

Optimal reliability-based design of bulk water supply systems

Ching-Chiao Chang

Dissertation submitted in partial fulfilment of the requirements for the degree of
Master of Science in Engineering

Supervisor: Associate Professor Jakobus E. van Zyl



Department of Civil Engineering
University of Cape Town

August 2011

Abstract

Bulk water supply systems are usually designed according to deterministic design guidelines. In South Africa, design guidelines specify that a bulk storage reservoir should have a storage capacity of 48 hours of annual average daily demand (AADD), and the feeder pipe a capacity of 1.5 times AADD (CSIR, 2000). Nel & Haarhoff (1996) proposed a stochastic analysis method that allowed the reliability of a reservoir to be estimated based on a Monte Carlo analysis of consumer demand, fire water demand and pipe failures. Van Zyl *et al.* (2008) developed this method further and proposed a design criterion of one failure in ten years under seasonal peak conditions.

In this study, a method for the optimal design of bulk water supply systems is proposed with the design variables being the configuration of the feeder pipe system, the feeder pipe diameters (i.e. capacity), and the size of the bulk storage reservoir. The stochastic analysis method is applied to determine a trade-off curve between system cost and reliability, from which the designer can select a suitable solution.

Optimisation of the bulk system was performed using the multi-objective genetic algorithm, NSGA-II. As Monte Carlo sampling can be computationally expensive, especially when large numbers of simulations are required in an optimisation exercise, a compression heuristic was implemented and refined to reduce the computational effort required of the stochastic simulation. Use of the compression heuristic instead of full Monte Carlo simulation in the reliability analysis achieved computational time savings of around 75% for the optimisation of a typical system.

Application of the optimisation model showed that it was able to successfully produce a set of Pareto-optimal solutions ranging from low reliability, low cost solutions to high reliability, high cost solutions. The proposed method was first applied to a typical system, resulting in an optimal reservoir size of approximately 22 h AADD and feeder pipe capacity of 2 times AADD. This solution achieved 9% savings in total system cost compared to the South African design guidelines. In addition, the optimal solution proved to have better reliability than one designed according to South African guidelines.

A sensitivity analysis demonstrated the effects of changing various system and stochastic parameters from typical to low and high values. The sensitivity results revealed that the length of the feeder pipe system has the greatest impact on both the cost and reliability of the bulk system. It was also found that a single feeder pipe is optimal in most cases, and that parallel feeder pipes are only optimal for short feeder pipe lengths.

The optimisation model is capable of narrowing down the search region to a handful of possible design solutions, and can thus be used by the engineer as a tool to assist with the design of the final system.

Table of Contents

Abstract	i
Table of Contents	ii
List of Figures	vi
List of Tables	viii
Acknowledgements	ix
1 Introduction	1
1.1 Background and motivation for research	1
1.2 Research scope, limitations and assumptions	2
1.3 Research goal and objectives	4
1.4 Dissertation layout	5
2 Literature Review	6
2.1 Water distribution system design and optimisation	6
2.1.1 Single objective versus multi-objective optimisation	7
2.1.2 Optimisation techniques	7
2.1.3 Genetic algorithms	8
2.2 Cost considerations	10
2.2.1 Cost modelling	11
2.2.2 Reliability of cost data and estimates	12
2.3 Reliability considerations	12
2.4 Stochastic analysis	14
2.4.1 Monte Carlo simulation for reliability analysis	15
2.4.2 Compression heuristic	15
2.4.3 Sampling techniques	16
2.5 Bulk water supply system and storage reservoir design	19
3 Method Overview	21
3.1 The compression heuristic	23
3.1.1 Purpose and motivation for developing the compression heuristic	23
3.1.2 Further development of the compression heuristic	23
3.2 Investigation of Latin Hypercube sampling	25
3.2.1 Determining the number of strata required	25
3.2.2 Convergence criterion for the simulation	26
3.2.3 Comparison of Latin Hypercube sampling to Monte Carlo sampling	26
4 Pre-Run	29
4.1 Introduction	29
4.1.1 Description of the pre-run algorithm	29
4.2 Representing the system variables	31

4.2.1	Input variables	31
4.2.2	Output variables	33
4.3	Recording demand failures	34
4.3.1	Exclusion of failure duration as a sizing parameter	35
4.4	Generating the reservoir level curve lookup table	35
4.5	Conditions for termination of the simulation	37
4.5.1	Convergence theorems and rules-of-thumb	37
4.5.2	Convergence criteria used in previous stochastic models	38
4.5.3	Termination criteria for this study	38
4.5.4	Exclusion of reservoir level convergence as a termination criterion	40
4.5.5	Summary of termination conditions	42
4.6	Curve fitting to demands failure frequency data	42
4.6.1	Argument against using a linear fit	43
4.6.2	Comparison of the two curve fitting techniques	44
5	Events Run	46
5.1	Introduction	46
5.1.1	Description of the events algorithm	46
5.2	Input variables and initialisation	49
5.3	Running an event simulation	54
5.3.1	Generating fire events	54
5.3.2	Generating pipe failure events	55
5.3.3	Determining the next event	56
5.3.4	Event simulation	56
5.3.5	Conditions for ending the event simulation	56
5.4	Obtaining the initial reservoir level	57
5.5	Effect of pipe configuration	59
5.5.1	Pipe configuration and reliability	59
5.5.2	Investigation of the effect of pipe configuration	59
5.6	Events failure frequency	67
5.6.1	Recording events failures	67
5.6.2	Effect of events on the system	67
5.7	Terminating conditions for the events run	68
5.8	Calculating total failure frequency for the system	69
5.8.1	Use of mean time-to-failure to determine failure frequency	69
5.8.2	Calculating the total failure frequency	70
6	Optimisation	71
6.1	Introduction	71
6.2	Defining the problem range	72
6.2.1	Range for the pipe configuration	72
6.2.2	Range for the pipe diameter	73
6.2.3	Range for the storage reservoir size	73

6.3	Population Initialisation	73
6.4	Evaluating the objective functions	74
6.4.1	Cost evaluation	74
6.4.2	Reliability evaluation	77
6.5	Sorting the population	78
6.5.1	Nondomination sorting	78
6.5.2	Crowding distance	78
6.6	The evolution process	79
6.6.1	Selection	80
6.6.2	Crossover	80
6.6.3	Mutation	83
6.6.4	Recombination and selection	85
6.7	Setting the population size and number of generations	85
6.7.1	Population sizing	85
6.7.2	Number of generations	86
7	Results and Discussion	88
7.1	Comparison of stochastic simulation results	88
7.2	Comparison of stochastic simulation times	92
7.2.1	Single stochastic simulation	92
7.2.2	Stochastic simulation of a large number of systems together	92
7.3	Application of the reliability-based optimisation model	94
7.3.1	General observations	99
7.3.2	Summary	105
7.4	Sensitivity analysis	105
7.4.1	Effect of length	106
7.4.2	Effect of head	109
7.4.3	Effect of demand pattern	111
7.4.4	Effect of fire rate	114
7.4.5	Effect of pipe failure rate	115
7.4.6	Effect of failure duration	118
7.4.7	Comparison of sensitivity results	120
7.4.8	Computational time savings	125
8	Conclusions	126
8.1	Summary of work	126
8.2	Main findings	127
8.2.1	Compression heuristic	127
8.2.2	Optimisation model	128
8.3	Contributions to research	129
8.4	Recommendations for future work	130
8.4.1	Speeding up the stochastic model	130
8.4.2	Expansion of the stochastic model	131

8.4.3	Additional optimisation model capabilities	131
8.4.4	Additional objectives for the optimisation problem	132
References		133
Appendix A: Comparison of reservoir level distributions		140
Appendix B: Comparison of stochastic simulation times		147
Appendix C: Sensitivity analysis results		151

List of Figures

Figure 1.1 Example network layout	2
Figure 3.1 Flowchart of proposed method	22
Figure 3.2 Comparison of failure frequency for Latin Hypercube sampling and Monte Carlo sampling	25
Figure 4.1 Flowchart for the demands-only pre-run	30
Figure 4.2 Typical grid of SR - RC matrix for failure results	33
Figure 4.3 Matrix for reservoir level results	34
Figure 4.4 Reservoir level distribution at Sunday 04h00 for the example system	36
Figure 4.5 Reservoir level lookup table at Sunday 04h00 for the example system	37
Figure 4.6 Curve fitting using the log transform and nonlinear least squares method on (a) all available data, and (b) the converged data	44
Figure 5.1 Flowchart for the events run	48
Figure 5.2 Reservoir level distribution curve for example, $SR = 1.2$ and $RC = 6$ h	52
Figure 5.3 Reservoir level distribution curve for example, $SR = 1.15$ and $RC = 8$ h	53
Figure 5.4 Reservoir level distribution curve for example, $SR = 1.15$ and $RC = 6$ h	54
Figure 5.5 Reservoir level distribution curve for example	58
Figure 5.6 Configuration (1) – Single pipe connecting the source to the storage reservoir	60
Figure 5.7 Configuration (2) – Two parallel pipes connecting the source to the storage reservoir	60
Figure 5.8 Configuration (3) – Three parallel pipes connecting the source to the storage reservoir	60
Figure 5.9 Configuration (4) – Two parallel pipes connecting the source to the storage reservoir, with one interconnection dividing the pipes into two equal lengths	60
Figure 5.10 Configuration (5) – Three parallel pipes connecting the source to the storage reservoir, with one interconnection dividing the pipes into two equal lengths	61
Figure 5.11 Configuration (6) – Two parallel pipes connecting the source to the storage reservoir, with two interconnections dividing the pipes into three equal lengths	61
Figure 5.12 Configuration (7) – Three parallel pipes connecting the source to the storage reservoir, with two interconnections dividing the pipes into three equal lengths	61
Figure 5.13 Effect of a pipe failure in configuration (1) with $h_f = 30$ m, $C_H = 120$, $L = 1000$ m, $D = 0.207$ m	62
Figure 5.14 Effect of a pipe failure in configuration (2) with $h_f = 30$ m, $C_H = 120$, $L = 1000$ m, $D = 0.207$ m for each pipe	63
Figure 5.15 Effect of a pipe failure in configuration (3) with $h_f = 30$ m, $C_H = 120$, $L = 1000$ m, $D = 0.207$ m for each pipe	63
Figure 5.16 Effect of a pipe failure in configuration (4) with $C_H = 120$, $L = 500$ m, $D = 0.207$ m for each pipe	64
Figure 5.17 Effect of a pipe failure in configuration (5) with $C_H = 120$, $L = 500$ m, $D = 0.207$ m for each pipe	65
Figure 5.18 Effect of a pipe failure in configuration (6) with $C_H = 120$, $L = 333$ m, $D = 0.207$ m for each pipe	65

Figure 5.19 Effect of a pipe failure in configuration (7) with $C_H = 120$, $L = 333$ m, $D = 0.207$ m for each pipe	66
Figure 6.1 Optimisation process using multi-objective genetic algorithm NSGA-II (Wu <i>et al.</i> , 2010)	71
Figure 6.2 Structure of each solution	73
Figure 6.3 Crowding distance calculation - points marked in filled circles are solutions of the same nondominated front (Deb <i>et al.</i> , 2002)	79
Figure 6.4 Comparison of distribution indices $\eta_c = 2; 10; 20$	83
Figure 6.5 Comparison of distribution indices $\eta_c = 2; 10; 20$ (close up)	83
Figure 6.6 Solution set after every ten generations	87
Figure 7.1 Comparison of failure results for system 1 ($SR = 1.2$; $RC = 3$ h)	89
Figure 7.2 Comparison of failure results for system 2 ($SR = 1.2$; $RC = 12$ h)	90
Figure 7.3 Comparison of failure results for system 3 ($SR = 1.5$; $RC = 3$ h)	91
Figure 7.4 Comparison of failure results for system 4 ($SR = 1.5$; $RC = 12$ h)	91
Figure 7.5 Comparison of simulation times for an increasing number of simulations	93
Figure 7.6 Pareto front of optimal solutions for the test system	96
Figure 7.7 Pareto front of optimal solutions for the test system which meet the design criterion	96
Figure 7.8 Pareto front for the test system including solutions with three pipes in parallel (close-up)	102
Figure 7.9 Pareto front for the test system including guidelines-based designs	103
Figure 7.10 Pareto fronts for systems with different pipe lengths	107
Figure 7.11 Pareto fronts for systems with different heads	110
Figure 7.12 Hourly demand patterns	112
Figure 7.13 Pareto fronts for systems with different peak hourly factors	113
Figure 7.14 Pareto fronts for different fire rates	115
Figure 7.15 Pareto fronts for different pipe failure rates	117
Figure 7.16 Pareto fronts for different pipe failure durations	118
Figure 7.17 Solutions meeting the design criterion for the different pipe failure durations	119
Figure 7.18 Impact of sensitivity parameters on system cost	120
Figure 7.19 Impact of sensitivity parameters on system cost (excluding length)	122
Figure 7.20 Trade-off of supply ratio against reservoir capacity for optimal solutions	123
Figure 7.21 Relative costs of individual components for each optimal solution	124

List of Tables

Table 2.1 Reservoir design criteria for low-density residential areas (van Zyl <i>et al.</i> , 2008)	19
Table 3.1 Comparison of results for Latin Hypercube and full Monte Carlo sampling	27
Table 3.2 Comparison of simulation times for Latin Hypercube sampling and full Monte Carlo sampling	28
Table 4.1 Example of designer input values	31
Table 4.2 Supply inflow rate and supply ratios calculated for the illustration example, indicating which diameters are eliminated	32
Table 4.3 Percentage difference of failure frequencies	41
Table 4.4 Percentage difference - reservoir level lookup table ($SR = 0.8$ & $RC = 8h$)	41
Table 4.5 Percentage difference - reservoir level lookup table ($SR = 0.9$ & $RC = 16h$)	42
Table 5.1 Full reservoir fraction for example	50
Table 5.2 Reservoir level distribution for example, $SR = 1.1$ and $RC = 4$ h	50
Table 5.3 Reservoir level distribution for example, $SR = 1.1$ and $RC = 8$ h	50
Table 5.4 Reservoir level distribution for example, $SR = 1.2$ and $RC = 4$ h	51
Table 5.5 Reservoir level distribution for example, $SR = 1.2$ and $RC = 8$ h	51
Table 5.6 Summary of remaining supply inflow to the storage reservoir as a percentage for each pipe configuration	66
Table 6.1 Service reservoir cost	75
Table 7.1 Average simulation times and simulated hours for each system	92
Table 7.2 Comparison of simulation times for a large number of simulations	93
Table 7.3 Range of available pipe diameters	94
Table 7.4 Supply ratios for different combinations of pipe configuration and diameter	95
Table 7.5 Demands failure frequency results from the pre-run	95
Table 7.6 Full reservoir fraction results from the pre-run	95
Table 7.7 Summary of final solutions for the test system	97
Table 7.8 Summary of costs for the test system solutions	98
Table 7.9 Costs for systems designed according to South African deterministic guidelines	102
Table 7.10 Systems designed according to South African deterministic guidelines	103
Table 7.11 Sensitivity analysis parameters	106
Table 7.12 Pipe diameter selection for each supply pipe system length	107
Table 7.13 Comparison of single solution for the different pipe lengths	109
Table 7.14 Pipe diameter selection for each system head	110
Table 7.15 Comparison of single solution for the different system heads	111
Table 7.16 Comparison the same solution with different peak hourly factors	113
Table 7.17 Comparison of single solution for the different hourly demand patterns	114
Table 7.18 Comparison of single solution for different fire rates	115
Table 7.19 Average failure rates for different pipeline materials (Nel, 2009)	116
Table 7.20 Comparison of single solution for different pipe failure rates	117
Table 7.21 Comparison of single solution for different pipe failure durations	119
Table 7.22 Near optimal solutions satisfying design criterion for each sensitivity parameter	121

Acknowledgements

I would like to express my appreciation for the guidance and assistance that I have received from:

- My supervisor, A/Prof Kobus van Zyl, without whom I would be completely lost throughout this research. Thanks to him for always having his door open, for being so enthusiastic about the research topic and for always giving me the opportunity to expand my knowledge in the field of water distribution systems.
- My office mates, Kirsty Carden, Sophia Pan and Nangolo Ashipala, for providing encouragement and advice.
- My family for being so supportive, and especially Paul for being a sounding board and for helping me at times when I hit a wall with coding.

Plagiarism Declaration

I know the meaning of plagiarism and declare that all of the work in the document, save for that which is properly acknowledged, is my own.

Signed by candidate

Ching-Chiao Chang
August 2011

1 Introduction

1.1 Background and motivation for research

The design of water distribution systems usually involves the optimisation of pipe diameters in a network to find the least-cost design. A least-cost design does not always guarantee the best solution as the system may not be reliable; therefore reliability needs to be considered as an objective in water distribution system optimisation to ensure customer satisfaction. Water distribution systems must be both reliable and cost-effective as pipe networks form a major part of urban infrastructure. Thus, the aim of water distribution system design optimisation involves a trade-off between the conflicting objectives of minimisation of cost and maximisation of reliability.

A significant amount of research has been done on the design and optimisation of water distribution systems, focusing mainly on pipe networks, but also including pump scheduling, and storage reservoir design. In contrast, there is little literature on the optimisation of bulk water supply systems. Bulk systems are usually designed according to deterministic guidelines. For example, South African design guidelines specify that the storage reservoir should have a storage capacity of 48 hours of annual average daily demand (AADD), while the capacity of the supply mains to the reservoir should not be less than 1.5 times AADD for the area served by the reservoir (CSIR, 2000). These guidelines are likely to be conservative and may lead to oversized storage reservoirs which would have a high reliability, but also a high cost. It is necessary to strike a balance between these two objectives: cost and reliability. Hence, a niche exists for developing a model to optimise the design of a bulk supply system so that a pre-specified reliability can be attained at minimum cost.

The reliability of a bulk supply system can be defined in terms of the reliability of its storage reservoir, as consumers will only notice a service interruption if the storage reservoir has failed, i.e. run dry. Work on a stochastic model for the reliability analysis of bulk systems was first started by Nel & Haarhoff (1996). Since then, research on the model has been ongoing (van Zyl & Haarhoff, 1999; Haarhoff & van Zyl, 2002; van Zyl *et al.*, 2008). An advantage of stochastic simulation is that it is more realistic since it is able to include uncertainty and randomness of factors (Yang *et al.*, 1996). On the other hand, stochastic simulation has the drawback of requiring excessive computational resources. However, with current advances in technology, this is becoming less of a problem and stochastic simulation is steadily gaining feasibility as a modelling technique. Nevertheless, an attempt has been made to reduce the computational effort required by the stochastic analysis model developed by van Zyl *et al.* (2008) through the use of a compression heuristic (Chang & van Zyl, 2010). It is recognised that the compression heuristic could be further developed and refined, which is one of the research objectives in this study.

Engineers have traditionally designed water distribution systems by trial-and-error using engineering judgment. In doing so, the main concern has been minimising the cost of the

system, while superficial thought was given to reliability considerations. Using a multi-objective optimisation model can rectify this practice and provide engineers with a range of design options that are based on reliability analysis as well as cost. The solutions that are produced from this optimisation model can be used to assist with the selection of the supply pipe configuration, pipe diameter, and reservoir capacity in the final design of the bulk system. The engineer can then decide how much reliability he is willing to sacrifice for the sake of a cheaper system, that is, he can decide how much more money he is willing to spend to produce a more reliable system.

1.2 Research scope, limitations and assumptions

The aim of this project was to develop a method to find the optimal design of a reliable and cost-effective bulk water supply system by the optimal selection of its pipe configuration, pipe diameter and reservoir size. A multi-objective genetic algorithm was used to determine the optimal values for the design variables such that the system reliability is maximised and cost is minimised. Throughout this study, the reliability of the bulk system was defined in terms of the failure frequency of its storage reservoir. Therefore, the objective of maximising reliability was equated to minimising failure frequency. The cost for the system was calculated using existing cost functions for each component, while a stochastic analysis model (van Zyl *et al.*, 2008) was used to determine the failure frequency of the system. To ensure that it would be practical to use the stochastic analysis method within the optimisation run, a compression heuristic was implemented.

The proposed design method was applied to a simple, yet typical, bulk water supply system, shown in the network layout in Figure 1.1. A supply pipe system provides water to the storage reservoir from a supply source, and a second pipe links the storage reservoir to the users.

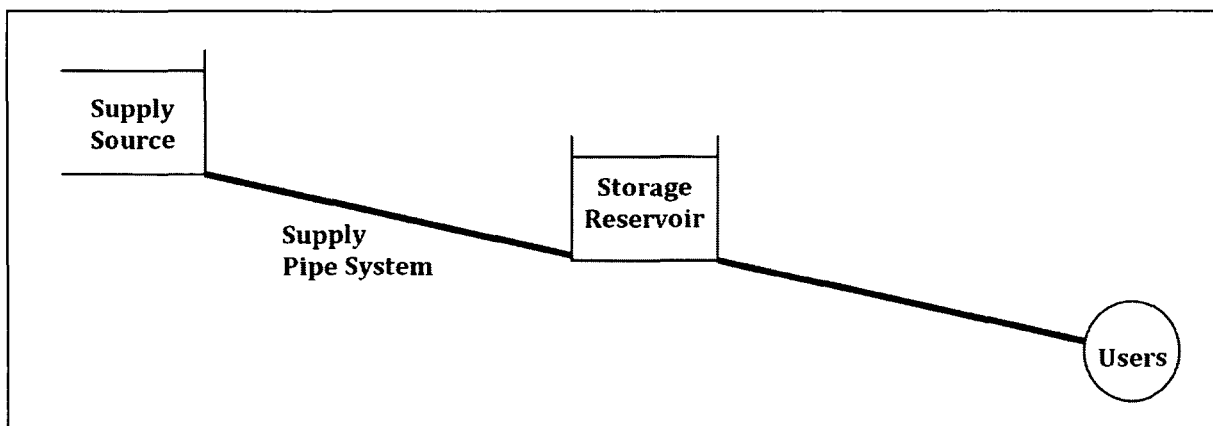


Figure 1.1 Example network layout

The bulk water supply system investigated was a gravity system; therefore it was assumed that the source is at a higher elevation than the reservoir, which is at a fixed hydraulic head. Since a pump system was not considered in this study, no pumps were required to transport

water from the source to the reservoir. This eliminated the need for inclusion of pump reliability, as well as operational and energy costs from the optimisation model. As a result, only capital costs were considered in the analysis.

It was assumed that the location of the source, the storage reservoir and thus the supply system was fixed, and only the pipe configuration and diameter of the supply system were varied, in addition to the size of the reservoir. The optimisation model was limited to seven different pipe configurations, ranging from a single pipeline to three pipes in parallel, with zero to two interconnections. The pipe diameter was selected from a discrete range of commercially available pipe sizes, while parallel pipes were assumed to have the same diameter. Parallel pipes were considered as being laid in separate trenches in this study, thus the cost of a supply system with two pipes in parallel is double that of a system with a single pipeline.

Reliability of the bulk system was defined in terms of the failure frequency of the storage reservoir, and not the duration of the reservoir failures. The stochastic model analyses the system for seasonal peak conditions. Hence, it models the behaviour at the most critical time in a year and estimates the minimum reliability of the system, instead of the annual average. A seasonal peak factor of 1.49 was estimated based on analyses, i.e. the seasonal peak demand is 1.49 (say 1.5) times the annual average daily demand. For the demand side of the storage reservoir, a known and stable demand pattern was assumed (i.e. no long term growth), whereas the reservoir was assumed to be supplied at a constant flow rate.

The stochastic analysis model is comprised of three components: consumer demand, fire events and pipe failure events. The consumer demand model consists of average demand, cyclic patterns, persistence and randomness (van Zyl *et al.*, 2008). Fire events are generated using an exponential distribution, while the fire flow and fire duration are modelled with a lognormal distribution. Similarly, pipe failure events and pipe failure duration are modelled with an exponential distribution and lognormal distribution respectively.

The reliability of the supply pipe system, which also influences the reliability of the storage reservoir, can be defined in terms of its ability to provide a constant and uninterrupted supply (van Zyl *et al.*, 2008). Failures of the water source, purification tanks, pumps and pipes can all cause the interruption of the supply inflow to the reservoir. In this study, it was assumed that the supply was limited only by the capacity of the pipes and source failures were not considered; while purification tanks, pumps and other storage reservoirs were ignored. Only pipe failures in the feeder pipe system supplying water into the storage reservoir were considered as they are most commonly dealt with in the literature. In the stochastic model, the supply system experiences only one pipe failure at a time, i.e. no pipe failures happen simultaneously. In addition, the effects of material, diameter and age on the pipe failure rate were not included. Modelling of pipe failures was limited to the pipe system supplying the inflow into the storage reservoir. It was considered conservative to not include the effect of pipe failures on the pipe system linking the storage reservoir to the consumers as a pipe

failure would only reduce or discontinue the demand from the reservoir for the duration of the failure (van Zyl *et al.*, 2008).

Van Zyl *et al.* (2008) recognise that two possible extreme values for the combination of supply and reservoir capacities can be estimated. No balancing storage would be required if the supply capacity matches the maximum demand of the system. In contrast, no reservoir would be large enough to provide a sustained service if the supply capacity is less than the average demand. The optimisation model aims to find intermediate solutions for the design problem.

1.3 Research goal and objectives

The goal of this study was to develop an optimisation model that will produce a trade-off curve of Pareto-optimal solutions for the design of a bulk water supply system based on cost and reliability analyses. The optimisation model used a multi-objective genetic algorithm, which incorporated cost modelling and stochastic analysis techniques to evaluate the respective cost and reliability objectives.

To achieve this research goal, a number of objectives are set out as follows:

- To compile a literature review on the state of knowledge regarding water distribution system design and optimisation, stochastic analysis of water supply systems, storage reservoir design and optimisation techniques, reliability and cost considerations for water distribution systems, specifically bulk water supply systems.
- To reduce the computational effort of the stochastic model developed by van Zyl *et al.* (2008) for the reliability analysis of municipal storage reservoirs, through additional work on the compression heuristic (Chang & van Zyl, 2010) to refine the model so that it ensures that accurate and reliable results are obtained quickly. Applying the compression heuristic will allow the practical use of the stochastic analysis model within the optimisation run.
- To test and verify the refined compression heuristic by comparing it to the original stochastic model which uses full Monte Carlo sampling, based on the correlation of the results and improvement in simulation times.
- To incorporate the stochastic model into an existing multi-objective genetic algorithm. Application of the stochastic model within the optimisation run was not an attempt to refine or improve the relationships defined by others (van Zyl & Haarhoff, 1999; Haarhoff & van Zyl, 2002), but rather to find optimal combinations of the design variables, based on both cost and reliability considerations.
- To apply the optimisation model to a typical system and analyse the resulting set of design solutions.
- To compare the optimisation results to deterministic design guidelines, without any intention of improving or replacing these existing guidelines.

- To investigate the effect of different parameters on the optimisation model through a sensitivity analysis, and comparison to a previous study on risk-based optimal design (Vlok, 2010).
- To draw conclusions based on the observations of the typical system and the sensitivity analysis.
- To propose recommendations for improvements to the method and provide suggestions for further extensions of the model.

1.4 Dissertation layout

A brief description of the chapters in this dissertation is provided:

- Chapter 2 consists of a literature study of relevant research topics, namely water distribution system design and optimisation, cost and reliability considerations, and stochastic analysis. These research topics provide a background and understanding of the current state-of-art in these fields.
- The research methodology is described in Chapters 3 to 6. Chapter 3 provides a basic overview of the methodology used in the study. This includes a brief description and motivation for the compression heuristic method (Chang & van Zyl, 2010), and an investigation of Latin Hypercube sampling. Chapter 4 and Chapter 5 deal with the two components of the compression heuristic, which are the pre-run and events run respectively. The optimisation model is described in Chapter 6.
- Following the methodology, the test results and discussion from the application of the reliability-based optimisation model can be found in Chapter 7. The results from a sensitivity analysis are also provided.
- In Chapter 8, the study is summarised and conclusions are drawn from the results and discussion in the previous chapter. Finally, proposals for improvements and expansions to both the stochastic method and optimisation model are suggested.

2 Literature Review

As previously mentioned, the aim of this project was to optimise the design of bulk water supply systems based on cost and reliability considerations. Before the proposed method to achieve this aim is described, it is necessary to first gain an understanding of the relevant research topics to this project. In this chapter, a review of the current state-of-knowledge on water distribution system design and optimisation, cost and reliability considerations, stochastic analysis, and bulk supply system and storage reservoir design is provided. In this review, established and recent work on each topic is discussed, research gaps are identified, and possible tools for use in this study are explored.

2.1 Water distribution system design and optimisation

Traditionally, water distribution systems have been designed by trial-and-error using basic hydraulic equations. Within the last few decades, many researchers have developed optimisation models as alternative tools for finding the optimal design of water distribution systems. Although a water distribution system design involves the design of components such as storage reservoirs and pumps, Walski (1995) notes that in many cases, optimisation models have equated pipe sizing to water distribution system design. This suggests that the objective of water distribution system design is to find pipe diameter sizes subject to hydraulic constraints. Basic pipe network optimisation considers the pipe layout to be fixed because it is usually constrained by the layout of roads.

An early definition of the optimal design of water distribution systems is provided as that which meets demands at least cost (Goulter, 1992). This definition has been used frequently in early literature as shown by Lansey & Mays (1989) who report that several optimisation models have been developed with cost minimisation as the single objective of water distribution system design.

In contrast, Walski (2001) states that there are a range of different objectives for water distribution system design. These objectives could include reliability and water quality. An alternative definition which more accurately reflects the aim of water distribution engineers is that optimal design is equivalent to the maximisation of the net benefits (benefits minus costs) of water distribution systems (Walski, 2001).

The main challenge of optimisation models is to move away from cost minimisation towards methods that maximise net benefits (Walski, 2001). Over the last decade and a half, this challenge has been addressed by many researchers who are moving away from single objective least-cost optimisation towards multi-objective optimisation.

2.1.1 Single objective versus multi-objective optimisation

In many previous studies, cost minimisation was the main objective in water distribution design. However, there is a problem with this single objective as modellers become too fixated with finding the least cost solution, resulting in the elimination of system capacity (and reduction of project benefits) to minimise cost (Walski, 2001). Farmani *et al.* (2005) agree that optimisation is inclined to reduce costs by decreasing the diameter of, or eliminating, some pipes, resulting in insufficient system capacity to deal with pipe breaks or demands that are greater than design values without violating required performance levels.

It is clear that water distribution system optimisation based only on cost minimisation is limited in value for practical design problems. Therefore multi-objective optimisation, which involves a trade-off between costs and benefits, should be used instead. Numerous objectives can be considered, such as capital and operational costs, reliability and water quality. Multi-objective optimal design of water distribution systems typically involves finding the maximum reliability at minimum cost.

2.1.2 Optimisation techniques

Various techniques have been applied to the problem of water distribution system optimisation. In this section, some of the past studies related to optimal design of water distribution systems are presented.

Some of the earlier studies on water distribution system design optimisation used linear programming (Alperovits & Shamir, 1977; Quindry *et al.*, 1981). The linear programming method takes a nonlinear problem and converts it into a linear one by taking pipe lengths as decision variables. However, stretches of pipe between adjacent nodes are divided into different segments and then the lengths of the segments are optimised to minimise the cost.

A nonlinear programming method (Su *et al.*, 1987; Duan *et al.*, 1990) has also been used for finding the optimal pipe diameter combinations. Constraints can be included explicitly in the model and costs can be expressed as any nonlinear function of pipe diameter and length. However, nonlinear programming has several limitations. In the method, the pipe diameters are continuous variables which are replaced with the nearest commercially available pipe sizes. This requires rounding off of the solution, which leads to a suboptimal solution. Nonlinear programming is only able to identify the local optimum. In addition, there is a limitation on the number of constraints used in the model, and thus the size of the network that can be handled (Simpson *et al.*, 1994).

Much work has been done in the literature on the application of genetic algorithms to the problem of optimal design of water distribution systems (Simpson *et al.*, 1994; Savic & Walters, 1997; Prasad & Park, 2004; Kadu *et al.*, 2008). Genetic algorithms have several advantages over other mathematical optimisation techniques (Goldberg, 1989). In the field of

water distribution system optimisation, genetic algorithms have been shown to overcome the limitations of traditional optimisation methods such as linear and nonlinear programming. Genetic algorithms are conceptually simple to understand. They have the advantage of working directly with discrete pipe sizes, and therefore no rounding of solutions is required. Another advantage is that they have global sampling capability. Therefore, the possibility of trial solutions becoming entrapped in local minima is reduced and dependency on a starting point is avoided (Savic & Walters, 1997). Finally, genetic algorithms have the benefit of being able to identify a set of near-optimal solutions, e.g. close to the minimum cost. These solutions may have different designs that can be compared in terms of other important objectives such as redundancy or reliability (Simpson *et al.*, 1994).

A benchmark problem that is often used to compare different optimisation techniques is the New York City water supply tunnels, which was first used in case study by Schaake & Lai (1969). Numerous studies have been conducted on the pipe optimisation problem, with each optimisation technique trying to improve on others to find the minimum cost. The optimisation techniques use different diameter design methods. Continuous diameter design was used by Schaake & Lai (1969), Quindry *et al.* (1981), Bhave (1985) and Fujiwara & Khang (1990). It was previously mentioned that continuous diameters require rounding off which leads to suboptimal solutions in the final design. Kessler (1988) used a split pipe design, which divides a continuous diameter into two partial lengths of adjacent discrete diameters to create a hydraulically equivalent pipe. The split pipe design is problematic as it is not realistic or practicable.

In contrast, Savic & Walters (1997) stated that for a more realistic and practical design, a single discrete pipe diameter between the adjacent nodes is required. Dandy *et al.* (1996) have shown that a genetic algorithm achieves a feasible discrete pipe solution with the lowest cost of \$38.80 million. Similarly, other studies have shown the efficiency of genetic algorithms applied to other benchmark problems, such as the Hanoi network and Two-loop network problem (Savic & Walters, 1997). A genetic algorithm is clearly an effective technique which can be used to solve optimisation problems.

2.1.3 Genetic algorithms

A genetic algorithm is a heuristic optimisation technique based on the mechanics of natural selection and genetics (Holland, 1975; Goldberg, 1989). It selects, combines and manipulates possible solutions in the same way that nature permits survival, reproduction and the combination of chromosomes in search of the best adaptation (Murphy & Simpson, 1992).

The first step in a genetic algorithm involves the generation of an initial population or set of solutions. The population consists of individuals or solutions which are randomly generated. Each solution to the problem is represented by a chromosome, which is a set of decision variables for a particular combination of decision variables, such as pipe sizes. The individuals in the population are allowed to evolve over a number of generations. In each

generation, the fitness of each individual is evaluated, i.e. each individual is measured to determine how good it is. Fitness evaluation is done by computing the value of the objective function. Based on the fitness value, individuals are selected and recombined to produce offspring for a new generation. Individuals with a high fitness value have a higher probability of being selected. The recombination of individuals is known as crossover. The final operator is mutation, which randomly alters a bit (or gene) in an individual with a small probability. The mutation operator is used to maintain diversity of the individuals, as it ensures that potentially useful genetic material is not lost. Once a new population has been generated, the fitness is evaluated. Generation of new populations continues until a stopping criterion has been met.

As previously discussed, WDS optimisation is a multi-objective problem, where several objectives such as cost, reliability, redundancy and water quality can be considered. In a multi-objective optimisation problem, a set of solutions called Pareto-optimal solutions exist. Each solution in the Pareto-optimal set is not dominated by any other solution. This means that it is not possible to improve one criterion without making another criterion worse. A solution A is thus nondominated if there is no other solution B in the set that optimises all criteria better than solution A.

Many single objective algorithms (e.g. linear programming and nonlinear programming) have been applied to solve these multi-objective water distribution systems problems. However, these algorithms have several limitations. Atiquzzaman *et al.* (2006) state that a problem with single objective algorithms is that they must change a multi-objective formulation into a single objective one using penalty factors. The choice of the penalty term is critical as it can distort the objective function. More precisely, a very big penalty coefficient causes the search to drift towards regions of local optima, whereas a very small penalty coefficient results in inferior solutions (Liong *et al.*, 2004). Another problem with the use of single objective algorithms is that they do not produce the required trade-off curve between costs and benefits to help the water engineer in better decision making (Atiquzzaman *et al.*, 2006).

In contrast, multi-objective genetic algorithms are able to deal with several objectives and both linear and nonlinear constraints. In addition, no penalty coefficients are required (Farmani *et al.*, 2004). Multi-objective genetic algorithms produce a Pareto front of the Pareto-optimal solutions. Prasad & Park (2004) emphasise that optimisation can only assist the engineer in the decision-making process, and that engineering judgment and experience are needed to provide a practicable solution. The Pareto front provides the engineer with more flexibility in the selection of a practicable solution (Prasad & Park 2004).

Several multi-objective genetic algorithms have been proposed in the literature (Goldberg, 1989; Fonseca & Fleming, 1993; Srinivas & Deb, 1995; Halhal *et al.*, 1997). Single objective and multi-objective genetic algorithms are similar, except for the allocation of fitness and the representation of the set of Pareto solutions (Dandy & Engelhardt, 2006).

