Department of Civil and Environmental Engineering
University of Strathclyde

HEAD-DEPENDENT MODELLING AND OPTIMISATION OF WATER DISTRIBUTION SYSTEMS

Thesis submitted in accordance with the requirements of

the University of Strathclyde

for the degree of

Doctor in Philosophy

by

ALEMTSEHAY GEBREMESKEL SEYOUM

June 2015

Signed: ALEMTSEHAY GEBREMESKEL SEYOUM

Date: 08-06-2015

HEAD-DEPENDENT MODELLING AND OPTIMISATION OF WATER DISTRIBUTION SYSTEMS

Alemtsehay Gebremeskel Seyoum

ABSTRACT

The construction, operation and maintenance of water distribution systems (WDSs) involve a huge capital investment and it is essential to design and manage them in a cost effective way. The optimisation approaches (e.g. evolutionary algorithms) for optimal design or operation of WDSs require simulation models to evaluate the hydraulic and/or water quality performances of solutions to the problem. An accurate performance assessment of solutions (feasible and infeasible) is crucial to guide the search towards the optimal solution efficiently. Conventional models are demand driven and consequently they are incapable of simulating pressure-deficient (infeasible) solutions accurately. When simulating a pressure-deficient network, results produced by demand-driven analysis model are highly unreliable and misleading. This thesis is concerned with the development and application of a new integrated head-dependent hydraulic and water quality model for pressure-deficient network modelling and optimisation of real-world systems.

The original and novel aspects of the work carried out in this research are stated next.

A new pressure dependent analysis (PDA) model has been developed for pressure-deficient network modelling. The model is an enhanced version of the pressure-dependent extension of EPANET (EPANET-PDX) that has an embedded logistic nodal head-flow function. The novelty of the proposed PDA is the formulation of a new, more efficient and robust implementation of the line search and backtracking procedure that greatly enhances computational properties for highly pressure deficient networks; and increases the algorithm's consistency over a wider range of operating conditions. The model performed consistently well when simulating hypothetical and real-life networks under normal and pressure deficient conditions. Comparison between results generated by the new model and the original EPANET-

PDX demonstrated that the two models produce identical hydraulic results. From a numerical perspective, a significant reduction in numbers of iterations to complete a simulation has been obtained for all pressure operating conditions. Also, for extremely pressure deficient conditions a significant reduction in computational time has been achieved in comparison to the original EPANET-PDX.

Water quality modelling under pressure-deficient conditions is addressed for the first time in this thesis. Hydraulic and water quality analyses based on two water supply zones in the UK were conducted for a range of simulated operating conditions including normal and subnormal pressure and pipe closures using PDA. It is shown that operating conditions with subnormal pressures, if severe and protracted, can lead to spatial and temporal distributions of the water age and concentrations of chlorine and disinfection by-products that are significantly different from operating conditions in which the pressure is satisfactory.

A new parallel model for the solution of multi-objective WDSs optimisation problems is developed. The model utilises a multi-objective genetic algorithm that has an embedded PDA hydraulic simulator. The PDA takes into account the pressure dependency of the nodal flows and thus avoids the need for penalties to address violations of the nodal pressure constraints. A controller-worker approach is implemented to parallelise the optimisation process in which a single controller processor executes the routine operation of the algorithm and employs the workers to carry out fitness evaluation. A real-life network that comprises multiple sources, multiple demand categories, many fire flows and extended period simulation is used to demonstrate the effectiveness of the model. Results show that the algorithm is stable and finds optimal and near-optimal solutions reliably and efficiently.

ACKNOWLEDGEMENT

First, I praise God the Almighty for the completion of my PhD research. I could not have done this without His help.

I am extremely grateful to my supervisor, Dr. Tiku Tanyimboh, for his intellectual guidance, enthusiastic encouragement and continuous support during the course of the PhD work. His deep knowledge in the field of network analysis and optimisation and his passion for research has continuously inspired me. I feel extremely privileged to have Dr. Tanyimboh as my supervisor.

I would like to thank the University of Strathclyde, the British Government (Overseas Research Students Award Scheme) and UK Engineering and Physical Sciences Research Council (EPSRC grant reference EP/G055564/1) for their financial support to my PhD programme. I am also grateful to Veolia Water UK (now Affinity Water) for the support they provided during my research and the HPC team at the University of Strathclyde for their technical support and high performance and parallel computing courses.

I am very thankful to Dr. Lewis Rossman of the United States Environmental Protection Agency, who offered insights on the EPANET-PDX program, Dr. Calvin Siew who provided assistance on the EPANET-PDX and PF-MOEA programs at the start of my PhD research and Dr. Euan Barlow for his assistance on high performance and parallel computing. Special thanks are also given to Dr. Girma Zawdie for his continuous encouragement and Anna Czajkowska, Dr. Salah Saleh, Dr. Upaka Rathnayake and Fidelis Abagulum for being so supportive in many ways.

My deepest gratitude goes to my beloved mother Mulu Mekonnen, my dear brother Kifleyacob, sisters Alemach, Atsede, Fikirte, Zufan and Mebrat and all other family members for their love, support, continuous prayers and encouragement. Also, I want to thank Pastor Yohans Tsigeyohans, Letekidan Tesfu, Tadesse Hailemariam and Meseret Petros for being so kind and caring throughout my stay in Glasgow.

DEDICATION

This thesis is dedicated to the memory of my beloved father Gebremeskel Seyoum and sisters Berhe, Fitlework and Aster.

TABLE OF CONTENTS

Contents *Page*

LIST OF FIGURES

Chapter 5

Chapter 6

Appendix A

Appendix B

Appendix C

LIST OF TABLES

Appendix A

Appendix C

NOTATION

$\underline{\Delta H}^{(k)}$	vector of the nodal heads corrections
ΔQp	loop-flow correction
Δt	time step between the current and the next steady state simulation
$\Delta' t$	time step from the current time to the time the status change occurs
ΔV_m	change of water volume in tank m
∇g	gradient of g
α^e	parameter used in Hazen-Williams equation
α_i	parameter to be calibrated in logistic pressure-dependent demand function
β^e	parameter used in Hazen-Williams equation
β_i	parameter to be calibrated in logistic pressure-dependent demand function
λ	over-relaxation coefficient for the line search and backtracking routine
$\lambda_{\min}$	minimum over-relaxation coefficient for the line search and backtracking routine
ω	dimensionless conversion factor used in the Hazen-Williams head loss equation (equals to 10.67 in S.I. units)
δHn	full Newton step for the nodal heads
δQp	full Newton step for the pipe flows
A_{10}	incidence matrix relating the pipes to nodes with known heads
A_{11}	diagonal matrix head loss in links
A_{12}	incidence matrix relating the pipes to nodes with unknown heads
A_{21}	the transpose of A_{12}

A_{22}	diagonal matrix whose elements are $Qn(Hn)/Hn$
B_m	set of links connected to tank m
b_i	node specific constant
C	reactant concentration in bulk flow
C_{cl}	chlorine concentration
C_{ext}	external source concentration entering into node
C_{ij}	Hazen-Williams roughness coefficient for pipe ij
C_L	limiting trihalomethane concentration
C_s	concentration within storage facility
C_{thm}	trihalomethane concentration
CR	cost ratio
CV	coefficient of variation
c_i	node specific constant
D_{11}	derivative of A_{11} and its elements
D_{22}	derivative of A_{22} and its elements
D_{ij}	Diameter of link ij
DBP	disinfection by-product
DDA	demand driven analysis
DF	demand factor
DSR	demand satisfaction ratio
DVM	discrete volume method
d_i	distance between the i^{th} non-dominated solution and the nearest solution of the optimal Pareto set
dHn	corrective steps of nodal head
dQp	corrective steps of pipe flow
EA	evolutionary algorithm
EPS	extended period simulation

$\underline{F}$	vector of respective functions of the nodal continuity equations
F_1	first objective function
F_2	second objective function
FEs	number of function evaluations
FSS	full solution space
f_1	pipe head loss conservation equation
f_2	flow conservation equation
f_{ij}	Darcy-Weisbach coefficient of friction in pipe ij
$f_m^{(i)}$	value of the m^{th} objective function for the i^{th} non-dominated solution
$f_m^{(k)}$	value of the m^{th} objective function for the k^{th} Pareto optimal solution
GA	genetic algorithm
GD	generational distance
GGM	global gradient method
GGA	global gradient algorithm
GIS	geographical information system
g	function to be minimised
$\underline{H}$	vector of unknown nodal heads
HGL	hydraulic gradient level
Hn	nodal head
Hn_i	head at node i
Hn_j	head at node j
$H_m(t)$	water level of tank m at time t
H_0	known nodal heads
HPC	high performance computer
H_s	head available at the source
$H_s^{\min}$	source head at which the most critical node of the network begins to deliver water

H_s^{des}	required source head to fully satisfy all the demand nodes
$H_{s,j}^{des}$	the head at the source above which the demand at node j is fully satisfied
$Hn_i^{\min}$	minimum required head at node i
Hn_i^{des}	desired head to satisfy the demand at node i
h_{ij}	head loss in pipe ij
IJ	set of all the links in the network
IJ_{lp}	set of all links in loop lp
ILP	integer linear programming
I_q	set of links with flow into node q
I_s	set of links providing flow into the facility
J	Jacobian
K_{ij}	resistance coefficient for link ij
k	iteration number
k_b	reaction rate constant in bulk flow
k_f	mass transfer coefficient
k_w	wall reaction rate constant
L_{ij}	length of pipe
l_{ij}	length of link ij
lp	loop
lp_{ij}	loops sharing link ij
Nn	number of nodes
N_p	population size
Nl	number of links
Nlp	number of loops in a network
NP	number of solutions in the Pareto front

NS	number of non-dominated solutions
n	reaction order
n_f	flow exponent
n_{ij}	Manning roughness coefficient
n_j	nodal exponent whose value ranges between 1.5 and 2
MOGA	multi-objective genetic algorithm
MPI	Message Passing Interface
O_s	set of links receiving flow from the facility
PDA	pressure dependent analysis
PFMOEA	Penalty-Free Multi-Objective Evolutionary Algorithm
Pr_i	pressure at node i
$\mathrm{Pr}_r^{\#}$	nodal pressure at which a given proportion of the demand is supplied
p_c	crossover probability
p_m	mutation probability
Q_{ext}	external source flow entering into node
Qn_i	flow at node i
$Qn_i^{base}(t)$	base demands for node i at time t
Qn_i^{req}	required supply at node i
Qn_j^{req}	required supply at node j
Qp	pipe flow
Qp_{ij}	pipe flow in link ij
$Qp_b(t)$	flow rates of links connected to tank m at time t
$QT_m(t)$	flow rates of tank m at time t
Q_s^{avl}	sum of all available nodal flows
Q_s^{req}	total nodal demands
RSS	reduced solution space

$r\,(C\,)$	rate of reaction as a function of concentration
r_h	hydraulic radius of pipe
R_s	Pipe Resistance constant
u	mean flow velocity
$V_m(t)$	water volumes of tanks at time t
V_s	volume in storage facility
v	pipe velocity
$v_{\max}$	maximum permissible velocity
t	time
THM	trihalomethane
WDS	Water Distribution System
WDN	Water Distribution Network

Chapter 1

Introduction

1.1 Background

Water distribution systems (WDSs) are key components of public infrastructures and it is essential to design and rehabilitate them in a cost effective manner without compromising the required performance and regulatory standards. Evolutionary algorithms (EAs) are a commonly applied approach to solve WDSs optimisation problems. The algorithms are often coupled with simulation models to evaluate the hydraulic and/or water quality performances of solutions to the problem. The simulation models evaluate the conservation of mass and energy constraints and check other constraints such as nodal pressures and water quality for any violation. However, EAs by nature generate both feasible and infeasible solutions. Evidence from literature on constraint or infeasible solutions handling approaches demonstrated the benefits of explicitly maintaining infeasible solutions for single and multi-objective constrained optimization problems (e.g., Singh et al., 2008; Ray et al., 2009). The presence of infeasible solutions guides the algorithm search towards an optimal solution from both feasible and infeasible side of the search space. The recent WDSs optimization algorithms that retain infeasible solutions (pressure-deficient designs) in full achieved superior results in terms of algorithms' convergence rate or quality of solutions

compared to those algorithms that reject or penalize infeasible solutions (Siew and Tanyimboh, 2012b; Saleh and Tanyimboh, 2013; Siew et al., 2014). This has given rise to the need of an efficient simulation model for evaluating the hydraulic and water quality performances of infeasible designs (pressure-deficient networks) more accurately.

Hydraulic and water quality simulations of WDSs can be performed under time-varying conditions by employing extended period simulation (EPS) models. The models include important time-varying features such as water levels in tanks, nodal demands and the scheduling of pumps. Conventional EPS models are demand driven and thus assume that all demands are fully satisfied. In reality, several events make WDSs pressure deficient. Unavailability of system components for maintenance or rehabilitation purposes, pipe bursts, pump failures and large increases in demand (e.g. for firefighting purposes) are some of the pressure-deficient condition that WDSs often experience (Tanyimboh et al., 1999; Kalungi and Tanyimboh, 2003; Siew and Tanyimboh, 2012a). During these events, not all demands can be satisfied in full and thus, the hydraulic and water quality performance of WDSs is reduced. When predicting the behaviour of a pressure-deficient system, results produced by demand-driven analysis (DDA) are highly unreliable and misleading (Tanyimboh and Templeman, 2010; Siew and Tanyimboh, 2012a). Consequently, EPS models based on DDA cannot simulate the performance of a pressure-deficient network realistically. In the aspect of evolutionary optimization of WDSs, accurate performance evaluation of solutions is highly essential in order to guide the search towards the optimal solutions efficiently. In this regard, the pressure dependent analysis (PDA) models that take into account the pressure dependent nature of nodal flows provide results that are more realistic (Chandapillai, 1991; Gupta and Bhave, 1996; Tanyimboh et al., 1999; Tabesh et al., 2009).

One of the difficulties in using EAs for WDSs optimisation problems is that the algorithms require huge computational time when applied to large optimisation problems such as real-life networks with large numbers of pipes and multiple operating

conditions (VenZyl et al., 2004). In the optimisation of large water distribution systems, for instance, a single optimisation run may involve thousands of hydraulic and water quality simulations (e.g Ghebremichael et al., 2008; Seyoum and Tanyimboh, 2014) and that may take many days on modern computers such as workstations. Such computational time, however, is usually unacceptable for water utilities applications. This has limited the algorithms' potential for practical applications to solve real-world WDSs optimisation problem. One way of addressing this difficulty is by utilising high performance and parallel computing approaches that are capable of reducing the computational time and makes the algorithms' convergence to optimal solutions faster. Despite the inherent characteristics of EAs as being easily parallelised, only a handful of WDS optimisation research has been carried out to investigate EAs parallel implementations (Balla and Lingireddy, 2000; Wu and Zhu, 2009; Ewald et al., 2008; Artina et al., 2012; Barlow and Tanyimboh, 2014).

The overall aim of this research is to help address the above-mentioned weaknesses, and in particular, to develop a fast and practical tool that can be utilised for WDNs analysis and optimisation effectively and efficiently.

1.2 Objectives of the Research

The objectives of this research are as follows:

1. To develop a computationally efficient PDA model for pressure-deficient network modelling

2. To study the spatial and temporal distribution of water quality under normal and abnormal pressure operating conditions in WDSs and make recommendations

3. To develop an efficient EA based model to solve computationally intensive WDSs optimisation problems

4. To assess the practical capability, robustness and computational efficiency of both the PDA and EA optimisation model by applying them to hypothetical and real-life networks

1.3 A Brief Description of the Methodology

The thesis presents a new PDA model for pressure-deficient network modelling. The model is an enhanced version of the pressure-dependent extension of EPANET i.e., EPANET-PDX (Siew and Tanyimboh, 2012a). EPANET-PDX integrates a logistic nodal head-flow relationship (Tanyimboh and Templeman, 2010) into EPANET hydraulic simulator coupled with a line search and backtracking procedure (Press et al., 1992) to facilitate convergence. The novelty of the proposed PDA is the formulation of a new, more efficient and robust implementation of the line search and backtracking procedure that greatly enhances computational properties for highly pressure deficient networks; and increases the algorithm's consistency over a wider range of operating conditions.

The thesis also presents a new parallel model for the solution of multi-objective WDSs optimisation problems. The new optimisation model utilises a multi-objective genetic algorithm (MOGA) that has an embedded PDA hydraulic simulator (Siew and Tanyimboh, 2012b). The MOGA is used to obtain least cost designs via optimization while PDA is used to evaluate the hydraulic performance and feasibility of each network design. The network hydraulic performance is formulated as an objective in the MOGA. The optimisation process is parallelised using a controller-worker approach where a single controller processor executes the routine operation of the algorithm and employs

the workers to carry out fitness evaluation. The fitness evaluation involves extended period simulations.

1.4 Layout of Thesis

The thesis is structured as follows:

Chapter 2 contains a review of water distribution networks (WDNs) hydraulic and water quality analysis. The chapter provides a detailed review on the fundamentals required in the modelling of WDNs along with the two WDNs analysis approaches i.e., demand driven analysis and pressure dependent analysis.

Chapter 3 presents a review of the existing approaches to enhance the efficiency and effectiveness of evolutionary algorithms (EAs), in particular genetic algorithm (GA) to solve WDSs optimisation problems. The chapter discusses the most common parallelization approaches that have been applied in WDS literature to improve the GA's computational efficiency.

Chapter 4 proposes a new algorithm for performing pressure-dependent modelling of WDNs. The chapter describes the method in detail. It also presents results obtained from the simulations of hypothetical and real life networks that verify the accuracy, robustness and computational efficiency of the algorithm.

Chapter 5 focuses on investigating the spatial and temporal distribution of water quality in WDNs under various pressure operating conditions. The chapter presents water quality results generated from a PDA model based on hypothetical and real-life networks. Comparison and verification of results are also presented.

Chapter 6 proposes a parallel optimisation model that utilises MOGA that has an embedded PDA. The parallel model is described in detail and applied to real-life network to demonstrate its computational efficiency and robustness. The comparisons of results generated by the parallel and the serial models are presented. The chapter also presents the extensive investigation carried out on the serial MOGA using a network in literature. Solutions that are both fully feasible and cheaper than the best-known solutions in the literature are presented and discussed.

Chapter 7 summarises the main ideas and conclusions from the research. Some suggestions for further research are discussed.

Chapter 2

Review of Water Distribution Network Analysis

2.1 Introduction

Water distribution models have become standard engineering tools of water utilities applications such as design, calibration, rehabilitation and operation. The models simulate the dynamics of the systems under a wide range of conditions (e.g. peak demands, pipe bursts, pump failure, fire flows). Using water distribution network (WDN) models, problems can be predicted and solutions can be evaluated prior to investing resources in real-world projects (Walski et al., 2003). The information provided by the models is valuable in assisting engineers in making appropriate and timely decisions. For instance, models can help in identifying potential problem areas where water quality is an issue. The concentration of disinfection and disinfection by-products throughout a network can be analysed (Ghebremichael et al., 2008; Seyoum and Tanyimboh, 2014); the impact of storage tanks on water quality can be evaluated (Clark and Grayman, 1998; Seyoum et al., 2014).

The analysis of WDNs can be of either steady state or extended period. Steady state analysis that represents a snapshot in time is used to determine the operating behaviour of a system under equilibrium conditions where nodal demands and water level of

storage reservoirs are constant over time. This can be useful in analysing the short-term effect of fire flows or average demand conditions on the WDN. The results of steady state analysis are instantaneous and may or may not be representative of the performance of the system over time. On the other hand, extended period analysis is used to evaluate WDN performance over time. This is a more realistic analysis as the operation of real-life WDNs varies with time. Extended period analysis is suitable in modelling tank filling and draining, regulating the operation of valves, and the changes in pressure and flow rates due to variation in demand.

WDN analysis model is developed by using a node-link formulation governed by the conservation of mass at nodes and conservation of energy around hydraulic links. Node represents the connection point of two or more links or the end-point of a single link. The governing hydraulic equations are solved numerically to obtain nodal heads and flow rates in links. The model results are then assessed in reference to the minimum performance standards (e.g. minimum node pressure requirement) to verify the feasibility of the solution.

There are two network analysis approaches in literature. The conventional approach, known as demand-driven analysis (DDA), is formulated on the basis that demands are fully satisfied irrespective of the network pressure. This assumption is valid only if the network performs under normal pressure conditions with adequate pressures at all nodes (Wu et al., 2009). However, in the presence of network irregularities such as pipe breaks, pump failures, temporary demand increase e.g. for fire fighting purposes, and system maintenance and repair restrictions, DDA provides unrealistic results (Tanyimboh and Templeman, 2010; Siew and Tanyimboh, 2012a). By contrast, the pressure dependent analysis approach (PDA) takes into account the pressure dependent nature of nodal flows and thus the model provides results that are more realistic (Chandapillai, 1991; Gupta and Bhave, 1996; Tanyimboh et al., 1999).

The primary objective of this chapter is to review the fundamentals involved in the modelling of WDNs. The next section (Section 2.2) presents the governing hydraulic equations for WDN analysis. Section 2.3 provides a review on the various numerical methods that are commonly used to solve DDA. The research carried out herein involves PDA implementation and thus, different PDA approaches have been discussed and their limitations highlighted in Section 2.4. Section 2.5 presents a review on water quality modelling approach of WDSs. Besides the PDA implementation, water quality modelling is one of the major parts of the research herein.

2.2 Governing Hydraulic Equations

Conservation of mass and energy are the two governing equations in the hydraulic modelling of the WDN. The principle of the conservation of mass is expressed using continuity equation while the principle of energy is expressed using the head loss equation. The governing equations have to be satisfied in network analysis. In order to satisfy the conservation of mass, the sum of nodal inflows and outflows in a network must be zero. The flow continuity equation for node j, $j = 1,\ldots, Nn$, is described as

$$\sum_{i:Hn_i>Hn_j} Qp_{ij} - \sum_{i:Hn_i<Hn_j} Qp_{ij} = Qn_j^{req} \tag{2.1}$$

where Nn is the number of nodes in the network; Qp_{ij} is the pipe inflow (if $Hn_i > Hn_j$) or outflow (if $Hn_i < Hn_j$) at node j; Qn_j^{req} is the demand (required supply) at node j; Hn_i and Hn_j are heads at nodes i and j respectively.

In order to fulfil the conservation of energy in a network, the sum of head losses in pipes forming each loop must be equal to zero.

$$\sum_{ij \in IJ_{lp}} h_{ij} = 0 \qquad lp = 1,...,Nlp \tag{2.2}$$

where IJ_{lp} represents the set of all links in loop lp. Nlp is the total number of loops in a network. The conservation of energy for a given path in the network can be written as Eq. 2.3, where the total head loss along the path should be equal to the difference in head between its starting and ending nodes. The equation for the path starting from node i with hydraulic gradient level (HGL) Hn_i and ending at node j with HGL Hn_j can be written as

$$\sum_{ij \in IJ_p} h_{ij} = Hn_i - Hn_j \qquad \forall ij \in IJ \tag{2.3}$$

where IJ is the set of all links in path p.

The pipe frictional head loss in Eq. 2.2 and Eq. 2.3 can be calculated using three widely used formulae (Bhave and Gupta, 2006). These are Hazen-Williams formula, Darcy-Weisbach formula, and Manning formula given by Eq. 2.4, Eq. 2.5 and Eq. 2.6 respectively.

$$h_{ij} = \frac{\omega L_{ij} Qp_{ij}^{1.852}}{C_{ij}^{1.852} D_{ij}^{4.87}} \qquad \forall ij \tag{2.4}$$

$$h_{ij} = \frac{8 f_{ij} L_{ij} Qp_{ij}^{2}}{\pi^2 g D_{ij}^{5}} \qquad \forall ij \tag{2.5}$$

$$h_{ij} = \frac{10.29 L_{ij} (n_{ij} Qp_{ij})^2}{D_{ij}^{5.333}} \qquad \forall ij \tag{2.6}$$

where ω is a dimensionless unit conversion factor and is equal to 10.67 in SI units; D_{ij}, h_{ij}, L_{ij}, and Qp_{ij} represent diameter, head loss, length and flow rate for pipe ij; C_{ij} and n_{ij} are Hazen-Williams roughness coefficient and Manning roughness coefficient respectively; f_{ij} is the coefficient of friction in pipe ij.

The frictional head loss equations (Eq. 2.4 - 2.6) are often expressed by a general head loss formula

$$h_{ij} = K_{ij} Qp_{ij}^{n_f} \qquad \forall ij \qquad (2.7)$$

where K_{ij} and n_f represent pipe resistance coefficient and flow exponent respectively. The value of n_f *is* 1.852 for Hazen-Williams formula and 2 for Darcy-Weisbach and Manning formulae. Eq. 2.8, Eq. 2.9 and Eq. 2.10 express the resistance coefficients of Hazen-Williams formula, Darcy-Weisbach formula, and Manning formula respectively.

$$K_{ij} = \frac{\omega L_{ij}}{C_{ij}^{1.852} D_{ij}^{4.87}} \qquad \forall ij \qquad (2.8)$$

$$K_{ij} = \frac{8 f_{ij} L_{ij}}{\pi^2 g D_{ij}^{5}} \qquad \forall ij \qquad (2.9)$$

$$K_{ij} = \frac{10.29 n_{ij}^{2} L_{ij}}{D_{ij}^{5.333}} \qquad \forall ij \qquad (2.10)$$

2.3 Demand Driven Network Analysis

The most common type of water distribution network (WDN) analysis is considering the demands as fully satisfied at all demand nodes. This conventional network analysis approach (Demand driven analysis) is well established and widely used in water industries for simulating WDN performance under normal operational conditions. The analysis of WDN can be of either steady state or extended-period depending upon whether the analysis determines the behaviour of a network at a given point of time or over a certain interval of time.

2.3.1 Steady State Analysis

In Steady state analysis, nodal demands and reservoir water levels are presumed constant (Bhave and Gupta, 2006). These assumptions are valid for a very short period to analyse, for instance, specific worst-case situations where the effects of time are insignificant. These include events such as peak demand times, fire protection usage, and system component failures. Steady state analysis enables engineers to predict the response of a network to a specific set of hydraulic conditions. During the analysis, nodal demands, pipe lengths, diameters and roughness are often pre-determined, while pipe head losses, nodal heads and pipe flow rates are unknown parameters (variables) and obtained from the analysis. The constitutive equations formulated as systems of equations using any of the variables as the basic unknown parameters.

2.3.1.1 Formulation of Hydraulic Equations

The analysis of WDN involves formulating systems of hydraulic equations that can be solved iteratively by employing numerical methods. There are various ways of

formulating the system of equations. When the hydraulic equations formulated with the pipe flow rates (Qp_{ij}) as the basic unknown variables, the equations formed are known as q-equations (Bhave, 1991). The unknown parameters, for example, for the flow continuity (Eq. 2.1) and head loss equations (Eq. 2.7) are Qp_{ij}. The hydraulic equations formulated with the nodal heads as unknown variables are known as H-equations (Bhave, 1991). For example, by using Eq. 2.3 and Eq. 2.7, the flow continuity equation (Eq. 2.1) can be rewritten as

$$\sum_{i:\ Hn_i > Hn_j} \left(\frac{Hn_i - Hn_j}{K_{ij}} \right)^{1/n_f} - \sum_{i:\ Hn_i < Hn_j} \left(\frac{Hn_j - Hn_i}{K_{ij}} \right)^{1/n_f} - Qn_j^{req} = 0;\ \ \forall j = 1,...,Nn \qquad (2.11)$$

To solve this system of equations, the number of unknown nodal heads must be equal to the number of continuity equations (Eq. 2.11). Also, the value of one nodal head should be known, which is generally taken as the fixed head at source node.

Equations formulated by considering loop-flow corrections as the basic unknowns are known as $\Delta Qp-$ equations (Bhave, 1991). These equations are set up by making an initial assumption of the pipe flow rates to satisfy nodal flow continuity (Eq. 2.1). However, the assumed pipe flow rates will not generally fulfil the loop-head loss relationship of Eq. 2.2. Thus, the loop-flow corrections are applied at each loop in order to adjust the pipe flow rates iteratively.

$$Qp_{ij}^{k} = Qp_{ij}^{k-1} + \sum_{lp \in lp_{ij}} \Delta Qp_{lp}^{k} \qquad \forall ij \in Nl \qquad (2.12)$$

in which ΔQp_{lp}^{k} represents the loop-flow correction applied for all pipe flows in loop lp according to the flow direction; Qp_{ij}^{k-1} is an estimated flow rate and Qp_{ij}^{k} is the corrected

flow rate. k represents the iteration number. lp_{ij} represents all loops sharing link ij. $\sum_{lp \in lp_{ij}} \Delta Qp_{lp}^{k}$ represents the summation of the corrections of all loops to which link ij belongs. By applying the loop-head loss relationship (Eq. 2.2), the $\Delta Qp-$equations can be formulated as

$$\sum_{ij \in IJ_{lp}} K_{ij} (Qp_{ij}^{(k-1)} + \sum \Delta Qp_{lp}^{(k)})^{n_f} = 0 \qquad \forall ij \in Nl; \qquad lp = 1,..., Nlp \tag{2.13}$$

where Nlp and Nl represent the number of loops and links in the network respectively. Eq. 2.13 produces a system of $\Delta Qp-$equations that can be simultaneously solved using an iterative method.

2.3.1.2 Solution of the Hydraulic Equations

Several numerical methods have been widely used to solve the conventional demand driven network analysis problem. These are Hardy-Cross method (Cross, 1936), Newton-Raphson method (Martin and Peters, 1963), Linear Theory method (Wood and Charles, 1972) and Global Gradient method (Todini and Pilati, 1988). The methods solve a system of non-linear equations iteratively starting with an initial trial solution. The newly obtained solution is compared with the trial solution and the procedure is repeated until the difference between successive solutions is less than a user specified value.

2.3.1.2.1 Hardy-Cross Method

The Hardy Cross method (Cross, 1936) is an iterative procedure for network analysis. The approach is based on loop-flow correction equations, i.e., $\Delta Qp-$equations. It

involves assuming initial pipe flow rates for a loop to satisfy the flow continuity. The procedure is formulated as

$$\sum_{ij \in IJ_{lp}} K_{ij} (Qp_{ij}^{(k-1)} + \sum \Delta Qp_{lp}^{(k)})^{n_f} = 0 \qquad \forall lp \tag{2.14}$$

Applying the first order Taylor's series expansion to Eq. 2.14 gives

$$\sum_{ij \in IJ_{lp}} K_{ij} (Qp_{ij}^{(k-1)})^{n_f} + \Delta Qp_{lp}^{(k)} \sum_{ij \in IJ_{lp}} \left| n_f \cdot K_{ij} (Qp_{ij}^{(k-1)})^{n_f - 1} \right| = 0 \qquad \forall lp \tag{2.15}$$

The loop-flow correction values can be obtained by rearranging Eq. 2.15 as follows:

$$\Delta Qp_{lp}^{(k)} = -\frac{\sum_{ij \in IJ_{lp}} K_{ij} (Qp_{ij}^{(k-1)})^{n_f}}{\sum_{ij \in IJ_{lp}} \left| n_f . K_{ij} (Qp_{ij}^{(k-1)})^{n_f - 1} \right|} \qquad \forall lp \tag{2.16}$$

The effect of adjacent loops on the loop-flow correction is neglected and thus each loop-flow correction equation (Eq. 2.16) contains only one variable i.e. the $\Delta Qp_{lp}^{(k)}$ value as unknown. The loop-flow correction obtained from Eq. 2.16 is used to update the pipe flow rates as follows:

$$Qp_{ij}^{(k)} = Qp_{ij}^{(k-1)} + \Delta Qp_{lp}^{(k)} \qquad \forall lp \ ; \qquad \forall ij \in IJ_{lp} \tag{2.17}$$

The first iteration ends once the loop-flow corrections are calculated for all loops and the pipe flow rates are all updated. The next iteration involves using the updated pipe flow rates as the new estimates for pipe flow rates in Eqs. 2.14 to 2.17. The process is repeated until pre-specified convergence criteria are satisfied (e.g. when the loop-flow correction values become very small).

2.3.1.2.2 Newton-Raphson Method

The Newton-Raphson method is a powerful numerical method for solving systems of non-linear equations. The method first proposed by Martin and Peters (1963) for the solution of WDN. The Newton-Raphson method can be formulated for a single variable non-linear function, $F(x) = 0,$ as follows:

$$x^{k+1} = x^k - \frac{F(x^{(k)})}{F'(x^{(k)})} \tag{2.18}$$

where $F'(x^{(k)})$ is the derivative of $F(x)$ evaluated at x^k. For a system of equations, Eq. 2.18 can be written as:

$$\underline{x}^{(k+1)} = \underline{x}^k - (J_x)^{-1}_{(k)} \underline{F}(\underline{x}^{(k)}) \tag{2.19}$$

where $\underline{x}$ is the vector of the variables; $\underline{F}$ is the vector containing function values; and J_x is the Jacobian or the matrix of the first partial derivatives of each F with respect to each $x's$. To apply the Newton Raphson method to the WDN analysis, the continuity equations at the nodes (Eq. 2.11) can be written as

$$\underline{F}(\underline{H}) = 0 \tag{2.20}$$

where $\underline{F}$ is the vector of the respective values of the nodal continuity equations and $\underline{H}$ is the vector of unknown nodal heads. From Eq. 2.19, the Newton Raphson formulation for the flow continuity can be expressed as

$$\underline{H}^{(k+1)} = \underline{H}^k - (J_H)^{-1}_{(k)} \underline{F}(\underline{H}^{(k)}) \tag{2.21}$$

where J_H is the Jacobian matrix for the unknown nodal heads. Eq. 2.21 can be expressed in terms of the vector of the nodal head corrections ($\underline{\Delta H}^k = \underline{H}^{k+1} - \underline{H}^k$) as follows:

$$(J_H)_{(k)} \underline{\Delta H}^{(k)} = -\underline{F}(\underline{H}^{(k)}) \tag{2.22}$$

The systems of equations in Eqs. 2.22 are solved simultaneously to obtain $\underline{\Delta H}^k$ which is then used to update $\underline{H}^{k+1}$. The iterative process ends when a pre-specified convergence criteria is attained for the value of $\underline{\Delta H}^k$ or $\underline{F}(\underline{H}^k)$.

2.3.1.2.3 Linear Theory Method

The Linear Theory method is another network analysis method that is developed by Wood and Charles (1972). In this method, there is no need for making initial estimates for flows that satisfy the nodal flow equation, although this is fulfilled after the first iteration. As described earlier, the continuity equations are linear but the head loss equations are non-linear. In the Linear Theory method, the head loss equations are linearized by merging the non-linear term with the pipe resistance coefficient as follows:

$$h_{ij}^{(k)} = \left| K_{ij} (Qp_{ij}^{(k-1)})^{n_f - 1} \right| \cdot Qp_{ij}^{(k)} \qquad \forall ij; \qquad k = 1,2 \tag{2.23a}$$

$$h_{ij}^{(k)} = \left[K_{ij} \left(\frac{Qp_{ij}^{(k-1)} + Qp_{ij}^{(k-2)}}{2} \right)^{n_f - 1} \right] \cdot Qp_{ij}^{k} \qquad \forall ij; \qquad k = 3,4,5... \tag{2.23b}$$

Eqs. 2.23 are used to form a set of loop-head loss and nodal flow continuity equations. The system of linear equations formed is solved simultaneously in an iterative manner to give an approximated pipe flow rates. Eq. 2.23a is used for the first and second iteration where the initial pipe flow rate (Qp_{ij}^{0}) is set to unity. For successive iterations Eq. 2.23b is used in which pipe flow rates Qp_{ij}^{k} are obtained by using the average of the flow rates obtained from previous iterations (i.e. $Qp_{ij}^{(k-1)}$ and $Qp_{ij}^{(k-2)}$). The iterative process is continued until the difference between two successive estimates of pipe flow rates (i.e. $Qp_{ij}^{(k-1)}$ and $Qp_{ij}^{(k-2)}$) is within a specified tolerance.

2.3.1.2.4 Global Gradient Method

Todini and Pilati (1988) proposed the Global Gradient Method (GGM), which is the application of the Newton-Raphson method to obtain both pipe flow rates and nodal heads simultaneously. The Newton-Raphson method solves for the corrections of either pipe flow rates or nodal heads. GGM directly obtains the improved pipe flow rates and nodal heads in an iterative procedure that is continued until no further improvement is observed. GGM is similar to the Linear Theory method in that it does not require satisfying the continuity equations at all nodes to start the solution procedure.

The GGM takes pipe flow rates and nodal heads as the basic unknowns to formulate the Q-H equations. The non-linear head loss equations for k^{th} iteration can be written as:

$$Hn_i^k - Hn_j^k - K_{ij}\left|Qp_{ij}^k\right|^{n_f-1} Qp_{ij}^k = 0 \qquad \forall ij \tag{2.24}$$

Also, the linear nodal flow continuity equations can be written as:

$$\sum_{ij\in j} Qp_{ij}^{k} + Qn_{j}^{req} = 0 \qquad j = 1,.....,Nn \tag{2.25}$$

Todini and Pilati (1988) represent the governing Eqs. 2.24 and 2.25 in the following matrix notation:

$$\begin{bmatrix} \mathbf{A}_{11} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Qp} \\ \cdots \\ \mathbf{Hn} \end{bmatrix} = \begin{bmatrix} -\mathbf{A}_{10}\mathbf{H}_{0} \\ \cdots \\ -\mathbf{Qn}^{\mathbf{req}} \end{bmatrix} \tag{2.26}$$

where $\mathbf{A_{11}}$ is a diagonal matrix whose elements are $K_{ij}(Qp_{ij})^{n_f-1}$. K_{ij} and n_f are the resistance coefficient and flow exponent in the head loss formula respectively. Qp_{ij} is the flow rate in pipe ij. $\mathbf{A_{12}}$ and $\mathbf{A_{10}}$ represent the overall incidence matrix relating the pipes to nodes with unknown and known heads respectively. Pipe flow leaving node is defined as -1, pipe flow into node as +1 and 0 if pipe is not connected to node. $\mathbf{A_{21}}$ is the transpose of $\mathbf{A_{12}}$. $\mathbf{Qp}$ denotes the column vector of unknown pipe flow rates. $\mathbf{Hn}$ and $\mathbf{H}_0$ are column vectors for unknown and known nodal heads respectively. $\mathbf{Qn^{req}}$ is the column vector for required nodal supplies.

The GGM representation of the WDN problem in Eq. 2.26 can be rearranged as follows:

$$\begin{bmatrix} \mathbf{A}_{11} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Qp} \\ \cdots \\ \mathbf{Hn} \end{bmatrix} + \begin{bmatrix} \mathbf{A}_{10}\mathbf{H}_{0} \\ \cdots \\ \mathbf{Qn}^{\mathbf{req}} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1 \\ \cdots \\ \mathbf{f}_2 \end{bmatrix} \tag{2.27}$$

where $\mathbf{f}_1 = \mathbf{f}_1(\mathbf{Qp}, \mathbf{Hn})$ and $\mathbf{f}_2 = \mathbf{f}_2(\mathbf{Qp})$ indicate how far from zero the relevant equations are for any given approximated solution $\mathbf{Qp}$, $\mathbf{Hn}$. Applying the Newton-Raphson method directly to this system results in the following set of equations for the flow and head corrections at iteration k of the process:

$$\begin{bmatrix} \mathbf{D}_{11}^{k} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{dQp} \\ \cdots \\ \mathbf{dHn} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1^k \\ \cdots \\ \mathbf{f}_2^k \end{bmatrix} \tag{2.28}$$

$\mathbf{D}_{11}$ is a diagonal matrix that represents the Jacobian of $\mathbf{A}_{11}\mathbf{Qp}$. Its diagonal element is $nK_{ij}(Qp_{ij})^{n_f-1}$. **dQp** and **dHn** represent the corrective steps of **Qp** and **Hn** respectively in successive iterations and can be defined as

$$\mathbf{dQp} = \mathbf{Qp}^{k} - \mathbf{Qp}^{k+1} \tag{2.29}$$

$$\mathbf{dHn} = \mathbf{Hn}^{k} - \mathbf{Hn}^{k+1} \tag{2.30}$$

From Eq. 2.27, $\mathbf{f}_1$ and $\mathbf{f}_2$ can be written as:

$$\mathbf{f}_1^k = \mathbf{A}_{11}^k \mathbf{Qp}^k + \mathbf{A}_{12}\mathbf{Hn}^k + \mathbf{A}_{10}\mathbf{H}_0 \tag{2.31}$$

$$\mathbf{f}_2^k = \mathbf{A}_{21}\mathbf{Qp}^k + \mathbf{Qn}^{\mathbf{req}} \tag{2.32}$$

By substituting Eq. 2.31 and Eq. 2.32 into Eq. 2.28, the iterative formulation of the GGM can be described as the following two equations.

$$\mathbf{Hn}^{k+1} = \left[\mathbf{A}_{21}(\mathbf{D}_{11}^k)^{-1}\mathbf{A}_{12}\right]^{-1}\left[\mathbf{A}_{21}\mathbf{Qp}^k - \mathbf{Qn}^{req} - \mathbf{A}_{21}(\mathbf{D}_{11}^k)^{-1}(\mathbf{A}_{11}^k\mathbf{Qp}^k + \mathbf{A}_{10}\mathbf{H}_0)\right] \tag{2.33}$$

$$\mathbf{Qp}^{k+1} = \mathbf{Qp}^k - (\mathbf{D}_{11}^k)^{-1}(\mathbf{A}_{11}^k\mathbf{Qp}^k + \mathbf{A}_{12}\mathbf{Hn}^{k+1} + \mathbf{A}_{10}\mathbf{H}_0) \tag{2.34}$$

To start the solution procedure, an initial guess of pipe flow rates is required. The initial pipe flow rate can be taken as unity or any arbitrarily chosen value. The corresponding

nodal heads are obtained using Eq. 2.33. The pipe flow rates are then updated using Eq. 2.34. As the algorithm progresses, the changes in pipe flow rates and nodal head values approach an insignificant value. The widely used EPANET 2 (Rossman, 2000) utilises the GGM as its network solver. Initial pipe flow rates in EPANET 2 are chosen to be equal to the flow corresponding to a velocity of 1ft/sec (0.3048 m/sec).

2.3.2 Extended Period Analysis

Section 2.3.1 has focused on the steady state analysis of WDNs in which nodal demands and reservoir water levels are presumed constant. However, the operation of real-life WDNs in general varies with time; and both nodal demand and tank water level fluctuate over a day. Using extended period analysis, a more realistic network analysis can be conducted since the network demands and water levels in tanks are allowed to change throughout the time period. It enables to understand the effects of changing water usage over time, the filling and draining cycles of tanks, or the response of pumps and valves to system changes. Also, it is an essential tool for analysing network water quality, optimizing pump scheduling, storage tank design and location.

Extended period analysis can be executed for a long period with network demands and operations changing from time to time. The total period of analysis is divided into several hydraulic time steps where the reservoir water levels and nodal demands are presumed constant in these time steps. For each time step, a steady state analysis is carried out. Results obtained at the end of these time steps, which involve the dynamics of tanks, pump scheduling and valve settings are used to update the inputs for the successive steady state analyses.

The basic steps involved in extended period analysis can be summarised as follows:

Step 1. At time t, the following data are available: (1) water levels of tanks $(H_m(t))$; (2) volumes of tanks $(V_m(t))$; (3) Nodal demand factors, $DF_i(t)$, and base demands, Qn_i^{base} for each node i. The required demand for each demand node at time t is calculated as

$$Qn_i^{req}(t) = Qn_i^{base}(t) \times DF_i(t) \tag{2.35}$$

Step 2. Using the data in step 1, the steady state hydraulic analysis of the network is obtained for time t. The flow rates of the tanks, $QT_m(t)$, are calculated as

$$QT_m(t) = \sum_{b \in B_m} Qp_b(t) \tag{2.36}$$

where B_m is the set of links connected to tank m and $Qp_b(t)$ represent the flow rates of links connected to tank m.

Step 3. Assuming the flow rate of a tank is constant during the time interval $(t, t+\Delta t)$, the volume change in a tank is obtained as

$$\Delta V_m(t, t+\Delta t) = QT_m(t)\Delta t \tag{2.37}$$

where $\Delta V_m(t, t+\Delta t)$ is the change of water volume in tank m during the time interval $(t, t+\Delta t)$, Δt represents the time step between the current and the next steady state simulation.

Step 4. The tank volume at time $t+\Delta t$ is then calculated as

$$V_m(t, t+\Delta t) = V_m(t) + QT_m(t)\Delta t \tag{2.38}$$

Step 5. The tank water level $H_m(t+\Delta t)$ is then predicted using the tank capacity-elevation curve by employing the tank volume in Eq. 2.38.

Step 6. If any status change occurs within the time interval $(t, t+\Delta t)$ (e.g tanks fully depleted or filled; valves opened or closed; pumps turned on or off), then the time $t+\Delta' t$ $(t < t+\Delta' t < t+\Delta t)$ at which the status occurs is determined. The time interval is reduced from Δt to $\Delta' t$. The volume change in tanks is calculated using Eq. 2.38 for the time $t+\Delta' t$ and water tank levels are updated.

Step 7. Time is advanced by the time step. Thus, $t = t+\Delta' t$ if any status change occurs. Else $t = t+\Delta t$.

The whole procedure is repeated until the entire period of the extended period simulation is analysed.

2.4 Pressure Dependent Network Analysis

As mentioned in Section 2.3, the main assumption of demand driven analysis (DDA) approach is that nodal demands are fully satisfied regardless of the pressure in the system. This method is satisfactory if the nodal heads are sufficient. However, there are several events when not all the nodal demands can be fully satisfied due to pressure deficient conditions in the systems. Pressure deficient conditions may arise due to pump failures, pipe bursts and the unavailability of components for system maintenance or rehabilitation purposes (Tanyimboh et al., 1999; Kalungi and Tanyimboh, 2003; Siew and Tanyimboh, 2012a). Under these circumstances, the performance of WDN is reduced. This requires water utilities to accurately simulate and analyse WDNs for crucial decision-making. Pressure dependent network analysis (PDA) is a realistic approach for evaluating the performance of WDNs under both normal and pressure

deficient conditions. PDA approximates the actual pressure shortfall in the system under any network operating conditions.

There are two PDA modelling approaches in literature for analysing WDNs under pressure-deficient conditions. The first method involves running DDA hydraulic simulator. The second method involves embedding the node head-flow relationship in the governing network equations.

2.4.1 Methods Based on Demand Driven Analysis

The existing DDA models for simulating WDNs do not have a procedure to integrate nodal head-flow relationships for pressure dependent modelling. To perform pressure dependent modelling multiple DDA runs are executed while adjusting specific parameters until sufficient hydraulic consistency is achieved. Although the method may seem to be straightforward for small WDNs, it is often time consuming and difficult for large systems applications (Jinesh Babu and Mohan, 2012). In addition, the method is not practical in situations requiring large number of hydraulic simulations. These include water quality modelling over an extended period simulation (e.g. Rossman, 2000; Seyoum et al., 2014) and design optimisation procedures based on evolutionary algorithms (Siew and Tanyimboh, 2012b).

Ang and Jowitt (2006) proposed an approach that involves iterative execution of DDA by introducing artificial reservoirs at demand nodes with insufficient pressure. The approach presumes that a low-resistance pipe connects each reservoir to its corresponding node. The method does not permit reverse flow from the artificial reservoir back into the network. Also, it does not allow flow to the reservoir to exceed the required demand of the corresponding node. In order to meet these conditions, the algorithm requires multiple runs of the EPANET solver where artificial reservoirs are added or removed from the model after each run. Jinesh Babu and Mohan (2012)

proposed an improvement to Ang and Jowitt (2006) method in order to carry out pressure-deficient network modelling in a single execution of the EPANET solver. Jinesh Babu and Mohan (2012) used artificial reservoir along with flow control valve and check valve. The flow control valve prevents surplus flow while the check valve prevents reverse flow from reservoir to demand node. The check valve will be closed if the nodal head is below the minimum required level. As noted by Gorev and Kodzhespirova (2013), the procedure by Jinesh Babu and Moha (2012) does not model the partial flow at a demand node satisfactorily. The approach is developed based on the head below which no flow is available. The head required for achieving a full demand satisfaction was not considered. By contrast, the approach proposed by Gorev and Kodzhespirova (2013) takes into account the partial flow at a demand node with proper resistance properties assigned to the artificial pipes. They used an artificial string made up of a flow control valve, a pipe with a check valve, and a reservoir at each node. The pressure- dependent demands are determined as the flows in the strings. The resistance of the artificial pipes is chosen such that the demands are satisfied in full at a desired nodal pressure. The Gorev and Kodzhespirova (2013) modelling approach assumed an identical nodal head-outflow relationship that applies to all networks. Abdy Sayyed et al. (2015) improved the Gorev and Kodzhespirova (2013) approach by replacing the artificial pipe and reservoir with an emitter to simulate partial flow condition. Also, they addressed the limitation in the nodal head-flow relationship to make it more generic. The artificial string proposed in Abdy Sayyed et al. (2015) to model pressure-deficient WDN is made up of a check valve, flow control valve and an emitter at each demand node. However, the procedure of adding artificial links and nodes at each demand node enlarges the network size that may lead to extra computational cost.

Rossman (2007) proposed the pressure-deficient analysis using the emitter component available in EPANET. Since the emitter operates based on unrestricted head-flow relationship, Rossman (2007) suggested modifying the EPANET 2 source code to add new status variables to the emitter to limit the nodal flow between zero flow and the fully assigned demand. The emitter approach was carried out in a single EPANET 2 run

while providing an improved computational efficiency in reference to the Ang and Jowitt (2006) approach. However, the algorithm has not been used to analyse large networks.

Kalungi and Tanyimboh (2003) developed a model for carrying out pressure-driven network analysis based on a technique that involves identifying zero-flow and partial flow nodes. The approach is based on joint nodal head and pipe flow system of equations, which does not make explicit use of the head-flow relationship. The technique involves iterative use of DDA.

Tanyimboh and Templeman (1995) and Tanyimboh and Tabesh (1997) developed a method of obtaining the available nodal outflow for pressure dependent network analysis based on the source head required to satisfying the network demand in full. They used the following source head-discharge relationship.

$$H_s = H_s^{\min} + R_s (Q_s^{avl})^{n_s} \tag{2.39}$$

where H_s is the head available at the source, Q_s^{avl} is the total flow from all demand nodes, R_s is the resistance constant and the exponent n_s was taken as 2. $H_s^{\min}$ represents the source head at which the most critical node of the network begins to deliver water. The expression for the sum of the nodal outflows is

$$Q_s^{avl} = \left(\frac{H_s - H_s^{\min}}{R_s} \right)^{1/n_s} \tag{2.40}$$

The total available nodal outflows Q_s^{avl} is equal to the total nodal demands Q_s^{req} when $H_s = H_s^{des}$. H_s^{des} is the required source head to satisfy all the demand nodes in full.

H_s^{des} is the total head losses (obtained from DDA) in links along a path from the source to the most critical node. Substituting H_s and Q_s^{avl} in Eq. (2.39) with H_s^{des} and Q_s^{req} respectively gives

$$H_s^{des} = H_s^{\min} + R_s (Q_s^{req})^{n_s} \tag{2.41}$$

From Eq. (2.41), an expression for R_s is obtained and applied into Eq. (2.40) to obtain the total flow delivered. Thus, the source head-discharge relationship for the source head method (SHM) is expressed as

$$Q_s^{avl} = Q_s^{req} \left(\frac{H_s - H_s^{\min}}{H_s^{des} - H_s^{\min}} \right)^{1/n_s} ; \quad H_s^{\min} \le H_s \le H_s^{des} \tag{2.42}$$

Although SHM is a simple approach, it tends to underestimate the total flow from pressure deficient networks. The method is based on DDA results. Very often, DDA gives smaller or even larger negative heads for the most critical node (Tanyimboh and Templeman, 2010; Kalungi, 2003). As a result, SHM gives a high value of H_s^{des} and thus an underestimated Q_s^{avl} value. Also, the SHM cannot be used for multiple-source networks. Tabesh (1998) proposed an enhancement of the SHM named the Improved Source Head method (ISHM). Unlike SHM, which estimates the network outflow based on just the critical node, the ISHM considers every demand node individually and calculate its flow as follows

$$Q_j^{avl} \approx Q_j^{req} \left(\frac{H_s - H_s^{\min}}{H_{s,\,j}^{des} - H_s^{\min}} \right)^{1/n_j} ; \qquad H_s^{\min} \le H_s \le H_{s\,j}^{des} \;,\; \forall j \tag{2.43}$$

where exponent n_j varies between 1.5 and 2 (Gupta and Bhave, 1996). $H_{s,j}^{des}$ is the head required at the source to satisfy the demand at node j in full. $H_{s,j}^{des}$ is obtained using DDA and it is taken as the nodal elevation plus the sum of the head losses in pipes connecting node j to the source. Tanyimboh et al. (2001) demonstrated that ISHM provides better estimates of the nodal flows than SHM approach. The main difficulty in the ISHM is that the exponent, n_j used requires an extensive field data for accurate calibration (Tanyimboh and Tabesh, 1997).

2.4.2 Methods Based on Nodal Head-Flow Relationship

Pressure deficient network analysis can be carried out by embedding a nodal head-flow relationship in the governing system of hydraulic equations. The nodal head-flow relationships used to estimate the actual flow at demand nodes based on nodal pressure (e.g Germanopoulos, 1985; Wagner et al., 1988; Udo and Ozawa, 2001; Tanyimboh and Templeman, 2010). The relationships are generally formulated on the basis that nodal demand is satisfied in full when the nodal head is equal to or greater than the desired head and zero when the nodal head is equal to or lower than the minimum head (Tanyimboh and Templeman, 2010). A major advantage of the methods where the head-flow relationships are embedded in the system of equations is that they solve the system of non-linear equations only once; contrary to the PDA approaches, that involves iterative DDA simulations (see Section 2.4.1). Several head-flow relationships have been proposed in literature to characterise the variation of nodal pressure and flows in WDNs.

Germanopoulos (1985) suggested the following head-flow relationship to estimate the available outflow at a node for a pressure deficient network.

$$Qn_i = Qn_i^{req}\left(1 - b_i\, e^{-c_i[\frac{\mathrm{Pr}_i}{\mathrm{Pr}_i^{\#}}]}\right) \tag{2.44}$$

where Qn_i and Qn_i^{req} are the available outflow and the required demand at node i respectively; b_i and c_i are coefficients to be calibrated for node i. Pr_i is the available pressure at node i and $\mathrm{Pr}_r^{\#}$ is the nodal pressure at which a proportion of the required demand of node i is supplied.

Germanopoulos (1985) suggested that the coefficients b_i and c_i may be taken as 10 and 5 respectively in the absence of field data and $\mathrm{Pr}_r^{\#}$ is taken as the pressure to satisfy 93.2% of the required nodal demand. However, the head-flow relationship does not fulfil the condition that $Qn_i = 0$ when $Hn_i = Hn_i^{\min}$ and $Qn_i = Qn_i^{req}$ when $Hn_i = Hn_i^{des}$; in which Hn_i is the head at node i; $Hn_i^{\min}$ is the nodal head at node i below which the outflow is zero; Hn_i^{des} is the desired head at node i for achieving a full demand satisfaction. Gupta and Bhave (1996) suggested an improvement to address the shortcomings. The modified head-outflow relationship is expressed as

$$Qn_i = Qn_i^{req}\left(1 - 10^{-c_i[\frac{Hn_i - Hn_i^{\min}}{H_{ni}^{des} - H_{ni}^{\min}}]}\right) \tag{2.45}$$

Wagner et al. (1988) and Chandapillai (1991) suggested a parabolic relationship (Eq. 2.46) for the HGL values between $Hn^{\min}$ and Hn^{des}.

$$Qn_i = Qn_i^{req}, \qquad if \quad Hn_i \geq Hn_i^{des} \tag{2.46a}$$

$$Qn_i = Qn_i^{req}\left(\frac{Hn_i - Hn_i^{\min}}{Hn_i^{des} - Hn_i^{\min}}\right)^{1/n_e}, \qquad if \quad Hn_i{}^{\min} < Hn_i < Hn_i^{des} \tag{2.46b}$$

$$Qn_i = 0, \qquad if \quad Hn_i \le Hn_i^{\min} \tag{2.46c}$$

The value of the exponent parameter, n_e varies between 1.5 and 2 (Gupta and Bhave, 1996).

Fujiwara and Ganesharaja (1993) proposed a differential function for the relationship between nodal head and flow.

$$Qn_i = Qn_i^{req}\frac{\int_{Hn_i^{\min}}^{Hn_i}(H - Hn_i^{\min})(Hn_i^{des} - H)dH}{\int_{Hn_i^{\min}}^{Hn_i^{des}}(H - Hn_i^{\min})(Hn_i^{des} - H)dH} \qquad Hn_i^{\min} < Hn_i < Hn_i^{des} \tag{2.47}$$

Eq. 2.47 can be applied to any network. However, the function is not straightforward to use and more computational effort may require for its evaluation (Tabesh et al., 2002).

