPTNET2 to network G is the fact of locating the global solution (£158233) in relatively few evaluations. In Figures 9.4.1/2/3, the best string costs, found during the courses of the three runs, are displayed against the number of function evaluations. Examination of these figures shows clearly the rapid convergence of OPTNET2 to the global solution in Figures 9.4.2 and 9.4.3 or to the very near global solution (Fig. 9.4.1) for fewer function evaluations compared to the generational replacement technique. The best solutions for each run, their numbers of function evaluations required and the differences from the global solution are tabulated in Table 9.10.

According to subsection 9.3.4, 27 generations were required to find the best solution (£159553). This corresponds to a total of 840 function evaluations. With OPTNET2, however, the total number of function evaluations (324) required is less than that corresponding to the application of OPTNET1. Therefore, for network G about half of the number of function evaluations are required by OPTNET2.

Run Number	Best Cost (£)	Number of Function Evaluations	Difference from Global (%)
1	160266	252	1.28
2	158233	294	0
3	158233	205	0

Table 9.10 Run Comparisons - Network G

As for OPTNET1, each complete run (930 function evaluations) of OPTNET2 has required for the same number of function evaluations 5 min on the same computer.

Before applying OPTNET2 to a greater network, it is worth noting that network G is symmetrical in layout and demand, whereas, the solutions produced by OPTNET2 and the global solution are unsymmetrical in pipe sizes. This can be explained by the fact that the optimisation technique used is a discrete technique. Furthermore the optimisation of the symmetrical network G without reliability constraints will be a tree configuration (Tree 1: 1, 2, 4 and 5 or Tree 2: 1, 2, 3 and 5) which is not symmetrical. As the GA program incorporates a measure of reliability, in the above global solution for instance, pipe 3 which is a link-forming loop has 150 mm which is greater than the minimum pipe size allowed (100 mm) for this example to produce the target reliability.

OPTNET2 finds also the "mirror-image" of unsymmetrical solutions. Examples are (see Table 9.9):

St_1: 350 225 225 200 250 and
St_2: 350 225 200 225 250

St_{28}: 350 250 150 175 250 and
St_{29}: 350 250 175 150 250

9.4.5.2 Network F

To further demonstrate the effectiveness of OPTNET2, a second numerical example was selected (Network F). The Data required for network F have been presented earlier in Chapter 8. As in chapter 8, the minimum pipe size considered here is again D_{min} = 50 mm. The number of candidate diameters N_d has been set to 16. In this situation, the maximum diameter is D_{max} = 450 mm (see Table 7.3). According to Chapter 8, a diameter of 350 mm was optimal and sufficient to convey the network maximum flow (Q_1 = 175 l/s). Therefore, 450 mm is a reasonable maximum diameter for link 1. Since the distribution of flow which may help in determining candidate pipe sizes for each link is unknown, N_d (50 mm-450 mm) is held constant for all links.

OPTNET2 was applied three times to network F with integer coding for 2000 function evaluations. The final results of the best run are shown in Table 9.11 where thirty alternative feasible solutions close to the optimum are presented.

Figures 9.5.1/2/3 show plots of the best string costs found during the course of trials for the three runs against the number of function evaluations.

As shown by Table 9.11, the genetic algorithm best solution corresponds to a network cost of £238170 found at the third run (Run #3) after 1947 function evaluations.

For run time, OPTNET2 has required approximately 2 Hrs 10 min on a 486 IBM compatible PC for 2030 function evaluations.

The following subsection addresses the problem of comparison of the results of the optimisation of network F by the two genetic programs: OPTNET1 and OPTNET2.