There have been several applications of multi-objective genetic algorithms to WDS problems. Prasad & Park (2004) applied the nondominated sorting genetic algorithm (NSGA) (Srinivas & Deb, 1994) in the design of a water distribution network with minimisation of cost and maximization of reliability. Khu & Keedwell (2004) found an optimal rehabilitation alternative using NSGA-II. Atiquzzaman *et al.* (2006) also used the NSGA-II to find the optimal solution for a 2-looped network.

The goals of Pareto multi-objective optimisation are: (1) closeness to the Pareto front; and (2) diversity among solutions in each front. Farmani *et al.* (2003) compared the application of multi-objective genetic algorithm techniques in water distribution systems, which showed that NSGA-II performed better than other techniques in reaching the optimisation goals. NSGA-II was shown to outperform two other contemporary multi-objective evolutionary algorithms (Deb *et al.*, 2002), the Pareto-archived evolution strategy (PAES) (Knowles & Corne, 1999) and the strength-Pareto evolutionary algorithm (SPEA) (Zitzler & Thiele, 1998).

An important concept in NSGA-II is that of domination (Liong *et al.*, 2004). Consider 2 feasible solutions: $\mathbf{x}^{(1)}$ & $\mathbf{x}^{(2)}$,

1. $\mathbf{x}^{(1)}$ is better than $\mathbf{x}^{(2)}$ if $\mathbf{x}^{(1)}$ is no worse than $\mathbf{x}^{(2)}$ in all objectives & at least one objective is better;
2. $\mathbf{x}^{(2)}$ is better than $\mathbf{x}^{(1)}$ if $\mathbf{x}^{(2)}$ is no worse than $\mathbf{x}^{(1)}$ in all objectives & at least one objective is better;
3. Otherwise, $\mathbf{x}^{(1)}$ & $\mathbf{x}^{(2)}$ are equally good.

Farmani *et al.* (2005) provides a comprehensive description of the NSGA-II (Deb *et al.*, 2002) approach. NSGA-II is a fast, non-dominated sorting approach. It randomly generates an initial population. Offspring solutions are created by selection, crossover and mutation operators. A selection operator creates a mating pool by combining parent and offspring populations and selecting the best solutions by considering fitness and spread. The next generation is populated starting with the best nondominated front, moving through the rest of the solutions until the population size is reached. If in the final stage there are more individuals in the nondominated front than there is available space, the individuals of the front are chosen based on the crowded distance-based niching strategy. The procedures are repeated until the maximum number of generations is reached. The final solution set is the Pareto optimal set.

2.2 Cost considerations

Water distribution system design cannot be based purely on functional requirements such as minimum pressure. Even more than half a century ago, Chilton (1949), as cited in Clark & Dorsey (1982), recognised that economic assessments are necessary so that alternatives that are not cost-effective can be eliminated, and research and engineering studies can be focussed on the designs that show the most potential. Hence, besides considering functional

requirements, the solution for the design of a water distribution system should be based on least-cost considerations (Swamee & Sharma, 2008).

A water distribution system consists of many components, which include pipes, pumps and pumping stations, storage reservoirs, and residential connections. Each of these components has its own capital costs. Apart from capital costs, the system components also have recurring operation and maintenance costs, for example, pumping stations have energy costs. It is advisable to calculate the life-cycle cost (overall cost) of the system over its lifespan so that a proper estimate of all the contributing expenses can be obtained. Since the capital and recurring cannot be added together, they can be combined through methods such as the capitalisation, annuity or net present value methods. Life-cycle cost analysis has not always been used in the optimisation of water distribution systems, and often only capital costs are considered, e.g. the Hanoi network (Fujiwara & Khang, 1990), and the New York Tunnels system (Schaake & Lai, 1969).

Various types of cost estimates exist; each of which is chosen based on the purpose and accuracy required at each stage of a project. Clark & Dorsey (1982) identified the following types of cost estimates:

- *Order-of-magnitude estimates* are used early in the study to provide a quick but rough estimate. Later estimates obtained from more comprehensive methods can be verified using this ballpark estimate.
- *Study estimates* are based on known costs relative to the capacity of the component or production quality. This can be estimated through a variety of models depending on the cost data.
- *Preliminary estimates* are obtained from the estimate of each individual item, and are therefore more accurate in comparison to study estimates. They are also used when cost data is unavailable to create a model for the study estimate.
- *Definitive estimates* need detailed information from engineering designs. These are highly accurate and therefore the preparation of this estimate is costly.
- *Detailed estimates* are similar to definitive estimates except they are site-specific. They require complete details of specifications and drawings. They can be used to control specific installations.

Costing of water distribution system designs requires a study estimate, since unit cost models are often used when evaluating and considering the feasibility of engineering designs.

2.2.1 Cost modelling

A large number of studies on cost modelling and analysis have been done over the years. Dickson (1972) produced cost curves based on actual construction costs of projects, while Lindsay & Walski (1982) and Walski (1985) calculated costs estimates for comparison of

facilities such as dams, pump stations, and open channels. Clark & Dorsey (1982) developed a cost model for treating drinking water.

In the field of water distribution, Clark & Males (1985) calculated the trade-offs between cost and quality of water in distribution systems, and Kim & Clark (1987, 1988) studied the general economics of water distribution systems. In particular, in South Africa a spreadsheet model of all the components of Rand Water's distribution network was developed (Fowler *et al.*, 1997), as well as a comprehensive econo-mathematical model of a water supply system (Barta & Rowse, 1998).

In addition, Swamee & Sharma (2008) developed practical cost functions specifically to model the unit costs of pumping stations, pipelines, storage reservoirs and residential connections. More detail on each cost function is provided in Section 6.4.

Vlok (2010) noted by examining the cost functions by Swamee & Sharma (2008) that unit cost models had not been developed for chambers that allowed interconnections between parallel pipes. Consequently, cost functions were developed for chambers that contain the necessary pipe work and valves required to connect two and three pipes in parallel (Vlok, 2010). These functions are also described in Section 6.4.

2.2.2 Reliability of cost data and estimates

Even though cost models can be developed based on cost data, it is necessary to consider how reliable the data is, and in turn how accurate the cost estimates produced from the cost models are. The reliability and accuracy of the cost estimates can be attributed to basic data availability, stage of the project development, experience of the analyst, and time spent on analysis (Clark & Dorsey, 1982). Although there may be thousands of possible solutions that vary by one or two percent, the costs of these solutions are accurate to approximately 20% (Haestad Methods, 2003). In design optimisation, it is important that the relative costs between these solutions are consistent so that the effects of this error are not severe.

2.3 Reliability considerations

While optimisation has focused predominantly on finding a least-cost design, researchers have gradually recognised the importance of including reliability considerations in the design of water distribution systems. The concept of reliability of a water distribution system is complex and multi-faceted; as a result, there is no universally accepted definition of reliability. For engineering systems, reliability is generally defined as the probability that a system performs its mission within specified limits for a given period of time in a specified environment (Bazovski, 1961, as cited in Cullinane *et al.*, 1992). With respect to water distribution systems, Farmani *et al.* (2005) suggests that reliability is the ability of the network to provide consumers with adequate and high-quality supply even under abnormal conditions.

An alternative definition is to define reliability in terms of network 'failure', which is an event in which the network is not able to provide adequate flow or pressure to meet demand (Goulter, 1992). This indicates that failures can represent a lack of reliability. Failures usually correlate to abnormal conditions such as extremely high demands or component failures. There are two types of failure: mechanical and hydraulic failure. Mechanical failure refers to the failure of system components, while hydraulic failure occurs when actual demands exceed design demand values for the system (Goulter, 1992).

The answer to the question 'what level of reliability is acceptable?' is as elusive as the definition of reliability. Goulter (1995) recognises that the specification of an acceptable level of reliability is dependent on the actual reliability measure used, and that it may be different for each scenario. Similar to the lack of an universal definition of water distribution system reliability, it was identified that there was no universal measure of reliability. Farmani *et al.* (2005) cites a paper by Walski *et al.* (1987) which suggests that a reliability measure should reflect the way in which water users are affected, such as the number of users with limited or no service and the length of time for which the event occurs. In addition, Goulter (1995) proposes that the reliability measure should include both the probability of the event causing failure or loss of service, as well as the magnitude of impact of the failure. This proposal relates to the basic principle of risk, which is often represented as the product of the likelihood and impact of an event.

In the literature, researchers have proposed a multitude of reliability measures and methods for quantifying reliability. Ostfeld (2004) provides a comprehensive review of these reliability assessment methods by categorising them into (1) connectivity/topological, (2) hydraulic, and (3) entropy as a surrogate measure.

Connectivity/topological reliability is associated with the probability of a given network remaining physically connected, keeping in mind the reliabilities of its components (Ostfeld, 2004). Examples of connectivity/topological reliability measures are connectivity and reachability (Wagner *et al.*, 1988a), and node pair reliability (Quimpo & Shamsi, 1991). These connectivity/topological measures fail to consider reliability from the consumers' point of view.

In contrast, hydraulic reliability is the probability of supplying the consumers' demands (Ostfeld, 2004). Measures of hydraulic reliability were investigated by Su *et al.* (1987), Wagner *et al.* (1988b), Bao & Mays (1990) and Cullinane *et al.* (1992). Xu & Goulter (1998) calculated the probability of meeting nodal demands at or above a minimum prescribed pressure, while Ostfeld (2001) calculated the probability of annual zero shortage. Since the system is vulnerable to random failures, the probability and impact of failure of each individual component must be considered to obtain the reliability of the entire system. Calculating this hydraulic reliability is a computationally impractical task, therefore it is predominantly evaluated using stochastic simulation (Ostfeld, 2004).

The third category is entropy, which is used as a surrogate measure of reliability. The basic principle is to measure the inherent redundancy of the network by using entropy measures of uncertainty to quantify the amount of information provided by a finite probability distribution (Ostfeld, 2004). Several researchers have made use of entropy as a reliability measures (Awumah *et al.*, 1990, 1991; Tanyimboh & Templeman, 1993, 2000). Within the last decade, a number of other reliability surrogate measures have been developed, such as the resilience index (Todini, 2000) and network resilience (Prasad & Park, 2004).

While three categories of reliability measures are identified by Ostfeld (2004), the methods used to quantify or evaluate these measures can be broadly classified into: analytical and simulation techniques. Analytical methods are based on deterministic modelling. These include heuristic approaches (Jacobs & Goulter, 1989), graph theory algorithms (Wagner *et al.*, 1988a), and minimum cut-set approaches (Quimpo & Shamsi, 1991; Su *et al.*, 1987). Simulation techniques typically involve the stochastic analysis of uncertainty, i.e. probabilistic modelling. Monte Carlo simulation is commonly used in these simulation methods (Wagner *et al.*, 1988b; Bao & Mays, 1990; Yang *et al.*, 1996). Analytical methods have the advantage of requiring less computational effort compared to stochastic simulation methods. However, simulation methods can consider a more extensive range of reliability measures, while providing a more realistic depiction of reliability.

2.4 Stochastic analysis

A water distribution system is a stochastic system as the failure of its components is random. According to Quimpo & Shamsi (1987), each component is either operative or it has failed. For example, a pipe is operative when it carries water and has failed when there is a pipe breakage.

Stochastic analysis is a technique whereby the deterministic and probabilistic parameters of a system are simulated to model the system's behaviour more accurately. This technique is relatively new to water distribution systems, even though it has long been used in fields like hydrology (Schultz, 1987 (cited by Haarhoff & van Zyl, 2002); Thomson *et al.*, 1997; Cui & Kuczera, 2003). Recently, residential water demand has been simulated with a stochastic end-use model by Blokker *et al.* (2010). Stochastic modelling is often used to analyse complex systems in which risk and uncertainty are significant, and where components are subject to stochastic failures (Yang *et al.*, 1996). Work on the application of stochastic analysis to water distribution systems has been done by Wagner *et al.* (1988b), Yang *et al.* (1996), and Ostfeld *et al.* (2001). In addition, Nel (2009) states that stochastic simulation is highly suitable for reliability assessment of bulk water distribution, because many of its variables have probabilistic characteristics.

2.4.1 Monte Carlo simulation for reliability analysis

As previously discussed, there are many ways of obtaining estimates of the reliability of water supply networks. Techniques used to estimate the reliability of water supply and distribution systems can be categorised into analytical methods, heuristic methods and Monte Carlo simulation (Xu & Powell, 1991). However, the use of simulation over an extended period of time or the use of stochastic techniques, such as Monte Carlo simulation, has been shown to be the most realistic of these measures of reliability.

Stochastic simulation, like the Monte Carlo technique, is useful because of the advantages that it provides in reliability analysis. The advantages are that the simulation results can be used to calculate the desired reliability factor; and the randomness of factors such as demand and failures can be easily included (Yang *et al.*, 1996). In fact, Zio (2009) suggests that Monte Carlo simulation ‘may be the only method that can yield solutions to complex multi-dimensional stochastic modelling problems such as those typically involved in reliability and availability analysis.’ On the other hand, a major disadvantage of stochastic simulation is that it is computationally expensive and lacks computing power; therefore its practical applications are limited. Nevertheless, Monte Carlo simulation is gaining feasibility as a technique for reliability analysis, as technological advances are made with the availability of faster computers. Its usefulness in reliability assessment has recently been demonstrated by van Zyl *et al.* (2008).

Van Zyl *et al.* (2008) developed a stochastic model for the reliability analysis of storage reservoirs. The initial work on the stochastic model was done by Nel (1993), summarised and published later by Nel & Haarhoff (1996). Haarhoff & van Zyl (2002) describe how this probabilistic modelling was applied to the sizing of bulk water supply systems. The stochastic model described in the study by van Zyl *et al.* (2008) consists of three unit models for consumer demand, fire demand and pipe failure. The stochastic model determines the failure characteristics of the reservoir. The failure characteristics are used to assess the reliability, which can then be used as a sizing criterion for the design of the storage reservoir. This model was then used to explore the effect of varying the user demand parameters on storage reservoirs (van Zyl *et al.*, 2010). Vlok (2010) applied this stochastic model to a bulk water supply system to investigate the financial implications associated with the use of risk-based analysis techniques. Since the stochastic model uses Monte Carlo simulation which is computationally intensive, it is worth investigating ways to increase its computational efficiency.

2.4.2 Compression heuristic

One way of speeding up the stochastic analysis of storage reservoirs is by using a compression heuristic proposed by Chang & van Zyl (2010). In the stochastic model (van Zyl *et al.*, 2008), the entire duration of the study period is simulated in full using the Monte Carlo method. This can be repetitive and time-consuming. In contrast, the compression heuristic

fully simulates the critical periods, when fire or pipe failure events occur, while using a heuristic meta-model to simulate the intermediate periods. The meta-model simulates the demand-only behaviour and determines the reservoir failure characteristics (Chang & van Zyl, 2010).

It was found in the study (Chang & van Zyl, 2010) that the compression heuristic achieved significant improvements in computation speed. It produced results to an accuracy of 95% with simulation times that were approximately 3 to 8.8 times faster than the full Monte Carlo method. The number of demand-only failures and the number of events that occurred affected the efficiency of the method. Best results were obtained for systems that are highly reliable with few occurrences of events, i.e. few demand-only failures, and low fire and pipe failure rates. Thus, the compression heuristic can bring about significant savings in the computational time required for the reliability analysis of storage tanks.

Similar work was done on a simple heuristic called replicate compression to improve Monte Carlo efficiency by Cui & Kuczera (2009). The replicate compression was applied to the optimisation of urban water supply headworks. It also takes advantage of the concept of a critical period and restricts simulation to these periods to reduce the computational effort.

2.4.3 Sampling techniques

The basic Monte Carlo sampling technique is computationally intensive. It involves the repeated calculation of the performance of the system, each time with a different combination of input parameters, which are randomly selected from probability distribution functions. The problem with this is that it requires long runs to produce precise estimates of reliability (Haarhoff & van Zyl, 2002). There are, however, sampling techniques that are potentially faster and more efficient than Monte Carlo sampling, i.e. they require fewer iterations and shorter simulation runtimes. Sampling techniques that are suggested are Latin Hypercube sampling, descriptive sampling and importance sampling.

Latin Hypercube Sampling

Latin Hypercube sampling is a stratified sampling technique that ensures that all portions of a sampling distribution are sampled (McKay *et al.*, 1979). Latin Hypercube sampling has been used in several applications, such as in the calculation of the probability of dam failure and flood risks (Thompson *et al.*, 1997), and in risk calculations for culverts (Lian & Yen, 2003).

The key to Latin Hypercube sampling lies in the stratification of input probability distributions. The range of the input variable is divided into N strata of equal marginal probability $1/N$. A sample is randomly drawn from each interval. Since the sampling is forced to represent values from each interval, it is forced to recreate the input probability distribution. A random value is then drawn from within the selected interval.

Latin Hypercube sampling has a number of advantages over the Monte Carlo method. The main difference between Latin Hypercube sampling and Monte Carlo sampling is in the number of iterations required until the sampled values approximate the input distributions. Several researchers in the literature agree that Latin Hypercube sampling has improved computational efficiency. It has the flexibility of Monte Carlo sampling with less computational effort (Manache & Melching, 2004). It reduces the number of computer runs to achieve the same precision as the conventional Monte Carlo method, i.e. it provides a more accurate estimate of the mean value of the function (Giunta *et al.*, 2003) This is because Latin Hypercube sampling forces sampling to select values over the whole range of the model parameter, thus reducing the total number of samples required to preserve the probability distribution (Khan *et al.*, 2008). Another advantage of Latin Hypercube sampling is that it ensures that the entire range is sampled (Lian & Yen, 2003).

A challenge for Latin Hypercube sampling is that generated values for each variable are not independently distributed, as they are considered to be drawn from different subranges. As a result the traditional methods for estimating sampling error are not applicable (Thompson *et al.*, 2003). Thompson *et al.* (2003) indicates that the primary difficulty with Latin Hypercube sampling is that most of the sampled values do not reflect large events, which are often the most critical events for simulation. However, it was found that if the number of strata N is large enough then the extreme events will be sampled (Thompson *et al.*, 2003). Therefore, extreme events can be accurately represented in the simulation outputs. Another drawback is that there is more than one possible arrangement of bins and samples that meet the criteria for Latin Hypercube sampling. If these samples are poorly arranged and are nearly co-linear, then the system will be ill-conditioned (Giunta *et al.*, 2003).

Descriptive Sampling

Saliby (1997) proposed the use of descriptive sampling as an improvement over Latin Hypercube sampling and an alternative to Monte Carlo sampling. Descriptive sampling is based on deterministic selection of input sample values and random permutations (Saliby, 1990). Saliby (1990) argues that descriptive sampling achieves the closest fit with the represented distribution and it yields, in general, much more precise estimators. However, the use of descriptive sampling is controversial as it proposes the abandonment of the concept that a random selection of sample values is necessary to describe random behaviour.

Two kinds of variation exist in a randomly generated sample in a Monte Carlo or simple random sampling approach, namely set and sequence variability. Set variability consists of the deviations between actual sample parameters and corresponding assumed parameters for the input design variables, whereas sequence variability is the randomness of the sample values. Saliby (1997) argues that sequence variability is inevitable, while set variability is unnecessary. Descriptive sampling tries to remove or reduce set variability.

The two sampling methods can be represented symbolically (Saliby 1997):

& Zhang, 2003). The outputs from simulations are then weighted so as to preserve the mean of the quantity being estimated. Importance sampling will not be effective unless the appropriate importance density function is known.

2.5 Bulk water supply system and storage reservoir design

Bulk supply pipelines and bulk storage reservoirs are some of the elements of a water distribution and storage system. The bulk water supply system must be designed so that it can provide adequate supply to the distribution network, and it is usually designed according to deterministic design guidelines. South African design guidelines suggest that the reservoir should have a storage capacity of 48 hours of annual average daily demand and that the supply mains should have a capacity capable of providing an inflow rate to the reservoir greater than or equal to 1.5 times the annual average daily demand for the area (CSIR, 2000).

There is limited work on the design of bulk water supply systems in the literature. Haarhoff & van Zyl (2002) describe work done on an alternative approach to guidelines-based design of bulk water supply systems, based on stochastic modelling, which was first proposed by Nel & Haarhoff (1996). This stochastic analysis method was further developed by van Zyl *et al.* (2008) and applied to a simple system consisting of a source feeding a reservoir via a single pipeline, and a separate pipe connecting the users to the reservoir.

Bulk supply system design includes the sizing of storage reservoirs. The purpose of a municipal storage reservoir is to balance supply and demand of water distribution systems, and provide additional storage for fire fighting and emergency purposes. Like pipe networks, there are various methods that can be used to design storage reservoirs in a water distribution system. A number of variables can be considered when designing reservoirs. These include the type of reservoir (ground or elevated), the location, the bottom elevation, the minimum and maximum operating levels, and the shape and size of the reservoir. Design methods for storage reservoirs can be broadly categorised into guidelines-based designs and optimisation models.

Table 2.1 Reservoir design criteria for low-density residential areas (van Zyl *et al.*, 2008)

Country	Storage requirements		
	Balancing (<i>B</i>)	Fire (<i>F</i>) (m ³)	Emergency (<i>E</i>) (days)
South Africa	Included in <i>E</i>	108	2 average
France	30% of peak day	120	2 average
United States	35% of peak day	680	1 peak

Note: Requirements represent van Zyl *et al.*'s (2008) interpretation of the guidelines for a typical low-density residential area. Where more than one option is given, the most stringent was usually selected. Where no fire duration was specified a value of 2 h was used.

3 Method Overview

In this study, a method is proposed to obtain a set of solutions for the design of a bulk water supply system. These Pareto-optimal solutions will assist the designer in selecting the configuration and diameter of the feeder pipe system, and the size of the bulk storage reservoir. These design variables can be selected according to a desired reliability, while concurrently bearing in mind the cost of the system.

One of the main objectives of any engineering design is ensuring that the system is as cost effective as possible. Water distribution system design, and likewise the design of bulk water supply systems, is the same in that the system is often optimised so that the cost can be minimised. To be able to achieve this objective, a cost model is required so that cost analysis can be performed on the bulk supply system. Since there is no point in reinventing the wheel, existing cost models (Swamee & Sharma, 2008; Vlok, 2010) were used in developing this method. The details for these cost models are presented later in Section 6.4.

Although for many years water distribution system design optimisation was regarded and modelled as a single-objective least-cost problem, engineers and decision makers have come to realise that it is in fact better represented as a multi-objective problem (Walski, 2001; Farmani *et al.*, 2005). Additionally, the importance of including objectives such as water quality, resilience, risk and reliability is recognised. Following this trend, the reliability of the bulk supply system was included as the second objective in this multi-objective optimisation problem. In this project, a stochastic analysis model (van Zyl *et al.*, 2008) was the chosen method for analysing the reliability of the system, as stochastic techniques such as Monte Carlo simulation have been declared to be one of the most realistic measures of reliability (Xu & Powell, 1996). Furthermore, the randomness of factors such as demand and failures can be included without difficulty (Yang *et al.*, 1996).

Even though stochastic techniques have the advantages of being realistic and have the ability to encompass the randomness and uncertainty of factors, they have a major drawback of being computationally expensive, given that they are time-consuming and memory-intensive. As a result, Chang & van Zyl (2010) proposed the use of a compression heuristic to reduce the computational effort of the stochastic model by van Zyl *et al.* (2008). The compression heuristic was shown to be an improvement of the original stochastic model in terms of computational time. However, it was felt to be deficient as each simulation was terminated according to fixed conditions based on several trials of the method and the judgment of the authors. Thus, one of the major objectives in this study was to further develop and refine the compression heuristic, so that the stochastic simulations terminate according to conditions which ensure that the produced results are reliable and accurate.

As a secondary objective, further reduction in the computational time of the stochastic model is desired. Alternative sampling techniques, such as Latin Hypercube sampling and importance sampling have been shown to be an improvement over Monte Carlo simple

random sampling. Section 3.2 provides a discussion of a brief investigation of Latin Hypercube sampling in the stochastic model. From the investigation, it was found that Latin Hypercube sampling did not produce the desired results, and thus further work in this area was disregarded in this study.

The cost model and stochastic model were both included in the optimisation procedure to analyse the cost and reliability for each bulk supply system. The optimisation technique selected for the proposed method is the well-known multi-objective genetic algorithm NSGA-II (Deb *et al.*, 2002). Although there are other multi-objective evolutionary algorithms, Deb *et al.* (2002) have shown that NSGA-II is able to perform better than other existing algorithms such as the Pareto-archived evolution strategy (PAES) and strength-Pareto evolutionary algorithm (SPEA). However, in a recent study by Raad (2010), the AMALGAM hyperheuristic was developed and shown to outperform NSGA-II for some water distribution system benchmarks. Yet for other benchmarks, the AMALGAM hyperheuristic was inferior to NSGA-II. Owing to its popularity in literature and easy implementation, NSGA-II was used as a tool in this method to produce the solutions for the optimal design of bulk water supply systems, based on cost and reliability analyses. A flowchart of the proposed method is shown in Figure 3.1

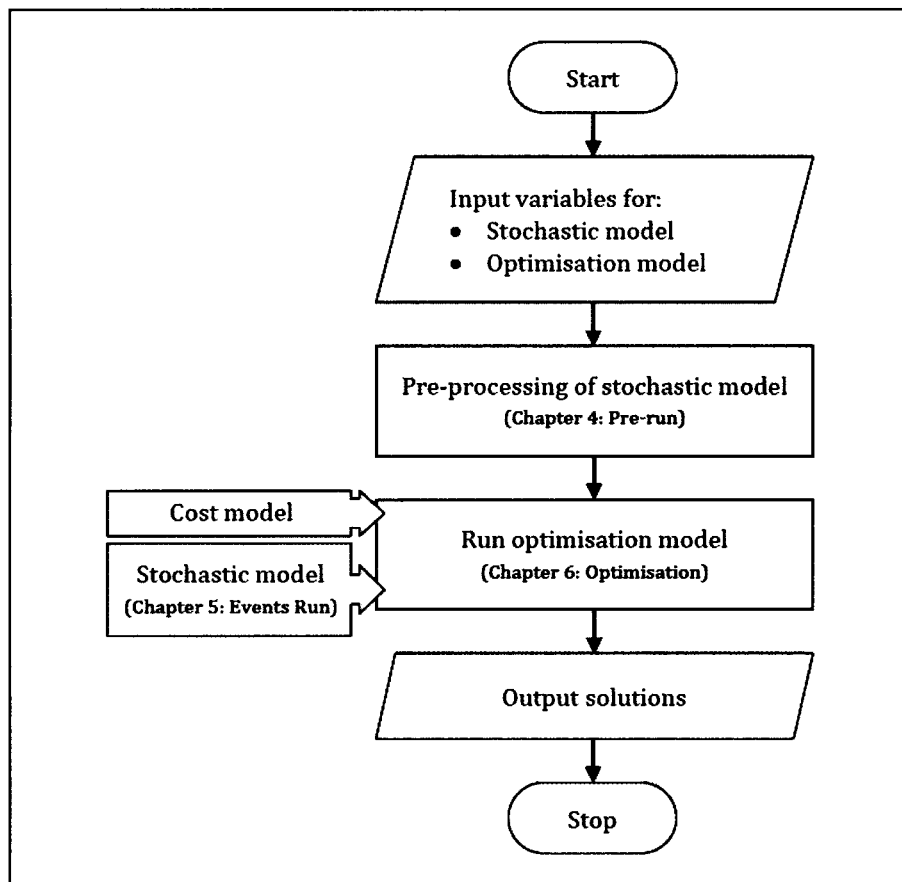


Figure 3.1 Flowchart of proposed method

3.1 The compression heuristic

This section describes the purpose of the compression heuristic, as well as the motivation behind it. The need for further development and refinement of the compression heuristic is also discussed.

3.1.1 Purpose and motivation for developing the compression heuristic

The full stochastic analysis method using Monte Carlo sampling simulates each individual time step (one hour) for the entire run duration. In each time step, the hourly consumer demand is calculated. In addition to the hourly demand, the fire demand and supply pipe inflow rate are determined in the same time step. If no fire event is happening, the fire demand is zero; and if no pipe failure has occurred, then the maximum inflow into the storage reservoir is supplied. It is evident that calculating the consumer demand, fire demand and inflow rate for each individual time step is computationally expensive and therefore highly time consuming. In order to reduce the computational effort required for the full stochastic analysis, a compression heuristic method was proposed by Chang & van Zyl (2010).

From previous work (van Zyl *et al.*, 2008; Chang & van Zyl, 2010), it was found that reservoir failures were generally more likely to coincide with the occurrence of fire and/or pipe failure events. Time steps where the reservoir is only affected by consumer demand have fewer failures, in comparison to those where there is an added fire demand or reduced supply inflow. Therefore, the time steps which are affected by fire or pipe failure events are the *critical periods*. It is important that these critical periods are simulated fully using the full Monte Carlo simulation so as to ensure that the effects of the events on the reservoir are accurately recorded. The *events run* (Chapter 5) is the main run of the compression heuristic responsible for simulating these critical periods.

Between the critical periods, the storage reservoir often experiences long periods of demands-only behaviour. In these intermediate periods, the reservoir repeatedly fills up completely and there are few failures. However, the consumer demand must still be considered, especially when high consumer demand coincides with fire or pipe failure events, which will increase the likelihood of a reservoir failure occurring. Thus, a heuristic meta-model referred to as the *pre-run* (Chapter 4), is used to simulate the system behaviour under consumer demands only.

3.1.2 Further development of the compression heuristic

In the original work on the compression heuristic, the pre-run simulation was run for a static length of time, which was chosen to be 1000 years (1.75 million hourly time steps). This was an arbitrary simulation period as it was selected based on obtaining 2000 reservoir failures for a bulk system with peak demand of 80 l/s, a supply mains capacity of 1.2 times the peak demand (i.e. supply ratio of 1.2), a reservoir size of 12 h storage capacity, a fire rate of 6

3.2 Investigation of Latin Hypercube sampling

In the literature, Latin Hypercube sampling (LHS) has been found to reduce the number of computer runs to achieve the same precision as conventional Monte Carlo sampling, with the added benefit of ensuring that the entire range is sampled. In this section, Latin Hypercube sampling is investigated as an alternative to Monte Carlo sampling.

The two main objectives are (i) to investigate how accurate Latin Hypercube sampling is in comparison to Monte Carlo sampling; and (ii) to investigate how much Latin Hypercube sampling is able to reduce the computational effort required for the simulation. As part of the investigation, it is necessary to formulate a method to implement LHS into the simulation. Thus, further objectives include (iii) determining how many strata, N , of equal probability $1/N$ would be required for the study; and (iv) determining stopping rules or convergence criteria for the simulation.

3.2.1 Determining the number of strata required

As part of the investigation, the full Monte Carlo simulation was run ten times using different seeds for the random number generator. These ten simulations were run so that the average failure frequency for the stochastic analysis could be found, and the 5% upper and lower boundaries could be established. The number of strata that were investigated were $N = 1000$, $N = 10,000$, $N = 100,000$ and $N = 1,000,000$. The LHS was run for these four strata sizes, with the random seed equal to zero for all. For each strata size, the LHS was run until 5000 failures were generated. The failure frequencies were then plotted against the number of failures on a graph (Figure 3.2) which also indicates the 5% upper and lower boundaries for the full MC method.

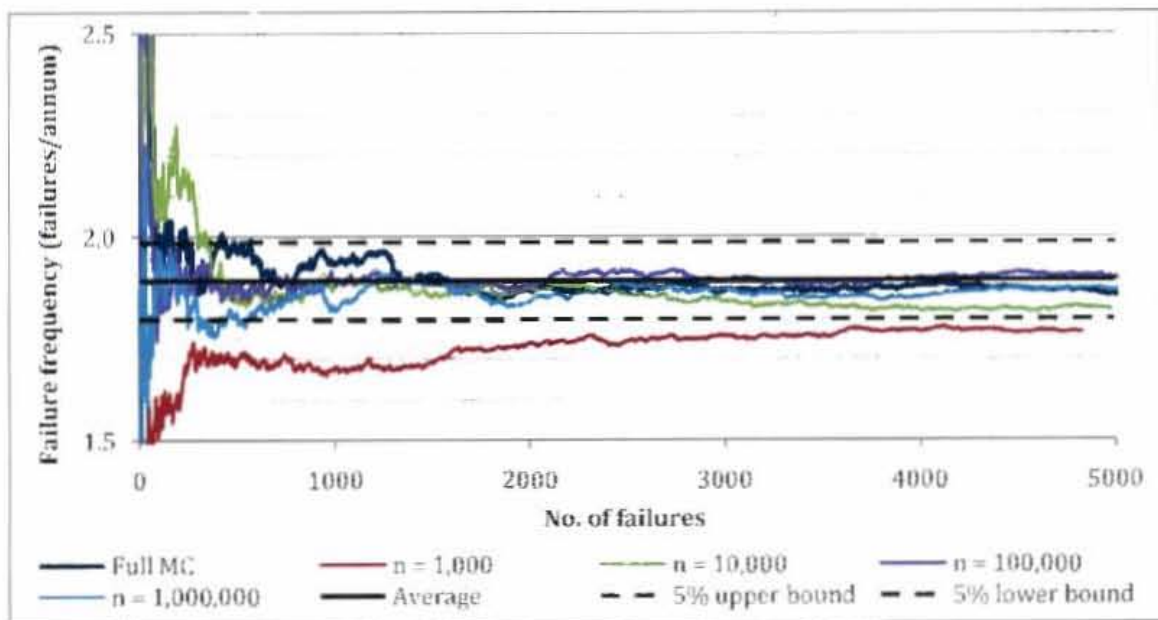


Figure 3.2 Comparison of failure frequency for Latin Hypercube sampling and Monte Carlo sampling

By plotting this graph, it can be determined for which strata sizes, the LHS is able to provide estimates of the failure frequency for which are reasonably close to that produced by Monte Carlo sampling. From the graph, it can be seen that with $N = 1000$, the results are completely out of bounds. This indicates that a strata size of 1000 intervals is too small to produce accurate results, and underestimates the failure frequency of the storage reservoir. $N = 10,000$ falls within the boundary lines; however it appears to tend towards the lower 5% boundary. In contrast, $N = 100,000$ and $N = 1,000,000$ appear to be fairly stable, and provide failure frequency results which are close to the full MC average value.

3.2.2 Convergence criterion for the simulation

The convergence criterion used for the stochastic simulation using either Latin Hypercube or Monte Carlo sampling is based on the confidence interval method, which is fully described in Section 4.5. In summary, the simulation failure results converge once

$$\frac{1.96S}{\sqrt{n}}/\bar{y} \leq 5\% \quad (3.1)$$

where \bar{y} is the mean time-to-failure, S is the time-to-failure standard deviation and n is the number of failures.

Using this convergence criterion ensures that the simulation stops when there is a 95% certainty that the mean time-to-failure is within 5% of its true value.

3.2.3 Comparison of Latin Hypercube sampling to Monte Carlo sampling

Once Latin Hypercube sampling was incorporated into the stochastic analysis method, the next step was to determine whether it would be feasible to use it as a replacement for Monte Carlo sampling. This was investigated by comparing the accuracy of the stochastic results for both simulations. It was also necessary to compare the simulation times required for both sampling techniques as it would be pointless to use Latin Hypercube sampling if it failed to reduce the computational effort of the Monte sampling for this application.

Table 3.1 shows the results for Latin Hypercube sampling with varying number of strata sizes, and they are compared to the results for the Monte Carlo sampling. For each sampling technique, the results for three different random seeds were obtained. The fire results, pipe failure results, and overall total reservoir failure results for the simulations using Latin Hypercube sampling are very close to those obtained for the simulations using Monte Carlo sampling. This indicates that Latin Hypercube sampling can be used instead of Monte Carlo sampling in the stochastic analysis.