Tanyimboh and Templeman (2010) suggested a nodal head-flow function based on the Logit function, which is expressed as

$$Qn_i = Qn_i^{req}\frac{\exp(\alpha_i + \beta_i Hn_i)}{1+\exp(\alpha_i + \beta_i Hn_i)} \tag{2.48}$$

The parameters α_i and β_i determine the shape of the function curve. Their values are obtained using field data. Eq. 2.48 can be rearranged to give

$$\frac{Qn_i}{Qn_i^{req}} = \frac{\exp(\alpha_i + \beta_i Hn_i)}{1 + \exp(\alpha_i + \beta_i Hn_i)} \tag{2.49}$$

where Qn_i / Qn_i^{req} represents the nodal demand satisfaction ratio (DSR). The nodal DSR value is 1.0 when $Hn_i \geq Hn_i^{des}$ and zero when $Hn_i \leq Hn_i^{\min}$. Tanyimboh and Templeman (2010) suggested that in the absence of field data the values of α and β can be found by taking Eq. 2.50 and 2.51 that effectively represent the conditions for full and zero demand satisfaction respectively.

$$Qn_i(Hn_i^{des}) = 0.999 Qn_i^{req} \tag{2.50}$$

$$Qn_i(Hn_i^{\min}) = 0.01 Qn_i^{req} \tag{2.51}$$

Solving Eq. 2.50 and 2.51 simultaneously will give

$$\alpha_i = \frac{-4.595 Hn_i^{des} - 6.907 Hn_i^{\min}}{Hn_i^{des} - Hn_i^{\min}} \tag{2.52}$$

$$\beta_i = \frac{11.502}{Hn_i^{des} - Hn_i^{\min}} \tag{2.53}$$

The major drawback that was noted in most of the nodal head-flow relationships in literature is the lack of continuity in the functions (and or their derivatives) at the transitions between zero and partial nodal flow and/or between partial and full demand satisfaction. These discontinuities may cause convergence difficulties when the functions are embedded in the governing systems of hydraulic equations (Gupta et al., 2003; Tanyimboh and Templeman, 2010). By contrast, the Tanyimboh and Templeman (2010) nodal head-flow relationship provides a smooth transition between zero and

partial nodal outflow and between partial and full demand satisfaction. Both the function and its derivative have no discontinuities that make the function suitable to be incorporated into the system of equations effectively. Recently, Kovalenko et al. (2014) indicated that the Tanyimboh and Templeman (2010) function provides better convergence properties in the computational solution of the system of equations in comparison to the Wagner et al. (1988) function when a line search procedure that control the iteration step is used. Ciapioni et al. (2015) studied the nodal head-flow relationship in two urban areas with different topographical characteristics (flat and mountainous sites). The procedure consists of simulating a large number of possible scenarios regarding the users supplied by a demand node of WDN. The results showed that the Tanyimboh and Templeman (2010) function performed better than the other nodal head-flow relationships considered in the study.

Several pressure dependent analysis works have been carried out by using the nodal head-flow relationship. Tabesh et al. (2002) proposed a method for solving head-dependent flows based on the Newton-Raphson technique that explicitly incorporates the Wagner et al. (1988) head–flow relationship into the continuity equations. The method ensures fast convergence based on a step length adjustment parameter whose value is obtained by trial and error. Recently, Tsakiris and Spiliotis (2014) proposed a PDA implementation of the Newton-Raphson method using a pipe flow formula computed by combining the Darcy-Weisbach pipe friction head loss formula and the Colebrook-White function for the friction factor (Spiliotis and Tsakiris, 2011). The approach utilised the Wagner et al. (1988) head–flow relationship for pressure-dependent network analysis. The approach has the advantages that it avoids the iterative solution of the Colebrook-White function and benefits from the extra accuracy of the Darcy-Weisbach pipe friction head loss formula compared to empirical equations such as the Hazen-Williams formula. Also, the approach obviates the use of the hydraulic resistance or pipe resistance coefficient. Ackley et al. (2001) utilised the Wagner et al. (1988) head-flow relationship for the pressure-dependent analysis. They formulated the pressure-dependent analysis as an optimization model where the sum of nodal outflows

was maximized. Also, Ackley et al. (2001) proposed a simple and efficient technique for verifying the accuracy of the PDA results. In this approach, the nodal flows obtained from the PDA model are used as nodal demands in a DDA model. The DDA model is run with all other network data remaining the same. If the PDA nodal flows are accurate, the nodal heads and pipe flows obtained from the DDA will be identical to the corresponding nodal heads and pipe flows generated by the PDA model. This PDA results verification technique has been adopted widely (e.g. Ackley et al., 2001; Tanyimboh et al., 2003; Kalungi and Tanyimboh, 2003; Siew and Tanyimboh, 2012a; Seyoum and Tanyimboh, 2014).

Giustolisi et al. (2008b) proposed a steady- state model that integrates the Wagner et al. (1988) nodal head-flow relationship and leakage (Germanopoulos 1985; Germanopoulos and Jowitt 1989) with the Global Gradient Method. The performance of the model was analysed using networks with pipes only. Giustolisi et al. (2008a) analysed the hydraulic performance of a WDN by employing a 24-hour extended period simulation. The analysis was performed using a PDA whose values of the required nodal head for full demand satisfaction (Hn_i^{des}) varied according to the daily demand pattern. However, referring to OFWAT (2004), the prescribed level of service does not vary throughout the day for water utilities within the UK. As reported in Tanyimboh and Siew (2012), a prescribed minimum residual pressure that tracks the daily demand pattern tends to underestimate the available flow when demands are above average and overestimate the available flow when demands are below average.

Tanyimboh and Templeman (2010) developed a PDA model based on the Newton Raphson method that has a line search and backtracking procedure. The model is known as PRAAWDS (Program for the Realistic Analysis of the Availability of Water in Distribution Systems) and it incorporates four nodal head-flow relationships, i.e the Wagner et al. (1988), Germanopoulos-Gupta-Bhave (Gupta and Bhave, 1996), Fujiwara and Ganeshrajah (1993), and Tanyimboh and Templeman (2010). It also enables

network simulation using DDA. The model incorporates the Ackley et al. (2001) PDA result verification technique. Several researches demonstrated the accuracy and robustness of PRAAWDS e.g. Setiadi et al. (2005), Tanyimboh and Setiadi (2008a, b) and Shan (2004). However, the model is restricted to simulation of steady state hydraulic condition. Also, it does not have a water quality modelling functionality.

Siew and Tanyimboh (2012a) developed a pressure-dependent extension to EPANET 2 called EPANET-PDX by integrating the Tanyimboh and Templeman (2010) nodal head-flow function into the global gradient algorithm (Todini and Pilati, 1988) that is the hydraulic analysis model of EPANET 2. EPANET-PDX employs line minimisation to optimise iterative corrections of pipe flow rates and nodal heads. The PDA model is thought to have preserved the EPANET 2 modelling functionality in full. Extensive testing conducted on the model revealed good modelling performances. EPANET-PDX was combined with a multi-objective genetic algorithm for optimization of WDNs and superior results were obtained compared to other previous solutions (Siew and Tanyimboh, 2012b; Siew et al., 2014). The research carried out herein involves enhancing the computational performance of EPANET-PDX further. Chapter 4 presents an alternative implementation of the line search and backtracking procedure to integrate the Tanyimboh and Templeman (2010) nodal head-flow function into the system of hydraulic equations in the global gradient algorithm.

2.5 Water Quality Modelling in Water Distribution Networks

One of the key concerns of water industries today is the deterioration of water quality in WDNs (Rossman, 1993; Clark and Grayman, 1998). Water quality changes as it transports from the treatment plant to the consumers. Some of these changes include, loss of disinfection residual, which can lead to bacterial re-growth (Clark and Haught, 2005), formation of potentially carcinogenic disinfection by products (DBPs) due to reaction of disinfectant with organic and inorganic substances in water (Rodriguez et al.,

2004), development of taste and odour, and corrosion. The water quality concerns in combination with the rigorous standards set by regulatory bodies have pressed water companies to depend increasingly on models in the quest to understand and control the dynamics of water quality processes. Water quality models can be used to investigate points in WDNs with long detention times, low disinfection residuals and excessive concentrations of disinfection by-products (DBPs). The models can also facilitate decision making for water quality management. This includes the selection of sampling locations and sampling frequency, optimisation of the operation and the locations of booster disinfection stations (Rossman et al., 1993). Water quality models have also been used to aid monitoring to help address concerns about possible deliberate contamination of water systems by terrorists (Skadsen et al., 2008).

The current water quality models for water distribution systems are coupled with hydraulic models to describe the quality of the system. Flow rates in pipes and the flow paths that define how water travels through the network are used to determine mixing, residence times, and other hydraulic characteristics affecting disinfectant transport and decay. Similar to hydraulic analysis of WDN, the water quality analysis can be of either steady state or extended-period.

2.5.1 Governing Principles of Water Quality Modelling

Most water- quality models utilize advective-reactive transport mechanism to predict the changes in constituent concentrations due to transport through a pipe and due to reactions within the bulk flow and along the pipe wall. The governing water-quality equations are based on the following principles. These are (1) conservation of mass within differential lengths of pipes; (2) complete and instantaneous mixing of the water entering at nodes and storage facilities; and (3) reaction kinetics for the growth or decay of the substance as it transports in pipes and storage facilities (Rossman and Boulos, 1996; Rossman, 2000; Clark and Grayman, 1998).

The expressions for the conservation of mass during transport in pipes is

$$\frac{\partial C_i}{\partial t} = -u_i \frac{\partial C_i}{\partial x} + r_i(C_i) \tag{2.54}$$

where C_i = concentration in link i at location x and time t; u_i = mean flow velocity in link i; and $r_i(C_i)$ = rate of reaction as a function of concentration. Eq. 2.54 shows that the rate at which the mass of material changes within a small section of pipe equals the difference in mass flow into and out of the section plus the rate of reaction within the section. The equation is based on the assumption that the bulk fluid is completely mixed and there no intermixing of mass between adjacent parcels of water traveling down a pipe (i.e., longitudinal dispersion in pipes is negligible). To solve Eq. 2.54, C_i at $x = 0$ for all times (a boundary condition) and a value for r_i should be known.

Eq. 2.55 represents the concentration of substance leaving a node and entering a pipe. Considering complete and instantaneous mixing at nodes, the concentration of a substance in water leaving node q is the flow-weighted sum of the concentrations from the inflowing pipes.

$$C_i \mid_{x=0} = \frac{\sum_{j \in I_q} Qp_j C_j \mid_{x=Lj} + Q_{ext} C_{ext}}{\sum_{j \in I_q} Qp_j + Q_{ext}} \tag{2.55}$$

where i = link with flow leaving node q; I_q = set of links with flow into q; L_j = length of link j; Qp_j = flow in link j; C_{ext} = external source concentration entering node q; and Q_{ext} = external source flow entering node q.

Assuming that the contents of the storage facilities (tanks or reservoirs) are completely mixed and the concentration in the tanks undertake reaction; a mass balance can be expressed as

$$\frac{\partial(V_s C_s)}{\partial t} = \sum_{i \in I_s} Qp_i C_i \mid_{x=L_i} - \sum_{j \in O_s} Qp_j C_s + r(C_s) \tag{2.56}$$

where V_s = volume in storage at time t, C_s = concentration within the storage facility, I_s = set of links providing flow into the facility, and O_s = set of links receiving flow from the facility; and $r(C_s)$ = rate of reaction as a function of concentration within the storage facility.

As shown in Eqs. 2.54 and 2.56 concentrations within pipes, storage tanks, and reservoirs are a function of reaction terms. Several complex physical and chemical processes take place when water transport into distribution systems. The most frequently used chemical processes in water quality simulation models are bulk flow reactions, reactions that occur on the pipe wall, and formation reactions that involve a limiting reactant.

When a substance travels in a pipe or resides in storage, it can undergo reaction with constituents in the bulk water. The rate of reaction can be described as a power function of concentration (Rossman, 2000).

$$r(C) = k_b C^n \tag{2.57}$$

where k_b is a bulk reaction constant; n is the reaction order; and C is concentration in bulk water. When a limiting concentration exists on the ultimate growth or loss of a substance then the rate expression in Eq. 2.57 becomes

$$r(C) = k_b (C_L - C) C^{(n-1)} \qquad for \qquad n > 0; k_b > 0 \tag{2.58a}$$

$$r(C) = k_b (C - C_L) C^{(n-1)} \qquad for \qquad n > 0; k_b < 0 \tag{2.58b}$$

Zero-, first-, and second-order reactions are commonly used to model chemical processes that occur in WDNs. Using the generalized expressions in Eq. 2.57 and 2.58, these reactions can be modelled by allowing n to equal 0, 1, or 2 and determining k_b experimentally.

During water flowing through pipes, the constituents in the bulk water can be transported to the pipe wall and react with materials at the pipe walls such as biofilm and corrosion materials. The rate of water-quality reaction occurring at the pipe wall is dependent on the concentration in the bulk water, the amount of wall area available for reaction and the rate of mass transfer between the bulk water and pipe wall. For first order kinetics, the rate of a pipe wall reaction can be expressed as

$$r(C) = \frac{k_w k_f C}{r_h (k_w + k_f)} \tag{2.59}$$

where k_w is wall reaction rate constant (length/time), k_f mass transfer coefficient (length/time), and r_h = hydraulic radius. The value of the mass transfer coefficient k_f depends on the molecular diffusivity of the reactive species and on the Reynolds number of the flow (Rossman et. al, 1994).

2.5.2 Solution Methods of the Water Quality Equations

Eulerian and Lagrangian approaches are two commonly used numerical methods to solve the dynamic water quality equations used in water quality models. In these solution methods, a hydraulic simulation must first be executed for extended-period simulation to determine the flow, flow direction and velocity in each pipe at all times during the simulation.

Grayman et al. (1988) and Rossman et al. (1993) suggested the application of Discrete Volume Method (DVM), which is an Eulerian approach, for water quality modelling of WDNs. The method divides each pipe into equal segments with completely mixed volumes. At each successive water quality time-step, the concentration within each segment is first reacted and then transferred to the adjacent downstream segment. At nodes, the concentration is updated considering a flow-weighted average of incoming inflows (as described in Eq. 2.51). The resulting concentration is then transported to all adjacent downstream segments. This process is repeated for each water quality time-step until a different hydraulic condition is occurred. When this occurs, the pipes are divided again under the new hydraulic conditions and the process continues. The accuracy of DVM depends on the size of the water quality time step used.

Liou and Kroon (1987) suggested the Lagrangian method for water quality modelling in WDNs. Similar to the Eulerian approach the method divides the pipes into segments. However, rather than using fixed control volumes as in Eulerian methods, the concentration and size of water parcels are tracked as they travel through the pipes. At each time step, the size of the farthest upstream parcel of each pipe increases as water enters into the pipe. At the same time, the farthest downstream parcel shortens as water leaves the pipe. Similar to the DVM, the reactions of a constituent within each parcel is calculated at each water quality time step. Also, the concentration at each node is updated taking a flow-weighted average of incoming inflows. If the resulting nodal

concentration is significantly different from the concentration of a downstream parcel, a new parcel will be created at the end of each link that receives inflow from a node. The process is then repeated for each water quality time-step until the next hydraulic change is occurred and the procedure begins again. It is worth mentioning that at the start of the procedure each pipe in the network consists of a single segment whose quality equals the initial quality assigned to the upstream node.

Lagrangian solutions can be either time-driven or event-driven. In a time-driven method, conditions are updated at a fixed time step. In an event-driven model, conditions are updated when the source water quality changes or when the front of a parcel reaches a node. A comparison of different solution methods by Rossman and Boulos (1996) indicated that the Lagrangian time-driven method is the most efficient approach for water quality modelling of WDNs. The accuracy of the method depends on the water quality time step and the concentration tolerance used to limit the creation of new segments.

Several computer models are available to simulate water quality processes in WDNs. EPANET 2, a public domain hydraulic and water quality model, is among the most widely used. EPANET's water quality simulator employs a Lagrangian time-based approach. The model enables simulation of non-reactive tracer materials, chlorine decay, disinfection by-products growth (e.g trihalomethanes) and water age (Rossman, 2000). However, the model is single species and limited to model the dynamics of chlorine, trihalomethane or water age. The model does not allow simulation of multiple interacting species. Shang et al. (2008) developed a multispecies extension of EPANET model (EPANET-MSX). EPANET-MSX enables simulation of multiple chemical species in bulk water and at pipe wall.

Both EPANET 2 and EPANET-MSX rely on DDA hydraulic model to obtain pipe flow data for subsequent water quality analysis. The use of DDA has restricted the application of the water quality models to normal operating conditions with satisfactory pressure.

The models breaks down when abnormal or irregular network operating conditions prevail due to considerable pressure reduction in the system (Seyoum and Tanymboh, 2013). The application of PDA model to evaluate the water quality performance of WDNs under abnormal (pressure-deficient) conditions is presented in Chapter 5.

2.6 Conclusions

WDN models are very valuable tools that allow engineers to analyse hydraulic as well as water quality performances of networks over a wide range of operational conditions. The conventional demand driven network analysis approach that presumes demands are fully satisfied irrespective of the network pressure is valid only if the network performs under normal operating conditions. The models may be inadequate when abnormal network operating conditions prevail. By contrast, the pressure dependent network analysis is a realistic approach for evaluating the performance of WDNs under both normal and abnormal operating conditions. A comprehensive review on the different PDA modelling approaches in literature for analysing WDNs under pressure deficient networks is provided. The limitations of the approaches were identified and discussed. The chapter has presented the principles of hydraulics and water quality analysis that are often employed in WDN models. A review of the most commonly used numerical methods to solve the network hydraulic and water quality analysis problems has been done. The limitation of steady-state analysis was highlighted in comparison to the extended period analysis that determines the behaviour of a network over time.

The next chapter comprises a review on genetic algorithm that is the most widely used evolutionary algorithm for optimisation of WDN design together with the current developments made to enhance the algorithm's efficiency and effectiveness.

Chapter 3

Review of the Recent Developments in Genetic Algorithms for Optimisation of Water Distribution Systems

3.1 Introduction

Water distribution systems (WDSs) comprise the networks of pipes, pumps, valves, reservoirs, tanks and nodes, which transport drinking water from the supply source to the users. The construction, operation and maintenance of these systems involve an enormous capital investment. It is, therefore, crucial to design and rehabilitate them in a cost effective manner without compromising the required performance and regulatory standards. Traditionally, engineers used rules of thumb to design WDSs; these methods are both time consuming and likely to produce sub-optimal solutions (Vairavamoorthy and Shen, 2004; Marchi et al., 2014). An effective solution method which is reliable, easy to implement and computationally efficient is required to enable the optimal design of real-life systems. Several optimization techniques have been proposed previously. Most of these techniques have employed mathematical programming approaches such as linear and non-linear programming where the design variables assumed to be continuous (Alperovits and Shamir, 1977; Morgan and Goulter, 1985; Su et al., 1987; Lansey and

Mays, 1989). These methods are computationally efficient as they approach the optimization problem in deterministic manner. Nevertheless, the performances of the mathematical programming techniques can become extremely complex due to the many non-linear constraints implemented while analysing networks with many pipes as well as hydraulic components such as pumps and storage reservoirs. These constraints are normally found in the form of nodal mass balance and energy conservation equations, which are commonly, satisfied using a hydraulic network solver (di Pierro, 2009). These difficulties associated with mathematical optimisation techniques have motivated several researchers to use stochastic search optimization methods such as evolutionary algorithms.

Evolutionary algorithms (EAs) are derivative-free random search methods that are suitable to solve non-linear, non-convex, and multimodal problems successfully (Nicklow et al., 2010). EAs have gained widespread acceptance in recent years for optimizing the design and operation of WDSs. The methods are capable of handling discrete variables, which is one of the key aspects in the WDS optimization. EAs are mathematically non-complex and their search approach is dependent only on objective function and corresponding fitness values. Also, the population–based nature of EAs allows them to explore a vast search space of solutions and increases the chance of obtaining the global optimum solution.

A number of EAs have been applied for optimizing the design and operation of WDSs. These include genetic algorithms (Simpson et al., 1994; Dandy et al., 1996; Savic and Walters,1997; Wu et al., 2001; Vairavamoorthy and Ali, 2000; and Montesinos et al., 1999), ant-colony optimisation (Maier et al., 2003), particle swarm optimization (Montalvo et al., 2008), simulated annealing (Cunha and Sousa, 1999), shuffled frog leaping (Eusuff and Lansey, 2003), differential evolution (Vasan and Simonovic, 2010), harmony search (Geem, 2006), tabu search algorithm (Cunha and Ribeiro, 2004). Amongst the stochastic optimisation techniques, genetic algorithms (GAs) are probably

the best known and most extensively applied method in the area of WDS optimization due to their robustness and capability in yielding optimal or near optimal solutions (Dandy et al., 1996; Savic and Walters, 1997; Wu et al., 2001 and Vairavamoorthy and Ali, 2000, 2005).

The main aim of this chapter is to review the existing approaches to enhance the efficiency and effectiveness of GAs to solve WDSs optimisation problems. The research carried out herein involves the implementation of GA and hence, the approach is discussed in detail in the next section (Section 3.2). Section 3.3 reviewed the main developments made on GAs for WDSs applications. Different approaches have been discussed and their limitations assessed. Also, the most common parallelization approaches that have been applied in WDS literature to improve the GAs' computational efficiency are reviewed and their benefits and drawbacks are discussed.

3.2 Genetic Algorithms

Genetic algorithm (GA) is a stochastic optimization approach inspired by Darwin's theory of evolutions that uses natural selection as the driving force (Holland, 1975; Goldberg, 1989). GA operates by creating a population of potential solutions for the given problem. It provides a fitness values to each solution. The fitness value determines the suitability of an individual solution (chromosome or strings) for a given problem. The higher the fitness is the more the probability of the individual to proceed to the next generation. The individuals of the population are recombined using operators such as crossover and mutation to produce new generation of individuals that are more likely to have better fitness value. A population of individuals undergoes a sequence of mutation and crossover transformations. A selection scheme biased towards selecting fitter individuals yields individuals for the next generation. The algorithm is then expected to

converge to the optimal solutions after several generations. The general framework of GA is described in Fig. 3.1.

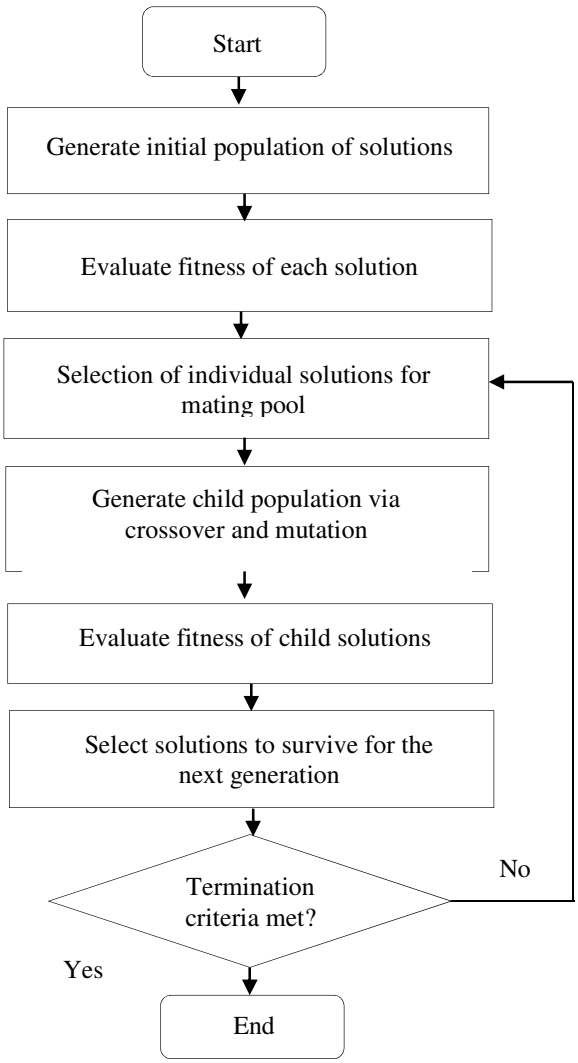


Figure 3.1 General framework of GA

GA begins by randomly generating a population of individuals. The individuals are encoded as chromosomes, each consisting of a set of genes that describe a solution. The chromosomes are evaluated based on fitness values assigned to them. Individuals are

then selected from the population to create a mating pool based on their fitness values. Individuals with higher fitness values have a higher probability of being selected to produce offspring that represent new solutions. A very small portion of the offspring will mutate after reproduction. The GA operation continues until a pre-specified termination criterion is met (e.g. number of generations).

3.2.1 Genetic Operators

Genetic operators consist of selection, crossover and mutation. In combination, the operators serve to create and explore new offspring at each generation. Crossover involves the transformation of two or more individuals into new offspring. It plays an important role in the reproduction phase and usually set to have a very high probability of occurrence during the evolution process to enable exploring more of the solution space and reducing the chance of falling into a local optimal solution. On the other hand, mutation operator randomly changes individuals to ensure that the whole search space has been thoroughly investigated and to facilitate the recovery of genetic diversity lost during reproduction and selection process. Selection drives the search toward the regions of best individuals.

3.2.1.1 Crossover Operator

In crossover, there is a partial exchange of some of the parents' genetic material to produce children. In doing so, it is expected that the GA will produce some offspring that may attain superior characteristics relative to their parents. In general, crossover happens to certain number of sub-strings in the chromosome after a point known as the crossover point. The simplest form of crossover (single-point crossover) works by selecting sections of one parent's chromosome to include in the child and then filling the

remaining parts of the child from the other parent's chromosome. A crossover point is randomly picked and genetic materials are exchanged between parents to produce offspring as illustrated in Fig. 3.2.

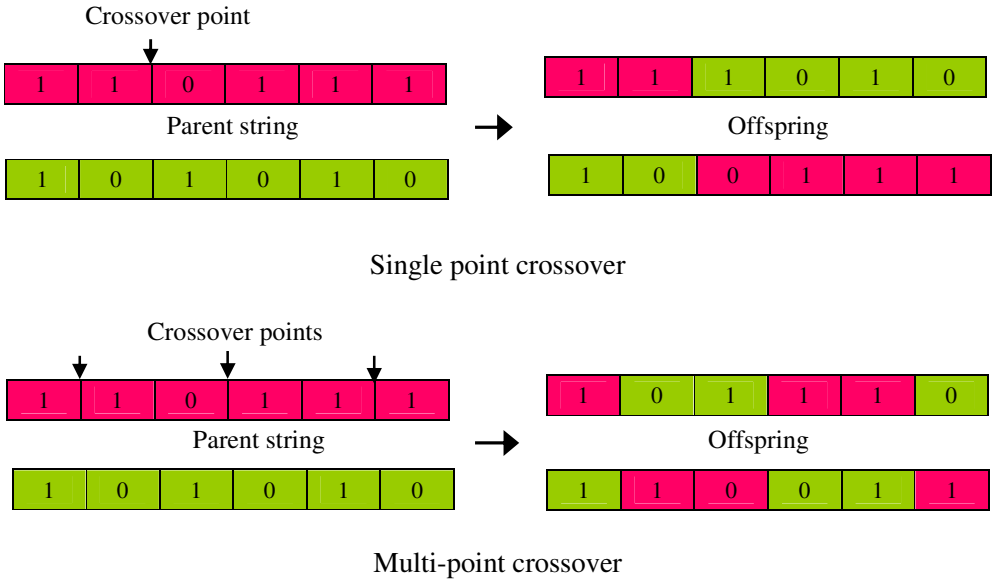


Figure 3.2 Operation of single and multi-point crossover

3.2.1.2 Mutation Operator

Mutation is an operator that randomly changes the value of a bit within the string. Normally it is applied with a very low probability. If the probability of mutation is too high, there will be frequent disruption of good genetic material that will hinder the algorithm from converging quickly. There are various forms of mutation such as the inversion mutation, insertion mutation, displacement mutation etc. Fig. 3.3 shows the operation of the basic single-point mutation that randomly flips a bit within the chromosomes.

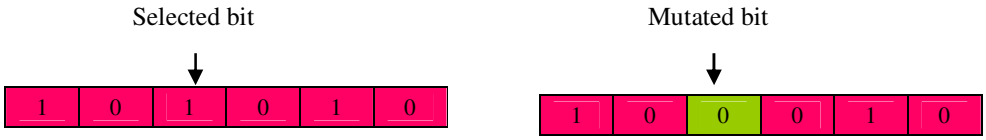


Figure 3.3 Operation of the basic GA mutation

3.2.1.3 Selection Operator

Selection provides the driving force in a GA that guides the GA search toward promising regions in the search space. Selection operators are characterized by selection pressure that represents the degree at which the best individual in the population are favoured by repeated application of the selection operator (Back et al., 2000). Selection process can control the level of exploration and exploitation by varying selection pressure (Back, 1994). A higher selection pressure can cause premature convergence of the algorithm due to quick loss of diversity in the population, while a lower selection pressure makes a slow convergence of the algorithm to find optimal solution (Back et al., 2000). Therefore, an appropriate selection pressure that maintains a good balance between exploitation and exploration is required for optimisation problems (Goldberg and Deb, 1991).

Roulette-wheel and tournament selection are the most commonly used selection operators in the literature. In the roulette-wheel selection, which was introduced by Holland (1975), each individual in the population occupies an area on the roulette-wheel proportional to its fitness. This representation makes high fitness individuals to have larger area and thus have higher selection probability. The roulette wheel is rotated as many times as the population size to select an individual that performs a genetic operation (e.g. crossover). Most spinning of the roulette wheel is likely to choose the fitter individuals. This may cause the population to lose genetic diversity and makes the algorithm to converge to suboptimal solution prematurely (Goldberg and Deb, 1991).

Goldberg and Deb (1991) proposed tournament selection where a specified number of individuals are selected from the population and their fitness compared. The fittest individual is the winner and will be selected to be part of the mating pool. The tournament is repeated with different individuals until the mating pool is sufficiently filled. In tournament selection, the selection pressure can be adjusted according to the tournament size. An increased in tournament size can provide an increase in selection pressure. A large tournament size will result in a mating pool consisting of a higher number of fitter solutions on average as compared to a smaller tournament.

3.2.2 Representation of Candidate Solution

When GA is applied to optimization problems, the candidate solutions are normally encoded using different approaches. Among these, binary coding (Holland, 1975; Goldberg, 1989), Gray coding (Caruana et al., 1989) and real coding are included. Binary coding is the most common encoding method where problem variables are represented by bit combinations of 0 and 1. As described in Fig. 3.2 and Fig. 3.3, the crossover and mutation operators for binary coding are easy to apply. However, one of the main drawbacks of this representation is the existence of redundant codes (Herrera et a., 1998). For example, in the design of a network of five pipes shown in Fig. 3.4, if the commercially available pipe diameter sizes are (100, 200, 300, 400 and 500 mm) then a three-bit substrings that has 8 (i.e. 2^3) possible bit combination will be used to map the diameters, therefore the mapping set will be (000, 001, 010, 011, 100, 101, 110, 111). It is worth noting that three redundant bit combinations are not corresponding to any of the available pipe diameters. One way of addressing this difficulty is by randomly remapping the redundant substrings to any of the five available pipe sizes.

Similar to binary coding, Gray-coding uses strings, but differs in the way the bits are represented. The key feature of this representation is that only a single bit changes

between neighbouring substrings. Binary coding, on the other hand, can have a large number of changes between adjacent strings. Table 3.1 shows the binary and Gray code representation of eight available pipe sizes using a 3-bit combination.

Table 3.1. Gray and binary code representation for 8 available pipe sizes

Pipe size (mm)	Gray coding	Binary coding
100	000	000
200	001	001
300	011	010
400	010	011
500	110	100
600	111	101
700	101	110
800	100	111

As can be seen in Table 3.1, in the binary representations, for the GA to change the representation from 400 mm and 500 mm, all three bits must alter simultaneously. Hence, the probability that crossover and mutation will occur to cross the Hamming cliff can be very small. Generally, the mutation probability is rather low and hence the chance of pipe size 400 mm becoming 500 mm is very low. Hamming cliff represents a phenomenon where neighbouring phenotypes (pipe sizes) are represented by completely different genotypes (binary code) (Caruana and Schaffer, 1988). The Hamming cliff may produce problems under some conditions, such as the convergence towards no global optimum (Herrera et a., 1998). Gray coding is therefore eliminates problems related to Hamming cliff.

In real coding, genes are represented as real numbers where each has a unique value. Real number encoding is suitable for optimization problems with variables in continuous search space (Herrera et a., 1998). The issue of redundancy associated with binary coding is addressed in this representation. In real coding representation, both genotype

space and phenotype space have identical topological structure (Gen et al., 2008). Therefore, the coding and decoding processes that are needed in the binary coding are avoided, thus increasing the GA's speed. Fig. 3.4 shows binary and real code representations of five randomly generated designs of a simple network.

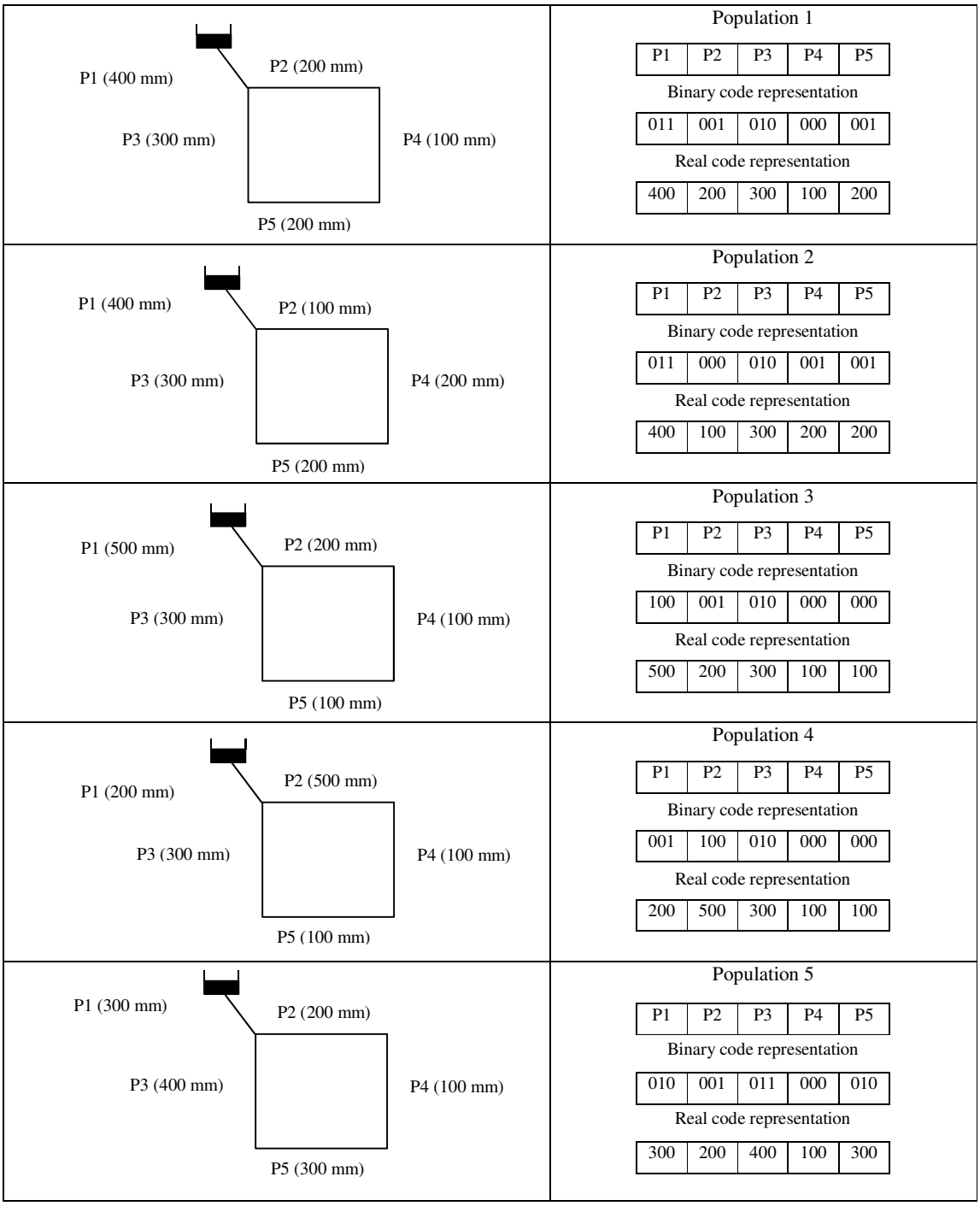


Figure 3.4 Binary and real code representations of five randomly generated water distribution network designs; available pipe sizes are (100, 200, 300, 400 and 500 mm)

3.3 Recent Developments in Genetic Algorithms for Water Distribution Systems Optimisation

GAs have been dominantly used in the literature of the WDS optimisation field. This includes pump operation scheduling (Goldberg and Kuo, 1987; Mackle et al.,1995; Rao and Salomons, 2007), network design and rehabilitation (Savic and Walters, 1997; Dandy and Engelhardt, 2001; Halhal et al., 1997; Siew et al., 2014), network calibration (Vitkovsky and Simpson, 1997), water quality optimization (Munavalli and Mohan-Kumar, 2003; Farmani et al., 2006) and tank siting and sizing (Vamvakeridou-Lyroudia et al., 2005; Prasad, 2010; Siew and Tanyimboh, 2011a). Many developments have been carried out to enhance GAs search capabilities and its computational efficiency in finding optimal and near optimal solutions. Among others implementing different coding systems, modifying genetic operators, reducing search space, applying various constraint handling methods, parallelization of GA are included.

3.3.1 Advances in Representation Scheme and Genetic Operators

Savic and Walters (1997) and Dandy et al. (1996) proposed the use of Gray coding as opposed to binary coding to overcome convergence problems related to the hamming cliff effect. Vairavamoorthy and Ali (2000) and Kadu et al. (2008) avoided the encoding and decoding of diameter variables and the issue of redundancy by implementing real coding. Dandy et al. (1996) proposed the use of an adjacency mutation operator based on the assumption that good solutions have the tendency to lie close each other. The mutation operator differs from the conventional bitwise mutation in that it mutates a complete decision variable substring to an adjacent decision variable substring located up or down the list of the design variable candidates, rather than randomly flipping bits from the solution string. Mutation plays an important role particularly in the later generations of the GA where the population is highly dense with good solutions. It

encourages exploring the search space further and increasing the chance of finding better solutions. Taking this into consideration Kadu et al. (2008) implemented a non-uniform mutation rate instead of a fixed mutation rate. The mutation probability increases from 1% to 10% as the search proceeds. Saleh and Tanyimboh (2014) proposed a seamless generic procedure for dealing with redundant codes. The approach was implemented for optimal design of WDSs based on entropy and topology where the redundant codes represent closed pipes whose flow-carrying capacity is zero. The closed pipes are assigned pipe sizes taken from just above the upper end of the real set of available pipe diameters. As the assumed diameters have no functional value, it is expected that they will become extinct through evolution and natural selection.

3.3.2 Enhancement to Computational Efficiency Related to the Use of the Network Solver

During the GA process, a network solver is required to satisfy the governing network constraints (i.e. mass balance and energy conservation) for each member of the population under consideration, hence resulting in a high computational cost. In attempt to reduce the computational time required by the network solver, Vairavamoorthy and Ali (2000) proposed the use of a regression model that avoids the need for a hydraulic solver for each member of the population. The model is derived from regression analysis of several network designs. The model approximates the hydraulic behaviour of the networks and it is less computationally demanding. Hindi and Hamam (1993) and Savic and Walter (1995) proposed changes to the stopping criteria of the network solver to reduce the total number of iterations required to satisfy the governing equations.

3.3.3 Search Space Reduction

The size of the search space is one of the main factors that influence the effectiveness and efficiency of the GA search in WDSs optimisation (Kadu et al. 2008). The size of the search space can be wider when the network size is larger and more candidate pipe sizes included. This can lead to an increase in computational time and reduce the chance of obtaining global optimum solution. Restricting the number of candidate pipe diameters involved is one way of reducing the search-space size significantly. However, this has to be carried out in a clever way as an incorrect choice of candidate pipe sizes can result in a sub-optimal solution. In an effort to reduce the search space size, Vairavamoorthy and Ali (2005) proposed a GA framework for the least-cost pipe network design problem that excludes regions of the search space where infeasible solutions are likely to exist. The formulation uses heuristics, which attempt to measure the relative importance of each pipe with regard to the overall performance of the network. This involves a hydraulic simulation of the network and the solution of a system of linear equations with multiple right hand terms both of which impose a very high computational burden. Kadu et al. (2008) reported a spectacular reduction in WDS GA execution times employing a heuristic based on a critical path concept to limit the GA search to a region of the solution space where near optimal solutions were thought to exist. The approach is built on the assumption that the cheapest mode of delivering water from the source to a demand node is through the shortest available path. They minimized the search space by reducing the number of pipe diameter options for each link. However, their method is limited in scope as the critical path concept is not applicable to the more industrially relevant problem of rehabilitation or networks with pumps, multiple operating conditions and time varying demands.

3.3.4 Constraint Handling Techniques

GAs are unconstrained optimisation procedure and they are incapable of differentiating feasible solutions from infeasible ones. Nevertheless, many real world optimisation problems such as WDSs are constrained optimisation problem, which necessitates appropriate handling of constraints. The most common approach to handle constraints in GAs is to use penalty functions (Savic and Walters, 1997; Vairavamoorthy and Ali, 2000). The idea of using penalty function is to convert a constrained-optimisation problem into unconstrained by penalizing infeasible solutions based on the amount of their constraint violation in order to concentrate the GA search within the feasible region of the solution space. For this, an additional penalty cost is applied to the actual WDS cost of the infeasible solution. The penalty cost is calculated based on penalty parameters that are formulated in such a way that it penalizes any violations of the constraints. The amount of violation determines the probability of solutions being discarded in the next generation. Too high penalty will confine the GA search to the feasible region of the solution space potentially resulting in very expensive solutions. On the other hand, too low penalty moves the GA search towards the infeasible region. Therefore, an extensive and exhaustive fine-tuning of penalty parameter requires before the penalty function can be effectively incorporated into the GA framework.

Many researchers have attempted to address the difficulties associated with penalty function. Khu and Keedwell (2005) converted each node pressure constraint into objective functions to avoid the use of penalty function. However, this procedure imposes huge computational demands when applied to real-life WDSs with large number of nodes. Prasad and Park (2004) adopted a constraint handling method, which does not require a penalty parameter (Deb, 2000). The approach uses a tournament selection operator in which feasible solutions are favourably selected over infeasible solutions. Also, when infeasible solutions are compared, those with smaller constraint violations are preferentially selected over those with larger violations. Among feasible

solutions, solutions with better objective functions preferentially selected over those with lower objective function values. However, this way of handling constraint ranks the most expensive feasible solution over a cost effective marginally infeasible solution, which can be acceptable in practice. Also, the method impedes the GA from using the good feature of marginally infeasible solutions to reach an optimal solution effectively. Wu and Simpson (2002) and Wu and Walski (2005) developed a self-adaptive penalty method based on a heuristic boundary search technique. The approach involved an evolving penalty factor that aimed to focus the GA search around the boundary between feasible and infeasible solution space. However, the implementation of this self-adaptive heuristic technique requires the prior calibration of several additional parameters. The self-adaptive penalty method proposed by Afshar and Marino (2007) uses a penalty parameter the value of which evolves during the execution of the algorithm. The ratio of the cost of the best feasible and infeasible solutions at each generation is used to guide the GA search towards the boundary of the feasible region by automatically adjusting the value of the penalty parameter. The procedure requires an initial value of the penalty parameter to be defined at the start of the optimization process. High and low initial values of the penalty parameter lead to searches inside and outside of the feasible region, respectively. Farmani et al. (2005) proposed a self-adaptive fitness formulation, which does not require any parameter calibration. The approach involves implementation of a two-stage penalty. The first stage ensures that the worst infeasible solution is assigned a penalty cost that is higher or at least equal to the cost of the best solution in the population. The penalty cost of the worst infeasible solutions is then further increased to the cost of the most expensive solution. The remaining infeasible solutions are penalised exponentially in proportion to their infeasibility. However, the approach allows the low-cost infeasible solutions to be selected over feasible solutions with higher costs. In doing so, there may be a possibility that the population will be over-dominated by infeasible solutions and the search may converge to infeasible solutions. Siew and Tanyimboh (2012b) have proposed an alternative approach that avoids the use of penalty functions or other special constraint handling techniques. In particular, the method does not

involve any parameters that require prior calibration. The procedure involves a pressure dependent analysis (PDA) within a multi-objective optimisation search. GAs, by their nature, are heuristic and generate enormous quantities of infeasible solutions, which are pressure deficient. Unlike the conventional method of analysis, PDA can simulate infeasible candidate solutions accurately. The approach allows all the feasible and infeasible solutions generated to compete in a way that is fundamentally bias-free with respect to constraint violation. The algorithm promotes full exploitation of all feasible and infeasible solutions generated to guide the GA search toward the boundary between infeasible and feasible regions. In most recent time, Saleh and Tanyimboh (2014) have adopted the penalty-free approach for optimal design of WDSs based on entropy and topology. The algorithm includes a special procedure for gauging the infeasibility of solutions. Overall, the penalty-free approach generated superior results for all the optimization problems solved in terms of cost, hydraulic performance and computational efficiency compared to other solutions in literature.

The penalty-free approach has been adopted in the research carried out herein due to its proven robustness and effectiveness. Comprehensive details on the method are provided in Chapter 6.

3.3.5 Multi-Objective Optimisation Developments

The optimisation of real-world WDSs involves several objectives, e.g, minimize cost, maximize performance, maximize reliability, etc., which are often conflicting in nature. The traditional single objective optimisation approach mainly aimed at finding the best solution, which corresponds to the minimum or maximum value of a single objective function that combines all different objectives. This type of optimisation approach cannot yield a set of alternative solutions that provide trade-off among different objectives. On the contrary, multi-objective optimisation approach produces a set of

non-dominated solutions, known as Pareto-optimal set. Each solution of the Pareto optimal set is not dominated by any other solution; the solutions are all of equal optimality. It is not possible to make improvement on one objective without making at least one of the other objectives worse. Pareto optimal solutions are found to be practical since the final solution of the decision maker is always a trade-off. For instance, engineers can evaluate the trade-off between different designs and make engineering judgment to obtain a practical solution based on the performance requirement and the available budget. Depending on the preference of a decision maker, a group of designs can be chosen from the Pareto optimal set for more detail analysis. For this reason, Pareto optimal solution sets are often preferred to single solutions (Konak et al., 2006).

The population-based evolutionary algorithms such as GAs are well suitable to solve multi-objective optimisation problems (Gen et al., 2008; Konak et al., 2006). The ability of the algorithms to explore different regions of a solution space simultaneously enables them to find a diverse set of solutions for difficult problems with non-convex, discontinuous, and multi-modal solutions spaces. Evolutionary multi-objective optimisation algorithms can be grouped as elitist and non-elitist (Deb, 2001). Examples of non-elitist algorithms include Vector Evaluated Genetic Algorithm (Schaffer, 1985), Multi-objective Genetic Algorithm (Fonseca and Fleming, 1993) and Non-dominated Sorting Genetic Algorithm (Srinivas and Deb, 1994). Strength Pareto Evolutionary Algorithm (SPEA) (Zitzler and Thiele, 1998), Elitist Non-Dominated Sorting Genetic Algorithm (NSGA-II) (Deb et al., 2002) and Pareto Archived Evolution Strategy (PAES) (Knowles and Corne, 2000) are the most common elitist methods. Elitism is one of the key factors for successful application of multi-objective EAs (Bekele and Nicklow, 2005; Kollat and Reed, 2006). It helps to prevent the loss of good solutions and in achieving better convergence in MOEA (Zitzler et al., 2000). Farmani et al. (2003) have assessed the application of some of the elitist and non-elitist multi-objective EAs in WDSs. The study indicated that the elitist Non-dominated Sorting Genetic Algorithm method (NSGA-II) (Deb et al., 2002) outperformed other approaches in terms

of finding a diverse set of solutions and generating Pareto-optimal solutions. NSGA-II is popular mainly due to its efficient non-dominated sorting procedure and strong global elitist approach implemented. It preserves elitism by combining parent and child populations before they are sorted and ranked into non-dominated fronts. By doing this, all elites from both the parent and child populations are preserved. This enables the algorithm to provide better spread of solutions and convergence near the optimal solutions. One of the advantages of NSGA-II is that it requires only few users' specified parameters (Dridi et al., 2008).

NSGA-II is one of the widely used multi-objective GA in many aspects of the WDS optimization researches. Atiquzzaman et al. (2006) presented a multi-objective genetic algorithm that couples NSGA-II with EPANET to produce optimal and less optimal solutions with the aim of providing alternative design solution within available budget and tolerated pressure deficit. Farmani et al. (2006) optimised the design and operation of a complex WDS ("Anytown" network) that involves multiple loading conditions, pump scheduling, tank siting and sizing using the NSGA II. The problem was formulated as a multi-objective optimization problem with total cost, water quality and reliability as the objectives. Jayaram and Srinivasan (2008) proposed a multi-objective approach for the optimal design and rehabilitation of a water distribution network, with minimization of whole-life cost and maximization of performance as objectives. In the area of WDS security, Jeong and Abraham (2006) proposed a multi-objective optimization model for an operational response strategy to mitigate consequences of deliberate attacks on WDS. Preis and Ostfeld (2008) and Weickgenannt et al (2010) employed NSGA-II for optimal sensor placement to detect contamination in WDSs. Nicolini et al. (2011) adopted a multi-objective optimisation approach for optimal leakage management in water distribution network. The objectives considered were minimization of the number of valves and minimization of the total leakage in the system. Siew and Tanyimboh (2012b) introduced a penalty-free multi-objective evolutionary approach (PF-MOEA) for design optimization of WDS. The approach

utilizes pressure dependent analysis (PDA) coupled with NSGA-II. PDA is able to simulate both normal and pressure deficient networks and provides the means to identify the feasible region of the solution space effectively. Siew et al. (2014) applied PF-MOEA for the phased whole-life design and rehabilitation of WDSs. The optimization model considers the initial construction, rehabilitation and upgrading costs. Repairs and pipe failure costs are included. The model also takes into consideration the deterioration over time of both the structural integrity and hydraulic capacity of every pipe.

Czajkowska and Tanyimboh (2013) proposed a maximum entropy-based multi-objective genetic algorithm approach based on NSGA-II for the design optimization of WDS under multiple operating conditions. Saleh and Tanyimboh (2013, 2014) developed a multi-objective optimisation approach using NSGA-II for simultaneous layout, pipe size and entropy-based optimisation of WDSs. The optimization method is based on a general measure of hydraulic performance that combines statistical entropy, network connectivity and hydraulic feasibility. Barlow and Tanyimboh (2014) developed a multi-objective memetic algorithm for least-cost design of WDSs. The algorithm hybridizes NSGA-II with local and cultural improvement operators. A description on the NSGA-II procedure is provided next.

In the NSGA II procedure, an initial population is generated randomly. Each individual in the population is then assigned fitness and ranked based on its non-domination level. Using this population, the operators of tournament selection, crossover and mutation are used to create an offspring population. To preserve elitism, the parent and offspring populations are combined into a single population that is then sorted into various levels of non-domination. The next generation is created by first selecting solutions belonging to the best non-dominated front of the combined parent and offspring population, and then subsequent non-dominated fronts in the order of their ranking. If the last accepted front has more solutions than required to achieve the population size, the solutions

having the largest crowding distances (higher diversity) are chosen. This whole procedure is repeated until a pre-specified number of generations are reached.

It is important to note that NSGA II employs a binary tournament selection operator where the selection criterion is based on the crowded-comparison operator. This operator requires both the non-domination rank and crowded distance value of each solution in the population. The diversity among non-dominated solutions is introduced by using the crowding comparison procedure, which is used in the tournament selection and during the population reduction stage. Given two solutions of S1 and S2, S1 is considered as the winner of the tournament if either S1 has a higher non-domination rank than S2 or if both have the same non-domination rank but S1 has a larger crowding distance than S2.

Crowding distance is very important in guiding the selection process at the various stages of NSGA II towards a uniformly distributed Pareto optimal front. The calculation of crowding distance in the NSGA II requires the solutions to be sorted in ascending order according to each objective considered. To maintain the boundary solutions for each objective function, solutions with the largest and smallest objective values are assigned an infinite distance value. The distance values for intermediate solutions are then calculated as the absolute normalised difference in the function values of two adjacent solutions. The overall crowding distance for the solution is the sum of the individual distance corresponding to each objective.

3.3.6 Global- Local Hybrid Search

One of the major advantages of GAs when applied to optimisation problem is that their capability to incorporate other local search and optimisation techniques within its framework to yield a hybrid, which achieves the best from the combination (Gen, 2008).

GAs are global search method and can rapidly locate the region where the global optimum exists. Nevertheless, the algorithms take a relatively long time to locate the exact global optimum in the region of convergence (De Jong, 2005; Preux and Talbi, 1999). On the other hand, local search methods can converge to local optimal solutions rapidly; however, they have no capacity to identify global optima. A hybrid method that couples the exploration of GAs with the exploitation of local methods, therefore, has the potential to enhance the search to locate the exact global optimum and reduce the computation time required (El-Mihoub et al., 2006). However, one of the key concerns in hybrid algorithms is that the efficiency and effectiveness of the search is dependent on the division of the global and local search time (Goldberg and Voessner, 1999). Choosing a proper balance between global and local search may not be a straightforward and can actually be considered as an optimisation problem itself (e.g. see Goldberg and Voessner, 1999).

Some of the recently proposed hybrid methods to solve WDSs optimisation problems include the algorithm developed by Haghighi et al. (2011). The algorithm combines GAs with integer linear programming (ILP). It requires defining a path from a source to each node arbitrarily. In doing so, a single pipe was excluded from each loop and the looped network was converted into a branched one. All excluded pipe sizes were determined by GAs while the rest of the pipes are optimized by ILP. ILP returns the optimum diameter to GAs and the evolution process continues until the convergence criteria are satisfied. The algorithm was applied on two networks from literature (each has 34 pipes) and a significant reduction in computational time has been reported. However, the algorithm has been applied on a looped network operating under single demand loading condition. The applicability of the algorithm in solving real-world WDSs problems such as networks operating under multiple operating conditions and the inclusion of pumps and storage reservoirs has not been assessed. Barlow and Tanyimboh (2014) introduced a multi-objective memetic algorithm that hybridises GA with local and cultural improvement operators for least-cost WDS design problem. The hybrid algorithm was

formulated in such a way that the GA operates as normal, except that at a certain generation interval the local and cultural operators are applied to create child population. The performance of the algorithm was assessed in reference to GA using two benchmark networks that have 34 and 454 pipes. Substantial improvement in the algorithm's computational efficiency has been reported. However, it was noted that the performance of a network in Barlow and Tanyimboh (2014) approach is measured based on the total nodal pressure deficit throughout the network. A drawback of this approach is that it generates only the least cost feasible solution. The approach is not capable of recognizing other feasible solutions. For this, only solutions that are located in the low-deficit region of objective space are more likely to be modified through local and cultural improvement operators to obtain the optimal solution. The local and cultural improvement operators are not applied to other individuals located in any other region of objective space. By contrast, a procedure that guides the search approaching the optimal solution from both sides of the region of the objective space is more efficient for exploring and exploiting the search space (Dong and Wang, 2014).

Wang et al. (2015) compared the performance of two hybrid search procedures named as high and low-level hybrid algorithms (Talbi, 2002) in reference to NSGA II. The high–level hybrid algorithm combines different algorithms that are operating independently within the optimisation framework while the low-level hybrid algorithm embedded various algorithms within the NSGA-II. The performances of the hybrid algorithms and that of NSGA II were assessed based on the non-dominated solutions generated by the algorithms. The authors concluded that the hybrid algorithms provide a more diversified Pareto front for small and medium size networks and better convergence and diversity for large size network in reference to NSGA II. However, the performance assessment was carried out using graphical plots only; and no rigours analysis (e.g. quantitative comparison) was carried out to examine the outcome of the algorithms.

3.3.7 Parallelization of Multi-Objective Evolutionary Algorithms

As described in Section 3.3.5, MOEAs have been widely used due to their potential in generating Pareto optimal solutions that provide decision makers flexibility in choosing a particular solution. Despite their benefit, MOEAs require a large number of function evaluations before generating good results (Shinde et al., 2011). This has limited the algorithms' potential for practical applications to solve real-world WDSs optimisation problem. One way of addressing this difficulty is by implementing parallel computing approach that is capable of reducing the computational time and makes the algorithms' convergence to optimal solutions faster. In parallel computing large problems are often broken into several ones that can be solved simultaneously on separate processors (Trobec et al., 2009). Since EAs are population based search approach, they are suitable for being implemented in parallel computing architectures (Cantu-Paz, 2000). The genetic operator such as crossover, mutation, and in particular the time-consuming fitness evaluation can be performed independently on different processors (Shinde et al., 2011). The most common parallelisation approaches that have been applied in WDS literature are the controller-worker model (single-population implementation) and island model (multiple-population implementation). The two types of parallel GAs have been extensively used to reduce the execution time of a variety of applications. The choice between the two parallel GAs' approaches is determined by factors such as ease of implementation and their potential to reduce execution time.