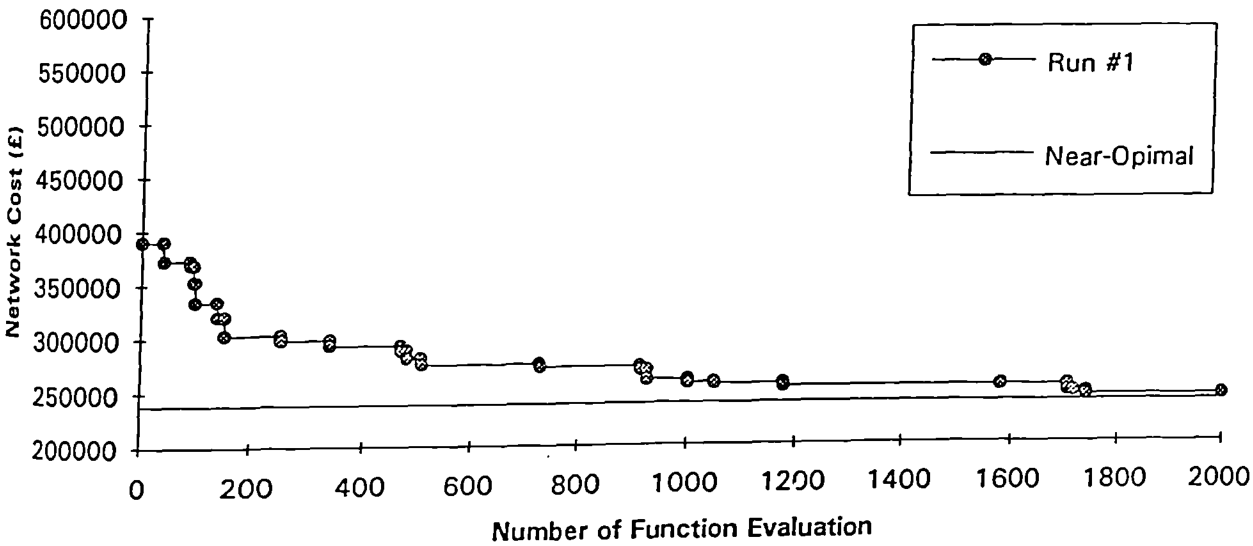


Figure 9.5.1 Best Network Costs for Run #1

- Application of OPTNET2 to Network F

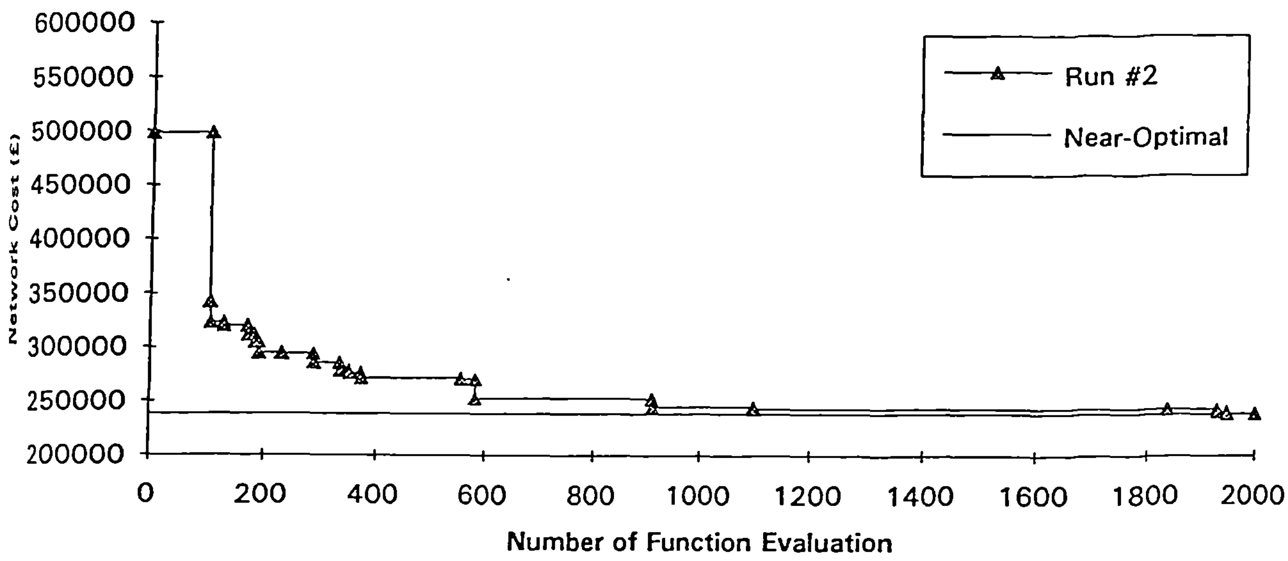


Figure 9.5.2 Best Network Costs for Run #2

- Application of OPTNET2 to Network F

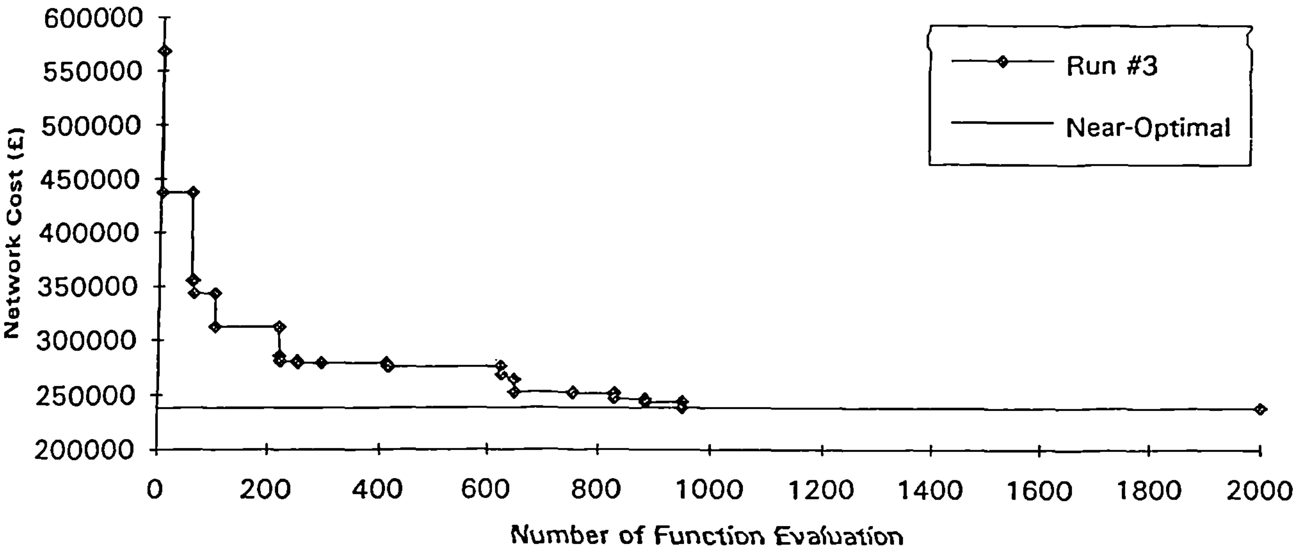


Figure 9.5.3 Best Network Costs for Run #3

- Application of OPTNET2 to Network F

St_i	Diameters (mm)													Total Cost of Strings (£)
1 :	375	250	250	50	175	125	80	50	150	150	80	100	300	248090.88
2 :	375	250	250	50	175	125	75	50	150	150	80	100	300	247789.94
3 :	350	250	250	50	175	100	80	150	175	150	100	100	300	247039.43
4 :	375	250	250	50	175	100	80	50	150	150	100	100	300	247038.47
5 :	375	250	250	50	175	100	80	80	150	150	80	100	300	246971.24
6 :	350	250	250	50	175	100	75	150	175	150	100	100	300	246738.48
7 :	375	250	250	50	175	100	75	50	150	150	100	100	300	246737.52
8 :	375	250	250	50	175	100	75	80	150	150	80	100	300	246670.29
9 :	350	250	250	50	175	150	80	80	175	150	100	100	300	245553.75
10:	350	250	250	50	175	100	80	150	175	150	80	100	300	245553.75
11:	375	250	250	50	175	100	80	50	150	150	80	100	300	245552.79
12:	350	250	250	50	175	150	75	80	175	150	100	100	300	245252.80
13:	350	250	250	50	175	100	75	150	175	150	80	100	300	245252.80
14:	375	250	250	50	175	100	75	50	150	150	80	100	300	245251.84
15:	350	250	250	50	175	150	80	50	175	150	100	100	300	244135.31
16:	350	250	250	50	175	150	80	80	175	150	80	100	300	244068.08
17:	350	250	250	50	175	150	75	50	175	150	100	100	300	243834.36
18:	350	250	250	50	175	150	75	80	175	150	80	100	300	243767.13
19:	350	250	250	50	175	150	80	50	175	150	80	100	300	242649.63
20:	350	250	*250*	*50*	*175*	*125*	*80*	*80*	*175*	*150*	*100*	*100*	*300*	*242193.65*
21:	350	250	250	50	175	125	75	80	175	150	100	100	300	241892.70
22:	350	250	250	50	175	125	80	50	175	150	100	100	300	240775.20
23:	350	250	250	50	175	125	80	80	175	150	80	100	300	240707.97
24:	350	250	250	50	175	125	75	50	175	150	100	100	300	240474.26
25:	350	250	250	50	175	125	75	80	175	150	80	100	300	240407.03
26:	350	250	250	50	175	100	80	80	175	150	100	100	300	239655.55
27:	350	250	250	50	175	100	75	80	175	150	100	100	300	239354.61
28:	350	250	250	50	175	125	80	50	175	150	80	100	300	239289.53
29:	350	250	250	50	175	100	80	50	175	150	100	100	300	238237.11
30:	350	250	250	50	175	100	80	80	175	150	80	100	300	238169.88

Table 9.11 Processing the Final Population - Network F

9.4.5.2.1 Comparison between OPTNET1 and OPTNET2

Network F was optimised three times with OPTNET1 for 100 generations. Fig. 9.6 shows the best-of-generation network costs against the number of generation. Table 9.12 presents the final results of running OPTNET1 and OPTNET2 with network F. The numbers of function evaluations presented in Table 9.12 did not include the first population. In the final column of Table 9.12, the 'near optimal' solution considered for computation is the best solution of the six runs.

Running OPTNET1 with network F has required 4 Hrs on the same PC for 3030 function evaluations.

Examination of Table 9.12 shows clearly that (1) the best solution for network F has been obtained by running OPTNET2. (2) Results of the three runs of OPTNET2 are all very close to the best network cost and are better than those corresponding to OPTNET1. (3) The total number of function evaluations (2030) required by OPTNET2 is less than that (3030) required by OPTNET1 and locates better solutions. (4) For the same network OPTNET2 has required lower run time (2 Hrs 10 min) than OPTNET1 (4 Hrs). These four points accord with the conclusion made earlier for the optimisation of network G. Increasing the number of generations may improve the quality of the solutions corresponding to OPTNET1 but at the cost of run time.

On the basis of the results presented in this Chapter, OPTNET2 outperforms OPTNET1, ie OPTNET2 not only locates near global solutions but also performs in a run time reduced relative to that required by OPTNET1.

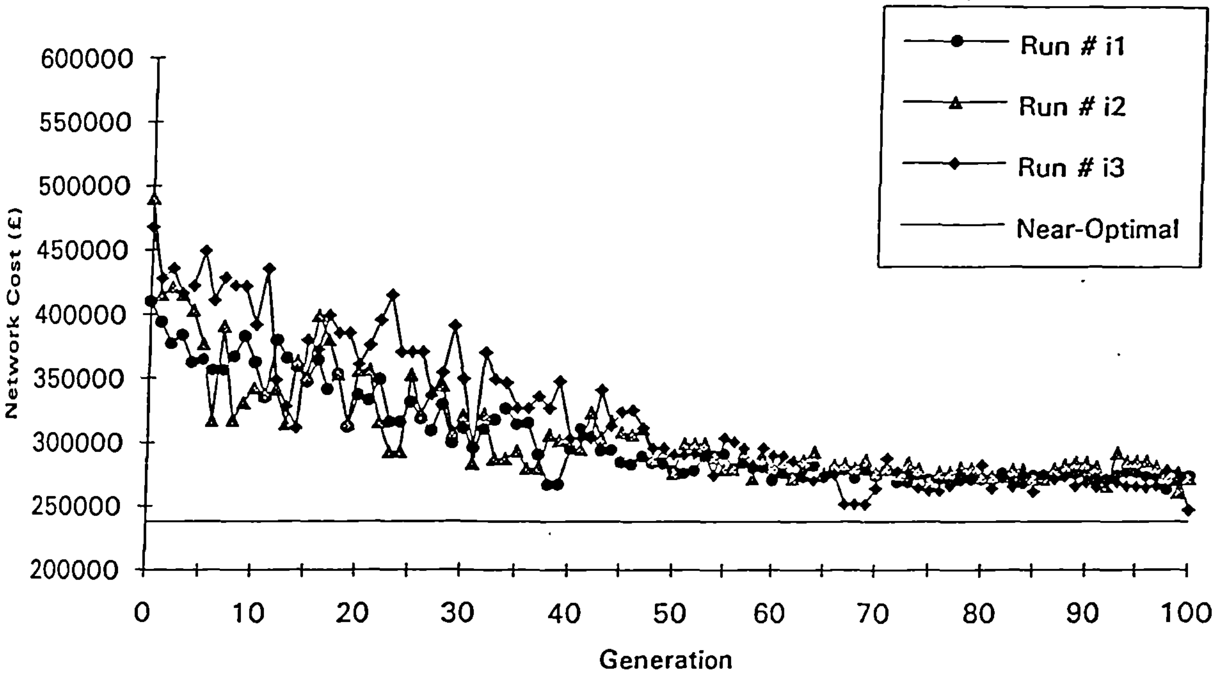


Figure 9.6 Best-of-generation Network Costs for the three Runs

- Application of OPTNET1 to Network F

Run Number	Diameters (mm)	Best Cost (£)	Number of Function Evaluations	Difference from near Optimal (%)
Results by OPTNET1				
1	375 300 225 175 175 75 80 175 100 125 200 100 225	262979	2970	10.42
2	375 300 225 80 275 80 150 100 125 75 200 75 175	260841	3000	9.52
3	350 275 225 50 175 150 50 150 125 75 250 150 225	246984	3000	3.70
Results by OPTNET2				
1	350 275 175 80 75 225 150 50 175 50 200 100 275	243433	1742	2.20
2	375 250 200 80 75 75 200 80 200 50 175 75 275	239130	1947	0.40
3	350 250 250 50 175 100 80 80 175 150 80 100 300	238170	951	0

Table 9.12 Run Comparisons - Network F

9.5 CONCLUSIONS

The efficient software developed in this project, OPTNET2, has demonstrated its efficacy as an improvement finding algorithm. Indeed, for each of the three runs, OPTNET2 was executed with network F to a total number of function evaluations of 2030. This may seem like a large number of function evaluations, for this case, until the size of the discrete space being searched is considered. For integer coding, each link of network F has the possibility of taking any diameter of the 16 diameters considered. This represents a total of $16^{13} = 4.50*10^{15}$ possible different alternatives in this huge search space. In this light, 2030 function evaluations is an insignificant fraction, 0.000000000044%, of the possible unique alternatives. To put this performance in perspective, if we were to search this efficiently for the best solution among say, 10000 billion solutions, we would only examine 5 solutions before making our selection.