Table 3.1 Comparison of results for Latin Hypercube and full Monte Carlo sampling

Sampling		Fires			Pipe failures			Total reservoir failures		
Method	Seed	per annum	Ave dur (h)	Sd dur (h)	per annum	Ave dur (h)	Sd dur (h)	per annum	Ave dur (h)	Sd dur (h)
Full MC	0	6.0206	0.8429	0.6110	1.9312	5.0311	2.5714	1.8641	3.0741	2.3012
	1	5.9889	0.837	0.6277	1.998	5.0087	2.5975	1.9028	3.1285	2.4345
	2	6.0514	0.8431	0.6168	1.9362	5.1151	2.5731	1.8463	3.1711	2.4067
LHS, N = 10 ⁴	0	5.9155	0.8475	0.6312	1.942	4.9653	2.554	1.8838	3.1617	2.4063
	1	5.9772	0.8379	0.6037	2.0389	4.9071	2.4006	1.902	3.107	2.3416
	2	5.9912	0.8291	0.6165	1.9043	5.0499	2.5175	1.854	3.1352	2.3371
LHS, N = 10 ⁵	0	6.0309	0.8466	0.6267	2.0059	4.9792	2.5494	1.9151	3.1557	2.4504
	1	5.9763	0.8391	0.6056	1.9599	4.8966	2.3933	1.8948	3.1239	2.332
	2	5.9786	0.8289	0.6182	2.0037	5.0598	2.5212	1.8903	3.1966	2.4521
LHS, N = 10 ⁶	0	6.0282	0.8467	0.629	1.9662	4.9783	2.561	1.8598	3.1696	2.4107
	1	6.0691	0.8386	0.6043	1.9452	4.9031	2.395	1.9392	3.1098	2.2992
	2	5.9348	0.8291	0.6183	2.0144	5.0643	2.5233	1.9297	3.1657	2.4195

Besides comparing the reproducibility of the results, the simulation time required for Latin Hypercube sampling was compared to that for Monte Carlo sampling. Two times were recorded: the number of time steps (hours) simulated by the program, and the actual real time taken to complete the simulation. In Table 3.2, it is clear that Latin Hypercube sampling generally simulates fewer hourly time steps than the Monte Carlo sampling for the three random seeds tested. However, Monte Carlo sampling is actually faster than Latin Hypercube sampling in real time, as the Monte Carlo simulation takes approximately 450 seconds while the Latin Hypercube simulation takes approximately 500 seconds. A possible reason for the longer real simulation time of Latin Hypercube sampling is that it requires extra computational effort to generate each Latin Hypercube sample. Another reason is that Latin Hypercube sampling is generally used in problems where multi-dimensional Latin Hypercube samples are generated, whereas only one-dimensional Latin Hypercube samples are needed in this problem. Consequently, the benefit that can be gained from using Latin Hypercube sampling is not maximised.

Table 3.2 Comparison of simulation times for Latin Hypercube sampling and full Monte Carlo sampling

Sampling		Time	
Method	Seed	Simulated (h)	Actual (s)
Full MC	0	10,832,293	432.2
	1	12,057,164	480.5
	2	11,401,293	449.6
LHS, N = 10 ⁴	0	10,081,698	482.2
	1	10,169,535	490.2
	2	10,451,444	500.8
LHS, N = 10 ⁵	0	10,703,686	512.8
	1	9,967,410	478.4
	2	10,195,052	490.3
LHS, N = 10 ⁶	0	10,541,294	511.9
	1	10,132,583	492.6
	2	10,341,213	500.7

From this investigation, it can be concluded that Latin Hypercube sampling does not provide a suitable alternative to Monte Carlo sampling for the stochastic analysis in this study. Even though it is able to produce similar results to Monte Carlo sampling, Latin Hypercube sampling does not appear to produce the required increase in computational speed. As a result, further work on Latin Hypercube sampling for this application of stochastic simulation was abandoned.

4 Pre-Run

4.1 Introduction

In this chapter, a detailed description of the methodology used in the demands-only pre-run is provided. The chapter also focuses on any problems that were encountered and solved while developing the method.

The demands-only pre-run serves two purposes. First of all, it determines the reservoir failure results due to consumer demands. The pre-run function computes the demands failure frequency for the storage reservoir, while the events run calculates the events failure frequency. In the events run, the demands failure frequency is extracted from the pre-run simulation and combined with the events failure frequency to give the total failure frequency for the bulk supply system.

In addition to computing the demand failure results, the pre-run function is responsible for generating a lookup table of the reservoir level distribution at the start of each week (Sunday 04h00) in the simulation. The events run will use this lookup table to set the initial reservoir level at the start of each fire or pipe failure event.

4.1.1 Description of the pre-run algorithm

A flowchart of the algorithm for the demands pre-run is provided in Figure 4.1. At the start of the pre-run, the reservoir data structure is initialised by taking a range of supply ratios and reservoir capacities as input variables.

Following the initialisation, for each supply ratio, the pre-run loops through each of the reservoir capacities, and performs the stochastic analysis using Monte Carlo sampling. In the stochastic analysis model, the system parameters for the supply inflow and the reservoir volume are initialised, and the time of day is set to zero. The consumer demand is calculated and the reservoir level is updated for each hourly time step. If a failure occurs, then it is recorded, and the demand failure results are updated. The storage level in the reservoir is also recorded every Sunday 04h00. When the stochastic analysis is completed for the current set of supply ratio and reservoir capacity, the lookup table for the reservoir level distribution curve is generated.

Once the demand failure results for all the supply ratios and reservoir capacities have been obtained, an exponential curve is fitted to the demands failure results. The coefficients for the exponential equations are saved for use in the events run.

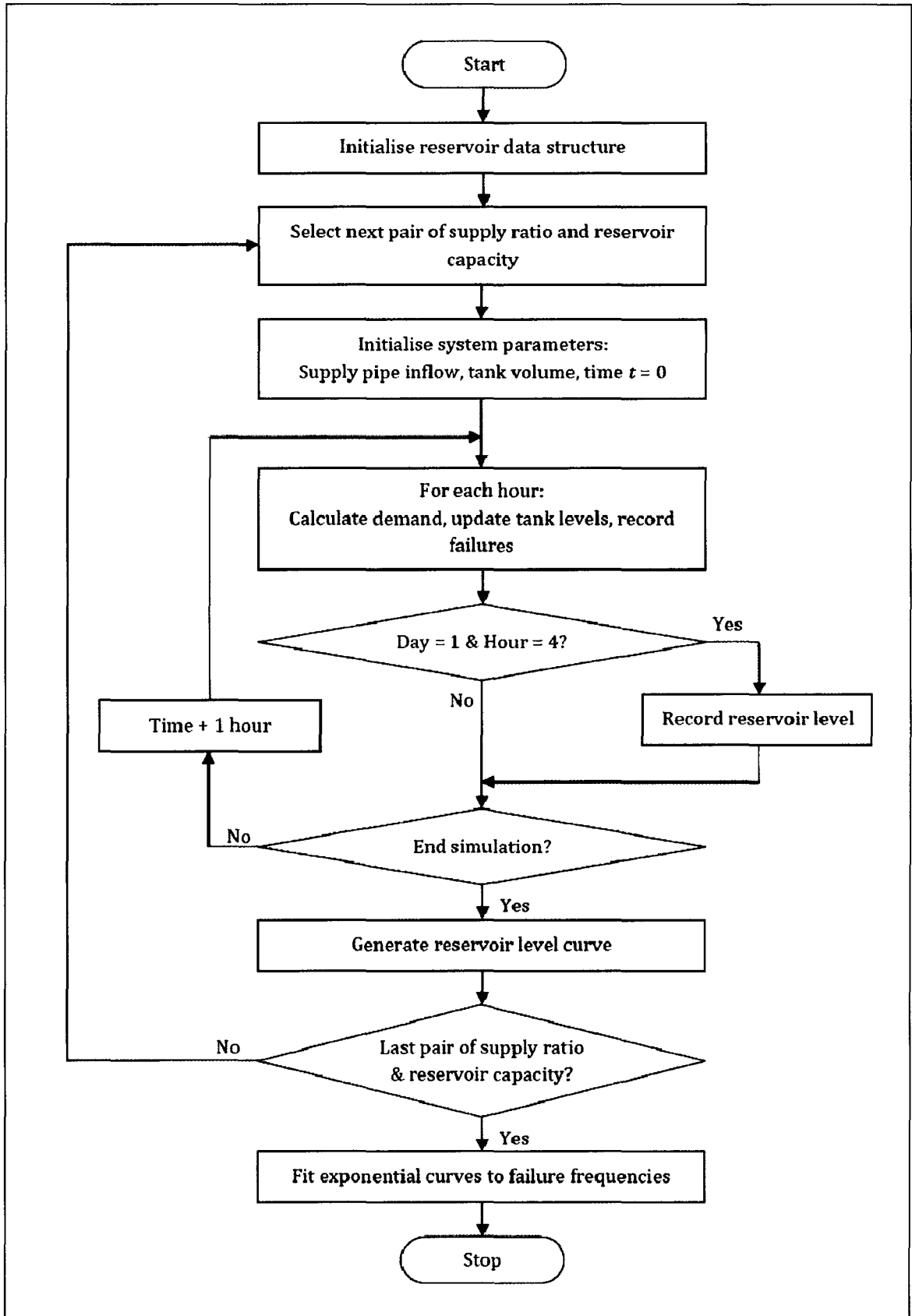


Figure 4.1 Flowchart for the demands-only pre-run

4.2 Representing the system variables

The purpose of the pre-run is to obtain the failure results and obtain curves of initial storage reservoir levels for a series of supply ratios and reservoir sizes. In this section, the required input variables for the pre-run function are discussed. Following this, a description of the output variables will be provided with an explanation of the data structures used.

4.2.1 Input variables

The pre-run takes in the stochastic demand parameters as input variables. These input variables are the seasonal peak demand (D_{peak}), the daily and hourly autocorrelation coefficients (ϕ_D and ϕ_H), and the daily and hourly standard deviations (σ_D and σ_H). Besides the stochastic parameters, the range of supply ratios and reservoir capacities to be analysed are also included as input variables.

Determining the range of supply ratios

Each bulk supply system will have its own pipe configuration and diameter. In this design model, seven different configuration options are available, ranging from a single pipe, to two and three pipes in parallel, to two and three pipes in parallel with a varying number of interconnection pipes. The pipe diameter options depend on the number of commercially manufactured pipe sizes that are available. Each pair of pipe configuration and pipe diameter options will have an associated total supply inflow into the storage reservoir, and thus an associated supply ratio.

Prior to the commencement of the pre-run, the designer is requested to input a total system head H , total pipe length for the system to be analysed L , and Hazen-Williams coefficient C_H . An illustration example of designer input values into the model is shown in Table 4.1.

Table 4.1 Example of designer input values

User Input:	
Seasonal peak demand, D_{peak} (L/s)	: 80
Total system head, H (m)	: 30
Pipe length, L (m)	: 1000
Hazen-Williams coefficient, C_H	: 120

For the range of commercially available pipe diameters D , the pipe flow Q through a single pipe can be calculated, using the Hazen-Williams headloss formula:

$$Q = 0.0885\pi C_H D^{2.63} \left(\frac{H}{L}\right)^{0.54} \quad (4.1)$$

The pipe flow Q calculated is for a single pipe; therefore, for two pipes in parallel it is multiplied by two, and for three pipes in parallel it is multiplied by three. This is applicable for pipes in parallel both with and without interconnections.

The corresponding supply ratios for each pipe diameter and the different number of pipes in parallel can be calculated by dividing each inflow rate by the average demand. The range of supply ratios obtained for each combination of configuration and diameter can be highly varied. At this stage, it is necessary to eliminate the pipe diameters which will provide infeasible supply ratios. There are two extremes of supply ratios which are infeasible to use in the stochastic analysis. The first extreme is a very small supply ratio, typically less than one. If the supply into the reservoir is less than the demand out of the system, then the supply inflow will never be able to fill the reservoir and it will often be empty. The other extreme is when the supply ratio is very large, thereby ensuring that the supply capacity is greater than or equal to the maximum system demand. In this case, no balancing storage is necessary.

Deterministic design guidelines often use a fixed supply ratio in the sizing of the feeder pipe. As noted by Nel & Haarhoff (1996), feeder capacity is generally taken as 1.5 times the annual average daily demand (AADD) in South Africa, equivalent to one times the seasonal peak demand. In the sensitivity analysis of a recent stochastic study, van Zyl *et al.* (2010) suggested the use of a typical supply ratio of 1.5 under seasonal peak conditions, with a lower bound value of 1.1 and an upper bound of 2. The lower and upper bound supply ratios were used as a guide to determine which pipe diameters would be feasible and retained for selection in the optimization run. This resulted in the use of the following two conditions where a diameter was eliminated if: (1) the supply ratio for a single pipeline was greater than the upper bound; or (2) the supply ratio is less than one for three parallel pipes. Condition (2) ensures that a small pipe size is included even if its supply ratio for a single pipeline is less than one, because of the possibility of a configuration with three parallel pipes being selected in the optimization. The pipe diameters that are not eliminated in the illustration example (with input values shown in Table 4.1) are highlighted in Table 4.2. For the illustration example, the range of supply ratios arranged in ascending order would be $SR = \{0.51, 0.91, 1.02, 1.52, 1.64, 1.83, 2.74, 3.28, 4.92\}$.

Table 4.2 Supply inflow rate and supply ratios calculated for the illustration example, indicating which diameters are eliminated

Program Calculates:							
D (mm)	Q ₁ (L/s)	Q ₂ (L/s)	Q ₃ (L/s)	SR ₁	SR ₂	SR ₃	Eliminate?
90	8.94	17.88	26.82	0.11	0.22	0.34	Y
110	15.15	30.3	45.45	0.19	0.38	0.57	Y
160	40.6	81.2	121.8	0.51	1.02	1.52	N
200	73.02	146.04	219.06	0.91	1.83	2.74	N
250	131.31	262.62	393.93	1.64	3.28	4.92	N
300	212.11	424.22	636.33	2.65	5.30	7.95	Y
350	318.16	636.32	954.48	3.98	7.95	11.93	Y
400	452.03	904.06	1356.09	5.65	11.30	16.95	Y

Choosing the range of reservoir storage capacities

Unlike the range for the supply ratios, the range for the reservoir sizes can be selected by the designer. The designer is requested to input a minimum and maximum reservoir capacity in hours of storage, and a step size for the intermediate reservoir capacities. If the designer inputs the minimum as 3 hours of storage and the maximum as 12 hours of storage, with a step size of 3 hours, then the range of storage capacities is: $RC = \{3, 6, 9, 12\}$ h.

Pre-run input grid

After they are selected, the range of supply ratios and reservoir capacities are entered as input variables into the demands pre-run function. The range of supply ratios and reservoir capacities provide the structure for the reservoir data output.

4.2.2 Output variables

For the range of supply ratios and reservoir capacities, the pre-run function produces matrices for the demand failure results and the reservoir level values. A typical grid is shown in Figure 4.2.

		RC			
		3 h	6 h	9 h	12 h
SR	0.51	<ul style="list-style-type: none"> • Number of failures • Simulated hours • Failure frequency • Convergence ratio • Full reservoir fraction 			
	0.91				
	1.02				
	1.52				
	1.64				
	1.83				
	2.74				
	3.28				
4.92					

Figure 4.2 Typical grid of *SR-RC* matrix for failure results

In each cell of the grid, there is a demand failure result corresponding to each row (supply ratio) and each column (reservoir capacity). The demand failure results that are recorded in this two-dimensional matrix structure are the number of demand failures that occur, the number of simulated hours, the failure frequency and the convergence ratio. How each of these failure results is obtained is described later in this chapter. The full reservoir fraction results also follow this structure.

Two three-dimensional matrices are created for the reservoir level results (Figure 4.3), one for the *top of bin* (i.e. the set of upper limits for each reservoir level category) for each pair of supply ratio and reservoir capacity, and one for the *relative cumulative frequency* (or probability of nonexceedance) of the reservoir levels. The top of bin and relative cumulative

frequency variables consist of a vector with ten elements. Therefore for a range of nine supply ratios and a range of four reservoir capacities, the three-dimensional matrix will contain 9 x 4 x 10 elements.

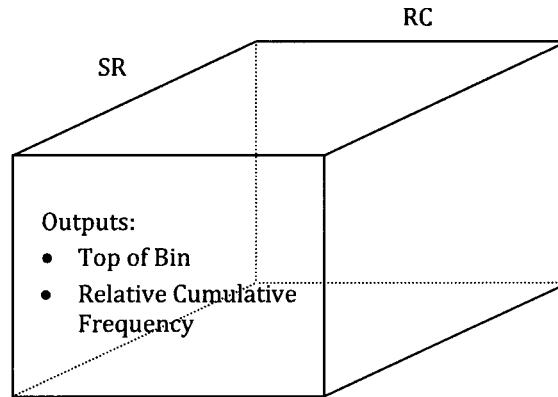


Figure 4.3 Matrix for reservoir level results

The output results obtained from the pre-run function will be entered as input variables into the main events function, keeping the same data structure. This data structure allows a simple and uniform format for easy referencing in the events function.

4.3 Recording demand failures

Reliability in this study is expressed in terms of the failure frequency of the storage reservoir, where high reliability corresponds to low failure frequency. The demands failure frequency from the pre-run and the events failure frequency from the events run are combined to give the total failure frequency for the compression heuristic. This section describes how demand failures are recorded and how the demands failure frequency is computed.

In the pre-run function, only demand failures are recorded and analysed. No fire or pipe failure events are simulated, therefore there are no failures caused by events in the pre-run. Only the consumer demand model is used for the pre-run simulation. In the consumer demand model, the daily demand is calculated at the start of each day. In each time step (hour), the hourly demand is then determined and used to update the volume of the reservoir. If the volume is less than zero, i.e. the reservoir is empty, then it has failed. When a failure occurs, the following quantities are recorded: the number of failures, the *time-to-failure* (TTF), the *mean time-to-failure* (MTTF) and the *time-to-failure standard deviation* (STTF). These quantities are used to determine when the failure results have converged and the pre-run is terminated. The failure ends once there is a net inflow into the reservoir.

The MTTF represents the average of the TTF between each system failure (i.e. years between the beginning of one failure and beginning of the next). The reciprocal of the MTTF is thus taken to convert the MTTF to the annual average number of failures or failure frequency, represented by failures per annum. This is expressed in the following equation:

$$\text{failure frequency} = \frac{1}{MTTF} \quad (4.2)$$

It is necessary to note here that while the failure frequency could be directly obtained from the simulation, the mean time-to-failure is used instead. The reason for using mean time-to-failure is discussed in Section 5.8.

4.3.1 Exclusion of failure duration as a sizing parameter

It is worthwhile to note that in previous studies on storage reservoir sizing using stochastic analysis (van Zyl *et al.* 2008; Chang & van Zyl, 2010), both the failure frequency (average annual number of failures) and the average failure duration were computed and recorded. In this study, however, it was felt that failure duration should be excluded as a parameter.

Van Zyl *et al.* (2008) referred to work done on formulating reliability requirements for water supply systems by Kwietniewski & Roman (1997). Kwietniewski & Roman (1997) identified three reliability criteria based on data analysis of consumers' feelings on threats to sanitary health, comfort and good living conditions. These reliability criteria were failure frequency, mean repair time, and proportion of time for which water is available. They surveyed 100 families to find that 74% of consumers considered a 24 h interruption to supply once a year 'tolerable'; while only 55% of consumers found two 12 h failures tolerable. This suggests that from the consumer's point of view, the number of failures is more significant than the failure duration.

In their own work, van Zyl *et al.* (2008) provide further motivation for why it is adequate to use failure frequency as the only reservoir sizing parameter. Their results indicated that the average annual number of failures is very sensitive to the reservoir capacity. In contrast, they found that the average failure duration does not vary much with storage capacity, and is not a suitable parameter for determining the required reservoir capacity. Following this reasoning, failure duration was therefore not included in the reliability analysis. By excluding the failure duration, the extra computational effort and time to calculate it can be eliminated.

4.4 Generating the reservoir level curve lookup table

Apart from simulating the reservoir failure behaviour caused by consumer demand parameters only, the pre-run function is also responsible for creating a distribution curve of the storage reservoir level at the beginning of the week, which is produced in the form of a lookup table. This distribution curve will be needed in the events run. When a fire or pipe failure event occurs, the events run proceeds to the beginning of that week in which the event takes place and extracts the initial reservoir level from the distribution curve or lookup table. Once the initial reservoir level has been determined, the events run continues to simulate the critical period in full.

The reservoir level distribution curve at the start of the week was estimated for Sunday at 04h00. This time was chosen as the starting point for each event simulation as it is on the first day of the week, and in the early hours of the morning there is little water usage while the reservoir is filling up, therefore there is a high likelihood that the storage reservoir is full.

The first step to producing the distribution curve is to record the reservoir level when the time is equal to Sunday 04h00. If the reservoir is full, then it is recorded as a proportion of the total number of data points, referred to in this study as the *full fraction*. Otherwise, if not full, the level is recorded in a table with ten bins. The frequency of reservoir levels in each bin is determined to produce a lookup table of *Reservoir level (% full)* against the *Probability of nonexceedance*. An illustration of the full data set for the reservoir level distribution in the base system is shown in Figure 4.4. Figure 4.5 shows a graphical representation of the resultant lookup table (red dashed line), which accurately portrays the reservoir level distribution when the reservoir is not full (black solid line).

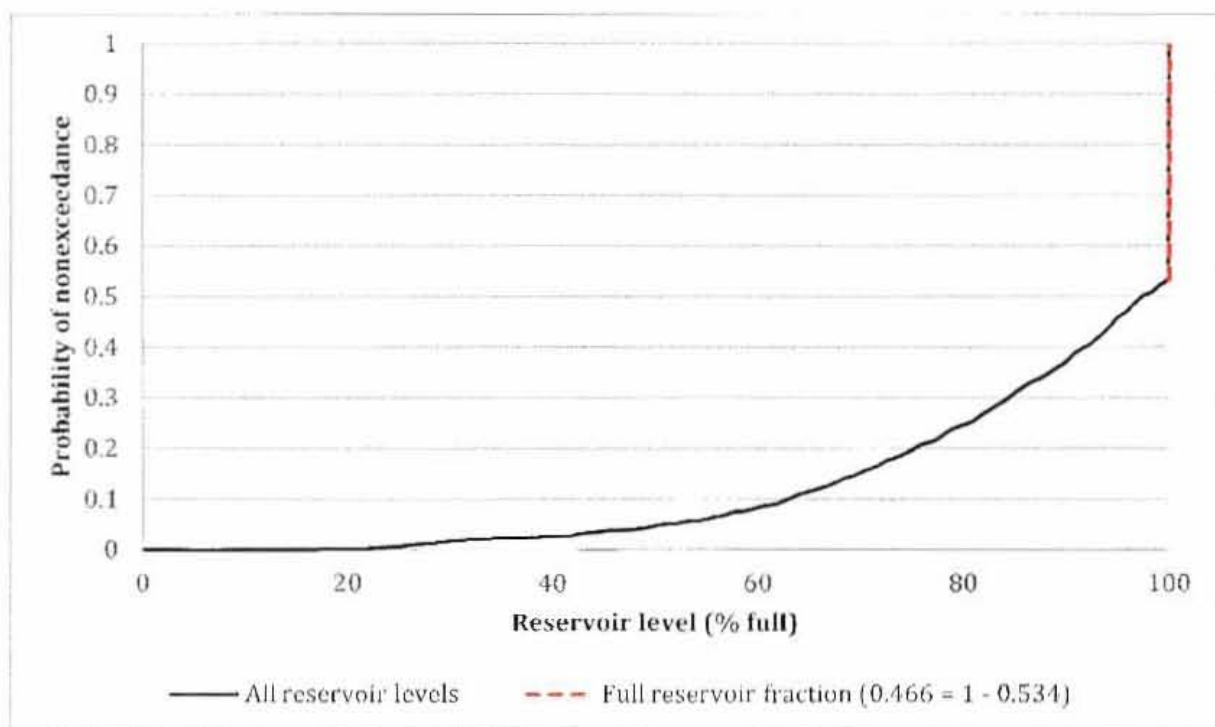


Figure 4.4 Reservoir level distribution at Sunday 04h00 for the example system

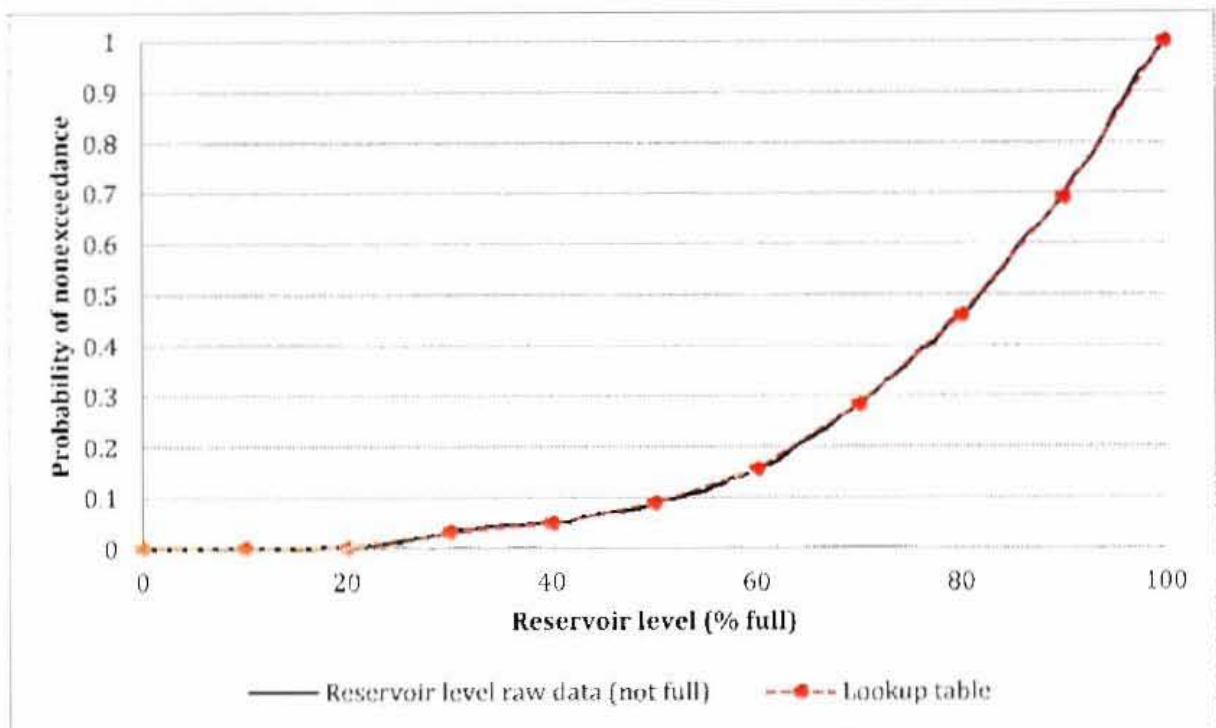


Figure 4.5 Reservoir level lookup table at Sunday 04h00 for the example system

4.5 Conditions for termination of the simulation

When using Monte Carlo simulation, one of the issues to decide on is: How many samples are required to give reliable results? Or rather, how many data points (or trials) are enough? By answering this question, it can be determined when the simulation of the demands-only pre-run or the events run should be terminated. An investigation into existing theorems was carried out to determine which convergence criteria would be most appropriate for this study.

To begin with, it is necessary to establish why an appropriately sized sample is required. If the sample size is too small, not enough samples will be obtained, meaning that the estimate will be too inaccurate to be useful. On the other hand, if the sample size is too large, the estimate may take a longer time to compute.

4.5.1 Convergence theorems and rules-of-thumb

The *law of large numbers* is one of several fundamental theorems expressing the idea that as the number of trials of a random process increases, the percentage difference between the actual and expected values tends to zero (Renze & Weisstein, 2011). This law indicates that when trying to estimate the failure frequency of a certain storage reservoir, the more failures that are generated in the simulation, the closer the estimated failure frequency will be to the real value.

Schuyler (1997) suggests the use of the *standard error of the mean (SEM)* statistic as a Monte Carlo stopping rule:

$$SEM = \frac{s}{\sqrt{n}} \quad (4.3)$$

where s is the standard deviation of the variable of interest, and n is the number of trials in the simulation. Schuyler (1997) recommends that a good rule-of-thumb stopping rule is to run the simulation until the standard error of the mean is less than 1% of the mean. Although this method seems like a good starting point, it is not commonly used in the literature. It also fails to provide an estimate of the confidence level of the results.

In a number of texts on probability (Snedecor & Cochran, 1989; Ott & Longnecker, 2008), it is stated that to find the sample size for a variable, the following equation can be used:

$$n = \left(\frac{z \cdot s}{E}\right)^2 \quad (4.4)$$

where n is the sample size, E is the allowable error, z is the z -value corresponding to the selected level of confidence and s is the standard deviation. This confidence interval method appeared to be worth investigating further as it allows the user to select an allowable error, as well as select the level of confidence for the estimate.

4.5.2 Convergence criteria used in previous stochastic models

In the stochastic analysis model by van Zyl *et al.* (2008), it was found that the reservoir failure properties were consistently within 5% of the ultimate values when the number of reservoir failures exceeds 2000; and consequently, all results were based on a minimum of 2000 reservoir failure events. Similarly, Chang & van Zyl (2010) terminated the stochastic analysis model in their study when 2000 failure results were obtained or the run duration was reached. Van Zyl *et al.* (2008) investigated the failure results further to find that the data followed a Weibull distribution which could be used to formulate an expression for the minimum number of reservoir failures required to ensure reliable results. This expression was, however, based on failure duration, which has been excluded from this study. Therefore, the expression was not considered for use in the simulation. Instead, an alternative expression based on the confidence interval method was used.

4.5.3 Termination criteria for this study

As previously discussed, it is necessary to determine how accurate or precise the estimate should be, and therefore a decision must be made on how large an error can be tolerated. This requires careful thinking about how the estimate will be used and also the consequences of a significant error. In most applications an allowable error of 5% is used.

The allowable error must then be expressed in terms of confidence limits. For 95% confidence limits computed from a sample mean, assumed normally distributed:

$$\bar{y} \pm \frac{1.96S}{\sqrt{n}} \quad (4.5)$$

where \bar{y} is the sample mean and S is the standard deviation.

For the purposes of this study, it was decided to reformulate the above expression so that it could be used as a convergence criterion for the simulation. Assuming that a 5% error would be allowed, the new convergence criterion is defined as follows:

$$\frac{1.96S}{\sqrt{n}}/\bar{y} \leq 5\% \quad (4.6)$$

This expression ensures that the simulation is terminated when the estimated failure frequency falls within 5% of its true value.

The above expression cannot be used as the sole criterion for termination of the simulation. This convergence criterion works effectively for bulk water supply systems with low reliability, and thus reservoirs with a large number of failures. In the case of a highly reliable system, the storage reservoir may seldom fail or not fail at all, and the failure result may take an extremely long time to converge to within 5% of its true value or it might never converge. As a result, additional termination criteria must be used.

The next termination criterion that was used depends on the maximum desired reliability of the system. The maximum desired reliability is a limit set by the designer, and ensures termination of the pre-run or events run for systems with high reliability. The termination criterion makes use of the upper confidence limit. If the upper confidence limit is greater than the maximum time-to-failure set by the designer, then the run is forced to stop. This criterion is expressed as:

$$\bar{y} + \frac{1.96S}{\sqrt{n}} > \max TTF \quad (4.7)$$

where \bar{y} is the mean time-to-failure, S is the time-to-failure standard deviation, n is the number of failures, and $\max TTF$ is the maximum time-to-failure.

In the study by van Zyl *et al.* (2008), a design criterion of one failure in 10 years under seasonal peak conditions was proposed. If this design criterion is used then the designer is interested in a level of reliability of 10 years between each failure. It is recommended then that the designer sets the maximum time-to-failure to be larger than the interested level of reliability, such as $\max TTF$ equal to 100 years (i.e. one failure in 100 years).

The designer is also given the option of directly dictating the simulation duration of the pre-run and events run. This run duration is entered in terms of the number of days of simulation. The designer is allowed this option so that a fixed period of interest can be analysed. However, allowing the designer to dictate the simulation duration also ensures that if neither of the above expressions causes the simulation to end, then it will when the duration selected by the designer is reached. This was the final condition for termination.

4.5.4 Exclusion of reservoir level convergence as a termination criterion

In the pre-run, two outputs are generated, namely the demands failure frequency and the reservoir level distribution curve. To ensure that the demands failure frequency result is reliable, an expression was developed to ensure that it converges. The same expression can be used to obtain convergence of the reservoir level distribution. Thus, if the termination criteria for the pre-run were to include convergence of the reservoir level distribution, then the program would only end once both the demands failure frequency and reservoir level distribution converge.

In this subsection, the feasibility of using the convergence of the reservoir level distribution as a termination condition for the pre-run is investigated. This investigation is necessary to determine whether the accuracy and reliability required for the reservoir level distribution curve outweighs the extra computational resources needed to achieve this accuracy.

For this investigation, the failure frequencies and the reservoir level distribution lookup tables were compared for the pre-run (i) when both the failure frequency and the reservoir level distribution curve converge, and (ii) when only the failure frequency converges. These comparisons were made for a set of supply ratios, $SR = \{0.8; 0.9; 1.0; 1.1; 1.2; 1.3; 1.4\}$, and a set of reservoir capacities, $RC = \{4; 6; 8; 10; 12; 14; 16\}$ h. Also noted, were the times taken for the pre-run to converge under both instances.

It was found that for moderate supply ratio values (SR between 1 and 1.4), either the failure results take longer to converge than the reservoir level, or both values do not converge by the time the period of analysis is completed. This means that both convergence scenarios (i) and (ii) will have the same demands failure frequency and the same reservoir level lookup table. For large supply ratios (SR approximately greater than 1.4), the reservoir is never not full, that is, it is always full on Sunday 04h00. As a result, no reservoir level can be measured for the distribution curve. When examining the results for very small supply ratio values (SR less than 1.0), it was found that the failure results converge before those of the reservoir level. It is for these small supply ratios that reservoir level convergence is relevant. The systems where the supply ratios are less than one are not feasible solutions; nevertheless they are investigated because they may appear in the pre-run grid for single pipe systems that have a small diameter.

The question that must be answered is whether using convergence of failure results only is sufficient, without allowing the reservoir level distribution curve to converge. In Table 4.3, the percentage difference of the failure frequencies between the two convergence scenarios is presented. For the systems tested, the percentage difference between results is less than 5%, with the maximum difference between the two convergence scenarios being 2.91% (when $SR = 0.9$ and $RC = 16$ h). This indicates that the demands failure frequency is not highly affected by the convergence of the reservoir level distribution.

Table 4.3 Percentage difference of failure frequencies

SR	RC (h of storage)						
	4	6	8	10	12	14	16
0.8	1.04	0.91	1.49	1.17	1.99	1.87	1.69
0.9	0.83	0.31	0.15	0.07	1.17	1.99	2.91
1.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.2	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00

A comparison of the lookup tables obtained shows that for a couple of the bins in some of the systems the percentage difference is quite large. For example, for a system with a supply ratio of 0.8 and reservoir capacity of 8 h, the difference between convergence scenario (i) and (ii) is 44.55% in bin 1 (Table 4.4). However, for the rest of the bins, the percentage difference is less than 5%. The percentage difference does not exceed 5% in other systems, as in the case for a system with supply ratio of 0.9 and reservoir capacity of 16 h (Table 4.5). Note that the large percentage difference of 44.55% in Table 4.4 actually corresponds to an absolute difference of 0.0495, which is fairly insignificant. Thus it appears that it is not necessary that the reservoir level has to converge, and that the reservoir level lookup table obtained from the pre-run just after the failure results converge is satisfactory. A full set of reservoir level distribution comparisons is available in Appendix A.

Table 4.4 Percentage difference - reservoir level lookup table ($SR = 0.8$ & $RC = 8$ h)

Bin	(i)	(ii)	%
1	0.0616	0.1111	44.55
2	0.7773	0.7407	4.94
3	0.9905	0.9630	2.86
4	0.9953	0.9630	3.35
5	0.9953	0.9630	3.35
6	0.9953	0.9630	3.35
7	0.9953	0.9630	3.35
8	0.9953	0.9630	3.35
9	1.0000	1.0000	0.00
10	1.0000	1.0000	0.00

Table 4.5 Percentage difference - reservoir level lookup table ($SR = 0.9$ & $RC = 16h$)

Bin	(i)	(ii)	%
1	0.2971	0.3082	3.60
2	0.9710	0.9686	0.25
3	0.9964	0.9937	0.27
4	0.9964	0.9937	0.27
5	0.9964	0.9937	0.27
6	0.9964	0.9937	0.27
7	0.9964	0.9937	0.27
8	1.0000	1.0000	0.00
9	1.0000	1.0000	0.00
10	1.0000	1.0000	0.00

From this investigation, it can be concluded that convergence of failure results is the most important factor to consider when determining the termination criteria of the pre-run simulation. It is not necessary to have the convergence of the reservoir level distribution as a termination criterion. To include it as such, might result in the pre-run taking longer to run under some circumstances, such as for larger supply ratio values.

4.5.5 Summary of termination conditions

The investigation of convergence theorems and Monte Carlo stopping rules assisted in determining how long each simulation should be. Through research into the use of confidence limits and sample sizes in statistical methods, and the testing of possible conditions for termination, the conditions for termination of simulations in this study were established. In summary, a simulation is terminated if one of the following conditions is met:

1. $\frac{1.96S}{\sqrt{n}}/\bar{y} \leq 5\%$;
2. $\bar{y} + \frac{1.96S}{\sqrt{n}} > \max TTF$; or
3. Time $t >$ duration chosen by the designer.

These termination conditions are used in the demands pre-run, as well as the main events run.

4.6 Curve fitting to demands failure frequency data

As part of the pre-run, failure frequency results were generated for a matrix of supply ratios and reservoir sizes. The program is then required to produce a failure frequency equation for each of these supply ratios, such that the equation's input variable will be the reservoir size and the output variable will be the demands failure frequency. These equations are obtained by fitting a curve to the demands failure results for reservoir sizes with the same supply ratio.