3.3.7.1 Controller-Worker Model (Single-Population Scheme)

Controller-worker model is one of the successful implementations of parallel GAs (Nowostawski and Poli, 1999; Cantu-Paz and Goldberg, 2000). It employs a single population and the evaluation of the individuals is performed in parallel. The selection and mating is done globally in which each individual may compete and mate with any

other. The approach is very efficient particularly for problems that require considerable computations for fitness evaluation (Cantu-Paz and Goldberg, 2000). Fitness evaluation requires only the knowledge of the individual being evaluated. No communication is required during this stage. Parallelisation of fitness evaluation is performed by assigning a fraction of the population to each of the processors available. Communication occurs only when each worker receives the individual to evaluate and when the workers return the fitness values. Controller-worker GA explores the search space in exactly the same manner as serial GA (Digalakis and Margaritis, 2003; Nowostawski and Poli, 1999). Consequently, the existing design guidelines for serial GA are directly applicable when employing the controller-worker GA. The method can achieve significant speedup if the communication costs are small in reference to the computation costs (Kumar et al. 2006; Cantu-Paz and Goldberg, 2000). Its ease of implementation in particular makes the model more popular with practitioners (Alba, 2005; Castillo et al., 2008; Cantu-Paz, 2000).

3.3.7.2 Island Model (Multiple-Population Scheme)

In island-model, the population is divided into few sub-populations that evolve serially and independently. Every processor (island) runs an independent GA using a separate sub-population. The processors communicate by exchanging some solutions (migrant solutions) occasionally (Nowostawski and Poli, 1999). In doing so, islands that were trapped in low-fitness regions of the search space can be taken over by individuals in the more successful islands. The island model fundamentally changes the GA search dynamics relative to the serial GA (Back et al., 1997). Thus, its difference from serial GA makes it hard to make a good and efficient comparison. Nevertheless, the model when designed well can be used to solve very complex problems for which the serial GA performs poorly regardless of the search duration (Tang et al., 2007). The communication over-heads required in the island model are much less than the

controller-worker model due to small population migration between islands (Digalakis and Margaritis, 2003). The model has high diversity as it has the working principle of isolated population with migration (Shinde et al., 2011). This produces a faster evolution of each island, which can cause a rapid convergence to optimal solution. However, despite the benefits of island model, it is much more difficult to design and assess the approach (Cantù-Paz and Goldberg, 2000). The application of the model is not straightforward, as it involves the choice of several parameters that require calibration (Artina et al., 2012). Among other things, one must decide the size of the island populations, the frequency of migration, and the number of migrant solutions and their destination. Traditionally, these parameters are set by trial and error. However, the trial and error approach may result in an inefficient use of computing resources as well as poor-quality solutions.

3.3.7.3 Parallel Genetic Algorithm for Optimal Design of WDSs

Despite the inherent characteristics of GA as being easily parallelised, only a handful of WDS optimisation research has been carried out to investigate its parallel implementations. Balla and Lingireddy (2000) reported a computational improvement by implementing a GA based parallel algorithm for optimal model calibration of WDS. A controller-worker approach was implemented on three networked PCs. One of the PCs was designed as a controller processor and the remaining as worker processors. The controller processor was dedicated to perform GA search, distribute solutions to and collect fitness evaluations values from the worker processors. The worker processors were in charge to perform hydraulic simulations and fitness calculation, and to return fitness values to controller process. Also, Wu and Zhu (2009) implemented the controller-worker model for pump scheduling optimisation to improve energy efficiency. Ewald et al. (2008) implemented an island model, which has an elitism strategy, for optimal location of booster chlorination stations in WDSs. Artina et al.

(2012) implemented both the controller-worker and island models for optimal design of networks. A good speed-up was reported from the two parallel implementations up to a limited number of processors. The researchers noted that for the island model, migration improves the final cost of feasible solutions in reference to a configuration that has no communication. Also, comparison with the serial version of the algorithm indicated that frequent and massive exchange of good solutions provides better results. Similarly, Barlow and Tanyimboh (2014) reported a substantial improvement in computational speed by applying the two parallel implementation methods for least cost design of WDSs. The controller-worker model was implemented to speed up the progress of the evolution on a single optimisation run, where the controller executes the routine operation of the GA and employs the workers to perform the fitness evaluations. Additionally, the island model was employed to perform independent optimisation runs simultaneously.

3.4 Conclusions

The optimisation problems of WDSs are complex and require computationally efficient algorithms, which are reliable and easy to implement. The classical optimization techniques though computational efficient have great limitations to solve highly constrained non-linear problems with discrete solution spaces. These techniques are incapable of solving multi-objective problems effectively. In contrast, the population-based evolutionary algorithms are well suitable to solve multi-objective optimization problems due to their capability to search different regions of a solution space simultaneously. This key feature makes them suitable for handling highly complex combinatorial problems with non-convex, multi-modal and discontinuous solution space. A comprehensive review on genetic algorithm that is the most popular of several EAs in WDSs optimisation is provided. The general procedure and operators involved in the standard GA have been detailed. The chapter examines the existing approaches

proposed to enhance the effectiveness and efficiency of GA for WDSs optimisation. The limitations of the approaches were identified and discussed. A review on two most commonly applied parallel GA implementations in WDSs optimisation is presented. Controller-worker is a straightforward and efficient and implementation of parallel GA. The method can achieve significant speedup if the communication costs are small in comparison to the computation costs. On the other hand, the island model implementation, which works on independent population with migration, provides a high diversity that enables a faster evolution of each island and leads to a rapid convergence to optimal solution. However, the application of the model is not straightforward, as it involves the choice of several parameters that require calibration.

Chapter 4

A New Pressure Dependent EPANET Extension Algorithm

4.1. Introduction

Water distribution models are used extensively in the design and operation of water distribution systems (WDSs) to help predict hydraulic and water quality behaviour within networks under a wide range of operating conditions. Nowadays, more attention has been given to the use of hydraulic models for solving optimisation problems in WDSs by combining them with evolutionary optimisation algorithms (EAs). The models evaluate the conservation of mass and energy constraints and check other constraints such as nodal pressures for any violation. EAs by nature are heuristic and generate a large number of infeasible designs, which are extremely pressure-deficient. This has given rise to the need for evaluating the performances of infeasible designs correctly. As explained in Chapter 2 (Section 2.4), the conventional demand driven analysis (DDA) models are incapable of simulating pressure-deficient networks. By contrast, the pressure dependent analysis (PDA) models gauge the performances of pressure-deficient designs realistically.

As discussed in Chapter 3 (Section 3.3.4), EAs (such as genetic algorithms) are unconstrained optimisation procedures and they are unable to differentiate feasible and infeasible solutions. Nevertheless, many real world optimisation problems such as

WDSs are constrained optimisation problem, which necessitates appropriate handling of constraints. Previous studies on constraint or infeasible solutions handling approaches demonstrated the benefits of explicitly maintaining infeasible solutions (pressure deficient designs) for single and multi-objective constrained optimization problems (Singh et al., 2008; Ray et al., 2009). The presence of infeasible solutions guides the algorithm search towards an optimal solution from both feasible and infeasible side of the search space. The recent WDSs optimization algorithms that retain infeasible solutions in full achieved superior results in terms of algorithms' convergence rate or quality of solutions compared to those algorithms that reject or penalize infeasible solutions (Siew and Tanyimboh, 2012b; Saleh and Tanyimboh, 2013; Siew et al., 2014). This further reinforces the importance of an efficient PDA that can simulate the pressure-deficient networks more accurately.

In this chapter, a computationally efficient and robust PDA model has been proposed. This is crucial particularly in the context of optimization of real-life networks with hundreds or thousands of pipes where the computational time for evaluating millions of designs could be limiting. The proposed PDA is an enhancement of the pressure dependent extension of the EPANET hydraulic simulator (EPANET-PDX) that was developed by Siew and Tanyimboh (2012a). As discussed in Chapter 2 (Section 2.4.2.1) EPANET-PDX has an integrated logistic nodal head-flow function proposed by Tanyimboh and Templeman (2010) coupled with a line search and backtracking procedure to facilitate convergence. Extensive testing conducted on the model previously using benchmark as well as real life networks revealed good modelling performance. Also, the model was combined with a multi-objective genetic algorithm for optimization of WDSs and superior results were obtained compared to other previous solutions (Siew and Tanyimboh, 2012b; Siew et al., 2014). Overall, no convergence issues were reported while the model executed millions of simulations.

Given the robustness and benefits of EPANET-PDX previously, including seamless integration in genetic algorithms, it seems beneficial to investigate ways of improving

the model further. In this chapter, a new line search and backtracking algorithm for integrating the logistic nodal head-flow function into the system of hydraulic equations in the global gradient algorithm is proposed. This has increased the robustness further by enhancing greatly the computational properties for low-pressure conditions (for extremely pressure deficient networks) and increasing the algorithm's consistency over a wider range of operating conditions. Several extended period simulations were executed, for a real life network that comprises multiple supply sources and various demand categories for ranges of normal and pressure-deficient operating conditions. The new algorithm has also been applied on two benchmark networks and overall good computational performances have been achieved. Details of the results and computational efficiency of the algorithm are presented.

4.2. Pressure Dependent Extension of EPANET

EPANET 2 model adopted the global gradient algorithm (GGA) proposed by Todini and Pilati (1988) to solve the flow continuity and head loss equations of pipe networks (Rossman, 2000). Siew and Tanyimboh (2012a) enhanced the model to enable it to simulate pressure dependent flows. The enhanced pressured dependent model (EPANET-PDX) integrates the continuous nodal head-flow function that Tanyimboh and Templeman (2010) proposed in the GGA. As described in Chapter 2 (Section 2.4.2), the major advantage of the nodal head-flow function is that, it consists of a single continuous function that applies to the full range of flow conditions; i.e., zero, partial, and full flow. It provides a smooth transition between normal and pressure deficient operating conditions. Moreover, the function and its first derivative have no discontinuities and created no convergence problems when integrated into systems of equations for WDSs. As shown in Chapter 2 (Section 2.4.2), the Tanyimboh and Templeman (2010) nodal head-flow function is expressed as

$$Qn_i(Hn_i) = Qn_i^{req} \frac{\exp(\alpha_i + \beta_i Hn_i)}{1+\exp(\alpha_i + \beta_i Hn_i)} \tag{4.1}$$

and its first derivative described as

$$\frac{dQn_i(Hn_i)}{dHn_i} = Qn_i^{req}\beta_i \quad \frac{\exp(\alpha_i + \beta_i Hn_i)}{(1+\exp(\alpha_i + \beta_i Hn_i))^2} \tag{4.2}$$

where, for node i, Qn_i and Hn_i are the flow and head respectively; Qn_i^{req} is the demand; α_i and β_i are parameters determined using field data.

As shown in Chapter 2 (Section 2.3.1.2.4), the systems of hydraulic equations in the GGA are solved simultaneously as:

$$\begin{bmatrix} \mathbf{A}_{11} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Qp} \\ \cdots \\ \mathbf{Hn} \end{bmatrix} = \begin{bmatrix} -\mathbf{A}_{10}\mathbf{H}_0 \\ \cdots \\ -\mathbf{Qn}^{\mathbf{req}} \end{bmatrix} \tag{4.3}$$

where $\mathbf{A_{11}}$ is a diagonal matrix whose elements are $K_{ij}(Qp_{ij})^{n_f-1}$. K_{ij} and n_f are the resistance coefficient and flow exponent in the head loss formula respectively. Qp_{ij} is the flow rate in pipe ij. $\mathbf{A_{12}}$ and $\mathbf{A_{10}}$ represent the overall incidence matrix relating the pipes to nodes with unknown and known heads respectively. Pipe flow leaving node is defined as -1, pipe flow into node as +1 and 0 if pipe is not connected to node. $\mathbf{A_{21}}$ is the transpose of $\mathbf{A_{12}}$. $\mathbf{Qp}$ denotes the column vector of unknown pipe flow rates. $\mathbf{Hn}$ and $\mathbf{H}_0$ are column vectors for unknown and known nodal heads respectively. $\mathbf{Qn^{req}}$ is the column vector for required nodal supplies.

To incorporate the nodal head-flow function, Eq. 4.3 is formulated as

$$\begin{bmatrix} \mathbf{A}_{11} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Qp} \\ \cdots \\ \mathbf{Hn} \end{bmatrix} = \begin{bmatrix} -\mathbf{A}_{10}\mathbf{H}_0 \\ \cdots \\ 0 \end{bmatrix} \tag{4.4}$$

where $\mathbf{A}_{22}$ is a diagonal matrix with the elements $Qn(Hn)/Hn$. $Qn(Hn)$ is the nodal head-flow function described in Eq. 4.1. Eq. 4.4 can be rearranged as follows:

$$-\begin{bmatrix} \mathbf{A}_{11} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Qp}^{k} \\ \cdots \\ \mathbf{Hn}^{k} \end{bmatrix} + \begin{bmatrix} -\mathbf{A}_{10}\mathbf{H}_0 \\ \cdots \\ 0 \end{bmatrix} = -\begin{bmatrix} \mathbf{f}_1 \\ \cdots \\ \mathbf{f}_2 \end{bmatrix} \tag{4.5}$$

where $\mathbf{f}_1 = \mathbf{f}_1(\mathbf{Qp}, \mathbf{Hn})$ and $\mathbf{f}_2 = \mathbf{f}_2(\mathbf{Qp}, \mathbf{Hn})$ indicate how far from zero the relevant equations are for any given approximated solution $\mathbf{Qp}, \mathbf{Hn}$.

The pressure dependent network analysis formulation indicated in Eq. 4.4 is solved based on successive linearization using a first order Taylor series expansion that gives

$$\begin{bmatrix} \mathbf{D}_{11} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{D}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Qp}^{k+1} - \mathbf{Qp}^{k} \\ \cdots \\ \mathbf{Hn}^{k+1} - \mathbf{Hn}^{k} \end{bmatrix} = -\begin{bmatrix} \mathbf{A}_{11} & \vdots & \mathbf{A}_{12} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{21} & \vdots & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Qp}^{k} \\ \cdots \\ \mathbf{Hn}^{k} \end{bmatrix} + \begin{bmatrix} -\mathbf{A}_{10}\mathbf{H}_0 \\ \cdots \\ 0 \end{bmatrix} \tag{4.6}$$

where $\mathbf{D}_{11}$ is the diagonal matrix of the flow derivatives of the pipe head losses whose elements are $nK_{ij}(Qp_{ij})^{n_f-1}$. $\mathbf{D}_{22}$ is the diagonal matrix of the head derivatives of the nodal head-flow function whose elements are described in Eq. 4.2. k represents the iteration number. Eq. (4.6) gives the following formulas to compute the heads and flows iteratively.

$$\mathbf{Hn}^{k+1} = \mathbf{A}^{-1}\mathbf{F} \tag{4.7}$$

$$\mathbf{Qp}^{k+1} = \mathbf{Qp}^{k} - \mathbf{D}_{11}^{-1}(\mathbf{A}_{11}\mathbf{Qp}^{k} + \mathbf{A}_{12}\mathbf{Hn}^{k+1} + \mathbf{A}_{10}\mathbf{H}_{0}) \tag{4.8}$$

where

$$\mathbf{A} = \mathbf{A}_{21}\mathbf{D}_{11}^{-1}\mathbf{A}_{12} - \mathbf{D}_{22} \tag{4.9}$$

$$\mathbf{F} = -\mathbf{A}_{21}\mathbf{D}_{11}{}^{-1}\mathbf{A}_{11}\mathbf{Qp}^{k} - \mathbf{A}_{21}\mathbf{D}_{11}{}^{-1}\mathbf{A}_{10}\mathbf{H}_{0} - \mathbf{D}_{22}\mathbf{Hn}^{k} + \mathbf{A}_{22}\mathbf{Hn}^{k} + \mathbf{A}_{21}\mathbf{Qp}^{k} \tag{4.10}$$

The GGA procedure described above is for steady state network analysis. Detail description on extended period analysis procedure is presented in Chapter 2 (Section 2.3.2).

EPANET-PDX utilized the line search and backtracking procedure (Press et al., 1992) to help ensure global convergence in the integration of the nodal head-flow functions in GGA. The application of the line search and backtracking procedure in EPANET-PDX, however, was somewhat limited, in an attempt to preserve the excellent computational properties of EPANET 2. An improved implementation herein that allows more iterations of the line search procedure has been developed.

4.3. A New Algorithm of the Line Search and Backtracking Procedure

This section describes a new algorithm of the line search and backtracking procedure (Press et al., 1992) in EPANET for pressure-deficient networks modelling. The equations for conservation of energy along hydraulic links and mass balance at nodes described in Eq. 4.5 above can be expressed as follows:

$$\mathbf{f}_1(\mathbf{Hn}, \mathbf{Qp}) = \mathbf{A}_{11}\mathbf{Qp}^{k} + \mathbf{A}_{12}\mathbf{Hn}^{k} + \mathbf{A}_{10}\mathbf{H}_{0} \tag{4.11}$$

$$\mathbf{f}_2(\mathbf{Hn},\mathbf{Qp}) = \mathbf{A}_{21}\mathbf{Qp}^k + \mathbf{A}_{22}\mathbf{Hn} \tag{4.12}$$

Together, Eq. 4.11 and 4.12 form a single system of simultaneous non-linear equations that is denoted by $\mathbf{F}(\mathbf{Hn},\mathbf{Qp})$ the solution of which is required. The aim of the line search and backtracking procedure is to find an appropriate Newton step size that decreases the function g, which is given in Eq. 4.13, in successive iterations sufficiently.

$$g = \frac{1}{2}\mathbf{F}.\mathbf{F} \tag{4.13}$$

To implement the line search and backtracking procedure, a scalar parameter λ for the Newton step has been introduced. Using λ (step size), the nodal heads $\mathbf{Hn}$ are updated iteratively as

$$\mathbf{Hn}^{k+1} = \mathbf{Hn}^{k} + \lambda.\boldsymbol{\delta}\mathbf{Hn} \tag{4.14}$$

λ satisfies $0 < \lambda \leq 1$; k represents the iteration number; $\delta\mathbf{Hn}$ is the full Newton step for the nodal heads. The pipe flow rates $\mathbf{Qp}^{k+1}$ are then updated by substituting the newly obtained nodal heads $\mathbf{Hn}^{k+1}$ into Eq. (4.8).

At each iteration of the GGA, the line search procedure checks the full Newton step first. If the Newton step does not reduce the value of g sufficiently, backtracking along the Newton direction is carried out in a series of minor iterations to obtain an acceptable step. The criterion for acceptance of a Newton step is

$$g(\mathbf{Hn}^{k+1},\mathbf{Qp}^{k+1}) \leq g(\mathbf{Hn}^{k},\mathbf{Qp}^{k}) + \alpha.\nabla g.\delta \tag{4.15}$$

where the parameter α satisfies $0 < \alpha < 1$; $\alpha = 10^{-4}$ (Press et al., 1992). ∇g is the

gradient of $g(\mathbf{Hn}, \mathbf{Qp})$ and $\delta = [\delta\mathbf{Hn}, \delta\mathbf{Qp}]$ is the Newton step. The initial rate of decrease of *g* is defined in Press et al. (1992) as

$$\nabla g.\delta = (\mathbf{F}.\mathbf{J}).(-\mathbf{J}^{-1}.\mathbf{F}) = -\mathbf{F}.\mathbf{F} \tag{4.16}$$

where **J** represents the Jacobian matrix. During backtracking that is when $g(\mathbf{Hn}^{k+1}, \mathbf{Qp}^{k+1})$ fails to meet the acceptance criterion in Eq. (4.16), $g(\mathbf{Hn}^{k+1}, \mathbf{Qp}^{k+1})$ is modelled as a quadratic function of λ by substituting $\mathbf{Hn}^{k+1}$ with $\mathbf{Hn}^{k} + \lambda.\delta\mathbf{Hn}$. The value of λ that minimize the function is obtained. For the second and any subsequent backtracks, $g(\mathbf{Hn}^{k+1}, \mathbf{Qp}^{k+1})$ is modelled as a cubic function of λ. The backtracking procedure continues until either Eq. (4.16) is satisfied or λ reaches the minimum set value $\lambda_{\min}$. Siew and Tanyimboh (2012a) used $\lambda_{\min}$ =0.2 to avoid the algorithm from taking too small steps that would result in long computational time. The same value has been used in this work as well.

Fig. 4.1 shows a flow chart that describes the line search and backtracking implementation. At the first iteration, the algorithm starts from an initial guess of nodal heads and pipe flow values. Nodal elevations are considered as the initial nodal heads values while the initial pipe flow rates are calculated based on an assumed velocity of 1 ft/sec (0.3048 m/sec). In each iteration of the computational solution, the Euclidean norm of the energy and mass balance was evaluated to ensure the algorithm did not converge spuriously and the real solution has been found. At the solution, the norm should approach a value of zero as an indication of the progress and accuracy of the algorithm.

It was noted that in some cases when λ reaches the minimum set value of 0.2 the norm fails to reduce in consecutive iterations. In such situations, a new nodal head is computed using the available Newton step (based on $\lambda = 0.2$) and the GGA iteration continues until the convergence criteria of the algorithm are fulfilled (See Fig. 4.1). The

algorithm may converge prematurely if this additional protection measure has not been provided. As the algorithm converges, the changes in pipe flow and nodal heads become insignificant. Thus, the EPANET-PDX's convergence criteria of 0.001 ft (3.048×10^{-4} m) for the maximum change in the nodal heads and 0.001 cfs (2.832×10^{-5} m^3/s) for the maximum change in pipe flows between successive iterations have been preserved herein.

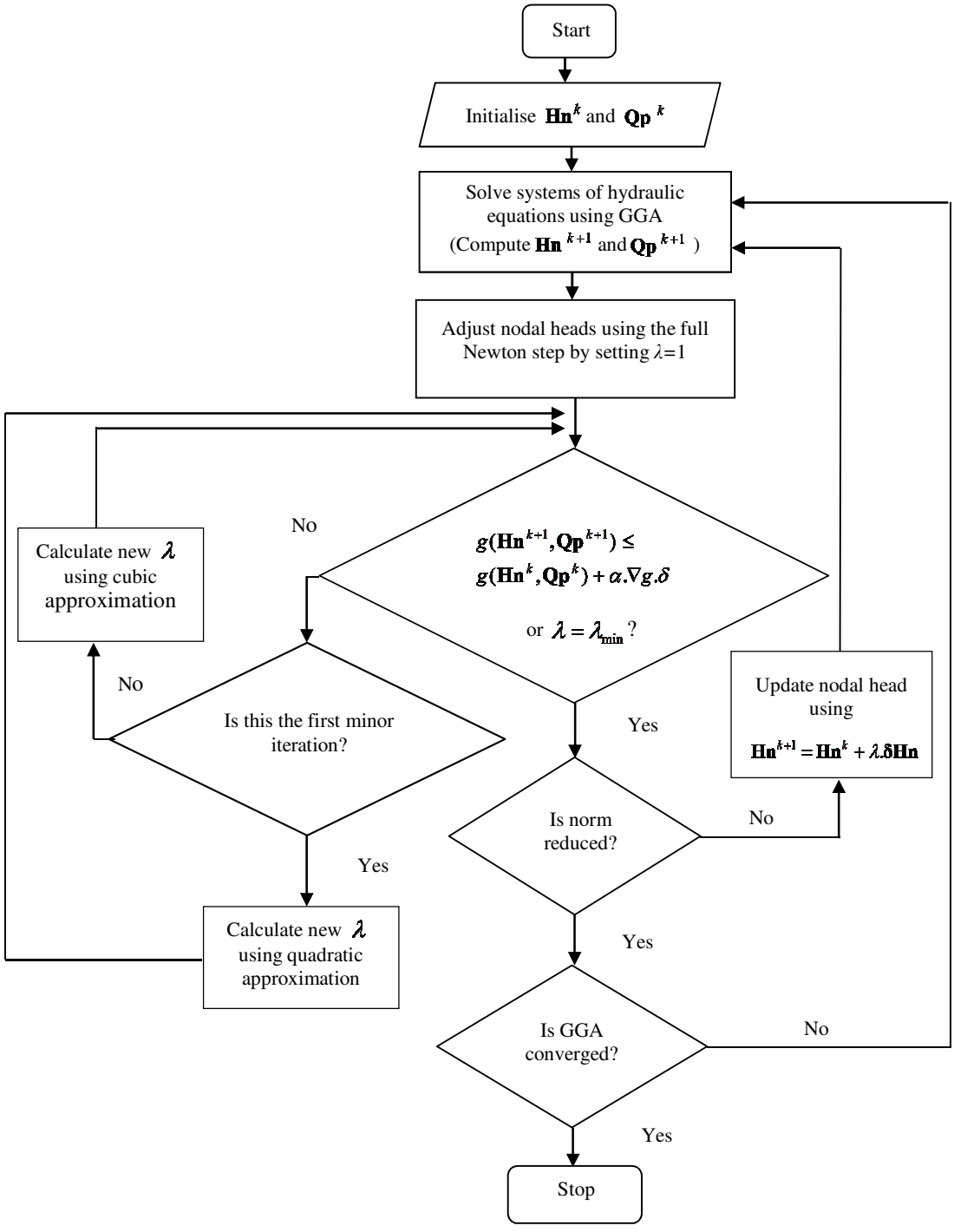


Figure 4.1 Line search implementation

The line search procedure implemented in EPANET-PDX (Siew and Tanyimboh, 2012a) deviates from the classical line search implementation in an attempt to exploit the excellent computational properties of EPANET-2. The approach intermingles the line search and backtracking procedure with the GGA. In other words, the line search routine is not only generating optimum step sizes but also it performs GGA iterations. By doing this, the opportunity to use GGA was maximized and very often, the algorithm makes progress using the full step sizes. As a result, the algorithm bypasses the line search in most of the cases. However, as explained earlier, the sole purpose of the line search procedure is to provide an optimum step size that decreases the function $g(\mathbf{Hn}, \mathbf{Qp})$ in successive iterations sufficiently. Major iteration (GGA) should not, therefore, be involved in line minimization.

The new line search implementation herein by far has a clear separation between the line search procedure and GGA. The algorithm checks progress for every iteration. If the algorithm does not make progress, the line minimization or the line search will be carried out. The new implementation provided more scope for the line search procedure to change the course of the algorithm. Overall, the current implementation utilized the line search in full.

4.4. Application of the Proposed Algorithm

The proposed pressure dependent algorithm is applied on three network examples. The main aim of the examples is to verify the accuracy, robustness and computational efficiency of the algorithm in the perspective of both small and large networks subjected to various levels of pressure operating conditions. The first two examples are benchmark networks taken from literature and used mainly to assess the computational performance of the proposed model in terms of number of iterations and CPU time required for convergence. Furthermore, the accuracy of the algorithm is demonstrated in the second

example by analysing the network's hydraulic results. The third example is based on a real-life network in the UK. It is used to illustrate the capability of the proposed algorithm to handle large networks with multiple operating conditions efficiently and accurately. Comprehensive assessment on the algorithm was carried out by executing several extended period simulations for various sources heads and pipe closure conditions. Most importantly, this example focuses on demonstrating the computational efficiency of the algorithm on modelling extremely low-pressure conditions. All simulations for the three examples were carried out on an Intel Xeon workstation (2 processors of CPU 2.4 GHz and RAM of 16 GB). For simplicity, the proposed algorithm is named hereinafter as EPANET-PDX (0.2) while EPANET-PDX is named EPANET-PDX (0.1). In all the simulations, the assumed head below which nodal flow is zero is equal to the nodal elevation.

4.4.1 Example 1

The example shown in Fig. 4.2 is a single source network taken from Todini (2003). The network consists of 11 demand nodes and 17 pipes of length 500 m with the Hazen-Williams roughness coefficient of 130. Details of nodes and pipes are provided in Table 4.1. Two sets of runs were carried out when applying the model to analyse the network. In the first set of runs, three steady state simulations were carried out by assuming various required pressures of 30 m, 20 m, and 10 m at all nodes. For all simulations, the source head was fixed at 150 m. In the second set of runs, four steady state simulations were executed by varying the source heads to 150 m, 125 m, 100 m and 75 m. The required pressure was fixed at 30 m at all nodes in all the simulations.

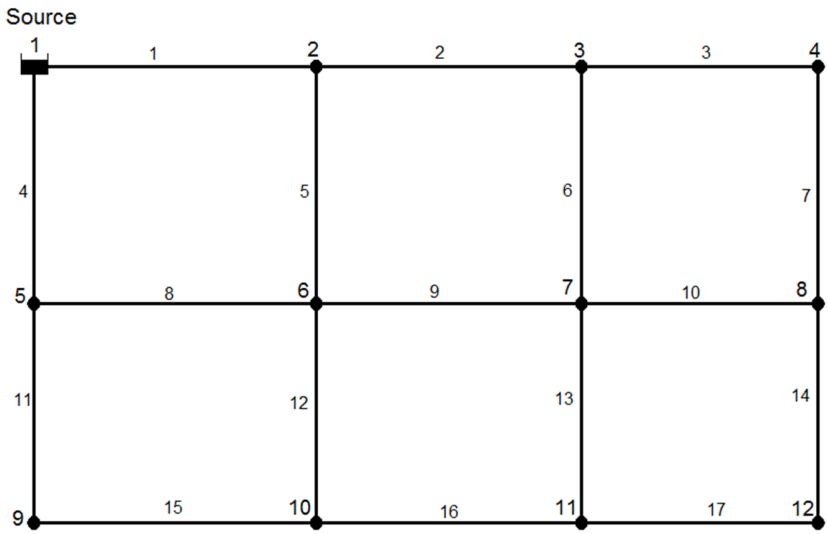


Figure 4.2 Layout of Example 1

Table 4.1 Node and pipe data for Example 1

Node	Elevation (m)	Demand (m3/s)	Pipe	Diameter (mm)
2	90	0.05	1	250
3	80	0.02	2	200
4	85	0.04	3	150
5	85	0.05	4	250
6	90	0.05	5	200
7	85	0.05	6	150
8	90	0.01	7	100
9	85	0.04	8	200
10	80	0.05	9	150
11	90	0.04	10	100
12	100	0.01	11	200
			12	150
			13	100
			14	100
			15	150
			16	100
			17	100

Table 4.2 summarises the computational performances of EPANET-PDX (0.1) and (0.2) for all the executed simulations. Generally, EPANET-PDX (0.2) provided smaller number of iterations, while for CPU time both models provided comparable results.

Table 4.2 Number of iterations and CPU time required to converge for Example 1

	First run			Second run			
	Required nodal pressure (m)			Reservoir head (m)			
	30 (88%)	20 (89%)	10 (91%)	150 (88%)	125 (71%)	100 (29%)	75 (0.04%)
	Number of iterations						
EPANET-PDX (0.1)	9	11	29	9	9	10	8
EPANET-PDX (0.2)	9	10	16	9	9	7	7
	CPU time (Seconds)						
EPANET-PDX (0.1)	0.03	0.03	0.04	0.03	0.04	0.03	0.03
EPANET-PDX (0.2)	0.05	0.04	0.04	0.05	0.03	0.03	0.03

Values in parentheses represent the network demand satisfaction ratio (DSR)

4.4.2 Example 2

This example is based on "Anytown" benchmark network shown in Fig. 4.3. The network presented previously as an optimisation problem to find cost effective design to upgrade the existing system to meet future demands (Walski et al., 1987). Siew and Tanyimboh (2012a) modified the original input data to enable pressure dependent analysis. The diameters of the six new pipes (10, 13, 14, 15, 16 and 25) were set to be 0.3048 m (12 in) and demands for nodes 2, 4, 5, 9, 10, 12 and 15 were reduced to 3.155 l/s. The modified network is used herein to execute a steady state simulation. Pipe and node data of the network can be found in Appendix A (Table A-1. (1)- (4)). Also, the

demand factors that represent the diurnal demand variation can be referred in Appendix A (Fig. A-1.1).

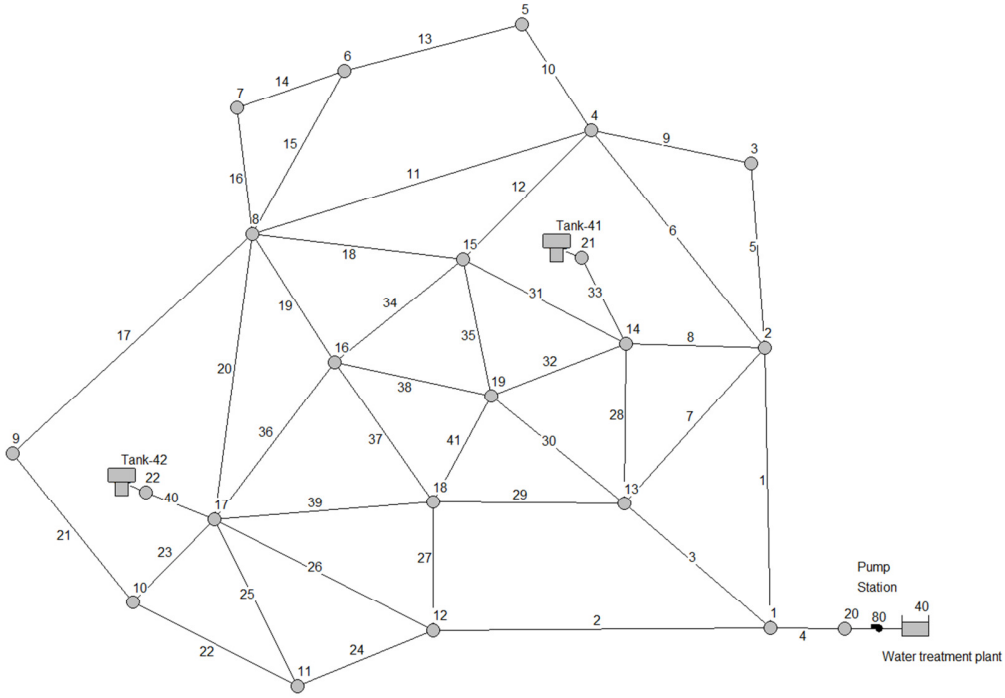


Figure 4.3 Layout of Example 2

The network is supplied from a treatment plant via three identical pumps operating in parallel. The water level in the treatment plant was fixed at 3.05 m (10 ft) while the required pressure for all nodes was 28.12 m (40 psi). The network has two storage tanks (Tanks 41 and 42); and the pipes connecting the tanks were closed off. The network was operating under pressure deficit conditions with DSR of 87%. Comparisons of nodal heads and nodal outflows for all nodes are presented in Fig. 4.4 and 4.5 respectively, where both EPANET-PDX (0.1) and (0.2) generated essentially identical results. The average numbers of iterations required for EPANET-PDX (0.1) and (0.2) to converge were 13 and 10 respectively. Both models required on average 0.04 seconds to complete the simulation.

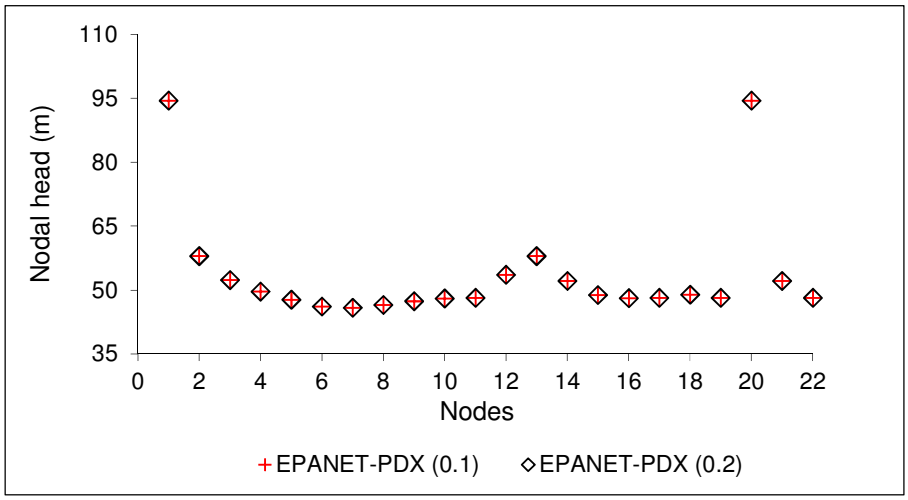


Figure 4.4 Nodal heads for Example 2

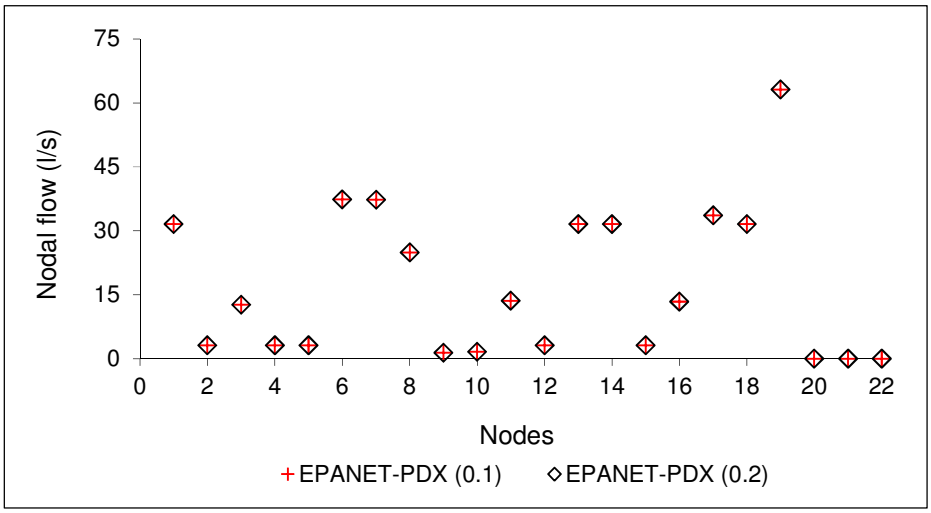


Figure 4.5 Nodal flows for Example 2

4.4.3 Example 3

In addition to the benchmark networks in Example 1 and 2, the performance of EPANET-PDX (0.2) was assessed in detail using a large network. The large network is a hydraulic demand zone (referred to hereafter as water supply zone) in the UK. Fig. 4.6

shows the network that consists of 251 pipes of various lengths, 228 demand nodes (including the fire hydrants), 3 demand categories and 29 fire hydrants at various locations; pipe diameters are 32-400 mm. The network obtains water entirely from the neighbouring water supply zones through five variable head supply nodes (i.e. nodes R1-R5 in Fig. 4.6). Diagrams that illustrate the temporal variations of the demand categories and the variable-head supply nodes can be found in Chapter 5. The range of variation in the head levels at each supply node is insignificant. Hence, to simplify interpretation of results, the supply nodes were modelled as constant-head nodes with water levels of 155 m each. The network and dynamic operational data used here for the hydraulic analyses are taken from calibrated EPANET model and a geographical information system (GIS) database. The EPANET model contains node data that include elevations and demands for individual nodes and the relevant demand categories. The demand categories comprise domestic demand, 10-hour commercial demand and unaccounted for water, for duration of 96 hours. The network also has 29 different fire demands of 1hour each. These are applied at the 29 fire hydrants located at different positions in network. Also in the EPANET model are link data, including pipe lengths, diameter and roughness values; the Darcy-Weisbach pipe friction head loss formula was used for the hydraulic analyses. The information in the GIS database includes data on the pipes such as age, material, diameter, length, renewal and / or rehabilitation year, burst rate and other data of a geospatial nature. The required residual head at all demand nodes is 20 m.

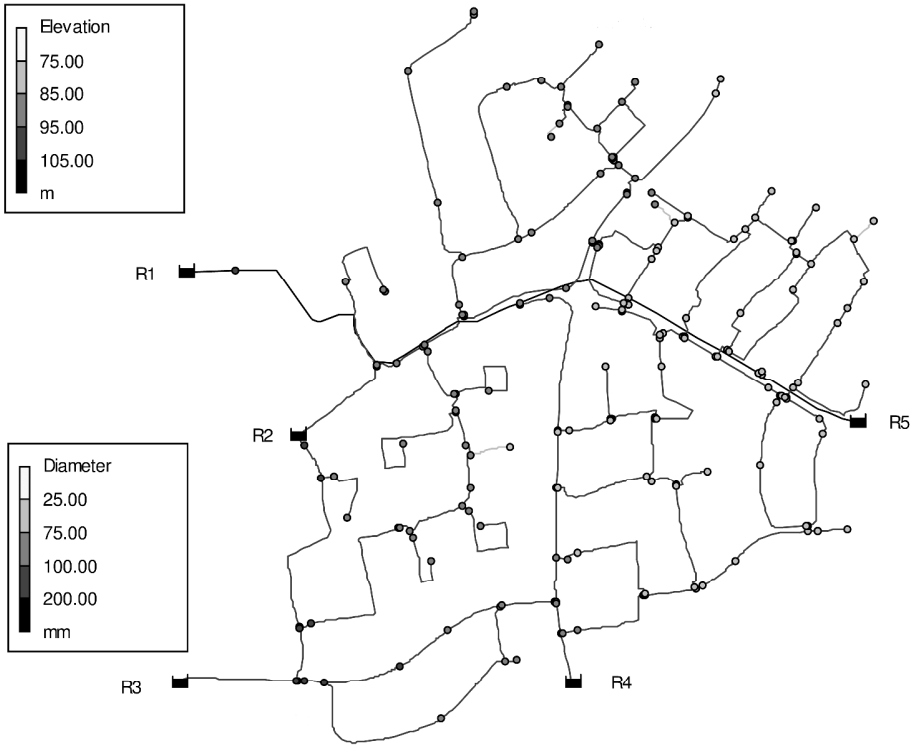


Figure 4.6 Network layout for Example 3

All the analyses reported on this network are extended period simulations (EPSs). Each EPS covered a period of 31 hours, based on a 1-hour hydraulic time step. For all three models considered here namely EPANET 2 and both versions of EPANET-PDX, 66 extended period simulations were executed in total for normal and pressure deficient conditions considering source head variations and pipe closure conditions.

4.4.3.1 Source Head Variation

The computational performance, the accuracy and robustness of EPANET-PDX (0.2) was evaluated by analysing the network under the entire range of demand satisfaction by varying the supply node heads from 75 m to 130 m in equal steps of 1 m. 56 EPSs were

carried out without convergence complications. Fig. 4.7 shows a comparison of EPANET-PDX (0.1) and (0.2), for the average hourly network demand satisfaction ratios. Identical results were obtained for the hydraulic simulations, for the entire range of demand satisfaction ratios.

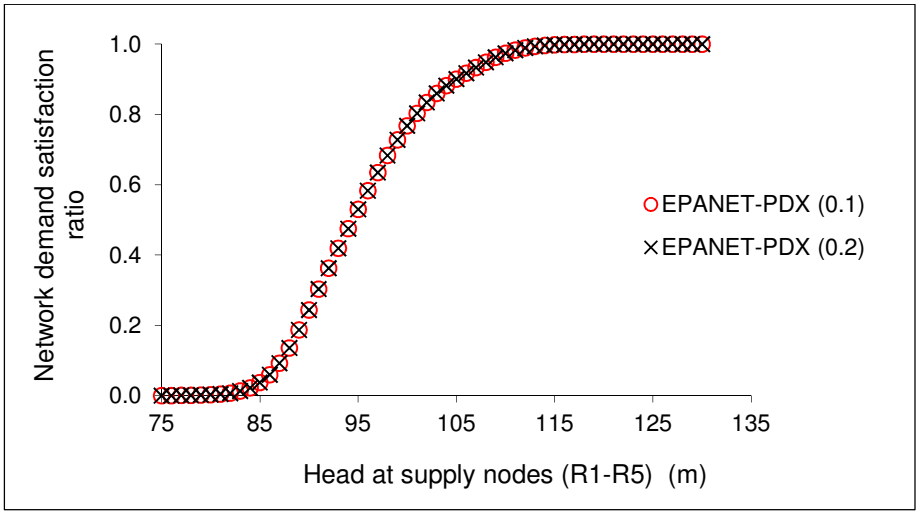


Figure 4.7. Influence of variations in supply node heads on the flow delivered for Example 3

For comprehensive analysis of the computational performance of EPANET-PDX (0.2), three scenarios were considered while evaluating the model. This was conducted by dividing the full network performance in Fig. 4.7 into 3 categories.

4.4.3.1.1 Normal Pressure Condition

The network is considered operating under normal condition when more than 99.9% of the total demand is satisfied. This corresponds to heads at supply node that ranges from 116 m to 130 m in Fig 4.7. The number of iterations needed to solve the system of hydraulic equations as a function of pressure in the network is shown in Fig. 4.8. The average numbers of iterations required per simulation were 7.00, 5.00 and 5.16 for

EPANET-PDX (0.1), EPANET-PDX (0.2) and EPANET 2 respectively. EPANET-PDX (0.2) required the smallest numbers of iterations. Fig. 4.9 compares the CPU time. On average EPANET-PDX (0.1) and (0.2) that use line minimization required about 0.27 seconds and 0.29 seconds, respectively, per EPS compared to 0.15 seconds for EPANET 2. EPANET 2 in general is more efficient and consistent. It is worth re-stating, however, that EPANET 2 and EPANET-PDX apply convergence criteria that are not identical (Siew and Tanyimboh, 2012a). The criterion used in EPANET 2 is the ratio of the sum of the absolute values of pipe flow changes to the total flow in all pipes should be less than 0.001. Therefore, it is worth emphasizing that the EPANET 2 results here provide a rough guide rather than an absolute direct comparison.

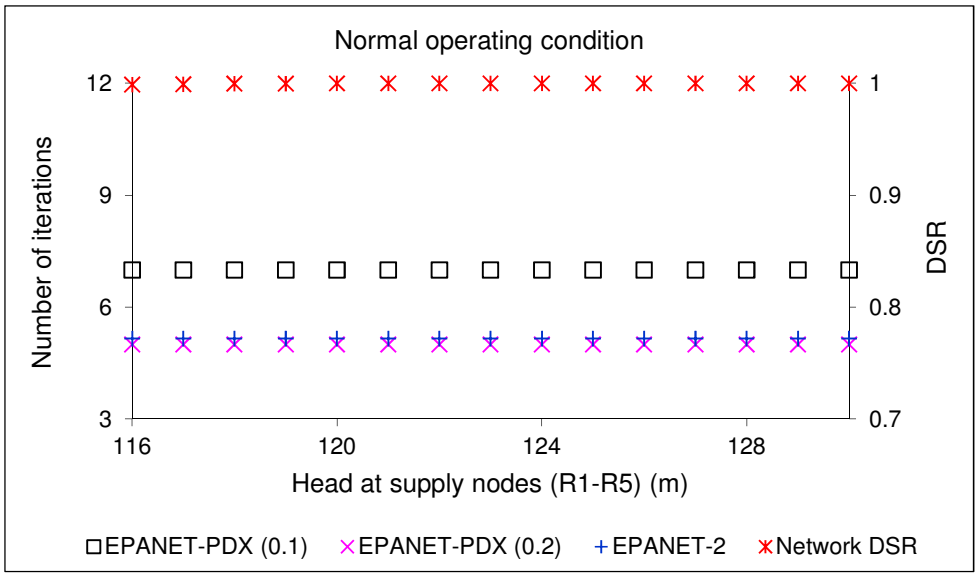


Figure 4.8 Number of iterations required as a function of the available pressure in the network for normal operating condition for Example 3

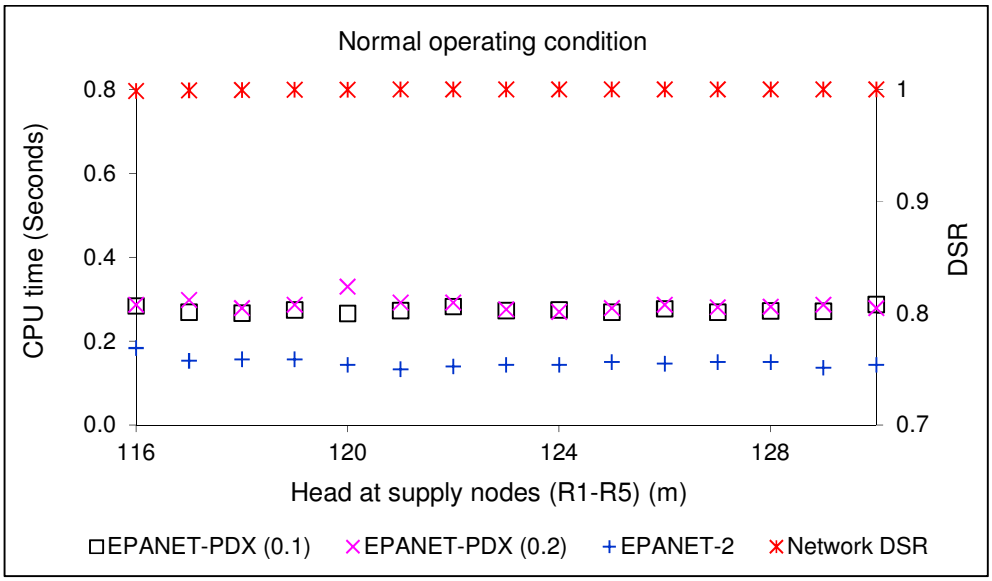


Figure 4.9 Comparison of CPU time for EPANET 2 and EPANET-PDX for normal operating condition for Example 3

4.4.3.1.2 Pressure Deficient Condition

The computational performance of EPANET-PDX (0.2) for a pressure deficient condition was assessed in reference to EPANET-PDX (0.1) for the network DSRs that range between 9.23% and 99.9%. The DSRs represented the network performances when operating just below full demand satisfaction and just above extremely low-pressure condition. The corresponding heads at supply nodes varies between 88 m and 115 m in Fig. 4.7. The average number of iterations and CPU time are compared in Fig. 4.10 and 4.11 respectively. It is worth mentioning that EPANET 2 results are not included in this and the next scenario as the model is not entirely suitable for pressure deficient operating conditions. Similar computational performance as the normal operating condition has been achieved for the pressure deficient operating condition. The average numbers of iterations required per simulations were 7.00 and 5.04 for EPANET-PDX (0.1) and EPANET-PDX (0.2) respectively. EPANET-PDX (0.2) required the smaller numbers of iterations. For CPU time, the models provided

comparable results; i.e. 0.27 seconds and 0.29 seconds for EPANET-PDX (0.1) and EPANET-PDX (0.2) respectively, per EPS.

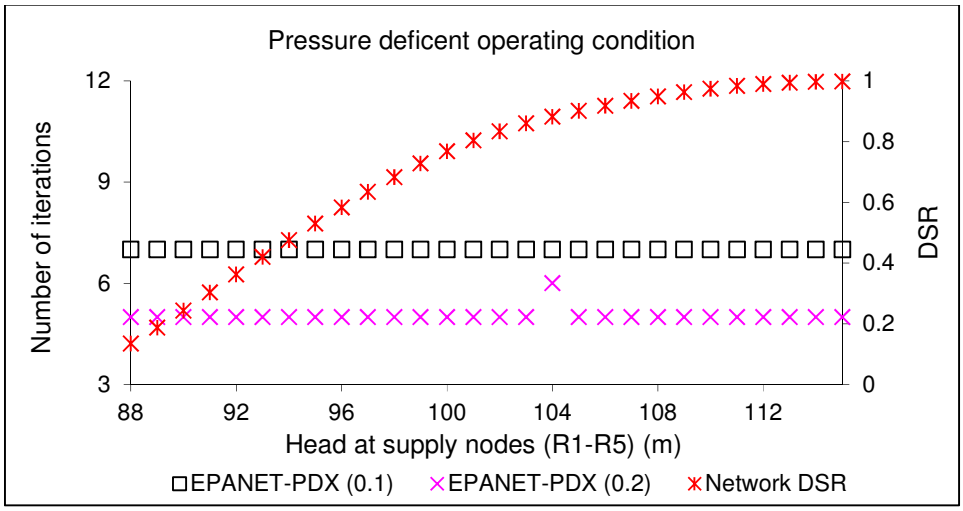


Figure 4.10 Number of iterations required as a function of the available pressure in the network for pressure deficient operating condition for Example 3

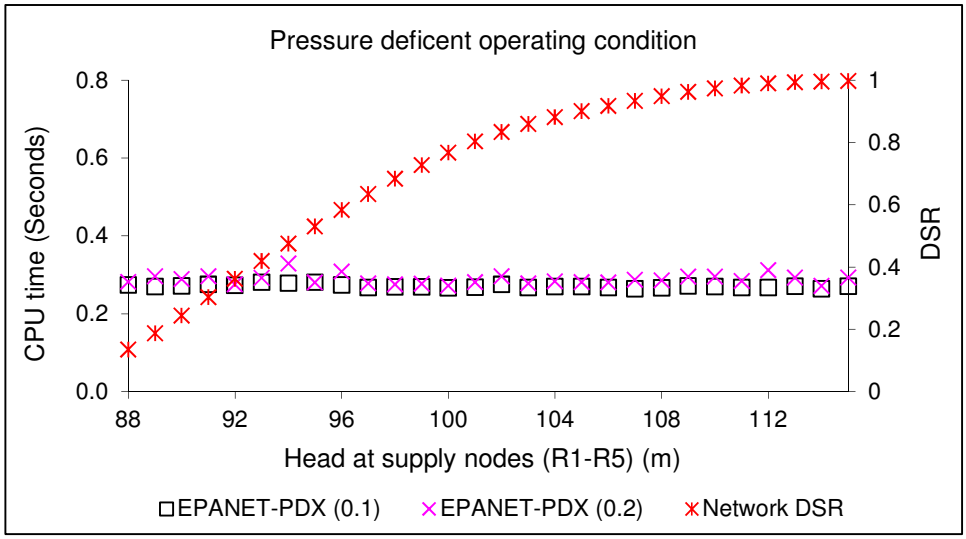


Figure 4.11 CPU time required as a function of the available pressure in the network for pressure deficient operating condition for Example 3

4.4.3.1.3 Extremely Low Pressure Condition

The efficiency of EPANET-PDX (0.2) has been evaluated for extremely low-pressure condition. This is essential from the perspective of design optimisation of WDSs using evolutionary algorithms (EAs). In the application of evolutionary algorithms in WDSs, a large number of extremely pressure deficient designs (infeasible designs) are generated. As described in Section 4.1, recent evidence on handling infeasible solutions (designs) in WDSs optimisation problem demonstrated the benefit of retaining the infeasible designs in full as they usually contain useful genetic materials that can lead to optimal designs (for example smaller pipe sizes can lead to minimum cost solution).

In order to evaluate these designs a computationally efficient PDA simulator requires. Most importantly, in the optimisation of real life networks with large number of pipe sizes the computational time to evaluate millions of designs over the progress of EAs could be limiting. Hence, any improvement on the computational performance of PDA simulator would be greatly beneficial.

EPANET-PDX (0.2) is primarily developed to enhance the computational properties for extremely pressure deficient networks. To assess the model's computational performances, the heads of the supply nodes were reduced to simulate an extremely low-pressure condition. A network DSR below 9.23% was considered as an indicative cut-off point for extremely low-pressure conditions in Fig 4.7. The corresponding heads at supply nodes varies between 75 m and 87 m. Fig. 4.12 and 4.13 show the number of iterations and CPU time needed to solve the system of hydraulic equations as a function of the pressure in the network. EPANET-PDX (0.2) achieved a significant improvement in both numbers of iterations and CPU time. The average numbers of iterations required per simulation were 6.85 and 4.08 for EPANET-PDX (0.1) and EPANET-PDX (0.2) respectively. The average CPU time required were 0.38 seconds and 0.28 seconds for EPANET-PDX (0.1) and EPANET-PDX (0.2) respectively. It was noted that EPANET-

PDX (0.2) performs consistently well for normal, pressure deficient and extremely low flow conditions on the whole. However, EPANET-PDX (0.1) is quite variable in performance during extremely low flow condition when the supply node heads are very low.

Fig. 4.14 shows the number of line search evaluations needed as a function of the pressure in the network. The line search procedure requires several evaluations (minor iterations) to determine an optimum step size that satisfies the line search criteria. In general, EPANET-PDX (0.1) does not need line search evaluations. For extremely low pressures condition, the algorithm requires only one line search evaluation per simulation. The result evidently indicated that the line search and backtracking procedure in EPANET-PDX (0.1) is restricted. On the other hand, EPANET-PDX (0.2) that has no restriction on the line search procedure requires on average 10.303 evaluations per simulations.