Since OPTNET2 is a stochastic optimisation software, there is no guarantee that it will obtain the mathematically optimum solution to the problem.

One possible disadvantage is the computer time requirement for large networks. In most cases, however, this is not a major consideration in view of the cost savings that may be obtained. Such implementation should result in the savings of many millions of pounds per year to the water industry.

Chapter X

DISCUSSION

Chapter X

DISCUSSION

10.1 INTRODUCTION

Most aspects of reliability and optimisation for water distribution networks have been discussed in the previous Chapters. However, it would be interesting to make a comparison between the methods developed herein for the same water distribution network. For this purpose, network F was selected and optimised using the two reliability based optimal design models: the first is the Entropy/LP (PATH_Q + LNOPTNET) method and the second is the GA/Reliability Tester (OPTNET2) approach. The comparison must include costs and reliability aspects for the same minimum and maximum bounds of velocity. For the entropy based method, since the distribution of flow was known, practical bounds on velocity were used to determine candidate diameters and hence the optimal solution will be within these bounds. The minimum and the maximum velocities considered were 0.5 m/s and 3.0m/s respectively. For the second approach however, since the distribution of flow was unknown a specified number of candidate diameters was held constant for all the pipes within the network. Limits on velocity can only be known once the optimal design has been found. The following section addresses these considerations.

10.2 Comparison between Flow Assignment plus LP Method and OPTNET2

In Chapter 8, Network F was optimised using the Entropy/LP method for the candidate diameters specified in Appendix B. Since the distribution of flow is known, pipe velocities can be found by the application of the continuity equation (steady state flow condition).

In Chapter 9 on the other hand, running OPTNET2 with network F gave the best network cost and availability of the optimal solution. This solution needs to be analysed for the distribution of flow and pipe velocities.

For clarity, Table 10.1 presents the solutions produced by the two methods for network F along with flows and velocities. Table 10.2 summarises the results of the analysis in terms of pressure for both solutions.

Examination of Table 10.1 shows that the minimum and the maximum velocities for both methods are practically the same. Velocities are in the range of 0.50-2.60 m/s. As far as the GA solution is concerned, it should be noted that the flow directions have changed in three pipes (Links 4, 7 and 8) while in the entropy based solution flows are not allowed to reverse.

Table 10.2 shows that the most pressure-critical nodes are different for the two solutions as a consequence of the different directions and distributions of flow. Node 5 for the entropy based solution and node 4 for the GA solution.

For the reliability aspect, it can be seen from Table 10.1 that both methods produce reliable schemes (≥99.95%). However the solution resulting from the application of the Entropy/LP method is more reliable (99.98%).

With respect to cost the GA/RT method produces for this example a more economical solution than that corresponding to the Entropy/LP method. This difference in cost can be explained by the fact that in addition to the specification of the values of flow in the Entropy based approach, the directions of flow are not allowed to reverse. This additional constraint prevents the optimisation search from exploring other solutions based on different directions.

Link	X (m)	Entropy/LP D (mm)	Entropy/LP Q (l/s)	Entropy/LP V (m/s)	GA/RT D (mm)	GA/RT Q (l/s)	GA/RT V (m/s)
1	1000.00	375	248.82	2.25	350	248.82	2.59
2	605.92	200	80.44	2.56	250	118.80	2.42
	394.08	225		2.02			
3	207.69	150	33.60	1.90	250	77.93	1.59
	792.31	175		1.40			
4	1000.00	225	50.41	1.27	50	1.03	-0.53*
5	1000.00	225	41.11	1.03	175	34.00	1.41
6	770.53	150	24.67	1.40	100	4.15	0.53
	229.47	175		1.03			
7	1000.00	200	22.88	0.73	80	4.75	-0.95*
8	37.14	80	8.58	1.71	80	4.03	-0.80*
	962.86	100		1.09			
9	501.93	50	2.86	1.46	175	43.11	1.79
	498.07	75		0.65			
10	764.43	175	42.18	1.75	150	36.57	2.07
	235.57	200		1.34			
11	1000.00	275	84.37	1.42	80	5.42	1.08
12	564.21	250	80.44	1.64	100	7.45	0.95
	435.79	275		1.35			
13	1000.00	300	168.38	2.38	300	130.02	1.84
Cost :		£345513				£238170	
A_{NET} :		99.98%				99.95%	
V_{min} :		0.65 m/s at Link 9				0.53 m/s at Link 4	
V_{max} :		2.56 m/s at Link 2				2.59 m/s at Link 1	

Table 10.1 Comparison between the two methods for Network F
(* the initial direction has reversed)

Node	Entropy/LP Pressure (m)	GA/RT Pressure (m)
source	90.00	90.00
1	78.25	73.56
2	52.59	52.03
3	38.06	42.18
4	33.02	30.13
5	30.00	44.39
6	45.41	33.87
7	61.35	63.09

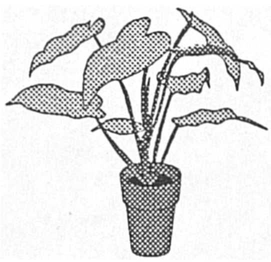
Table 10.2 Pressures of the two solutions

With respect to run time, the first method is very quick (about 5 min for network F) since it involves simple mathematical computations (Simplex algorithm) compared to the second (2 Hrs 10 min for network F) which requires frequent use of the reliability tester, ie, for each member of the population and for the number of function evaluations selected.

On the other hand, OPTNET2 may produce directly, not only the required system availability of the near global solution, but also a set of near global solutions that do not require to be adjusted as in the case of the solution produced by the first approach.

Chapter XI

RECOMMENDATIONS FOR FURTHER WORK



Chapter XI

RECOMMENDATIONS FOR FURTHER WORK

With the Reliability Tester, it is now possible to examine systematically in what sense entropy based design are reliable. It is recommended that future works investigate the impact of different directions of flow on the reliability of water networks based on the entropy principle.

The efficient optimisation program, OPTNET2, developed in this research is a global search algorithm for single source networks. OPTNET2 incorporates the Reliability Tester, which is a general frame for the calculation of reliability of water distribution networks. The use of OPTNET2 for the optimisation of complex water systems relies on the extension of the Reliability Tester. For this purpose, it is recommended that the Reliability Tester should be extended to accommodate more complex water networks.

The possibilities for further research in the general area of water distribution system optimisation with reliability considerations are numerous. These possibilities include the development of optimisation models for large and complex water distribution systems that incorporate a wide range of hydraulic components. It is likely that any successful work in this area will involve some form of Genetic Algorithm due to their robustness and efficiency in finding near global solutions. As seen earlier in this report, in addition to being easy to use, GAs do not rely on the restrictive assumptions of other optimisation methods (unimodality, existence of derivatives, piecewise linearity, etc.). Indeed, GAs stand ready, today, to optimise water distribution networks and other engineering systems.

Finding optimal locations of hydraulic components such as pumps, valves and water towers, including their optimal sizes and heights, may be possible areas for further work although some very few studies have been published in the literature (eg, Lansey and Mays, 1989).

Research should be carried out into the selection of optimal layouts of water distribution systems since very little work has been published in this area (eg, Rowell and Barnes, 1982; Morgan and Goulter, 1989).

For the reliability aspects, reliability models have been developed for pumps and links in the system. However no effort, to the knowledge of the author, has been made to develop the same models for joints and fittings in water systems which are more prone to failure (ie leakage) than pipes. Computation of nodal reliabilities also involves the junction of pipes, where two or more pipes are joined together by joints and fittings. In the absence of reliability models for the joints and fittings, the network nodes are assumed to be perfectly reliable. This assumption will give an upper bound to the overall system reliability. Therefore it is recommended that reliability models be developed for all components of water networks to obtain more realistic estimates of the network reliability.

The past work on rates of breakage of pipes, information on pump failures and failures of other network components, has shown that historical data in this area are scarce. Since these data are essential for the accuracy of reliability models, it is further recommended that a complete data base for all network components that records the break and repair history including date and time of breaks, date and time of repair, probable cause of break, type of break (eg a circumferential or a longitudinal crack in a pipe) etc, should be developed.

Chapter XII

CONCLUSIONS

Chapter XII

CONCLUSIONS

Basically three main models for the design of water networks have been developed in this study. The first model is related to the reliability aspects while the two other models handle the reliability-based optimal design of water networks. These models are: (1) the Reliability Tester, (2) the Entropy based flow assignment plus Linear Programming model and (3) the Genetic Algorithm search coupled with the Reliability Tester.