A couple of curve fitting methods were explored to determine which would be best for this application. From the study by van Zyl *et al.* (2008), it was found that an exponential curve could be fitted to the number of failures per annum versus the reservoir capacity. Therefore, it was decided that the use of an exponential model, which is a nonlinear model with the form $y = Ce^{Ax}$, would be explored. Initially, the in-built curve fitting tools in Matlab were used to fit an equation to the failure data. To verify the equations produced by Matlab, the Excel curve fitting tool was also used to fit an equation. It was surprising to find that different equations were given as solutions to this problem. After further investigation, it was discovered that Excel and Matlab make use of different curve fitting methods for nonlinear models.

Excel uses the data linearisation method. This involves using a log transform and fitting a linear fit on the log scale. So an exponential curve of the form $y = Ce^{Ax}$ is transformed into a linear equation $Y = AX + B$, where $Y = \ln y$, $X = x$, $B = \ln C$. On the other hand, Matlab fits using a nonlinear fit on the original scale, but with the linear fit as an initial estimation. This curve fitting technique is known as the nonlinear least squares method.

This section investigates which of the two methods is most appropriate for the application of fitting an exponential equation to the demands failure frequency.

4.6.1 Argument against using a linear fit

One of the flaws of fitting nonlinear models by transforming to linearity is that it can be highly affected by outliers. In some cases, an observation might not be an outlier in the original data, but once it undergoes a log transform, the y -value of that observation may be ‘stretched out’ more than the neighbouring points. This is because the log is a very nonlinear transform. As a result symmetric measurement errors on the original scale become more asymmetric on the log scale (MathWorks, 2011). A linear fit on the log scale will thus be affected greatly by that outlier.

This discussion provides an explanation for why the two fits are different. However, a question which follows naturally from this discussion is: which fit is ‘correct’? MathWorks’ (2011) answer is that in the log transform, the y -values of the observations are affected by multiplicative errors which are not symmetric, and therefore least squares on the original scale would not be appropriate. However, if a log transform was used, the errors would be symmetric on the log scale, and the linear least squares fit on the log scale is appropriate (MathWorks, 2011). Accordingly, the ‘correct’ method then depends on the assumptions made about the data. If the noise term (error) is small in comparison to the trend, the log transform is ‘locally linear’ insofar as the y -values near the same x -values will not be stretched out too asymmetrically, which means that the two methods would basically give the same would fit. However, if the noise term is significant, sensible assumptions must be made and a suitable fitting method must be chosen (MathWorks, 2011).

The question at this point is to determine which curve fitting method produces the better fit and is the most appropriate for use in predicting the failure results.

4.6.2 Comparison of the two curve fitting techniques

A set of failure frequency results was obtained for the purpose of the investigation. For each supply ratio in the set $SR = \{1.1, 1.2, \dots, 1.6\}$, two sets of data were selected for curve fitting, namely (1) all the available data for reservoir sizes from 1 hour to 24 hour capacity – where the failure frequency is not zero; and (2) the converged data only – where the convergence ratio of the failure frequency is less than or equal to 5%. These two datasets were selected so that it could be determined how accurate each curve fitting technique is in comparison to the other.

The log transform and nonlinear least squares methods were used to fit an exponential model to the two sets of data for each supply ratio, and the resulting exponential equations were recorded. Both exponential equations were plotted on a graph with the observed failure data, so that the linear fit and nonlinear fit could be compared visually.

Although six supply ratios ($SR = 1.1$ to 1.6) were investigated, for the sake of conciseness, only the results for $SR = 1.2$ will be discussed (as a typical example). These observations are generally applicable to all six cases. From Figure 4.6, it is clear that for reservoirs with smaller capacity from approximately 1 to 4 hours of storage, the nonlinear fit provides a closer approximation to the observed failure frequency than the linear fit. For larger reservoir sizes (4 to 18 hours of storage), the linear fit provides a much closer approximation, while the nonlinear fit produces slightly more conservative failure frequency values. Generally, it appears that the nonlinear least squares method provides a better overall fit than the log transform method.

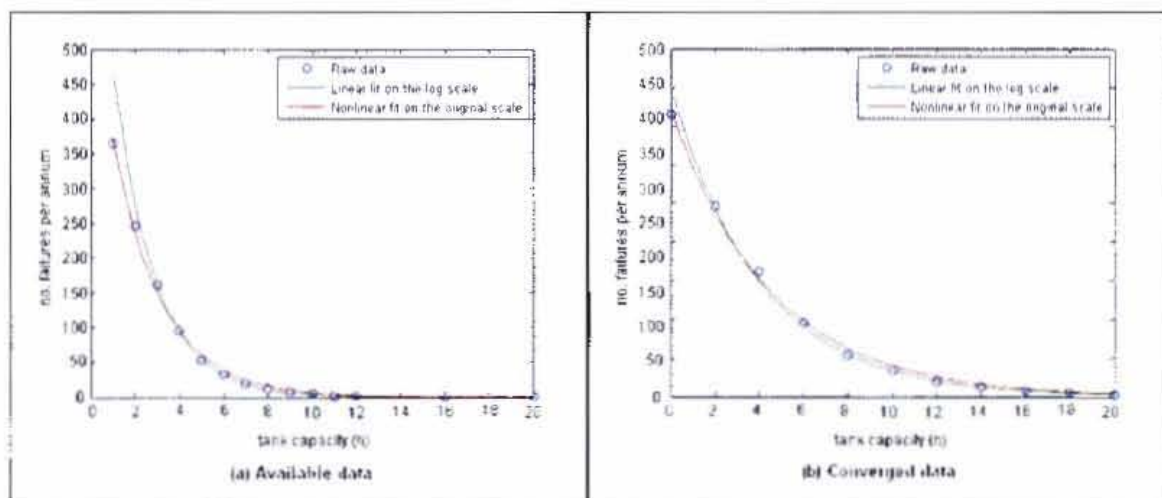


Figure 4.6 Curve fitting using the log transform and nonlinear least squares method on (a) all available data, and (b) the converged data

Besides seeing how closely the predicted values related to the observed values for the failure frequency, the actual exponential equations that were generated were studied. These were examined so that a confirmation could be made whether the nonlinear least squares method would be more appropriate or not. In addition, it would be used for further confirmation of whether the data used for curve fitting needs to have converged or not.

The exponential equations for curves fitted to all the available data are:

$$\text{Linear fit:} \quad y = 816.0783e^{-0.5364x} \quad (4.8)$$

$$\text{Nonlinear fit:} \quad y = 582.2837e^{-0.4494x} \quad (4.9)$$

The exponential equations for curves fitted to the converged data only are:

$$\text{Linear fit:} \quad y = 661.7805e^{-0.4986x} \quad (4.10)$$

$$\text{Nonlinear fit:} \quad y = 582.1811e^{-0.4493x} \quad (4.11)$$

As can be seen above, the linear fit with the log transform method provides two very different exponential equations for curves fitted to the available data (Equation 4.8) and converged data (Equation 4.10). This is a marked contrast to the exponential equations produced by the nonlinear least squares method (Equations 4.9 and 4.11), as the coefficients in both nonlinear fit equations are very similar. This result indicates that the nonlinear least squares technique provides a more stable and accurate fit, and that it is not as volatile as the log transform method. This result is also an indication that even if all the failure results have not converged fully, they can still be used for the purposes of curve fitting if the nonlinear fit is used. However, it is recommended that, where possible, only converged data is used for curve fitting purposes. It is concluded that the most appropriate curve fitting technique for this application is the nonlinear least squares method.

5 Events Run

This chapter gives an overview of the events run. It provides a brief description of the methodology used to find the events failure frequency. The sections in this chapter present a more detailed understanding of the method, as well as a report of the issues encountered during the development of the events function.

5.1 Introduction

The compression heuristic model consists of an events run in addition to the pre-run. The premise behind the compression heuristic is that the computational effort taken to find the reliability of a bulk water supply system using full stochastic analysis can be reduced. It aims to achieve this by using the main events run to fully simulate the critical time periods in which there is a fire or pipe breakage affecting the operation of the storage reservoir. It is assumed that most reservoir failures happen in these periods because of the combined effect of high consumer demand, and demand for fire-fighting purposes or loss of supply due to a pipe break. These critical periods form only a fraction of the simulation duration for the full stochastic analysis, therefore it is expected that the computational effort and time taken to run a full simulation can be diminished.

Since only the critical periods are simulated, it means that the intermediate time between events is not simulated in the events run. The output of the events run is thus the events failure frequency. Even though the events run does not simulate the time between events, this intermediate time, when only consumer demand affects the reservoir operation, cannot be disregarded. Therefore, a pre-run function, which was discussed in length in Chapter 4, is used to determine the demands failure frequency for the system. To measure the bulk supply system's total reliability in terms of failure frequency, the demands failure frequency is combined with the events failure frequency obtained from the events run. The development of the events run is presented in this chapter.

5.1.1 Description of the events algorithm

Figure 5.1 shows a flowchart of the algorithm for the events run. Unlike the demands pre-run which simulates the system behaviour for a matrix of supply ratios and reservoir capacities, the events run finds the events failure results for a single supply ratio and reservoir capacity.

The first step in the events run consists of the initialisation of system parameters related to the results obtained from the pre-run function. The demands failure frequency for the current system is calculated from the exponential curves fitted to the pre-run demand failures. In addition, the full fraction for the reservoir and the lookup table of the reservoir level distribution are obtained. The supply inflow into the storage reservoir and the reservoir volume are also initialised.

After initialisation, the first event is generated. The events run goes to the first event and the time t is set to the hour of the beginning of the event. At this point, the initial reservoir level at the start of the event (Sunday 04h00 = day 1, hour 4) is retrieved. The hourly consumer demand, fire demand and supply inflow into the reservoir are then calculated. The reservoir level is updated and any system failures are recorded. A check is made to see if the event has ended. If the conditions for ending the current event are met, then the stochastic model continues with the next event.

Once the events run meets the conditions for terminating the simulation, then the events failure frequency and the demands failure frequency are combined to give the total failure frequency for the system.

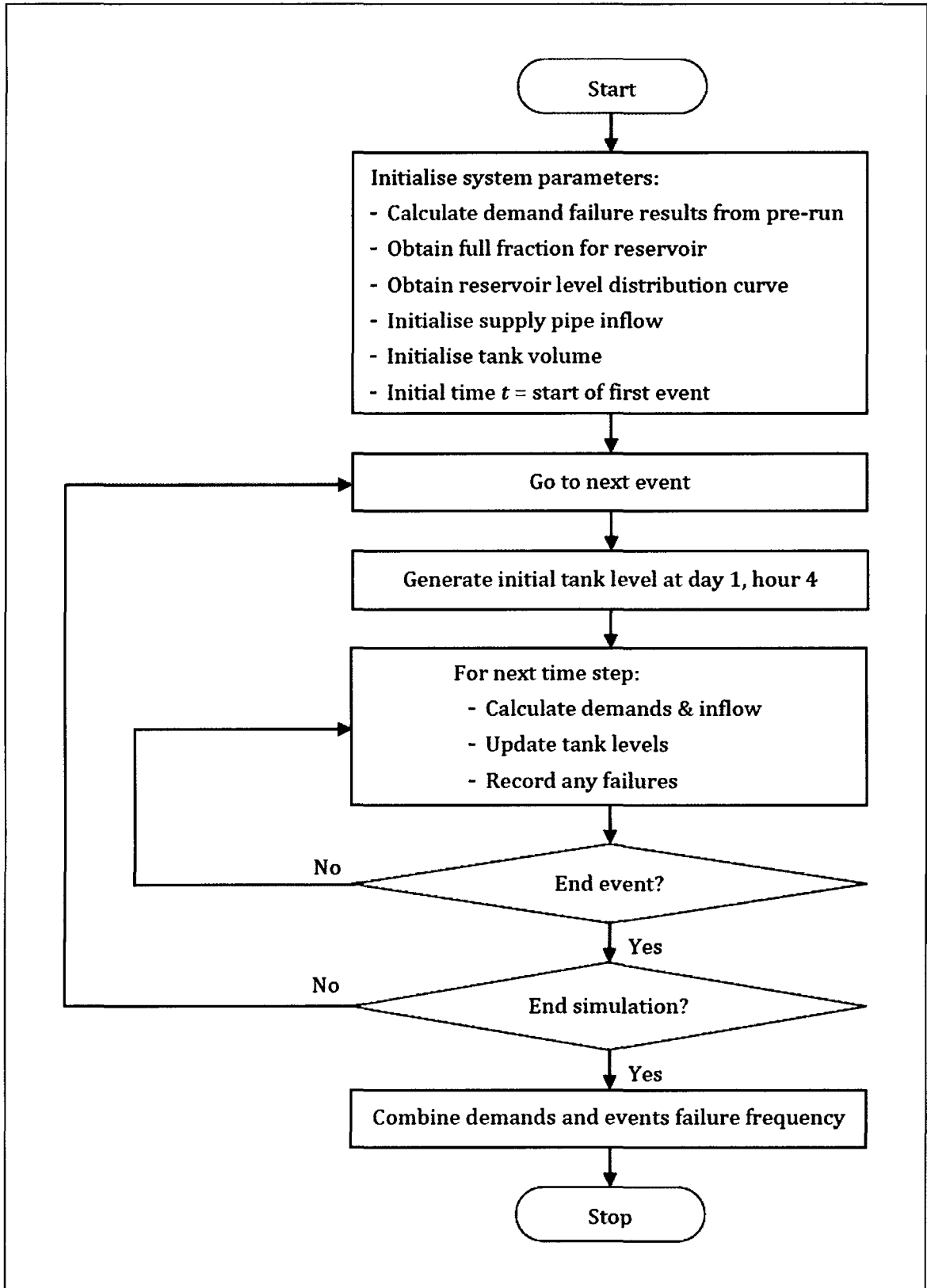


Figure 5.1 Flowchart for the events run

5.2 Input variables and initialisation

The events run is different to the demands-only pre-run in that it does not perform reliability analysis on a grid of systems. It analyses the reliability of a single system with an individual supply ratio and reservoir capacity as input values instead. However, the pre-run output results for the grid of systems are needed as input variables in the events run. These input variables are the tables for the demands failure frequency, the full reservoir fraction and the reservoir level distribution curve.

The demands failure frequency, full reservoir fraction and lookup table for the reservoir level distribution must be initialised for the system that is to be analysed. The system will have a given supply ratio and reservoir capacity that lies in the range of values evaluated in the pre-run. The range of supply ratios and reservoir capacities analysed in the pre-run are discrete, therefore the results are also discrete. It is possible that the system analysed in the events run has supply ratio and reservoir capacity values that do not lie precisely on one of the discrete values in the pre-selected range. Consequently, the events run must find the demands failure frequency, full reservoir fraction and reservoir level lookup table for the given supply ratio and reservoir capacity.

For the given supply ratio, two scenarios were identified:

- The given supply ratio is equal to a discrete value in the pre-run grid; or
- The given supply ratio lies between two discrete values in the pre-run grid.

Similarly, two scenarios were identified for the given reservoir capacity:

- The given reservoir capacity is equal to a discrete value in the pre-run grid;
- The given reservoir capacity lies between two discrete values in the pre-run grid.

Depending on which scenario is applicable for the supply ratio and reservoir capacity, different methods were used to find the demands failure frequency, full fraction and reservoir level lookup table.

There are four possible combinations of the different scenarios described above:

- 1) Both the given supply ratio and reservoir capacity are equal to discrete values in the pre-run grid;
- 2) Only the supply ratio is equal to a discrete value in the pre-run grid;
- 3) Only the reservoir capacity is equal to a discrete value in the pre-run grid;
- 4) Both the given supply ratio and reservoir capacity lie between two discrete values in the pre-run grid.

As an example to explain the parameter initialisation in the events run, suppose the pre-run grid has the supply ratios, $SR = \{1.1; 1.2\}$, and the reservoir capacities, $RC = \{4; 8\}$ h. The

pre-run computes the exponential equations for the demands failure frequency as the following:

- For $SR = 1.1$: $y = 603.833e^{-0.3311x}$ (5.1)

- For $SR = 1.2$: $y = 726.1939e^{-0.5092x}$ (5.2)

The full reservoir fraction for each pair of supply ratio and reservoir capacity is shown in Table 5.1.

Table 5.1 Full reservoir fraction for example

		RC (h)	
		4	8
SR	1.1	0.2890	0.1621
	1.2	0.5725	0.5012

The reservoir level distribution lookup table for each pair is shown in Table 5.2 to Table 5.5.

Table 5.2 Reservoir level distribution for example, $SR = 1.1$ and $RC = 4$ h

Bin	Probability of nonexceedance	Initial level (h of storage)
1	0.4	0
2	0.8	0
3	1.2	0.0036
4	1.6	0.0036
5	2.0	0.0468
6	2.4	0.1906
7	2.8	0.4496
8	3.2	0.7590
9	3.6	0.9209
10	4.0	1

Table 5.3 Reservoir level distribution for example, $SR = 1.1$ and $RC = 8$ h

Bin	Probability of nonexceedance	Initial level (h of storage)
1	0.8	0
2	1.6	0.0014
3	2.4	0.0637
4	3.2	0.3029
5	4.0	0.4523
6	4.8	0.5742
7	5.6	0.6939
8	6.4	0.7997
9	7.2	0.9148
10	8.0	1

Table 5.4 Reservoir level distribution for example, $SR = 1.2$ and $RC = 4$ h

Bin	Probability of nonexceedance	Initial level (h of storage)
1	0.4	0
2	0.8	0
3	1.2	0
4	1.6	0
5	2.0	0.0056
6	2.4	0.0280
7	2.8	0.0924
8	3.2	0.3221
9	3.6	0.6692
10	4.0	1

Table 5.5 Reservoir level distribution for example, $SR = 1.2$ and $RC = 8$ h

Bin	Probability of nonexceedance	Initial level (h of storage)
1	0.8	0
2	1.6	0
3	2.4	0.0086
4	3.2	0.0805
5	4.0	0.2019
6	4.8	0.3050
7	5.6	0.4430
8	6.4	0.6001
9	7.2	0.7835
10	8.0	1

To illustrate what would happen for combination (1), assume that the given supply ratio is 1.2 and the given reservoir capacity is 8 h. The exponential equation for $SR = 1.2$ (Equation 5.2) is retrieved from the pre-run and the reservoir capacity is substituted into the equation to give the demands failure frequency of 12.3569. The full reservoir fraction and reservoir level lookup table are both the same as that produced in the pre-run for the given supply ratio and reservoir capacity.

Now, let the given supply ratio be 1.2 and the given reservoir capacity be 6 h. This is an example of combination (2), where the supply ratio is equal to a discrete value in the pre-run grid, while the reservoir capacity lies between two discrete reservoir capacity values in the pre-run grid. To get the demands failure frequency for the system, the given reservoir capacity is substituted directly into the exponential equation for the supply ratio ($SR = 1.2$), giving a value of 34.2134. To get the full fraction for the system, the full fraction is found for the supply ratio and the first reservoir capacity value from the pre-run grid, in addition to the full fraction for the supply ratio and the second reservoir capacity. From Table 5.1, this gives 0.5725 for $SR = 1.2$ and $RC = 4$ h, and 0.5012 for $SR = 1.2$ and $RC = 8$ h. The final full fraction is obtained through linear interpolation between these two full fractions, giving

0.5369 for $SR = 1.2$ and $RC = 6$ h. The reservoir level lookup table is also obtained in the same way through linear interpolation of the two lookup tables for each of the two reservoir capacities. Figure 5.2 shows the reservoir level distribution curve for $SR = 1.2$ and $RC = 6$ h.

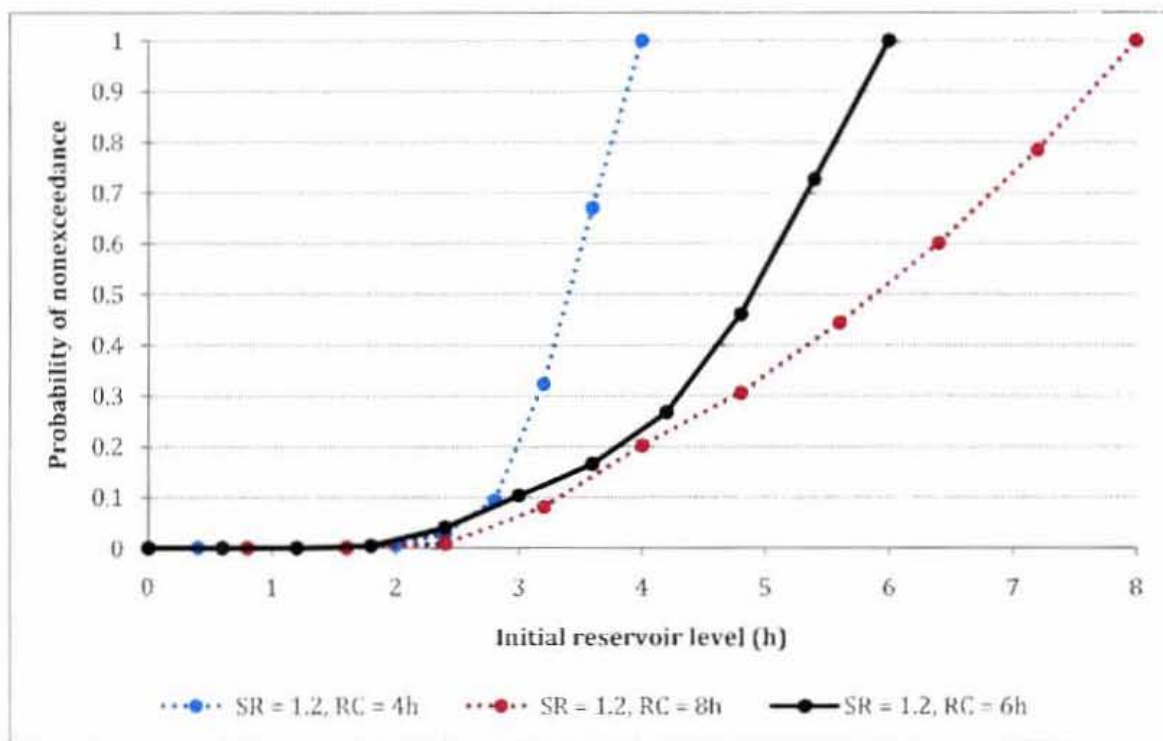


Figure 5.2 Reservoir level distribution curve for example, $SR = 1.2$ and $RC = 6$ h

Combination (3) is similar to combination (2), except there is now one reservoir capacity equal to a discrete value in the pre-run grid, and two discrete supply ratio values that the given supply ratio lies between. As an example, assume the given supply ratio is 1.15 and the given reservoir capacity is 8 h. For each supply ratio, the exponential equation is used to find the demands failure frequency from the given reservoir capacity. The demands failure frequency for a reservoir capacity of 8 h is 42.7128 and 12.3569, for $SR = 1.1$ and $SR = 1.2$ respectively. The system's demands failure frequency is found by linearly interpolating between these two demands failure frequencies to give 27.5349. The full fraction for the first supply ratio and the reservoir capacity (0.1621) is interpolated with the full fraction for the second supply ratio and the reservoir capacity (0.5012), so that the system's full reservoir fraction of 0.3317 can be found. In the same way, using linear interpolation, the reservoir level distribution curve for the system is obtained (Figure 5.3).

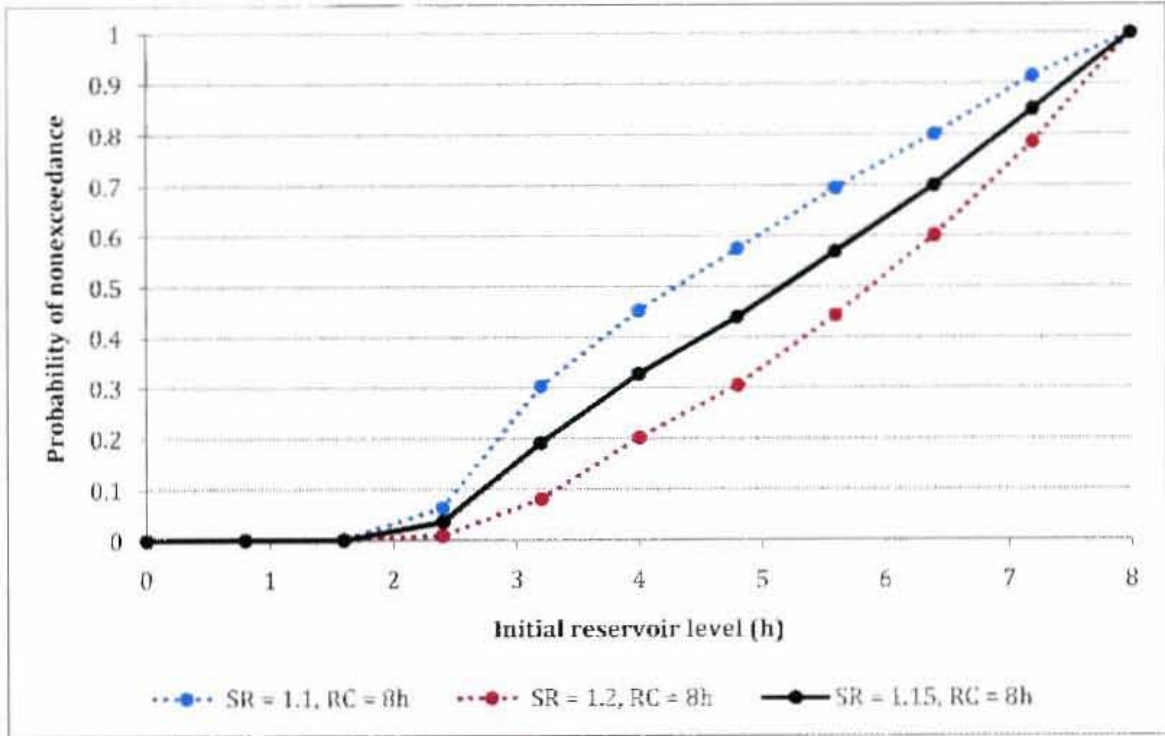


Figure 5.3 Reservoir level distribution curve for example, $SR = 1.15$ and $RC = 8$ h

Both the given values lie between two discrete grid values for combination (4). Let the given supply ratio be 1.15 and the given reservoir capacity be 6 h. The given reservoir capacity of 6 h is substituted into the exponential equation for each of the two supply ratios ($SR = 1.1$ and $SR = 1.2$), to give demands failure frequencies of 82.8223 and 34.2134 respectively. These are linearly interpolated to give the final demands failure frequency of 58.5179. Getting the full reservoir fraction and lookup table for this system is slightly more complicated than above. The full fractions of 0.2890 and 0.1621, for $SR = 1.1$ and each of the two reservoir capacities, $RC = 4$ h and $RC = 8$ h, are interpolated to get the first full fraction 0.2256. The full fractions for the second supply ratio, $SR = 1.2$, and each of the two reservoir capacities are 0.5725 and 0.5012. They are interpolated to get the second full fraction 0.5369. Finally, the first full fraction and the second full fraction are interpolated to get the final value of 0.3813 for the system's full reservoir fraction. The same process is followed to get the system's reservoir level distribution curve (Figure 5.4).

When the demands failure frequency, full fraction and reservoir lookup table for the system have been initialised, then the stochastic analysis model for the events run can be used. The full reservoir fraction and the reservoir level lookup table are both used each time an event occurs in the model, while the demands failure frequency is used at the end of the stochastic model to find the total failure frequency for the system.

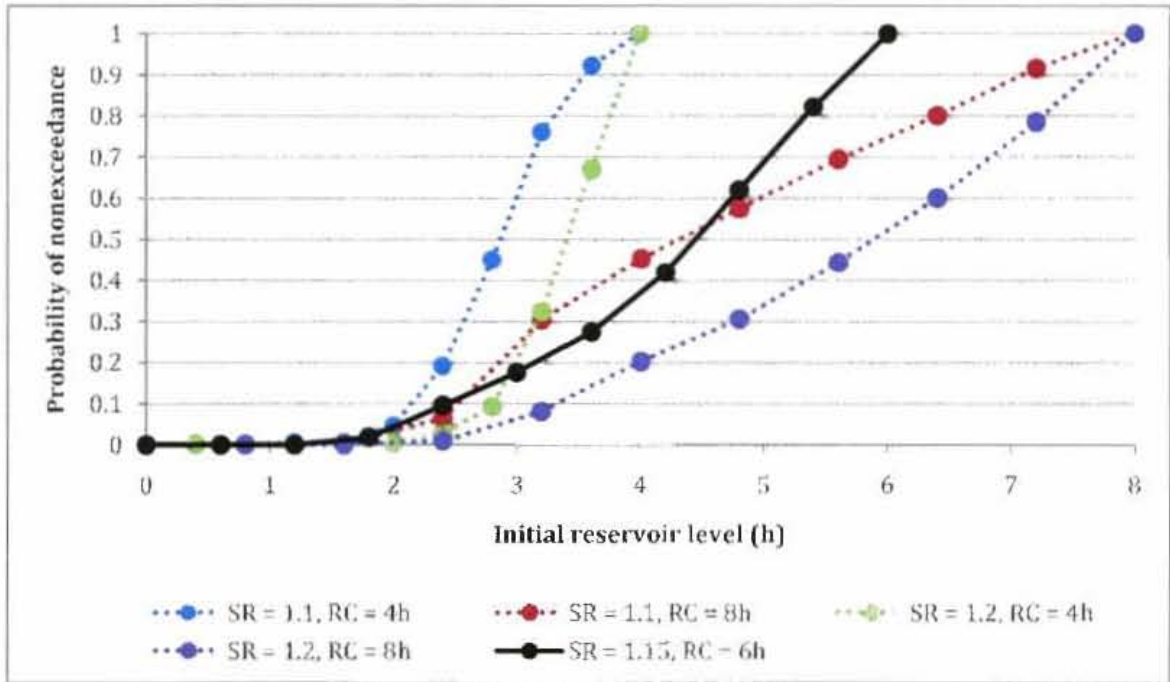


Figure 5.4 Reservoir level distribution curve for example, $SR = 1.15$ and $RC = 6$ h

5.3 Running an event simulation

This section describes how events are generated. There are two types of events that are analysed in the stochastic model along with consumer demand, namely fires and supply pipe failures.

5.3.1 Generating fire events

A fire event occurs when a fire breaks out on the demand side of the storage reservoir. Water is needed for fire fighting purposes; therefore the water demand is increased by the additional fire demand. There are three functions that are necessary for modelling the fire demand in the stochastic model. These are responsible for finding the time to the next fire event, the fire duration and the fire flow. All three functions are based on probability distributions.

A function, referred to in the model as `calc_timeToNextFire`, calculates the time between successive fire events. This function uses an exponential distribution. The probability density function of the exponential distribution is:

$$f(\Delta t; \lambda) = \lambda e^{-\lambda \Delta t}, \Delta t \geq 0 \quad (5.3)$$

where Δt is the time between successive fires (h), and λ is the rate parameter. In this case, the rate parameter is the fire event rate or number of fires per hour, which can be determined by

$$\lambda = 1 / \left(\frac{\text{firesPerAnnum}}{365 \times 24} \right) \quad (5.4)$$

The number of fires per annum (`firesPerAnnum`) can be varied. For the base example in this study, `firesPerAnnum` is taken to be six, which results in a rate parameter λ of 1460.

For each fire event, the fire duration is determined by the function `calc_fireDuration`. It uses a log normal distribution with probability density function:

$$f(T; \mu, \sigma) = \frac{1}{T\sigma\sqrt{2\pi}} e^{-(\ln T - \mu)^2 / 2\sigma^2}, \quad T \geq 0 \quad (5.5)$$

where T is the duration of the event, μ is the mean of the natural logarithms of the duration, and σ is the standard deviation of the natural logarithms of the duration. In this project, the same statistical parameters are used as the study by van Zyl *et al.* (2008), where $\mu = -0.393$ and $\sigma = 0.66$.

The fire flow for each fire event is also modelled with a log normal distribution in the function `calc_fireFlow`. The statistical parameters for the fire flow function are $\mu = 1.31$ and $\sigma = 1.31$, as determined by van Zyl *et al.* (2008).

5.3.2 Generating pipe failure events

A pipe failure event occurs when a pipe breakage happens in the feeder pipe system supplying water into the storage reservoir. When there is a pipe failure, the supply inflow into the storage reservoir is either partially reduced to a fraction of its full flow rate or completely reduced to zero, depending on the pipe configuration. As a result, even if the water demand remains constant, the storage reservoir will still be unable to provide adequate supply to the users.

In the stochastic model, the time between successive pipe failures is determined by the function `calc_timeToNextFailure`, which is also based on an exponential distribution (Equation 5.3). Resembling the expression for the fire event, the rate parameter λ can be determined with the formula:

$$\lambda = 1 / \left(\frac{\text{failuresPerAnnum}}{365 \times 24} \right) \quad (5.6)$$

where `failuresPerAnnum` = 2, for the base example, which gives a rate parameter λ of 4380.

`calc_failureDuration` is the function that calculates the duration of each pipe failure event using a lognormal distribution (Equation 5.5). The statistical parameters in this function are $\mu = 1.49$ and $\sigma = 0.48$ for the base example (van Zyl *et al.*, 2008).

5.3.3 Determining the next event

After both the next fire event and next pipe failure event have been generated, it is necessary to decide which of the two happens first, that is, which event is the first to be simulated in the stochastic analysis model (or `nextEvent`). The next event in the model is the fire event if its start time is earlier than the start time of the pipe failure event, otherwise the next event is the pipe failure event. The beginning of the next event (`beginEvent`) is then set to the beginning of that hour. Thereafter, the end of the event (`endEvent`) must be determined.

Events generated by the fire event functions and the pipe failure event functions may overlap. For instance, it is possible that the pipe failure event may begin before the end of the fire event. If so, then the end of the next event is not the end of the fire event, but rather the end of the pipe failure event.

The three variables `nextEvent`, `beginEvent` and `endEvent` are included in the stochastic analysis model for the computation of the events failure frequency. These variables are updated each time a new fire or pipe failure event is generated.

5.3.4 Event simulation

It is worth emphasising again that the purpose of the compression heuristic is to reduce the computational effort and simulation time required to analyse the reliability of a system. The compression heuristic intends to achieve this by only performing the full stochastic analysis on the critical periods of fire or pipe failure events. This means that the events run in the compression heuristic ‘jumps’ over the intermediate periods between the events.

In the events run, after the next event has been generated, the model jumps to the start of the week of `beginEvent` at Sunday 04h00. At this point the initial volume in the storage reservoir is found using the full reservoir fraction and the reservoir level lookup table (explained in the next section). For each time step until the end of the current event simulation, the consumer demand, fire demand and supply inflow are calculated; after which the reservoir volume is updated and any reservoir failures are recorded. Once the current event simulation is complete, the next event is generated, and a new event simulation is performed.

5.3.5 Conditions for ending the event simulation

Termination conditions are required for each event simulation to ensure that the event simulation does not run indefinitely. Three conditions were identified. All of the conditions, stated below, must be met before the events run can continue to the next event simulation:

1. The event simulation must end at Sunday 04h00, i.e. a complete number of weeks must be simulated;
2. There must be no events in the current time step, i.e. the event must have ended;
3. The reservoir is full again.

Each of these conditions is explained individually in more detail below.

The first condition stipulates that the event simulation must end at Sunday 04h00. Since each event simulation starts at Sunday 04h00, it is necessary that the current event simulation ends at the same time of the potential start time for the next event. This condition guarantees that if the next event starts in the same week as the current event, then the event simulation will just continue until after the end of the next event. It ensures that the start time of the next event does not overlap with the current event, so that no hours are simulated for a second time.

The second condition stipulates that the current time step does not have any fire or pipe failure event. The event must end before the simulation can be completed. If a new fire or pipe failure event occurs before the end of the current event, then the end of that new fire or pipe failure event is assigned to be the new `endEvent`.

When an event simulation starts, it is possible that the storage reservoir is no longer full, if a supply pipe has failed or the combined consumer and fire demand exceeds the supply. A Boolean variable `reservoirFull` is set to zero (false). Condition 3 stipulates that the event simulation can only be complete provided that the storage reservoir has filled up after the end of the last event. This condition ensures that the indirect impact of an event on the reservoir behaviour due to a reduced reservoir level at the end of the event is modelled. Only once the storage level is back to the maximum level is the Boolean variable `reservoirFull` set to one (true).

In each time step of the event simulation, a check is performed to assess whether all the conditions have been met. If the conditions are met, then the current event simulation is complete and the model proceeds to simulate the next event. Otherwise, if any of the three conditions is not satisfied, then the simulation continues to the next time step.

5.4 Obtaining the initial reservoir level

Each event simulation in the events run requires an initial reservoir level as a starting point. Section 5.2 already explained how the events run obtains the full reservoir fraction and lookup table of the reservoir level distribution curve for the current system. This section describes how those values are used to find the initial reservoir level at the beginning of each event.