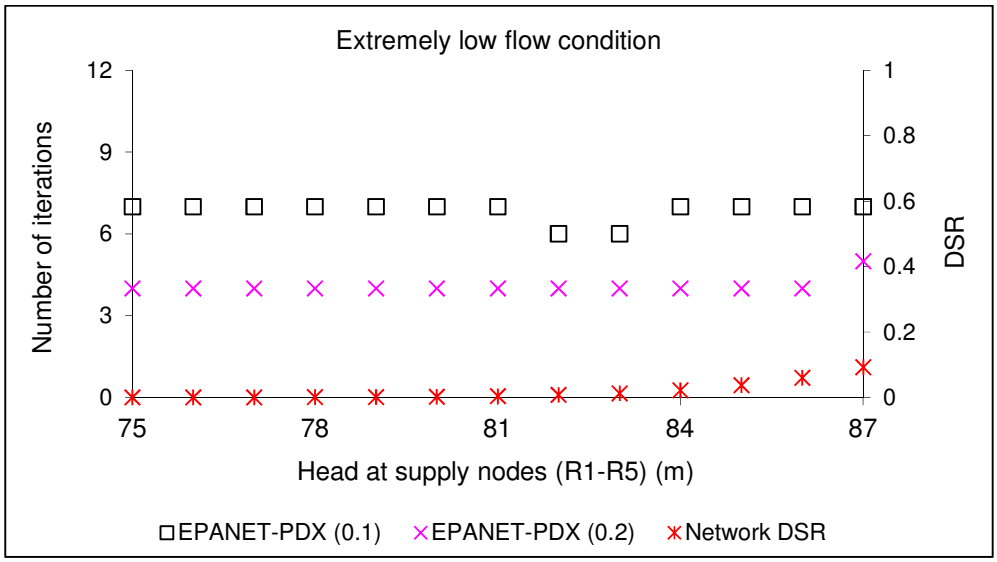


Figure 4.12 Number of iterations required as a function of the available pressure in the network for extremely low-pressure condition for Example 3

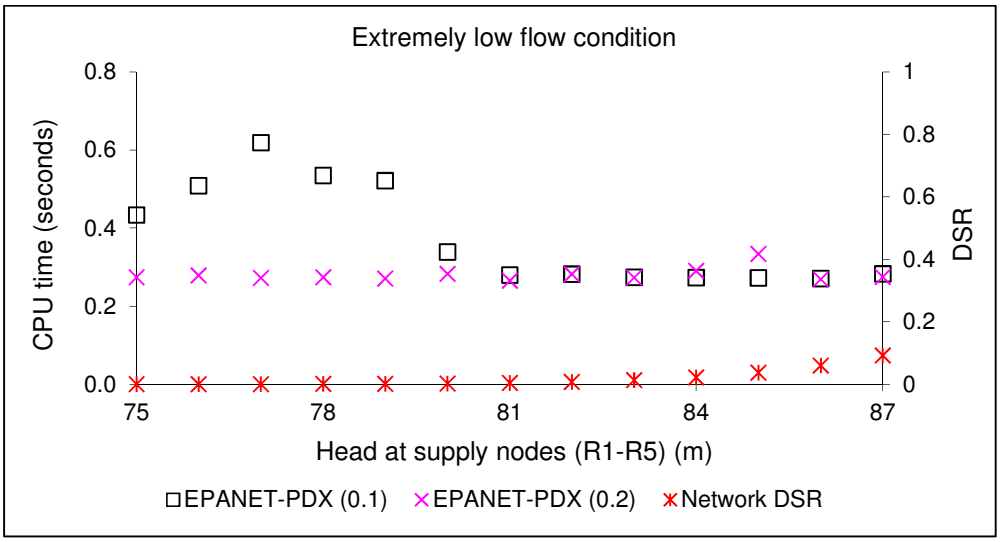


Figure 4.13 CPU time required as a function of the available pressure in the network for extremely low-pressure condition for Example 3

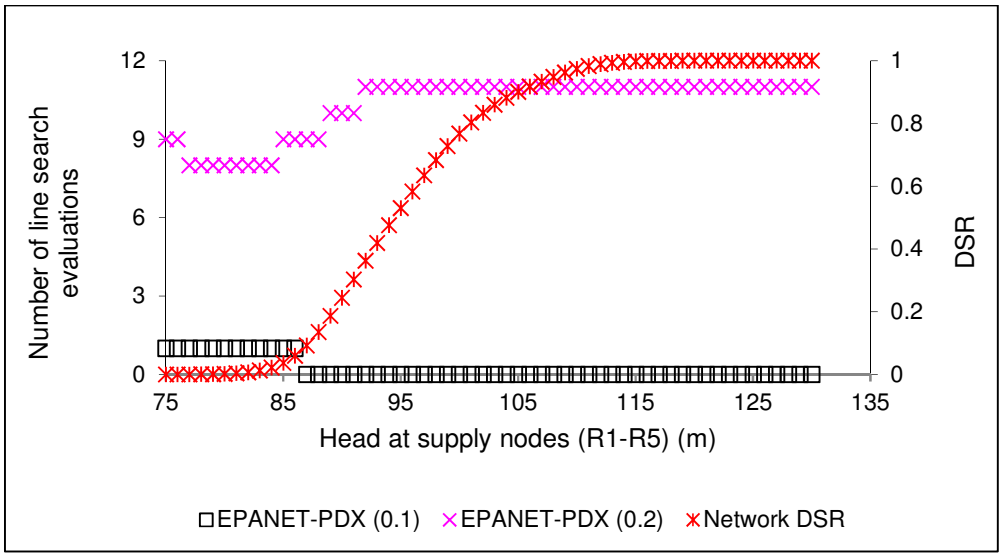


Figure 4.14 Number of line search evaluations needed as a function of the available pressure in the network for Example 3

4.4.3.2 Pipe Closure Conditions

The performance of EPANET-PDX (0.2) was assessed in terms of simulating major supply mains failures by closing simultaneously the supply pipes from three supply nodes out of five (R1-R5). A total of 10 such 'supply failures' resulting from multiple simultaneous supply mains failures were simulated. In these simulations, only two supply nodes out of five supplied the network and the nodal demands were fully satisfied in each case (i.e. all the network DSRs were 1.0). Table 4.3 summarises the average number of iterations and CPU time for the pipe closure simulations. The average numbers of iterations required per EPS were 6.77, 5.22 and 4.90 for EPANET-PDX (0.1), EPANET-PDX (0.2) and EPANET 2, respectively. The corresponding average CPU time for EPANET-PDX (0.1), EPANET-PDX (0.2) and EPANET 2 were 0.20 seconds, 0.21 seconds and 0.14 seconds, respectively. With network DSR of 1.0 for each pipe closure simulation, the CPU time for the EPANET-PDX (0.1) and EPANET-PDX (0.2) models were about the same while EPANET-PDX (0.2) provided smaller number of iterations in comparison to EPANET-PDX (0.1). Hence, overall the implementation of EPANET-PDX (0.2) would appear to be successful.

Fig. 4.15 shows the number of line search evaluations needed for the supply nodes closure simulations. On average EPANET-PDX (0.2) requires 10.431 evaluations per simulation, while EPANET-PDX (0.1) needs only 0.16 evaluations per simulation. The results are consistent with the previous result for the source head variation simulations, where the line search procedure of EPANET-PDX (0.1) is limited.

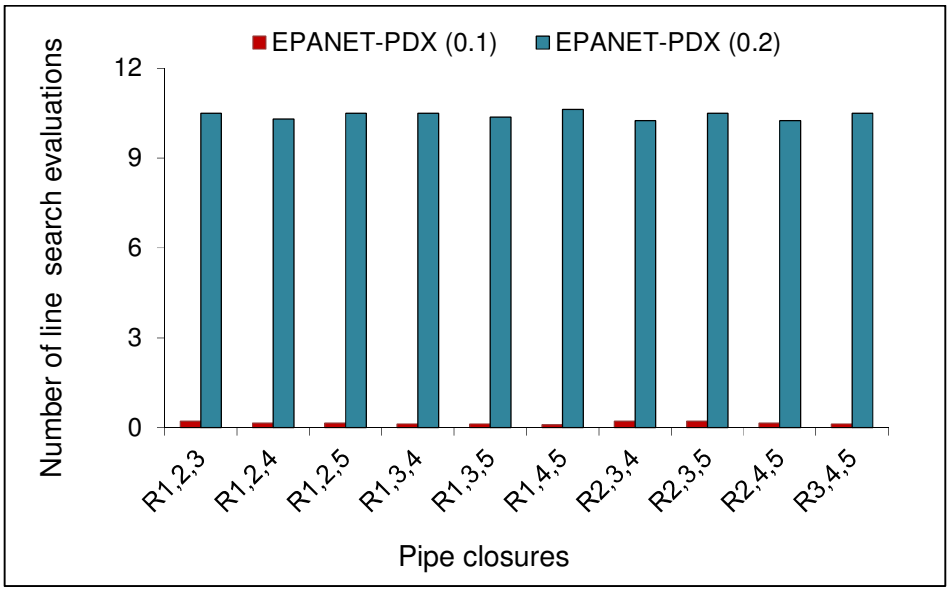
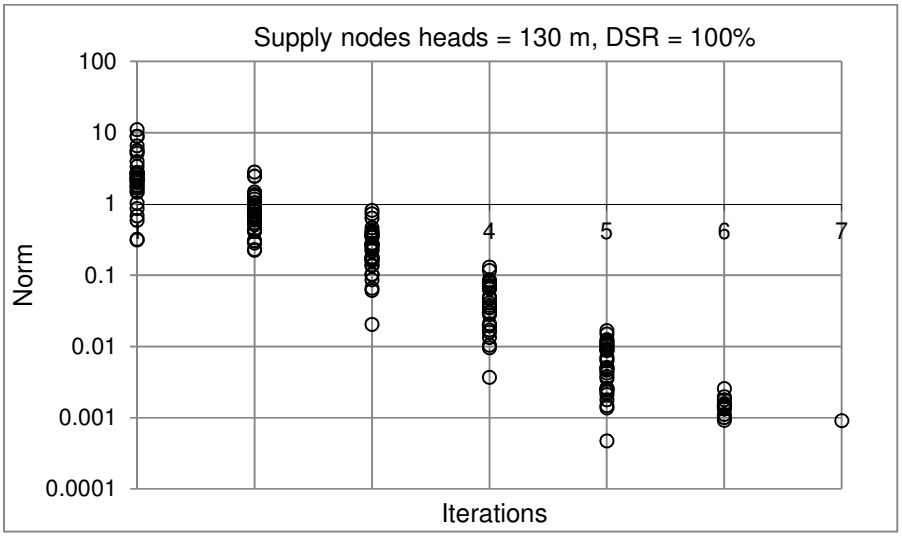


Figure 4.15 Number of line search evaluations for pipe closure conditions for Example 3

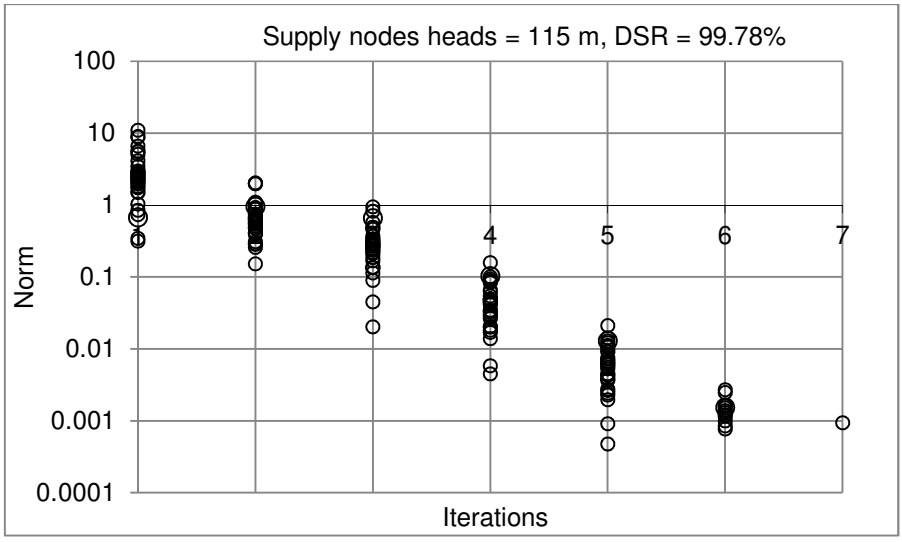
Table 4.3 Performance of simulators for pipe closure conditions for Example 3

Closed supply nodes	Average number of iterations			Average CPU time (seconds)		
	EPANET-PDX (0.1)	EPANET-PDX (0.2)	EPANET 2	EPANET-PDX (0.1)	EPANET-PDX (0.2)	EPANET 2
R1,2,3	6.969	5.250	4.875	0.197	0.202	0.137
R1,2,4	6.719	5.156	5.000	0.200	0.209	0.137
R1,2,5	6.781	5.250	4.875	0.193	0.207	0.140
R1,3,4	6.750	5.250	4.875	0.210	0.238	0.140
R1,3,5	6.563	5.188	4.844	0.197	0.202	0.147
R1,4,5	6.719	5.313	4.875	0.190	0.204	0.147
R2,3,4	6.781	5.125	4.844	0.200	0.205	0.147
R2,3,5	6.969	5.250	4.938	0.197	0.197	0.137
R2,4,5	6.656	5.125	4.969	0.200	0.202	0.140
R3,4,5	6.750	5.250	4.875	0.200	0.209	0.153

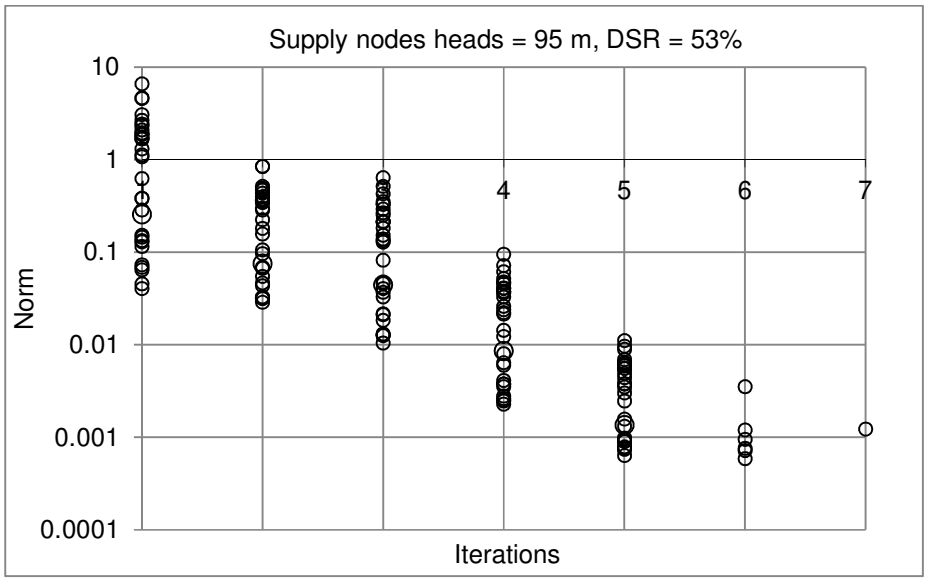
Fig. 4.16 shows the norm of the energy and mass balance at successive iterations for the arbitrarily selected four unbiased sample simulations (EPSs) of EPANET-PDX (0.2) based on Example 3. Each EPS comprises 31 steady state simulations. The two EPSs represent the endpoints of the network performance curve indicated in Fig. 4.7, while the remaining two denote intermediate points in the curve.



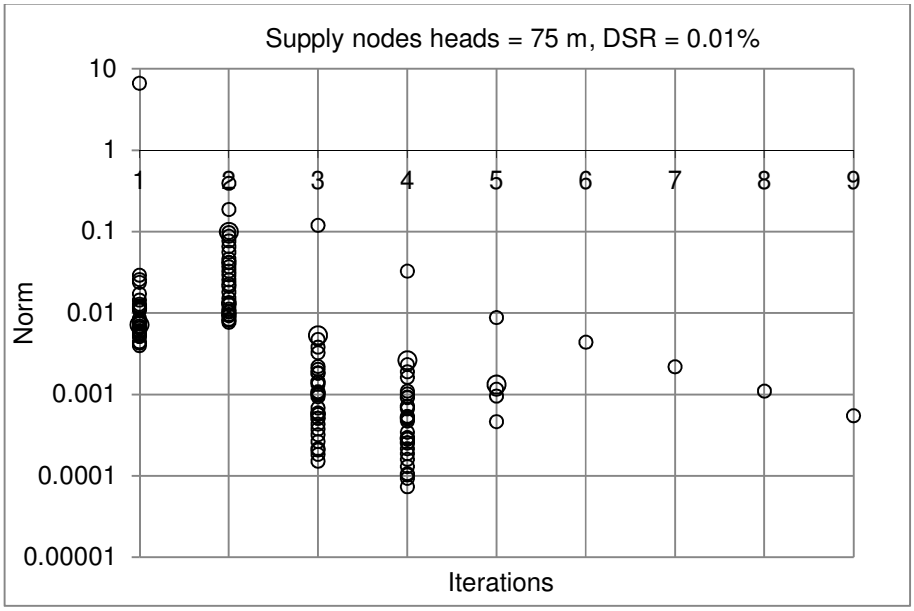
(a)



(b)



(c)



(d)

Figure 4.16 Norm at successive iterations for 31 steady state simulations of an extended period simulation for Example 3

The heads at the supply nodes for the extreme points were 75 m and 130 m with the respective network DSR of 0.01% and 100%. For the intermediate points, the heads at

the supply nodes were 95 m and 115 m and the corresponding network DSRs were 53% and 99.78%. Overall, the norm values decreases as the iteration progresses. At the 6^{th} iteration, the norm values for all simulations were below 0.01. It is worth mentioning that the norm values shown in Fig. 4.16 are the sum of the mass and energy balance and are based on the imperial unit, i.e. cfs and ft for mass and energy balance respectively. Thus, the corresponding values in SI units (m^3s^{-1} and m) would be much smaller. It was noted that for network DSR = 0.01% (Fig. 4.16d), the norm increased at the 2^{nd} iteration and continued to decrease at the consecutive iterations. The result is consistent with the algorithm procedure described in Section 4.3. In such circumstances, the algorithm improves convergence by updating nodal head using the available Newton step and continues solving the system of hydraulic equations iteratively until the convergence criteria of the algorithm are fulfilled.

4.5. Conclusions

A new line search and backtracking algorithm for integrating the logistic nodal head-flow function into the system of hydraulic equations in the global gradient algorithm is proposed to enable pressure dependent modelling. The proposed algorithm performed consistently well when simulating two benchmark networks and a real-life network that comprises multiple operating conditions. A total of 66 EPSs (31 hour duration) and 8 steady state simulations were carried out assuming various required nodal residual heads and considering source head variations and pipe closure conditions. The algorithm has no convergence issues and proves to be robust while simulating the full range of pressure operating conditions effectively.

Comparison between results generated by the model and EPANET-PDX demonstrated that the two models produce exactly the same hydraulic results for both normal and pressure deficient conditions. From a numerical perspective, a significant reduction in numbers of iterations to complete a simulation has been obtained for all pressure

operating conditions. Also, for extremely pressure deficient conditions an average of 26% reduction in computational time has been achieved in comparison to EPANET-PDX.

Overall, the obtained results indicated that the proposed line search and backtracking algorithm would appear to be successful. The results further highlight suggestion for future work on using the algorithm for WDSs optimization.

The next chapter presents water quality modelling of WDSs using PDA for various pressure operating conditions.

Chapter 5

Pressure-Dependent Network Water Quality Modelling

5.1 Introduction

Water distribution networks (WDNs) are designed and operated to provide water that is wholesome to consumers at an adequate pressure. However, a major challenge in the operation of WDNs today arises from pressure-deficient conditions caused by events such as pipe breaks, pump failures or large increases in demand (e.g. for firefighting purposes). These situations affect not only the hydraulic performance but also the quality of the water. Several studies have been conducted to evaluate the performance of WDNs under pressure-deficient conditions. These studies have focused on hydraulic analysis and addressed issues such as hydraulic reliability and design optimisation (Germanopoulos et al., 1986; Giustolisi et al., 2008; Tanyimboh and Kalungi, 2009; Siew and Tanyimboh, 2012b) without considering water quality.

Pressure deficiency plays a major role in the deterioration of water quality in WDNs. The problems caused by lack of pressure include low velocities, which result in long water travel and detention times that contribute to the loss of disinfection residual. This

may lead to bacterial regrowth (Clark and Haught, 2005) and, ultimately, water-borne diseases. Previous studies (Ghebremichael et al., 2008; Rodriguez et al., 2004) have also indicated that long detention times are a significant contributing factor in the formation of disinfection by products (DBPs). DBPs are formed when the disinfectant reacts with organic and inorganic substances in water. They can cause reproductive and developmental problems in humans and are thought to be carcinogenic (Nieuwenhuijsen, 2005; Richardson et al., 2002). Therefore, given the self-evident deleterious effects of subnormal pressure on water quality, urgent action is required to integrate the hitherto separate water quality and pressure-dependent hydraulic analysis models.

Hydraulic and water quality analysis of WDNs can be performed under time-varying conditions by employing extended period simulation (EPS) models. The models include important time-varying features such as water levels in tanks, nodal demands and the scheduling of pumps. Conventional EPS models are demand driven and thus assume that all demands are fully satisfied even if a network is in a pressure-deficient condition. Consequently, EPS models based on demand-driven analysis (DDA) cannot simulate the performance of a pressure-deficient network realistically. EPANET 2 and EPANET-MSX are DDA-based EPS models whose use is widespread throughout the world. EPANET 2 (Rossman, 2000) is public domain software that can model non-reactive tracer materials, chlorine decay, DBPs growth and water age. Also in the public domain, EPANET-MSX (Shang et al., 2008) is an extension of EPANET 2 that can simulate multiple chemical species concurrently.

In this chapter, the pressure-dependent analysis (PDA) model, EPANET-PDX (Siew and Tanyimboh, 2012a) is investigated and discussed with particular reference to water quality. In Chapter 4, an alternative implementation of the line search procedure for integrating the nodal head-flow function into the system of hydraulic equations is proposed to enhance the computational efficiency of EPANET-PDX. The EPS in EPANET-PDX consists of successive time intervals in which pressure-dependent steady-state analyses are performed that account for changes in nodal demands, water

levels in tanks, the operation of pumps and the status of valves (Siew and Tanyimboh, 2010b). EPANET-PDX is thought to have preserved the EPANET 2 modelling functionality in full. Thus, EPANET-PDX performs both hydraulic and water quality modelling under both normal and low-pressure conditions entirely seamlessly. The overriding objective here is to demonstrate this.

To assess EPANET-PDX, hydraulic and water quality analyses were conducted on two water supply zones of a network in the UK for a range of simulated operating conditions including normal and subnormal pressure and pipe closures. The performance of EPANET-PDX has also been assessed using a network in the literature. The properties considered are temporal and spatial variations in water age, chlorine and trihalomethane (THM) concentrations under various hydraulic conditions. EPANET 2 and EPANET-MSX results (for operating conditions with sufficient pressure) are also included for comparison and verification purposes. Given that EPANET 2 and EPANET-MSX are demand driven analysis (DDA) based models, they are unsuitable for pressure-dependent water quality investigations. They are, however, included in this study to verify that EPANET-PDX water quality results for operating conditions with sufficient pressure are accurate.

Pressure-dependent demand functions are used in pressure-dependent WDN models to achieve realistic estimates for operating conditions with insufficient pressure and these functions require calibration for each demand node. When there is insufficient pressure, the flow delivered will be less than the demand. It has been shown that if the PDA predictions are accurate, then the PDA predictions of the nodal flows if used as demands in a DDA model will result in identical pipe flow rates, nodal flows and nodal heads in the DDA model (Ackley et al., 2001).

The aim of this chapter is to demonstrate that, given identical hydraulic conditions of flow and pressure, EPANET-PDX yields essentially the same results for water quality as EPANET 2 and EPANET-MSX. Furthermore, by establishing that EPANET-PDX

provides accurate hydraulic analyses for PDA and DDA subject to the pressure-dependent demand function imposed, it is demonstrated that EPANET-PDX can carry out both normal and pressure-deficient water quality modelling. In other words, given a set of demands for which the actual WDN pressure is insufficient, PDA identifies the feasible set of demands for which the available pressure would be sufficient that is 'closest' to the specified demands. Accordingly, these reduced demands can be modelled entirely satisfactorily using DDA (as stated above). It is thus sufficient here to show that EPANET-PDX yields accurate PDA predictions and that its water quality results are essentially the same as EPANET 2 and EPANET-MSX under normal operating conditions. A secondary aim is to demonstrate that, during pressure-deficient operating conditions, spatial and temporal variations in water quality can be very different from those in operating conditions with fully satisfactory pressure. The third aim of the chapter is to illustrate the water quality modelling difficulties when low pipe flow velocities prevail due to excessive pressure reduction. As previous research (Tzatchkov et al., 2002) indicates, advection-driven models such as EPANET 2 may yield inconsistent water quality results if dispersion is significant due to low flow velocities.

5.2 Water Quality Modelling

A literature review on water quality models and the computational solution procedures has been presented in Chapter 2. This chapter provides an overview of the EPANET 2 water quality model. The water quality model in EPANET 2 is dynamic and includes relevant fate processes such as transport (mainly advection) and various transformation processes (e.g. chlorine decay). In addition to pipe flow, some of the processes considered are mixing at pipe junctions, mixing in storage facilities, pipe wall reactions and a range of bulk flow reactions. Several modelling approaches for the chemical reactions are available, such as first- and second order reaction models. The first-order reaction models are summarised briefly here. For chlorine decay

$$\frac{\partial C_{cl}}{\partial t} = -k_b C_{cl} \tag{5.1}$$

in which C_{cl} is chlorine concentration, k_b is a reaction rate constant in bulk flow and $t$is time. For THM production, the limited first-order model used is

$$\frac{\partial C_{thm}}{\partial t} = k_b (C_L - C_{thm}) \tag{5.2}$$

where C_{thm} *is* THM concentration and C_L is the ultimate THM concentration. For dissolved substances in water that react with materials at the pipe wall such as biofilm, the first-order model used is

$$\frac{\partial C}{\partial t} = \frac{2k_w k_f C}{r_h (k_w + k_f)} \tag{5.3}$$

where k_w is the wall reaction rate constant, k_f is the mass transfer coefficient that represents the rate of dissolved substances transported between the bulk flow and the pipe wall, r_h is the hydraulic radius of the pipe and C is the reactant concentration in bulk flow.

The system of equations for conservation of mass that constitutes the water quality model can be set up by superposition of the relevant transport and transformation equations. For a pipe, for example, the equation that results is

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} + r(C) \tag{5.4}$$

in which u = mean flow velocity, x is distance along the pipe and $r(C)$ is rate of reaction (e.g. Eq. (5.1)-(5.3)).

5.3 Case Studies

Three network examples have been used to investigate the spatial and temporal distribution of water quality in WDNs under various pressure operating conditions. The examples are based on a simple network in literature (Network 1; see Fig. 5.1) and two large real-life networks. The two real-life networks that are named here as Networks 2 and 3 (Fig. 5.6 and 5.13 respectively) are water supply zones of a network in the UK. The water quality analyse for these networks were conducted based on demands the water utility provided for the purposes of hydraulic design optimisation. Network 2 has previously been described in Chapter 4 to demonstrate PDA modelling. A brief description on the network has also been provided in this chapter for completeness.

Both Networks 2 and 3 obtain water entirely from neighbouring water supply zones through supply nodes (i.e. nodes R1–R5 in Fig. 5.6 and nodes R1–R4 in Fig. 5.13). The network and dynamic operational data for Networks 2 and 3 used for the hydraulic analyses were taken from calibrated EPANET models and a geographical information system (GIS) database. The EPANET models contain node data that include elevations and demands for individual nodes and the relevant demand categories. The demand categories in Network 2 that comprise domestic demand, 10-hour 'commercial' demand and unaccounted for water are shown in Fig. 5.6(b)-5.6(d)**.** Network 2 also has 29 different fire demands of 1 hour each. These are applied at the 29 fire hydrants located at different positions in Network 2. Network 3 has domestic demand, 10-hour and 16-hour 'commercial' demands and unaccounted for water as shown in Fig. 5.13(b)-5.13(e). For all these demand types, demand multipliers are available every 15 minutes for 24 hours total duration.

Also in the EPANET models are link data, including pipe lengths, diameters and roughness values; the Darcy–Weisbach pipe friction head loss formula was used for the hydraulic analyses. The information in the GIS database includes data on the pipes such as age, material, diameter, length, renewal and/or rehabilitation year, burst rate and other data of a geospatial nature. The calibrated hydraulic models notwithstanding, due to difficulties in obtaining up-to-date operational water quality data, some typical or limiting values from the literature were assumed, as explained later. More fundamentally, this initial assessment, under artificially controlled operating conditions, is in fact a prerequisite to the fieldwork and subsequent parameter calibration that will be required subsequently to address even more realistic and network-specific conditions that are inherently more complex and for which exact solutions will not be available. Collectively, the subset of assumed modelling (i.e. reaction rate constants) and operational (i.e. chlorine concentration) values, when taken together with mandatory values stipulated in a selection of leading international standards for drinking water, are intended to represent the 'most favourable' scenario.

Accordingly, values of k_b = 0.5/day and k_w= 0.1 m/day (Carrico and Singer, 2009; Helbling and Van Briesen, 2009) were used. To achieve a detectable chlorine residual of 0.2 mg/l (WHO, 2011) at remote points in the system, the chlorine concentration at each supply node was assumed constant at 1 mg/l. Moreover, World Health Organization guidelines on drinking water quality (WHO, 2011) recommend a minimum residual chlorine concentration of 0.2 mg/l at the point of delivery; no minimum concentration value is stipulated in UK and EU drinking water standards. Also, the UK and EU standards do not specify a maximum concentration for chlorine; however, the values given in the WHO guidelines (WHO, 2011) and the US Safe Drinking Water Act, 1996 (USEPA, 1996) are 5 mg/l and 4 mg/l respectively (see e.g. Twort et al., 2000). It is recognised here that the taste and/or odour threshold is much lower; it may be noted that the typical concentration in most disinfected drinking water is 0.2–1.0 mg/l (WHO, 2011). A maximum total THM concentration of C_L = 100 µg/l was adopted in Eq. 5.2

based on EU and UK drinking water standards (EC, 1998; HMG, 2001, 2010); with a maximum concentration of 80 µg/l, the US EPA regulations are more stringent. Indeed, the EU and US standards advise that, where possible, a lower value should be aimed for without compromising disinfection. One of the drawbacks of the EPANET 2 THM model is that it requires modellers to specify the limiting THM concentration in advance. Sohn et al. (2004) suggested an alternative model for THM that avoids the need to pre-specify a limiting concentration (see e.g. Seyoum and Tanyimboh, 2013). The same water quality operational data and reaction rate constants as Networks 2 and 3 were used (assumed) for Network 1 as well. For all the three networks, 5-minute water quality time step and 0.01 mg/l concentration tolerance were used for water quality analysis. Also, the initial THM concentration at all nodes were assumed zero.

For modelling the pressure deficient networks, the head below which the nodal flow is zero was taken as the nodal elevation while the head above which the demand is satisfied in full was taken as the elevation plus the minimum required residual head.

5.3.1 Network 1

The first example (Fig. 5.1) is a simple two-loop network taken from Fujiwara and Ganesharajah (1993) for demonstration purposes. The network consists of a single source, eight pipes of length 1000 m with the Hazen-Williams roughness coefficient of 140 and six demand nodes with the desired pressure heads of 60 m. The network was considered to be at steady state condition where the nodal demands were not varied with time. A total duration of 24 hours was specified to run water quality analysis. Details of nodes and pipes are listed in Table 5.1.

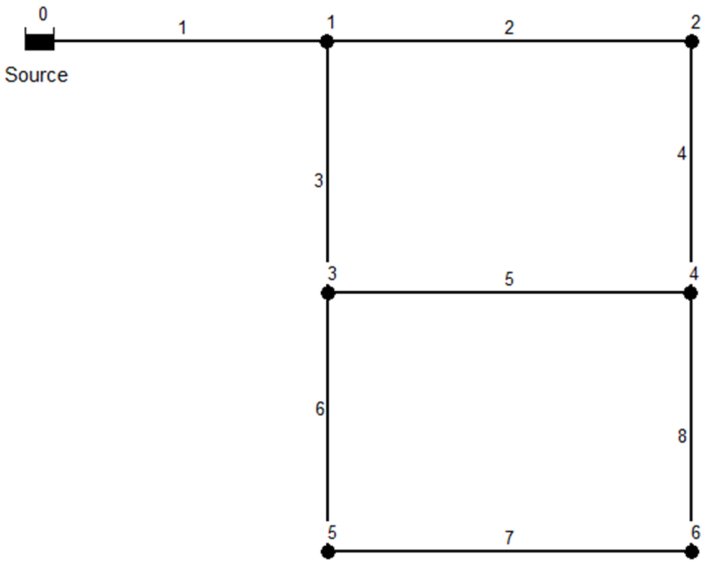


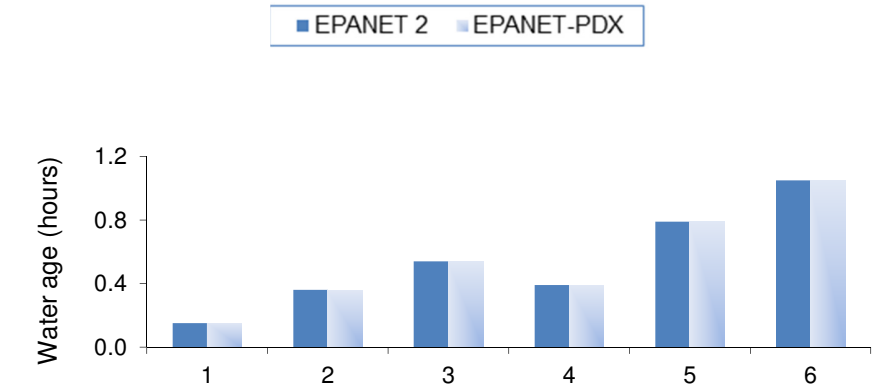
Figure 5.1 Layout of Network 1

Table 5.1 Node and pipe data for Network 1

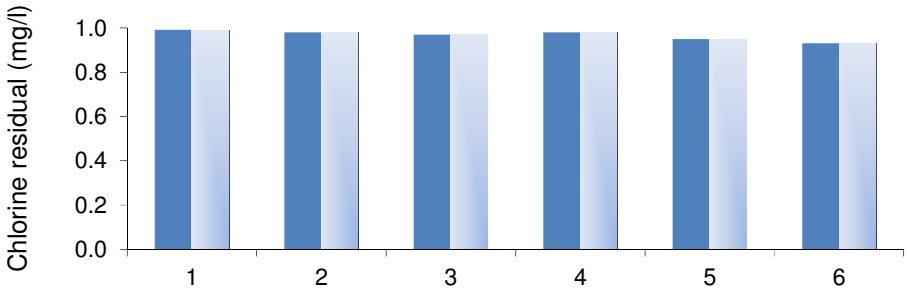
Node	Elevation (m)	Demand (m^3/s)	Pipe	Diameter (mm)
1	50	0.0471	1	500
2	50	0.0471	2	400
3	45	0.0778	3	400
4	45	0.0471	4	400
5	55	0.0556	5	250
6	55	0.0889	6	250
			7	250
			8	250

The source head was fixed at 90 m for normal pressure condition with full demand satisfaction. Various source heads (60 m, 65 m and 70 m) were considered for pressure deficient conditions. All simulations were carried out on an Intel Xeon workstation (2 processors of CPU 2.4 GHz and RAM of 16 GB). Fig. 5.2 shows that both EPANET-

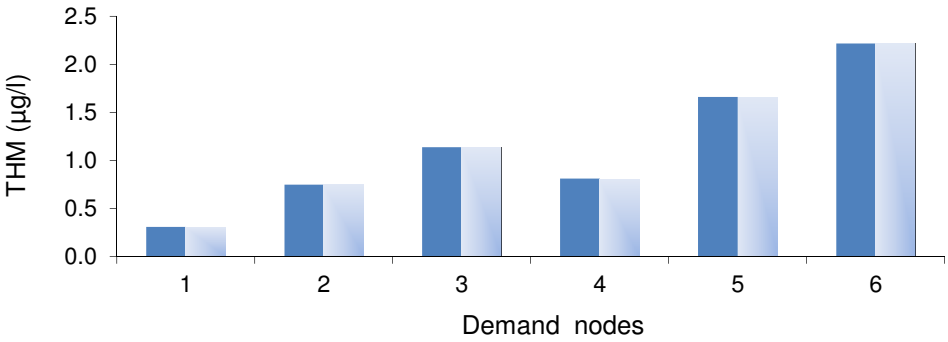
PDX and EPANET 2 provide essentially identical water quality results when the pressure in the system is sufficient.



(a) Water age



(b) Chlorine residual

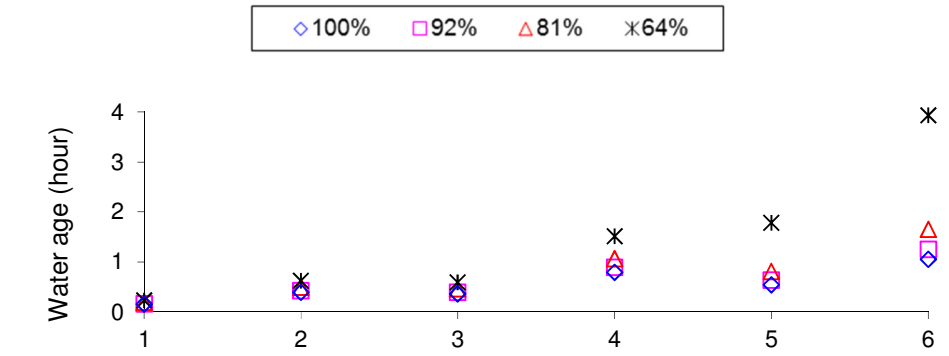


(c) THM

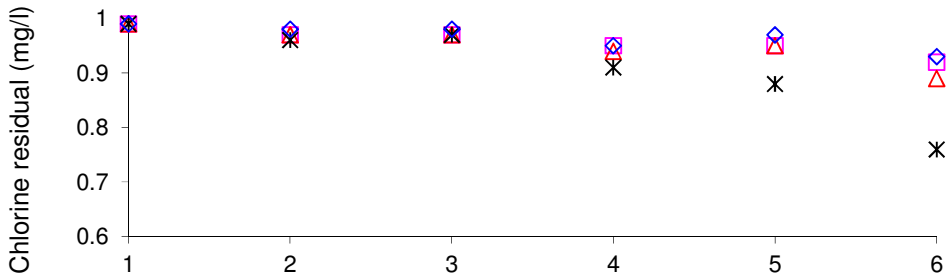
Figure 5.2 Comparison of water quality results for normal pressure

For pressure deficient network simulations, the source water level was reduced from 90 m to 70 m, 65 m and 60 m for which the network satisfied only 92%, 81% and 64% of the total demand respectively. Rather unsurprisingly, EPANET 2 provided identical water quality results to normal pressure conditions in which there is enough pressure to satisfy all demands. In practical terms, a pressure-deficient network cannot satisfy demands in full. In this regard, EPANET 2 results are unrealistic. This limitation is attributable to the underlying DDA modelling approach. By contrast, EPANET-PDX that has PDA functionality provided different water quality results that reflected the actual pressure in the network (Fig. 5.3). It was observed that when the pressure in the system decreases, the THM concentration and water age increase while the chlorine residual decreases. This evidently reveals the fact that when the pressure in the network is low, the flow will correspondingly be low (Fig. 5.4) and the hydraulic residence time (water age) will be greater. An increase in residence time will enable the THM concentration to increase and the chlorine concentration to decrease. It is worth noting in Fig. 5.3 that the differences in the model results for the normal and pressure deficient conditions are significant at the distant nodes (e.g. Nodes 5 and 6) from the source during high-pressure deficiency with low demand satisfaction condition. In general, nodes that are far from the sources have long residence time in the network and thus require accurate model prediction.

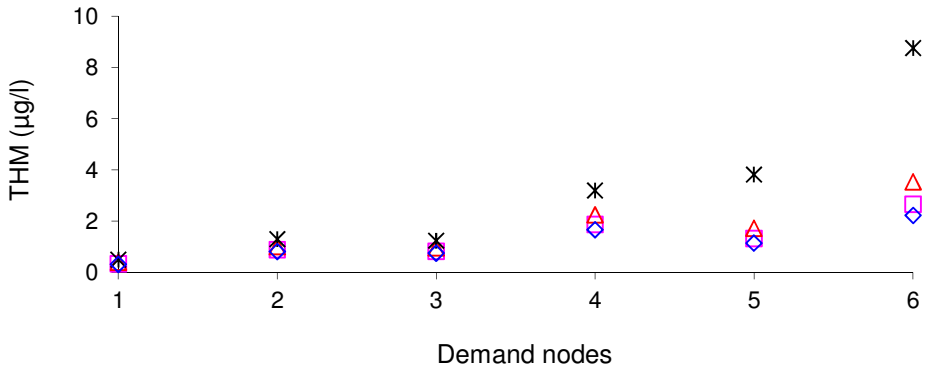
A confirmation test was carried out on the PDA results as described in Section 5.1 (Ackley et al., 2001; Tanyimboh and Templeman, 2010; Siew and Tanyimboh, 2012a) based on 41 source heads from 40 m to 80 m in equal steps of 1 m. Nodal heads obtained from EPANET 2 were compared with those of EPANET-PDX. A correlation between the heads gives R^2 value of 0.9999995 (i.e. $1\text{-}R^2 = 5\text{x}10^{-7}$) as it is shown in Fig. 5.5. This clearly confirms that EPANET-PDX results are accurate.



(a) Water age



(b) Chlorine residual



(c) THM

Figure 5.3 EPANET-PDX water quality results for normal and subnormal pressure. The percentages refer to the percentages of total demand satisfied under conditions of normal and low pressure.

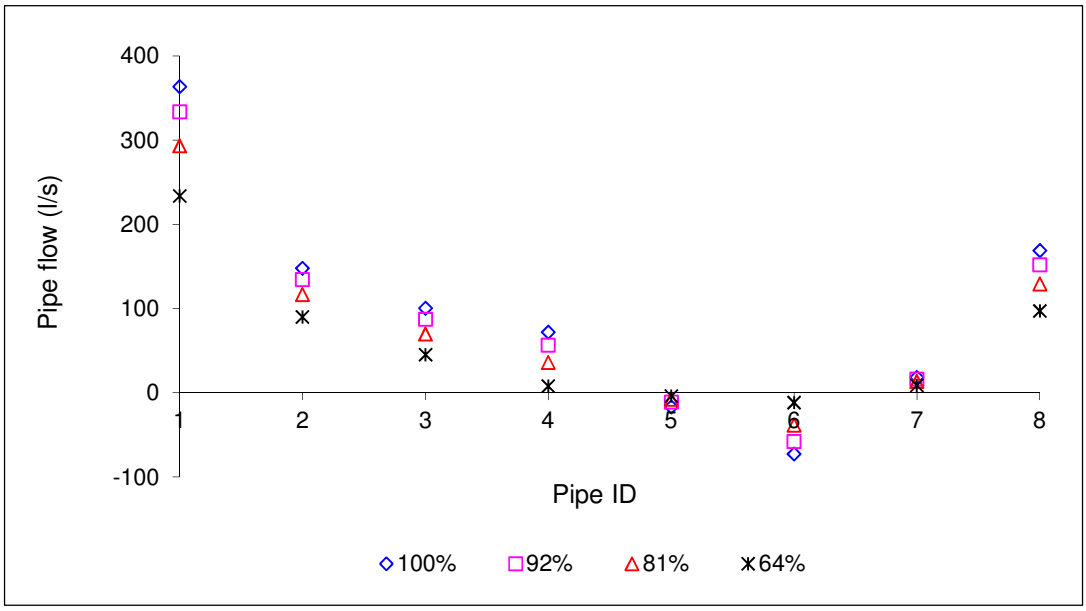


Figure 5.4 EPANET-PDX pipe flow results for pressure deficient conditions. The percentages refer to the percentages of total demand satisfied under conditions of normal and low pressure.

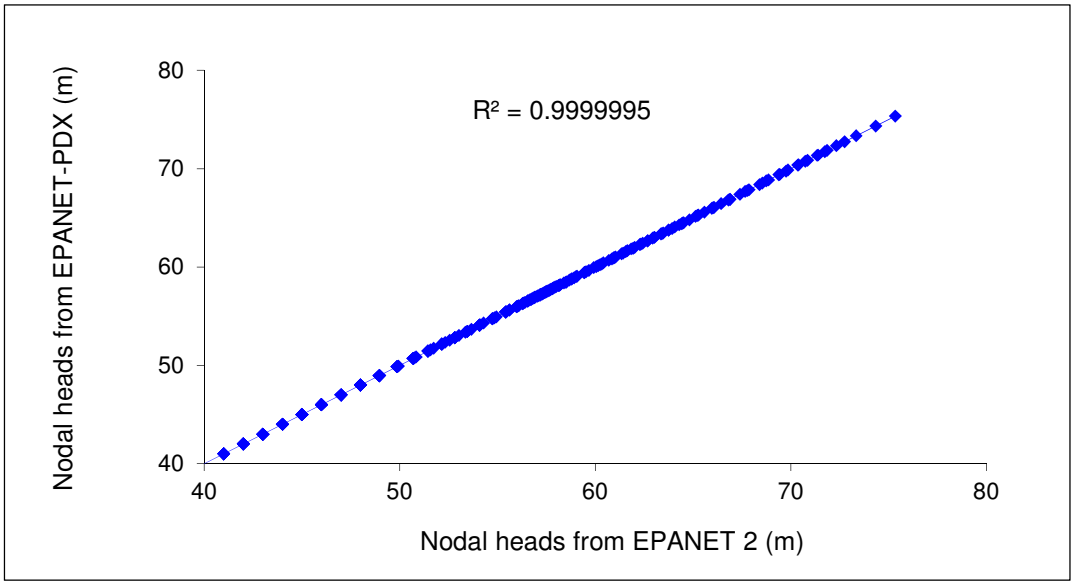


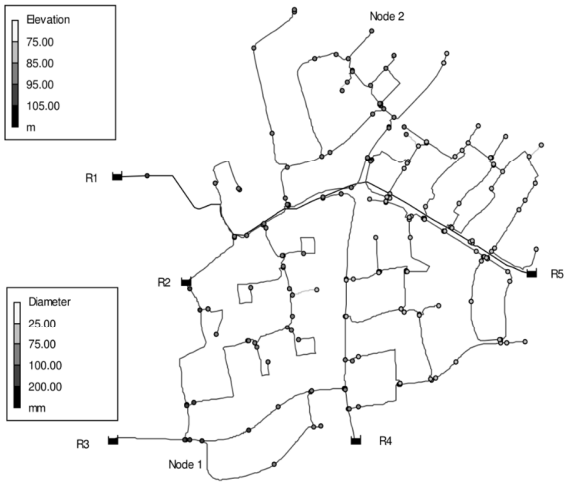
Figure 5.5 Hydraulic feasibility check for source heads 40 m-80 m

5.3.2 Network 2

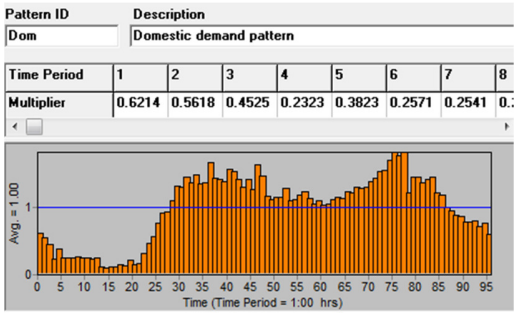
Extended period simulations (EPSs) were conducted on Network 2 (Fig. 5.6) on an Intel Core 2 Duo personal computer (CPU = 3.2 GHz, RAM= 3.21 GB). The network consists of 251 pipes of various lengths, 228 demand nodes (including fire hydrants), five variable-head supply nodes (R1 to R5) and 29 fire hydrants at various locations; pipe diameters are 32-400 mm. Each variable-head supply node has an average level of 155 m along with head level multipliers at 1–hour intervals for duration of 96 hours as presented in Appendix B (Fig. B-1.1). The range of variation in the head levels at each supply node is insignificant. Hence, the supply nodes were modelled here as constant-head nodes with water levels of 155 m each. The residual head for full demand satisfaction is 20 m. To allow for the observed inconsistencies in the water quality results at the start of the simulations, the chosen EPS duration was 93 hours. The water quality results become stable after sufficient time has elapsed. The results reported here are for the last 30 hours.

Both normal and low-pressure conditions were considered. Pressure-deficient conditions were created artificially by setting the water levels at the supply nodes to satisfy, in turn, only 90%, 75%, 50% and 30% of the total demand. A demand satisfaction ratio (DSR) of 30% is included to investigate water quality modelling in EPANET 2 under low-flow conditions with significant dispersion likely. The effects of closing individual pipes were also investigated. Results are highlighted for two typical demand nodes that represent the nodes closest to the supply nodes and the remote points in the networks. Node 1 is close to Supply node R3 while Node 2 represents a remote point in the network (Fig. 5.6). Results for the other demand nodes are not shown explicitly for brevity. Additionally, variations in water age in the whole of the network are included for two different operating conditions – normal pressure (DSR = 100%) and a pressure deficient condition (DSR = 75%). For all the EPSs 1-hour hydraulic time step was used.

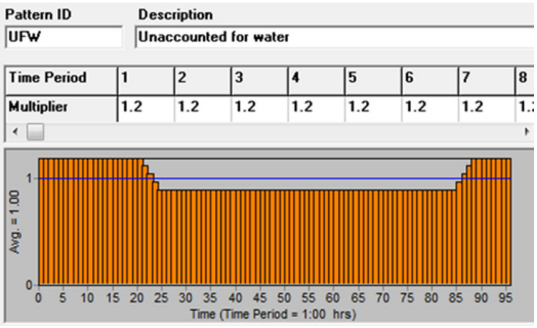
Considering various operating scenarios, 1,838 EPSs of 93 hours were carried out (see Table 5.2).



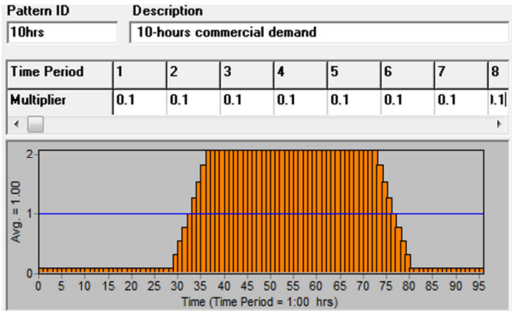
(a) Layout with diameters and elevations



(b) Domestic demand factors



(c) Unaccounted for water factors



(d) 10-hour commercial demand factors

Figure 5.6 Network 2: layout and demand factors

Table 5.2 Number of EPSs for Network 2

	EPANET 2	EPANET-PDX	EPANET-MSX
Normal pressure			
Water age	1	1	1
Chlorine	1	1	1
THM	1	1	1
Water age, chlorine and THM concurrently	NA[a]	NA	1
Pressure-deficient conditions[b]			
Water age	4	4	NA
Chlorine	4	4	NA
THM	4	4	NA
Water age, chlorine and THM concurrently	NA	NA	4
Pipe closures			
Water age	251	251	NA
Chlorine	251	251	NA
THM	251	251	NA
Supply-node head variations[c] (Fig. 5.12(a))	81	81	NA
PDA confirmation tests	66	66	NA
Total	915	915	8

[a] Not applicable

[b] One each for DSR= 90%, 75%, 50% and 30%

[c] Identical demands and supply-node heads used for both EPANET 2 and EPANET-PDX

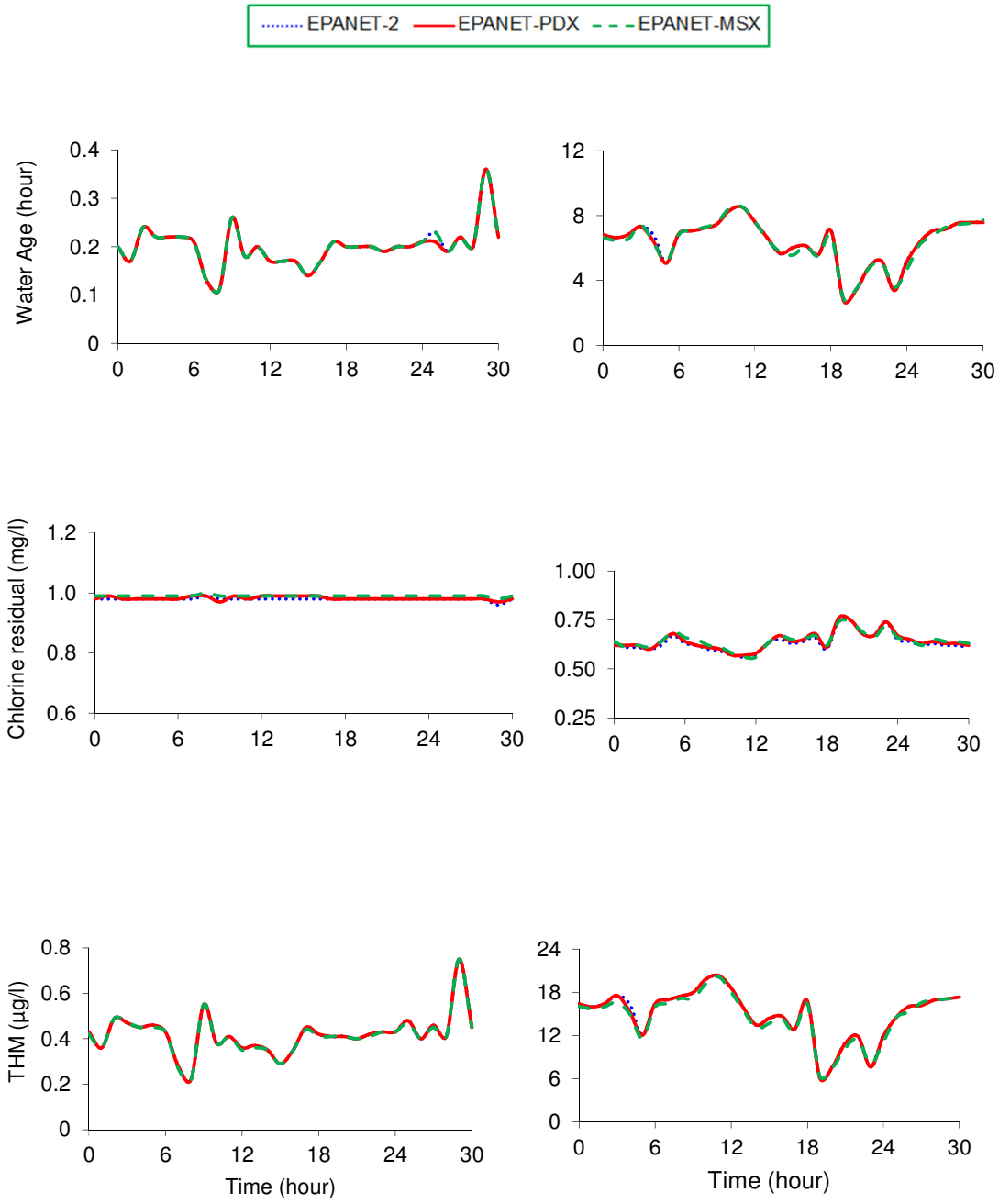
5.3.2.1 Normal Pressure

For the initial verification purposes, water age, chlorine residual and THM were simulated using the EPANET 2, EPANET-MSX and EPANET-PDX under normal pressure (i.e. DSR = 100%), with the supply node heads fixed at 155 m. The three models provide essentially identical results, as shown in Fig. 5.7, and the results are consistent with the demand fluctuations. An increase in the residence time due to a reduction in the demand is accompanied by a reduction in the concentration of chlorine and an increase in the water age and concentration of THMs. The chlorine residual at Node 1 is effectively constant due to its proximity to a supply node, where the

concentration is kept constant. Taking water age as an example, the correlation coefficients between EPANET-PDX and EPANET 2 were $R^2 = 1$ for Node 1 and $R^2 = 1$ for Node 2. Similarly, the correlation between EPANET-PDX and EPANET-MSX was $R^2 = 1$ for Node 1 and $R^2 = 0.99$ for Node 2 (See Appendix B (Fig. B-1.2 and B-1.3)). EPANET-MSX provides the water age, chlorine residual and THM concentration in a single simulation, while EPANET 2 and EPANET-PDX can simulate only one species at a time.

The nodal heads and pipe flows from EPANET-PDX and EPANET 2 for the entire 93-hour EPS were compared. The data points plotted were 21,204 for the nodal heads (228 demand nodes × 93 hydraulic time steps) and 23,343 for the pipe flow rates (251 pipes × 93 hydraulic time steps). The correlation coefficients were $R^2 = 1$ for the nodal heads and $R^2 = 1$ for the pipe flow rates (See Appendix B (Fig. B-1.4)). This further strengthens previous evidence (Siew and Tanyimboh, 2012a) that the results from the two models are essentially identical when there is sufficient pressure in the network.

To complete the 93-hour EPS, EPANET 2 and EPANET-PDX respectively required an average of 0.7 seconds and 1.0 seconds for chlorine, 1.0 seconds and 1.4 seconds for THM and 0.6 seconds and 1.0 seconds for water age. Corresponding values for EPANET-MSX were 10.5 seconds for chlorine, 2 seconds for THM and 2 seconds for water age. The longer simulation time for chlorine is due to the simulation time associated with the wall reaction component of the chlorine decay. To simulate chlorine, THM and water age concurrently, EPANET-MSX required an average of 13 seconds (Table 5.3).