12.1 THE RELIABILITY TESTER

The reliability tester, that incorporates the major factors of reliability aspects such as the randomness of both the demands and pipe failures and the concept of repair time, is a powerful tool for fast assessment of nodal and system availabilities of water distribution networks.

The model can be utilised for the analysis of an existing water system and the identification of critical nodes with serious supply problems, and also identify the major causes of unreliability.

In the design of water networks, where the focus is on provision of a system of specified reliability or where the reliability is to be maximised, the model can also be used.

Due to its modest computational requirements compared to Monte-Carlo simulations, the model has been efficiently and successfully incorporated into the genetic searches: OPTNET1 and OPTNET2, for determining least cost design of water distribution networks under reliability constraints.

12.2 PATH_Q, SIZED AND LNOPTNET

The unsatisfactory methods currently available for the design of water networks, with reliability indices, require the introduction of new approaches such as the quick and efficient first approach developed in this work, and termed the flow assignment plus linear programming method. In this method the distribution of flow is based on the entropy principle.

PATH_Q finds the 'reliable' distribution of flow, SizeD determines the candidate diameters for each link within the water system according to the velocity constraints, while LNOPTNET solves for the network least cost and the optimal set of diameters. Execution of this sequence of programs with a water distribution network requires very low run time compared to that required by the genetic algorithm programs developed such as OPTNET2.

Application of the three computer programs: PATH_Q, SizeD and LNOPTNET to the design of water networks results in the production of more reliable schemes than those based on an arbitrary flow distribution. Results of the use of the entropy and LP based method for the design of three samples of networks showed that the maximum network availability obtained is a high availability of about 99.99% when tested using the Reliability Tester with standard assumption of demand distribution and pipe breakage statistics.

12.3 OPTNET2

The optimisation model, OPTNET2 developed herein is a robust stochastic search algorithm that can overcome the limitations of the previous models. The technique applied is based on the genetic algorithm and incorporates the Reliability Tester for the computation of nodal and network availabilities.

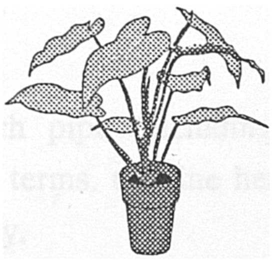
OPTNET2 uses two genetic operators: crossover and mutation that involve nothing more complex than string copying and partial string swapping.

Numerical results of the examples discussed have demonstrated that OPTNET2 is a rapid search algorithm in the huge spaces explored.

OPTNET2 locates not only a near-global design but identifies a region of a population of solutions close to the near-global solution that offers to the designer the flexibility of choosing any close alternative solution to the optimum, which may actually be preferred to the optimum solution, based on other non-quantifiable measures, eg environmental considerations.

One possible drawback of OPTNET2 is related to run time ie, OPTNET2 is time consuming. However, this is not a major consideration compared to the cost savings that may be obtained. Moreover, due to the rapid development in the world of PCs, if for instance, OPTNET2 now takes 10 Hrs to converge for a network to the best solution, in the near future, it will probably take just 10 min or less on faster machines.

APPENDICES



APPENDIX A

SOLUTION OF THE NETWORK EQUATIONS USING THE LINEAR METHOD

There is a considerable amount of published material for treating the problem of analysis of pressure and flows of water distribution networks. An efficient algorithm to solve this problem was proposed by Wood and Charles (1972) in which a linearization scheme was developed for the energy equations. The basic strategy is simple: the non-linear equations are first linearized using the tangent approximation for each pipe belonging to a closed circuit and/or path between the fixed head nodes. Then, the system of N_p equations is solved (N_J continuity equations are included) for the flows. The newly calculated flows are compared to the previous ones to check for convergence. Next, if the pre-defined criterion for convergence is satisfied the algorithm stops. Otherwise, these new values are substituted into the linearized part and the process continues. The global head equation for a path, expressed as a function f_Q of the flowrate Q, is:

$$\sum f_Q(Q) = \Delta E = \sum (h_L) - \sum (H_{pump}) \qquad \text{(A.1)}$$

h_L is the head loss in each pipe (including minor losses) which can be expressed as the sum of two terms, the line head loss in a pipe and the minor loss, h_{LP} and h_{LM} respectively.

$$h_{LP} = K_P Q^n \qquad \text{(A.2)}$$

$$h_{LM} = K_m Q^2 \qquad \text{(A.3)}$$

K_p is a pipe-line constant which is a function which depends on length, diameter and roughness of the pipe, and n is an exponent. K_p depends on the head loss equation used in the analysis. Given a pipe of diameter D and length L expressed in SI Units (m) K_p for the Hazen-Williams equation is:

$$K_p = K_x L / C^{1.852} D^{4.87} \tag{A.4}$$

Where

K_x = 10.70 for SI Units;

C = Hazen-Williams Coefficient;

n = 1.852.

K_m is the minor loss constant which is a function of the sum of the minor loss coefficients for the fittings in the pipe ($\sum m$) given by:

$$K_m = \sum (m) / 2gA_d^2 \tag{A.5}$$

Where

m = minor loss coefficient;

g = acceleration due to gravity, m/s^2;

A_d = $\pi D^2/4$, m^2.

Although pump head can be expressed in several ways, usually, it is described by a concave curve which is fitted to actual pump operating data. For the normal operating range, this curve is commonly approximated by a quadratic or exponential equation:

$$H_{pump} = a_p Q^2 + b_p Q + c_p$$

Or

$$H_{pump} = H_c - d_p Q^{e_p} \tag{A.6}$$

With a_p, b_p, c_p, d_p, e_p are coefficients and H_c the cutoff head. Irrespective of whichever relationship is used to link the pump head E_p to the discharge Q, one finishes with a function f_p of Q such that:

$$E_p \quad = \quad f_p(Q) \tag{A.7}$$

So, (A.1) can be written as:

$$\sum f_Q(Q) \quad = \quad \Delta E \quad = \quad \sum ((K_p Q^n + K_m Q^2) - f_p(Q)) \tag{A.8}$$

Eq. A.8 is then linearized in terms of an approximate flowrate, Q_i, in each pipe. This is achieved by developing a Taylor expansion truncated to the first order. The function f_Q and its gradient G_i evaluated at $Q = Q_i$ are:

$$f_Q(Q_i) \quad = \quad K_p Q_i^n + K_m Q_i^2 - f_p(Q_i) \tag{A.9}$$

$$G_i \quad = \quad f_Q'(Q_i) = \partial f_Q / \partial Q \Big|_{Q=Q_i} = nK_p \, Q_i^{n-1} + 2K_m \, Q_i - f_p'(Q_i) \tag{A.10}$$

and for a path, the following linearized equation results:

$$\sum f_Q(Q) \quad = \quad \Delta E \; = \; \sum (f_Q(Q_i) + G_i(Q - Q_i)) \tag{A.11}$$

Or

$$\sum (G_i Q) \quad = \sum (G_i Q_i - f_Q(Q_i)) + \Delta E \tag{A.12}$$

The resulting $(N_L + N_F - 1)$ linearized equations (A.11) added to N_J linear continuity equations form a set of simultaneous linear equations, in N_P unknown pipe flows. These can be determined by an efficient computer matrix routine, eg LU Decomposition for solving a set of linear equations.

APPENDIX B: Candidate Diametrs for Networks E &F

Network E

Link	Candidate Diameters (mm) Path_Q	Arbitrary Q Distribution
1	500 525 600	500 525 600
	675 700	675 700
2	300 350 375	350 375 400
	400 450	450 500
3	175 200 225	200 225 250
	250 275	275 300
4	350 475 400	300 350 375
	450 500	400 450
5	200 225 250	250 275 300
	275 300	350 375
6	125 150 175	150 175 200
	200 250	225 250
7	200 225 250	100 125 150
	275 300	175 200
8	150 175 200	125 150 175
	225 250	200 225
9	250 275 300	250 275 300
	350 375	350 375
10	200 225 250	175 200 225
	275 300	250 275
11	125 150 175	150 175 200
	200 225	225 250
12	150 175 200	150 175 200
	225 250	225 250
13	125 150 175	125 150 175
	200 225	200 225
14	150 175 200	150 175 200
	225 250	225 250
15	125 150 175	125 150 175
	200 225	200 225
16	100 125 150	75 80 100
	175 200	125 150
17	75 80 100	80 100 125
	125 150	150 175
18	75 80 100	100 125 150
	125 150	175 200

Table B1. Candidate Diameters for the Two Flow Distributions for Network E

Path	Number of Links in Path	Link Number					
	Node Equation						
1	1	1					
2	2	1	2				
3	3	1	2	3			
4	4	1	2	3	6		
5	3	1	2	5			
6	2	1	4				
7	3	1	4	9			
8	4	1	4	9	12		
9	5	1	4	9	12	13	
10	6	1	2	3	6	11	16
11	5	1	4	9	14	17	
12	4	1	4	9	14		
	Loop Equation						
13	4	-2	4	7	-5		
14	4	-3	5	8	-6		
15	4	-8	10	13	-11		
16	4	-7	9	12	-10		
17	4	-12	14	17	-15		
18	4	-13	15	18	-16		

Table B2. Strings of Pipes for Pressure and Loop Constraints

For Network E.