An event simulation starts at Sunday 04h00 in the week of the start of the current event under analysis. It is at this point that the initial reservoir level of the simulation is required. The

value is found using probabilistic means. A uniform random number between zero and one is generated. If the random number is less than or equal to the full reservoir fraction, then the initial reservoir level is set to the maximum level. If the random number is greater than the full reservoir fraction, then the storage reservoir is not full and the initial level is acquired from the lookup table.

An example will be used to further explain the process of determining whether the reservoir is full or not. Let the full reservoir fraction be 0.4877, where the probability of the reservoir not being full is 0.5123. If the generated random number is 0.48, then the reservoir is at full capacity. However, if the generated random number is 0.50, then the lookup table is consulted for the initial reservoir level.

The lookup table for the reservoir level distribution curve consists of ten bins with associated initial levels in terms of hours of remaining storage. Another uniform random number between zero and one is generated. Depending on which bin the random number lies in, the initial reservoir level can be found from the corresponding range of storage levels in the lookup table. Each bin has an upper and lower limit for the reservoir level. The initial reservoir level for the system is found using linear interpolation for the generated random number.

Figure 5.5 represents the lookup table for the reservoir level distribution. It shows graphically how the initial reservoir level is obtained. If the uniform random number is 0.3, then it lies in the bin that ranges from 0.2935 to 0.4671. By linear interpolation, the initial reservoir level for the system would be 8.4449 hours of storage.

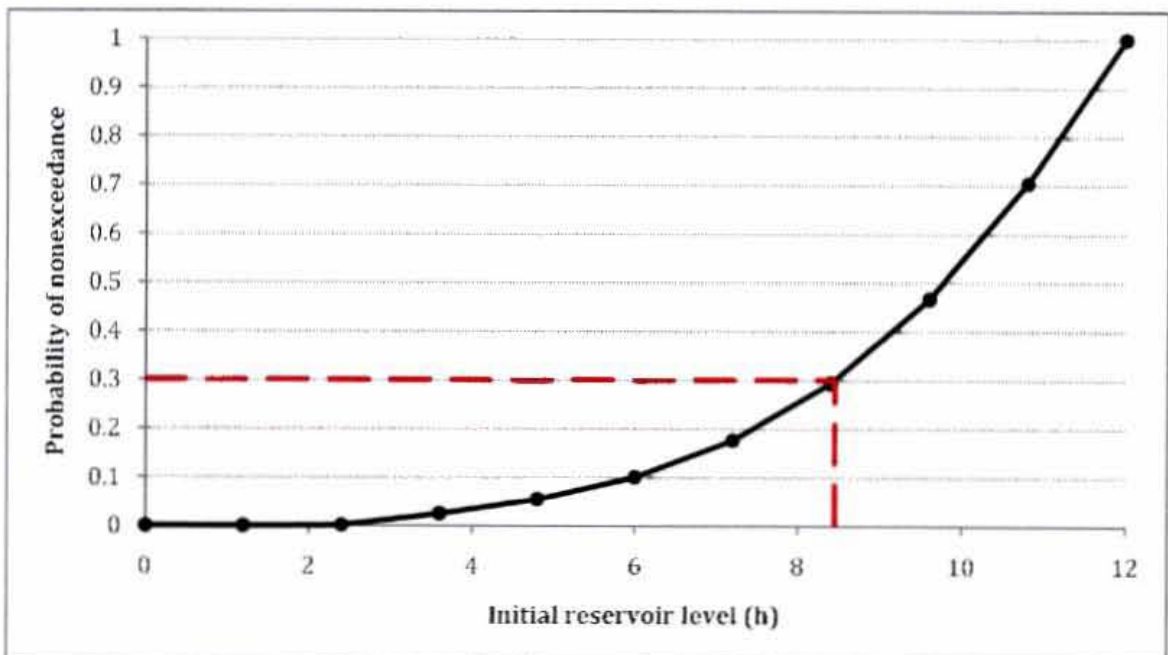


Figure 5.5 Reservoir level distribution curve for example

5.5 Effect of pipe configuration

When designing a bulk water supply system, the designer may want to optimise the configuration of the feeder pipe system into the storage reservoir. Since only the effect of consumer demand on the storage reservoir is examined in the pre-run, the pre-run is not affected by any change in the pipe configuration. In contrast, the supply pipe configuration may have a large impact on the reliability behaviour of the reservoir in the events run. This section explains how the pipe configuration is modelled in the events simulation. It also explores how the reliability of the system is affected by different configurations.

5.5.1 Pipe configuration and reliability

In practice, the pipe layout of a water supply system is usually constrained by the topography of the site, such as the layout plan of roads and streets. A supply pipe connecting a supply source to a storage reservoir in a bulk water supply system may be restricted by similar constraints. In this study, the pipe layout refers to the physical location of the pipe system, while the pipe configuration refers to whether there are parallel pipes or interconnections between parallel pipes.

Depending on the layout of the area, a distribution network may have different configurations, namely looped or branched. Like a distribution network, the pipeline system connecting a supply source to a bulk storage reservoir may have a fixed layout, but its configuration can be varied. Bulk supply pipe systems can consist of single pipes, pipes in parallel, or interconnected parallel pipes. Van Zyl & Haarhoff (1999) researched the effect of feeder pipe configuration on the reliability of bulk water supply systems. They found that the reliability of bulk water supply systems can be significantly improved by increasing the number of parallel pipes, or by increasing the number of interconnections between two or more pipes. This investigation revealed that the increase in the number of interconnecting pipes resulted in a similar improvement in order of magnitude in the system reliability as an increase in the number of parallel pipes; however the option of increasing the number of interconnecting pipes was much cheaper than increasing the number of parallel feeder pipes (van Zyl & Haarhoff, 1999).

5.5.2 Investigation of the effect of pipe configuration

In this study, the effect of different supply pipe configurations on the reliability of the bulk supply system will be examined. The pipe configuration of the system will be varied according to number of parallel feeder pipes and number of interconnections, as suggested by van Zyl & Haarhoff (1999). Seven pipe configurations, illustrated in Figure 5.6 to Figure 5.12, will be explored. The base configuration (1) consists of a single pipeline connecting the supply source to the storage reservoir. Configurations (2) and (3) show two and three pipes in parallel respectively. Configurations (4) and (5) have two and three pipes in parallel with one interconnection dividing each pipe into two equal lengths, while configurations (6) and (7)

have two and three parallel pipes respectively, with two interconnections dividing each pipe into three equal lengths.

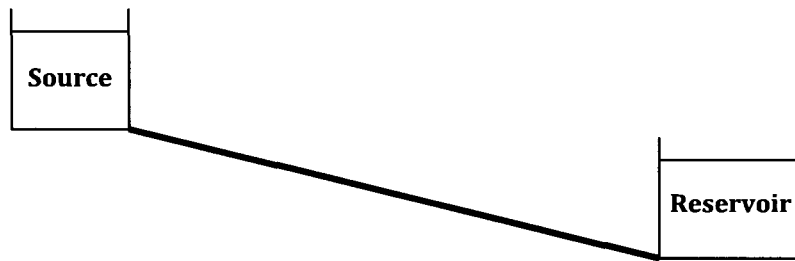


Figure 5.6 Configuration (1) – Single pipe connecting the source to the storage reservoir

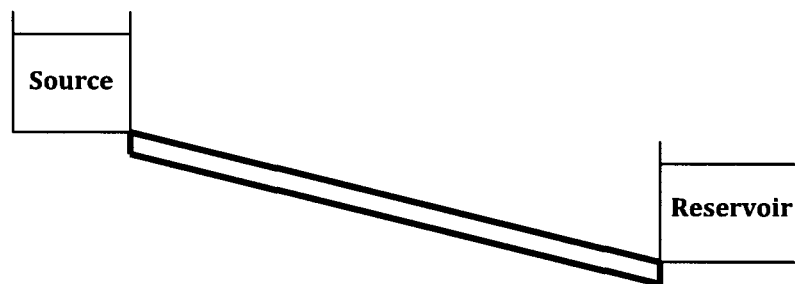


Figure 5.7 Configuration (2) – Two parallel pipes connecting the source to the storage reservoir

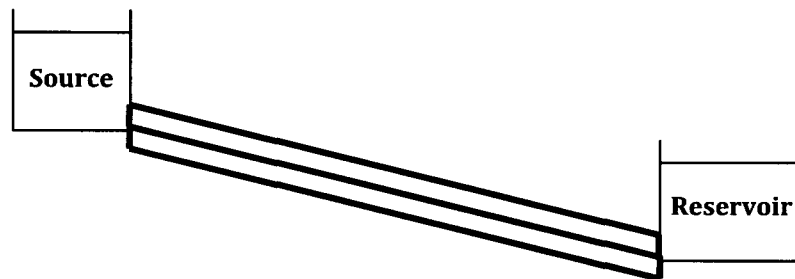


Figure 5.8 Configuration (3) – Three parallel pipes connecting the source to the storage reservoir

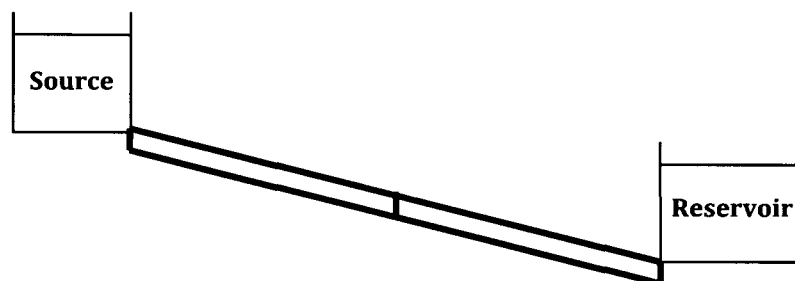


Figure 5.9 Configuration (4) – Two parallel pipes connecting the source to the storage reservoir, with one interconnection dividing the pipes into two equal lengths

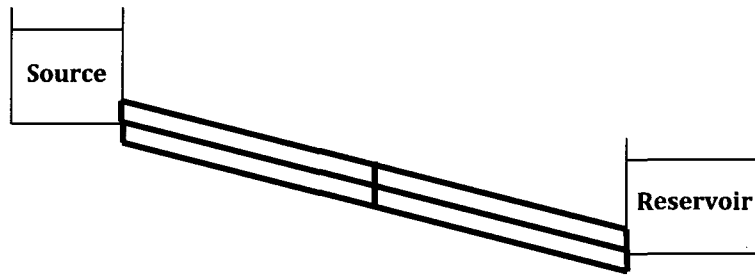


Figure 5.10 Configuration (5) – Three parallel pipes connecting the source to the storage reservoir, with one interconnection dividing the pipes into two equal lengths

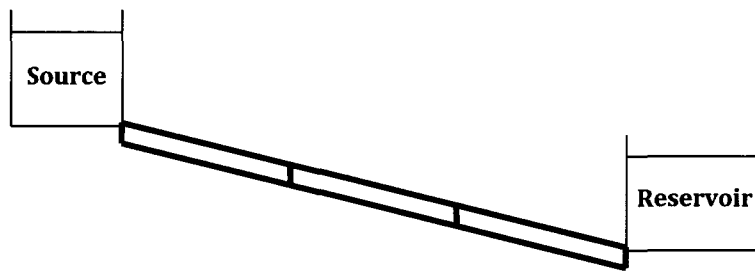


Figure 5.11 Configuration (6) – Two parallel pipes connecting the source to the storage reservoir, with two interconnections dividing the pipes into three equal lengths

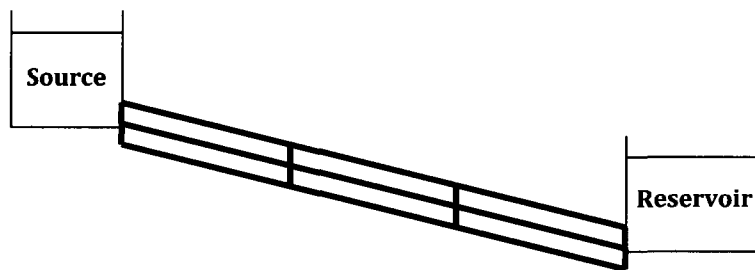


Figure 5.12 Configuration (7) – Three parallel pipes connecting the source to the storage reservoir, with two interconnections dividing the pipes into three equal lengths

It is assumed that the probability of two or more pipe failures occurring at the same time to be negligibly small, thus only one pipe break was allowed at a time. This assumption can be shown to be reasonable. If a 10 km long pipe has a failure rate of five failures per annum with each failure having an average duration of five hours, then the probability of a pipe failure occurring in a single hour is 0.00285. The probability of two failures happening at the same time in a single hour is then 0.000008144, which is equivalent to 0.01426 failures/annum. The probability of two pipe failures happening at the same time in a system with interconnecting pipes between pipes in parallel is negligibly small, and will therefore be disregarded in the simulation. Using this assumption, for each of the configurations it is possible to calculate the total supply inflow into the storage reservoir should a pipe failure happen. The method used to determine the total inflow rate (Q_{in}) will be described in this section.

Firstly, one of the seven configuration options will be selected for each event simulation. For each supply system, the velocity (V) for a single pipeline between the source and reservoir is calculated from the Hazen-Williams headloss formula:

$$V = 0.354C_H D^{0.63} \left(\frac{h_f}{L}\right)^{0.54} \quad (5.7)$$

where V is velocity (m/s), C_H is the Hazen-Williams coefficient dependent on surface roughness, D is the pipe diameter (m), h_f is the friction loss (m), and L is the length of pipe (m). The supply inflow rate can then be calculated from the continuity equation:

$$Q_{in} = V \times A = V \times \frac{\pi D^2}{4} \quad (5.8)$$

where Q_{in} is the supply inflow rate (m^3/s), and A is the cross-sectional area of the pipe (m^2).

At the beginning of the simulation, the user inputs the head difference between the source and reservoir (equivalent to the headloss h_f), as well as the Hazen-Williams coefficient C_H , the length of pipe L , and diameter of pipe D . These input values will be used to calculate the inflow rate Q_{in} . As an example, if $h_f = 30$ m, $C_H = 120$, $L = 1000$ m and $D = 0.207$ m, then $Q_{in} = 80$ l/s. This means that for a pipe configuration (1) consisting of a single pipeline connecting the source to the reservoir, the inflow rate into the reservoir will be 80 l/s. If a pipe failure occurs, then there is zero supply into the reservoir and $Q_{in} = 0$ l/s (Figure 5.13).

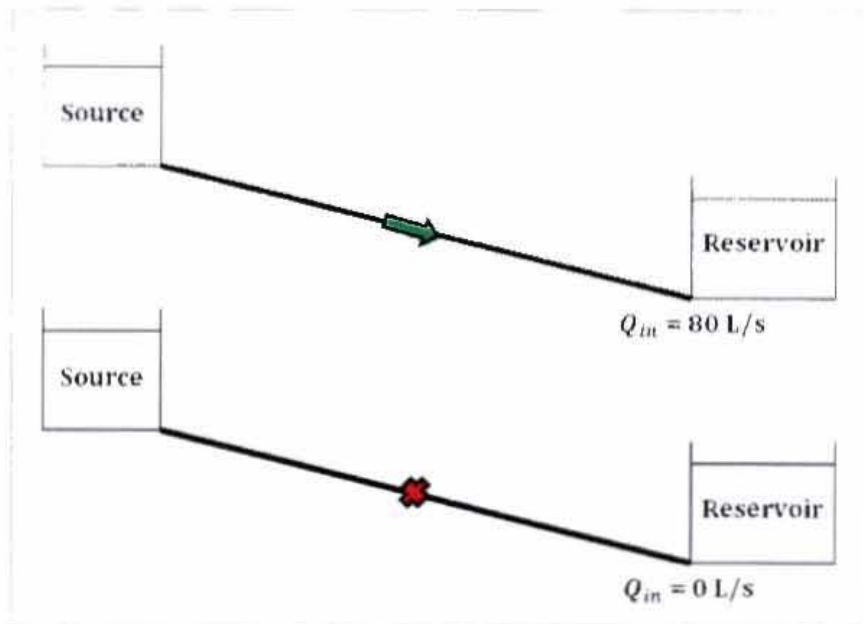


Figure 5.13 Effect of a pipe failure in configuration [1] with $h_f = 30$ m, $C_H = 120$, $L = 1000$ m, $D = 0.207$ m

For pipes in parallel, it is assumed that all of the pipes have the same characteristics. Using the same input values as above, two pipes in parallel will supply water at twice the inflow

rate of the single pipeline at $Q_{in} = 160$ l/s. If a pipe failure occurs and the other pipe is still operating, then the total supply inflow rate will be halved to $Q_{in} = 80$ l/s (Figure 5.14). Similarly, three pipes in parallel will supply at three times the inflow rate of a single pipeline, providing $Q_{in} = 240$ l/s. When a pipe failure occurs, there are still two pipes operating and the total supply inflow rate will be reduced by a third to $Q_{in} = 160$ l/s (Figure 5.15).

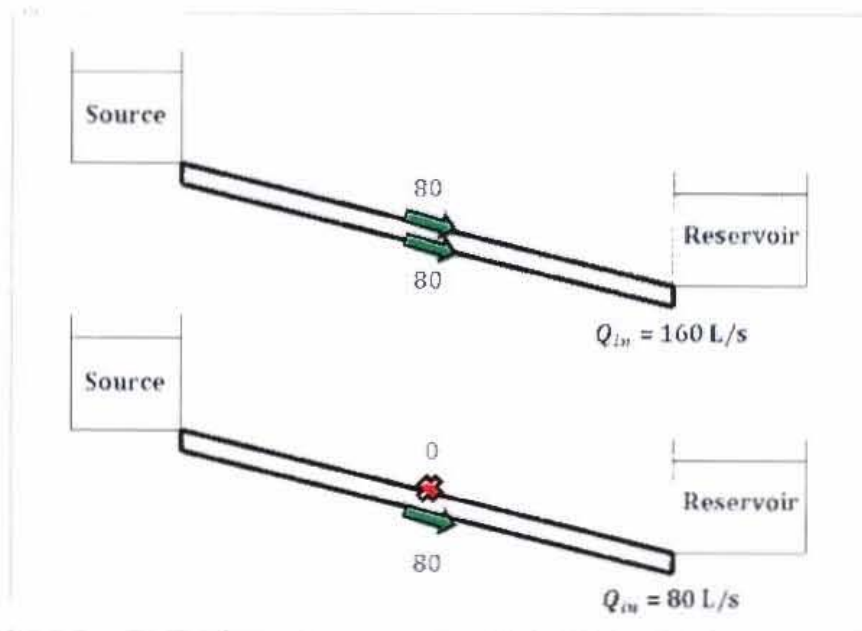


Figure 5.14 Effect of a pipe failure in configuration (2) with $h_f = 30$ m, $C_H = 120$, $L = 1000$ m, $D = 0.207$ m for each pipe

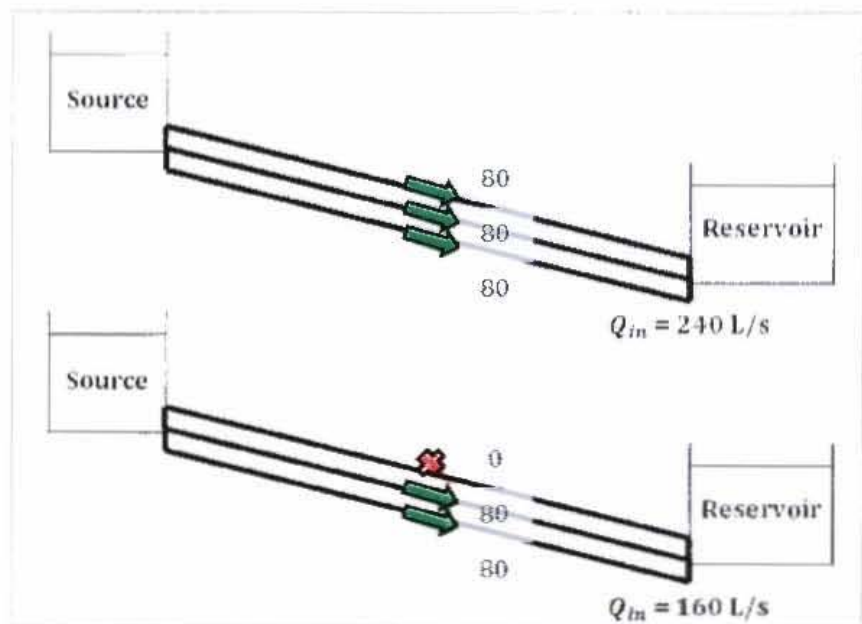


Figure 5.15 Effect of a pipe failure in configuration (3) with $h_f = 30$ m, $C_H = 120$, $L = 1000$ m, $D = 0.207$ m for each pipe

The effect of a pipe failure in a system with parallel pipes has been illustrated. Here, the effect of a pipe failure in systems with two and three pipes in parallel with interconnecting pipes is demonstrated. Any interconnecting pipes added to the system are modelled such that it divides the parallel pipes into equal lengths, with identical pipe characteristics. Once again, it is assumed that in systems with interconnections between parallel pipes, only one pipe failure occurs at any one time.

For configuration (4) and (5), one interconnection pipe divides the parallel pipes into two equal halves. Each pipe in the system has the same characteristics: $C_H = 120$, $L = 500$ m and $D = 0.207$ m. The total pipe length is still $L = 1000$ m and total system head is 30 m. The supply inflow into the reservoir for configuration (4) is the same as for configuration (2). $Q_{in} = 160$ l/s, when all pipes are fully operational. Using Epanet to model a failure in one of the pipes, it was found that the inflow rate into the reservoir is reduced to 102 l/s, which is 63.7% of the total supply inflow (Figure 5.16). When configuration (5) is fully operational, it has a supply inflow of 240 l/s, the same configuration (3). In the event of one pipe failure, the inflow rate into the reservoir is cut down to 189 l/s (Figure 5.17). This means that 78.7% of the supply capacity remains.

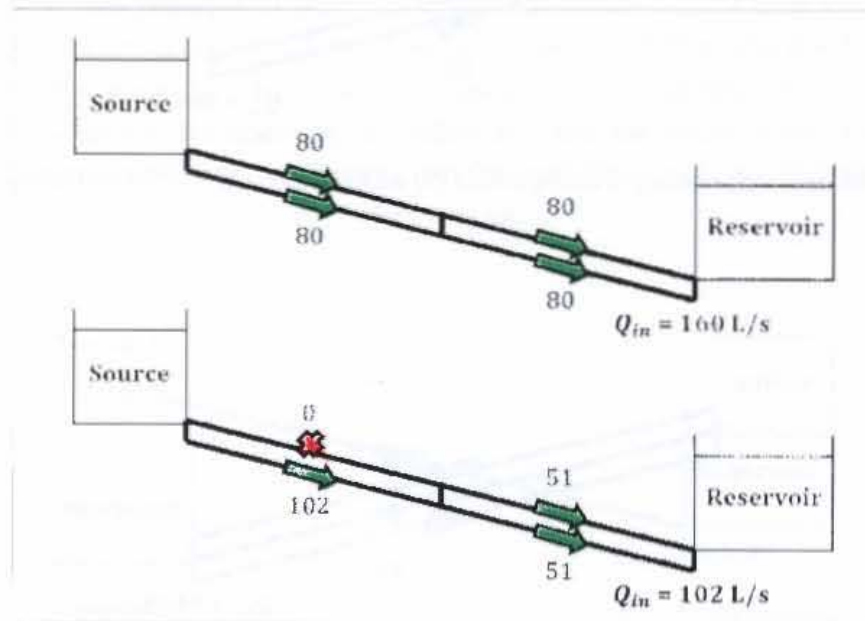


Figure 5.16 Effect of a pipe failure in configuration (4) with $C_H = 120$, $L = 500$ m, $D = 0.207$ m for each pipe

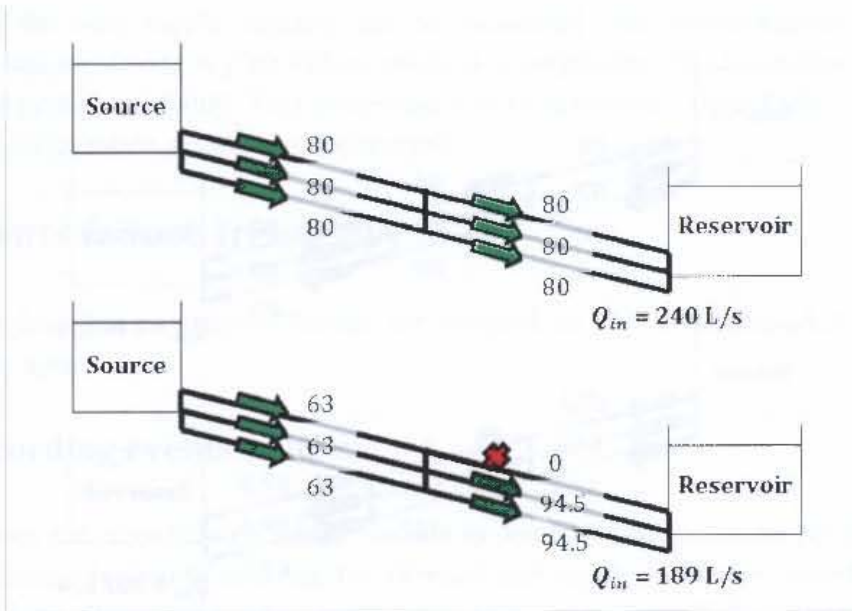


Figure 5.17 Effect of a pipe failure in configuration (5) with $C_H = 120$, $L = 500$ m, $D = 0.207$ m for each pipe

Two short interconnecting pipes divide the parallel pipes in configurations (6) and (7) into three equal sections. The system head is 30 m and the total pipe length remains $L = 1000$ m. The individual pipe sections have the same characteristics: $C_H = 120$, $L = 333$ m and $D = 0.207$ m. Figure 5.18 and Figure 5.19 show what happens when a pipe failure occurs in the system for two and three parallel pipes respectively. Again, with the use of Epanet to model the pipes failure, it is found that the inflow rate is reduced from 160 l/s to 114 l/s (71.3%) for configuration (6), and from 240 l/s to 202 l/s (84.3%) in configuration (7).

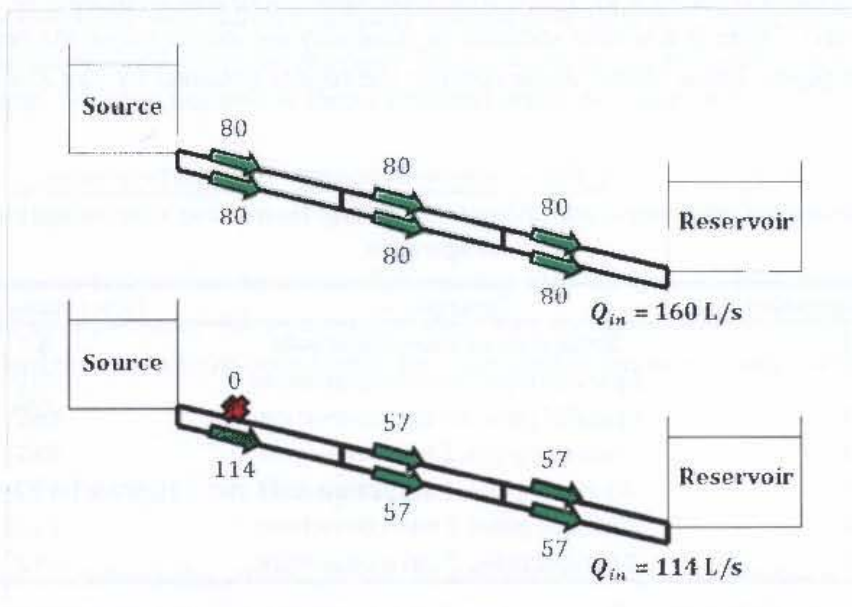


Figure 5.18 Effect of a pipe failure in configuration (6) with $C_H = 120$, $L = 333$ m, $D = 0.207$ m for each pipe

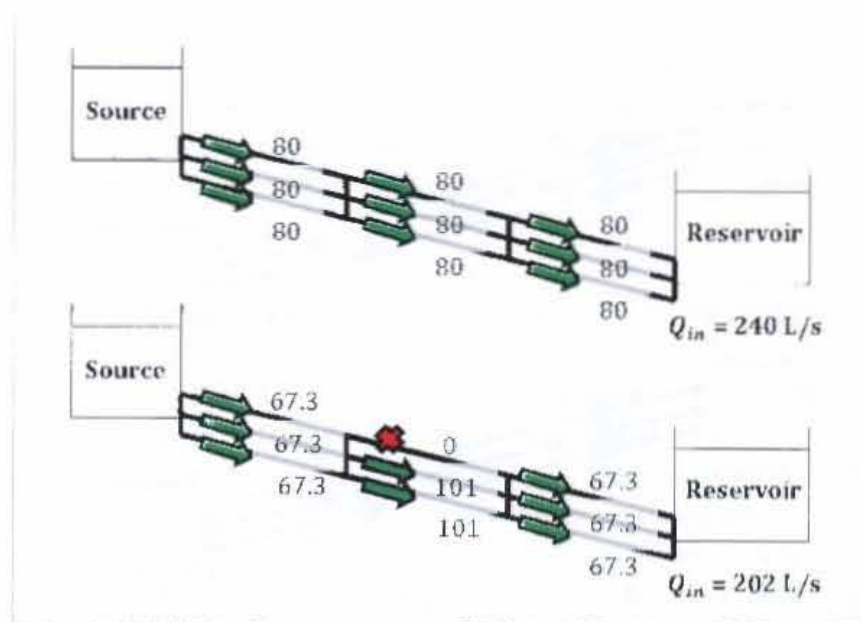


Figure 5.19 Effect of a pipe failure in configuration (7) with $C_H = 120$, $L = 333$ m, $D = 0.207$ m for each pipe

Table 5.6 summarises the above discussion. It indicates how much inflow into the storage reservoir remains after a pipe failure event. It is clear that increasing the number of pipes in parallel increases the reliability of the bulk water supply system, as for two pipes in parallel there is always 50% of the flow remaining, while for three pipes in parallel there is always two thirds of the total supply capacity entering the storage reservoir. If only the systems with two parallel feeder pipes are considered, namely configurations (2), (4) and (6), it is evident that by increasing the number of interconnections between the parallel pipes, the remaining inflow rate increases from 50% to 63.7% (one interconnection) and from 50% to 71.3% (two interconnections). There is a similar increase in reliability for interconnections between three parallel feeder pipes. These observations confirm the results obtained by van Zyl & Haarhoff (1999).

Table 5.6 Summary of remaining supply inflow to the storage reservoir as a percentage for each pipe configuration

Pipe Configuration	Description	Flow remaining (%)
1	Single pipe, no interconnections	0.0
2	2 parallel pipes, no interconnections	50.0
3	3 parallel pipes, no interconnections	66.7
4	2 parallel pipes, 1 interconnection	63.7
5	3 parallel pipes, 1 interconnection	78.7
6	2 parallel pipes, 2 interconnections	71.3
7	3 parallel pipes, 2 interconnections	84.3

In the events run, a pipe configuration for the system is selected as an initial input value. The total system head, total length of pipe, and Hazen-Williams coefficients for the feeder pipes are also input variables, as well as the pipe diameter (the same for all the pipes in the system),

from which the total supply capacity can be calculated. The percentage of total supply capacity remaining during a pipe failure event is incorporated into the events run used to calculate the system reliability. This percentage will be determined from Table 5.6, according to the pipe configuration selected by the designer.

5.6 Events failure frequency

This section describes how events failures are recorded, as well as the impact of the different events on the system.

5.6.1 Recording events failures

Events failures are recorded in a similar fashion to demands failures in the pre-run. For each event simulation, the hourly demand, fire demand and supply inflow are calculated in each time step. These values are used to update the reservoir volume. If the reservoir is empty, then a reservoir failure has occurred.

When a reservoir failure takes place, the number of failures, the time-to-failure (TTF), mean time-to-failure (MTTF) and time-to-failure standard deviation (STTF) are recorded as it was done before in the pre-run. The way in which the time-to-failure is computed for the events run is different to that of the pre-run. To calculate the time-to-failure for the current reservoir failure, two values must be recorded: (1) the time between the hours of the start of the previous and current failure, which is the difference between the hour in which the current failure occurs (*hour*) and the hour in which the previous failure occurred (*prevTimeFail*); and (2) the time omitted or 'jumped over' between the previous and current failure, i.e. the difference between the total number of hours jumped over thus far (*jumpCount*) and the total number of hours omitted when the previous failure occurred (*prevJumpCount*). The time-to-failure (in years between failures) is then calculated using the equation:

$$TTF = \frac{(hour - prevTimeFail) - (jumpCount - prevJumpCount)}{24 \times 365} \quad (5.9)$$

The mean time-to-failure for the events failures can also be used to find the events failure frequency for the system, by taking its reciprocal. Once the mean time-to-failure and time-to-failure standard deviation have been found, they are used in the terminating conditions for the main events run.

5.6.2 Effect of events on the system

The events run simulates the critical periods in which system failures are more likely to exist because of the occurrence of fire and pipe failure events. Changing the rate of events will obviously affect the number of system failures that occur. The effect of doing so will be examined later in the sensitivity analysis. It is important to note, however, that the effect of

an individual fire event on a bulk supply system is different to that of a pipe failure event, in terms of severity. The severity of impact of each type of event on the system is discussed here.

It is clear that the severity of a fire event depends on the duration and fire flow, as these factors determine the fire demand. If one time step of the event duration is observed, equivalent to one hour in the stochastic simulation, then the fire flow will be the only factor deciding the severity of the event. In each hour of the simulation, the change in volume for the reservoir is computed by:

$$\Delta Volume = Q_{in} - (D_h + D_{fire}) \quad (5.10)$$

where $\Delta Volume$ is the change in volume, Q_{in} is the supply inflow rate, D_h is the hourly consumer demand and D_{fire} is the fire demand.

If a high fire demand coincides with a high consumer demand, then it is likely that the total demand will exceed the supply into the reservoir, which will result in a net negative change in volume. On the other hand, if the fire demand is low, and the total demand is less than the supply, then there will be a net positive change in volume. Thus the severity of a fire event will depend predominantly on the fire flow generated from the lognormal distribution.

A pipe failure event will have a significant impact on the behaviour of the storage reservoir. Referring to the above equation for change in volume, a pipe failure will reduce the supply inflow. For a supply pipe system consisting of a single pipeline, there is no supply into the reservoir when a pipe fails. The change in volume will undoubtedly be negative, and it will increase the chance of a reservoir failure occurring. It is clear that in the case of a single supply pipe, a pipe failure event will have a greater impact than a fire event. However, if there are two pipes in parallel, then when a pipe fails, half of the supply will remain. The severity of the pipe failure compared to a fire event in this case will vary depending on what change in volume it causes. A pipe failure in a system with two parallel pipes will be less severe than a system with single pipeline only. From this discussion, it can be seen that the pipe configuration will determine the severity of the impact of a pipe failure event on the storage reservoir.

5.7 Terminating conditions for the events run

Since the events run involves stochastic analysis using Monte Carlo simulation, it is also necessary to establish an appropriate sample size and thus suitable terminating conditions. The stochastic analysis in the events run is the same as that in the pre-run except that fire and pipe failure events are considered and included in the analysis. Reservoir failures happen for the same reason in both cases, and the failures are recorded in a similar fashion in both runs. Therefore, it was logical to use the same terminating conditions as the pre-run for the events run.

As a reminder, the terminating conditions, which are also applied in the events run are:

1. $\frac{1.96S}{\sqrt{n}}/\bar{y} \leq 5\%$;
2. $\bar{y} + \frac{1.96S}{\sqrt{n}} > \max TTF$;
3. $t > \text{user inputted duration}$; or

where n is the number of reservoir failures, \bar{y} is the mean time-to-failure, S is the time-to-failure standard deviation, and t is the time.

These conditions ensure that the events run will be concluded when one of the following occurs:

1. The events failure frequency is accurate to within 5% of its actual value with a 95% confidence level;
2. The system has a greater reliability than desired by the designer; or
3. The system is analysed for the period of time required by the designer.

In the events run, the number of failures, time-to-failure, mean time-to-failure and time-to-failure standard deviation are updated each time there is a reservoir failure. These are used to calculate the events failure frequency, in addition to determining when the events run should be terminated.

5.8 Calculating total failure frequency for the system

After the events failure frequency has been calculated, it must be combined with the demands failure frequency to obtain the total failure frequency of the system. The total failure frequency can be used to describe the bulk system's reliability.

5.8.1 Use of mean time-to-failure to determine failure frequency

Failure frequency can also be referred to as the annual average number of failures or failures per annum. In the pre-run and events run, the demands and events failure frequencies are not directly recorded, and mean time-to-failure is recorded in its place in the simulation. Time-to-failure analysis is used in reliability analysis when the performance of a system is analysed over time. Mean time-to-failure is a commonly used reliability measure of system performance (Tung *et al.*, 2006). Mean time-to-failure is used in this study to determine the failure frequency of the storage reservoir and thus the reliability of the bulk system.

The time-to-failure is defined as the length of time during which a component or system under consideration remains operational (Tung *et al.*, 2006). In this case, it is the length of time during which the storage reservoir is not empty.