(a) Node 1 (b) Node 2

Figure 5.7 Water quality variations in Network 2

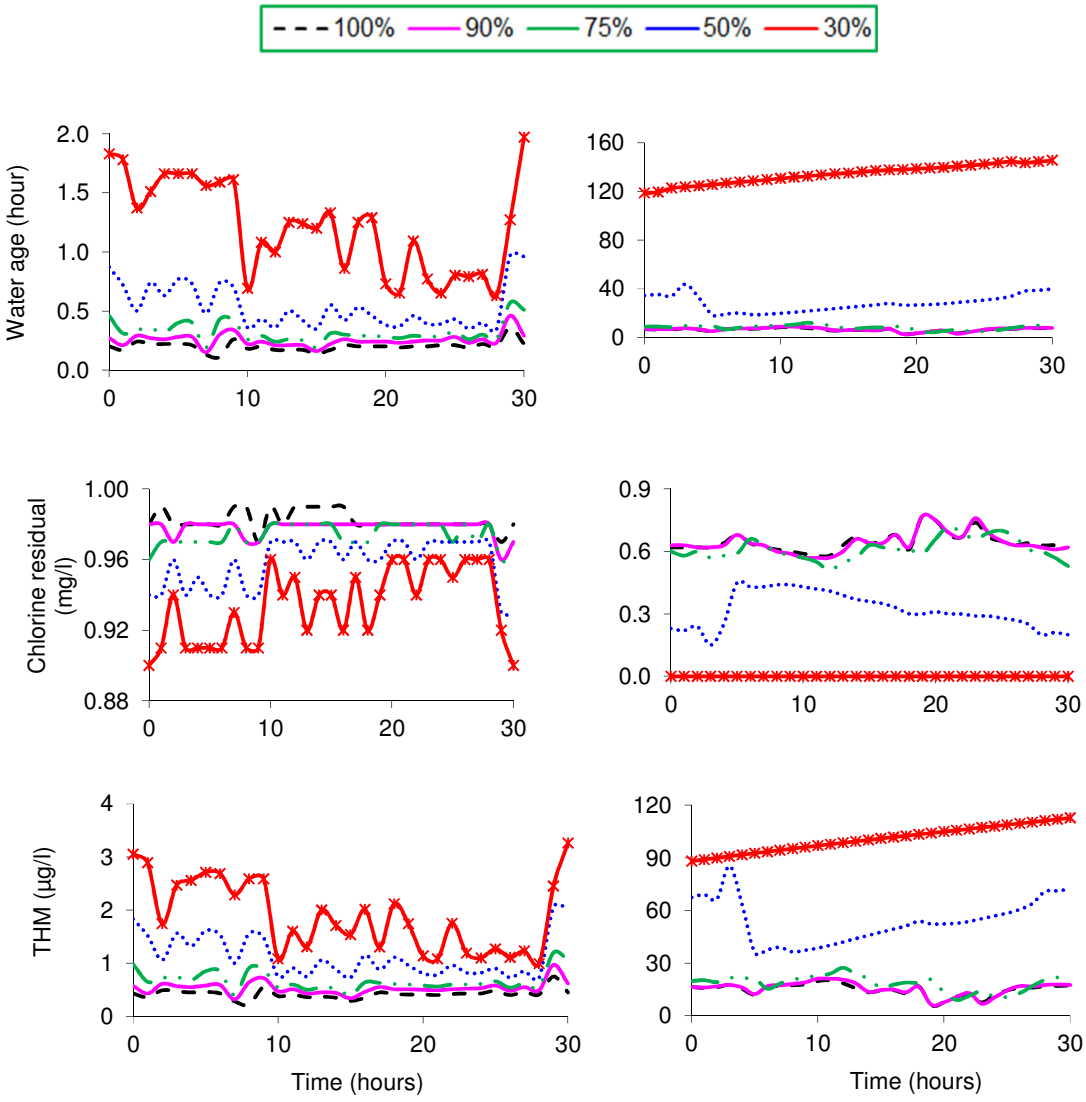
Table 5.3 CPU times for Network 2

	Mean CPU time per EPS (seconds)		
	EPANET 2	EPANET-PDX	EPANET-MSX
Chlorine	0.7	1.0	10.5
THM	1.0	1.4	2.0
Water age	0.6	1.0	2.0
Chlorine, THM and Water age concurrently	NA[a]	NA	13.0

[a]Not applicable

5.3.2.2 Pressure-Deficient Conditions

Simulations were carried out for several pressure deficient conditions using EPANET-PDX. The constant heads at the supply nodes were reduced from 155 m to 105 m, 100 m, 95 m and 90 m to achieve network DSRs of 90%, 75%, 50% and 30% respectively. Fig. 5.8 shows EPANET-PDX provides different values of water age, chlorine residual and THM concentrations for the different low-pressure conditions. Similar explanations to Network 1 can be made that the greater the pressure deficiency the greater the water age and THM concentration and the smaller the chlorine residual concentration. For Node 1, it can be seen that there are large fluctuations in the water age and chlorine and THM concentrations for DSR values of 50% and 30% for which the flow velocities are low in general. These fluctuations are consistent with previous results of the advection-driven EPANET 2 model, under conditions in which velocities are low (Tzatchkov et al., 2002). Tzatchkov et al. (2002) demonstrated that an advection-dispersion model reduces the fluctuations and consequently provided more realistic EPS results.

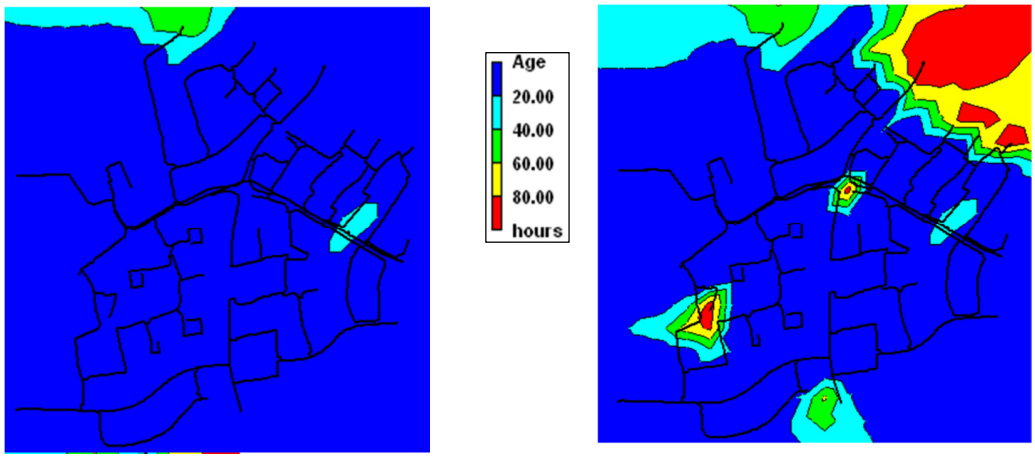


(a) Node 1 (b) Node 2

Figure 5.8 Water quality in Network 2 under normal and pressure deficient conditions (EPANET-PDX). The percentages refer to the percentages of total demand satisfied under conditions of normal and low pressure.

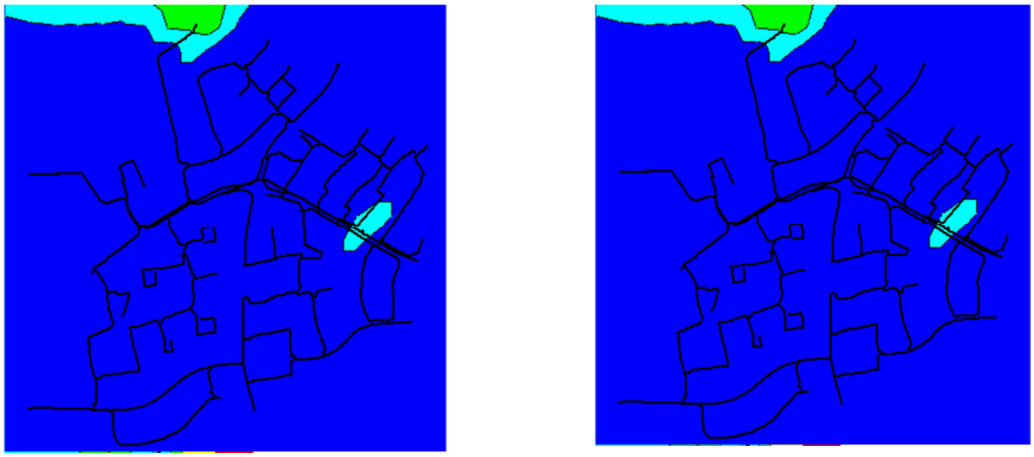
Fig 5.8 also shows that, from a water quality perspective, the effects of low pressure are more significant at remote points (Node 2) than in an area close to a supply node (Node 1). This is a direct consequence of the spatial distribution of the water age. When DSR = 30%, the majority of the nodes in Network 1, including Node 2, have very little or zero

nodal flow. The results for Node 2 when DSR = 30% would appear to reveal an anomaly. Zero- and low-flow nodes require special care and may indicate the need for more improvements in the underlying (EPANET 2) water quality model in the context of PDA. EPANET 2 and EPANET-MSX, however, provided the same results as the normal pressure condition shown previously in Fig 5.7 because they lack PDA functionality. This is illustrated in Fig 5.9, which shows EPANET 2 and EPANET-PDX predictions of water age throughout the network at 93 hour for DSR values of 75% and 100%.



(a) EPANET 2 (DSR = 75%)

(b) EPANET-PDX (DSR = 75%)



(c) EPANET 2 (DSR = 100%)

(d) EPANET-PDX (DSR = 100%)

Figure 5.9 Water age at 93 hour in Network 2. The EPANET 2 results for DSRs of 75% and 100% are the same.

The effects of closing pipes one at a time were also investigated, with the heads at the supply nodes maintained at the normal level of 155 m. Fig 5.10 shows EPANET-PDX estimates of the shortfall in the flow delivered (i.e. 1 – *DSR*). Whereas Fig 5.10 represents the entire network, the effects of low pressure can be greater in areas near closed pipes.

Fig. 5.11 depicts the water quality for EPANET 2 and EPANET-PDX for the individual pipe closures. The results are comparable, apart for some slight differences in the chlorine residuals. Fig 5.10 shows that overall the pressure in the network is mostly satisfactory, with the DSR close or equal to 1. This explains the similarity between EPANET 2 and EPANET-PDX in Fig. 5.11. In practice, it may not be possible to isolate individual pipes. Where multiple pipes must be taken out of service, it can be expected that any discrepancies due to the DDA modelling errors (i.e. in EPANET 2 for example) would be greater.

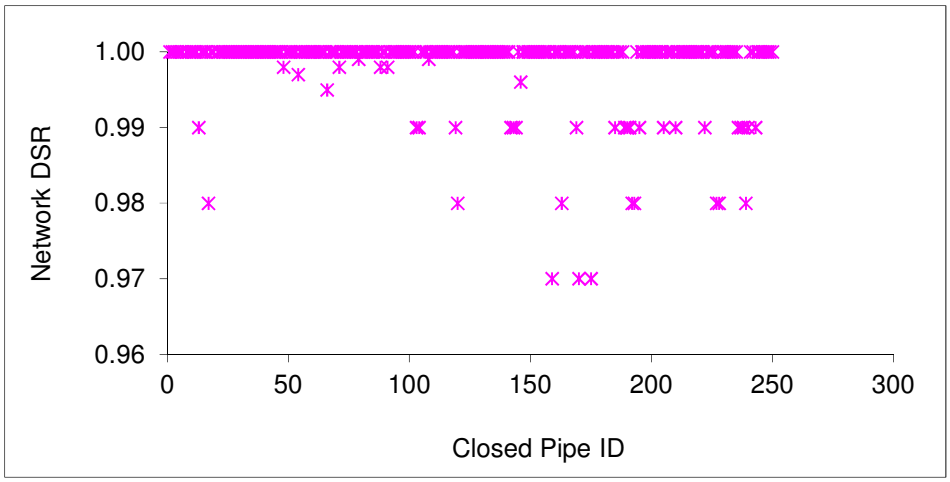
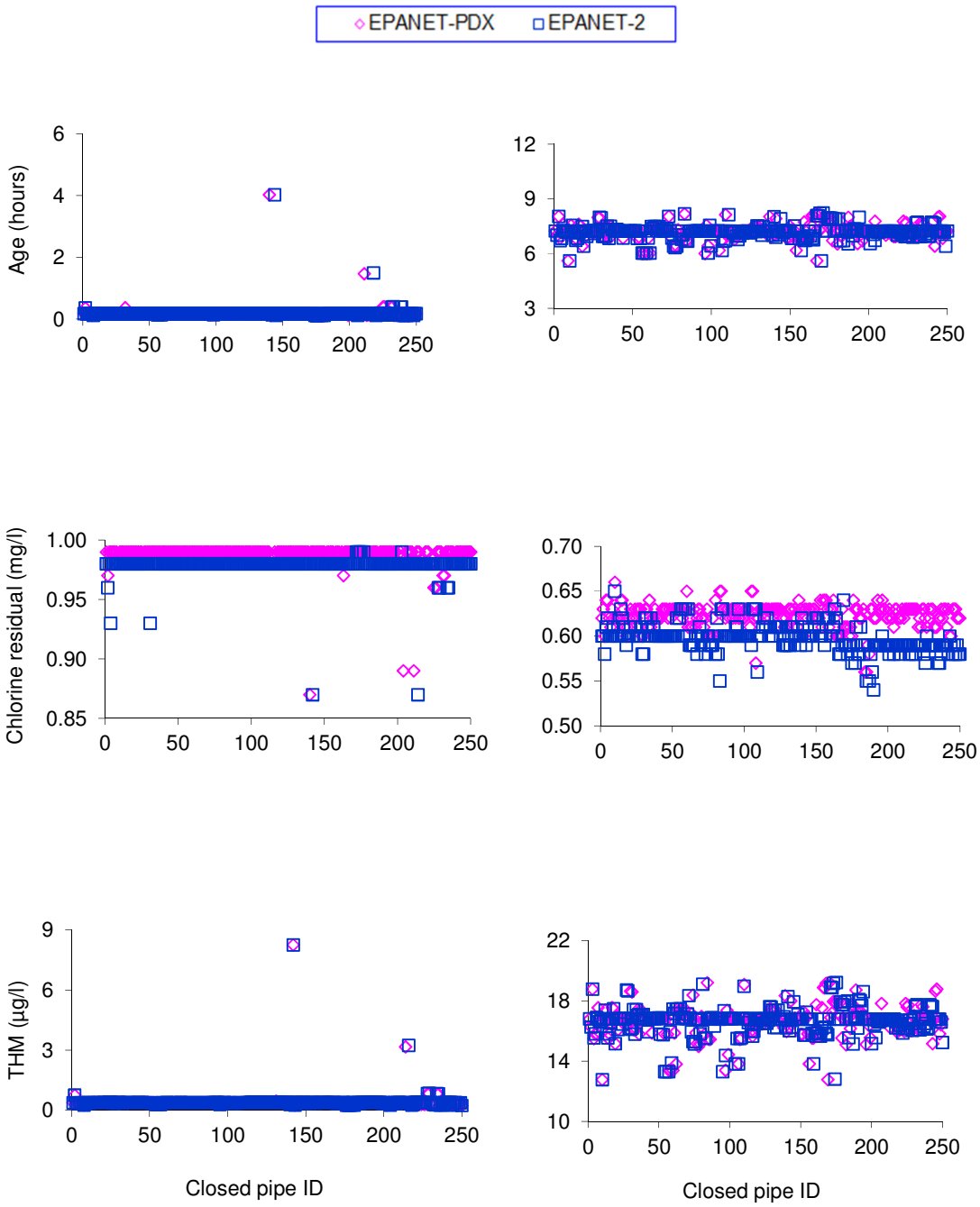


Figure 5.10 EPANET-PDX: pipe closure effects on the flow supplied in Network 2

The computational efficiency of EPANET-PDX was assessed also with reference to EPANET 2 for the above-mentioned water quality simulations. Fig 5.12 (a) shows the results for a range of constant-head levels at the supply nodes. The average CPU time

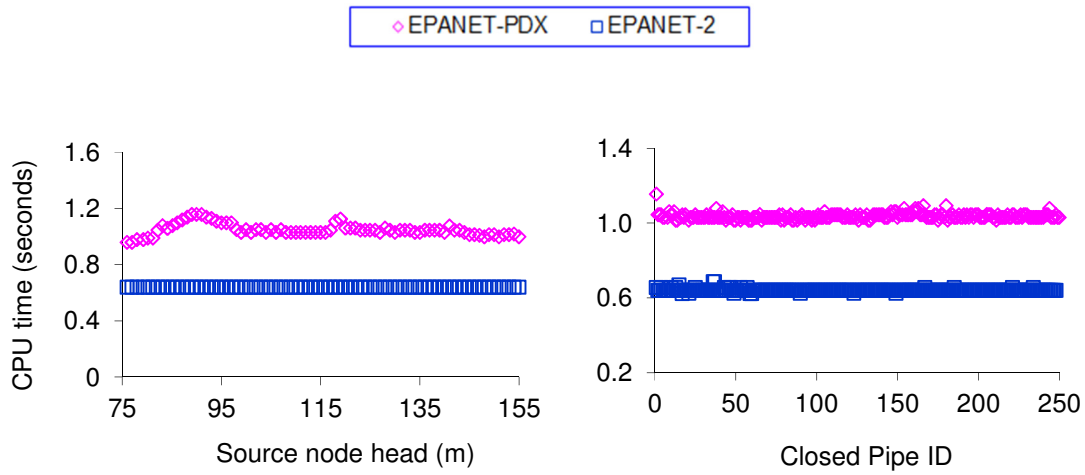
was 1.048 seconds per 93-hour EPS simulation for EPANET-PDX and 0.640 seconds for EPANET 2 based on water age. Fig. 5.12 (b) shows the results for pipe closures. The average CPU time per 93-hour EPS simulation for water age was 1.037 seconds for EPANET-PDX and 0.641 seconds for EPANET 2. The results here together with the CPU times for the normal operating conditions that are based on a real system are an indication that EPANET-PDX is somewhat slower than EPANET 2. It is worth reiterating, however, that EPANET-PDX provides realistic results for pressure deficient conditions whereas EPANET 2 does not. A confirmation test was carried out on the PDA results based on 66 constant supply-node heads from 90 m to 155 m in equal steps of 1 m for all the supply nodes. Siew and Tanyimboh (2012a) have previously tested the accuracy of EPANET-PDX. Therefore (unlike the test in Section 5.3.2.1 that included all hydraulic time steps), only the results of the last hydraulic time steps in each of the EPSs were included in the present test. The correlation coefficients obtained were R^2 = 0.999996 (or more simply $1\text{-}R^2 = 4\times10^{-6}$) for nodal heads and R^2 = 0.996 for pipe flow rates based on 15,048 demand-node heads (66 supply-node heads × 228 demand nodes) and 16,566 for pipe flow rates (66 supply-node heads × 251 pipes) as presented in Appendix B (Fig. B-1.5). This confirms the accuracy of the EPANET-PDX hydraulic analysis results for subnormal pressures.

In Section 5.3.2.1, the accuracy of the EPANET-PDX water quality results was investigated. It was shown that the EPANET-PDX results are essentially the same as those produced by EPANET 2 and EPANET-MSX, for identical regimes of flow and pressure. Therefore, confirmation here of the accuracy of the PDA results of EPANET-PDX also confirms the accuracy of the EPANET-PDX water quality results for operating conditions with subnormal pressure.



(a) Node 1 (b) Node2

Figure 5.11 Pipe closure effects on water quality in Network 2

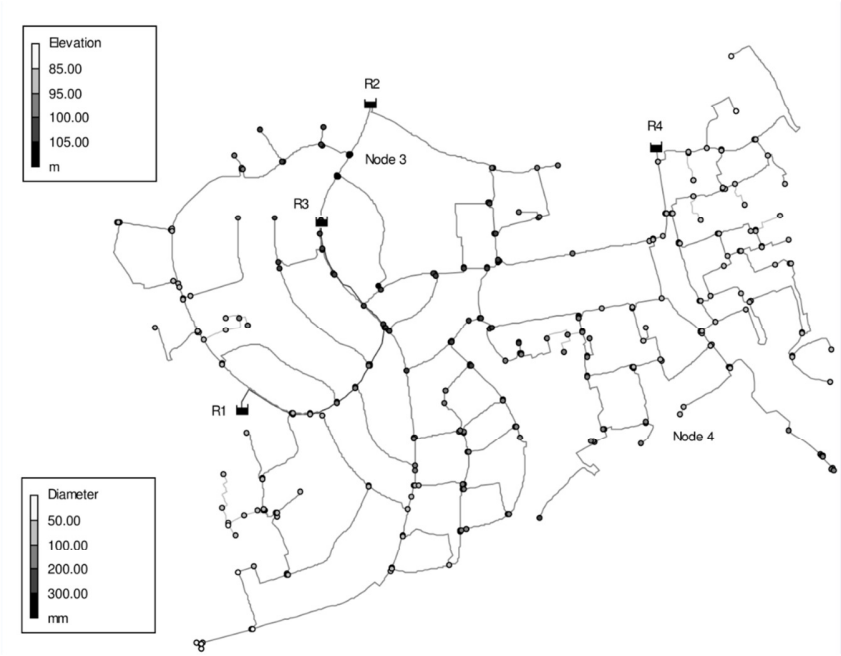


(a) supply-node head (b) pipe closure

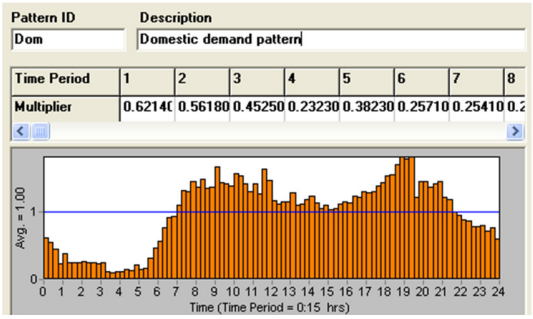
Figure 5.12 CPU times for Network 2 based on water age

5.3.3 Network 3

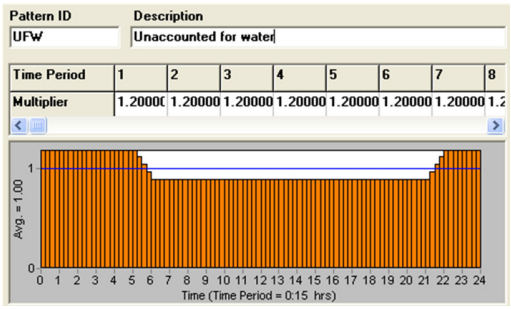
Network 3 (Fig. 5.13) consists of 416 pipes of various lengths, 380 demand nodes and 4 supply nodes (R1 to R4). The pipe sizes range from 50 mm to 500 mm in diameter. R2, R3 and R4 are constant-head supply nodes with a water level of 133 m. R1 is a variable-head supply node. The head level multipliers for R1 are available at 15-minute intervals (see Appendix B (Fig. B-2.1)). R1 was modelled here as a constant-head supply node using its average water level of 133 m (as explained previously for Network 2). Nodes 3 and 4, respectively, were selected to represent nodes close to a supply node and remote points in the network. Network 3 has no hydrants or fire-fighting flows.



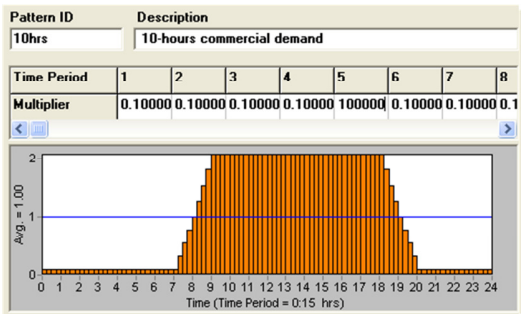
(a) layout with diameters and elevations



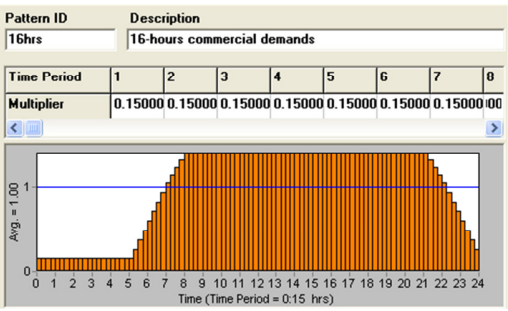
b) Domestic demand factors



(c) Unaccounted for water factors



(d) 10-hour commercial demand factors



(e) 16-hour commercial demand factors

Figure 5.13 Network 3: layout and demand factors

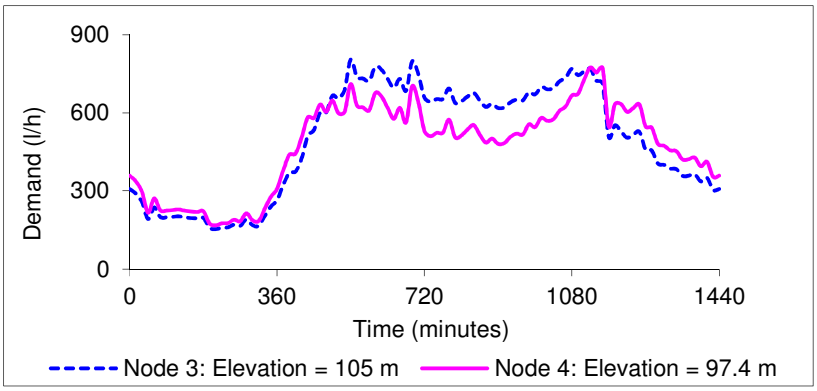
Extended period simulations (EPSs) were conducted on an Intel Core 2 Duo personal computer (CPU = 3.2 GHz, RAM = 3.21 GB) considering normal and pressure-deficient conditions. For all the EPSs, 15-minute hydraulic time step and 240 hours duration were used. The residual head for full demand satisfaction is 20 m. The results reported here are for the last 24 hours. In total, 2,534 EPSs of 240 hours were performed (see Table 5.4). For the chlorine, total THM concentration and water age, the agreement between EPANET 2, EPANET-MSX and EPANET-PDX was excellent for normal operating conditions with sufficient pressure. As in Network 2, the heads at the supply nodes were assumed constant. They were then reduced from 133 m to 112 m, 107 m, 102 m and 97 m, in turn, to obtain DSRs of 90%, 75%, 50% and 30% respectively. The daily demand and nodal flow variations at Nodes 3 and 4 can be seen in Fig. 5.14.

Table 5.4 Number of EPSs for Network 3

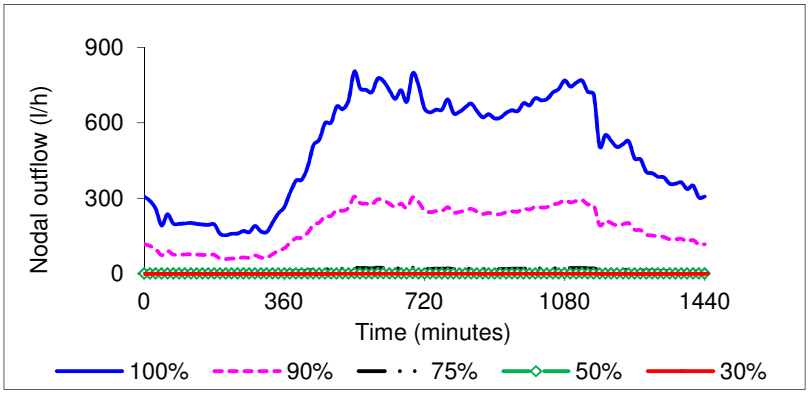
	EPANET 2	EPANET-PDX	EPANET-MSX
Normal pressure			
Water age	1	1	1
Chlorine	1	1	1
THM	1	1	1
Water age, chlorine and THM concurrently	NA[a]	NA	1
Pressure-deficient conditions[b]			
Water age	4	4	NA
Chlorine	4	4	NA
THM	4	4	NA
Water age, chlorine and THM concurrently	NA	NA	4
Pipe closures			
Water age	416	416	NA
Chlorine	416	416	NA
THM	416	416	NA
Total	1263	1263	8

[a] Not applicable

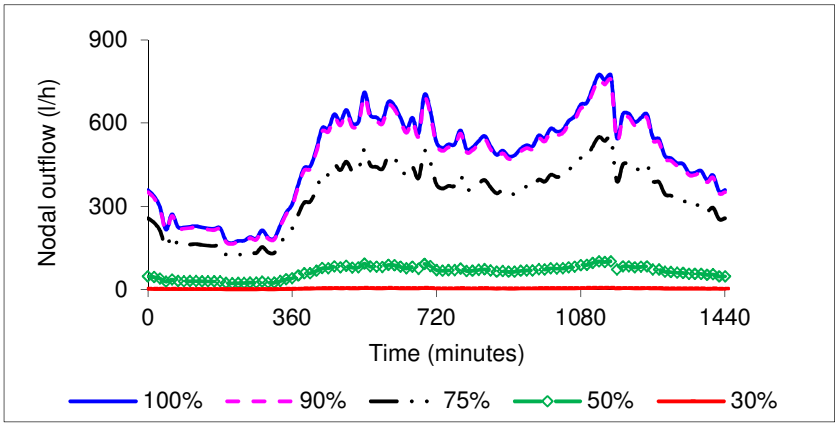
[b] One each for DSR= 90%, 75%, 50% and 30%



(a) Daily demand variations at Nodes 3 and 4



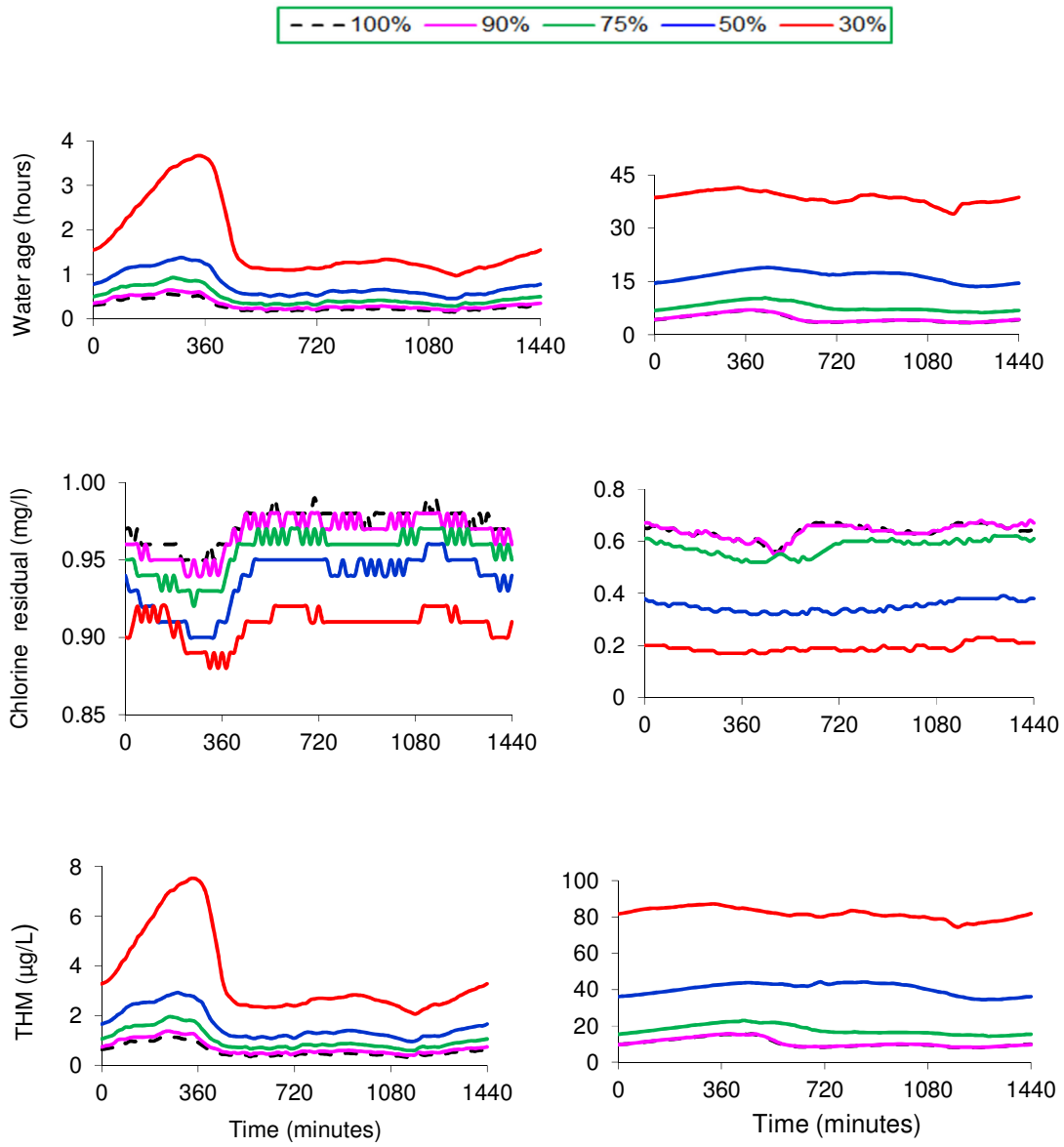
(b) Nodal flow variations at Node 3 for normal and low pressure conditions



(c) Nodal flow variations at Node 4 for normal and low pressure conditions

Figure 5.14 Network 3: Daily demand and nodal flow variations at Nodes 3 and 4. The percentages refer to the percentages of the total demand satisfied under normal and pressure-deficient conditions.

Fig. 5.15 provides a graphical summary of the water quality results for EPANET-PDX.



(a) Node 3 (b) Node 4

Figure 5.15 Water quality in Network 3 under normal and pressure-deficient conditions (EPANET-PDX). The percentages refer to the percentages of the total demand satisfied under normal and pressure-deficient conditions.

In general the effects of pressure deficiency (from water quality perspective) are greatest at the extremities of the network; the severity increases as the shortfall in pressure increases and temporal variations track the overall demand pattern (see, for example, Fig. 5.13 (b)-5.13 (e)). Unlike Node 2 in Network 2 for DSR = 30% that has almost zero flow, the water quality results at the remote node (Node 4) follow the demand pattern. Also, the individual pipes were closed to simulate pipe failures. Results are presented in Appendix B (Fig. B-2.2). The obtained results were very similar to the corresponding results for Network 2 shown in Fig. 5.11. Table 5.5 compares the computational speeds of EPANET-PDX, EPANET 2 and EPANET-MSX for the water quality analyses under normal pressure conditions for the 240-hour EPS. Based on these results, EPANET 2 is fastest and EPANET-MSX is slowest. Considering that the EPS covers a period of 240 hours (with a relatively small time step of 15 minutes), the results suggest that EPANET-PDX may be fast enough for regular use.

Table 5.5 CPU times for Network 3

	Mean CPU time per EPS (seconds)		
	EPANET 2	EPANET-PDX	EPANET-MSX
Chlorine	8.8	12.8	116.0
THM	12.6	16.5	23.0
Water age	9.1	12.5	17.0
Chlorine, THM and water age concurrently	NA	NA	146.0

[a]Not applicable

5.4 Conclusions

Besides water age, water quality modelling for low-pressure conditions in water distribution systems has been addressed in this chapter. The approach is an extension of EPANET 2 that integrates pressure dependency with hydraulic and chemical analyses while preserving the modelling functionality of EPANET 2. Convergence difficulties or failures were not experienced with EPANET-PDX for the various cases considered in this study. The accuracy of the model has been verified using a hydraulic consistency test for pressure deficient conditions, along with EPANET 2 and EPANET-MSX for normal operating conditions.

Sample results based on a simple network in literature and two water supply zones in the UK are included. In total 4,372 EPSs were performed for the two water supply zones using EPANET 2, EPANET-MSX and EPANET-PDX. The results show that, if pressure is low, the conventional demand driven modelling approach can provide misleading results that in turn can lead to inappropriate water quality policy decisions. An important corollary worth stating is that, under conditions of low pressure, poor demand-driven analysis estimates of the spatial distribution of the water age could mislead efforts intended to identify the source of accidental or intentional contamination.

Finally, the results here have highlighted the need for more PDA related work including the incorporation of dispersion in the water quality model and the collection of field data under conditions of low pressure, low flow rates and low velocities.

Chapter 6

High Performance and Parallel Computing for Complex Water Distribution Systems Design Optimisation Problems

6.1 Introduction

Water distribution systems (WDSs) are key components of public infrastructures and it is essential to design and rehabilitate them in a cost effective manner without compromising the required performance and regulatory standards. Evolutionary algorithms (EAs) are a commonly applied optimisation approach. EAs such as genetic algorithms (GAs) are popular in providing optimal or near optimal solutions to WDSs optimisation problems (Simpson et al., 1994). A review on the current developments in GAs for WDSs optimisation is presented in Chapter 3 (Section 3.3.)

One of the main issues in GAs is that the algorithms are time consuming when applied to large optimisation problems such as real life networks with large numbers of pipes and multiple operating conditions (VenZyl et al., 2004). For example, in the optimisation of large water distribution systems, a single optimisation run may involve thousands of hydraulic and water quality simulations (see e.g. Ghebremichael et al., 2008; Seyoum and Tanyimboh, 2014) and that may take many days on modern

computers such as workstations. Such computational time, however, is usually unacceptable for water utilities applications. One way to address this difficulty is by utilising high performance and parallel computing techniques.

The aim of this chapter is twofold. The first aim is to assess a multi-objective evolutionary algorithm introduced by Siew and Tanyimboh (2012b) that is based on the Non-dominated Sorting Genetic Algorithm II (Deb et al., 2002). EAs for WDSs often use penalties to assess the merits of infeasible solutions when solving optimization problems that have constraints. By contrast, the penalty-free multi-objective evolutionary algorithm (PF-MOEA) proposed by Siew and Tanyimboh (2012b) uses pressure-dependent analysis that accounts for the pressure dependency of the nodal flows and thus avoids the need for penalties to address violations of the nodal pressure constraints. The pressure-dependent analysis that PF-MOEA uses is the pressure-dependent extension of EPANET 2 (EPANET-PDX) described in Chapter 4 and 5 (Siew and Tanyimboh, 2012a). EPANET-PDX simulates WDSs with insufficient flow and/or pressure more realistically.

PF-MOEA has been discussed previously in terms of the least-cost solution and the smallest number of function evaluations achieved on some standard test problems (Siew and Tanyimboh, 2010a, 2011b, 2012b). However, operators used in genetic algorithms (e.g. mutation) are probabilistic in nature. Thus, a statistically more robust assessment is used here that reflects the stochastic nature of the algorithm. New solutions that are hydraulically feasible and cheaper than the current best solution in the literature were found, for the Kadu et al. (2008) network that represents one of the challenging benchmark problems in the literature. An improvement in cost of almost 5% was achieved.

The second aim of this chapter is to implement the high performance and parallel computing approach to solve real-life network design problem. The 251-pipe network

described in Chapter 4 and 5 was considered. The network comprises multiple sources, multiple demand categories, many fire flows and involves extended period simulation. Due to the size and complexity of the optimisation problem, a high performance computer that comprises multiple cores was used for the computational solution. Multiple optimisation runs were performed concurrently using PF-MOEA. Overall, the algorithm performs well; it consistently provides least cost solutions that satisfy all the system requirements quickly. The least-cost design obtained was significantly cheaper than (*48%*) the existing network in terms of the pipe costs. Also, the computational performance of PF-MOEA has been enhanced by parallelising the algorithm using a Message-Passing Interface parallel programming model. The controller-worker parallelisation approach has been implemented in this chapter. The primary aim of this work is to demonstrate the computational performance that can be achieved by using a simple and straightforward parallel implementation approach. Results in general indicated that the parallel algorithm consistently outperforms the serial algorithm with respect to computational time while providing comparable solutions. The average *speedup* achieved by the parallel algorithm is *15* on an *eight*- core workstation. A review on the most common parallelisation approach that has been applied in WDS literature is provided in Chapter 3 (Section 3.3.7).

In the remainder of this chapter, the PF-MOEA approach and its application is described in Section 6.2. In Section 6.3, the application of high performance and parallel computing of EAs for multi-objective optimization of WDSs is presented. Two network examples were considered for this chapter. The results achieved are discussed in Section 6.4.

6.2 Penalty-Free Multi-Objective Evolutionary Optimization Approach

Evolutionary algorithms (EAs) by nature start with a randomly generated set of solutions that may include both feasible and infeasible solutions. To address the node pressure constraints, penalty methods have been applied widely (Savic and Walters, 1997; Vairavamoorthy and Ali, 2000; Broad et al., 2005; Ostfeld and Tubaltzev, 2008). The major drawback of the penalty-based approach is that additional case-specific parameters are required whose calibration is generally challenging (Siew and Tanyimboh, 2012b; Siew et al., 2014; Saleh and Tanyimboh, 2013; Prasad and Park, 2004). A review on different constraint handling techniques for EAs is provided in Chapter 3 (Section 3.3.4).

In an attempt to alleviate the difficulties on handling the node pressure constraints, Siew and Tanyimboh (2012b) proposed a penalty-free multi-objective evolutionary algorithm (PF-MOEA) that eliminates the use of penalty method. The approach allows all the feasible and infeasible solutions generated to compete in a way that is fundamentally bias-free with respect to constraint violation. PF-MOEA uses pressure-dependent analysis to assess each individual in the population of solutions. Unlike the conventional demand-driven analysis approach, pressure-dependent analysis takes proper account of the relationship between the flow and pressure at a node. By definition, feasible solutions satisfy all nodal demands in full. Conversely, infeasible solutions do not and the shortfall in the water they supply represents a real measure of the infeasibility of the water distribution system. In this way, pressure-dependent analysis addresses the node pressure constraints as an integral part of the hydraulic analysis. PF-MOEA employs the pressure-dependent extension of EPANET 2 (Siew and Tanyimboh, 2012a) to carry out pressure-dependent analysis seamlessly. A full description on the pressure dependent model (EPANET-PDX) is provided in Chapter 4 and 5.

6.2.1 Previous Applications of PF-MOEA

PF-MOEA has been previously applied to various aspects of WDSs optimization that include design, operation and long-term rehabilitation and upgrade of WDSs (Siew and Tanyimboh, 2011; Siew and Tanyimboh, 2012b; Siew et al., 2014). The algorithm used to solve the three well-known optimization problems in literature, i.e. the design of the Two-Loop (Alperovits and Shamir, 1977) and Hanoi (Fujiwara and Khang, 1990) WDSs and the expansion of the New York Tunnels. Also, it was utilised for the phased whole-life design and rehabilitation of WDSs using the real life network Wobulenzi (Siew et al., 2014). The algorithm used to address WDSs optimisation problems that involve multiple loadings, storage tanks and pumps of the benchmark "Anytown" network. Overall, PF-MOEA generated superior results for all the optimization problems solved in terms of cost, hydraulic performance and computational efficiency compared to other solutions in literature.

6.2.2 PF-MOEA Formulation

Minimising the total network cost (capital and operation) and maximising the network hydraulic performance are the two conflicting objectives of PF-MOEA. The network hydraulic performance ensures that all nodal demands are satisfied by maximising the total available flow of the most critical node in the network. The conflicting objectives produce a set of non-dominated solutions where no one solution in the set can be considered superior to the others. The objective functions are expressed by equations below.

$$\text{Minimise} \quad F_1 = (CR)^2 \tag{6.1}$$

$$\text{Maximise} \quad F_2 = (DSR)^4 \tag{6.2}$$

where F_1 and F_2 represent the first and second objective functions respectively. *CR* is the cost ratio i.e. the ratio of the cost of a particular solution to the cost of the most expensive solution in the whole population within a single generation. *DSR* is the demand satisfaction ratio i.e. the ratio of the available flow to the required flow and measures the feasibility of a solution. Both objective functions are thus normalised and have values between zero and one. A solution that has a *DSR* value that is less than one is infeasible and cannot satisfy the demands in full. The objective functions in Eq. 6.1 and 6.2 focus the search around the frontier between feasible-infeasible solutions where optimal solutions are commonly found. PF-MOEA seamlessly couples the widely used Non-dominated Sorting Genetic Algorithm II (Deb et al., 2002) with the hydraulic analysis model EPANET-PDX. A description on the NSGA II procedure is provided in Chapter 3 (Section 3.3.5). The decision variables in PF-MOEA are represented using binary coding. Single-point crossover, single-bit mutation and binary tournament selection are the genetic operators used in the algorithm.

6.3 High Performance and Parallel Computing of Evolutionary Algorithms

The utilisations of high performance computing techniques are increasingly becoming important to solve computationally intensive applications. High performance computing is a computing system where multiple computers connected together as a cluster to solve large-scale problems that are difficult or impossible to execute on standard desktop computers (Wilkinson and Allen, 2004). In parallel computing large problems are often divided into several smaller ones that can be solved simultaneously on parallel processors in a shorter time (Trobec et al., 2009). EAs are one of the areas that can benefit from parallel computing. EAs such as genetic algorithms work with a population of independent solutions. This makes the algorithms suitable to be implemented in parallel computing architectures effectively (Cantú-Paz and Goldberg, 2000).

Evolutionary operators such as crossover, mutation and fitness evaluation can be executed in parallel across different processors. In this chapter, a controller-worker approach is applied to parallelise PF-MOEA. Controller-worker is the widely used application of parallel EAs where a single controller processor executes the routine operation of the algorithm and employs the workers to carry out fitness evaluation. It is the most straightforward approach with a potential of improving computational performances significantly (Nowostawski and Poli, 1999; Cantú-Paz and Goldberg, 2000).

Fitness evaluation is the most computationally expensive operator in EAs (Alba and Tomassini, 2002; Schutte et al., 2004). Other evolutionary operators such as selection, crossover and mutation by far require small CPU time. In this work, the fitness evaluation of the child population is parallelized. It is worth noting here that fitness evaluation in this parallel implementation context refers to the evaluation of solutions using the second objective function described in Eq. 6.2. This involves extended period simulations and fitness calculation over generations. The task is divided equally among all the processors. The controller processor retains a portion of the population to carry out evaluation in parallel with the worker processors. In addition, the controller processor is in charge of evolutionary operators such as selection, crossover and mutation. It also performs all the remaining tasks in PF-MOEA.

Fig. 6.1 shows a flowchart that illustrates the implementation of the controller-worker approach for parallelizing PF-MOEA. The proposed parallel program is written in C++ by using Message Passing Interface (MPI) communication routines. Microsoft HPC pack 2008 is used to run the program in Microsoft Visual Studio (version 2010). Parallelization of fitness evaluation was done by assigning equal portion of the child population to each of the participating processors. Communication among processors takes place when each worker processers receives the individuals to evaluate and when the workers send the fitness value to the controller. As can be seen in the flowchart, after

generation of child population via crossover and mutation, the controller sends individuals to worker processes for evaluation and fitness calculation. Subsequently, both the controller and worker processors evaluate their portion of population. The workers return the evaluations to the controller as soon as they finish. After the controller processor receives all the fitness values for the child population, it combines the parent and child solutions (each of which has size N_P) to perform non-domination ranking and crowding distance calculation for the combined solutions. Based on the non-domination ranks and crowding distances, N_P solutions are then selected for the next generation. Finally, the algorithm stops and terminates the MPI execution environment when the maximum number of generations is reached. It is worth mentioning that in a situation where the number of solution in the best non-dominated front exceeds the population size (N_P), 30% of the best feasible solutions (least cost-feasible solutions) from the front are retained in each generation in the PF-MOEA procedure. The remaining solutions (70%) are selected based on crowding distance.

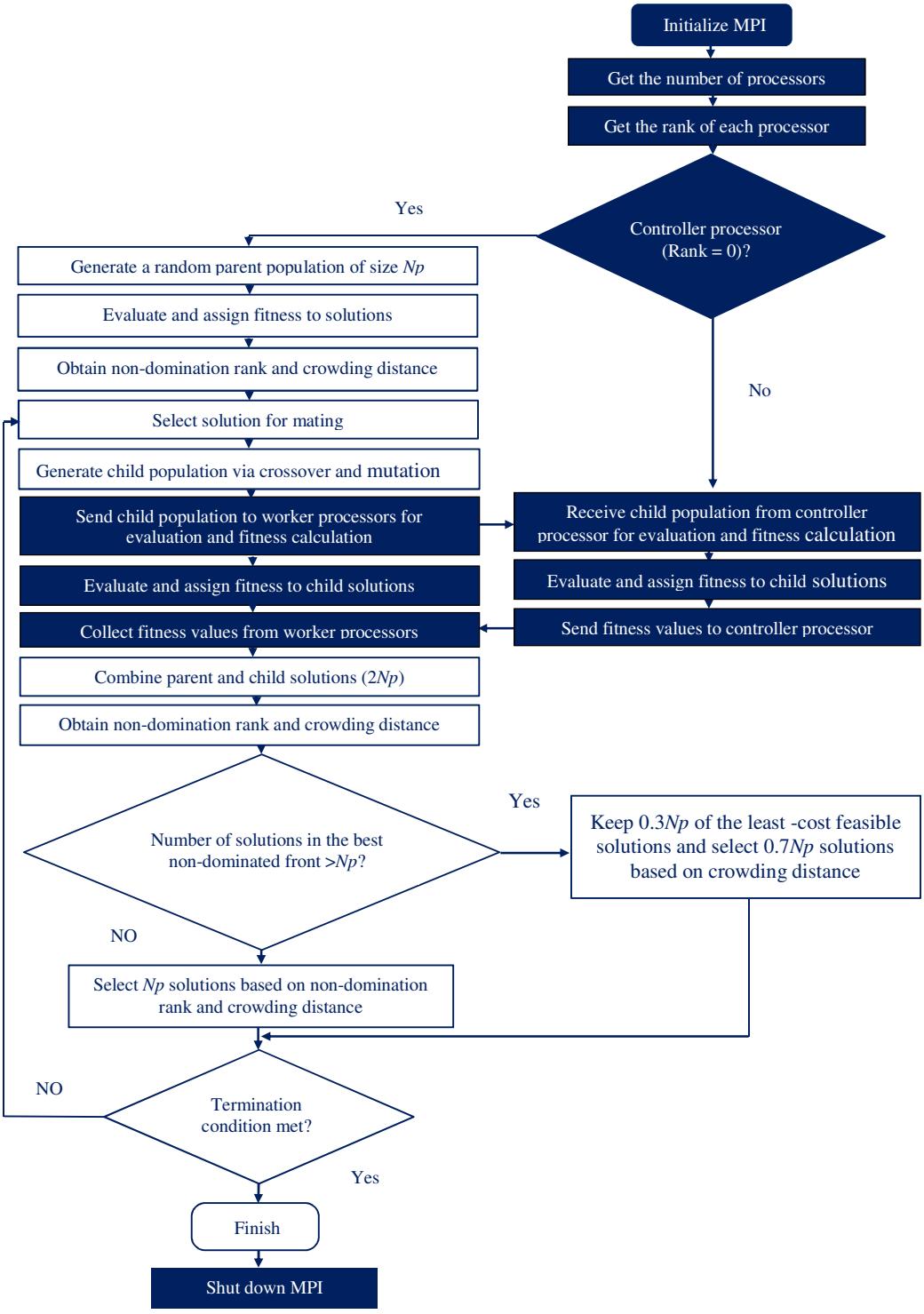


Figure 6.1 Flowchart of the parallelized PF-MOEA

6.4 Case Studies

Two network examples have been used in this chapter. Both examples are design optimisation problems where their objectives were to obtain the cheapest possible combination of pipe sizes that satisfy all the system requirements. The first example (Network 1) was taken from Kadu et al. (2008) that has the details of the optimization problem. In this example, extensive statistical investigation on the performance of PF-MOEA was conducted. The second example (Network 2) is one of the water supply zones of a network in the UK that was described in full in Chapter 4. The network is used as a case study to implement the high performance and parallel computing approach to solve real-life network design problem that comprises multiple loadings and supply sources. It involves extended period simulations that take into account the demand variation in time.

The design optimization problem of Network 2 is executed both sequentially and in parallel. For the sequential computing, the high performance facility at the University of Strathclyde was used to perform multiple serial optimisation runs concurrently. The parallelised version of PF-MOEA was implemented on a workstation to execute parallel optimisation runs. The effectiveness of the parallelised algorithm was evaluated considering both the quality of solution and speedup together in reference to results from sequential computing.

6.4.1 Network 1

Network 1 (Fig. 6.2) is fully looped and consists of 24 demand nodes, 34 pipes and 9 loops. There are two reservoirs with constant water levels of 100 m and 95 m, respectively. The network data including the minimum required nodal heads can be found in Appendix C (Table C-1.1). Fourteen candidate pipe sizes are available for this

network. The diameters and costs of these pipes can be referred in Appendix C (Table C-1.2). Therefore, with 34 pipes and 14 pipe sizes, there are $14^{34} \approx 9.3\times10^{38}$ feasible and infeasible solutions. The Hazen-Williams equation with parameters as shown was used.

$$h_{ij} = \frac{\omega L_{ij} Qp_{ij}^{\alpha^e}}{C_{ij}^{\alpha^e} D_{ij}^{\beta^e}} \tag{6.3}$$

where h_{ij}, L_{ij}, Qp_{ij}, C_{ij} and D_{ij} represent the head loss, length, flow rate, roughness coefficient and diameter of pipe ij respectively; ω is a dimensionless conversion factor for the units used; α^e and β^e are exponents. The values for C_{ij}, α^e, β^e and ω used are 130, 1.85, 4.87 and 10.68 respectively.

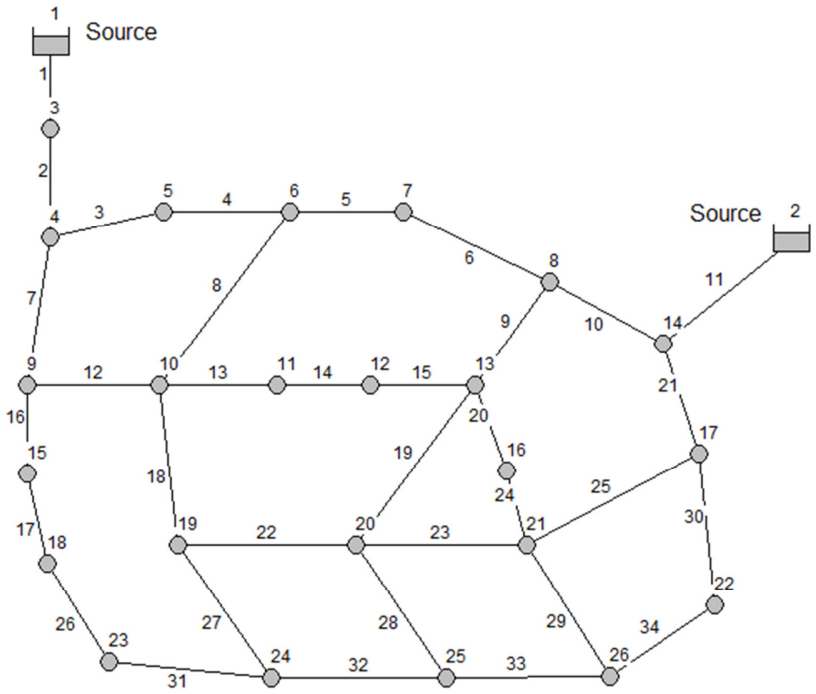


Figure 6.2 Layout of Network 1

The initial populations were generated randomly. The maximum number of function evaluations (i.e. hydraulic simulations) permitted per optimization run was 500,000 and the crossover probability was $p_c = 1$ in all cases. Other parameters, including the population size, mutation rate and number of optimization runs were as shown in Table 6.1. A total of *eight* cases were considered. One case (out of 8) had the default values of the coefficients of the Hazen-Williams formula in EPANET 2 (i.e. $\alpha^e = 1.852$, $\beta^e = 4.871$, $\omega = 10.667$). Summarised results are shown in Table 6.1 based on sample sizes (i.e. the total number of optimization runs) of 100 (in 6 cases out of 8) and 30 (in 2 cases out of 8) as the initial results suggested the smaller sample size might also be statistically satisfactory.

The cheapest solution obtained was 125,460,980 Rupees (i.e. with $Np = 500$, $p_m = 0.05$ in Table 6.1), within 436,000 function evaluations. Other researchers have not found this solution previously and it is the cheapest hitherto. Also, the *minimum-cost* solutions found in the eight cases considered were close to the *smallest* minimum-cost. The *means* of the minimum-cost differ from the *smallest* minimum-cost achieved here by only 2.31–4.15%. The average number of function evaluations to obtain convergence (within the specified maximum of 500,000 function evaluations) ranged from 354,160 to 397,083. The average CPU time to achieve convergence was 1.11–1.25 hours. To complete a single optimisation run consisting of 500,000 function evaluations, due to differences in the number of iterations per hydraulic simulation, the average CPU time required was 1.57 hours and the standard deviation was 0.10 hours on a personal computer (Intel Core 2 Duo with 2.5 GHz CPU and 1.95 GB RAM).

Further limited sensitivity analysis was also conducted to check the influence of the mutation rate. Nine different additional mutation rates spread approximately evenly in the range $p_m = [0.01, 0.7]$ were used; $p_m = 0.05$ and $p_m = 0.07$ that feature in Table 6.1 were excluded. The population size was $Np = 200$ and the maximum number of function evaluations permitted was 500,000. Only five optimization runs were executed for each

mutation rate. Therefore, with only five trials per mutation rate for the limited sensitivity analysis, a fixed set of five different initial populations (each with Np = 200) that were generated randomly was used for all the mutation rates considered. The *minimum-cost* feasible solution for the nine mutation rates ranged from 126,035,000 to 129,832,000 Rupees. For the nine mutation rates, the *average* minimum-cost (based on 5 optimization runs per mutation rate) ranged from 127,615,000 to 131,353,800 Rupees. Based on the results in Table 6.1, it can be expected that a population size of Np = 500 would provide even better results. Overall, for the mutation rates and population sizes attempted, the performance of the optimization algorithm was consistently reliable and satisfactory.

Kadu et al. (2008) proposed a critical path concept to reduce the number of candidate diameters for each pipe as mentioned in Chapter 3 (Section 3.3.3). In this way, they reduced the solution space to 8.65×10^{20} feasible and infeasible solutions. PF-MOEA was also tested as summarised in Table 6.2 using the same reduced solution space of size 8.65×10^{20} solutions as in Kadu et al. (2008). The reduced candidate pipe diameters can be referred in Appendix C (Table C-1.3). The *minimum cost* achieved for a feasible solution was 125,826,425 Rupees within 82,400 function evaluations. This is only 0.29% more costly than the smallest cost that was achieved for the full solution space of size $14^{34} \approx 9.3\times10^{38}$ solutions. Overall, the values of the cost and function evaluations are improved, on average, by reducing the size of the solution space. On average, approximately 27% fewer function evaluations, i.e. hydraulic simulations, were required to find a near-optimal solution when the solution space was reduced, in comparison to the full solution space.