Network F

Link	Candidate Diameters (mm) Path_Q	Arbitrary Q Distribution
1	350 375 400 450 500	350 375 400 450 500
2	175 200 225 250 275	250 275 300 350 375
3	150 175 200 225 250	150 175 200 225 250
4	200 225 250 275 300	50 75 80
5	175 200 225 250 275	75 80 100 125 150
6	125 150 175 200 225	125 150 175 200 225
7	150 175 200 225 250	50 75
8	80 100 125 150	50 75
9	50 75	125 150 175 200 225
10	150 175 200 225 250	125 150 175 200 225
11	250 275 300 350 375	150 175 200 225 250
12	225 250 275 300 350	100 125 150 175 200
13	225 250 275 300 350	250 275 300 350 375

Table B3. Candidate Diameters for the Two Flow Distributions for Network F.

Path	Number of Links in Path	Link Number			
	Node Equation				
1	1	1			
2	2	1	2		
3	3	1	2	3	
4	4	1	2	3	5
5	4	1	2	11	8
6	3	1	2	11	
7	2	1	13		
	Loop Equation				
8	3	13	12	-2	
9	3	10	-11	-12	
10	3	11	4	-3	
11	3	6	-5	-4	
12	3	8	-7	-6	
13	3	9	-8	-10	

Table B4. Strings of Pipes for Pressure and Loop Constraints

For Network F.

APPENDIX C

RESULTS OF MINIMISATION OF

$$F = 3X_1^2 + 4X_1.X_2 + 5X_2^2$$

$$\text{S.T} \quad X_1, \quad X_2 >= 0$$

$$X_1 + X_2 >= 4$$

Generation 0

String No.	x_1	x_2	F	Fitness
1	5	4	235	2465
2	11	9	1164	1536
3	6	4	284	2416
4	3	13	1028	1672
5	5	2	135	2565
6	12	4	704	1996
7	5	12	1035	1665
8	14	1	649	2051
9	7	4	339	2361
10	5	3	180	2520

$\sum$ Fitness = 21247;
Max Fitness = 2565;
Avg Fitness = 2124.7;

Generation 7

String No.	x_1	x_2	F	Fitness
1	6	0	108	2592
2	7	6	495	2205
3	7	5	412	2288
4	6	0	108	2592
5	5	6	375	2325
6	5	4	235	2465
7	6	4	284	2416
8	13	3	708	1992
9	12	5	797	1903
10	4	3	141	2559

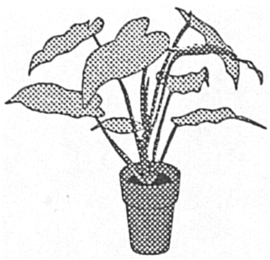
$\sum$ Fitness = 23337;
Max Fitness = 2592;
Avg Fitness = 2333.7;

Generation 10

String No.	x_1	x_2	F	Fitness
1	7	3	276	2424
2	6	1	137	2563
3	7	1	180	2520
4	3	13	1028	1672
5	7	2	223	2477
6	6	8	620	2080
7	7	5	412	2288
8	7	1	180	2520
9	3	1	44	2656 (optimum)
10	7	0	147	2553

$\sum$ Fitness = 23753;
Max Fitness = 2656;
Avg Fitness = 2375.3;

REFERENCES



REFERENCES

Alperovits, E. and U. Shamir (1977) Design of Optimal Water Distribution Systems, Water Resources Research, Vol. 13, No. 6, p. 885-900.

Altman, T. and P. Boulos (1992) Convergence of Newton Method in Nonlinear Analysis, Tech. Report TR-14, Dept. of Computer Science and Engineering, Univ. of Colorado at Denver, March.

Anon, (1968) *Pumping Manual.* third edition, Trade and Technical Press.

Anon, (1982) *The British Pump Market.* 1981-1982, C.H.W. Roles and Associates.

Awumah, K., S. Bhatt and I.C. Goulter (1989) An Integer Programming Model for Layout Design of Water Distribution Networks, Eng. Opt., Vol. 15, p. 57-70.

Awumah, K., I.C. Goulter and S. Bhatt (1991) Entropy Based Redundancy Measures in Water Distribution Network Design, J. of Hyd. Div., ASCE, Vol. 117, No. 10, p. 595-614.

Awumah, K. and I.C. Goulter (1992) Maximising Entropy-Defined Reliability of Water Distribution Network Design, Eng. Opt., Vol. 20, p. 57-80.

B. S. 3811 (1974) *Maintenance Terms in Terotechnology.* British Standards Institution, London.

Bao, Y. and L.W. Mays (1990) Model for Water Distribution System Reliability, J. of Hyd. Eng., ASCE, Vol. 116, No. 9, Sep., p. 1119-1137.

Bazaraa, S.M. and J.J. Jarvis and H.D. Sherali (1990) *Linear Programming and Network flows.* John Wiley and Sons, Inc., New York, N. Y.

Beim, G.K. and B.F. Hobbs (1988) Analytical Simulation of Water System Capacity Reliability, 2. A Markov Chain Approach and Verification of the Models, Water Resources Research, Vol. 24, No. 9, Sep, p. 1445-1458.

Bhave, P. (1978) Noncomputer Optimization of Single-Source Networks, J. of Env. Eng. Div., ASCE, Vol. 104, No. EE4, p. 799-813.

Billinton, R., (1972) Bibliography on the Application of Probability Methods in Power System Reliability Evaluation, IEEE Trans. Power Appar. Syst., PAS-91 (2), p. 649-660.

Billinton, R. and R.N. Allan (1983) *Reliability Evaluation of Engineering Systems: Concepts and Techniques.* Plenum Press, New York, N. Y.

Billinton, R. and R.N. Allan (1984) *Reliability Evaluation of Power Systems.* Plenum Press, New York, N. Y.

Bouchart, F. and I.C. Goulter (1991) Reliability Improvement in Design Networks recognizing Valve Location, Water Resources Research, Vol. 27 No. 12, p. 3029-3040.

Booker, L.B. (1982) *Intelligent behaviour as an adaptation to the task environment.* PhD thesis, University of Michigan, at Ann Arbor, Mich.,

Booker, L.B. (1987) Improving Search in Genetic Algorithms, in Davis (ed), *Genetic Algorithms and Simulated Annealing,,* Pitman, London, and Morgan Kaufmann Publishers, Inc. p. 61-73.

Box, M.J. (1966) A comparison of Several Current Optimization Methods of Transformation in Constrained Problems, The Computer Journal, London, England, Vol. 9, p. 67-77.

Brindle, A. (1981) *Genetic Algorithms for Function Optimization.* PhD and Technical Report TR81-2, Edmonton: Univ. of Alberta, Dept. of Computer Science.

Bunday, B.D. and G.R. Garside (1987) *Optimisation Methods in Pascal.* Edward Arnold, London.

Charnes, A. and W. Cooper (1963) Deterministic Equivalents for Optimizing and satisfying Under Chance Constraints, Operations Research, Vol. 11, No. 1, p. 18-39.

Charnes, A. and A.C. Sterdy (1966) A Chance-Constrained Model for Real-Time Control in Research and Development Management, Management Science, Vol. 12, No. 8, Apr., p. 335-362.

Chiu, C. L. (1987) Entropy and Probability Concepts in Hydraulics, J. of Hyd. Eng., ASCE, Vol. 113, No. 5, p. 583-600.

Chiu, C. L. (1988) Entropy and 2-D Velocity Distribution in Open Channel, J. of Hyd. Eng., ASCE, Vol. 104, No. 7, p. 738-756.

Coals, A.V. and I.C. Goulter (1985) Approaches to the Consideration of Reliability in Water Distribution Networks, International Symposium of urban Hydrology, Hydraulics, Infrastructures and Water Quality Control., Univ. of Kentucky, Lexington, Kentucky, p. 287-295.

Collins, A.G. and R.L. Johnson (1975) Finite-Element Method for Water Distribution Networks, J. of A.W.W.A., Vol. 67, July, p. 385-389.

Collins, A.G., L. Looper and J.L. Kennington (1976) Solving the Pipe Network Analysis Problem using Optimization Techniques, Technical Report, IEOR 76008, School of Engineering and Applied Science, Southern Methodist Univ.

Critchley, R. and J.J. Habershon (1981) An Investigation into the Influence of Pipe Characteristics on the Failure Rate of uPVC and Cast Iron Water Mains, Water Research Centre, Report 7E.

Cross, H. (1936) Analysis of Flow in Networks of Conduits or Conductors, Bulletin No. 286, Univ. of Illinois Eng. Experimental Station, Urbana, III.

Cullinane, M.J. (1986) Hydraulic Reliability of Urban Water Distribution Systems, Water Forum '86', ASCE, New York, N. Y., p. 353-358.

Cullinane, M.J. (1987) Hydraulic Availability Evaluation of Urban Water Distribution Systems, Proc. of the 1987 National Conf. on Hyd. Eng., ASCE, p. 463-468.

Davidon, W.C. (1959) Variable Metric Method For Minimization, Atomic Energy Commission Research and Development Report, ANL-5990 (Revised), Argonne Labo., U.S., Atomic Energy Commission.

Davis, L. (1989) Adapting Operator Probabilities in Genetic Algorithms, Proc. of the third International Conf. on Genetic Algorithms, Morgan Kauffman, Los Altos, CA, p. 61-69.