The mean time-to-failure (MTTF) would then be the expected value of time-to-failure (Mays & Tung, 1992), expressed mathematically as

$$MTTF = \int_0^{\infty} t f_T(t) dt \quad (5.11)$$

which has a unit of time (e.g. hours, minutes, etc.).

The time-to-failure between each failure is recorded (in years) in order to calculate the mean and standard deviation of the time-to-failure, which can be used to determine convergence of the failure results. The demands failure frequency and the events failure frequency can then be calculated by taking the reciprocal of the demands mean time-to-failure and events mean time-to-failure respectively.

5.8.2 Calculating the total failure frequency

Once both the demands and events failure frequencies have been calculated, they must be combined to obtain the total failure frequency of the bulk supply system. Once again, note that the failure frequency represents the average number of failures per annum. During the events run, the fraction of the total simulation time that consisted of consumer demand only and was not simulated fully was recorded, i.e. the time 'jumped over'. Using the fractions for the demands-only periods and the fully simulated critical periods, as well as the demands mean time-to-failure and the events mean time-to-failure; the total failure frequency can be calculated as follows:

$$Total\ MTTF = \frac{Demands\ fraction}{Demands\ MTTF} + \frac{Events\ fraction}{Events\ MTTF} \quad (5.12)$$

In the above expression, *Demands fraction* is the simulation time consisting of consumer demands as a proportion of the total simulation period; and *Events fraction* is the proportion of the total simulation time during which fire and pipe failure events are simulated.

6 Optimisation

6.1 Introduction

The goal of this project is to use multi-objective optimisation to find Pareto-optimal solutions that can be used in the design of bulk water supply systems. These solutions will be presented in the form of trade-off curves of cost against failure frequency (or lack of reliability), where the cost is determined using cost models developed for the bulk system, and the failure frequency is assessed through a stochastic analysis model. It is assumed that the supply is only limited by the capacity of the pipes, and source failures are not considered in addition to pipe failures. Consumer demand, fire and pipe failure events are outside the control of the designer. What the designer can control when designing a bulk system is the pipe configuration and pipe diameter for the supply pipe system, as well as the size of the storage reservoir.

A nondomination-based genetic algorithm, called the *Nondominated Sorting Genetic Algorithm II* (NSGA-II), was developed by Deb *et al.* (2002) for use in multi-objective optimisation. Figure 6.1 shows a flowchart of the optimisation process using NSGA-II. In NSGA-II, the initial population is generated, followed by an evaluation of the objective functions using the simulation models. Until the stopping criteria have been met (e.g. the number of generations are complete), the algorithm continues to loop through the process of generating a global (or parent) population, ranking the population, performing comparison and selection, and finally generating a child population. The Pareto-optimal solutions consist of the child population produced in the final generation of the algorithm.

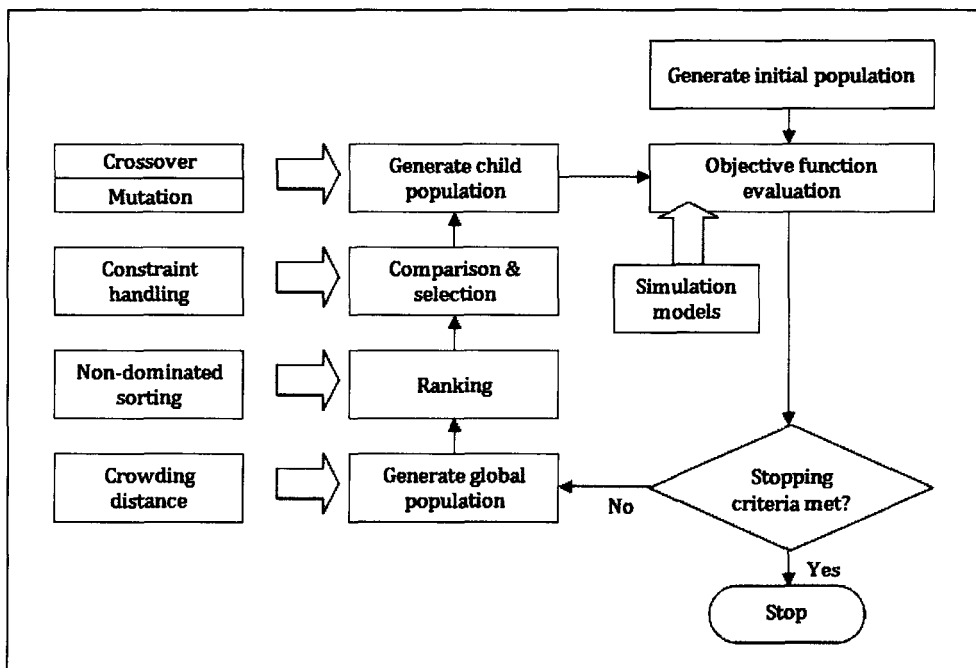


Figure 6.1 Optimisation process using multi-objective genetic algorithm NSGA-II (Wu *et al.*, 2010)

Seshadri (2009) programmed the NSGA-II code in Matlab to work for real-coded problems and continuous variables. Seshadri's code was adapted to suit the objectives in this optimisation problem. Most of the original code by Seshadri remained unchanged as the decision variables also consisted of real-coded numbers. Changes were made in the functions for the initialisation of the population, the evaluation of the objective function, and in the genetic operator function. The functions used to initialise the population and evaluate the objective functions obviously required modification for each specific problem. In contrast, the function containing the genetic operators had to be adapted, not just because it is a different problem, but because two of the decision variables consist of discrete values.

In this chapter, each step of the NSGA-II is examined to see how each of the functions fulfil their purpose, and how they work together to perform the optimisation. Where Seshadri's Matlab code was modified, the functions are described in more detail to provide a greater understanding of why and how the changes were made.

6.2 Defining the problem range

Before the population is initialised, the program requires the number of objective functions, the number of decision variables and the range of the decision variables to be defined. In this optimisation problem, there are two objectives: (1) cost minimisation and (2) minimisation of failure frequency (or maximisation of reliability). The cost of a system is found using a cost model, while its reliability is determined using the stochastic analysis model. For this system, there are three decision variables, namely (1) pipe configuration, (2) pipe diameter, and (3) reservoir size.

The range of each decision variable must be defined to prevent its limits from being exceeded. This is because some system components, such as pipe diameters, are limited by what is commercially available; or only certain options are offered to the user, for example, set pipe configuration patterns.

6.2.1 Range for the pipe configuration

There are seven pipe configuration options for the system. These were described in Section 5.5. Each of the pipe configurations is given a number from one to seven. During the optimisation process, the selection, crossover and mutation operators are modified to deal with this discrete variable. These configurations were included in the optimisation because it was felt that the selection provided an adequate variety of options as well as a fair representation of typical configurations. If the designer wishes to include additional pipe configuration options, the decision variable can easily be modified to do so.

6.2.2 Range for the pipe diameter

The range for decision variable two, the pipe diameter, depends on what commercially produced pipe sizes are available. Before the optimisation, the user is required to input a set of commercially available pipe sizes. Depending on the total system head, the length of the pipe system and the Hazen-Williams coefficient, the supply ratio for each available pipe diameter is obtained. The pipe diameters which give impractical supply ratios (eliminating conditions in Section 4.2) are disregarded. The remaining pipe diameters are ranked from smallest to largest, and they are each given an index number starting from one. The pipe index numbers are integers and therefore decision variable two also consists of discrete values.

6.2.3 Range for the storage reservoir size

The size of the storage reservoir is the only decision variable whose range is directly controlled by the designer. The designer is required to input a minimum and maximum reservoir size in terms of hours of storage capacity, which is then converted to volume in cubic metres. These upper and lower reservoir size limits are used in the analysis of demands failures in the pre-run. This decision variable is also the only one whose value is selected from a continuous range.

6.3 Population Initialisation

The initial population is randomly generated. This population is initialised based on the problem range. Each solution in the population is structured such that it has seven elements (or substrings) as shown in Figure 6.2. The first three elements contain the decision variables: pipe configuration, diameter, and reservoir size. Elements four and five contain the objective function values: cost and reliability. These five elements are initialised in the function `initialize_population`. The last two elements in the solution are the rank and crowding distance which are obtained through the nondomination sorting.

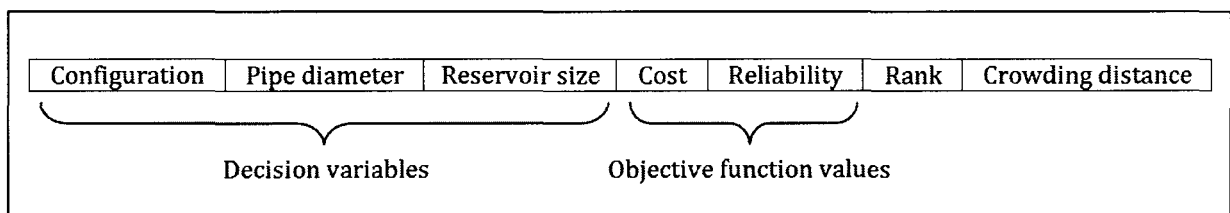


Figure 6.2 Structure of each solution

For each solution in the initial population, the pipe configuration is initialised by randomly selecting an integer from the range one to seven. This integer corresponds to one of the pipe configuration options set previously. Following this, the pipe diameter for the solution is initialised by randomly selecting an integer from the range for the discrete pipe index number,

which starts from one and ends at the total number of available pipe diameters sizes. This integer is stored as the second decision variable. Finally, a value is randomly selected from the continuous range between the minimum and maximum reservoir sizes. This value is the reservoir size stored as the third decision variable.

As part of initialising the population, the values of the objective functions for each solution must be found. These values are computed using the `evaluate_objective` function.

6.4 Evaluating the objective functions

The `evaluate_objective` function takes in a given input solution which consists of an array of decision variables and outputs the values for the objective functions. In this function, the objective functions are defined and evaluated using simulation models. The simulation models used are a cost model for the cost analysis, and the stochastic model for the reliability analysis.

It should be noted that the NSGA-II optimisation algorithm used in this study is structured in such a way that it always minimises the objective function. If it is required that the objective function is maximised, then it must be multiplied by negative one. For the purposes of optimisation in this study, both the system cost and the failure frequency need to be minimised. As previously mentioned, reliability (or lack thereof) of a system is defined by its failure frequency. Therefore minimisation of the system's failure frequency represents maximisation of its reliability.

6.4.1 Cost evaluation

The cost of a system is more often than not the main design aspect that is focussed on after ensuring that the system meets the performance constraints. Therefore, the goal of single-objective optimisation is usually minimisation of cost. Over the years, designers have realised the value of multi-objective optimisation to include other objectives; however, cost remains a priority.

It is important that costs for a water distribution system are modelled accurately. Unfortunately it is sometimes not possible to get exact values. However, when comparing costs of systems for optimisation purposes, what is important is that the accuracy of the relative costs between solutions.

Early studies on water distribution system optimisation generally used capital costs in the cost analyses, for example in benchmark optimisation problems such as the Hanoi (Fujiwara & Khang, 1990) and New York Tunnels (Schaake & Lai, 1969) systems. The benefits of using capital costs in cost modelling are that they are easy to use and it is fairly simple to collect the cost data. Other studies have developed cost models based on *life cycle costing*

(LCC), because it is useful to know the cost of a system over its lifetime. LCC includes fixed capital costs in addition to recurring costs such as the operation and maintenance costs.

The cost model used in this optimisation process consists of capital costs only. It was felt that LCC was unnecessary because of the assumption that there are no pumping stations in the system. The system components essentially consist of the supply pipe system and the storage reservoir, which makes developing a cost model for the system straightforward. In the text by Swamee & Sharma (2008), cost functions for pipes and storage reservoirs are described based on an Australian dataset from Samra & Essery (2003). These cost functions were also used in this project.

Cost function for pipelines

A completed pipeline can be modelled to have a cost C_m based on the relationship between the pipe length L and pipe diameter D (Swamee & Sharma, 2008):

$$C_m = k_m L D^m \tag{6.1}$$

where k_m is a coefficient, and m is an exponent. The pipe cost parameters k_m and m are dependent on the pipe material, the monetary unit of the cost and the economy (Swamee & Sharma, 2008). Assuming that the pipes in this study are made of cast iron and using the cost data set from Samra & Essery (2003), the pipe cost parameters are $k_m = 480$ and $m = 0.935$. The cost function for the supply pipes used in the optimisation model is:

$$C_m = 480 L D^{0.935} \tag{6.2}$$

Cost function for the storage reservoir

Swamee & Sharma (2008) also developed a cost function for a surface concrete reservoir from the Australian data (Samra & Essery, 2003) in Table 6.1.

Table 6.1 Service reservoir cost

Reservoir Capacity (m ³)	Cost (A\$)
100	28,000
200	55,000
400	125,000
500	160,000
1000	300,000
2000	435,000
4000	630,000
5000	750,000
8000	1,000,000
10,000	1,150,000
15,000	1,500,000
20,000	1,800,000

A cost function for the surface reservoir was developed using analytical methods. This cost function is used as part of the cost model in the optimisation:

$$C_R = \frac{290V_R}{\left[1 + \left(\frac{V_R}{1100}\right)^{5.6}\right]^{0.075}} \quad (6.3)$$

where C_R is the service reservoir cost (A\$), and V_R is the volume of the reservoir (m^3).

Cost function for chambers

Various pipe configurations were considered, which included interconnections between pipes in parallel. Chambers are required to house the necessary valves and pipe work. Swamee & Sharma (2008) did not consider the use of parallel pipes or interconnections in their study, and thus did not develop a cost function for the costs of the chambers. As a result, Vlok (2010) developed an equation to represent the chamber costs, which were de-escalated so that it could be used in conjunction with the unit cost models for the pipelines and storage reservoir that are based on 2003 Australian monetary terms.

Unit cost models were developed for four types of chambers (Vlok, 2010):

- 2 x split chamber – Chamber with pipe work and valves to aid the transition of one pipe to two pipes in parallel;
- 3 x split chamber – Chamber with pipe work and valves to aid the transition of one pipe to three pipes in parallel;
- 2 pipe interconnection – Chamber with pipe work and valves to aid the interconnection of two parallel pipes;
- 3 pipe interconnection – Chamber with pipe work and valves to aid the interconnection of three parallel pipes.

Transitions from one pipe to two or three pipes in parallel are excluded from the pipe configuration options. Only parallel pipes and parallel pipes with interconnections are considered, and consequently only the unit cost models for the last two chamber types are used. For the interconnection of two pipes, the cost function of the chamber is:

$$C_c = 0.0457D^2 + 16.519D + 2624.1 \quad (6.4)$$

and the cost function of the chamber for the interconnection of three pipes is:

$$C_c = 0.07D^2 + 24.898D + 2790.2 \quad (6.5)$$

where C_c is the cost of the chamber (A\$), and D is the pipe diameter (mm).

Using the cost model in the optimisation process

After the decision variables for each solution have been established, they must be converted to the correct units for input into the cost functions. The required input variables for the total cost function are the pipe configuration, pipe length and diameter, and storage reservoir volume.

The first decision variable, that is, the system's pipe configuration, will clearly have an effect on how expensive the system is. If the system has two parallel pipes then the pipeline cost will be twice as much as a single pipe, assuming that the pipes are laid in separate trenches. Similarly a system with three parallel pipes will be three times as expensive as a single pipe. Besides affecting the pipeline cost, the pipe configuration also increases the total cost with additional chamber costs when there are interconnections.

The pipe diameter of the supply system also affects the pipeline cost. Since decision variable two is in the form of a diameter index number, it must be converted to a pipe diameter in metres, from the range of commercially available pipes sizes. The pipe length, in metres, is the final input for the cost of the pipeline component of the system.

The last input variable for the cost function is the reservoir volume. The reservoir size in hours of storage capacity must be converted to a volume in cubic metres, so that the storage reservoir cost can be calculated, and added to the pipeline cost and the chamber cost.

Once the individual cost components are calculated using the above cost functions, the total cost C_T (A\$) is obtained by adding up the individual components:

$$C_T = C_m + C_R + C_c \quad (6.6)$$

The total cost is then stored as the first objective function value for the solution.

6.4.2 Reliability evaluation

A multi-objective optimisation consists of a minimum of two objectives. A number of objectives other than cost could be evaluated. Possible objectives include *inter alia* reliability, resilience, water quality and security. The focus of this study is to find the optimal design of a bulk water supply system based on reliability, which is the second objective function value that must be found for each solution.

Reliability of the system is expressed in terms of the failure frequency of its storage reservoir, which is evaluated by a simulation model using stochastic analysis. The stochastic analysis model consists of both a demands pre-run and main events run. This compression heuristic has been fully described in Chapter 4: Pre-run and Chapter 5: Events Run.

The total failure frequency for the solution is found using the events run function, whose input variables were described in Section 5.2. The input values that need special mention are the configuration and the current solution's supply ratio. Previously in Section 4.2, the method used to determine the range of practical pipe diameters for the system was described. Included in the description were the supply ratio values for each combination of pipe diameter and number of pipes. Before the events run can be evaluated, the current supply ratio must be determined. This is done by finding the supply ratio that is associated with the number of pipes, and the pipe diameter.

Once the events run is complete, the total failure frequency computed is stored as the second objective function value for the solution.

6.5 Sorting the population

The population is sorted based on the fast nondominated sorting algorithm and the crowding distance assignment algorithm (Deb *et al.*, 2002). The standard algorithms were used in the optimisation without any changes. These algorithms are described below.

6.5.1 Nondomination sorting

The initialised population is sorted based on nondomination (Deb *et al.*, 2002), using the fast sorting algorithm. For each solution p in the main population P , two values are calculated: (1) a set S_p of solutions which are dominated by solution p , and (2) a nondomination count n_p , which is the number of solutions that dominate p . If solution p 's nondomination count n_p is zero, no solutions dominate p , then p belongs in the first nondominated front. This process is carried out for all the solutions in the main population P . The front counter is then initialised to one, $i = 1$.

While the i^{th} front is nonempty, Q is initialised as an empty set that stores the solutions for the $(i + 1)^{th}$ front. For each solution p in the i^{th} front with $n_p = 0$, the domination count of each member q in its set S_p is decreased by one. If the domination count of any member q is reduced to zero, then it is put in the set Q . Members in set Q belong to the $(i + 1)^{th}$ front. The front counter is increased by one. This process continues until all the fronts are identified.

6.5.2 Crowding distance

Aside from desiring convergence to a Pareto-optimal set, there should also be a good spread of solutions. After completion of the nondominated sorting, the crowding distance is assigned. The crowding distance is a measure of the diversity of the solutions in a population. Large average crowding distance will result in better diversity in a population (Seshadri, 2009).

Solutions in the population are selected based firstly on rank, and then crowding distance. Crowding distance between solutions in different fronts cannot be compared as it would be pointless to do so, since crowding distance is assigned according to each front.

The distance of each solution in each front F_i is set to be zero, i.e. $F_i(d_j) = 0$, where j is the j th individual in front F_i . This distance represents an estimate of the perimeter of the cuboid (Figure 6.3) formed by using the nearest neighbours as vertices (referred to as the *crowding distance*) (Deb *et al.*, 2002).

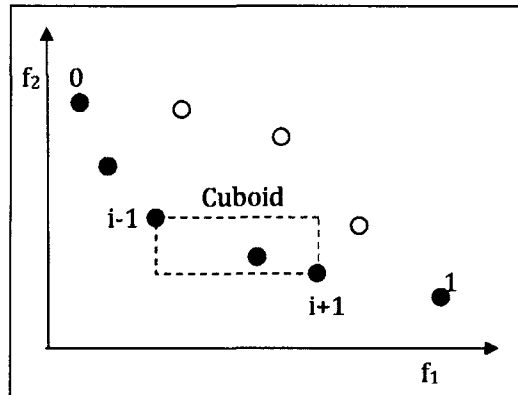


Figure 6.3 Crowding distance calculation - points marked in filled circles are solutions of the same nondominated front (Deb *et al.*, 2002)

The solutions in front F_i are sorted from smallest to largest for each objective function value, and stored as set I . The boundary solutions are then assigned an infinite distance value, i.e. $I(d_1) = \infty$ and $I(d_n) = \infty$. All the other intermediate solutions are assigned a distance value calculated with the following equation:

$$I(d_k) = I(d_k) + \frac{I^{(k+1).m} - I^{(k-1).m}}{f_m^{max} - f_m^{min}} \quad (6.7)$$

where $I(k).m$ is the value of the m^{th} objective function of the k^{th} individual in I , and f_m^{max} and f_m^{min} are maximum & minimum values of m^{th} objective function. The individual distance values corresponding to each objective are summed up to give the overall crowding distance value.

6.6 The evolution process

Each generation undergoes the following process:

- Selection of the parents for reproduction;
- Crossover and mutation are performed on the selected parents;
- Selection is performed from the parents and the children;
- Replacement of the unfit solutions with the fit solutions to maintain a constant population size.

6.6.1 Selection

Once each solution in the population has been assigned a nondomination rank i_{rank} and crowding distance $i_{distance}$, a parent population is selected, from which a child population can be generated. The solutions of the parent population are selected for reproduction using tournament selection based on a *crowded-comparison operator*.

The original NSGA-II used a binary tournament selection, which means that the fitness of two solutions is compared and the better one is selected as the parent. Tournament selection is performed until the pool size, equal to half the population size, is filled.

The NSGA-II optimisation algorithm uses a crowded-comparison operator \prec_n (Deb *et al.*, 2002) to direct the selection process towards a uniformly spread out Pareto-optimal front. It does so by using the nondomination rank i_{rank} and crowding distance $i_{distance}$ to indicate which solution is preferred.

A partial order \prec_n is defined as

$$i \prec_n j$$

if

$$(i_{rank} < j_{rank})$$

or

$$(i_{rank} = j_{rank}) \& (i_{distance} > j_{distance})$$

This means that if the two solutions belong to different fronts, then the solution with lower rank is preferred. Otherwise, if both solutions belong to the same front, then the solution that is less crowded is preferred.

6.6.2 Crossover

NSGA-II follows the scheme whereby each variable (either discrete or continuous) undergoing crossover is changed. For this problem, crossover is performed with 90% probability, the variables are coded directly and a real-coded *simulated binary crossover* (SBX) operator (Deb & Agrawal, 1995) is used. SBX was developed such that it simulates single-point binary crossover and respects *interval* schema processing.

In simulated binary crossover, children solutions $x_i^{(1,t+1)}$ and $x_i^{(2,t+1)}$ are computed from the parent solutions $x_i^{(1,t)}$ and $x_i^{(2,t)}$ using the following polynomial probability distribution:

$$P(\beta) = \begin{cases} 0.5(\eta_c + 1)\beta^{\eta_c}, & \text{if } 0 \leq \beta \leq 1 \\ 0.5(\eta_c + 1)\frac{1}{\beta^{\eta_c+2}}, & \text{if } \beta > 1 \end{cases} \quad (6.8)$$

where β is the spread factor and η_c is the distribution index.

The spread factor β is defined as the ratio of the absolute difference in the children values to that of the parent values (Deb & Goyal, 1996):

$$\beta = \left| \frac{x_i^{(2,t+1)} - x_i^{(1,t+1)}}{x_i^{(2,t)} - x_i^{(1,t)}} \right| \quad (6.9)$$

To create the children solutions using crossover of the parent solutions, a random number u between 0 and 1 is chosen. The spread factor β is found by equating the area under the probability curve equal to u using the expression:

$$\beta = \begin{cases} (2u)^{\frac{1}{\eta_c+1}}, & \text{if } u \leq 0.5 \\ \left(\frac{1}{2(1-u)}\right)^{\frac{1}{\eta_c+1}}, & \text{if } u > 0.5 \end{cases} \quad (6.10)$$

Thereafter, the children solutions are calculated by:

$$x_i^{(1,t+1)} = 0.5 \left[(1 + \beta)x_i^{(1,t)} + (1 - \beta)x_i^{(2,t)} \right] \quad (6.11)$$

$$x_i^{(2,t+1)} = 0.5 \left[(1 - \beta)x_i^{(1,t)} + (1 + \beta)x_i^{(2,t)} \right] \quad (6.12)$$

The procedure described is for the SBX operator applied to continuous variables coded directly, such as the reservoir size in this optimisation problem. For crossing discrete variables which are coded directly, such as the pipe diameter and configuration, the discrete version of SBX is suggested where a discrete probability distribution function is used. Even though that is the recommended method, for the sake of ease of use, the continuous version of SBX was used for the discrete variables in this thesis, and the points obtained were simply rounded off to the nearest permissible value.

Where lower and upper bounds (x^l and x^u) of a variable are specified, Deb & Gulati (2001) describe the calculation of children solutions with:

$$\beta = \begin{cases} (\alpha u)^{\frac{1}{\eta_c+1}}, & \text{if } u \leq \frac{1}{\alpha} \\ \left(\frac{1}{2-\alpha u}\right)^{\frac{1}{\eta_c+1}}, & \text{if } u > \frac{1}{\alpha} \end{cases} \quad (6.13)$$

where $\alpha = 2 - \beta^{-(\eta_c+1)}$ and β is calculated as follows (assuming that $x_i^{(1,t)} < x_i^{(2,t)}$):

$$\beta = 1 + \frac{2}{x_i^{(2,t+1)} - x_i^{(1,t+1)}} \min \left[\left(x_i^{(1,t)} - x^l \right), \left(x^u - x_i^{(2,t)} \right) \right] \quad (6.14)$$

An alternative method of enforcing the bounds of a variable is to round off any children solution outside the prescribed range to the appropriate boundary value, as was done in this project.

Distribution index for crossover, η_c

The distribution index η_c can be any non-negative real number. Large η_c values cause children points closer to the parents to be selected. Small η_c values mean that the children solutions are more likely to be distant from the parents. The simulation results using the SBX operator closely match those of single-point binary crossover if moderate values ($\eta_c = 2$ to 5) are selected (Deb & Kumar, 1995). It is recommended that this parameter is initially small when a broad search is desired, and later a larger η_c is selected as the solutions converge so that a narrower search is obtained (Deb & Kumar, 1995; Deb & Goyal, 1997).

For the purposes of this optimisation problem, the distribution index was kept constant. A small value of η_c equal to 2 was chosen. Since two of the decision variables are discrete, the selection of a large η_c value would result in a high probability of creating children that are equal to the parent solutions. However, if a small η_c is selected, the children solutions at least have a greater chance of changing to a different value.

To confirm that an adequate distribution index was chosen in terms of providing enough spread as well as ensuring convergence, various η_c values were tested and compared. The optimisation procedure was carried out for distribution indices η_c equal to 2, 10 and 20. The Pareto-optimal solution set is shown for each of these η_c values in the graph below (Figure 6.4). A close-up of the range of solutions with cost between A\$5 million and A\$6 million, and failure frequency between zero and twenty-five failures per annum is provided in Figure 6.5.

It was found that the solutions for $\eta_c = 10$ and $\eta_c = 20$ are densely located in the range of costs between A\$500,000 to A\$550,000. $\eta_c = 2$ appears to provide a greater spread of solutions, as it is able to find the range of solutions with costs between A\$425,000 and A\$500,000. After fifty generations, all the distribution indices managed to find the same front for most solutions. However, for the most unreliable, but most cost-effective solution, $\eta_c = 20$ is unable to find the cheaper and more reliable solution that both $\eta_c = 2$ and $\eta_c = 10$ were able to. In addition, both $\eta_c = 2$ and $\eta_c = 10$ were able to find the cheapest reliable solution at A\$553,867, with failure frequency less than the suggested design criterion of one failure in ten years (see Figure 6.5). This optimal solution is superior to the equivalent solution produced by $\eta_c = 20$, which is 1.55% more costly at A\$564,233. Thus, a distribution index of η_c equal to 2 was chosen for this optimisation procedure, since it was able to produce the necessary convergence and spread of solutions. Although the distribution index η_c was fixed at a value of 2 for all optimisation runs, the user has the option of experimenting with and changing it.

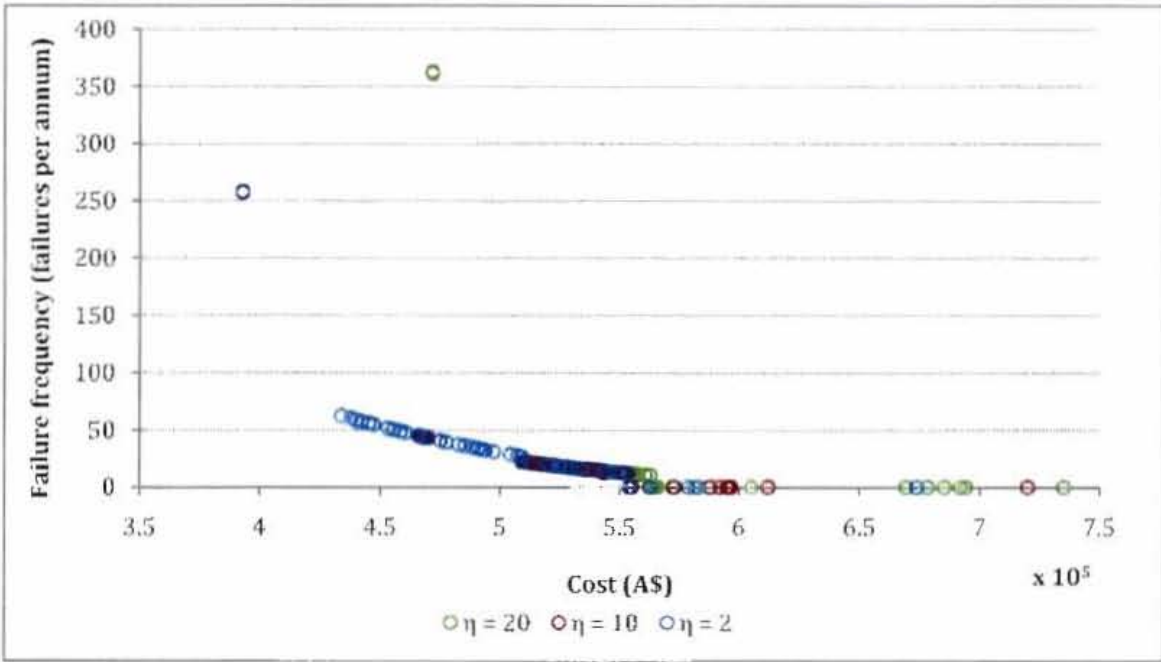


Figure 6.4 Comparison of distribution indices $\eta_i = 2; 10; 20$

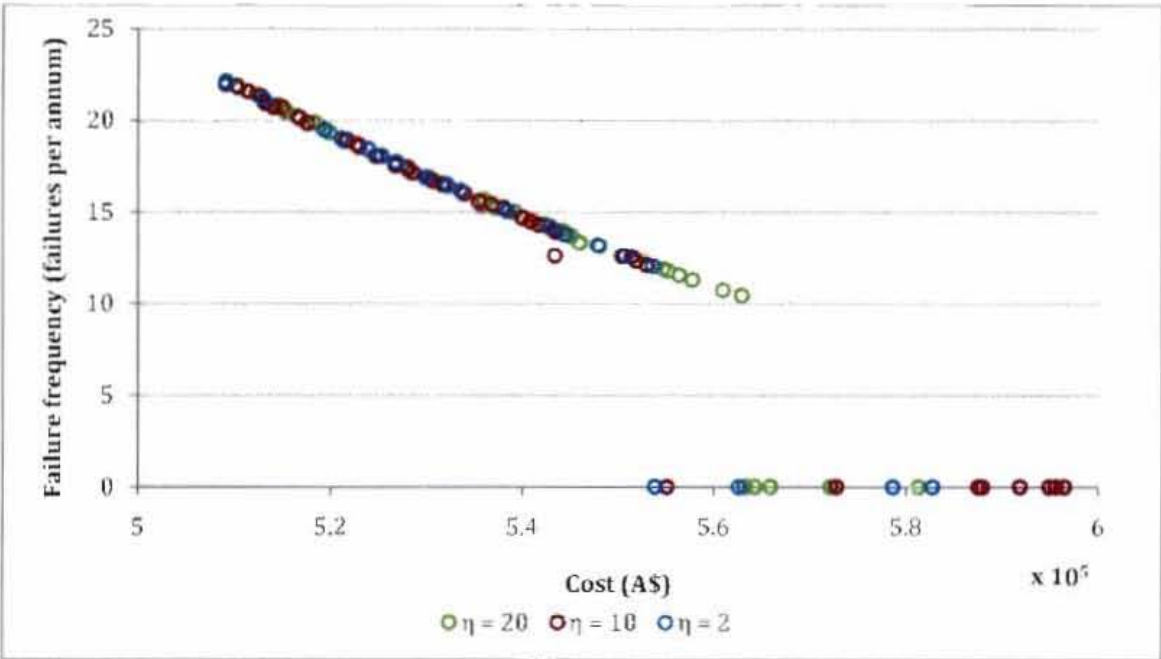


Figure 6.5 Comparison of distribution indices $\eta_i = 2; 10; 20$ (close up)

6.6.3 Mutation

Mutation is applied to maintain diversity of the solutions. NSGA-II uses the *polynomial mutation* operator (Agrawal, 1995), which is applied to each variable in the solution that is selected to undergo mutation. In this optimisation, mutation occurs with a 10% probability.

For a continuous variable, polynomial mutation changes its current value to a nearby value. Like the crossover operator, this is done with a polynomial probability distribution. The current value is the mean for the probability distribution, whereas the variance is a function of the distribution index η_m .

The mutated (or child) value is calculated with the probability distribution

$$P(\delta) = 0.5(n + 1)(1 - |\delta|)^{\eta_m} \quad (6.15)$$

where δ is the perturbation factor defined by

$$\delta = \frac{c-p}{\Delta_{max}} \quad (6.16)$$

In the above expression, Δ_{max} represents the maximum allowable perturbation in the parent (or current) value p , and c is the child (or mutated) value.

To create the child value, a random number u between 0 and 1 is generated, which is used to determine the perturbation factor δ :

$$\delta = \begin{cases} (2u)^{\frac{1}{\eta_{m+1}}} - 1, & \text{if } u < 0.5 \\ 1 - [2(1 - u)]^{\frac{1}{\eta_{m+1}}}, & \text{if } u \geq 0.5 \end{cases} \quad (6.17)$$

The child value is then created with the expression:

$$c = p + \delta\Delta_{max} \quad (6.18)$$

This procedure is followed when applying the polynomial mutation operator to the reservoir size.

For discrete variables directly coded, the discrete version of the above probability distribution $P(\delta)$ is usually used. However, for the pipe configuration and diameter variables in this project, the continuous version was used, and obtained mutated values were just rounded off to the nearest permissible value.

Like the SBX operator, lower and upper bounds (x^l and x^u) may be specified for variables. The above equation would then be altered such that:

$$\delta = \begin{cases} [2u + (1 - 2u)(1 - \bar{\delta})^{\eta_{m+1}}]^{\frac{1}{\eta_{m+1}}} - 1, & \text{if } u < 0.5 \\ 1 - [2(1 - u) + 2(u - 0.5)(1 - \bar{\delta})^{\eta_{m+1}}]^{\frac{1}{\eta_{m+1}}}, & \text{if } u \geq 0.5 \end{cases} \quad (6.19)$$

where

$$\bar{\delta} = \frac{\min[(x-x^l), (x^u-x)]}{(x^u-x^l)} \quad (6.20)$$

and Δ_{\max} would be set to

$$\Delta_{\max} = x^u - x^l \quad (6.21)$$

Once again, these equations are not used in this study. Instead, if the mutated value falls outside of the range $[x^l, x^u]$, then it is rounded off to the corresponding boundary value.

Distribution index for crossover, η_m

The distribution index η_m for the mutation operator has an equivalent definition to what is outlined for the crossover distribution index. It defines the shape of the mutation operator's probability distribution. A small η_m value means that there is a high likelihood that the child value is far away from the original parent value. While a large value means that it is probable that the child value will be closer to the parent value.

In this optimisation, a small mutation distribution index was selected ($\eta_m = 2$), for similar reasons as why a small crossover distribution index was used.

6.6.4 Recombination and selection

After crossover and mutation have been performed, the parent and child population of the current generation are combined. Including both the previous and current generation's best individuals in the population ensures elitism. The population is once again sorted based on nondomination. Thereafter the best solutions are selected based on rank and crowding distance until the population size is reached. The evolution process repeats to generate the subsequent generations until the maximum number of generations is reached.

6.7 Setting the population size and number of generations

6.7.1 Population sizing

Part of setting up the multi-objective genetic algorithm is selecting the population size. Careful consideration must be given to this selection because a large population will produce better quality of solutions but the downside of it is that it will be slower to converge and waste computational resources. The opposite is true for a small population size, which will converge faster, but may become trapped in localised regions of the search space, as stated by Lobo & Lima (cited in Raad, 2010).