Haghighi et al. (2011) also solved the same optimization problem using a hybrid approach consisting of a genetic algorithm and integer linear programming as stated in Chapter 3 (Section 3.3.6). A comparison between the PF-MOEA and the best results reported previously in the literature is shown in Table 6.3. For the full solution space,

the solution obtained by Kadu et al. (2008) was 131,678,935 Rupees, within 120,000 function evaluations. This is 4.96% more expensive than the new best solution of 125,460,980 Rupees. Haghighi et al. (2011) achieved 131,312,815 Rupees, within 4,440 function evaluations. This is 4.66% more expensive than the new best solution. For the reduced solution space, Kadu et al. (2008) obtained a solution of 126,368,865 Rupees, within 25,200 function evaluations. This is 0.72% more expensive than the new best solution.

However, the feasibility of the Kadu et al. (2008) and Haghighi et al. (2011) solutions is questionable. Based on EPANET 2, the Kadu et al. (2008) and Haghighi et al. (2011) solutions were deemed infeasible (as shown in Table 6.4). The Kadu et al. (2008) solutions violate the minimum node-pressure requirement at Nodes 12, 24 and 25 (for the full solution space) and Node 26 (for the reduced solution space). Similarly, the Haghighi et al. (2011) solution violates the minimum node-pressure requirement at Nodes 13, 24 and 25. By contrast, the new solutions of PF-MOEA are all feasible.

Table 6.1 Computational characteristics of PF-MOEA for Network 1 for the full solution space

Full solution space					
Maximum number of function evaluations allowed	$500x10^3$				
Number of GA runs[d]	100[a]			30[a,d]	30[b,d]
Population size	200	500			
Mutation rate	0.005		0.05	0.07	0.05
Cost ($Rupeesx10^6$)					
Minimum	126.740 (1.02)[c]	126.067 (0.48)	125.461 (0.00)	125.820 (0.29)	125.551 (0.00)
Maximum	134.243 (7.00)	133.005 (6.01)	132.986 (6.00)	131.797 (5.05)	128.222 (2.13)
Mean	130.662 (4.15)	129.479 (3.20)	128.623 (2.52)	128.363 (2.31)	127.019 (1.17)
Median	130.904 (4.34)	129.391 (3.13)	128.588 (2.49)	128.253 (2.23)	127.054 (1.20)
Standard deviation	1.707	1.689	1.506	1.43	0.716
Number of better solutions from the best run than previous solutions					
Cheaper than Haghighi et al. (2011)	19	31	30	29	16
Cheaper than Kadu et al. (2008)	20	32	30	30	17
Number of function evaluations to achieve convergence					
Minimum	140,000	198,500	112,000	195,500	204,500
Maximum	497,800	498,500	498,000	497,500	500,000
Mean	354,160	396,505	395,515	397,083	370,483
Median	364,500	413,000	418,750	402,750	402,250
Standard deviation	101,000	80,000	95,000	82,000	98,000

[a]Hazen Williams parameters: $\alpha^e = 1.85$; $\beta^e = 4.87$; and $\omega = 10.68$

[b]Based on the default EPANET 2 Hazen Williams parameters: $\alpha^e = 1.852$; $\beta^e = 4.871$; and $\omega = 10.667$

[c]The deviations from the least cost achieved for the full solution space are shown in the parentheses (%)

[d]30 optimisation runs were carried out in two tests after six preceding tests with 100 optimisation runs revealed highly consistent results

Table 6.2 Computational characteristics of PF-MOEA for Network 1 for the reduced solution space

Reduced solution space			
Maximum number of function evaluations allowed	$500x10^3$		
Number of GA runs	100[a]		
Population size	100	200	
Mutation rate	0.005		0.05
Cost ($Rupeesx10^6$)			
Minimum	125.919 (0.36)[b]	125.826 (0.29)	125.826 (0.29)
Maximum	130.868 (4.31)	127.281 (1.45)	127.322 (1.48)
Mean	127.435 (1.57)	126.182 (0.57)	126.007 (0.44)
Median	127.244 (1.42)	126.112 (0.52)	125.883 (0.34)
Standard deviation	1.035	0.328	0.291
Number of better solutions from the best run than previous solutions			
Cheaper than Haghighi et al. (2011)	30	32	31
Cheaper than Kadu et al. (2008)	30	32	31
Number of function evaluations (FEs) to achieve convergence			
Minimum	32,000	43,400	54,600
Maximum	497,100	499,000	493,800
Mean	278,145	285,998	269,554
Median	293,000	293,600	265,700
Standard deviation	130,000	124,000	125,000

[a]Hazen Williams parameters: $\alpha^e = 1.85$; $\beta^e = 4.87$; and $\omega = 10.68$

[b]The deviations from the least cost achieved for the full solution space are shown in the parentheses (%)

Table 6.3. Comparison of alternative solutions for Network 1

Pipe ID	Diameter (mm)				
	Full Solution Space (FSS)			Reduced Solution Space (RSS)	
	Kadu et al. 2008	Haghighi et al. 2011	PF-MOEA	Kadu et al. 2008	PF-MOEA
1	1000	1000	900	1000	900
2	900	900	900	900	900
3	400	400	350	350	400
4	350	350	300	250	250
5	150	150	150	150	150
6	250	250	250	250	200
7	800	800	800	800	800
8	150	150	150	150	150
9	400	400	450	600	600
10	500	500	500	700	600
11	1000	1000	900	900	900
12	700	700	700	700	700
13	800	800	500	500	500
14	400	400	500	450	500
15	150	150	150	150	150
16	500	500	500	450	500
17	350	350	350	350	350
18	350	350	400	400	350
19	150	150	150	450	450
20	200	150	150	150	150
21	700	700	700	600	600
22	150	150	150	150	150
23	400	450	450	150	150
24	400	400	350	400	350
25	700	700	700	500	600
26	250	250	250	200	250
27	250	250	250	350	300
28	200	200	300	250	300
29	300	300	200	150	200
30	300	300	250	300	300
31	200	200	150	150	150
32	150	150	150	150	150
33	250	200	150	150	150
34	150	150	150	200	150
Cost (Rupees)	131,678,935	131,312,815	125,460,980	126,368,865	125,826,425
Function evaluations	120,000[a]	4,440[a]	436,000	25,200[a]	82,400

[a] Infeasible solutions (based on EPANET 2 with $\alpha^e = 1.85$; $\beta^e = 4.87$; $\omega = 10.68$)

Table 6.4. Comparison of nodal heads from alternative solutions of Network 1

Node ID	Required head (m)	Available head (m) (based on EPANET 2 with $\alpha^e = 1.85$; $\beta^e = 4.87$; $\omega = 10.68$)				
		Full Solution Space (FSS)			Reduced Solution Space (RSS)	
		Kadu et al. 2008[a]	Haghighi et al. 2011[a]	PF-MOEA	Kadu et al. 2008[a]	PF-MOEA
1[b]	100.00	100.00	100.00	100.00	100.00	100.00
2[b]	95.00	95.00	95.00	95.00	95.00	95.00
3	85.00	98.95	98.96	98.28	98.98	98.26
4	85.00	95.65	95.66	95.04	95.76	94.98
5	85.00	90.85	90.85	87.47	88.79	90.68
6	85.00	89.40	89.41	85.63	85.28	85.07
7	82.00	87.75	87.73	85.83	87.99	82.95
8	82.00	89.99	89.96	88.82	91.62	89.35
9	85.00	91.77	91.79	91.12	91.83	91.05
10	85.00	89.05	89.08	88.30	88.88	88.22
11	85.00	88.85	88.88	86.42	87.1	86.38
12	85.00	84.98* (0.02)	85.01	85.13	85.13	85.12
13	82.00	82.02	81.88* (0.12)	83.25	86.73	84.85
14	82.00	94.49	94.49	94.14	94.13	94.15
15	85.00	88.44	88.46	87.97	87.11	87.92
16	82.00	84.53	84.81	83.11	82.05	83.04
17	82.00	90.88	90.88	90.69	90.29	90.04
18	85.00	85.46	85.47	85.39	85.24	85.39
19	82.00	85.11	85.24	86.14	85.93	83.82
20	82.00	82.1	83.78	83.15	83.72	82.07
21	82.00	87.39	87.38	87.37	83.98	87.00
22	80.00	86.45	86.55	80.69	84.8	85.50
23	82.00	82.09	82.07	82.96	82.17	83.05
24	80.00	79.94* (0.06)	79.89* (0.11)	80.28	83.63	80.86
25	80.00	79.96* (0.04)	79.77* (0.23)	81.10	80.15	80.54
26	80.00	82.87	84.04	80.04	78.38* (1.62)	80.39
Total shortfall in head (m)		0.12[a]	0.46[a]	0.00	1.62[a]	0.00
Cost (Rupees)		131,678,935	131,312,815	125,460,980	126,368,865	125,826,425
Function evaluations		120,000	4,440	436,000	25,200	82,400

[a] Infeasible solutions; the asterisk indicates a shortfall in head (shown in parentheses)
[b] Reservoir (i.e. source with constant head)

Barlow and Tanyimboh (2014) have also solved Network 1 using a memetic algorithm that combines a genetic algorithm with a local improvement and cultural learning operators, which is described in Chapter 3 (Section 3.3.6). They used the default EPANET 2 Hazen–Williams head loss parameters that are less restrictive than the parameters used by PF-MOEA to obtain the best solutions (see Table 6.1 and 6.2). Both memetic algorithm and genetic algorithm (i.e. without the local and cultural improvement operators) were used to solve the network. Each algorithm carried out 100 optimisation runs. For each optimisation runs, 10 million function evaluations (i.e. hydraulic simulations) were allowed. A least cost solution of 124,690,000 Rupees within 142,000 function evaluations for the memetic algorithm and within 2,572,200 for the genetic algorithm was reported. As can be seen in Table 6.1, PF-MOEA has also solved the network using the EPANET 2 standard head loss parameters by carrying out 30 optimisation runs. The least-cost solution obtained was 125,551,000 Rupees within 500,000 function evaluations. This is only 0.7% costlier than the best solution by Barlow and Tanyimboh (2014) s' best solution. Barlow and Tanyimboh (2014) have also assessed their algorithm in terms of the number of function evaluations required to find a feasible solution within 1% of the solution reported by Haghighi et al. (2011). Accordingly, the memetic algorithm required 84,000 function evaluations while the genetic algorithm required 665,600 function evaluations. The corresponding number of function evaluations required by PF-MOEA was 86,567. However, it is worth noting here that the network performance in the Barlow and Tanyimboh (2014) approach is measured based on the total nodal pressure deficit throughout the network. A drawback of this approach is that it generates only the least cost feasible solution. The approach is not capable of recognising other feasible solutions. By contrast, PF-MOEA, which uses demand satisfaction ratio as a performance indicator of a solution, is efficient in providing ranges of feasible and infeasible solutions that provides more flexibility in choosing a particular solution. Additionally, Barlow and Tanyimboh (2014) method utilised many operators (e.g. random mutation, creeping mutation, local and cultural

improvement operators) and user-specified parameters (e.g. frequency of applying local and cultural improvement operators, number of individuals selected for cultural improvement, percentage of non-dominated front available for local improvement, etc.), which affect the efficiency and effectiveness of the algorithm. The PF-MOEA approach is by far simple and straightforward to implement. It utilised the most basic operators such as single point crossover and single bit wise mutation.

Fig. 6.3 shows the pipe velocities of the best PF-MOEA designs that were obtained based on the full and reduced solution spaces. Although velocity constraints were not required in the original problem specification, the pipe velocities for the two optimised best designs were checked. The minimum and maximum pipe velocities in both designs were 0.1 m/s and 2.7 m/s respectively.

Fig. 6.4 (a) illustrates the progress of PF-MOEA for the best solutions achieved for both the full and reduced solution spaces. For the full solution space, a cost reduction from 294,152,000 Rupees at the start of the optimization to (131,003,000 Rupees and 73,500 function evaluations) was achieved. The algorithm converged at (125,460,980 Rupees and 436,000 function evaluations). Also, fast reductions in cost from 175,535,000 Rupees at the start, i.e. at zero function evaluation, to (131,184,000 Rupees and 8,000 function evaluations) and then (127,850,000 Rupees and 25,000 function evaluations) were achieved for the reduced solution space. The algorithm finally converged at (125,826,425 Rupees and 82,400 function evaluations).

It is worth emphasizing, also, that the algorithm found dozens of feasible solutions that are cheaper than the solutions found by Kadu et al. (2008) and Haghighi et al. (2011) as summarised in Table 6.1 and 6.2. For example, for the full solution space and a population size of Np = 500, each of the mutation rates i.e. p_m = 0.005 and p_m = 0.05 achieved 30 *solutions* that are cheaper than Kadu et al. (2008) and Haghighi et al. (2011) in a *single* optimization run. For the total 705 optimization runs executed, the PF-

MOEA discovered more than 4,000 individual solutions approximately (with mean, median and standard deviation of 129,125,871 Rupees, 129,195,000 Rupees and 1,523,810 Rupees, respectively) that are feasible and cheaper than the previous best solution of 131,678,935 Rupees by Kadu et al. (2008). The distribution of these new solutions is shown in the Fig. 6.5. These results illustrate clearly the *high evolutionary sampling efficiency* of the PF-MOEA. In other words, the number of solutions evolved and analysed on average before finding a near-optimal solution is small in comparison to the size of the solution space. Also, the small distance between the graphs in Fig. 6.4(a) for the full and reduced solution spaces is worth a mention, considering that the reduced solution space is approximately a factor of 10^{18} smaller than the full solution space.

The Pareto-optimal fronts for the best optimization runs are shown in Fig. 6.4(b) for both the full and reduced solution spaces. In the full solution space, all permissible pipe sizes were included. However, in the reduced solution space, for each pipe, the pipe sizes that are unlikely to be feasible and/or competitive were excluded, leaving only 3–5 options in each case (see Appendix C (Table C-1.3)). It can be seen in Fig. 6.4(b) that the front for the full solution space (in which Np = 500) has a higher density of solutions than the front for the reduced solution space (in which Np = 200), and some solutions with the smallest and largest cost values are missing in the front for the reduced solution space.

It is worth mentioning that the solutions shown in Fig. 6.5 include the solutions found in the 30 optimisation runs with the default values of the coefficients of the Hazen-Williams formula in EPANET 2. Approximately 20-25 solutions in total (out of 4135) may be considered theoretically borderline in terms of feasibility due to modelling and unit conversion errors including minor differences between EPANET 2 and EPANET-PDX. For these theoretically borderline solutions, the shortfall in the residual head at the node with the smallest pressure may be up to about 2 cm approximately.

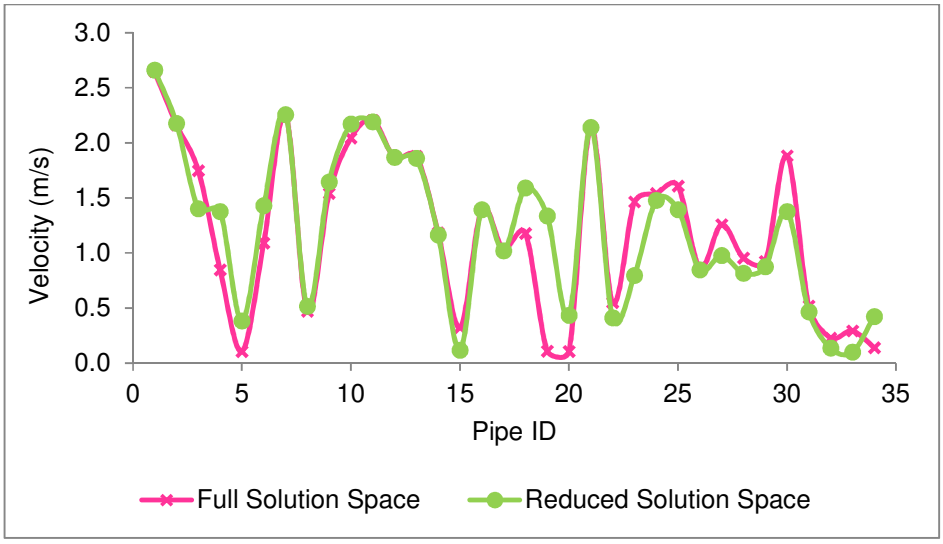
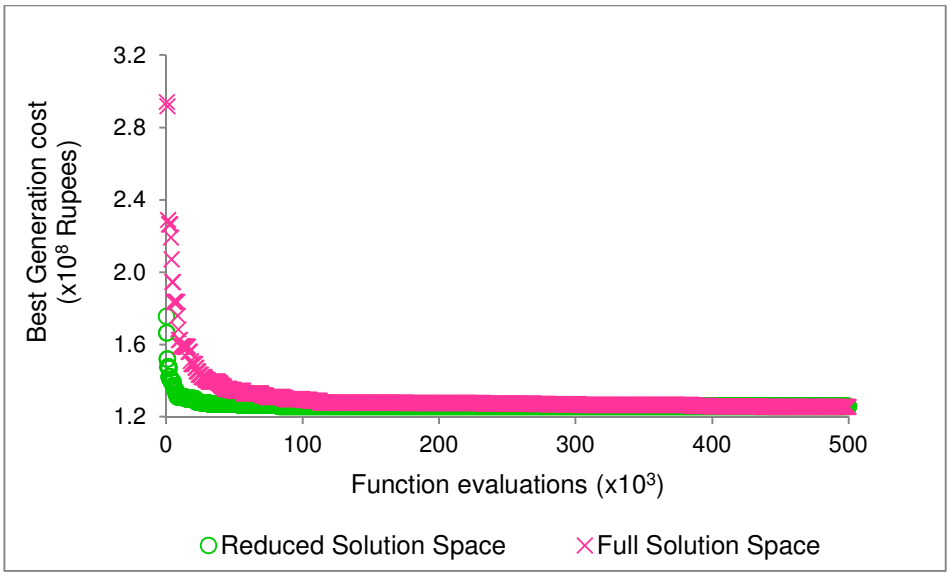
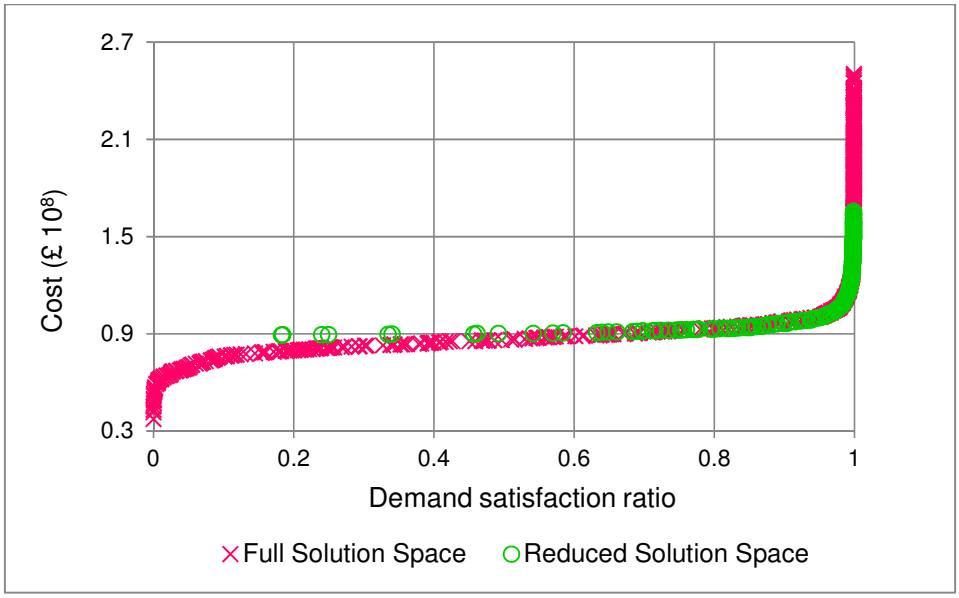


Figure 6.3 Pipe velocities of the best PF-MOEA designs for Network 1



(a) Evolution of the cost of the cheapest feasible solutions



(b) Pareto-optimal fronts

Figure 6.4 Network 1: Progress graphs and Pareto-optimal fronts for the best runs

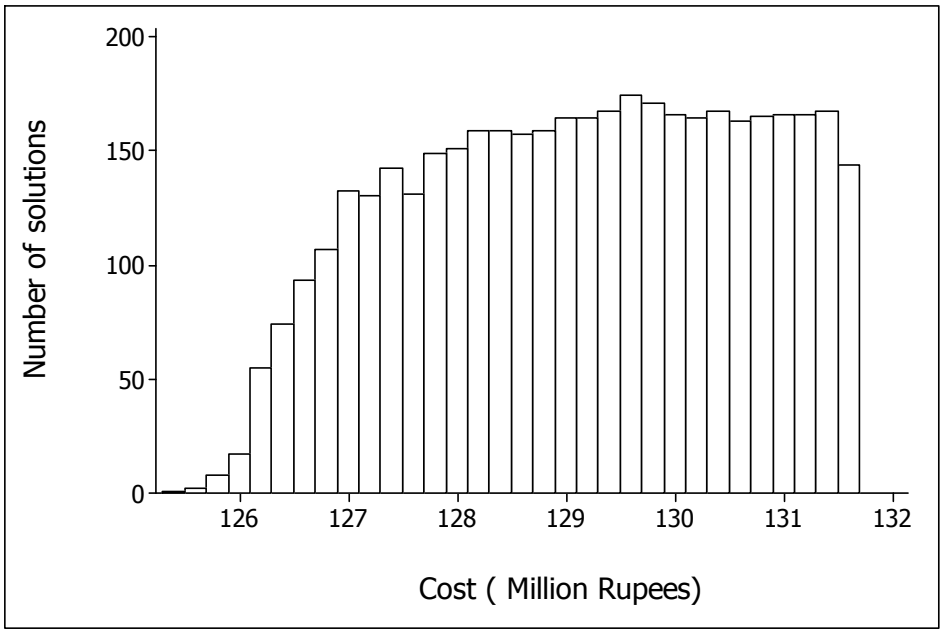


Figure 6.5 Distribution of the individual solutions found (4135 in total) in 705 optimisation runs that are feasible and cheaper than the previous best solution

6.4.2 Network 2

Network 2 is a water supply zone of a network in the UK that was described in Chapter 4 and 5. The network consists of 251 pipes of various lengths, 228 demand nodes (including the fire hydrants), 5 variable-head supply nodes, 29 fire hydrants at various locations and 3 demand categories. The network layout and the demand categories are presented in Chapter 5. The minimum pressure requirement to be fulfilled was 20 m at all demand nodes. Also, the minimum residual pressure requirement at all fire hydrants (with a fire flow of 8 l/sec) was 3 m. A maximum velocity of 1 m/s is also additional requirement that needs to be satisfied. However, referring to British Standard for Water supply-Requirements for Systems and Components outside buildings (BS EN 805:2000) the 1 m/s maximum velocity requirement is too conservative. According to the standard, the velocity range between 0.5 m/s and 2 m/s may be considered appropriate; and in special circumstances (e.g. fire flow), velocities up to 3.5 m/s can be acceptable. It is worth noting here that velocity constraints were not implemented in PF-MOEA procedure. Nevertheless, a post optimisation analysis was carried out to assess the pipe velocities of the optimised designs based on the maximum velocity constraint of 3.5 m/s to account for the fire flow conditions based on British Standards.

Extended period simulation (EPS) was used to cover all the 29 different fire demands and the normal demands. It is worth mentioning here that for the purposes of comparison of results with the existing network, both the network and dynamic operational data that were taken from a calibrated EPANET model were used without making any changes in the original data. Accordingly, the EPS adopted covered a period of 31 hours based on a 1-hour hydraulic time step. At each hour of the EPS period, except at the first and last hours, one fire demand is operational. The Darcy–Weisbach pipe head-loss formula was used for the hydraulic analysis. The pipe roughness coefficients range from 0.01mm to 3mm.

Network 2 needed rehabilitation and upgrading. Rehabilitation of the network was previously carried out using genetic algorithm based commercial optimisation software (Optimizer™). In this chapter, the network was optimised as a new design using PF-MOEA and results were compared with the existing network. The network and pipe unit cost data were supplied by a water utility. The complete network data is presented in Appendix C (Fig. C-2.1 and Table C-2.1). The pipe unit cost data consists of 28 pipe sizes with their respective unit cost. With 28 candidate pipe sizes, the solution space is huge (28^{251}) where computer cannot handle. In this work, some of the historical pipe sizes that are not commercially available anymore were removed. By doing this, not only the solution space can be reduced but also comparison of results with other optimal solutions can be possible. This can also help to implement the optimised design. Accordingly, 10 commercially available pipe sizes were selected based on the existing network pipe diameters that range from 32 mm to 400 mm. This has provided a massive reduction in solution space. The commercial pipe sizes utilised along with their cost can be found in Appendix C (Table C-2.2). The 10 candidate pipe sizes provide 10^{251} feasible and infeasible solutions in total. A four-bit binary string was used to represent the discrete candidate pipe sizes. This provided 2^4 i.e. 16 four-bit combinations of which six were redundant. The redundant codes were allocated one each to the two smallest and two largest candidate pipe sizes; and one each to the two middle candidate pipe sizes. Alternative approaches for dealing with redundant code are available in the literature (e.g. Saleh and Tanyimboh, 2014). Since the network is composed of 251 pipes, a chromosome that has a 1004-bit binary string represents each design. The crossover and mutation probability used were 1 and 0.005 respectively for all the optimisation runs conducted on Network 2.

It will be recalled in Eq. 6.2 that the second objective function (F_2) of PF-MOEA is based on the *DSR* of the worst performing node as used in Network 1. Nevertheless, due to the complexity of Network 2, the objective function considered for optimising Network 2 is based on the *DSR* of the entire network (i.e. maximising the *DSR* of the

network). Network 2 has a solution space that is in big order of magnitude in comparison to Network 1. Also, the spatial and temporal variation of the demand in Network 2 is huge in addition to its size. It is, therefore, realistic to consider a performance measure that assesses the entire network. This can provide a baseline for future work to assess the network further in order to have a complete understanding of its performance.

Network 2 is a computationally intensive optimisation problem. It was thus solved using a high performance computer (HPC) sequentially. In addition, the parallel algorithm was employed to solve the network.

6.4.2.1 Sequential Computing Using a High Performance Computer

In total, 30 PF-MOEA runs were executed in serial mode using the HPC facility at the University of Strathclyde. In serial (sequential) computing, independent optimization runs are performed in different processors concurrently. The high performance computing facility has 276 compute nodes; each has dual Intel Xeon 2.66 GHz CPU (six cores each) and 48 GB RAM running Linux operating system. The 20 optimization runs were performed using a population size of 200 and PF-MOEA was allowed to progress through 2500 generations i.e. a maximum of 500,000 function evaluations. The remaining ten runs were carried out based on a population size of 1000 and the algorithm was permitted to progress through 1000 generations i.e. a maximum of 1,000,000 function evaluations. The initial populations were generated randomly. Summarized results are shown in Table 6.5. The *minimum-cost* solution obtained was £419,514 within 985,000 function evaluations for population size of 1000 and £419,900 within 499,000 function evaluations for population size of 200. The average, median and maximum value of the minimum cost for the 20 optimization runs were £439,311, £436,129, £478,356 respectively. The corresponding values for the 10 optimization runs

were £421,938, £420,408 and £432,643 respectively. The standard deviation and coefficient of variation of the minimum cost for the 20 runs were 15,074 and 0.034 respectively while the respective values for the 10 runs were 4,038 and 0.01. A smaller coefficient of variation in both sets of runs demonstrates the consistency of results.

On average, the number of function evaluations and the CPU time to achieve convergence within the specified maximum of 500,000 and 1,000,000 function evaluations were (493,190, 6.7 hours) and (981,000, 12.81 hours) respectively. To complete a single optimisation run consisting of 500,000 function evaluations, the average CPU time required was 6.7 hours and the standard deviation was 0.4 hours. The average CPU time required to complete a single optimisation run consisting of 1,000,000 function evaluations was 13.17 hours and the standard deviation was 0.94 hours. It is worth emphasising that a *single* optimization run, with 500,000 function evaluations allowed, which takes approximately 15 days on a workstation (with two quad-core 2.4 GHz CPU and 16 GB RAM) was performed in less than seven hours using HPC facility. Also, *a single* optimization run, with 1,000,000 function evaluations allowed, which takes about 30 days on the workstation was performed in about 13 hours using the HPC. As can be seen in Table 6.5, the deviations of the *smallest* and *average* minimum-costs based on population size of 200 are 0.09% and 4.72% respectively from the *minimum-cost* (£419,514) obtained using population size of 1000. Results in general shows that population size of 1000 provided a better result. Nevertheless, in a situation where time is limited or when quick design is needed, population size of 200 (maximum allowed function evaluations of 500,000) may be appropriate to use.

The Pareto-optimal fronts of the 20 runs (population size = 200) were combined from which the final set of non-dominated solutions (199 solutions in total) was selected. In a similar manner, the final set of non-dominated solutions (1083 solutions in total) was selected by combining the Pareto–optimal fronts of the 10 runs (population size = 1000). Fig. 6.6. compares the final set of the non-dominated solutions obtained from the two

sets of runs. The similarity of the fronts illustrated the consistent performance of the algorithm in different independent runs and population sizes. The non-dominated fronts of the individual run from the two sets of optimisations runs can be found in Appendix C (Fig. C-2.2 and C-2.3). The small distances among the fronts re-confirm the reliability of the algorithm.

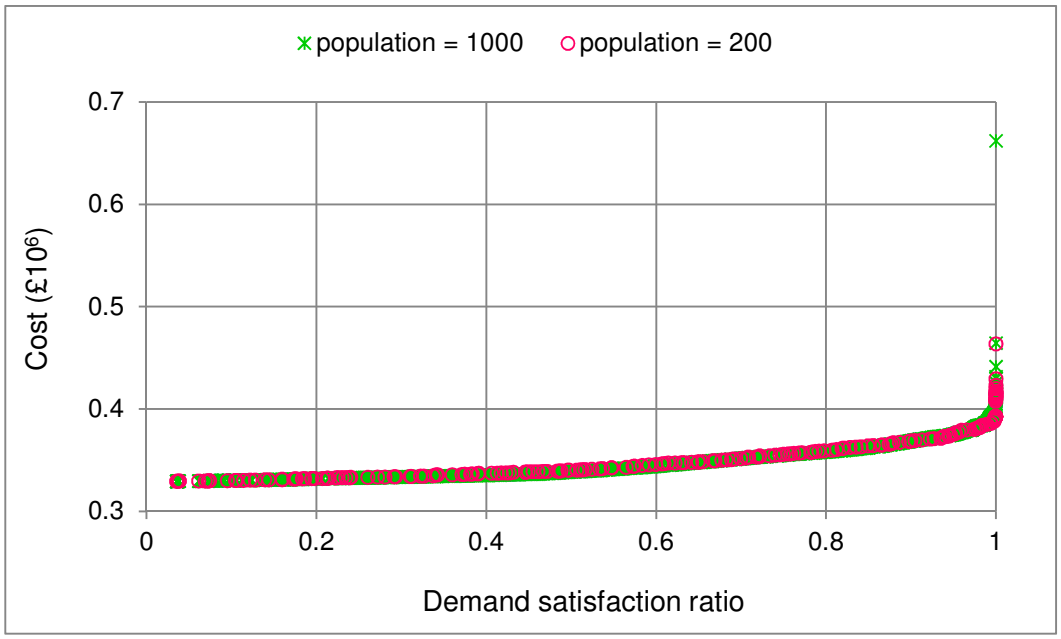


Figure 6.6 Non-dominated solutions from the union of 20 optimisation runs and from the union of 10 optimisation runs for Network 2

Table 6.5 Computational characteristics of PF-MOEA for Network 2

Mutation rate	0.005		(% deviation in cost from the minimum cost solution of £419,514)*	
Maximum function evaluations allowed	500,000	1,000,000		
Number of GA runs	20	10		
Population size	200	1000	200	1000
Cost (£)				
Minimum	419,900	419,514	0.09	0.00
Maximum	478,356	432,643	14.03	3.13
Mean	439,311	421,938	4.72	0.58
Median	436,129	420,408	3.96	0.21
Standard deviation	15,074	4,038		
Coefficient of variation (CV)	0.034	0.01		
Number of function evaluations (FEs) to achieve convergence				
Minimum	476,209	960,000		
Maximum	500,000	994,000		
Mean	493,190	981,000		
Median	497,000	982,000		
Standard deviation	7,544	9,557		
Coefficient of variation (CV)	0.015	0.01		
CPU time to achieve convergence (hours)				
Minimum	5.2 (5.2)**	11.20 (11.37)		
Maximum	7.2 (7.3)	13.43 (13.72)		
Mean	6.7 (6.7)	12.81 (13.17)		
Median	6.7 (6.8)	13.13 (13.58)		
Standard deviation	0.4 (0.4)	0.82 (0.94)		
Coefficient of variation (CV)	0.058 (0.059)	0.064 (0.072)		

* The percentage deviation in cost of solutions obtained using population sizes of 200 and 1000 from the minimum cost solution of £419,514
** values in parentheses show the CPU time required to complete a single optimisation run

Fig. 6.7 shows the evolution of the cost of the cheapest feasible solution obtained from the optimisation runs for population sizes of 200 and 1000. A rapid cost reduction from £1,682,340 at the start of the optimisation to (£798,653 and 25,200 function evaluations) was achieved for the best solution based on population size of 200. The algorithm converged at (£419,900 and 499,000 function evaluations). In a similar observation, a cost reduction from £1,641,210 at the start of the optimisation to (£792,874 and 34,000 function evaluations) was obtained for the best solution based on population size of 1000. The algorithm converged at (£419,514 and 985,000 function evaluations). The progresses of the minimum-cost solutions of the individual optimisation runs based on population sizes of 200 and 1000 are presented in Appendix C (Fig. C-2.4 and C-2.5). It is worth noting the similarity of the progress graphs.

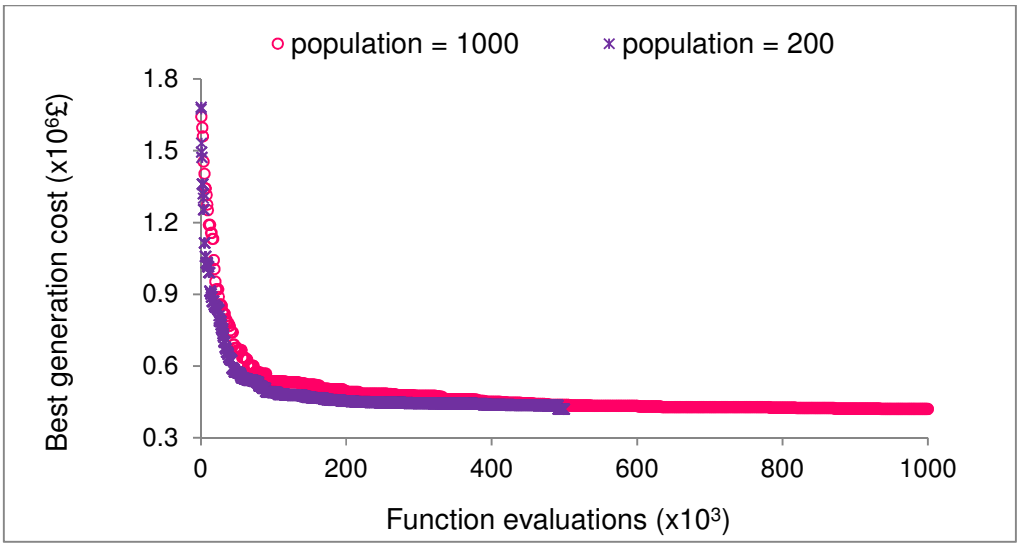


Figure 6.7 Progress of the best optimisation runs for Network 2. The maximum numbers of function evaluations allowed per run were 500,000 (population size = 200) and 1,000,000 (population size = 1000)

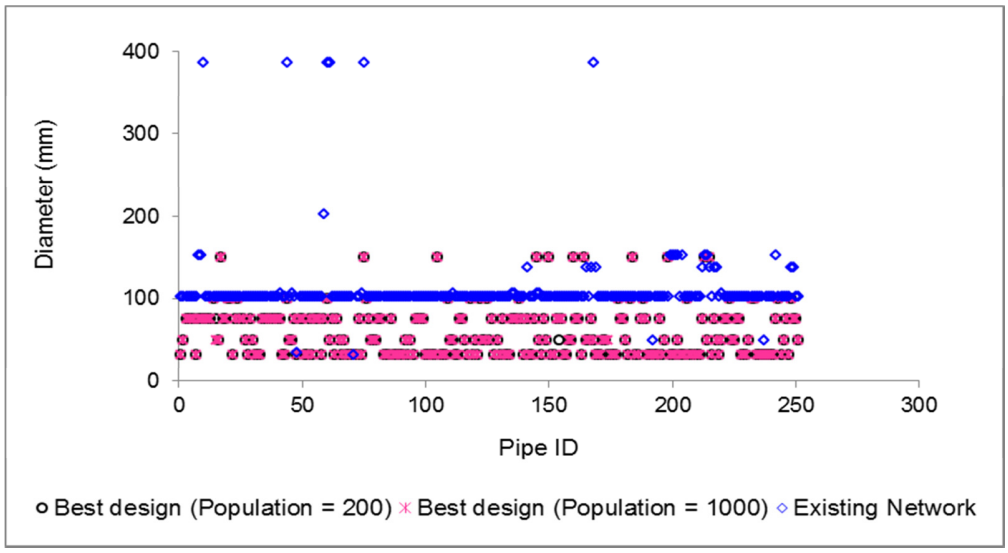
The optimisation results have been evaluated with reference to the existing network cost (£809,700). On average, 45.7% cost reduction was achieved for population size of 200 and 47.9% for population size of 1000. Also, comparison of the cheapest feasible

solutions with the existing network design shows a cost reduction of 48.1% and 48.2% for population size of 200 and 1000 respectively. The cheapest solutions were also simulated using EPANET 2 to confirm their feasibility. The pipe diameters, pipe velocities and nodal heads of the existing network and the optimised designs are compared in Fig. 6.8. It can be observed in Fig 6.8(a) that the PF-MOEA solutions in general consist of smaller pipe sizes compared to the existing network. Conversely, Fig. 6.8(b) shows that in general the PF-MOEA solutions have lower residual pressures than the existing network. It may be noted that the existing network has some pipe sizes that are not commercially available any more. Also, the minimum pressure requirements at all demand nodes and fire hydrants were fulfilled for the entire operating cycle. Fig. 6.8(b) shows that the pressures at all demand nodes including fire hydrants for all time steps of EPS are above 20 m. It is worth emphasizing that the pressures at fire hydrants were not close to 3 m (the minimum pressure requirement) due to the proximity of the fire hydrants to the demand nodes. The pipe velocities of the optimised designs and the existing network were evaluated in reference to the maximum velocity constraint of 3.5 m/s. The maximum velocity in the existing network is 1.1 m/s. However, in the optimised designs, 28 pipes have violated the maximum velocity constraint. It was noted that the velocity violation occurred in short pieces of pipes (about 1.0 m length each) located at the network junctions. Although fittings details of Network 2 were not provided by the water utility, these short pieces of pipes might be installed for connecting fittings with the distribution pipes. The *average* velocity violation in the network was calculated as follows.

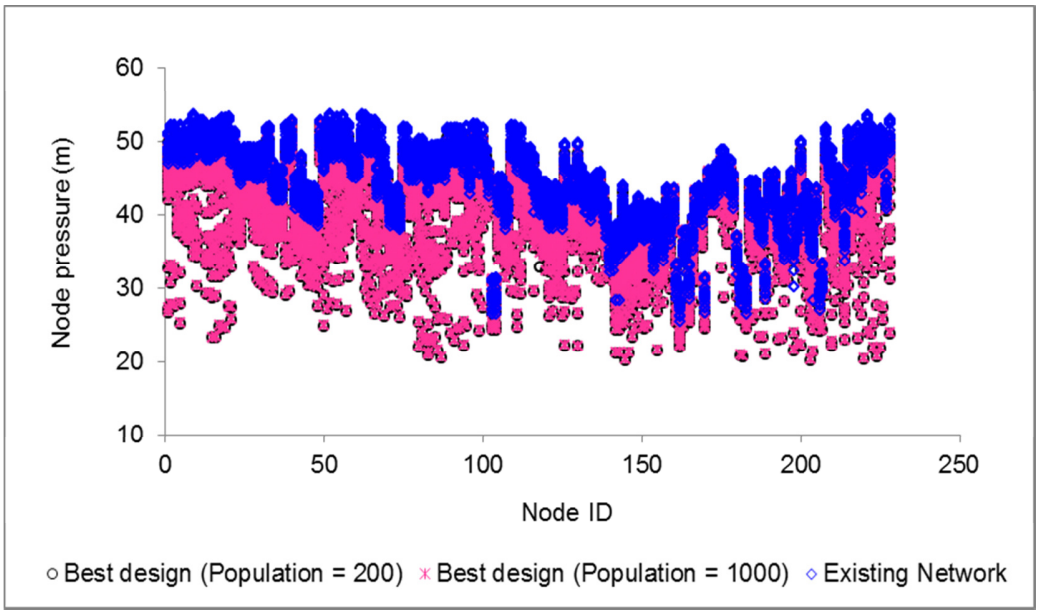
$$Velocity\ violation = \frac{\sum (v - v_{\max}; v > v_{\max}) * duration\ of\ exceedance}{Total\ duration\ of\ exceedance} \tag{6.4}$$

where v is pipe velocity and $v_{\max}$ is the maximum permissible velocity. Duration of exceedance refers to the number of times a pipe exceeded the maximum velocity limit in the extended period simulation. For the whole system over 31-hour operation cycle, the

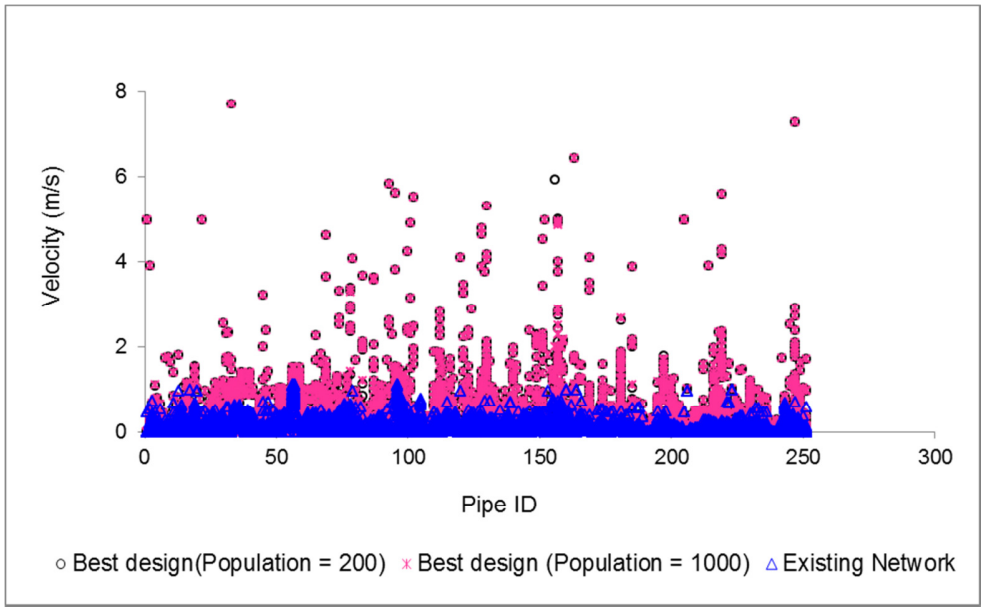
average velocity violation for the optimised designs were only 1.1 m/s (population size = 1000) and 1.2 m/s (population size = 200) by giving each pipe the same importance in terms of length. These results seem realistic given that the pipe length is not taken into effect. However, velocity constraint handling is an area for future work.



(a) Existing and optimised pipe diameters



(b) Nodal pressures for all time steps of the extended period simulation



(c) Pipe velocities for all time steps of the extended period simulation

Figure 6.8 Network 2: Pipe diameters, nodal pressures and pipe velocities of the existing network and optimised designs (serial runs)

6.4.2.2 Parallel Computing

The parallelized version of PF-MOEA has been applied to Network 2 using a workstation that comprises dual Intel Xeon 2.4 GHz CPU (four cores each) and 16 GB RAM running Windows operating system. The workstation has eight cores in total and all were utilized for parallel computing. Ten randomly generated optimization runs were carried out. All runs have a population size of 1000 given that for the serial algorithm population size of 1000 provided better result in comparison to population size of 200. The maximum numbers of function evaluations allowed per run were 1,000,000. The performance of the parallelized algorithm is evaluated in terms of computational efficiency and quality of solution with reference to the serial PF-MOEA results (for population size of 1000). The computational efficiency was assessed based on CPU time

(process time) required to complete a single optimization run (maximum allowed function evaluations = 1,000,000) using the serial PF-MOEA on the same workstation. Also, the quality of the solution of the parallel algorithm was evaluated in terms of the serial PF-MOEA results that was discussed in Section 6.4.2.1 based on 10 random runs. It is worth mentioning here that each pair of optimization runs (serial and parallel) started from the same initial population. Also, for proper comparison of the serial and parallel algorithms, the same genetic parameters (i.e. crossover and mutation rates and population size) were used.

Table 6.6 shows the comparison of the *minimum–cost* solutions and the number of function evaluations for convergence for the parallel and serial PF-MOEA algorithm. The *smallest* minimum-cost obtained from the parallel algorithm is £418,685 within 975,000 function evaluations. This is the cheapest solution obtained for Network 2. The solution is 0.2% cheaper than the *smallest* minimum-cost that was obtained from the serial algorithm (£419,514 and 985,000 function evaluations). Also, the *means* of the minimum-cost of the parallel algorithm (£425,334) differ from the *means* of the minimum-cost of the serial algorithm (£421,938) by only 0.8%. The average number of function evaluations to obtain convergence (within the specified maximum of 1,000,000 function evaluations) is 981,000 and 973,700 for serial and parallel algorithms respectively. The consistency of the results of the parallel algorithm has been demonstrated in Table 6.6 in which the values of the coefficient of variations of the cost and function evaluations (0.024 and 0.014 respectively) are small and comparable with the serial algorithm (0.01 and 0.01 respectively).

The consistent performance of the parallel algorithm is described in Fig. 6.9 in which the non-dominated fronts of the 10 random parallel runs are similar with the corresponding serial runs. Fig. 6.10 provides a graphical comparison on the progress of the cost of the serial and parallel runs over a generation. It is worth observing the similarity of the graphs for all the optimisation runs.

Table 6.6 Serial and parallel PF-MOEA results for Network 2

Maximum number of function evaluation allowed			1,000,000	
Number of optimisation runs			10 serial and 10 parallel runs	
Population size			1000	
Mutation rate			0.005	
	Cost (£x10^6)		Number of function evaluations to achieve convergence	
	Parallel	Serial	Serial	Parallel
Minimum	418,685	419,514	960,000	951,000
Maximum	453,643	432,643	994,000	998,000
Mean	425,334	421,938	981,000	973,700
Median	422,265	420,408	982,000	972,000
Standard deviation	10,139	4,038	9,557	13,849
Coefficient of variation	0.024	0.01	0.01	0.014

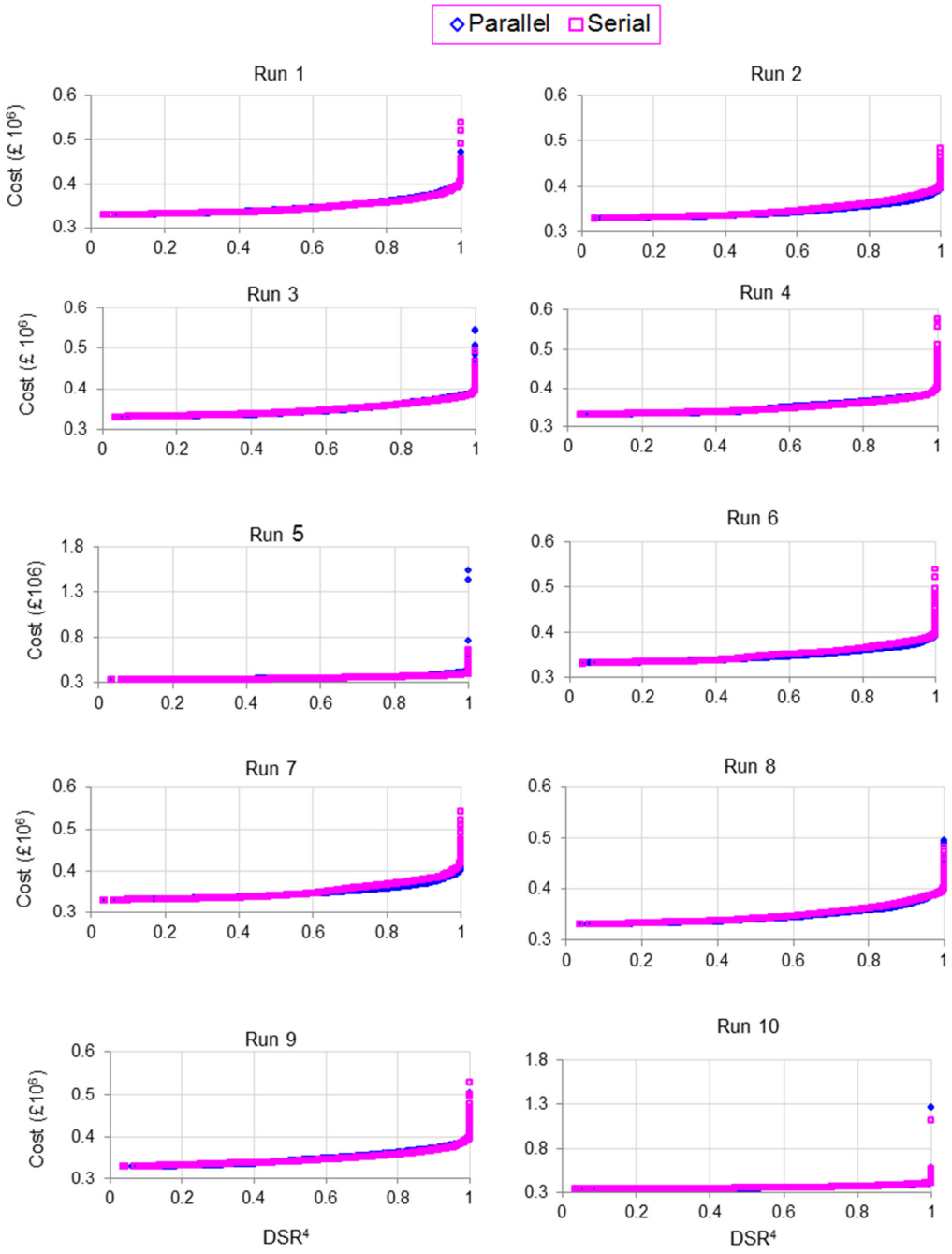


Figure 6.9 Consistency of the non-dominated fronts from the parallel and serial runs for Network 2

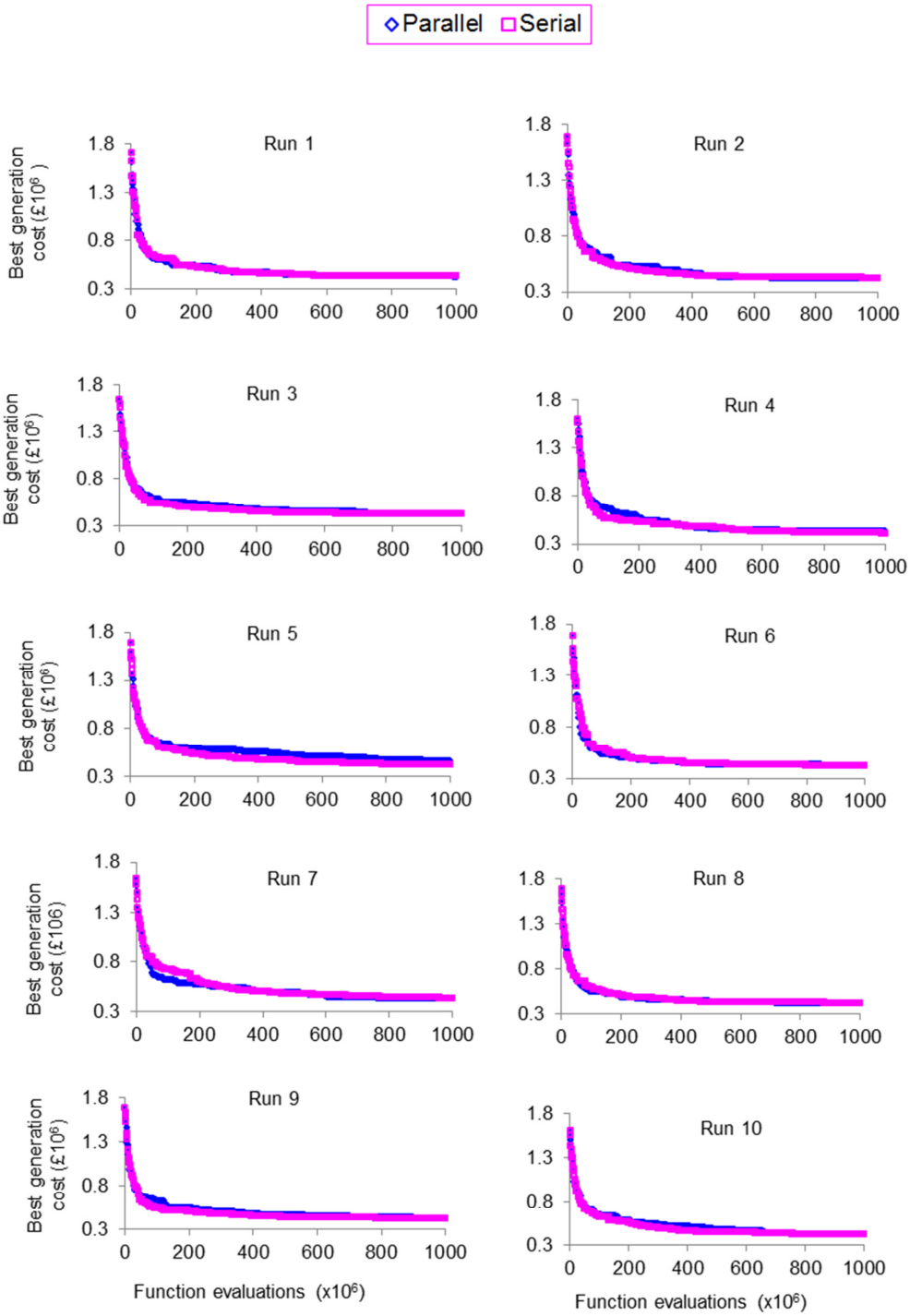


Figure 6.10 Progress of the cost of the serial and parallel runs. Each pair of runs started from the same initial population

The computational efficiency of the parallel algorithm is evaluated in terms of the ratio of the computational time the serial algorithm takes to complete a single run by using a single processor of the workstation to the computational time the parallel algorithm required to complete a single run by using all the eight processors of the workstation concurrently. To complete a single optimization run of 1,000,000 function evaluations, the serial PF-MOEA required an average CPU time of 30 days while the parallel algorithm requires *only* 2 days. The use of parallel algorithm, therefore, leads to a considerable improvement in execution time. The speedup that was achieved from the ten parallel runs ranges from 10.96 to 17.15 as shown in Fig. 6.11. Speedup refers to the ratio of serial algorithm execution time to that of the parallel algorithm to complete a single optimization run. On average, the parallel algorithm has achieved a speedup of 15. In another word, the average parallel computing time is equal to 1/15 of the serial computing time.