Davis, L. (1991) *The Handbook of Genetic Algorithms*. Var Nostrand Reinhold, New York.

Deb, A. (1978) Optimization in Design Pumping Systems, J. of Env. Eng. Div., ASCE, Vol. 104, No. EE1, p. 127-136.

De Jong, K.A. (1975) *An Analysis of the Behaviour of a Class of Genetic Adaptive Systems*. PhD Thesis, Univ. of Michigan.

De Jong, J.J., J.A.C. Snijders and A. Wedmijer (1983) Reliability of the Rotterdam Drinking Water System, H_2O, 16(1), p. 10-16 (in Dutch).

Dennis, J.E. and R.B. Schnabel (1983) *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs, N. J. p. 112-129 and 203-207.

Didier, G. (1980) *Hydraulique appliquée*. Tome 1, Alimentation.

Duan, N. and L.W. Mays (1987) Reliability Analysis of Pumping Stations and Storage Facilities, Proc. of the 1987 National Conf. on Hyd. Eng., ASCE, p. 600-605.

Duan, N. and L.W. Mays (1990) Reliability Analysis of Pumping Systems, J. of Hyd. Eng., ASCE, 116(2), p. 230-248.

Eigen, M. (1987) Stufen Zum Leben, Piper Verlag, München.

Endrenyi, J. (1978) *Reliability Modelling in Electric Power Systems Wiley Inter-Science Publications*. New York, N. Y.

Erlander, S. (1977) Accessibility, Entropy and the Distribution and Assignment of Traffic, Transportation Research 11, p. 149-153.

Erlander, S. (1980) Optimal Spatial Interaction and the Gravity Model, Lectures Notes in Economics and Mathematical Systems (Springer Verlag, Berlin, 1980).

Even, S. (1979) *Graph Algorithms*. Computer Science Press, Potomac, Maryland, p. 62-63.

Even, S. and R.E. Tarjan (1976) Computing an s-t numbering, The computer Science, (2), p. 339-344.

Fujiwara, O. and D. Dey (1987) Two Adjacent Pipe Diameters at The Optimal Solution in the Water Distribution Network Models, Water Resources Research, Vol. 23, No. 8, Aug., p. 1457-1460.

Fujiwara, O., B. Jenchaimahakoon and N.C.P. Edirisinghe (1987) A modified Linear Programming Gradient Method for optimal Design of Looped Water Distribution Networks, Water Resources Research, Vol. 26, No.6, p. 977-982.

Fujiwara, O. and D.B. Khang (1990) A Two-Phase Decomposition Method for Optimal Design of Looped Water Distribution Networks, Water Resources Research, Vol. 26, No. 4, Apr., p. 539-549.

Fujiwara, O. and A.U. Silva (1990) Algorithm for Reliability-Based Optimal Design of Water Distribution Networks, J. of Env. Eng. Div., ASCE, Vol. 116, No. 3, May/June, p. 575-587.

Fujiwara, O. and H.D. Tung (1991) Reliability Improvement for Water Distribution Networks Through Increasing Pipe Sizes, Water Resources Research, Vol. 27, No. 7, p. 1395-1402.

Gessler, J. (1982) Optimisation of Pipe Networks, International Symposium on Urban Hydrology, Hydraulics and Sediment Control, Univ. of Kentucky, Lexington, KY.

Gessler, J. (1985) Pipe Network Optimization by Enumeration, Proc., ASCE, Computer Applications for Water Resources, Buffalo, NY, p. 572-581.

Gessler, J. and T.M Walski (1985) Water Distribution System Optimization, Final Report, Environmental Laboratory, Dept. of the Army, Waterways Experiment Station, Corps of Engineers PO Box 631, Vicksburg, Mississippi 39180-0631, Oct.

Goldberg, D.E, (1983) Computer-aided Gas Pipeline Operation using Genetic Algorithms and Rule Learning, PhD Thesis, Univ. of Michigan, at Ann Arbor, Mich.

Goldberg, D.E, (1987) Genetic Algorithms in Pipeline Optimization, ASCE, J. of computing in Civil Engineering, Vol. 1, No. 2, April, p. 128-141.

Goldberg, D.E, (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addision-Wesley, Reading, MA.

Goldberg, D.E, B. Korb and K. Deb (1989) Messy Genetic Algorithms: Motivation, Analysis, and First Results, Complex Systems, Vol. 3, p. 793-530.

Goulter, I.C. (1987) Current and Future Use of System Analysis in Water Distribution Network Design, Civil Engineering Systems, Vol. 4, Dec., p. 175-183.

Goulter, I.C. (1992) Modern Concepts of a Water Distribution System: Policies for Improvement of Networks with Shortcoming, Los abastecimientos de agua urbanos. Estado actual y tendencias futuras, U.D. Mecanica de Fluidos, Universidad politecnica de Valencia, Valencia, 19-23 October, p. 83-99.

Goulter, I.C. and F. Bouchart (1987) Joint Consideration of Pipe Breakage and Pipe Flow Probabilities, Proc. of 1987 National Conf. on Hyd. Eng., Williamsburg, Virginia, p. 469-474.

Goulter, I.C. and A.V. Coals (1986) Quantitative Approaches to Reliability Assessment in Pipe Networks, J. of Transportation Eng., Vol. 112, No. 3, p. 287-301.

Goulter, I.C. and D. Morgan (1985) An Integrated Approach to the Layout and Design of Water Distribution Networks, Civil Engineering Systems, Vol. 2, June, p. 104-113.

Grefenstette, J.J. (1986) Optimization of Control Parameters for Genetic Algorithms, IEEE, Transactions on Systems, Man and Cybernetics, SMC-16, 1, p. 122-128

Gupta, I. (1969) Linear Programming Analysis of Water Supply System, Transaction of the American Inst. of Industrial Engineers, Vol. 1, No. 1, p. 56-61.

Haarhoff, P.C. and J.D. Buyes (1970) A New Method for The Optimization of Nonlinear Functions Subject to Nonlinear Constraints, The Computer Journal, London, England, Vol. 13, No. 2, May, p. 178-184.

Headquarters Department of the Army (1987) MAPS User's Guide, Part 1 of 2, Change 6 (Draft). EM 1110-2-502, Washington, DC.

Henley, E.J. and H. Kumamoto (1981) *Reliability Engineering and Risk Assessment.* Prentice hall, Inc. Englwood cliffs. N. J.

Hobbs, B.F. (1985a) Reliability Analysis of Urban Water supply, In Computer Applications in Water Resources, American Society of Civil Engineers, p. 1229-1238.

Hobbs, B.F. (1985b) Reliability Analysis of Water System Capacity, in Hydraulic and Hydrology in the Small Computer Age, American Society of Civil Engineers, Vol. 1, p. 341-346.

Hobbs, B.F. and G.K. Beim (1986) Verification of Supply Reliability Model, Case Western Reserve Univ. Cleveland, Ohio.

Holand, J.H. (1975) *Adaptation in Natural and Artificial Systems.* Univ. of Michigan Press, Ann Arbor.

Hooper, B.D. (1981) Survey of water tourist consumption, South West Authority.

Ibid (1982) Unplasticised PVC Pipes and Fittings, Information and Guidance Note, 4-31-01.

IEEE Subcommittee on the Application of Probability Methods (1978) Bibliography on the Application of Probability in Power System Reliability Evaluation, IEEE Trans. Power Appar. Syst. PAS-97(6),p. 2235-2242.

Institution of Water Engineers and Scientists. (1983) *Water Practice Manual.* Book 3, I.W.E.S. London.

Institution of Water Engineers and Scientists. (1984) *Water Practice Manual.* Book 4, I.W.E.S. London.

Itai, A. and M. Rodeh (1984) The Multi-tree Approach to Reliability in Distribution Networks, Found. of Computer Science, Proc. of the 25th Symposium, p. 137-147.

Jacoby, S. (1968) Design of Optimal Hydraulic Networks, J. of Hyd. Div., ASCE, Vol. 94, No. HY3, p. 641-661.

Jacobs, P. and I.C. Goulter (1989) Optimisation of Redundancy in water Distribution Networks using Graph Theoretic Principles, Eng. Opt., Vol. 15, No. 1, p. 71-82.

Jaynes, E. T. (1957) Information Theory and Statistical mechanics, The Physical Review, Vol. 106, p. 620-630 and Vol. 108, p. 171-190.

Jeppson, R.W. (1983) *Analysis of Flow in Pipe Networks.* Ann Arbor Science. The Butterworth Group.

Kally, E. (1972) Computerized planning of the Least Cost Water Distribution Network, Water and Sewage Works, p. 121-127.

Kapur, J.N. (1983) *Maximum Entropy Model in Science and Engineering*. South Asia Publisher, New Delhi, India.

Karmeli, D., Y. Gadish and S. Meyers (1968) Design of Optimal Water Distribution Networks, J. of Pipe Div., AMCE, 94 (PL1), p. 1-9.

Kennington, J.L. and R.V. Helgason (1980) *Algorithms for Network Programming*. John Wiley and Sons.

Kesavan, H.K. and M. Chandrashekar (1972) Graph-Theoretic Models for Pipe Network Analysis, J. of Hyd. Div., ASCE, Vol. 98, Feb., p. 345-364.