Traditionally no rule is followed to select the population size for multi-objective evolutionary algorithms (MOEAs), and it is usually informally selected. Although there are now studies that have explored adaptive population sizing, a review of which was done by Lobo & Lima (2005), most current MOEAs use a fixed population size. Lobo & Goldberg's study (as cited in Raad, 2010) recognised that population sizing in the region of 50-100 solutions are most commonly used. In the original study on NSGA-II, Deb *et al.* (2002) tested the algorithm on a number of problems using a population size of 100. For the purposes of this optimisation problem, a population size of 50 solutions is selected. The reason for this is that the evaluation of the reliability objective function takes a substantial amount of time for each solution. A smaller population size will ensure faster convergence towards the optimal solution.

Since the population size affects the performance of a particular algorithm, the sensitivity analysis should be conducted with respect to the population size used (Raad, 2010). Therefore, a static population size of 50 solutions is used for all the simulations in this study.

6.7.2 Number of generations

Genetic algorithms also require stopping criteria to be defined. The number of generations is often used as a stopping condition, but other criteria such as the fitness of the best solution or population convergence can be used.

The final result is often improved by increasing the number of generations. In the simulation results from Deb *et al.* (2002), NSGA-II was run for a maximum of 500 generations, which was recognised as a rather large number. They found, however, that a reasonably good spread of solutions is generated by 200 generations. Balakrishnan & Jacob (1996) argue, however, that using the number of generations as the sole stopping criterion is insufficient. If few generations are specified then the algorithm may be not given enough opportunity to generate the optimal solutions. In contrast, if the algorithm converges too quickly to a set of solutions, then additional generations will be executed with no improvement in the solutions, thereby wasting valuable computational resources.

Another possible stopping criterion for the genetic algorithm is convergence of the population. Seeing as the true Pareto-set is not available, convergence is defined with respect to a static population, which may take an extremely long time (Raad, 2010). In his dissertation, Raad (2010) chose to specify a threshold of 0.05% change in hypervolume per generation, which must have occurred for 200 consecutive generations for the algorithm to have 'converged'. If this convergence criterion was used in this study, it would most likely take an excessive length of time to converge considering the computational effort required when performing the stochastic analysis. Also, the number of different combinations of pipe configuration and pipe diameter for the bulk system is fairly small in comparison with a water distribution network. Once these two decision variables converge, the genetic algorithm will essentially just be testing various reservoir sizes; therefore it is expected that convergence to the Pareto-front

will happen fairly quickly anyway. It was decided that explicitly specifying a convergence criterion would not be done. The number of generations would be used as the only stopping criterion, against the argument by Balakrishnan & Jacob (1996). In addition, the same number of generations is for each optimisation in the sensitivity analysis.

To investigate how many generations would be sufficient, the optimisation is performed for a single system and the solutions after every ten generations are plotted on the same graph. Up to fifty generations were performed by the algorithm. As can be seen in Figure 6.6, the solutions converge to the same front even after twenty generations. Using these results as a guide, it was decided to set the number of generations at a somewhat conservative value of fifty generations.

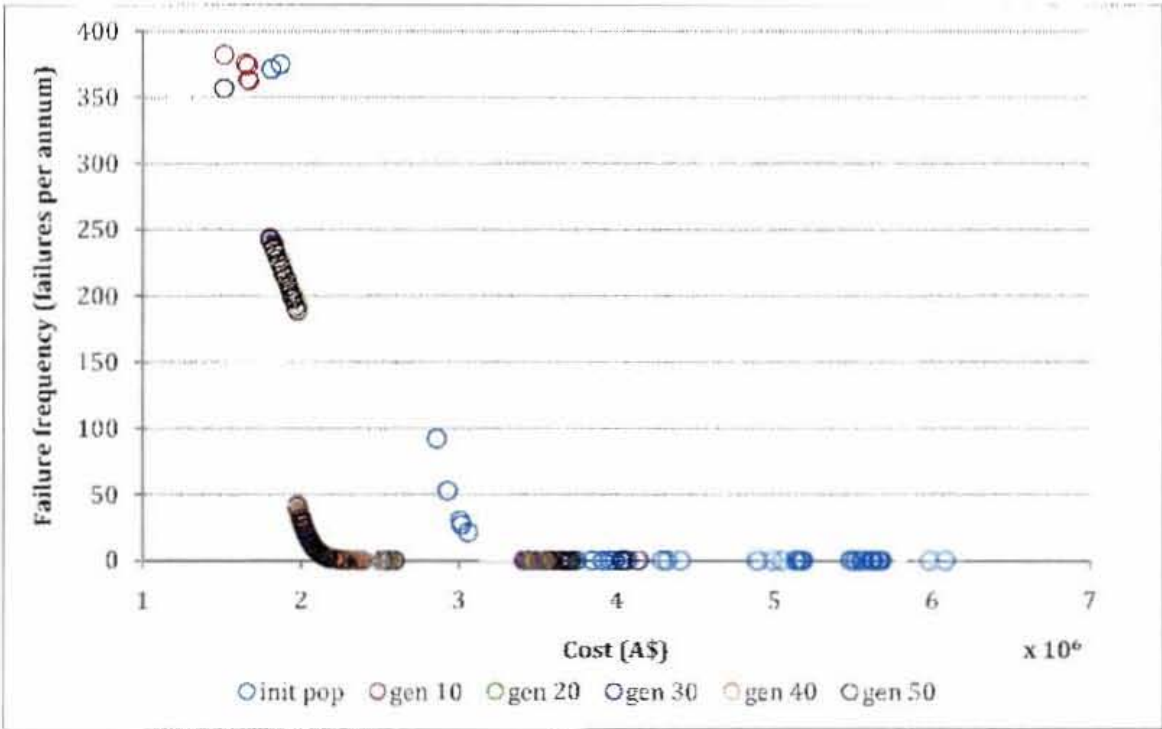


Figure 6.6 Solution set after every ten generations

7 Results and Discussion

In this chapter, results and simulation times for the compression heuristic and full stochastic method are compared and discussed. Following the confirmation of the practical usability of the compression heuristic, the reliability-based optimisation model is applied to an example bulk supply system, with the same system parameters as those used by van Zyl *et al.* (2008). A sensitivity analysis is then performed using the example system as the *typical* system.

7.1 Comparison of stochastic simulation results

This section investigates the accuracy of the compression heuristic failure results. Both the pre-run and the events run are run until convergence. Their individual failure results are combined to give the total system failure frequency for the compression heuristic. Likewise, the full stochastic analysis method (full Monte Carlo method) is used to determine the reliability of the same bulk system in terms of failure frequency. The same convergence criterion is applied. The failure frequencies obtained by these two methods are then compared.

The compression heuristic and the full Monte Carlo method were each run for ten simulations using different random seeds. The average and standard deviation of the failure frequency for both methods were calculated and evaluated. An important determinant of how successful the compression heuristic was in providing accurate and reproducible results was whether the failure frequency for each simulation fell within the 5% bounds of the average for the full Monte Carlo stochastic analysis.

It was recognised that investigating just one system would be inadequate to provide a fair representation of how accurately the compression heuristic performs. Four systems were thus analysed to investigate the range of possible bulk water supply systems that could be used. These systems were varied in terms of small or large supply ratio, and small or large reservoir capacity. The following systems were analysed:

1. Small supply ratio (1.2) and small reservoir capacity (3 h);
2. Small supply ratio (1.2) and large reservoir capacity (12 h);
3. Large supply ratio (1.5) and small reservoir capacity (3 h);
4. Large supply ratio (1.5) and large reservoir capacity (12 h).

An analysis and discussion of the results for each bulk water supply system is provided. This is followed by a summary of the investigation and conclusion on the performance of the compression heuristic.

Analysis of system 1

The first system that was analysed had a small supply ratio of 1.2 and a small reservoir size of 3 h under seasonal peak flow conditions. These parameters represent a highly unreliable bulk supply system, with a large number of reservoir failures. The simulation results for the compression heuristic and full Monte Carlo method are plotted in Figure 7.1.

The compression heuristic and full Monte Carlo simulations have very close averages of 157.68 and 157.76 failures/annum respectively. The percentage difference between the averages of the two methods is 0.048%. This result indicates that the compression heuristic is able to provide a good approximation to the actual failure frequency for reservoirs with low reliability. Figure 7.1 shows the average and 5% bounds for the full stochastic analysis (marked in black). As can be seen from the graph, the failure frequencies for both methods fall within the 5% bounds of the full Monte Carlo method. It is clear, however, that the variability of the full Monte Carlo results is greater than that of the compression heuristic. This is confirmed by the standard deviation calculated for each method, where the standard deviation for the compression heuristic is 1.48 failures/annum and 3.53 failures/annum for the full Monte Carlo method. These results suggest that for a system with low reliability, the compression heuristic provides more stable failure results.

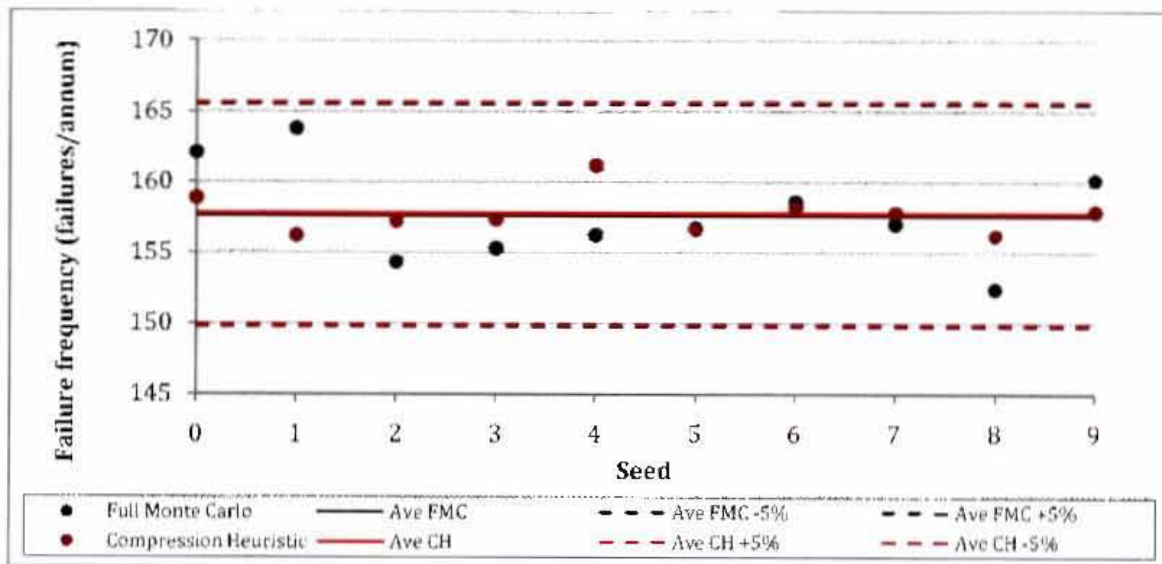


Figure 7.1 Comparison of failure results for system 1 ($SR = 1.2$; $RC = 3$ h)

Analysis of system 2

The next system analysed had a small supply ratio of 1.2 and a larger reservoir capacity of 12 h. These values are the same ones used in the example system tested in the stochastic analysis study by van Zyl *et al.* (2008). The failure frequency results of ten simulations for both the compression heuristic and the full Monte Carlo method are plotted in Figure 7.2. The average failure frequency for the compression heuristic is 1.965 failures/annum, greater than

that of the full Monte Carlo method (1.890 failures/annum). Figure 7.2 illustrates that the average failure frequency for the compression heuristic still lies within the 5% bounds of the average Monte Carlo method, although it provides a more conservative estimate. It can also be seen from the figure that the full Monte Carlo method has a slightly greater variability than the compression heuristic. It has a standard deviation of 0.044 failures/annum, while the compression heuristic has a standard deviation of 0.022 failures/annum.

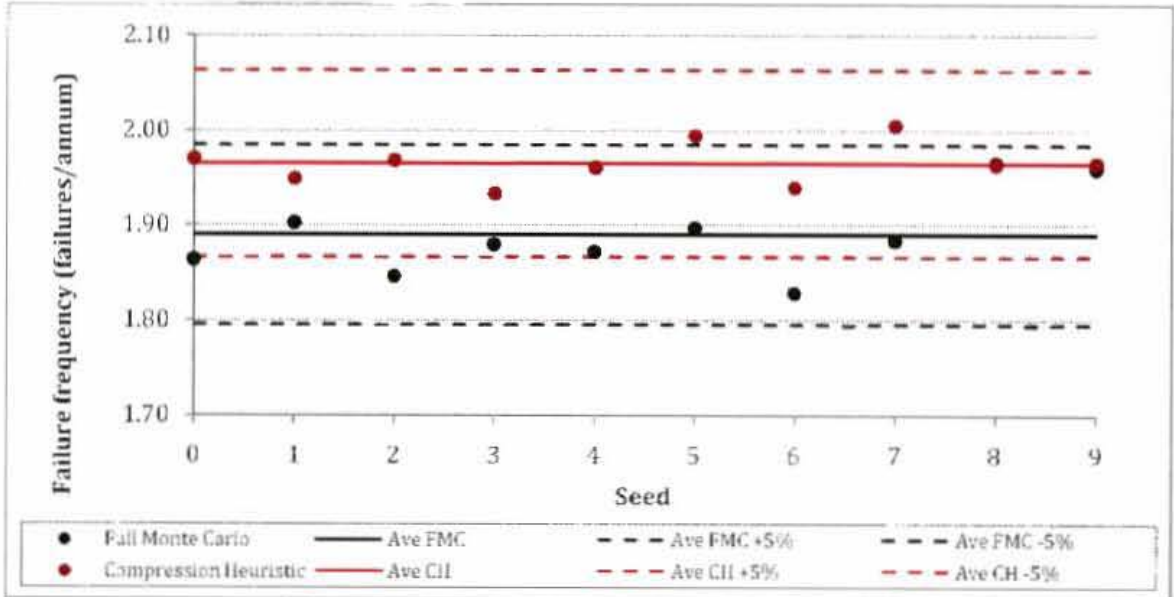


Figure 7.2 Comparison of failure results for system 2 (SR = 1.2; RC = 12 h)

Analysis of system 3

For case 3, a system with a small reservoir size of 3 h and a larger supply ratio of 1.5 was analysed. The failure results are shown in Figure 7.3. The compression heuristic has an average failure frequency of 26.87 failures/annum, which is slightly more conservative compared to the full Monte Carlo method with an average of 26.67 failures/annum. All failure results fall within the 5% bounds of the average value for the full Monte Carlo method. However the compression heuristic is less variable than the full Monte Carlo method. The standard deviation is 0.39 failures/annum and 0.73 failures/annum respectively.

Analysis of system 4

The final system analysed had a supply ratio of 1.5 and a reservoir capacity of 12 h. This represents a more reliable system. From the ten simulations for each method, the graph in Figure 7.4 was produced. The average value for the failure frequency of the compression heuristic is 0.1277 failures/annum, which is 1.44% larger than that of the full Monte Carlo method (0.1259 failures/annum). Again, the compression heuristic seems to provide a more conservative estimate of failure frequency. However, in this instance its results appear to have marginally greater variability, as its standard deviation is 0.0030 failures/annum and the standard deviation of the full Monte Carlo method is 0.0028 failures/annum.

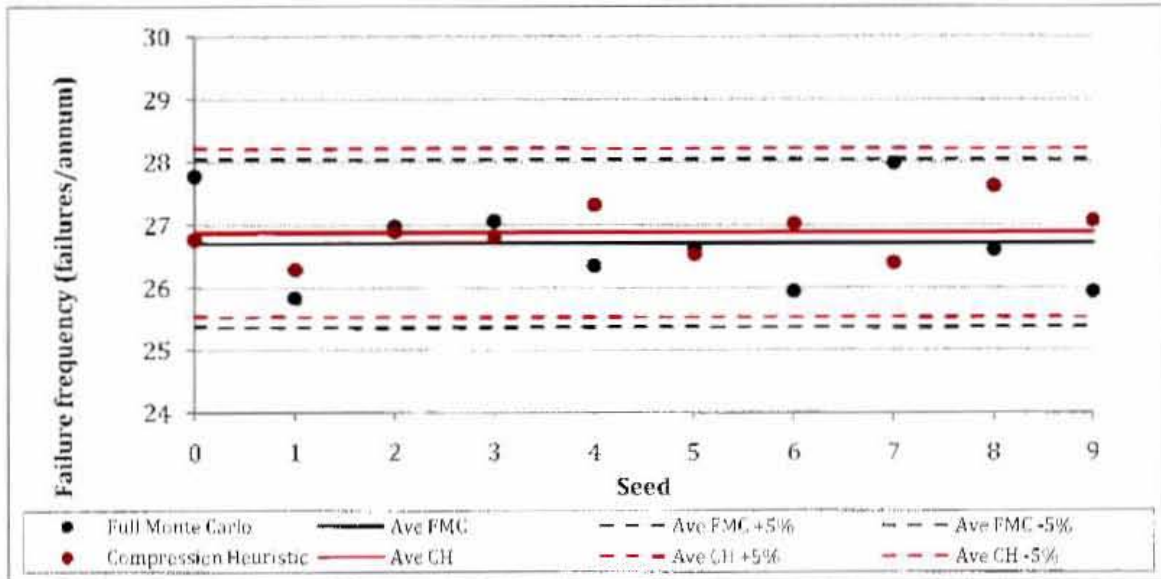


Figure 7.3 Comparison of failure results for system 3 ($SR = 1.5$; $RC = 3$ h)

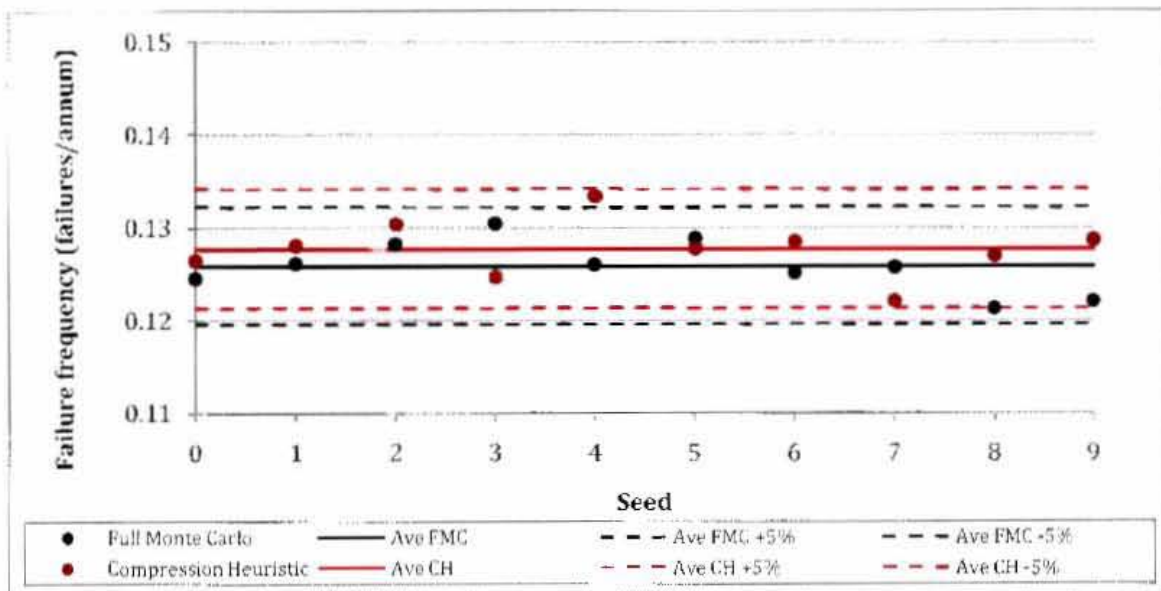


Figure 7.4 Comparison of failure results for system 4 ($SR = 1.5$; $RC = 12$ h)

Summary

In the above analysis, ten different random seeds were used for the simulations of the full Monte Carlo method and the compression heuristic. The results for all four cases show that the compression heuristic's failure frequency is very close to that of the full Monte Carlo, as its average value consistently falls within the 5% bounds of the average failure frequency for the full Monte Carlo method. In addition, the compression heuristic produces results which are usually more conservative. This analysis indicates that the compression heuristic produces similar results to the full Monte Carlo method, with good reproducibility and accuracy. It can thus be used as an alternative means of stochastically analysing the system's reliability.

7.2 Comparison of stochastic simulation times

7.2.1 Single stochastic simulation

In addition to comparing the failure results for both methods, a comparison of the simulation times for each of the above systems was performed. Each simulation was run until the stopping criteria discussed in Section 4.5 were met, where the designer selected duration is 2 million days. The average simulation times of ten different seeds for each system are highlighted in Table 7.1.

Table 7.1 Average simulation times and simulated hours for each system

SR	RC (h)	Simulation time (s)				Simulated hours (h)		
		Pre-run	Events	Total CH	Full MC	Pre-run	Events	Full MC
1.2	3	1.3	1.4	2.6	1.8	84541	519658	83573
1.2	12	192	96	288	95	48000000	32574477	10711889
1.5	3	86	8	94	15	5678773	3178571	660606
1.5	12	712	127	839	425	48000000	48001355	48000000

For all four systems tested individually, the total simulation time of the compression heuristic (combined times for the pre-run and the events run) is longer than the full Monte Carlo method. This result can be explained by the number of hours simulated in each run. Since the fire and pipe failure events are simulated separately in the events run, more time steps must be simulated so as to generate enough failures for the results to converge. This increased computational effort is exacerbated by the fact that both the pre-run and events run have to converge or meet the termination criteria separately, whereas only one run is required for the full Monte Carlo method. This comparison of simulation times and simulated hours demonstrates that the compression heuristic does not provide the desired increase in computational speed over the full Monte Carlo method for an individual simulation.

7.2.2 Stochastic simulation of a large number of systems together

Although it was at first discouraging that the compression heuristic took longer than the full Monte Carlo method for a single system, it was hypothesised that an improvement in computational speed could be achieved by the compression heuristic if it was applied to a large number of systems at the same time. This means that the pre-run is simulated only once at the beginning of the analysis, and the events run is then run for each individual system. In order to examine the improvement of speed achieved by the compression heuristic over the full Monte Carlo method, both stochastic methods were applied to an increasing number of systems from 10 up to 100 systems. These systems were randomly selected for analysis. The supply ratio was selected from the range 1.0 to 2.0, while the reservoir capacity was selected from the range of 4 h to 24 h of storage capacity.

The pre-run in the compression heuristic was simulated for a set of supply ratios, $SR = \{1.0; 1.2; 1.4; 1.6; 1.8; 2.0\}$ and a set of reservoir capacities, $RC = \{4; 8; 12; 16; 20; 24\}$ h. The full simulation time for the pre-run was 9411 seconds. This is the total time for 36 systems, and not just the time for one stochastic simulation as reported in the previous subsection. No matter how many systems are analysed in the events run, the pre-run simulation time always remains the same as it is only performed once. The simulation time for the events run is then added to the pre-run time to get the total simulation time for the compression heuristic, which is then compared to the simulation time for the full Monte Carlo method. The simulation times are shown in Table 7.2.

Table 7.2 Comparison of simulation times for a large number of simulations

No. simulations	CH Time (s)			Full MC Time (s)	CH/FMC
	Pre-run	Events run	Total run		
10	9411	875	10286	4958	2.1
20	9411	1832	11243	9256	1.2
50	9411	4151	13562	19411	0.7
100	9411	8845	18256	38455	0.5

For the analysis of 10 systems, the compression heuristic takes approximately twice as long as the full Monte Carlo method to complete the simulations. For 20 systems, the total simulation time for the compression heuristic is about 1.2 times longer than the full Monte Carlo method. At 50 and 100 systems, the benefit of using the compression heuristic becomes apparent, as it takes about 0.7 and 0.5 times the length of simulation time required for the full Monte Carlo method. Figure 7.5 shows that for this set of system parameters tested, the compression heuristic starts to be faster than the full Monte Carlo method after approximately 30 simulations. A full list of simulation times for all 100 systems can be found in Appendix B.

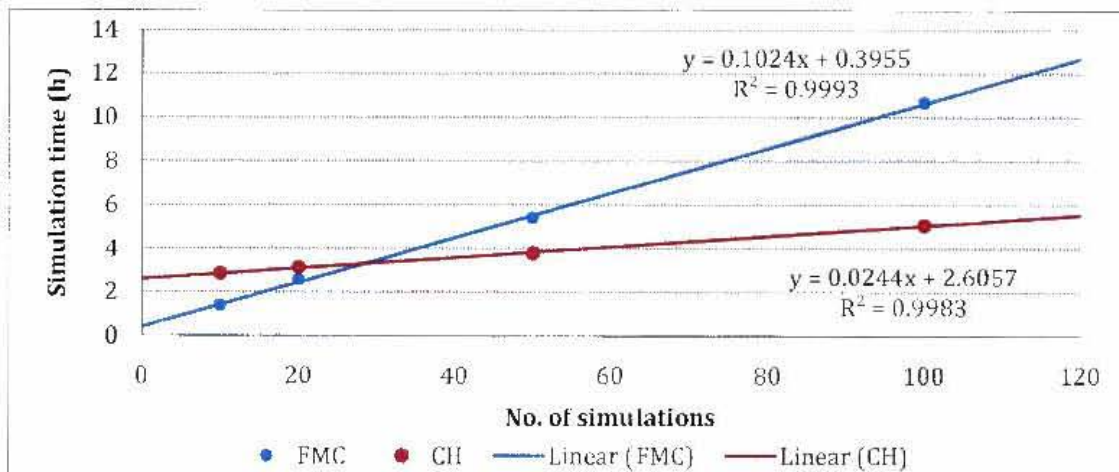


Figure 7.5 Comparison of simulation times for an increasing number of simulations

If the optimisation model was applied to a problem with a population size of 50 solutions, for 50 generations, then a total of 2550 reliability evaluations would be required. For the full

Monte Carlo method, if each system evaluation takes approximately 380 seconds, then the reliability evaluations would take 269 hours to complete. If the compression heuristic was used, assuming that the pre-run takes 9411 seconds and each events run takes 90 seconds on average, then the reliability evaluations would take 66 hours to complete. The compression heuristic would thus provide a 75% saving in computation time.

This investigation shows that the compression heuristic is capable of reducing the computational time required for the full Monte Carlo method for a large number of evaluations. For a fixed pre-run simulation, the more reliability evaluations carried out, the greater the computational savings in time.

7.3 Application of the reliability-based optimisation model

In this section, the optimisation model is tested for a single bulk water supply system and the resulting solution set is analysed. The system input values are as follows: the seasonal peak demand D_{peak} is 80 L/s, the system head H is 60 m, the length L of the feeder pipe system is 10 km and the Hazen-Williams coefficient is 120. The stochastic demand parameters are the same as described in the study by van Zyl *et al.* (2008), and the fire rate is 6 fires/annum, while the pipe failure rate is 0.2 failures/km/annum.

To recap, the aim of the optimisation procedure is to determine possible design solutions for the bulk water supply system, in terms of the supply pipe configuration, the pipe diameter and the size of the reservoir. Any one of the seven pipe configurations as described in Section 5.5 may be selected, whereas the possible pipe diameters are determined from the range of commercially available pipe sizes. The available pipe sizes used in this project were extracted from Raad (2010) and are shown in Table 7.3. The range for the reservoir size is determined by the user, which for the test system ranges from 4 h to 16 h of storage capacity.

Table 7.3 Range of available pipe diameters

Diameter (m)			
0.127	0.428	0.777	1.366
0.145	0.479	0.828	1.568
0.182	0.530	0.878	1.773
0.227	0.574	0.929	1.970
0.286	0.626	0.976	2.174
0.322	0.675	1.074	
0.363	0.726	1.176	

The optimisation model immediately starts by eliminating the pipe diameters that are impractical (i.e. too small or too large) for the test system, leaving the remaining pipes with diameters: 0.227 m, 0.286 m, 0.322 m and 0.363 m. The resulting supply ratios for each combination of pipe configuration and pipe diameter are summarised in Table 7.4.

Table 7.4 Supply ratios for different combinations of pipe configuration and diameter

Diameter (m)	Configuration		
	1 pipe	2 pipes	3 pipes
0.227	0.5330	1.0660	1.5990
0.286	0.9787	1.9573	2.9360
0.322	1.3367	2.6735	4.0102
0.363	1.8321	3.6641	5.4962

The demands-only pre-run is executed, and the results from the pre-run are stored for use in the main events run, namely, the demands failure frequency (Table 7.5), the full reservoir fraction (Table 7.6), and the lookup table for the reservoir level distribution curve.

Table 7.5 Demands failure frequency results from the pre-run

Supply Ratio	Reservoir Capacity (h)			
	4	8	12	16
0.5330	373.7425	362.8012	369.1505	375.7111
0.9787	254.7129	171.3845	151.1635	141.8123
1.0660	189.0312	67.5519	24.7372	9.8882
1.3367	39.4290	2.7690	0.1304	0.0025
1.5990	5.0717	0.1464	0.0032	0
1.8321	0.5959	0.0106	0	0
1.9573	0.1619	0.0059	0	0
2.6735	0.0003	0	0	0
2.9360	0	0	0	0
3.6641	0	0	0	0
4.0102	0	0	0	0
5.4962	0	0	0	0

Table 7.6 Full reservoir fraction results from the pre-run

Supply Ratio	Reservoir Capacity (h)			
	4	8	12	16
0.5330	0	0	0	0
0.9787	0.0809	0.0073	0.0022	0
1.0660	0.1754	0.0805	0.0791	0.0791
1.3367	0.9159	0.8558	0.8526	0.8535
1.5990	0.9999	0.9956	0.9958	0.9956
1.8321	1	0.9999	0.9999	0.9999
1.9573	1	1	1	1
2.6735	1	1	1	1
2.9360	1	1	1	1
3.6641	1	1	1	1
4.0102	1	1	1	1
5.4962	1	1	1	1

After the pre-run is completed, the optimisation procedure is carried out for 50 generations for a population size of 50 solutions. The final result is shown in the form of a trade-off curve between failure frequency and cost in Figure 7.6, with a close-up of the systems which meet the design criterion in Figure 7.7. A summary of the full solution set is also provided in Table 7.7.

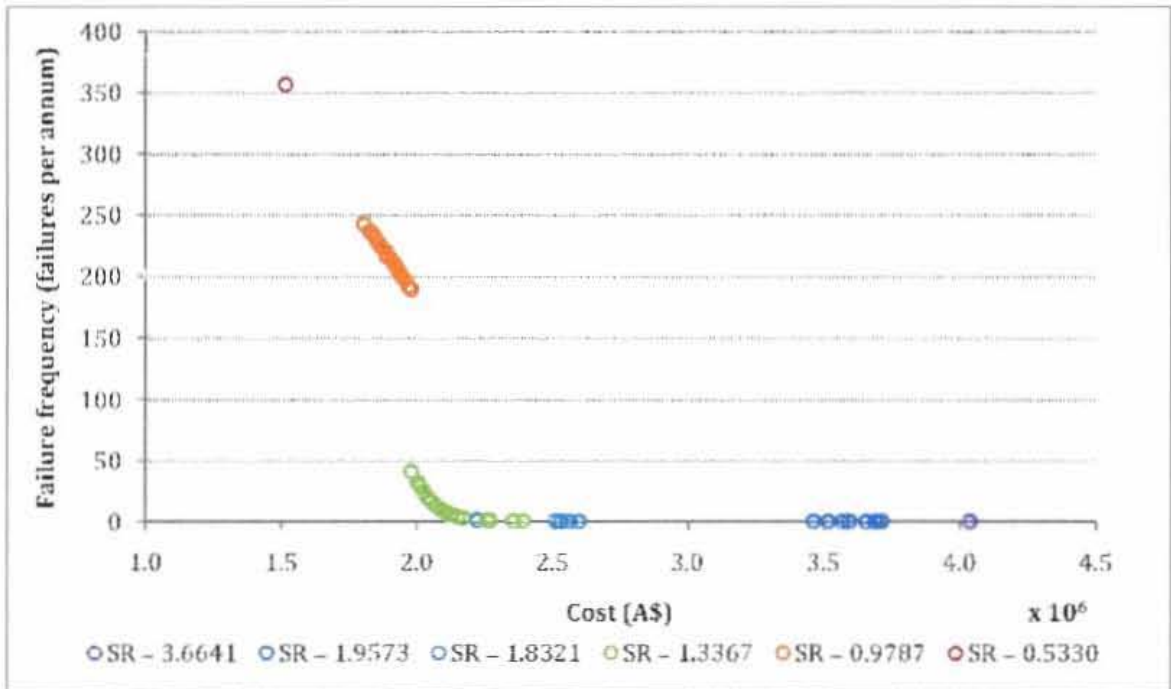


Figure 7.6 Pareto front of optimal solutions for the test system

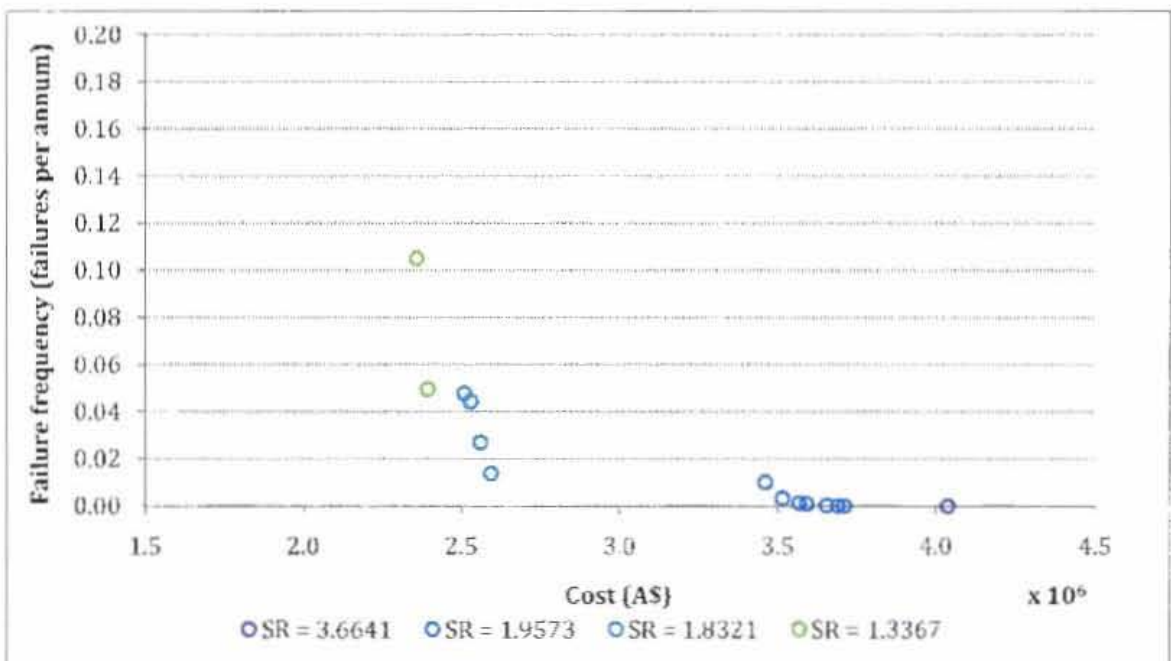


Figure 7.7 Pareto front of optimal solutions for the test system which meet the design criterion

* of seasonal peak demand

Configuration (pipes/interconn.)	Diameter (m)	SR*	RC [†] (h)	Cost (A\$)	Failure frequency (failures/annum)
2/0	0.363	3.6641	4.0000	4035963.12	< 0.0002
2/0	0.363	3.6641	4.0000	4035963.12	< 0.0002
2/0	0.286	1.9573	16.0000	3710500.37	< 0.0002
2/2	0.286	1.9573	14.4270	3690006.76	< 0.0002
2/1	0.286	1.9573	13.5553	3654427.78	0.0002
2/1	0.286	1.9573	11.4486	3592311.44	0.0008
2/2	0.286	1.9573	10.3197	3568094.11	0.0011
2/1	0.286	1.9573	9.0565	3513449.05	0.0031
2/1	0.286	1.9573	7.5223	3461256.69	0.0100
1/0	0.363	1.8321	16.0000	2593198.24	0.0136
1/0	0.363	1.8321	14.7819	2560322.53	0.0267
1/0	0.363	1.8321	13.6609	2529892.89	0.0441
1/0	0.363	1.8321	12.9527	2508717.31	0.0476
1/0	0.322	1.8367	15.8966	2393160.69	0.0494
1/0	0.322	1.8367	14.5358	2366254.61	0.1048
1/0	0.322	1.8367	11.6115	2271601.70	0.4311
1/0	0.322	1.8367	11.1963	2258878.03	0.5468
1/0	0.363	1.8321	4.8897	2223045.52	1.0748
1/0	0.322	1.8367	8.5375	2171954.45	2.4837
1/0	0.322	1.8367	7.8940	2149185.27	3.5715
1/0	0.322	1.8367	7.1606	2122043.58	5.5949
1/0	0.322	1.8367	6.6589	2102663.93	7.6610
1/0	0.322	1.8367	6.3295	2089494.15	9.2225
1/0	0.322	1.8367	5.9856	2075292.13	11.5775
1/0	0.322	1.8367	5.8456	2069362.53	12.5632
1/0	0.322	1.8367	5.4828	2053515.04	15.8343
1/0	0.322	1.8367	5.1822	2039776.64	19.1830
1/0	0.322	1.8367	4.7968	2021150.02	24.5558
1/0	0.322	1.8367	4.5614	2009089.86	28.9622
1/0	0.322	1.8367	4.4139	2001219.36	31.6906
1/0	0.322	1.8367	4.0000	1977642.25	40.4481
1/0	0.286	0.9787	7.9773	1977575.72	188.9838
1/0	0.286	0.9787	7.6809	1966835.53	191.6911
1/0	0.286	0.9787	7.2285	1950036.37	197.8197
1/0	0.286	0.9787	7.0724	1944116.22	199.6380
1/0	0.286	0.9787	6.8307	1934812.61	201.9789
1/0	0.286	0.9787	6.7544	1931037.57	204.3048
1/0	0.286	0.9787	6.5359	1923215.16	207.1175
1/0	0.286	0.9787	6.3112	1914172.92	210.3233
1/0	0.286	0.9787	5.9996	1901303.58	214.1469
1/0	0.286	0.9787	5.6396	1885878.59	216.3900
1/0	0.286	0.9787	5.5985	1884072.01	219.9988
1/0	0.286	0.9787	5.4705	1878391.47	220.9281
1/0	0.286	0.9787	5.1594	1864135.60	224.5219
1/0	0.286	0.9787	4.9196	1852651.14	228.3321
1/0	0.286	0.9787	4.7402	1843725.45	232.1735
1/0	0.286	0.9787	4.4853	1830483.79	235.1221
1/0	0.286	0.9787	4.3833	1824975.92	236.7835
1/0	0.286	0.9787	4.0000	1803065.60	242.4032
1/0	0.227	0.533	4.0000	1513744.08	356.9697