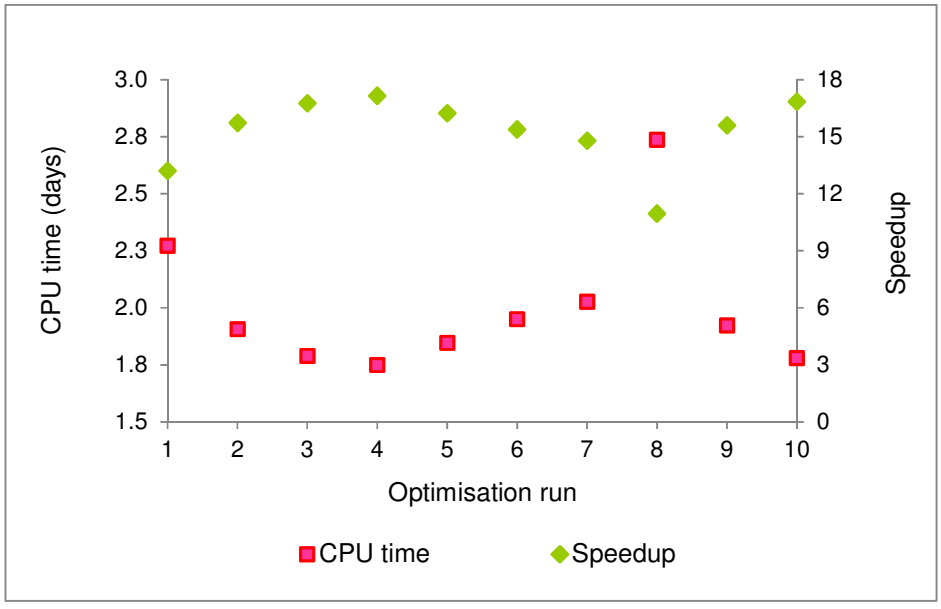
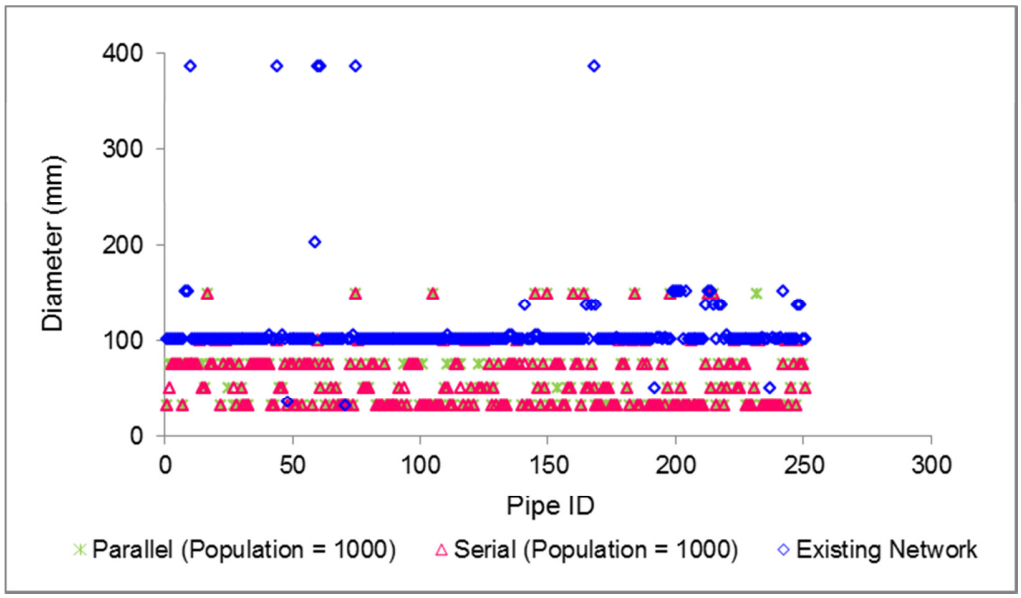
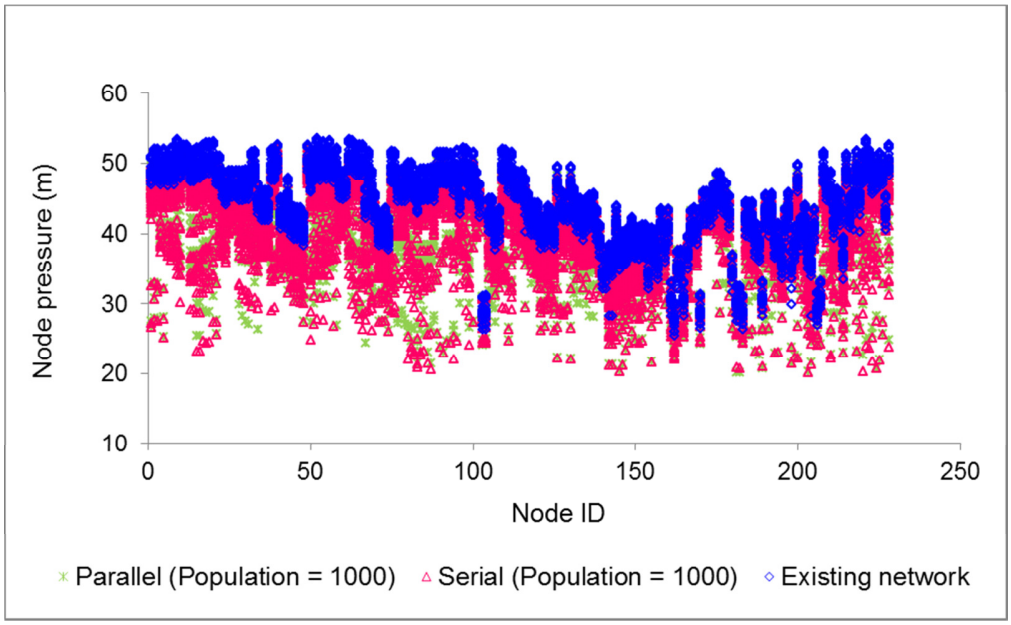


Figure 6.11 CPU time required to complete a single run using parallel PF-MOEA and the speedup achieved from each run for Network 2

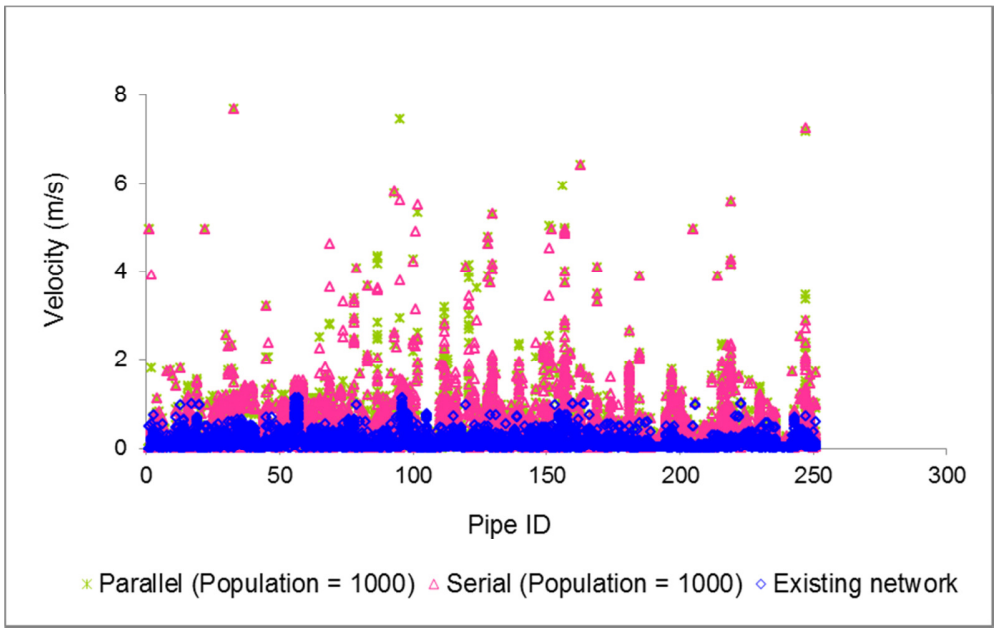
Also, in Fig. 6.11 the computational time the parallel algorithm required to complete a single run ranges from 1.75 to 2.74 days. Overall, the parallel result is encouraging and more improvement in computational efficiency can be achieved by implementing the algorithm on HPC facilities. Fig 6.12 compares pipe diameters, nodal pressures and pipe velocities of the best design of the parallel algorithm (£418,685) with the best design of the serial algorithm (£419,514) and the existing network. The best design of the parallel run is fully feasible and does not violate the minimum pressure requirement at all nodes for all operating conditions. It provides smaller pipe sizes in comparison to the existing network. The maximum velocity constraint was violated in 27 pipes; and the average velocity violation for the 31-hour operation cycle is 1.2 m/s.



(a) Existing and optimised pipe diameters



(b) Nodal pressures for all time steps of the extended period simulation



(c) Pipe velocities for all time steps of the extended period simulation

Figure 6.12 Network 2: Pipe diameters, nodal pressures and pipe velocities of the existing network and optimised designs (serial and parallel runs)

6.4.2.3 Performance Assessment of PF-MOEA for Convergence to the Pareto Front

In total, 40 optimisation runs were performed for Network 2 as discussed earlier. In the previous sections, the PF-MOEA results were analysed statistically. The efficiency of the algorithm was assessed in terms of CPU time and number of function evaluations to achieve convergence. Also, the progress of PF-MOEA over generations has been analysed to understand the algorithm's behaviour well. In this section, the algorithm's convergence to the set of solutions close to the global Pareto front has been assessed. Generational distance that was proposed by Veldhuizen and Lamont (1998) was taken into account as a performance measure to quantify the distance between a given set of non-dominated solutions and the Pareto front. The generational distance (GD) is expressed as follows:

$$GD = \frac{1}{NS}\left(\sum_{i=1}^{NS} d_i^{\,2}\right)^{1/2} \tag{6.5}$$

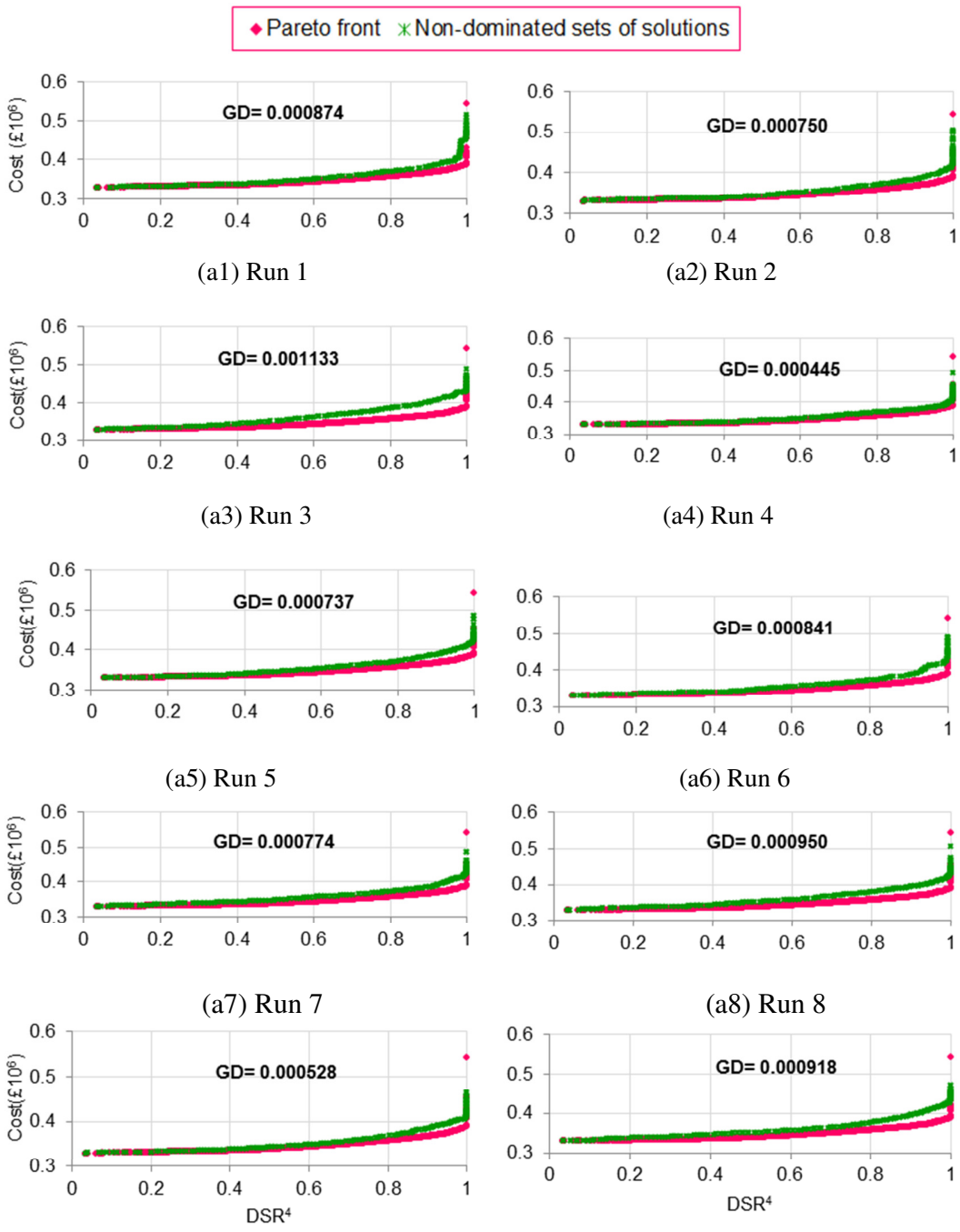
where NS = number of non-dominated solutions and d_i = is the distance in the objective space between the i^{th} non-dominated solution and the nearest solution of the optimal Pareto set. A value of $GD = 0$ indicates that all solutions are placed on the Pareto front. Any other value will indicate how far the solutions are from the Pareto front.

$$d_i = Min\left(\left[\sum_{m=1}^{2}(f_m^{\,(i)} - f_m^{(k)}\,)^2\right]^{1/2};\;\; k = 1,...,NP\right) \tag{6.6}$$

$f_m^{(i)}$ = value of the m^{th} objective function for the i^{th} non-dominated solution; $f_m^{(k)}$ = value of the m^{th} objective function for the k^{th} Pareto optimal solution; NP = number of solutions in the Pareto front.

The approximation to the Pareto front has been found by combining all the non-dominated solutions obtained from the 40 PF-MOEA runs (24,000 solutions). The final set of non-dominated solutions (989 solutions) was then selected as Pareto optimal solutions. The coordinates of the solutions in the objective space (i.e. cost and DSR) were normalised and have values between 0 and 1. Fig. 6.13 visually illustrates the PF-MOEA performance by plotting each of the non-dominated fronts of the 40 runs and the Pareto front (reference front). The first 20 fronts in Fig. 6.13 were based on population size of 200 (Fig. 6.13 (a1-a20)) while the other 20 fronts (Fig. 6.13 (b1-b10): parallel and Fig. 6.13 (b11-b20): serial) were based on population size of 1000. From the plots, it can be seen that the non-dominated fronts are close in distance to the Pareto front in general. In particular, the non-dominated fronts based on population size of 1000 Fig. 6.13 (b1-b20) provided smaller distances from the Pareto or reference front.

In addition to the visual inspection, the performance of PF-MOEA for convergence to the Pareto front has been evaluated in terms of GD in Fig. 6.13. The Pareto front comprises 989 non-dominated solutions (i.e. NP = 989). The number of non-dominated solutions in each of the 20 fronts in Fig 6.13 (a1-a20) is 200 (i.e. NS = 200). The remaining 20 fronts in Fig. 6.13 (b1-b20) consist of 1000 non-dominated solutions in the individual fronts (i.e. NS = 1000). The minimum, average, median and maximum values of GD for the 40 fronts were 0.000049, 0.000487, 0.000469 and 0.001623 respectively. It is worth noting here that the maximum GD between any two solutions in Fig. 6.13 would be $\sqrt{2}$. Overall results show that PF-MOEA provided consistently small GDs that demonstrate how close the algorithm has come to the Pareto front. Also, uniform distribution of solutions along the Pareto front has been obtained.



(a1) Run 1

(a2) Run 2

(a3) Run 3

(a4) Run 4

(a5) Run 5

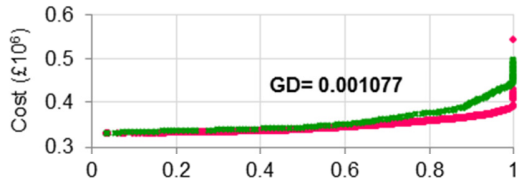
(a6) Run 6

(a7) Run 7

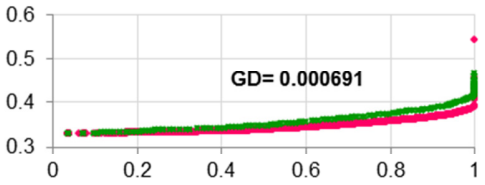
(a8) Run 8

(a9) Run 9

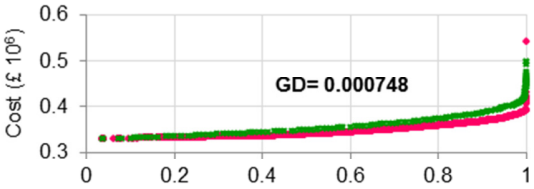
(a10) Run 10



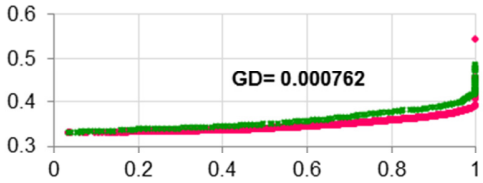
(a11) Run 11



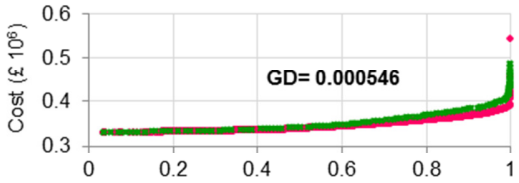
(a12) Run 12



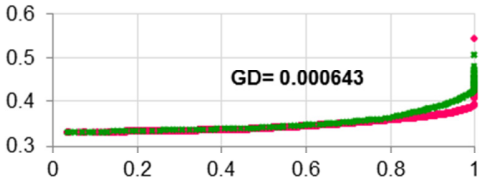
(a13) Run1 3



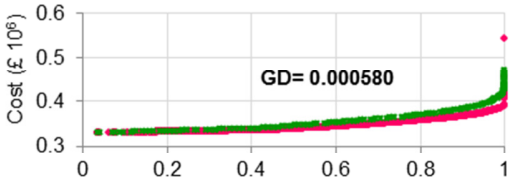
(a14) Run 14



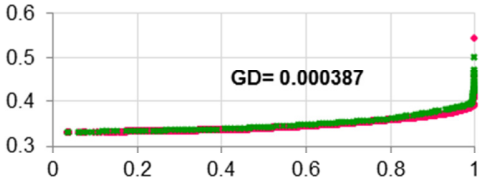
(a15) Run15



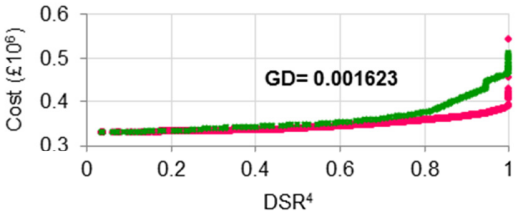
(a16) Run 16



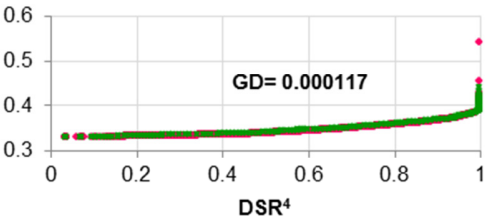
(a17) Run 17



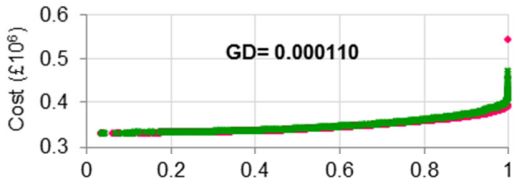
(a18) Run 18



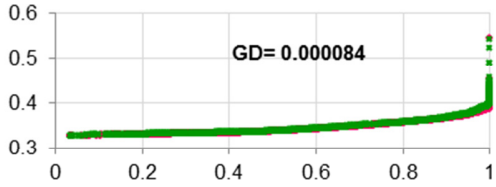
(a19) Run 19



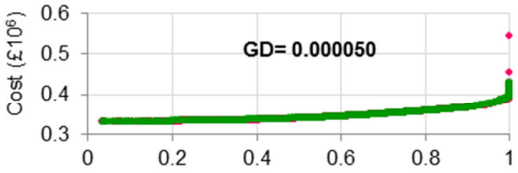
(a20) Run 20



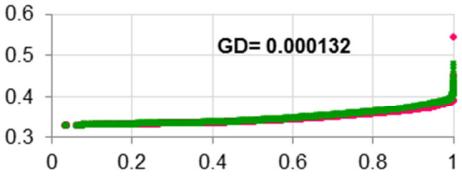
(b1) Run 21



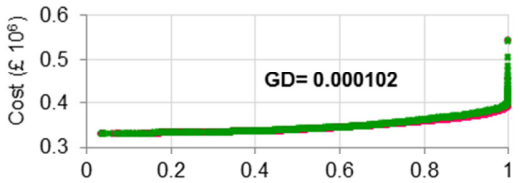
(b2) Run 22



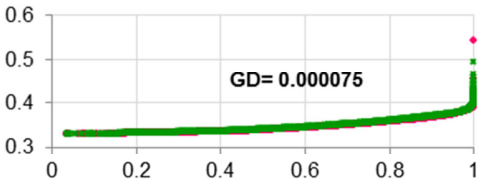
(b3) Run 23



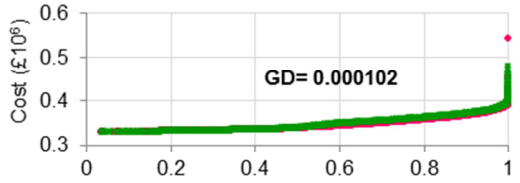
(b4) Run 24



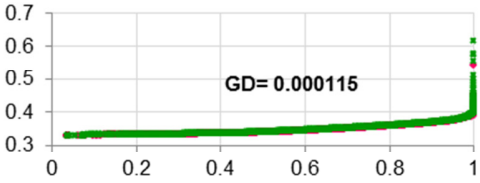
(b5) Run 25



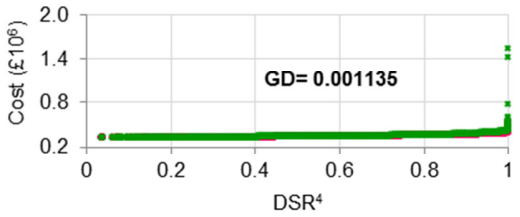
(b6) Run 26



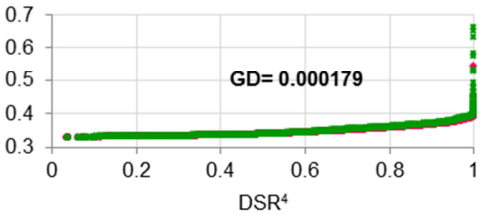
(b7) Run 27



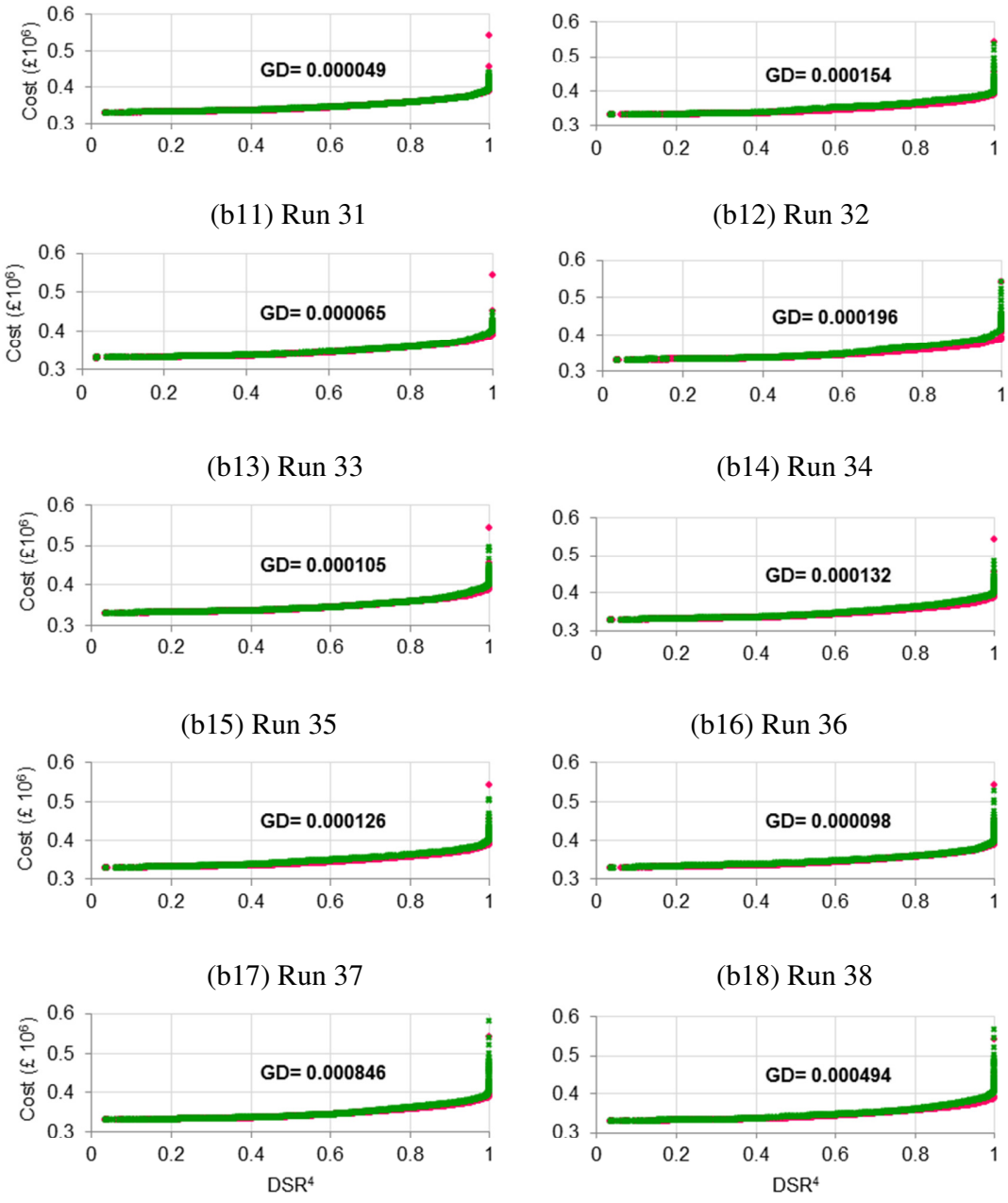
(b8) Run 28



(b9) Run 29



(b10) Run 30



(b11) Run 31 (b12) Run 32

(b13) Run 33 (b14) Run 34

(b15) Run 35 (b16) Run 36

(b17) Run 37 (b18) Run 38

(b19) Run 39 (b20) Run 40

Figure 6.13 Performance assessment of PF-MOEA for convergence to the Pareto front

6.5 Conclusions

This chapter assesses the penalty-free multi-objective evolutionary optimization approach for the design optimisation of water distribution systems. The approach uses pressure-dependent analysis that accounts for the pressure dependency of the nodal flows and obviates the need for penalties to address violations of the nodal pressure constraints. Results for two network examples show the algorithm is stable and finds optimal and near-optimal solutions reliably and efficiently. The results also suggest that the evolutionary sampling efficiency is very high. In other words, the number of solutions evolved and analysed on average before finding a near-optimal solution is small in comparison to the total number of feasible and infeasible solutions. Only one solution in every 10^{33} solutions for Network 1 and in every 10^{245} solutions for Network 2 was assessed.

In total, 705 optimization runs were executed for Network 1 for which there were 352.5 million hydraulic simulations. Thousands of solutions that are both fully feasible and cheaper than the best-known solutions in the literature were found. The cheapest solution obtained was 4.96% cheaper than the least cost solution published in literature (Kadu et al. 2008). Results for the reduced solution space demonstrated a significant reduction in the number of function evaluations needed to find optimal solutions. This strongly suggests a need for further research to develop an efficient solution space reduction technique that is capable of selecting appropriate candidate pipe sizes dynamically. It is encouraging, also, that the algorithm seems reasonably stable with respect to the mutation rate. This suggests that fine-tuning of the mutation rate may not be essential.

The optimisation algorithm PF-MOEA performed well for the real-life optimisation problem (Network 2) that involves multiple supply sources, multiple demand categories

and extended period simulations. The algorithm provides a least cost design that satisfies the pressure requirements. The least cost design obtained was significantly lower in cost compared to the existing design (i.e. *48%* reduction approximately). In total 30 million *extended period simulations* of the network were carried out in PF-MOEA (using the serial and parallel versions). In all cases, the algorithm performs reliably well. The pressure dependent analysis algorithm EPANET-PDX (Siew and Tanyimboh, 2012a) that is embedded in PF-MOEA performed reliably well also.

Also, the developed parallel algorithm based on a controller-worker approach has been successful. With respect to computational time, the algorithm consistently outperforms the serial algorithm significantly. The average *speedup* achieved was *15* on an eight-core workstation. In terms of quality of solution, the performance of the parallel algorithm is very satisfactory in which comparable results were obtained in reference to the serial algorithm. Overall, the parallel results are encouraging and further studies on other parallel implementation approaches (e.g. Island model) can be beneficial to evaluate the improvement on the computational performance.

In PF-MOEA, only node pressure constraints were explicitly considered. Given that the algorithm is efficient and robust in finding optimal/near optimal solutions, it would be beneficial to address other constraints (e.g. maximum and minimum velocities, water quality etc.) to widen the algorithm's application in real world optimisation problem.

Chapter 7

Conclusions and Recommendations for Future Work

7.1 Introduction

Water distribution systems (WDSs) are an invaluable component of urban infrastructure. Thus, the systems need to be designed and managed in a cost effective way while ensuring that the required performance and regulatory standards are satisfied. In WDSs optimisation, simulation models are usually coupled to evolutionary algorithms (EAs) to evaluate the hydraulic and water quality performances of solutions to the problem. However, EAs by nature are stochastic and generate a large number of infeasible designs, which are pressure-deficient. Previous studies on constraint or infeasible solutions handling approaches demonstrated the benefits of explicitly maintaining infeasible solutions (pressure deficient designs) in full. An accurate network performance assessment of infeasible designs is therefore extremely essential in order to guide the evolutionary search towards the optimal solution effectively and efficiently. Unfortunately, the conventional demand-driven analysis (DDA) method is invalid to evaluate the both the hydraulic and water quality performances of WDNs under pressure-deficient conditions. When predicting the behaviour of a pressure-deficient system, results produced by DDA are highly unreliable and misleading. The performance of WDNs can only be assessed accurately by considering demands to be

pressure dependent. This method of analysis that is called pressure dependent analysis (PDA) measures the performances of pressure-deficient networks more realistically.

Another issue of EAs in WDSs optimisation problem is that the algorithms require huge computational time when applied to large optimisation problems such as real-life networks with large numbers of pipes and multiple operating conditions. This has limited the algorithms' potential for practical applications to solve real-life WDSs optimisation problems.

In an attempt to address the aforementioned issues, the present research has proposed a robust PDA model that enhances greatly the computational properties for low-pressure conditions. The model has been successfully applied to analyse ranges of normal and pressure-deficient operating conditions for extended period simulations. The research has thoroughly investigated the water quality performance of WDNs under all pressure operating conditions. Also, a parallel multi-objective evolutionary algorithm has been proposed to solve computationally intensive WDSs design problems. The focus in this chapter is to provide an overall summary and general conclusion of the research carried out herein. This is followed by several recommendations for future works.

7.2 Summary of the Present Research

7.2.1 Pressure Dependent Network Analysis Model

A computationally efficient and robust PDA model has been proposed in the current research. The model is an enhanced version of the pressure-dependent extension of EPANET i.e., EPANET-PDX (Siew and Tanyimboh, 2012a). EPANET-PDX has an embedded logistic nodal head-flow function (Tanyimboh and Templeman, 2010) coupled with a line search and backtracking procedure. The line search and backtracking

is a deterministic approach that optimises the size of the Newton step iteratively to ensure both the head loss and flow continuity functions are improved progressively. The line search implementation in EPANET-PDX, however, deviates from the classical line search implementation in an attempt to exploit the excellent computational properties of EPANET 2. The approach intermingles the line search procedure with the global gradient algorithm (GGA) and very often the algorithm makes progress using the full Newton step size and avoids the line minimisation. An alternative implementation of the line search and backtracking procedure for integrating the logistic nodal head-flow function into the system of hydraulic equations is proposed in Chapter 4. This has increased the robustness of the algorithm by enhancing greatly the computational properties for low-pressure conditions and increasing the algorithm's consistency over a wider range of operating conditions. The new implementation has a clear separation between the line search procedure and GGA. It provided more scope for the line search procedure to change the course of the algorithm.

The proposed PDA model has been demonstrated to be efficient, robust and accurate in analysing ranges of normal and pressure-deficient operating conditions. Several extended period simulations were executed for a real-life network that comprises multiple supply sources and various demand categories. Comparison between results generated by the model and EPANET-PDX demonstrated that the two models produce exactly the same hydraulic results for both normal and pressure deficient conditions. From a numerical perspective, a significant reduction in numbers of iterations to complete a simulation has been obtained for all pressure operating conditions using the proposed model. Also, a considerable reduction in computational time has been achieved for extremely pressure deficient conditions in comparison to EPANET-PDX.

7.2.2 Water Quality Modelling of Pressure Deficient Networks

Water quality models are increasingly being routinely used to help ascertain the quality of water in drinking WDSs for design and operational management purposes. Conventional water quality models are demand driven and consequently do not incorporate the effects of any deficiency in pressure on the water quality throughout the distribution network. Chapter 5 has addressed water quality modelling of WDSs under pressure-deficient conditions. EPANET-PDX has been investigated and discussed with particular reference to water quality. The model is thought to have preserved the EPANET 2 modelling functionality in full. Thus, EPANET-PDX performs both hydraulic and water quality modelling under both normal and low-pressure conditions entirely seamlessly. This has been demonstrated in Chapter 5 by carrying out hydraulic and water quality analyses for a range of simulated operating conditions including normal and subnormal pressure and pipe closures based on two water supply zones in the UK and a network in literature. Temporal and spatial variations of water age, chlorine and trihalomethane (THM) concentrations were evaluated under various hydraulic conditions. Convergence difficulties or failures were not experienced with EPANET-PDX for the various cases considered. The accuracy of the model has been verified using a hydraulic consistency test (Ackley et al., 2001) for pressure deficient conditions, along with EPANET 2 and EPANET-MSX for normal operating conditions.

7.2.3 High Performance and Parallel Computing for Complex Water Distribution Systems Design Optimisation Problems

High performance and parallel computing techniques were utilised to solve computationally intensive WDSs design problems. In Chapter 6, a controller-worker parallel algorithm based on the penalty-free multi-objective evolutionary algorithm (Siew and Tanyimboh, 2012b) was proposed. The penalty-free multi-objective evolutionary algorithm (PF-MOEA) uses pressure-dependent analysis (EPANET-PDX) that accounts for the pressure dependency of the nodal flows and thus avoids the need

for penalties to address violations of the nodal pressure constraints. In the proposed controller-work parallel algorithm, a single controller processor executes the routine operation of the algorithm and employs the worker processors to carry out fitness evaluation.

Two network optimisation problems were used in Chapter 6 as case studies where their objectives were to obtain the cheapest possible combination of pipe sizes that satisfy all the system requirements. Extensive statistical investigation on the performance of PF-MOEA was conducted using the first network (Kadu et al., 2008). Thousands of solutions that are both fully feasible and cheaper than the best-known solutions in the literature were found. The second network is a real-life network that involves multiple supply sources, multiple demand categories and extended period simulations. The network was optimised both sequentially and in parallel. For the sequential computing, a high performance computer (HPC) was used to perform multiple serial optimisation runs concurrently. The parallel algorithm was applied to execute parallel optimisation runs. The effectiveness of the parallel algorithm was evaluated in reference to results from sequential computing. The least cost design obtained for the network was significantly lower in cost compared to the existing design (i.e. 48% reduction approximately). In total 30 million extended period simulations of the network were carried out using the serial and parallel versions of PF-MOEA. In all cases, PF-MOEA performed reliably well. The pressure dependent analysis algorithm EPANET-PDX that is embedded in PF-MOEA performed reliably well also. The developed parallel algorithm has been successful. With respect to computational time, the algorithm consistently outperforms the serial algorithm significantly. The average speedup achieved was 15 on an eight-core workstation. In terms of quality of solution, the performance of the parallel algorithm is very satisfactory in which comparable results were obtained in reference to the serial algorithm.

In conclusion,

1. The proposed PDA model has consistently performed well while simulating the full range of pressure operating conditions effectively. The new implementation of the line search and backtracking procedure has greatly enhanced the computational properties for low-flow conditions and increased the algorithm's consistency over a wider range of operating conditions.

2. It has been shown that operating conditions with subnormal pressures, if severe and protracted, can lead to spatial and temporal distributions of water age, concentrations of chlorine and disinfection by-products that are significantly different from operating conditions in which the pressure is satisfactory.

3. The present research has demonstrated that under conditions of low-pressure, the conventional demand driven modelling approach can provide misleading results that in turn can lead to inappropriate water quality policy decisions.

4. The application of PF-MOEA for the design optimisation of a real-life network and a network in literature demonstrated that the algorithm is stable and finds optimal/ near-optimal solutions reliably and efficiently. The developed parallel algorithm based on a controller-worker approach has been successful. In terms of computational time, the algorithm consistently outperforms the serial algorithm significantly while providing comparable quality of solutions.

7.3 Scope for Future Work

Various issues related to WDSs modelling and optimisations have been addressed in the present research. However, some aspects of the research require further investigations and enhancements. This section summarises potential areas for further research.

EPANET-PDX has been extensively used in this research to perform water quality analyses under normal and low-pressure conditions entirely seamlessly. However, EPANET's water quality model is single-species and does not allow simulation of multiple interacting species concurrently. The first order kinetic model depends on the concentration of only one reactant. This prevents the model from including other potential reactants. Hence, it is essential to improve the model in order to solve real-life water quality problems that involve multiple reactive substances. This can be achieved by integrating the PDA hydraulic simulator into EPANET-MSX (Shang et al., 2008) water quality solver.

The water quality modelling difficulties when low-flow velocities prevail due to excessive pressure reduction in WDSs has been illustrated in Chapter 5. Previous research indicated that advection-driven models such as EPANET 2 might yield inconsistent water quality results if dispersion is significant due to low flow velocities (Tzatchkov et al., 2002; Rossman et al., 1994). Low-flow velocities are a common occurrence in WDSs. These situations, therefore, require special care and may indicate the need for more improvements in the underlying (EPANET 2) water quality model in the context of PDA. This includes incorporation of dispersion in the water quality model and the collection of field data under conditions of low pressure, low flow rates and low velocities.

With respect to WDSs optimisation, PF-MOEA has proven to be efficient and robust in finding optimal/near optimal solutions. However, the algorithm needs to be improved further to widen its application in real-world optimisation problem. In its current form, the algorithm considers only node pressure constraints explicitly. Other important constraints such as maximum and minimum velocities and water quality were not addressed. Given that the integrated PDA model (EPANET-PDX) has both hydraulic and water quality modelling functionalities, the suggestion of implementing other important WDSs constraints seems to be highly feasible.

Research has shown the effectiveness of reducing the size of the search-space on the performance of the genetic algorithm. As illustrated in Chapter 6 limiting the number of candidate pipe sizes for each link results in a significant reduction in the number of function evaluations needed to find optimal solutions. These results are encouraging and suggest further research in developing an efficient search-space reduction technique that is capable of selecting appropriate candidate pipe sizes dynamically.

The parallel algorithm developed in this research based on the controller-worker approach has demonstrated to be computationally very efficient. The results obtained are encouraging and further studies on other parallel implementation approaches (e.g Island model) are beneficial to evaluate the improvement of the computational performance.

The proposed PDA model in the current research has proven to be efficient and robust as detailed in Chapter 4. Given that the model has achieved significant improvement in computational performances for extremely pressure-deficient conditions and the multiple use of PDA in PF-MOEA, integrating the proposed model into PF-MOEA is beneficial. Also, further study is needed to assess the performance of the integrated water quality model in the proposed PDA model.

References

Abdy Sayyed, M. A.H., Gupta, R., and Tanyimboh, T. T. (2015) Noniterative application of EPANET for pressure dependent modelling of water distribution systems. *Water Resources Management*, doi:10.1007/s11269-015-0992-0.

Ackley, J. R. L., Tanyimboh, T. T., Tahar, B. and Templeman, A. B. (2001) Head driven analysis of water distribution systems. In Water Software Systems: Theory and Applications (Ulanicki, B., Coulbeck, B. and Rance, J. (eds.)). Taunton, UK: Research Studies Press, 183–192.

Afshar, M. H. and Marino, M. A. (2007). A parameter-free self-adapting boundary genetic search for pipe network optimization. Computational Optimization and Applications, 37(1), 83-102.

Alba, E. (2005) *Parallel metaheuristics: a new class of algorithms.* (Vol. 47). John Wiley & Sons.

Alba, E. and Tomassini, M. (2002) Parallelism and evolutionary algorithms. *Evolutionary Computation, IEEE Transactions on*, 6(5), 443-462.

Alperovits, E. and Shamir, U. (1977) Design of optimal water distribution systems. *Water Resource Research*, 13(6), 885-900.

Ang, W. H. and Jowitt, P. W. (2006) Solution for water distribution systems under pressure-deficient conditions. *Journal of Water Resources Planning and Management*, 132(3), 175-182.

Artina S., Bragalli C., Erbacci G., Marchi A. and Ri-vi M. (2012) Contribution of parallel NSGA-II in optimal design of water distribution networks. *Journal of Hydroinformatics*, 14(2), 310-323.

Atiquzzaman, M., Liong, S. Y. and Yu, X. (2006) Alternative decision making in water distribution network with NSGA-II. *Journal of Water Resources Planning and Management*, 132(2), 122-126.

Back, T. (1994) Selective pressure in evolutionary algorithms: A characterization of

selection mechanisms. In *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence. Proceedings of the First IEEE Conference on* (57-62). IEEE.

Back, T., Fogel, D. and Michalewicz, Z. (1997) *Handbook of evolutionary computation*. Bristol, U.K: IOP Publishing Ltd. and Oxford University Press.

Back, T., Fogel, D. B. and Michalewicz, Z. (Eds.) (2000) *Evolutionary computation 1: Basic algorithms and operators.* (Vol. 1). CRC Press.

Balla, M. C. and Lingireddy, S. (2000) Distributed genetic algorithm model on network of personal computers. *Journal of Computing in Civil Engineering*, 14(3), 199-205.

Barlow, E. and Tanyimboh, T. T. (2014) Multi-objective memetic algorithm applied to the optimisation of water distribution systems. *Water Resources Management*, 28(8), 2229-2242.

Bekele, E. G. and Nicklow, J. W. (2005) Multiobjective management of ecosystem services by integrative watershed modeling and evolutionary algorithms. *Water Resources Research*, 41(10).

Bhave, P.R. (1991) *Analysis of Flow in Water Distribution Networks*. Lancaster, PA: Technomic Publishing.

Bhave, P.R. and Gupta, R. (2006) *Analysis of Water Distribution Networks*. UK: Alpha Science International Ltd.

British Standards Institution (2000) *Water supply. Requirements for systems and components outside buildings*, BS EN 805:2000.

Broad, D. R., Dandy, G. C. and Maier, H. R. (2005) Water distribution system optimization using metamodels. *Journal of Water Resources Planning and Management*, 131(3), 172-180.

Cantu-Paz, E. (2000) *Efficient and accurate parallel genetic algorithms* (Vol. 1). Springer Science & Business Media.

Cantu-Paz, E. and Goldberg, D. E. (2000) Efficient parallel genetic algorithms: theory and practice. *Computer methods in applied mechanics and engineering*, 186(2), 221-238.

Carrico, B. and Singer, P.C. (2009) Impact of booster chlorination on chlorine decay and THM production: simulated analysis. *Journal of Environmental Engineering,* 135(10), 928–935.

Caruana, R. A., Eshelman, L. J. and Schaffer, J. D. (1989) Representation and hidden bias II: Eliminating defining length bias in genetic search via shuffle crossover. In *Proceedings of the 11th international joint conference on Artificial intelligence-Volume 1* (750-755). Morgan Kaufmann Publishers Inc..

Caruana, R.A. and Schaffer, J.D. (1988) Representation and hidden bias: Gray vs. binary coding for genetic algorithms. In *Proceedings of the 5th International Conference on Machine Learning*, 153-161, Los Altos: CA, Morgan Kaufmann.

Castillo, O., Melin, P. and Pedrycz, W. (Eds.) (2008) *Soft computing for hybrid intelligent systems.* (Vol. 154). Springer Science & Business Media.

Chandapillai, J. (1991) Realistic simulation of water distribution system. *Journal of Transportation Engineering*, 117(2), 258-263.

Ciaponi, C., Franchioli, L., Murari, E. and Papiri, S. (2015) Procedure for defining a pressure-outflow relationship regarding indoor demands in pressure-driven analysis of water distribution networks. *Water Resources Management*, 29(3), 817-832.

Clark, R.M. and Grayman W.M. (1998). Modelling Water Quality in Drinking Water Distribution Systems. Denver, Colo. : AWWA

Clark, R.M. and Haught, R.C. (2005) Characterizing pipe wall demand: implications for water quality modelling. *Journal of Water Resources Planning and Management* 31(3), 208–217.

Cross, H (1936) Analysis of flow in networks of conduits or conductors. Bulletin No. 286, Univ. of Illinois Engineering, Experimental Station, Urbana, 111.

Cunha, M. C. and Sousa, J. (1999) Water distribution network design optimization: simulated annealing approach. *Journal of Water Resources Planning and Management*, 125(4), 215-221.

Cunha, M. C. and Ribeiro, L. (2004) Tabu search algorithms for water network optimization. *European Journal of Operational Research*, 157(3), 746-758.

Czajkowska, A. M. and Tanyimboh, T. T. (2013) Water distribution network optimization using maximum entropy under multiple loading patterns. *Water Science & Technology: Water Supply*, 13(5), 1265-1271.

Dandy, G. C., Simpson, A. R. and Murphy, L. J. (1996) An improved genetic algorithm for pipe network optimization. *Water Resources Research*, 32(2), 449-458.

Dandy, G.C. and Engelhardt, M.O. (2001) Optimal scheduling of water pipe replacement using genetic algorithms. *Journal of Water Resources Planning and Management*, 127(4), 214-223.

Deb, K. (2000) An efficient constraint handling method for genetic algorithms. *Computer methods in applied mechanics and engineering*, 186(2), 311-338.

Deb, K. (2001) *Multi-objective optimization using evolutionary algorithms* (Vol. 16). John Wiley & Sons.

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. A. M. T. (2002) A fast and elitist multi-objective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, 6(2), 182-197.

De Jong, K. (2005) Genetic algorithms: a 30 year perspective. In *Perspectives on Adaptation in Natural and Artificial Systems,* (Booker et al. (eds.)). Oxford University Press.

Digalakis, J. and Margaritis, K. (2003) Parallel evolutionary algorithms on MPI. In *Proceedings of the International Conference on Computer, Communication and Control Technologies*, CCCT'2003, Orlando, USA.

di Pierro, F., Khu, S. T., Savic, D., and Berardi, L. (2009) Efficient multi-objective optimal design of water distribution networks on a budget of simulations using hybrid algorithms. *Environmental Modelling & Software*, 24(2), 202-213.

Dong, N. and Wang, Y. (2014) A Memetic Differential Evolution Algorithm Based on Dynamic Preference for Constrained Optimization Problems. *Journal of Applied Mathematics*, *2014*.

Dridi, L., Parizeau, M., Mailhot, A. and Villeneuve, J.P. (2008) Using evolutionary optimization techniques for scheduling water pipe renewal considering a short planning horizon. *Computer-Aided Civil and Infrastructure Engineering*, 23(8), 625-635.

EC (European Community) (1998) Council Directive 98/83/EC on the quality of water intended for human consumption. *Official Journal of the European Communities* L330.

El-Mihoub, T. A., Hopgood, A. A., Nolle, L. and Battersby, A. (2006) Hybrid Genetic Algorithms: A Review. *Engineering Letters*, 13(2), 124-137.

Eusuff, M. M. and Lansey, K. E. (2003) Optimization of water distribution network design using the shuffled frog leaping algorithm. *Journal of Water Resources Planning and Management*, 129(3), 210-225.

Ewald, G., Kurek, W. and Brdys, M. A. (2008) Grid implementation of a parallel multiobjective genetic algorithm for optimized allocation of chlorination stations in drinking water distribution systems: Chojnice case study. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(4), 497-509.

Farmani, R., Savic, D. A. and Walters, G. A. (2003) Multi-objective optimization of water system: a comparative study. *Cabrera et al.(eds.) Pumps, Electromechanical Devices and Systems Applied to Urban Water Management*, vol.1, 247-256.

Farmani, R., Wright, J. A., Savic, D. A. and Walters, G. A. (2005) Self-adaptive fitness formulation for evolutionary constrained optimization of water systems. *Journal of Computing in Civil Engineering*, 19(2), 212-216.

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

Fonseca, C. M. and Fleming, P. J. (1993) Genetic algorithms for multiobjective optimization: formulation, discussion and generalization. In *Proceedings of the International Conference on Genetic Algorithms*, vol. 5, 416–423, Morgan Kaufmann, Los Altos, CA.

Fujiwara, O. and Khang, D. B. (1990) A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resource Research*, 26(4), 539-549.

Fujiwara, O. and Ganesharajah, T. (1993) Reliability assessment of water supply systems with storage and distribution networks. *Water Resources Research*, 29(8), 2917-2924.

Geem, Z. W. (2006) Optimal cost design of water distribution networks using harmony search. *Engineering Optimization*, 38(03), 259-277.

Gen, M., Cheng, R. and Lin, L. (2008) *Network models and optimization: Multiobjective genetic algorithm approach*. Springer Science & Business Media.

Germanopoulos, G. (1985) A technical note on the inclusion of pressure dependent demand and leakage terms in water supply network models. *Civil Engineering Systems*, 2(3), 171-179.

Germanopoulos, G., Jowitt, P.W. and Lumbers, J.P. (1986) Assessing the reliability

of supply and level of service for water distribution system. *Proceedings of the Institution of Civil Engineers Part 1,* 80(2), 413–428.

Germanopoulos, G. and Jowitt, P.W. (1989) Leakage reduction by excess pressure minimization in a water supply network. *Proceedings of the Institution of Civil Engineers Part 2*, 87(2), 195-214.

Ghebremichael, K., Gebremeskel, A., Trifunovic, N. and Amy, G. (2008) Modelling disinfection by-Products: coupling hydraulic and chemical models. *Water Science and Technology: Water Supply,* 8(3), 289-295.

Giustolisi, O., Kapelan, Z. and Savic, D. A. (2008a) Extended period simulation analysis considering valve shutdowns. *Journal of Water Resource Planning and Management*, 134(6), 527-537.

Giustolisi, O., Savic, D. A., and Kapelan, Z. (2008b) Pressure-driven demand and leakage simulation for water distribution networks. *Journal of Water Resource Planning and Management*, 134(5), 626-635.

Goldberg, D. E. and Kuo, C. H. (1987) Genetic algorithms in pipeline optimization. *Journal of Computing in Civil Engineering*, 1(2), 128-141.

Goldberg, D. E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning.* Boston, MA: Addison-Wesley Longman Publishing Co., Inc.,

Goldberg, D. E. and Deb, K. (1991) A comparative analysis of selection schemes used in genetic algorithms. *Foundations of genetic algorithms*, *1*, 69-93.

Goldberg, D. E. and Voessner, S. (1999) Optimizing global-local search hybrids. In *Proceedings of the Genetic and Evolutionary Computation Conference,* vol.1, (Banzhaf et al.(eds.)). San Francisco, CA: Morgan Kaufmann, 220–228.

Gorev, N.B. and Kodzhespirova, I.F. (2013) Noniterative implementation of pressure-dependent demands using the hydraulic analysis engine of EPANET 2. *Water Resources Management* , 27(10), 3623-3630.

Grayman, W. M., Clark, R. M. and Males, R. M. (1988) Modeling distribution-system water quality; dynamic approach. *Journal of Water Resources Planning and Management*, 114(3), 295-312.

Gupta, R. and Bhave, P. R. (1996) Comparison of methods for predicting deficient-network performance. *Journal of Water Resources Planning and Management*, 122(3), 214-217.

Gupta, R., Sawarkar, V.R. and Bhave, P.R. (2003) Application of Newton-Raphson

method in optimal design of water distribution networks. *Journal of Indian Water Works Association,* 35(1), 31–37.

Haghighi, A., Samani, H. M. and Samani, Z. M. (2011) GA-ILP method for optimization of water distribution networks. *Water Resources Management*, 25(7), 1791-1808.

Halhal, D., Walters, G. A., Ouazar, D. and Savic, D. A. (1997) Water network rehabilitation with structured messy genetic algorithm. *Journal of Water Resources Planning and Management*, 123(3), 137-146.

Hebert, A., Forestier, D., Lenes, D., Benanou, D., Jacob, S., Arfi, C., Lambolez, L. and Levi, Y. (2010) Innovative method for prioritizing emerging disinfection by-products (DBPs) in drinking water on the basis of their potential impact on public health. *Water Research*, 4(10), 3147-3165.

Helbling, D.E. and Van Briesen, J.M. (2009) Modelling residual chlorine response to a microbial contamination event in drinking water distribution systems. *Journal of Environmental Engineering,* 135(10), 918–927.

Herrera, F., Lozano, M. and Verdegay, J. L. (1998) Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis. *Artificial Intelligence Review*, 12(4), 265-319.

Hindi, K. S. and Hamam, Y. M. (1993) Locating pressure control elements for leakage minimisation in water supply networks by genetic algorithms. In *Artificial Neural Nets and Genetic Algorithms* (583-587). Springer Vienna.

HMG (Her Majesty's Government) (2001) Water Supply (Water Quality) (Scotland) Regulations. Her Majesty's Stationery Office, Edinburgh, UK, Scottish Statutory Instrument 2001 No. 207.

HMG (2010) Water Supply (Water Quality) Regulations. The Stationery Office, London, UK, Statutory Instrument 2010 No. 994 (W. 99).

Holland, J. H. (1975) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Oxford, England: U Michigan Press.

Jayaram, N. and Srinivasan, K. (2008) Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. *Water Resources Research*, 44(1).

Jeong, H.S. and Abraham, D.M. (2006) Operational response model for physically attacked water networks using NSGA-II. *Journal of Computing in Civil Engineering*, 20(5), 328-338.

Jinesh Babu, K. S. and Mohan, S. (2012) Extended period simulation for pressure-deficient water distribution network. *Journal of Computing in Civil Engineering*, 26(4), 498-505.

Kadu, M. S., Gupta, R. and Bhave, P. R. (2008) Optimal design of water networks using a modified genetic algorithm with reduction in search space. *Journal of Water Resources Planning and Management*, 134(2), 147-160.

Kalungi, P. (2003) *A holistic approach to the optimal long-term upgrading of water distribution networks*. A Thesis Submitted in partial fulfilment of the Requirements of University of Liverpool for the Degree of Doctor of Philosophy. England: University of Liverpool.

Kalungi, P. and Tanyimboh, T. T. (2003) Redundancy model for water distribution systems. *Reliability Engineering and System Safety*, 82(3), 275-286.

Khu, S. T. and Keedwell, E. (2005) Introducing more choices (flexibility) in the upgrading of water distribution networks: the New York City tunnel network example. *Engineering optimization*, 37(3), 291-305.

Knowles, J. D. and Corne, D. W. (2000) Approximating the non-dominated front using the Pareto archived evolution strategy. *Evolutionary computation*, 8(2), 149-172.

Kollat, J. B. and Reed, P. M. (2006) Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. *Advances in Water Resources*, 29(6), 792-807.

Konak, A., Coit, D. W. and Smith, A. E. (2006) Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering and System Safety*, 91(9), 992-1007

Kovalenko, Y., Gorev, N.B., Kodzhespirova, I.F., Prokhorov, E., Trapaga, G. (2014) Convergence of a hydraulic solver with pressure-dependent demands. *Water Resources Management,* 28(4), 1013-1031.

Kumar, S. V., Doby, T. A., Baugh Jr, J. W., Brill, E. D. and Ranjithan, S. R. (2006) Optimal design of redundant water distribution networks using a cluster of workstations. *Journal of Water Resources Planning and Management*, 132(5), 374-384.

Lansey, K. E. and Mays, L. W. (1989) Optimization model for water distribution system design. *Journal of Hydraulic Engineering*, 115(10), 1401-1418.

Liou, C. P. and Kroon, J. R. (1987) Modeling the Propagation of Waterborne

Substances in Distribution Networks. *Journal of the American Water Works Association*, 79(11), 54-58.

Mackle, G., Savic, D. A. and Walters, G. A. (1995) Application of genetic algorithms to pump scheduling for water supply. In *Genetic Algorithms in Engineering Systems: Innovations and Applications, GALESIA*. First International Conference on (Conf. Publ. No. 414) (400-405). IET.

Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seah, H. Y. and Tan, C. L. (2003). Ant colony optimization for design of water distribution systems. *Journal of Water Resources Planning and Management*, 129(3), 200-209.

Marchi, A., Dandy, G., Wilkins, A., and Rohrlach, H. (2014) Methodology for Comparing Evolutionary Algorithms for Optimization of Water Distribution Systems. *Journal of Water Resources Planning and Management*, 140(1), 22–31.

Martin, D. W., and Peters, G. (1963) The application of Newton's method to network analysis by digital computer. *Journal of Institute of Water Engineers*, 17, 115-129.

Montalvo, I., Izquierdo, J., Perez, R. and Tung, M. M. (2008) Particle swarm optimization applied to the design of water supply systems. *Computers and Mathematics with Applications*, 56(3), 769-776.

Montesinos, P., Garcia-Guzman, A. and Ayuso, J. L. (1999) Water distribution network optimization using a modified genetic algorithm. *Water Resources Research*, 35(11), 3467-3473.

Morgan, D.R. and Goulter, I.C. (1985) Optimal urban water distribution design. *Water Resources Research*, 21(5), 642-652.

Munavalli, G. R. and Kumar, M. M. (2003) Optimal scheduling of multiple chlorine sources in water distribution systems. *Journal of Water Resources planning and Management*, 129(6), 493-504.

Nicklow, J., Reed, P., Savic, D., Dessalegne, T, Harrell, L., Chan-Hilton, A., Karamouz, M, Minsker, B., Ostfeld, A., Singh, A. and Zechman, E. (2010) State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management. *Journal of Water Resources Planning and Management*, 136(4), 412-432.