Kessler, A., L. and U. Shamir (1989) Analysis of the Linear Programming Gradient Method for Optimal Design of Water Supply Networks, Water Resources Research, Vol. 25, No.7, Jul., p. 1469-1480.

Kessler, A., L. and U. Shamir (1991) Decomposition Technique for Optimal Design of Water Supply Networks, Eng. Opt., Vol. 17, p. 1-19.

Kessler, A., L. Ormsbee and U. Shamir (1990) A Methodology for Least Cost Design of Invulnerable Water System, Vol. 7, No. 1, March, p. 20-28.

Kettler, A.J. and I.C. Goulter (1983) Reliability Consideration in the Least Cost Design of Looped Water Distribution Systems, Proc. International Symposium on Urban Hydrology, Hydraulics and Sediment Control, Lexington, Kentucky, p. 305-312.

Kettler, A.J. and I.C. Goulter (1985) An Analysis of Pipe Breakage in Urban Water Distribution Networks, Canadian J. of Civil Engineering, p. 286-293.

Khomsi, D. (1991) Looped Water Distribution Network Optimisation with Reliability Consideration, Internal Report, School of Engineering, City University, August, 75p.

Khomsi, D., G.A. Walters, A.R.D. Thorley and D. Ouazar (1993) A Reliability Tester for Water Distribution Networks, Paper submitted to the Journal of Computing in Civil Engineering, ASCE.

Khinchin, A.I. (1953) The Entropy Concept in Probability Theory, Uspekhi Matematich eskikh, 8, (3), 3-20. Translation in A.I. Khinchin (1957) Mathematical Foundations of Information Theory, Dover, New York, p. 1-28.

Kim, Y.H., K.E. Case and P.M. Ghare (1972) A Method for Computing Complex System Reliability, IEEE Transactions on Reliability, R12, (4), p. 215-219.

Kirby, P.C. (1981) PVC Pipe Performance in water mains and Sewers, International Conf. on underground Plastic Pipes, New Orleans.

Kullback, S. (1959) *Information Theory and Statistics.* Wiley, New York.

Kumar, V. (1987) Entropic Measures of Manufacturing Flexibility, International. J. of Production Res., 25(7), p. 957-966.

Labye, Y. (1966) Etude des Procedés de Calcul Ayant pour But de Rendre Minimal le Coût d' Un Réseau de Distribution d' Eau sous Pression, La Houille Blanche, No. 5, p. 577-583.

Lackington, D.W. (1983) Survey of the Renovation of Water Mains, Proc. of the I.W.E.S Symposium on Deterioration of Underground Assets.

Lasdon, L.S., A.D. Waren and M.S. Ratner (1984) "GRG User's guide", Univ. of Texas at Austin, Tex.

Lai, D. and Schaake (1969) Linear Programming and Dynamic Programming Applied to Water Distribution Network Design, Massachusetts Inst. of Tech., Hydrodynamics Lab., Report 116, Cambridge, MA.

Lansey, K.E. and L.W. Mays (1987) Optimal Design of Large Scale Water Distribution Systems, Proc. of the 1987 National Conf. on Hyd. Eng., ASCE, Williamsburg, Virginia, August 3-7, p. 475-480.

Lansey, K.E. and L.W. Mays (1989) Optimization Model for Water Distribution System Design, J. of Hyd. Eng., Vol. 115, No. 10, October, p. 1401-1418.

Latham, B. (1990) *Water Distribution.* The institution of Water and Environmental Management, London.

Lin, C.Y and P. Hajela (1992) Genetic Algorithms in Optimization Problems with Discrete and Integer Design Variable, Eng. Opt., Vol. 109, p. 309-327.

Loganathan, G.V., M.P. Shah and H.P. Sherali (1990) A Two-Phase Network Design Heuristic for Minimum Cost Water Distribution Systems Under a Reliability Constraint, Eng. Opt., Vol. 15 , No. 4, p. 311-336.

Male, W., T.M. Walski and A.H. Slutsky (1990) Analyzing Water Main Replacement Policies, J. of J. of Water Resources Planning and Management, Vol. 102, No. 116, Dec., p. 1539-1555.

Martin, D.W. and G. Peters (1963) The application of Newton's Method to Network Analysis by Digital Computer, J. of Inst. of Water Engineers, Vol. 17, p. 115-129.

Mays, L.W. (1989) New Methodologies for the Reliability-Based Analysis and design of Water Distribution Systems, Technical Report CRWR 227, Center for Research in Water Resources, Univ. of Texas, Austin, Tex.

Mays L.W. and M.J. Cullinane (1986) A Review and Evaluation of Reliability Concepts for Design of Water Distribution Systems, Miscellaneous Paper EL-86-1, US Army Engineer Waterways Experiment Station, Vicksburg, Miss.

Mays, L.W., N. Duan and Y.C. Su (1986) Modelling Reliability in Water Distribution Network Design. In Water Forum '86', World Water Issues in Evolution, American Society of Civil Engineers, Vol. 2, p. 1272-1278.

Morgan, D.R. and I.C. Goulter (1985) Optimal Urban Water Distribution Design, Water Resources Research, Vol. 21, No. 5, p. 642-652.

O'Day, D.K. (1982) Organizing and Analysing Leak and Break Data for Making Main Replacement Decisions, J. of A.W.W.A., 74(11), Denver Co., p. 589-596.

Orth, H.M. (1986) *Model-based Design of Water Distribution and Sewage systems*. John Wiley, New York.

Park, C.C. (1982) The Supply of and demand for Water, in Johnson, R. J. and Doornkamp, J. C. (Ed.). *The Changing Geography of the United Kingdom*, Methuen, London, p. 192-148.

Park, C.C. (1986) Water Demand Forecasts and the Social Sciences, Water Demand Forecasting, Proc. of workshop, Vince Gardiner and Paul Herrington, Univ. of Leicester, p. 17-23.

Patton, J.L. and M.B. Horsly (1980) Curbing the Distribution Energy Appetite, J. of A.W.W.A, 72, 315.**

Pettey, C.B., M.R. Leuze and J.J. Grefenstette (1987) A parallel Genetic Algorithm, in Proc. of an International Conf. on Genetic Algorithms and their Applications, in Grefenstette (ed), Carnegie-Melon University, Pittsburgh, p. 155-161.

Power, N.A., R.E. Volker and K.P. Stark (1981) Deterministic Models for Predicting Residential Water Consumption, Water Resources Bulletin, 17, 1042.

Press, W., B. Flannery, S. Teukolsky and W. Vetterling (1988) *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, New York.

Quimpo, R. and U. Shamsi (1987) Network Reliability for Water Supply Reliability Determination, Proc. of the 1987 National Conf. on Hyd. Eng., Williamsburg Hilton National Conf., Center Williamsburg, Virginia, August 3-7, p. 716-721.

Quindry, G.E., E.D. Brill and J.C. Liebman (1981) Optimization of Looped Water Distribution systems, J. of Env. Eng. Div., ASCE, Vol. 107, No. EE4, p. 665-679.

Quindry, G.E., E.D. Brill, J.C. Liebman and A. Robinson (1979) Comment on 'Design of Optimal Water Distribution systems' by Alperovits and Shamir, Water Resources Research, Vol. 15, No. 6, p. 642-652.

Rajeev, S. and C.S. Krishnamoorthy (1992) Discrete Optimization of Structures Using Genetic Algorithms, J. of Structural Eng., ASCE, Vol. 118, No. 835, p. 1233-1250.

Richardson, J.T., M.R. Palmer, G. Liepins and M. Hilliard (1989) Some Guidelines for Genetic Algorithms with Penalty Functions, Proc. of the third International Conf. on Genetic Algorithms and Their Applications, Morgan Kauffman, Los Altos, CA, p. 191-197

Robert, N.P. and T. Regan (1974) Causes of Fractures in Grey Cast Iron Water Mains, City University.

Robinson, R. and T. Austin (1976) Cost Optimization of Rural Water Systems, J. of Hyd. Div., ASCE, Vol. 102, No. HY8, p. 1119-1134.

Rosenthal, A. (1977) Computing the Reliability of Complex Networks, SIAM, J. Appl. Math., 32(2), p. 384-393.

Rowell, W. and J. Barnes (1982) Obtaining the Layout of Water Distribution Systems, J. of Hyd. Div., ASCE, Vol. 102, No. HY8, p. 1119-1134.

Satyanarayana, A. and R.K. Wood (1982) Polygon-to chain Reductions and Network Reliability, ORC 82-4, Operations Research Center, Univ. of California, Berkeley, Calif.

Schaffer, J.D., R.A. Caruana, L.J. Eshelman and R. Das (1989) A Study of Control Parameters Affecting Online Performance of Genetic Algorithms for Function Optimization, Proc. of the third International Conf. on Genetic Algorithms. Morgan Kauffman, Los Altos, CA, p. 51-60.

Schrage, L. (1987) USER's Manual for Linear, Integer and Quadratic Programming with LINDO, Third Edition, The Scientific Press.

Shamsi U.M. and R.G. Quimpo (1988) Network Reliability for Water Distribution System Management, Proc. of Fifth IAHR International Symposium on Stochastic Hydraulics, August, p. 298-302.

Shamir, U. (1974) Optimal Design and Operation of Water Distribution Systems, Water Resources Research, Vol. 10, No. 1, p. 27-37.