Nicolini, M., Giacomello, C. and Deb, K. (2011) Calibration and optimal leakage management for a real water distribution network. *Journal of Water Resources Planning and Management*, 137(1), 134-142.

Nieuwenhuijsen, M.J. (2005) Adverse reproductive health effects of exposure to chlorination disinfection by-products. *Global NEST Journal*, 7(1), 128-144.

Nowostawski, M. and Poli, R. (1999) Parallel genetic algorithm taxonomy. In *Proceedings of the Third International Conference on Knowledge-Based Intelligent Information Engineering Systems, KES'99, Adelaide*, 88-92.

OFWAT (2004) *Levels of Service for the Water Industry in England and Wales.* 2002-2003 Report. OFWAT Centre, 7 Hill Street, Birmingham B5 4UA, UK.

Ostfeld, A. and Tubaltzev, A. (2008) Ant colony optimization for least-cost design and operation of pumping water distribution systems. *Journal of Water Resources Planning and Management*, 134(2), 107-118.

Prasad, T. D. and Park, N. S. (2004) Multiobjective genetic algorithms for design of water distribution networks. *Journal of Water Resources Planning and Management*, 130(1), 73-82.

Prasad, T. D. (2010) Design of pumped water distribution networks with storage. *Journal of Water Resources Planning and Management*, 136(1), 129-132.

Preis, A. and Ostfeld, A. (2008) Multiobjective contaminant sensor network design for water distribution systems. *Journal of Water Resources Planning and Management*, 134(4), 366-377.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992) Numerical Recipes in FORTRAN: The Art of Scientific Computing. New York, USA: Cambridge University Press.

Preux, P. and Talbi, E. G. (1999) Towards hybrid evolutionary algorithms. *International Transactions in Operational Research*, 6(6), 557-570.

Rao, Z. and Salomons, E. (2007) Development of a real-time, near-optimal control process for water-distribution networks. *Journal of Hydroinformatics*, 9(1), 25-37.

Ray, T., Singh, H.K., Isaacs, A. and Smith, W. (2009) Infeasibility driven evolutionary algorithm for constrained optimization. In *Constraint Handling in Evolutionary Optimization Studies in Computational Intelligence, vol.198, 145-165.*

Richardson, S.D., Simmons, J.E. and Rice, G. (2002) Disinfection by-products: the next generation. *Environmental Science and Technology*, 36(9), 198A–205A.

Rodriguez, M.J., Serodes, J.B. and Levallois, P. (2004) Behaviour of trihalomethanes

and haloacetic acids in a drinking water distribution system. *Water Research,* 38(20), 4367–4382.

Rossman, L.A. (2000) EPANET 2 Users Manual. Water Supply and Water Resources Division, National Risk Management Research Laboratory, US EPA, Cincinnati, OH, USA.

Rossman, L. A. (2007) Discussion of 'Solution for water distribution systems under pressure-deficient conditions'. *Journal of Water Resources Planning and Management*, 133(6), 566-567.

Rossman, L.A. and Boulos, P.F. (1996) Numerical methods for modeling water quality in distribution systems: A comparison, *Journal of Water Resources Planning and Management*, 122(2), 137-146.

Rossman, L. A., Boulos, P. F. and Altman, T. (1993) Discrete volume element method for network water quality models. *Journal of Water Resources Planning and Management*, 119(5), 505–517.

Rossman, L.A., Clark, R.M. and Grayman, W.M. (1994). Modeling chlorine residuals in drinking-water distribution systems, *Journal of Environmental Engineering*, 120(4), 803-820.

Saleh, S. H. and Tanyimboh, T. T. (2013) Coupled topology and pipe size optimization of water distribution systems. *Water Resources Management*, 27(14), 4795-4814.

Saleh, S. H. and Tanyimboh, T. T. (2014) Optimal Design of Water Distribution Systems Based on Entropy and Topology. *Water Resources Management*, 28(11), 3555-3575.

Savic, D. A. and Walters, G.A. (1995) *Genetic Algorithm Techniques for Calibrating Network Models*. Report No. 95/12.Center for Systems and Control Engineering, University of Exeter, UK.

Savic, D. A. and Walters, G. A. (1997) Genetic algorithms for least-cost design of water distribution networks. *Journal of Water Resources Planning and Management*, 123(2), 67-77.

Schaffer, J. D. (1985) Multiple objective optimization with vector evaluated genetic algorithms. In *Proceedings of the 1st International Conference on Genetic Algorithms, Pittsburgh, PA, USA,* 93-100.

Schutte, J. F., Reinbolt, J. A., Fregly, B. J., Haftka, R. T. and George, A. D. (2004) Parallel global optimization with the particle swarm algorithm. *International Journal for Numerical Methods in Engineering*, 61(13), 2296-2315.

Setiadi, Y., Tanyimboh, T. T. and Templeman, A. B. (2005) Modelling errors, entropy and the hydraulic reliability of water distribution systems. *Advances in Engineering Software*, *36*(11), 780-788.

Seyoum, A.G. and Tanyimboh, T.T. (2013) Pressure-dependent multiple-species water quality modelling of water distribution networks. In *Proceedings of 8th International Conference of European Water Resources Association,* Porto, Portugal, 553–558.

Seyoum, A.G. and Tanyimboh, T.T. (2014) Pressure dependent network water quality modelling. *Proceedings of ICE: Water Management*, 167(6), 342-355.

Seyoum, A.G., Tanyimboh, T.T. and Siew, C. (2014) Optimal tank design and operation to improve water quality in distribution systems. In *Proceedings of the 11th International Conference on Hydroinformatics*, New York City, USA, 16-21 August 2014.

Shan, N. (2004) *Head dependent modelling of water distribution network*. A Thesis Submitted in partial fulfilment of the Requirements of University of Liverpool for the Degree of Master of Science. England: University of Liverpool.

Shang, F., Uber, J.G. and Rossman, L.A. (2008) Modelling reaction and transport of multiple species in water distribution systems. *Environmental Science and Technology* 42(3), 808–814.

Shinde, G.N., Jagtap, S.B. and Pani, S. K. (2011) Parallelizing Multi-objective Evolutionary Genetic Algorithms. In *Proceedings of the World Congress on Engineering* (Vol. 2).

Siew, C and Tanyimboh, T.T. (2010a) Penalty-free multi-objective evolutionary optimization of water distribution systems. In *Proceedings of 12th International Conference on Water Distribution Systems Analysis, Tucson, AZ, USA*, CD-ROM.

Siew, C. and Tanyimboh, T.T. (2010b) Pressure-dependent EPANET extension: extended period simulation. In *Proceedings of 12th International Conference on Water Distribution Systems Analysis, Tucson, AZ, USA*, CD-ROM.

Siew, C. and Tanyimboh, T.T. (2011a) Design of the "Anytown" network using the penalty-free multi-objective evolutionary optimization approach. In *Proceedings of the 13th Annual Water Distribution Systems Analysis Conference, WDSA,* 22-26.

Siew, C. and Tanyimboh, T.T. (2011b) Penalty-free evolutionary algorithm

optimization for the long-term rehabilitation and upgrading of water distribution systems. In *Proceedings of the World Environmental and Water Resources Congress, Palm Springs, California*, ISBN 978-0-7844-1173-5, 214–223.

Siew, C. and Tanyimboh, T.T. (2012a) Pressure-dependent EPANET extension. *Water Resources Management* 26(6), 1477–1498.

Siew, C., and Tanyimboh, T.T. (2012b) Penalty-free feasibility boundary convergent multi-objective evolutionary algorithm for the optimization of water distribution systems. *Water Resources Management* 26(15), 4485–4507

Siew, C., Tanyimboh, T. T. and Seyoum, A. G. (2014) Assessment of Penalty-Free Multi-Objective Evolutionary Optimization Approach for the Design and Rehabilitation of Water Distribution Systems. *Water Resources Management*, 28(2), 373-389.

Singh, H.K., Isaacs, A., Ray, T. and Smith, W. (2008) Infeasibility Driven Evolutionary Algorithm (IDEA) for Engineering Design Optimization. *In Proceedings of 21st Australiasian Joint Conference on Artificial Intelligence AI-08*, 104–115.

Simpson, A. R., Dandy, G. C. and Murphy, L. J. (1994) Genetic algorithms compared to other techniques for pipe optimization. *Journal of water resources planning and management*, 120(4), 423-443.

Skadsen, J., Janke, R., Grayman, W., Samuels, W., Tenbroek, M., Steglitz, B. and Bahl, S. (2008) Distribution system on-line monitoring for detecting contamination and water quality changes. *Journal of American Water Works Association,* 100(7), 81–94.

Sohn, J., Amy, G., Cho, J., Lee, Y. and Yoon, Y. (2004) Disinfectant decay and disinfection by-products formation model development: chlorination and ozonation by-products. *Water Research,* 38(10), 2461–2478.

Spiliotis, M. and Tsakiris, G. (2011) Water distribution system analysis: Newton-Raphson method revisited. *Journal of Hydraulic Engineering*, 137(8), 852-855.

Srinivas, N. and Deb, K. (1994) Muiltiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary computation*, 2(3), 221-248.

Su, Y. C., Mays, L. W., Duan, N. and Lansey, K. E. (1987) Reliability-based optimization model for water distribution systems. *Journal of Hydraulic Engineering*, 113(12), 1539-1556.

Tabesh, M. (1998) *Implications of Pressure Dependency of Outflows on Data Management, Mathematical Modelling and Reliability Assessment of Water Distribution Systems*. A Thesis Submitted in partial fulfilment of the Requirements of University of Liverpool for the Degree of Doctor of Philosophy. England : University of Liverpool.

Tabesh, M., Asadiyani Yekta, A.H., Burrows, R. (2009) An Integrated Model to Evaluate Losses in Water Distribution Systems. *Water Resources Management*, 23(3), 477–492.

Tabesh, M., Tanyimboh, T.T. and Burrows, R. (2002) Head-driven simulation of water supply networks. *International Journal of Engineering,* 15(1), 11–22.

Talbi, E. G. (2002). A taxonomy of hybrid metaheuristics. *Journal of heuristics*, 8(5), 541-564.

Tang, Y., Reed, P. M. and Kollat, J. B. (2007) Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications. *Advances in Water Resources*, 30(3), 335-353.

Tanyimboh, T. T., Burd, R., Burrows, R. and Tabesh, M. (1999) Modelling and reliability analysis of water distribution systems. *Water Science and Technology*, 39(4), 249-255.

Tanyimboh, T.T. and Kalungi, P. (2009) Multi-criteria assessment of optimal design, rehabilitation and upgrading schemes for water distribution networks. *Civil Engineering and Environmental Systems,* 26(2), 117–140.

Tanyimboh, T.T. and Setiadi, Y. (2008a) Joint layout, pipe size and hydraulic reliability optimization of water distribution systems. *Engineering Optimization*, 40(8), 729-747.

Tanyimboh, T.T. and Setiadi, Y. (2008b) Sensitivity analysis of entropy constrained designs of water distribution systems. *Engineering Optimization*, 40(5), 439-457.

Tanyimboh, T.T. and Siew, C. (2012) Reference pressure for extended period simulation of pressure deficient water distribution systems. *ASCE World Environmental and Water Resources Congress 2012: Crossing Boundaries* (Loucks, E.D. (ed.)). ASCE, Albuquerque, NM, USA, 3257–3264.

Tanyimboh, T.T. and Tabesh, M. (1997) Discussion of Comparison of methods for predicting deficient-network performance. *Journal of Water Resources Planning and Management*, 124(6), 369-370.

Tanyimboh, T. T., Tabesh, M. and Burrows, R. (2001). An appraisal of source head

methods for calculating the reliability of water distribution networks. *Journal of Water Resources Planning and Management*, 127 (4), 206-213.

Tanyimboh, T.T., Tahar, B. and Templeman, A. B. (2003) Pressure-driven modelling of water distribution systems. *Water Science and Technology Water Supply*, 3 (1-2), 255-262.

Tanyimboh, T.T. and Templeman, A. B. (1995) A new method for calculating the reliability of single-source networks. *Developments in Computational Techniques for Civil Engineering* (Topping, B.H.V. (ed.)). Civil-Comp Press.

Tanyimboh, T.T. and Templeman, A.B. (2010) Seamless pressure deficient water distribution system model. *Proceedings of ICE: Water Management*, 163(8), 389–396.

Todini, E. (2003) A more realistic approach to the extended period simulation of water distribution networks. *Advances in water supply management* (Maksimovic, C., Butler, D. and Memon, F. A. (eds.)). Balkema, Lisse, The Netherlands, 173–184.

Todini, E. and Pilati, S. (1988) A gradient algorithm for the analysis of pipe networks. *Computer applications in water supply: Systems analysis and simulation*, vol. 1 (Coulbeck, B. and Chun-Hou, O. (eds.)). Research Studies Press, Taunton, UK, 1–20.

Trobec, R., Vajtersic, M. and Zinterhof, P. (2009) *Parallel computing: numerics, applications, and trends*. Springer Science & Business Media.

Tsakiris, G. and Spiliotis, M. (2014) A Newton-Raphson analysis of urban water systems based on nodal head-driven outflow. *European Journal of Environmental and Civil Engineering*, 18(8), 882-896.

Twort, A.C., Ratnayaka, D.D. and Brandt, M.J. (2000) *Water Supply*. Arnold, London, UK.

Tzatchkov, V.G., Aldama, A.A. and Arreguin, F.I. (2002) Advection– dispersion–reaction modeling in water distribution networks. *Journal of Water Resources Planning and Management,* 128(5), 334–342.

Udo, A. and Ozawa, T. (2001) Steady-state flow analysis of pipe networks considering reduction of flow in the case of low water pressures. *Water Software Systems: Theory and Applications* (Ulanicki, B., Coulbeck, B. and Rance, J. (eds.)). Research Studies Press, Taunton, UK, vol. 1, 73-182.

US EPA (United States Environmental Protection Agency). (1996) *Safe Drinking*

Water Act, *1996*. [Online] Available from: http://water.epa.gov/lawsregs/rulesregs/sdwa/index.cfm [accessed: 18th July 2013].

Vairavamoorthy, K. and Ali, M. (2000) Optimal design of water distribution systems using genetic algorithms. *Computer-Aided Civil and Infrastructure Engineering*, 15(5), 374-382.

Vairavamoorthy, K. and Ali, M. (2005) Pipe index vector: A method to improve genetic-algorithm-based pipe optimization. *Journal of Hydraulic Engineering*, 131(12), 1117-1125.

Vairavamoorthy, K. and Shen, Y. (2004) Least cost design of water distribution network using particle swarm optimization. In: Phoon, Liong, B. (Eds.), *Sixth International Conference on Hydroinformatics*. World Scientific Publishing Company, vol. 1, 834–841.

Van Veldhuizen, D. A. and Lamont, G. B. (1998) *Multiobjective evolutionary algorithm research: A history and analysis*. Technical Report TR-98-03. Department of Electrical and Computer Engineering, Graduate School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, Ohio.

Van Zyl, J. E., Savic, D. A. and Walters, G. A. (2004) Operational optimization of water distribution systems using a hybrid genetic algorithm. *Journal of Water Resources Planning and Management*, 130(2), 160-170.

Vamvakeridou-Lyroudia, L. S., Walters, G. A. and Savic, D. A. (2005) Fuzzy multiobjective optimization of water distribution networks. *Journal of Water Resources Planning and Management*, 131(6), 467-476.

Vasan, A. and Simonovic, S. P. (2010) Optimization of water distribution network design using differential evolution. *Journal of Water Resources Planning and Management*, 136(2), 279-287.

Vitkovsky, J. P. and Simpson, A. R (1997) Calibration and leak detection in pipe networks using inverse transient analysis and genetic algorithms. *Res. Rep. No. R157*, Department of Civil and Environmental Engineering, University of Adelaide: Adelaide, Australia.

Wagner, J. M., Shamir, U. and Marks, D. H. (1988) Water distribution reliability: simulation methods. *Journal of Water Resources Planning and Management*, 114(3), 276-294.

Walski, T. M., Brill, E. D., Gessler, J., Goulter, I. C., Jeppson, R. M., Lansey, K.,

Lee, H. L., Liebman, J. C., Mays, L., Morgan, D. R. and Ormsbee, L. (1987) Battle of the network models: epilogue. *Journal of Water Resource Planning and Management*, 113(2), 191-203.

Walski, T. M. et al. (2003) *Advanced water distribution modeling and management*. Haestad press.

Wang, Q., Creaco, E., Franchini, M., Savic, D. and Kapelan, Z. (2015) Comparing Low and High-Level Hybrid Algorithms on the Two-Objective Optimal Design of Water Distribution Systems. *Water Resources Management*, 1-16.

Weickgenannt, M., Kapelan, Z., Blokker, M. and Savic, D. A. (2010) Risk-based sensor placement for contaminant detection in water distribution systems. *Journal of Water Resources Planning and Management*, 136(6), 629-636.

WHO (World Health Organization) (2011) Guidelines for Drinking-water Quality. WHO, Geneva, Switzerland.

Wilkinson, B. and Allen, M. (2004) *Parallel programming: techniques and applications using networked workstations and parallel computers*, 2nd Edition, Pearson/Prentice Hall.

Wood, D. J. and Charles, C. O. (1972) Hydraulic network analysis using linear theory. *Journal of the Hydraulics Division*, 98(7), 1157-1170.

Wu, Z. Y., Boulos, P. F., Orr, C. H. and Ro, J. J. (2001) Using genetic algorithm to rehabilitate distribution systems. *Journal American Water Works Association*, 93(11), 74-85.

Wu, Z. Y. and Simpson, A. R. (2002) A self-adaptive boundary search genetic algorithm and its application to water distribution systems. *Journal of Hydraulic Research*, 40(2), 191-203.

Wu, Z.Y. and Walski, T. (2005) Self-adaptive penalty approach compared with other constraint-handling techniques for pipeline optimization. *Journal of Water Resources Planning and Management*, *131*(3), 181-192.

Wu, Z.Y., Wang, R. H., Walski, T. M., Yang, S. Y., Bowdler, D. and Baggett, C. C. (2009) Extended global-gradient algorithm for pressure-dependent water distribution analysis. *Journal of Water Resources Planning and Management*, 135(1), 13-22.

Wu, Z. Y. and Zhu, Q. (2009) Scalable parallel computing framework for pump scheduling optimization. *ASCE, Proceedings of World Environmental and Water Resources Congress 2009*, 1-11, doi: 10.1061/41036(342)42.

Zitzler, E. and Thiele, L. (1998) *An Evolutionary Algorithm for Multiobjective Optimization: The Strength Pareto Approach*. Technical Report 43. Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH). Zurich, Switzerland.

Zitzler, E., Deb, K. and Thiele, L. (2000) Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary Computation,* 8(2), 173-195.

Appendix A

Network Data for Case Studies in Chapter 4

A-1 Input Data for Network in Example 2

Table A-1.1 Node data for network in Example 2

Node	Elevation (m)	Demand (l/s)
1	6.1	31.57
2	15.2	3.15
3	15.2	12.64
4	15.2	3.15
5	24.4	3.15
6	24.4	37.89
7	24.4	37.89
8	24.4	25.25
9	36.6	3.15
10	36.6	3.15
11	36.6	25.25
12	15.2	3.15
13	15.2	31.57
14	15.2	31.57
15	15.2	3.15
16	36.6	25.25
17	36.6	63.14
18	15.2	31.57
19	15.2	63.14
20	6.1	0.00
21	15.2	0.00
22	36.6	0.00

Table A-1.2 Reservoir data for network in Example 2

Reservoir	Head (m)
40	3.0488

Table A-1.3 Tank data for network in Example 2

Tank	Elevation (m)	Initial level (m)	Minimum level (m)	Maximum level (m)	Diameter (mm)
41	65.5	3.0	3.0	10.7	829.3
42	65.5	3.0	3.0	10.7	829.3

Table A-1.4 Pipe data for network in Example 2

Pipe	Length (m)	Diameter (mm)	Roughness
1	3658.54	305	120
2	3658.54	305	70
3	3658.54	406	70
4	30.49	762	130
5	1829.27	254	120
6	2743.90	254	120
7	2743.90	305	70
8	1829.27	254	120
9	1829.27	254	120
11	3658.54	203	120
12	1829.27	254	120
17	3658.54	203	120
18	1829.27	254	120
19	1829.27	203	120
20	1829.27	203	120

21	1829.27	203	120
22	1829.27	203	120
23	1829.27	254	120
24	1829.27	203	120
26	1829.27	254	120
27	1829.27	203	70
28	1829.27	305	70
29	1829.27	305	70
30	1829.27	254	70
31	1829.27	305	70
32	1829.27	254	70
33	30.49	305	120
34	1829.27	254	70
35	1829.27	254	70
36	1829.27	203	120
37	1829.27	305	70
38	1829.27	254	70
39	1829.27	203	120
40	30.49	305	120
41	1829.27	254	70
142	0.30	305	120
143	0.30	305	120
10	1829.27	305	130
13	1829.27	305	130
14	1829.27	305	130
15	1829.27	305	130
16	1829.27	305	130
25	2743.90	305	130

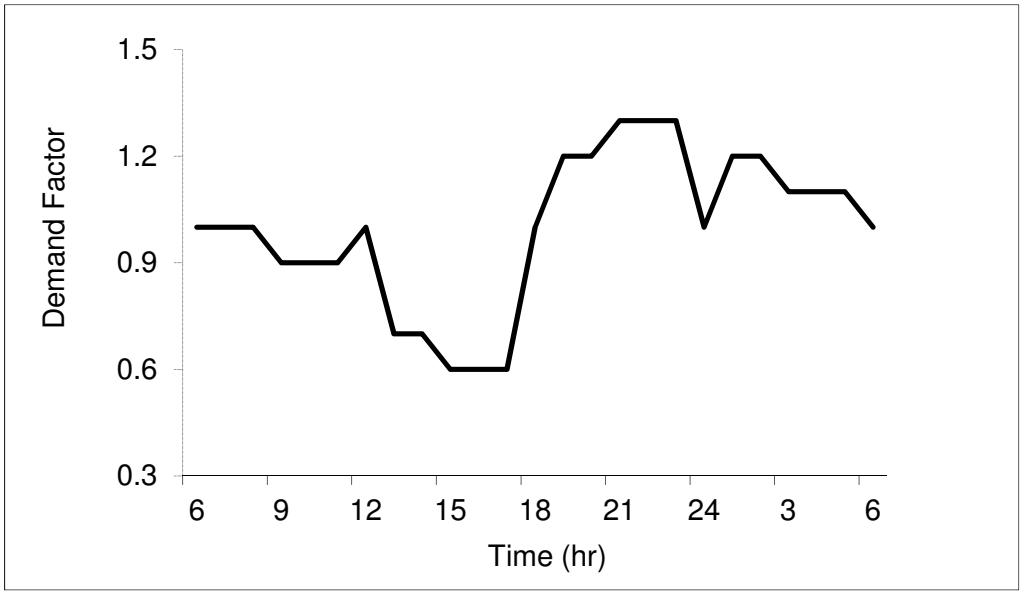
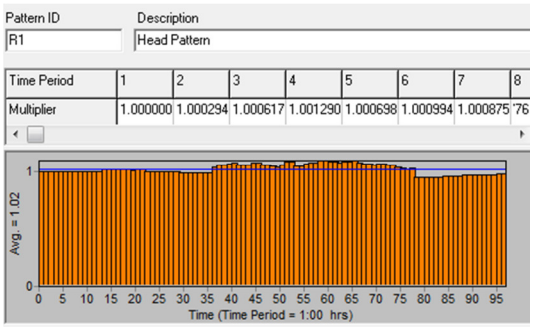


Fig A-1.1 Demand factor for network in Example 2

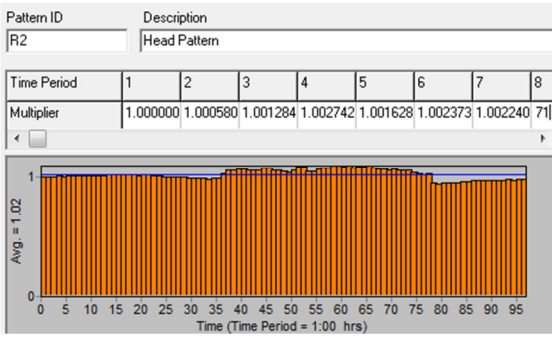
Appendix B

Additional Network Data and Results for Case Studies in Chapter 5

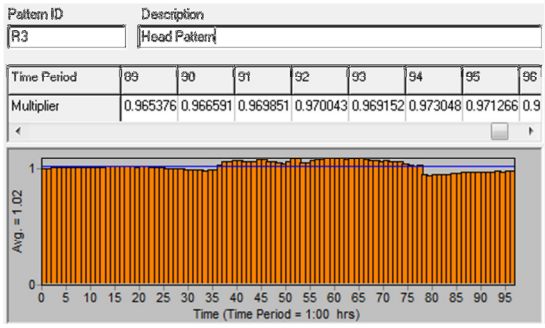
B-1 Network 2



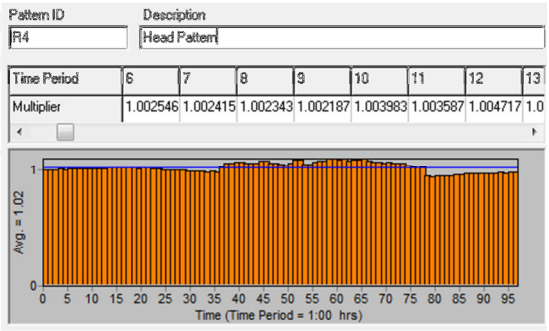
(a) R1



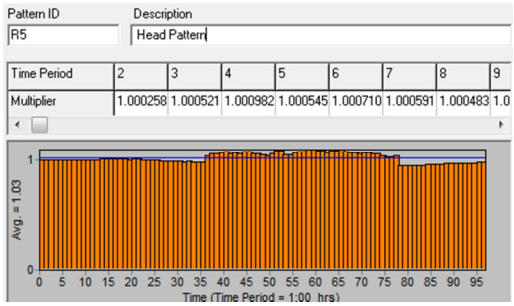
(b) R2



(c) R3

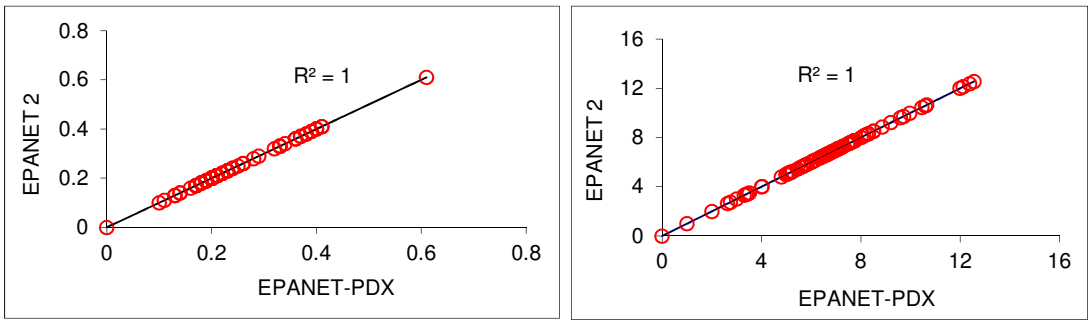


(d) R4



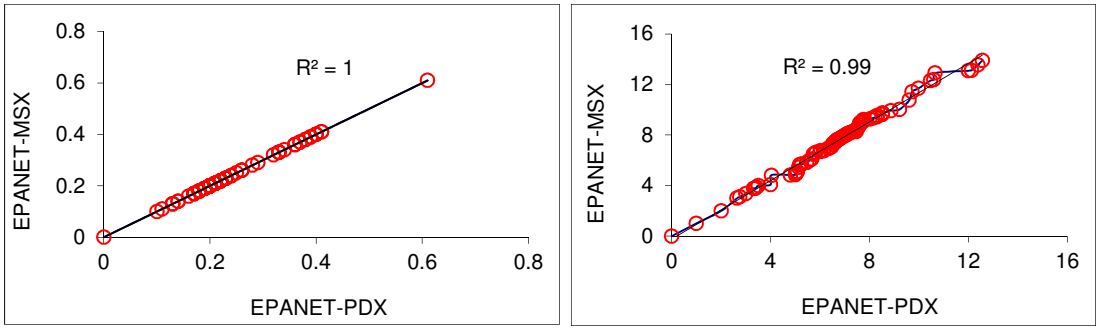
(e) R5

Figure B-1.1 (a)-(e) Variable-head supply nodes R1-R5



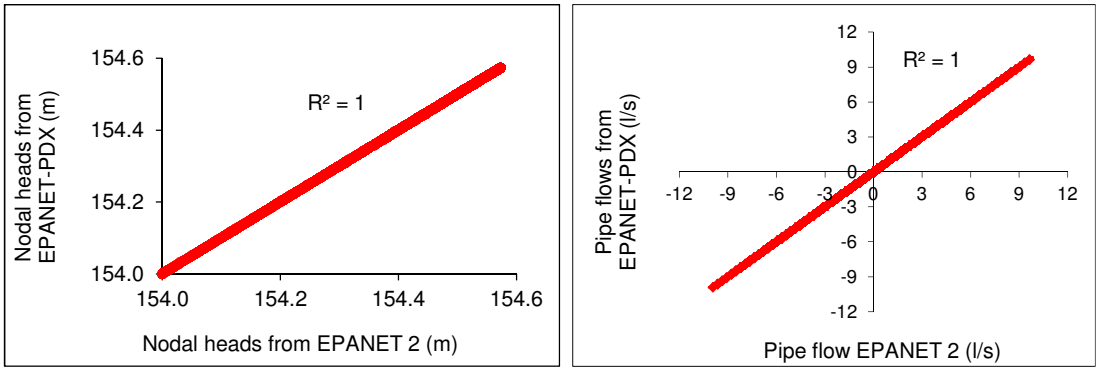
(a) (b)

Figure B-1.2 Water age consistency check for normal pressure using EPANET 2 and EPANET-PDX: (a) Node 1; (b) Node 2



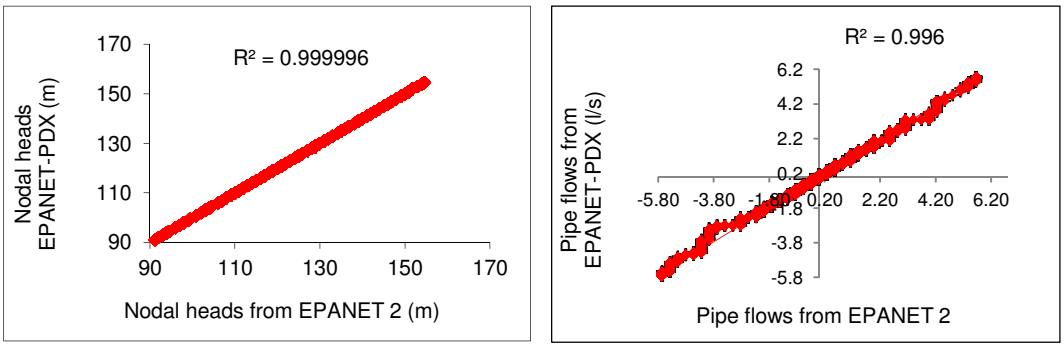
(a) (b)

Figure B-1.3 Water age consistency check for normal pressure using EPANET-MSX and EPANET-PDX: (a) Node 1; (b) Node 2



(a) (b)

Figure B-1.4 Hydraulic feasibility check for normal pressure: (a) nodal heads; (b) pipe flows



(a) (b)

Figure B-1.5 Hydraulic feasibility check for supply node heads 90 m-155 m (a) nodal heads; (b) pipe flows

B-2 Network 3

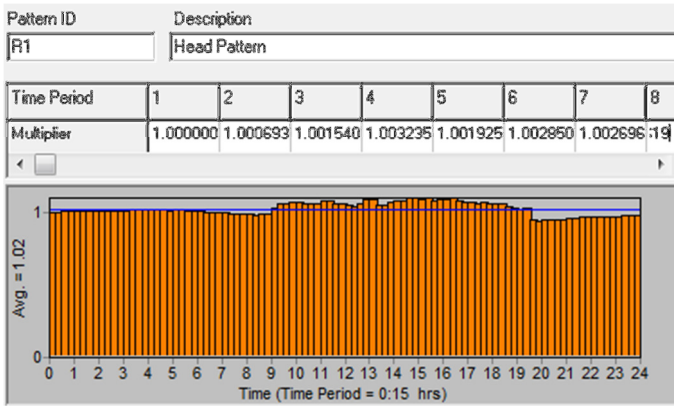
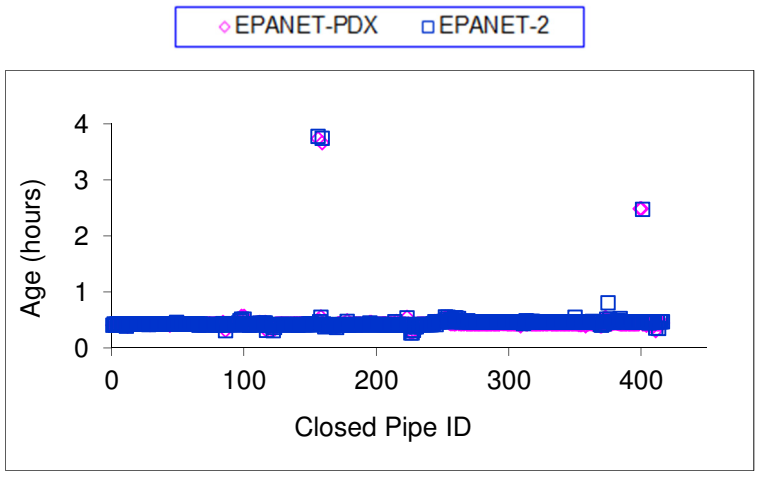
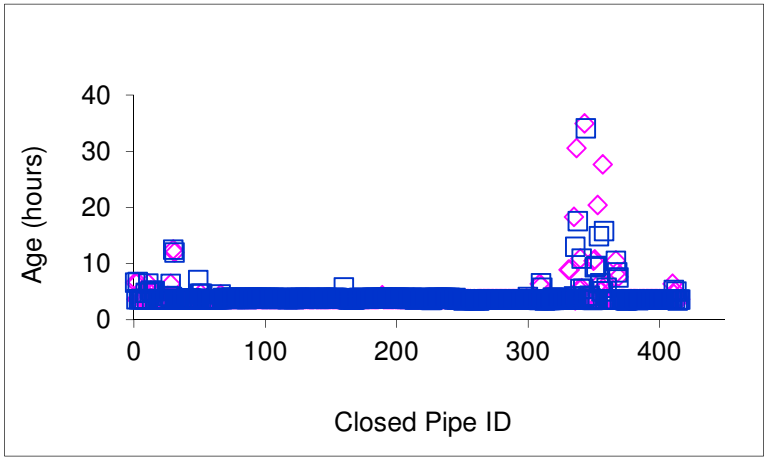


Figure B-2.1 Supply node R1



(a)



(b)

Figure B-2.2 Pipe closure effects on water age: (a) Node 3; (b) Node 4

Appendix C

Network Data and Results for Case Studies in Chapter 6

C-1 Network 1

Table C-1.1 Pipe and node data for Network 1

Pipe details		Node details		
Pipe	Length (m)	Node	Minimum required nodal head (m)	Demand (m3/min)
1	300	1	100	–
2	820	2	95	–
3	940	3	85	18.4
4	730	4	85	4.5
5	1620	5	85	6.5
6	600	6	85	4.2
7	800	7	82	3.1
8	1400	8	82	6.2
9	1175	9	85	8.5
10	750	10	85	11.5
11	210	11	85	8.2
12	700	12	85	13.6
13	310	13	82	14.8
14	500	14	82	10.6
15	1960	15	85	10.5
16	900	16	82	9
17	850	17	82	6.8
18	650	18	85	3.4
19	760	19	82	4.6

20	1100	20	82	10.6
21	660	21	82	12.6
22	1170	22	80	5.4
23	980	23	82	2
24	670	24	80	4.5
25	1080	25	80	3.5
26	750	26	80	2.2
27	900			
28	650			
29	1540			
30	730			
31	1170			
32	1650			
33	1320			
34	3250			

Table C-1.2 Commercial pipe sizes and unit costs for Network 1

Pipe diameter (mm)	Unit cost (rupees)
150	1,115
200	1,600
250	2,154
300	2,780
350	3,475
400	4,255
450	5,172
500	6,092
600	8,189
700	10,670
750	11,874
800	13,261
900	16,151
1000	19,395

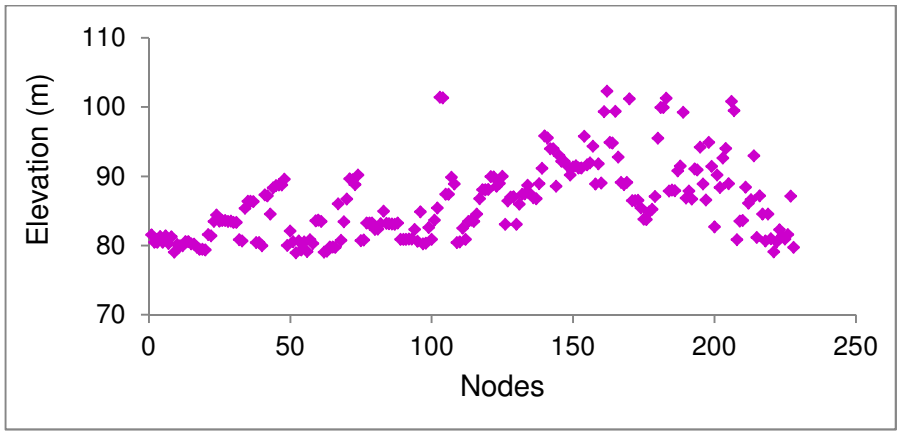
Table C-1.3 Pipe candidate diameters for reduced search space for Network 1

Pipe	Candidate diameters (mm)*				
1	800,	900,	1000,	1000,	1000
2	750,	800,	900,	1000,	1000
3	300,	350,	400,	450,	500
4	150,	200,	250,	300,	350
5	150,	150,	150,	200,	250
6	150,	150,	200,	250,	300
7	700,	750,	800,	900,	1000
8	150,	150,	150,	200,	250
9	450,	500,	600,	700,	750
10	500,	600,	700,	750,	800
11	750,	800,	900,	1000,	1000
12	500,	600,	700,	750,	800
13	400,	450,	500,	600,	700
14	350,	400,	450,	500,	600
15	150,	150,	150,	200,	250
16	400,	450,	500,	600,	700
17	250,	300,	350,	400,	450
18	250,	300,	350,	400,	450
19	350,	400,	450,	500,	600
20	150,	150,	150,	200,	250
21	450,	500,	600,	700,	750

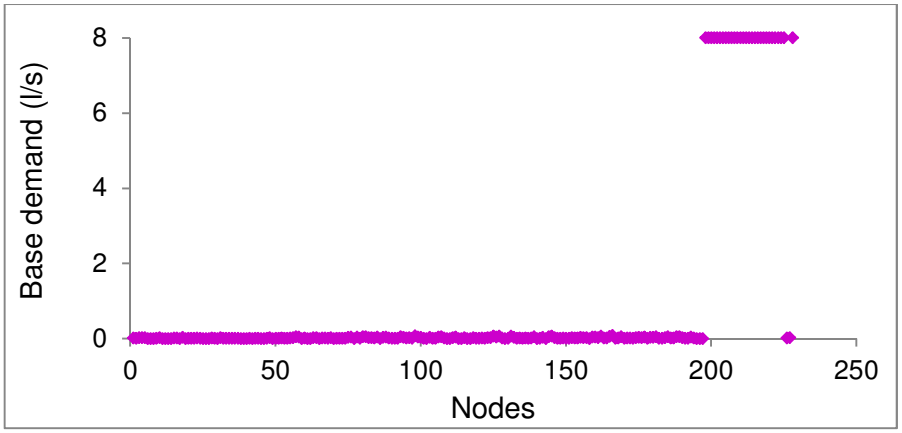
22	150,	150,	150,	200,	250
23	150,	150,	150,	200,	250
24	250,	300,	350,	400,	450
25	400,	450,	500,	600,	700
26	150,	200,	250,	300,	350
27	150,	200,	250,	300,	350
28	150,	200,	250,	300,	350
29	150,	150,	150,	200,	250
30	250,	300,	350,	400,	450
31	150,	150,	150,	200,	250
32	150,	150,	150,	200,	250
33	150,	150,	150,	200,	250
34	150,	150,	200,	250,	300

*Candidate diameters used by Kadu et al. (2008) for reduced search space

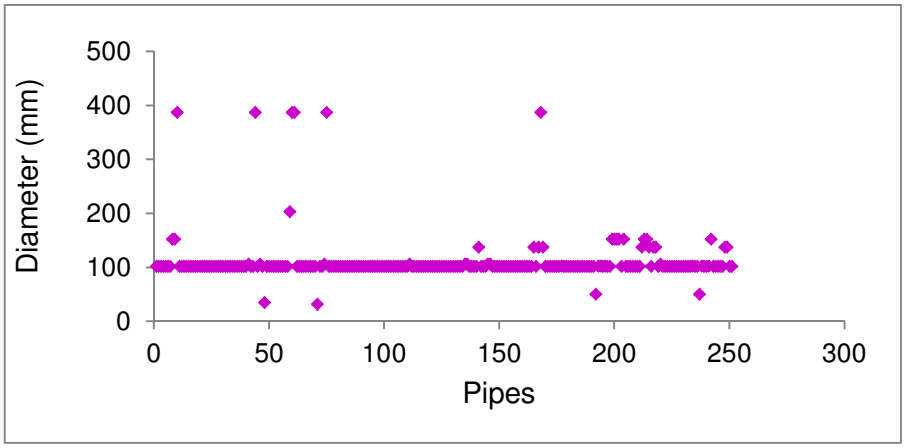
C-2 Network 2



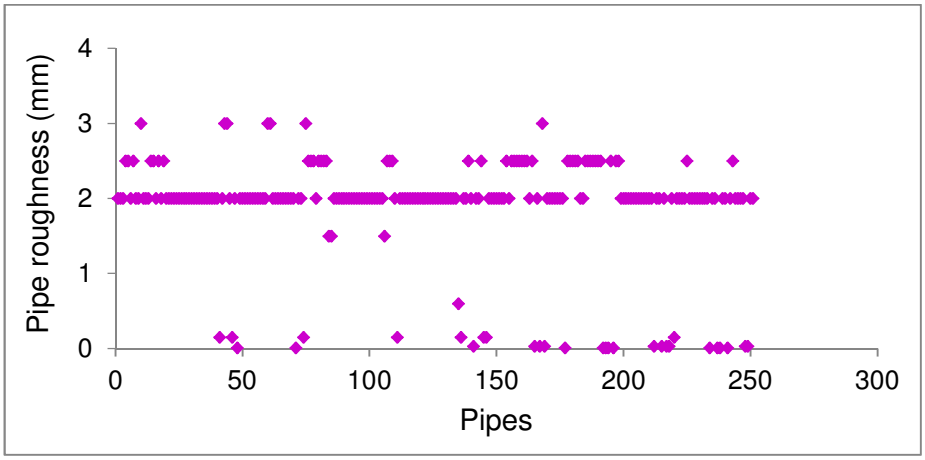
(a) Node elevation



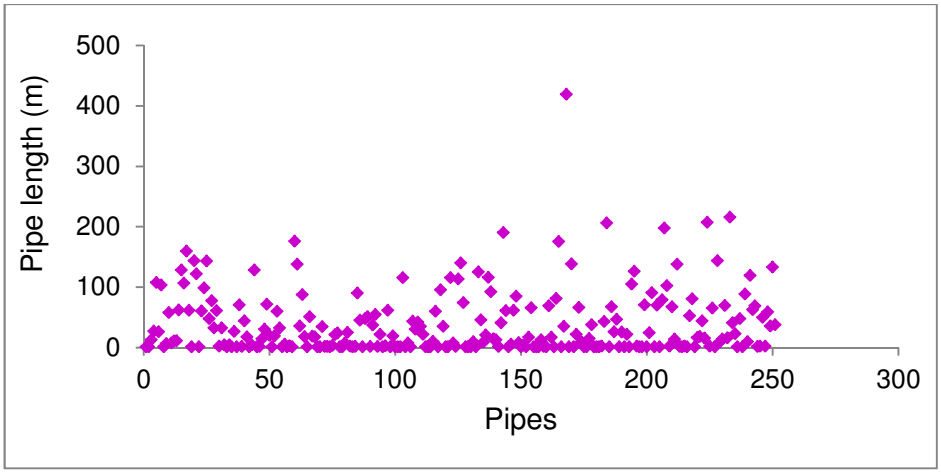
(b) Base demand



(c) Pipe diameter



(d) Pipe roughness



(e) Pipe length

Figure C-2.1 Pipe and node data for Network 2

Table C-2.1 Supply nodes data for Network 2

Supply nodes	Head (m)
R1	130.33
R2	129.94
R3	129.85
R4	129.88
R5	130.32

Table C-2.2 Commercial pipe sizes and unit costs for Network 2

Pipe diameter (mm)	Unit cost (£)
32	33.73
50	37.99
75	44.77
100	52.86
150	73.54
200	102.15
250	142.12
300	197.74
350	275.11
400	382.64

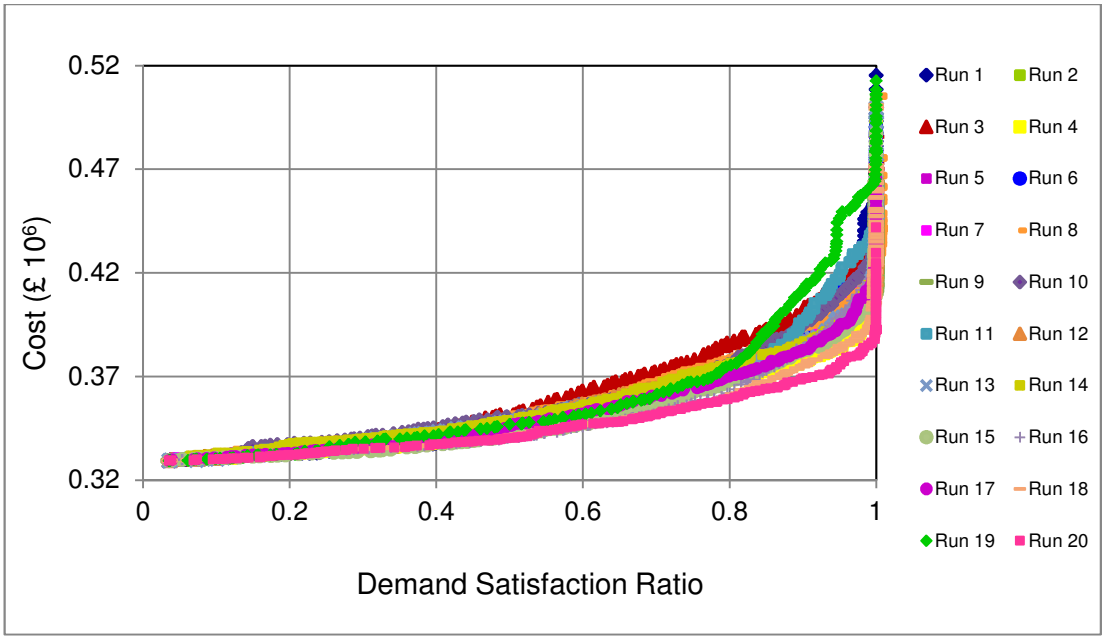


Figure C-2.2 Pareto optimal fronts for Network 2 (population = 200)

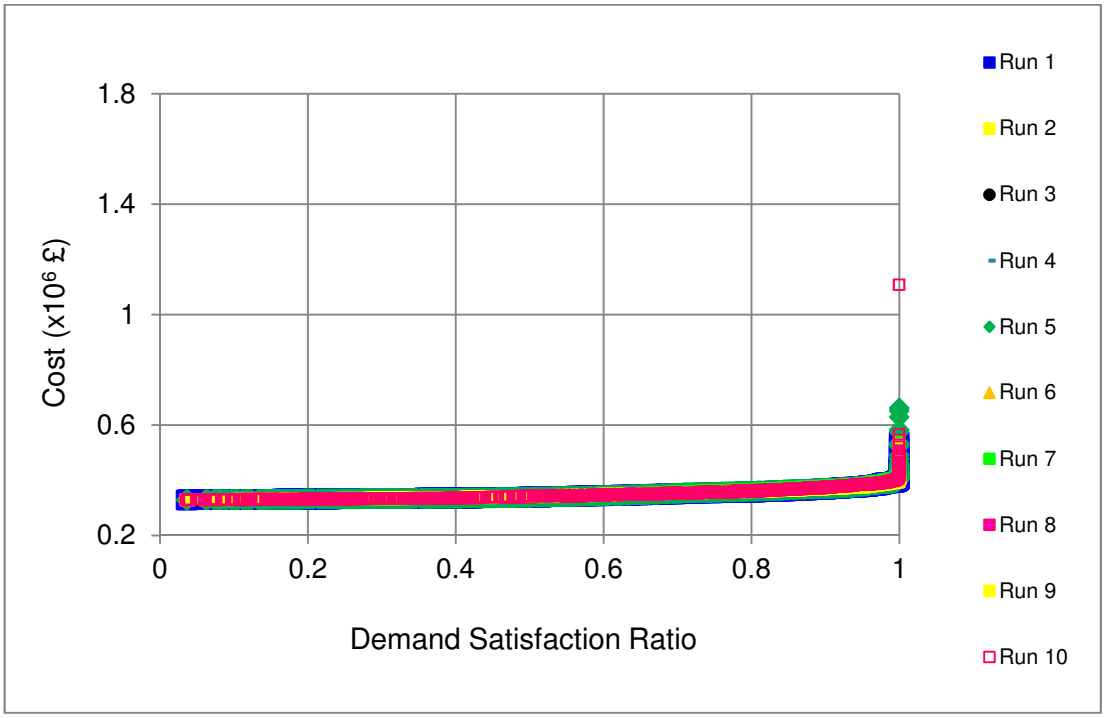


Figure C-2.3 Pareto optimal fronts for Network 2 (population = 1000)

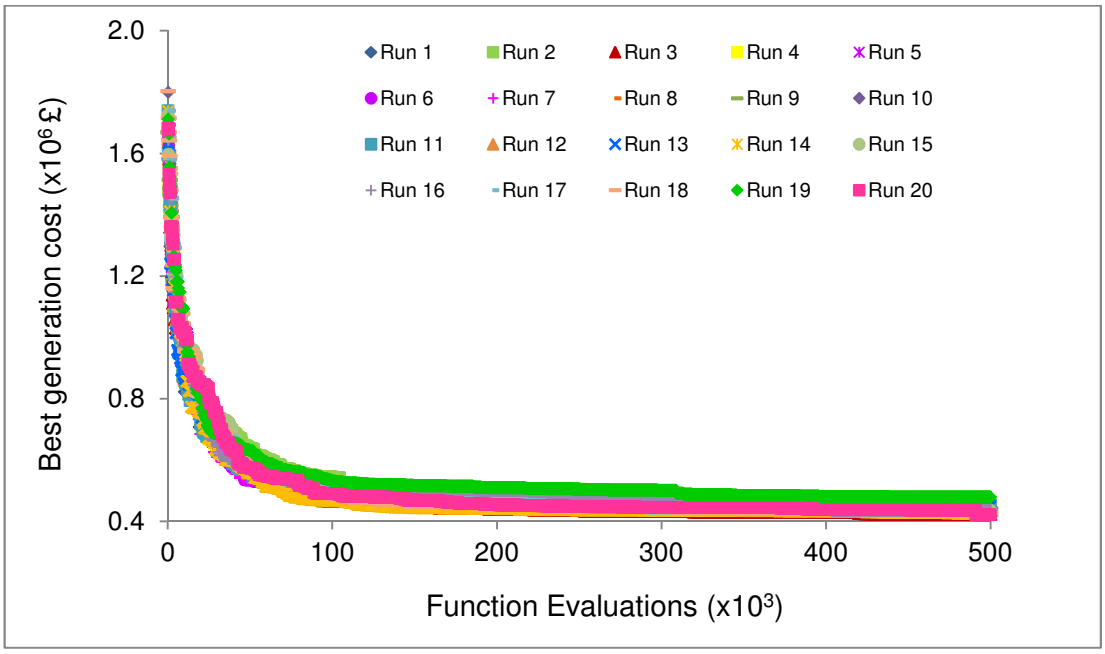


Figure C- 2.4 Progress of optimisation runs for Network 2 (population = 200)

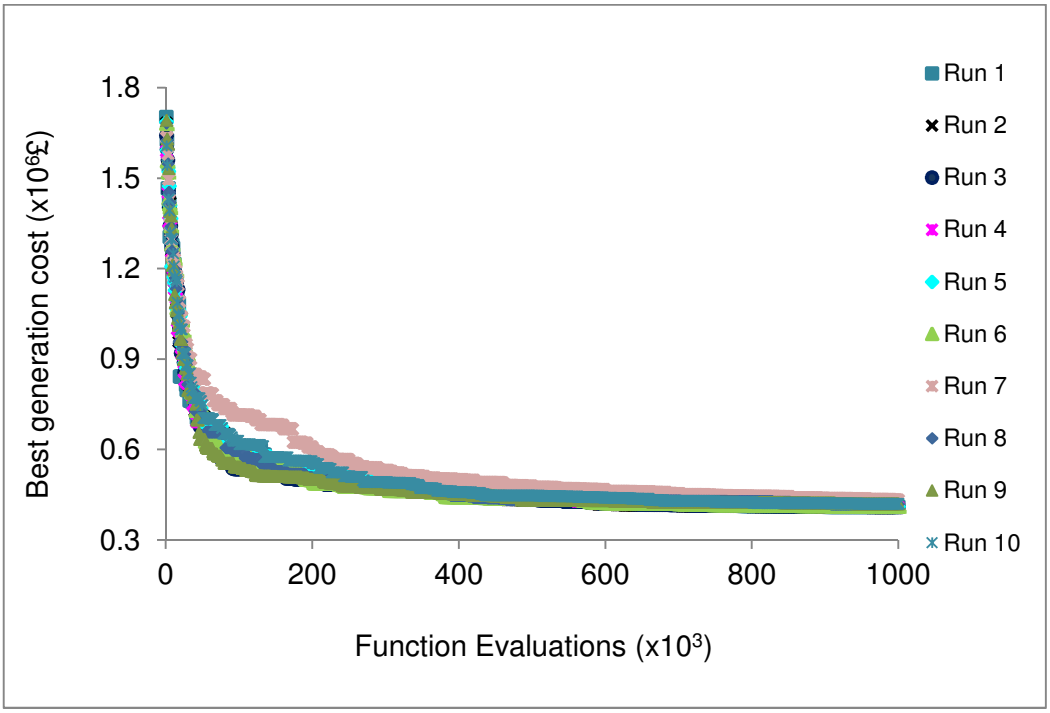


Figure C- 2.5 Progress of optimisation runs for Network 2 (population = 1000)

LIST OF PUBLICATIONS

Peer-Reviewed International Journal Papers

Seyoum, A.G., Tanyimboh, T.T. and Siew, C. (2013) Assessment of Water Quality Modelling Capabilities of EPANET Multi-species and Pressure Dependent Extension Models. *Water Science and Technology: Water Supply*, 13(4), 1161-1166.

Seyoum, A.G. and Tanyimboh, T.T. (2014) Pressure dependent network water quality modelling. *Proceedings of ICE: Water Management*, 167(6), 342-355.

Siew, C., Tanyimboh, T.T. and **Seyoum, A.G.** (2014) Assessment of Penalty-Free Multi-Objective Evolutionary Optimization Approach for the Design and Rehabilitation of Water Distribution Systems. *Water Resources Management,* 28(2), 373-389.

Peer-Reviewed International Conference Papers

Seyoum, A.G., Tanyimboh, T. T. and Siew, C. (2011) Comparison of demand driven and pressure dependent hydraulic approaches for modelling water quality in distribution networks. *Computing and Control for the Water Industry 2011: Urban Water Management - Challenges and Opportunities*, Exeter, UK, 5-7 September 2011.

Seyoum, A.G., Tanyimboh, T.T. and Siew, C. (2012) Assessment of Water Quality Modelling Capabilities of EPANET Multi-species and Pressure Dependent Extension Models. In *Proceedings of the International Water Association World Water Congress*, Busan, South Korea, 16-21 September 2012.

Seyoum, A.G. and Tanyimboh, T.T. (2013) Pressure-dependent Multiple-species Water Quality Modelling of Water Distribution Networks. In *Proceedings of the 8th International Conference of European Water Resources Association,* Porto, Portugal, 26-29 June 2013.

Seyoum, A.G. and Tanyimboh, T.T. (2014) Application of pressure-dependent EPANET extension. In *Proceedings of the 11th International Conference on Hydroinformatics*, New York City, USA, 16-21 August 2014.

Seyoum, A.G., Tanyimboh, T.T. and Siew, C. (2014) Optimal tank design and operation to improve water quality in distribution systems. In *Proceedings of the 11th International Conference on Hydroinformatics*, New York City, USA, 16-21 August 2014.

Seyoum, A.G., Tanyimboh, T.T. and Siew, C. (2015) Practical application of penalty-free evolutionary multi-objective optimisation of water distribution systems. In *Proceedings of t*he *9th World Congress of European Water Resources Association*, Istanbul, Turkey, 10-13 June 2015.