Shamir, U. (1979) Optimization in Water Distribution Systems Engineering, Mathematical Programming No. 11, p. 65-75.

Shamir, U. and C.D.D. Howard (1968) Water Distribution System Analysis, J. of Hyd. Div., ASCE, Vol. 94, No. HY1, p. 219-235.

Shamir, U. and C.D.D. Howard (1981) Water Supply Reliability theory, J. of A.W.W.A., 73(7), p. 379-384.

Shamir, U. and C.D.D. Howard (1985) Reliability and Risk Assessment for Water Supply Systems, Computer Applications in Water Resources, H. Tourno, ed., ASCE, New York, N. Y.

Shannon, G. E. (1948) A Mathematical Theory of Communication, Bell System Technical Journal, Vol. 27, No. 3, p. 379-428.

South West Water Authority (1978) Study of Domestic Water Consumption.

Stephens, J.W. and B.W. Gill (1982) Service Failure Experience of uPVC Pressure Pipes in the Water Industry, Plastic Pipes, 5th International Conf. Univ. of York.

Su, Y.C., L.W. Mays, N. Duan and K.E. Lansey (1987) Reliability-Based Optimization Model for Water Distribution Systems, J. of Hyd. Div., ASCE, Vol. 102, No. 12, Dec. p. 1539-1555.

Syswerda, G. (1989) Uniform Crossover in Genetic Algorithms, Proc. of the third International Conf. on Genetic Algorithms. Morgan Kauffman, Los Altos, CA, p. 2-9.

Tangena, B.H. and P.K. Koster (1983) Reliability of Drinking Water Supply Systems, Proc. International Water Supply Assoc. Specialized Conf. on Security in Water Supply, International Water Supply Assoc., Brussels, Belgium.

Tanyimboh, T. and A.B. Templeman, (1993a) Calculating Maximum Entropy Flows in Networks, J. of Operational Research Society, Vol.44, No.4, p. 383-396.

Tanyimboh, T. and A.B. Templeman, (1993b) Maximum Entropy Flows for Single-Source Networks, Eng. Opt., Vol. 22, No. 1, p. 49-63.

Tanyimboh, T. and A.B. Templeman (1993c) Optimum Design of Flexible Water Distribution Networks, Civil Engineering Systems, Vol. 10, p. 243-258.

Tavallaee, A. (1974) Inclusion of Pumps, Pressure Reducing Valves and Reservoirs in Pipe Networks Solved by Linear Theory Method, M. S. Thesis, Dept. of Civil and Environmental Engineering, Utah State Univ.

Templeman, A.B. (1982) Discussion of 'Optimization of Water Distribution Systems' by Quindry et al., J. of Env. Eng. Div., ASCE, Vol. 108, No. 3, p. 599-602.

Templeman, A.B. (1989) Entropy and Civil Engineering Optimization, Proc. of N.A.T.O. Advanced study Institute, Heriot-Watt Univ., Edinburgh, Jul.

Templeman, A. B. and **L. Xingsi** (1985) Entropy Duals, Eng. Opt. Vol. 9, p. 107-120.

Templeman, A. B. and L. Xingsi (1987) A Maximum Entropy Approach to Constrained Non-Linear Programming, Eng. Opt., Vol. 12, No. 3, p. 191-205.

Theil, H. (1967) *Economics and Information Theory.* North-Holland, Amsterdam.

Thorley, A.R.D. (1991) *Fluid Transients in Pipeline Systems*. D. & L. George Ltd.

Thorley, A.R.D. and D.J. Wood (1986) Microcomputer Modelling of Flow Conditions in Pipe Distribution Systems, Hydrosoft 86-Hydraulic Engineering Software, Proc. of The 2nd International Conf. Southampton, UK, September, p. 741-468.

Tribus, M. (1969) *Rational Descriptions, Decisions and Designs*. Pergamon Press, New York.

Tung, Y.K. (1985) Evaluation of Water Distribution Network Reliability, In Hydraulic and Hydrology in the Small Computer Age, American Society of Civil Engineers, Vol. 1, p. 359-364.

Tung, Y.K. (1986) Model for Optimal-Risk Based Water Distribution Network Design, Water Forum '86', p. 1280-1284.

U.S. Army Corps of Engineers (1980) *Methodology for Areawide Planning*, Engineering Manual 1110-2-502, Washington DC.

Wagner, J.M., U. Shamir and D.H. Marks (1986) Reliability of Water Distribution Systems, R86-18, Ralph M. Parsons Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.

Wagner, J.M., U. Shamir and D.H. Marks (1988a) Water Distribution Reliability: Analytical Methods, J. of Water Resources Planning and Management, ASCE, Vol. 114, No. 3, p. 253-276.

Wagner, J.M., U. Shamir and D.H. Marks (1988b) Water Distribution Reliability: Simulation Methods', J. of Water Resources Planning and Management, ASCE, Vol. 114, No. 3, May, p. 276-294.

Walski, T.M. (1985) State-of-the-Art Pipe Network Optimization, Proc. of the Speciality Conf. On Computer Applications in Water Resources, ASCE, New York, p. 559-568.

Walski, T.M. and A. Pelliccia (1982) Economic Analysis of Water Main Breaks, J. of A.W.W.A., Vol. 74, N.3.

Walski, T.M. and R. Wade (1986) New York City Water Supply Infrastructure Study: Vol. III The Bronx and Queens, Technical Report El 87-9, U.S. Army Engineer Waterways Experiment Station , Vicksburg, Mississippi.

Walski, T.M., R. Wade, and W.W. Sharp (1987) New York City Water Supply Infrastructure Study: Vol. IV Staten Island, Draft Technical Report El 87, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

Walters, G.A. (1988) Optimal Design of Pipe Networks: A Review. Proc. First International Conf. on Computer Methods and Water Resources, Morocco, Vol. 2, Computational Hydraulics, Computational Mechanics Publications, Southampton, p. 21-31.

Walters, G.A. and R.G. Cembrowicz (1993) Optimal Design of Water Distribution Networks, in 'Water Supply Systems: State of the Art and Future Trends', ed. Cabrera E. and Martinez F., Computational Mechanics Publications, Southampton, U.K, p. 91-117.

Walters, G.A. and J. Knezevic (1988) The Influence of Reliability on the Layout and Design of Water Distribution Networks, Proc. of Symposium on Reliability-Based Design in Civil Engineering, Ecole Polytechnique Fédérale de Lausane, Switzerland, Vol. 1, July, p. 237-243.

Walters, G.A. and J. Knezevic (1989) Discussion on Su et al. 'Reliability-Based Optimisation Model for Water Distribution Systems', J. of Hyd. Eng., ASCE, Vol. 115, No. 8, Aug., p. 1157-1158.

Watanatada, T. (1973) Least-Cost Design of Water Distribution Systems, J. of Hyd. Div., ASCE, Vol. 89, p. 1497-1513.

Water Authority Association Advisory Committee (1982) Asbestos Cement Pipe and Fittings, Information and Guidance Note, 4-12-03.

Water Authority Association Advisory Committee (1983) Ductile Iron Pipes and Fittings, Information and Guidance Note, 4-21-01.

Water Research Centre (1977) Cost Information for Water Supply and Sewage Disposal, Technical Report TR 61.

Weiss, R.A, T.M. King, D.K. O'Day, S.K. Lion, J.F. Hood and N.E. Johnson (1985) Philadelphia Water Supply Infrastructure Study. U.S. Army Engineer District, Philadelphia, Pennsylvania.

Whitley, L. D. and J. Kauth (1988) GENITOR: A Different Genetic Algorithm, Proc. Rocky Mountain Conf. on Artificial Intelligence, Denver, Colo., p. 118-130.

Willie, R.R. (1978) Computer Aided Fault Tree Analysis, Operations Research Center, Univ. of California, Berkeley, California.

Wilson, A. G. (1970) *Entropy in Urban and Regional Planning*. Pion Ltd., London.

Wood, D.J. (1980) KYPIPE, Computer Analysis of Flow in Pipe Networks Including Extended Period Simulation, User's MANUAL, Office of Engineering Continuing Education, Univ. of Kentucky, Kentucky, KY 40506.

Wood, D.J. and C. Charles (1972) Hydraulic Network Analysis using Linear Theory, J. of Hyd. Div., ASCE, Vol. 98, No. HY7, July, p. 1157-1170.

Wood, D.J. and A.G. Rayes (1981) Reliability of Algorithms for Pipe Networks Analysis, J. of Hyd. Div., ASCE, Vol. 107, HY10, Oct., p. 1145-1161.

Wood, D.J. and J.E. Funk (1993) Hydraulic Analysis of Water distribution Systems, Part 1 Network Equations and Reliability Solutions, in 'Water Supply Systems: State of the Art and Future Trends', ed. Cabrera E. and Martinez F., Computational Mechanics Publications, Southampton, U.K, p. 43-68.

Yao, D.D. (1985) Material Flow and Information Flow in Flexible Manufacturing systems. Material Flow, 2, p. 143-149.

Yates, D.F., A.B. Templeman and T.B. Boffey (1984) The Computational Complexity of the Problem of Determining Least Capital Cost Design for Water Supply Networks, Eng. Opt., Vol. 7, p. 143-155.