Table 7.7 Summary of final solutions for the test system

Table 7.8 Summary of costs for the test system solutions

Configuration (pipes/interconn.)	Diameter (m)	RC ^a (h)	Pipe Cost (A\$)	Chamber Cost (A\$)	Reservoir Cost (A\$)	Total Cost (A\$)
2/0	0.363	4.000	3,722,064.78	0.00	313,898.34	4,035,963.12
2/0	0.363	4.000	3,722,064.78	0.00	313,898.34	4,035,963.12
2/0	0.286	16.000	2,978,334.51	0.00	732,165.85	3,710,500.37
2/2	0.286	14,427.0	2,978,334.51	22,173.22	689,498.92	3,690,006.65
2/1	0.286	13,553.3	2,978,334.51	11,086.61	665,007.93	3,654,429.05
2/1	0.286	11,448.6	2,978,334.51	11,086.61	602,889.00	3,592,310.12
2/2	0.286	10,319.7	2,978,334.51	22,173.22	567,587.88	3,568,095.61
2/2	0.286	9,056.5	2,978,334.51	11,086.61	526,029.08	3,515,450.21
2/1	0.286	7,522.3	2,978,334.51	11,086.61	471,834.94	3,461,256.06
1/0	0.363	16,000	1,861,032.39	0.00	732,165.85	2,593,198.24
1/0	0.363	14,781.9	1,861,032.39	0.00	699,290.24	2,560,322.63
1/0	0.363	13,690.9	1,861,032.39	0.00	668,860.50	2,529,892.89
1/0	0.363	12,952.7	1,861,032.39	0.00	647,686.23	2,508,718.62
1/0	0.322	15,896.6	1,861,032.39	0.00	362,015.17	2,223,047.56
1/0	0.322	14,535.8	1,663,743.90	0.00	729,417.11	2,393,161.02
1/0	0.322	11,611.5	1,663,743.90	0.00	692,511.29	2,356,255.20
1/0	0.322	11,196.3	1,663,743.90	0.00	607,857.10	2,271,601.00
1/0	0.363	4,889.7	1,663,743.90	0.00	595,134.27	2,258,878.18
1/0	0.322	8,537.5	1,663,743.90	0.00	508,211.67	2,171,954.97
1/0	0.322	7,894.0	1,663,743.90	0.00	485,410.24	2,149,154.14
1/0	0.322	7,160.6	1,663,743.90	0.00	458,360.86	2,122,044.77
1/0	0.322	6,658.9	1,663,743.90	0.00	438,921.20	2,102,665.11
1/0	0.322	6,329.5	1,663,743.90	0.00	425,275.01	2,089,494.03
1/0	0.322	5,985.6	1,663,743.90	0.00	411,549.62	2,075,293.52
1/0	0.322	5,845.6	1,663,743.90	0.00	405,618.11	2,069,362.02
1/0	0.322	5,482.8	1,663,743.90	0.00	389,272.87	2,053,516.78
1/0	0.322	5,182.2	1,663,743.90	0.00	376,034.12	2,039,778.02
1/0	0.322	4,796.8	1,663,743.90	0.00	357,406.81	2,021,150.72
1/0	0.322	4,561.4	1,663,743.90	0.00	345,344.43	2,009,088.33
1/0	0.322	4,413.9	1,663,743.90	0.00	337,474.79	2,001,218.69
1/0	0.322	4,000	1,663,743.90	0.00	313,898.34	1,977,642.25
1/0	0.286	7,977.3	1,489,167.26	0.00	488,409.85	1,977,577.10
1/0	0.286	7,680.9	1,489,167.26	0.00	477,666.50	1,966,833.75
1/0	0.286	7,228.5	1,489,167.26	0.00	460,867.65	1,950,034.90
1/0	0.286	7,072.4	1,489,167.26	0.00	454,947.73	1,944,114.99
1/0	0.286	6,830.7	1,489,167.26	0.00	446,432.79	1,934,811.05
1/0	0.286	6,754.4	1,489,167.26	0.00	442,669.84	1,931,837.09
1/0	0.286	6,535.9	1,489,167.26	0.00	434,048.19	1,923,215.45
1/0	0.286	6,311.2	1,489,167.26	0.00	425,066.68	1,914,173.94
1/0	0.286	5,999.6	1,489,167.26	0.00	412,137.70	1,901,304.95
1/0	0.286	5,639.6	1,489,167.26	0.00	396,270.82	1,885,878.08
1/0	0.286	5,598.5	1,489,167.26	0.00	394,906.13	1,884,073.39
1/0	0.286	5,470.5	1,489,167.26	0.00	389,222.41	1,878,389.66
1/0	0.286	5,159.4	1,489,167.26	0.00	374,966.05	1,864,133.30
1/0	0.286	4,919.6	1,489,167.26	0.00	363,481.46	1,852,648.72
1/0	0.286	4,740.2	1,489,167.26	0.00	354,534.59	1,843,725.85
1/0	0.286	4,483.3	1,489,167.26	0.00	341,316.13	1,830,483.38
1/0	0.286	4,383.3	1,489,167.26	0.00	335,809.52	1,824,976.78
1/0	0.286	4,000	1,489,167.26	0.00	313,898.34	1,803,065.60
1/0	0.227	4,000	1,199,845.74	0.00	313,898.34	1,513,744.08

^a of seasonal peak demand

7.3.1 General observations

The Pareto front shows the optimal set of solutions obtained, ranging from cheap but unreliable solutions on one end of the curve to reliable but costly solutions on the other end. It clearly illustrates that there is a trade-off between cost and failure frequency. If the designer desires a more affordable solution, then a cheaper solution can be selected, but at the risk of a lower reliability. On the other hand, the designer may be willing to spend more money on a system so that a greater reliability can be obtained.

The red solution indicated in the trade-off curve in Figure 7.6 is the cheapest solution. To be able to achieve such a low cost, the supply ratio was minimised to 0.53 through the selection of a single 0.227 m diameter pipeline. The smallest reservoir capacity of 4 h of seasonal peak demand was also selected. This cost of A\$ 1,513,744 was obtained to the detriment of the reliability of the solution, as it has a high failure frequency of 357 failures/annum, which is nearly equivalent to one failure per day.

The cost of a pipeline is a function of its diameter and length. Looking at the single pipelines in Table 7.8, a clear trend of increasing cost is observed with increasing pipe diameter. In addition, the cost of pipelines increases with length, and so two parallel pipes are double the cost of a single pipeline with the same diameter, assuming that the pipes are laid in separate trenches. Thus, there is a very distinct gap between solutions with a single pipeline and solutions with two parallel pipes. The majority of solutions in the Pareto front consist of systems with single pipelines. This configuration is a ‘popular’ option because of its low cost.

Systems with single pipelines generally have a high failure frequency, especially if the selected pipe configuration and diameter provide a supply ratio less than one under seasonal peak conditions. A bulk system with supply ratio less than one would not be able to meet the design criterion of one failure in ten years under seasonal peak flow conditions, as the supply inflow would struggle to fill the storage reservoir; therefore this option is not a viable solution, even though it is more affordable.

In some instances, systems with single pipelines are able to meet the desired reliability design criterion, as illustrated by the green and light blue solutions in Figure 7.7. For example, the solution with a single pipeline of diameter 0.322 m and reservoir capacity of 14.5 h of seasonal peak demand has a failure frequency of 0.1048 failures/annum, which is nearest to the desired reliability of 0.1 failures/annum under seasonal peak conditions. For any other system with a single pipeline, the system should either have a reservoir capacity greater than 14.5 h or pipe diameter of 0.363 m.

Keep in mind that the selection of practical pipe diameters was based on the elimination of a diameter if a single pipeline provided a supply ratio greater than 2 times seasonal peak demand. For this typical system, the largest diameter allowed was 0.363 m, thus for a single pipe configuration the supply ratio is 1.83. It is recognised that a Pareto-optimal solution

consisting of a single pipeline with the next commercially available pipe size of 0.428 m, and a smaller reservoir capacity could exist. This solution would have a supply ratio of 2.83 times seasonal peak demand. In this project, however, the maximum allowable supply ratio was limited to 2 times seasonal peak demand for all applications of the optimisation model. It is recommended that the maximum allowable supply ratio should be investigated further in future studies.

There are other solutions (dark blue) which are able to meet the desired reliability. These consist of systems with pipe configurations consisting of two pipes in parallel. Pipes in parallel provide additional reliability to the system, as at least half the supply into the reservoir remains when there is a pipe outage. It was observed that the reservoir sizes for these solutions are notably smaller than for systems with single pipeline configurations. For this application to a typical system, it was found that parallel pipe systems are more reliable, yet more expensive than single pipe systems. This once again emphasises that cost and reliability of a system are conflicting objectives.

The purple solution in Figure 7.6 and Figure 7.7 represents the most reliable, yet most expensive system at A\$ 4.037 million (Table 7.8). It consists of two pipes in parallel and no interconnections, with the largest available pipe diameter of 0.363 m, but with the smallest allowable reservoir capacity of 4 h of seasonal peak demand. Since the inflow capacity is 3.66 times greater than the seasonal peak, the storage reservoir experiences a negligibly small number of failures. The high cost of the system can be attributed to the fact that the supply pipe consists of two 0.363 m diameter pipes in parallel, which contributes to 92% of the total system cost. It is nearly twelve times the cost of the bulk storage reservoir.

For bulk systems with a feeder pipe length of 10 km between its source and the storage reservoir, the supply pipe system contributes between approximately 70% to 90% of the total cost of the bulk system. The size of the reservoir appears to be of little consequence in comparison to the pipe system with regards to cost considerations. From Table 7.8, it can also be seen that the costs of the chambers for the interconnection of parallel pipes are insignificant as they make a marginal difference to the total cost.

It is also interesting to note that for solutions where the pipe cost is at a minimum, at approximately 70% of the total cost, these solutions also correspond to the region of solutions that would be examined more closely by the engineer, i.e. cheapest solution meeting the desired reliability criterion.

Gaps between solutions in the Pareto front

The solutions in Figure 7.6 and Table 7.7 are highlighted according to their supply ratios. As can be seen in Figure 7.6, there are large gaps between solutions caused by the discrete nature of the optimisation problem, as both the pipe configuration and pipe diameter are discrete variables. Only the reservoir capacity is a continuous variable. The gaps represent the change in supply ratio, which is determined by the solutions' pipe configuration and pipe diameter.

The supply ratios of the solutions have a direct impact on the system reliability. The gaps are also linked to changes in the system cost.

Working from left to right of the Pareto front in Figure 7.6, each gap in the front is explained. The red, orange, green and light blue solutions all have systems with single pipe configurations. The gaps between the red and orange solutions, the orange and green solutions, and the green and light blue solutions are caused by increases in the pipe diameter. In contrast, the significant gap between the light blue to the dark blue solutions is explained by the change in pipe configuration from a single pipeline to two pipes in parallel. Although both pipe configurations are capable of producing systems with similar reliability, the additional pipe has a large impact on the cost of the system. The gap represents a 33.5% increase in cost. Lastly, the gap between the dark blue and purple solutions is explained by an increase in the pipe diameter from 0.286 m to 0.322 m for two pipes in parallel, which represents an increase in the supply ratio from 1.96 to 3.66.

Using the optimisation model, a Pareto front can be obtained for any set of system parameters. Regardless of what parameters are used, gaps will exist in all trade-off curves due to the discrete nature of the optimisation problem. These gaps can be explained by either a change in pipe diameter or pipe configuration between adjacent solutions. From the analysis of the solutions and the gaps between them, a trend in optimal variables is observed. Single pipe solutions are included in the front as they usually provide the cheapest systems, but also the least reliable ones. Solutions with single pipeline configurations can be reliable and meet the desired design criterion so long as a large enough diameter is provided to produce an acceptable supply ratio. From the Pareto front, it is clear that parallel pipe solutions dominate the right-hand end of the curve, illustrating that a change in pipe configuration from single to parallel pipes improves the reliability of the system, at the expense of a higher cost.

Absence of systems with three parallel pipes

It was at first surprising that no systems with three pipes in parallel appeared in the final solution set. Only solutions with single pipelines and two pipes in parallel with a varying number of interconnections were produced. The reason for this is that some bulk systems with two parallel pipes have similar reliability to what systems with three pipes in parallel would have, but at a reduced cost. To confirm this observation, a number of systems with three parallel pipes and 0.227 m diameter were evaluated. These systems have a supply ratio of 1.599 under seasonal peak conditions. The cost and failure frequency for each system is shown in Table 7.9, and compared to the most expensive and reliable system obtained through the optimisation procedure.

Table 7.9 illustrates that systems with three pipes in parallel are generally more costly than the most expensive solution in the Pareto front. Where a system with three parallel pipes is cheaper, it is unable to meet the reliability of the most expensive solution in the optimisation. The solutions with two parallel pipes consequently outperform the solutions with three parallel pipes in the multi-objective optimisation, as illustrated in Figure 7.8.

Table 7.9 Costs for systems designed according to South African deterministic guidelines

Configuration (pipes/inter.)	Diameter (m)	SR*	RC* (h)	Total Cost (A\$)	Failure Frequency (failures/annum)
2/0	0.363	3.6641	4	4035963	< 0.0002
3/0	0.227	1.599	4	3913436	5.0851
3/0	0.227	1.599	8	4088762	0.1460
3/0	0.227	1.599	12	4219124	0.0073
3/0	0.227	1.599	16	4331703	< 0.0002

* of seasonal peak demand

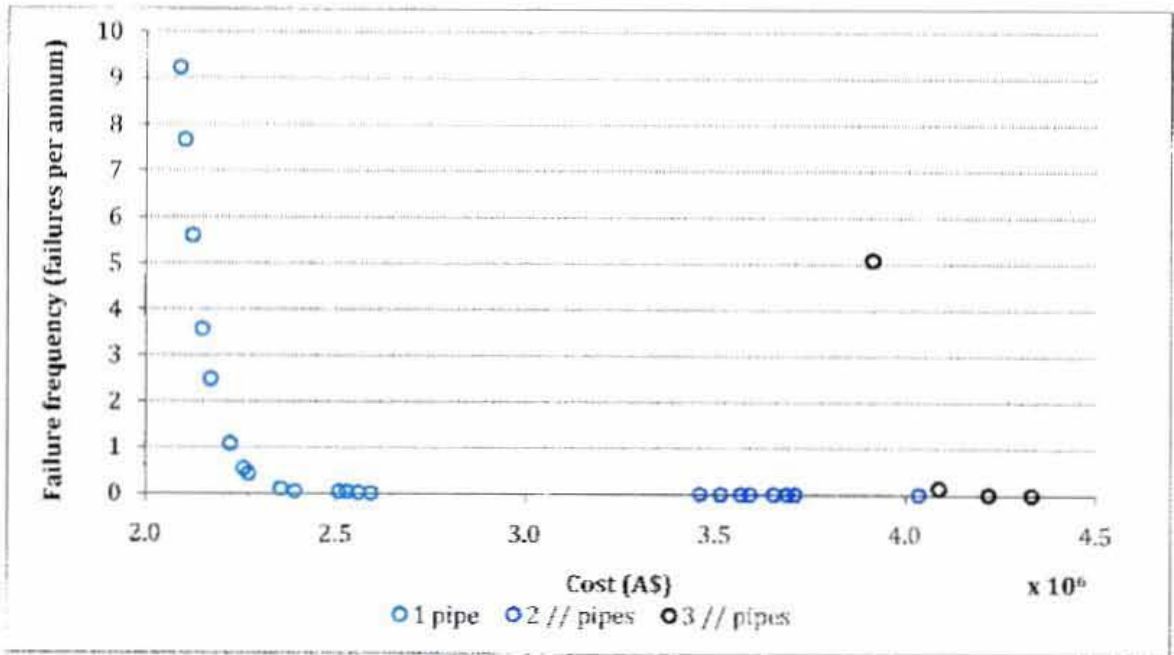


Figure 7.8 Pareto front for the test system including solutions with three pipes in parallel (close-up)

Comparison with deterministic design guidelines

The region that the designer would be concerned with is the solutions with single pipelines with large diameters greater than or equal to 0.322 m and a fairly large reservoir size with storage capacity greater than 12 h. The designer would probably be most interested in using the solution with the pipe configuration which gives a supply ratio of 1.34 and has a reservoir capacity of 14.54 h, since this is the cheapest solution (A\$ 2.36 million) that is able to achieve the reliability design criterion with just over 0.1 failures/annum under seasonal peak conditions. This supply ratio of 1.34 (i.e. the inflow rate to the reservoir is 1.34 times the seasonal peak demand), is equivalent to 2 times AADD. This is greater than 1.5 times AADD as stipulated in the Red Book (CSIR, 2000) for the capacity of the supply main to the reservoir. In contrast, the reservoir capacity is 14.54 h of seasonal peak demand, equivalent to 21.8 h of AADD, is less than the Red Book's suggested storage capacity of 48 h of AADD (CSIR, 2000). The Red Book also allows for a reservoir capacity of 24 h of AADD, although it does not explicitly state when this reservoir size can be used.

Suppose that the bulk system was designed according to the deterministic guidelines stated in the Red Book, where the capacity of the supply main is 1.5 times AADD (i.e. one times seasonal peak demand) and the storage reservoir capacity is 48 h of AADD (i.e. 32 h of seasonal peak demand), then two possible design options are available. Two design options are also available for a system where the supply capacity is approximately 1.5 times AADD and the reservoir capacity is 24 h of AADD. These options with their associated costs are shown in Table 7.10. The supply capacity of 1.5 times AADD is equivalent to one times seasonal peak demand. This supply ratio of one is too low for a reliable design, and it is thus a critical point. The guidelines-based solutions from Table 7.10 are plotted in the same graph as the Pareto front in Figure 7.9. The figure clearly demonstrates that the design solutions obtained using the deterministic guidelines are dominated by the solutions obtained through the optimisation model.

Table 7.10 Systems designed according to South African deterministic guidelines

Configuration (pipes/inter.)	Diameter (m)	SR (AADD)	SR*	RC [AADD] (h)	RC* (h)	Total Cost (A\$)	Failure Frequency (failures/annum)
1/0	0.286	1.4681	0.9787	48	32	2,583,671	140.0671
2/0	0.227	1.5990	1.0660	48	32	3,494,195	0.3098
1/0	0.286	1.4681	0.9787	24	16	2,221,333	145.6258
2/0	0.227	1.5990	1.0660	24	16	3,131,857	10.3485

* of seasonal peak demand

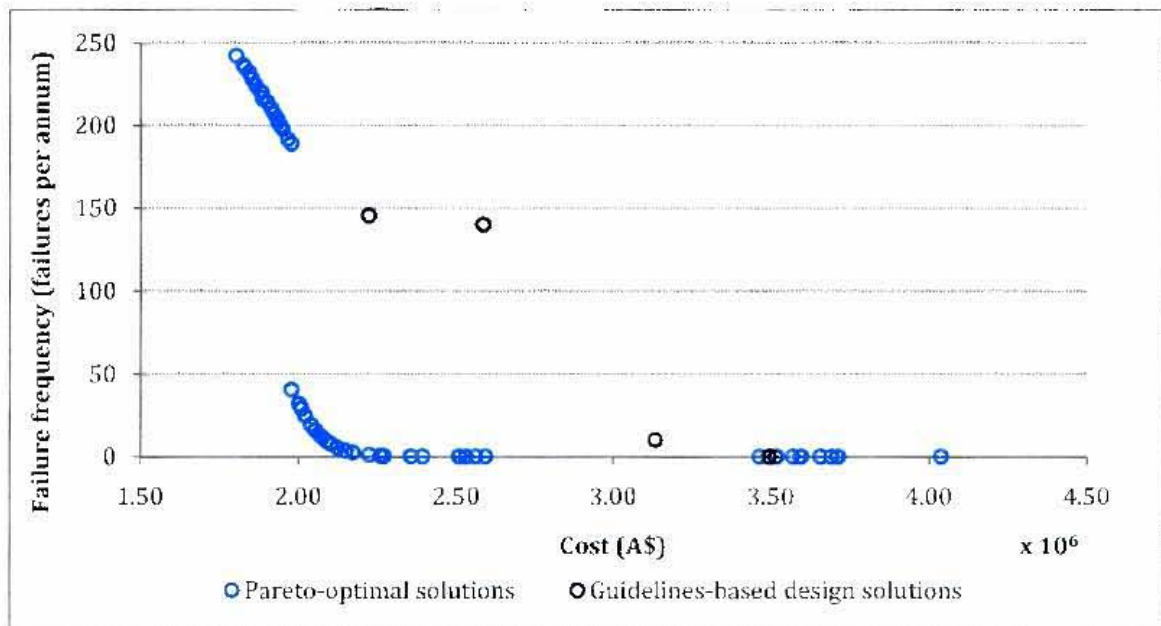


Figure 7.9 Pareto front for the test system including guidelines-based designs

From the table, it is apparent that the South African deterministic guidelines produce a more costly system than if it was designed using the optimisation model, as the total cost for both systems with reservoir capacity of 32 h of seasonal peak demand, or 48 h of AADD,

(A\$ 2.6 million and A\$ 3.5 million) is greater than the optimised solution described above with a system cost of A\$ 2.36 million. The large reservoir capacity of 32 h accounts for the increased cost. Where the reservoir capacity is 24 h of AADD, i.e. 16 h of seasonal peak demand, the deterministic guidelines provide one system that is slightly cheaper (A\$ 2.2 million) than the optimised solution, but at a much higher failure rate. The other system is more expensive at a cost of A\$ 3.1 million. This result suggests that a more economical system can be designed if a larger supply capacity is selected in preference to a bulk reservoir with large storage capacity, for this set of system parameters.

It is stressed that the bulk systems designed according to the deterministic guidelines, shown in Table 7.10, actually fail to meet the design criterion of 0.1 failures/annum under seasonal peak flow conditions. This is because the supply ratio is either slightly less or slightly greater than the critical value of one. For a single pipeline, the supply ratio is 0.98, which means that the seasonal peak demand is greater than the supply inflow into the reservoir. This leads to an extremely high failure frequency of 140 failures/annum. The pipe diameter selected in this case is too small and a larger diameter should be selected. A pipe that is one size larger would have a diameter of 0.322 m, and would provide a supply ratio of 1.34 of seasonal peak demand or 2 times AADD. This pipe configuration and new diameter is coincidentally the same as the Pareto-optimal solution that the designer would be most interested in.

For the systems with a parallel pipe configuration in Table 7.10, the supply ratio is 1.066 times seasonal peak demand. The failure frequency is 0.3098 and 10.3485 failures/annum for the reservoir size of 32 h and 16 h of seasonal peak demand respectively. These systems still fail to meet the design criterion of 0.1 failures/annum.

These failure results suggest that the minimum supply capacity of 1.5 times AADD as specified by the Red Book is inadequate. The supply capacity should be increased so that the design criterion can be met, like it was done for the optimised solution. The obvious question leading from this is: if bulk storage reservoirs are designed according to the Red Book in practice, then why do they not fail all the time in the field? A possible reason for this is that seasonal peak conditions only happen for a short period of the year, thus not as many failures would occur in any year. Another explanation is that the failure rate of the system is highly sensitive to the supply ratio (van Zyl *et al.*, 2010), and the lack of failures in practice can be attributed to the design of systems with higher supply ratios. Engineers design the bulk system for a 20 to 40 year design period and thus design the system using a higher AADD value than the current one. In addition, only discrete pipe diameters are available and so the supply ratio in practice is likely to be higher than the minimum Red Book specification.

Note that the assumed seasonal demand pattern of 1.5 is not an unrealistically or unreasonably high value in practice. Since a supply ratio of 1.5 is clearly inadequate under such conditions, it is suggested that the minimum supply ratio specified by the Red Book should be reviewed.

7.3.2 Summary

The optimisation model was applied to a test system with certain parameters. It was effective in excluding the impractical pipe diameters, and preserving the commercially available pipe diameters of 0.227 m, 0.286m, 0.322 m and 0.363 m. Once the optimisation procedure was completed, a final set of 50 solutions was generated, each of which has a certain pipe configuration, pipe diameter and reservoir size. Each of these solutions has an associated reliability and cost. The designer can determine which solutions are of the most interest based on how much reliability he is willing to sacrifice in favour of a cheaper system. Assuming that the designer wants the bulk system to meet the design criterion of one failure in ten years under seasonal peak demand conditions; the most attractive solution would be a system with a single 0.322 m supply pipe, giving a supply ratio of 1.34 times seasonal peak demand and reservoir storage capacity of 14.54 h of seasonal peak demand.

It is evident that the optimisation model is able to assist the designer in narrowing down thousands of possible design options to just a handful as indicated by the findings in this section. From these solutions, the search can be reduced and refined so that the final system design can be reliable and cost-effective.

7.4 Sensitivity analysis

Thus far, a single bulk water supply system has been optimised and the results for the system were presented and discussed in the previous section. Using this system as the base system, a sensitivity analysis on the method is performed by investigating the effects of varying individual system parameters or stochastic parameters. The system parameters refer to the physical aspects of the system, such as the length of the feeder pipe system, and the system head. The stochastic parameters refer to external factors affecting the system, such as the consumer demand patterns, the fire occurrence rate, the pipe failure rate and the average pipe failure duration.

The base system is assumed to have parameters with typical values, most of which were selected based predominantly on the parameters as described by van Zyl *et al.* (2010). Where necessary, these typical values were also identified based on literature and accepted values in practice. A similar method was followed for identifying the low and high values for each parameter. In the sensitivity analysis, each individual parameter is taken and the effect of changing it to a low and high value is analysed and compared to the typical value. The values selected for each sensitivity parameter are summarised in Table 7.11.

The approach used to analyse the results from the sensitivity analysis followed two steps. Firstly, for each sensitivity parameter, the trade-off curves for the low, typical and high values were plotted on the same graph. Any shifts of the curves to the left or right of the typical one were observed and discussed, as well as any changes in the spread of results. Secondly, a comparison was made by selecting the solutions with a failure frequency closest

to 0.1 failures/annum under seasonal peak conditions. Each solution is most likely not going to have an exact failure frequency of 0.1 failures/annum, therefore it is necessary to keep this variation in mind when comparing the results. It must also be remembered that stochastic simulations have an inherent variability in their results as illustrated earlier. The full set of solutions for each sensitivity parameter is shown in Tables C.1 to C.6 in Appendix C.

Table 7.11 Sensitivity analysis parameters

Sensitivity Parameter	Low	Typical	High
Length (km)	1	10	100
Head (m)	30	60	120
Hourly demand pattern: Peak factor	1.25	1.49	1.75
Fire rate (fires/a)	0	6	24
Pipe failure rate (failure/km/a)	0.1	0.2	0.5
Average failure duration (h)	3	4.5	9

In the optimisation process, the range of design variables was kept constant. The pipe configuration was limited to the seven different options as described before. The full selection of commercially available pipe sizes was allowed, but the range of practical pipe diameters was selected using the Hazen-Williams equation and the eliminating conditions described in Section 4.2. Finally, the reservoir capacity values ranged from 4 h to 16 h, as previously specified for the base system. During the execution of the sensitivity analysis, it was found that it was necessary to explore reservoir capacities greater than 16 h. The reliability analysis of these systems with larger reservoir capacities was performed using the full stochastic analysis model for each individual system. In each optimisation, 50 Pareto-optimal solutions were produced after 50 generations.

7.4.1 Effect of length

Obviously the length of the feeder pipe system depends on the distance between the supply source and bulk storage reservoir. For the typical value, a pipe length of 10 km was assumed. This value seems reasonable as the Drakenstein Pipeline which conveys water from the Drakenstein Pump Station to the Berg River Dam in the Western Cape is 10 km long. Bulk supply pipelines can, however, be much longer than 10 km, for example the Western Aqueduct in eThekweni is 73 km long. Thus the high value for the length parameter was chosen as 100 km. A short pipe length of 1 km was selected as the low value.

The length of the feeder pipe system has a direct impact on the pipe diameters selected for the optimisation process. The Hazen-Williams head-loss formula indicates that for a constant flow rate and a constant head, if the length of the pipe is increased, then the diameter must increase. Therefore, to obtain acceptable supply ratio values between 1.0 and 2.0, a 100 km long pipe system requires larger diameter pipes than a 10 km and 1 km long system. The allowed pipe diameters for each pipe length are shown in Table 7.12.

Table 7.12 Pipe diameter selection for each supply pipe system length:

L = 1 km	L = 10 km	L = 100 km
0.127 m	0.227 m	0.322 m
0.145 m	0.286 m	0.363 m
0.182 m	0.322 m	0.428 m
0.227 m	0.363 m	0.479 m
		0.530 m
		0.574 m

The trade-off curves for the different pipe lengths are illustrated in Figure 7.10, with the cost displayed on a log scale. For the 100 km long system, there was a significantly large shift to the right of the original curve for the 10 km system; while there is a distinct shift to the left for the 1 km long system. These shifts of the trade-off curves are caused by the large differences in cost for each system. The curves show that an increase in order of magnitude of the pipe length results in an associated increase in an order of magnitude of the cost for the system. To illustrate this point, the most expensive solution costs A\$ 553,868 for the 1 km system, A\$ 4,035,963 for the 10 km system, and A\$ 53,444,689 for the 100 km system (see Table C.1, Appendix C). The increase in cost for each system is directly related to the length of the supply system. The cost of a pipeline is a function of its diameter and its length. Thus for a system that has an increased length, a larger diameter is required, leading to a more expensive system. In addition, a longer pipe system experiences more total failures per annum, resulting in the need for a larger reservoir, which is more costly.

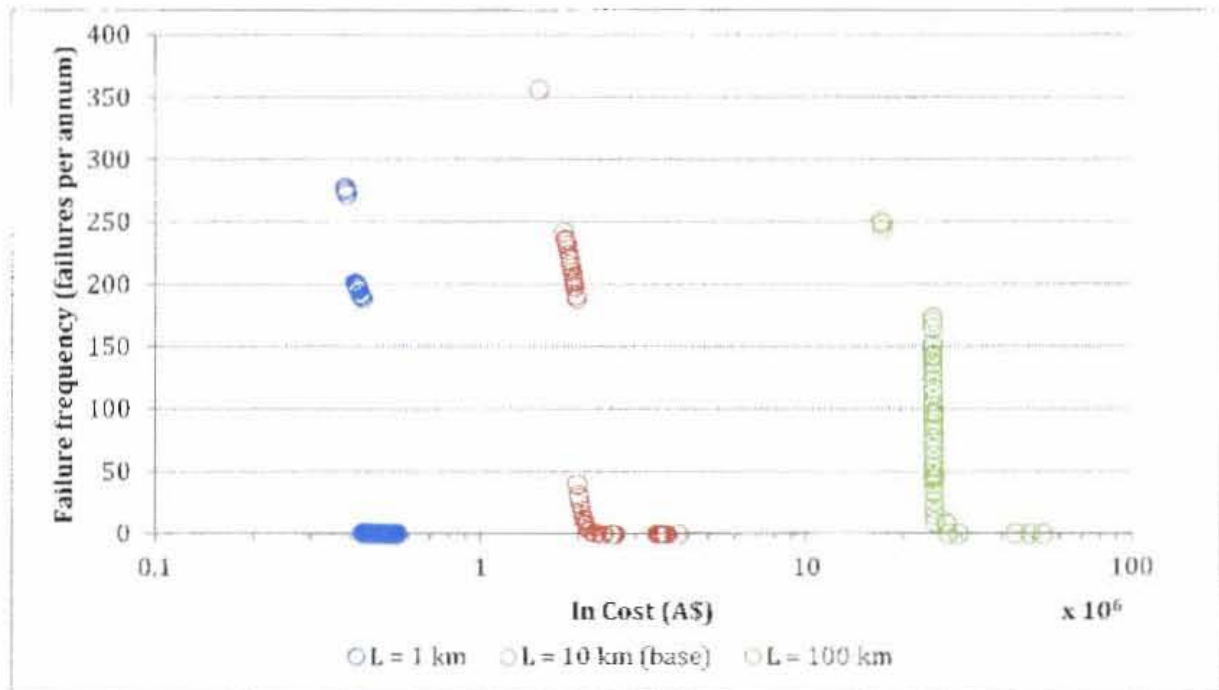


Figure 7.10 Pareto fronts for systems with different pipe lengths

It was noticed that the location and size of the gaps between adjacent solutions are different for each pipe length. As discussed in Section 7.3, gaps exist between adjacent solutions because of the discrete nature of the optimisation problem. These gaps are caused by changes

in the pipe diameter or pipe configuration, which result in a change in supply ratio. For the typical system, there are three distinct gaps in the Pareto front. The first two gaps are caused by an increase in pipe diameter for solutions with a single pipe configuration. Notice that the gaps are predominantly in the vertical direction thereby representing a sudden change in failure frequency. In contrast, the third gap lies in the horizontal direction and represents a sudden change in the cost of the solution which is caused by a change in pipe configuration from a single pipeline to two pipes in parallel.

For the short pipe length of 1 km, the changes in the costs of the solutions are gradual, unlike the failure frequencies. There is a sudden jump from solutions with a high failure frequency of approximately 190 failures/annum down to solutions with failure frequencies less than one failure/annum. This jump represents a sudden increase in the supply ratio of the solutions. The solutions that are highly unreliable have a supply ratio less than or equal to one, which is accounted for by the single pipe configuration with small diameter. Over half of the Pareto-optimal solutions have a fairly low failure frequency of less than one failure a year under seasonal peak conditions, which can be attributed to the supply ratio values which are greater than 1.8481. Most of these solutions have high supply ratios predominantly due to the parallel pipe configuration of the system.

The Pareto front for the 100 km long pipe length is a stark contrast to the front for the 1 km long pipe. The large vertical gap that was present in the 1 km long Pareto front has been eliminated, and there is a gradual change in the failure frequency of solutions in this 100 km front. The most significant gap in this front is the one in the horizontal direction which represents a large increase in cost. It portrays the change from a single pipe configuration to a parallel pipe configuration.

For each length, the design variables for the solutions that were able to meet the specified design criterion of 0.1 failures/annum were examined. From the previous discussion of the typical system with a 10 km pipe length, it was found that solutions that met the criterion had mixed pipe configurations, either having a single or parallel pipe configuration; and the reservoir sizes ranged from small (4 h) to large (16 h). For a pipe length of 1 km, the acceptable solutions consisted of configurations with two parallel pipes, both with and without interconnections, while most of the reservoir sizes are fairly small, with storage capacities less than 5 h. In contrast, all the solutions that met the design criterion for a 100 km system had two parallel pipes, which resulted in extremely expensive systems. No single pipe configurations were acceptable. However, this was due to the fact that the maximum allowed reservoir capacity explored was 16 h. If the reservoir range was expanded, then single pipeline solutions with reservoir sizes greater than 18 h would dominate and be included in the Pareto front. Even a solution with a single pipeline and reservoir capacity of 32 h of seasonal peak demand (equivalent to 48 h of AADD) would meet the design criterion at a more economical price than a system with two pipes in parallel. From the optimisation run, it was not possible to observe a clear trade-off between interconnections and extra pipes. The main observation is that in many cases single pipelines are more cost effective, except where there is a very short pipe length.

A comparison of one solution for each system length that has a failure frequency close to 0.1 failures/annum was made. These solutions are summarised in Table 7.13. The solution for the 1 km system is quite reliable as one failure in 15 years under seasonal peak conditions is expected, even though it has a small reservoir size of just over 4 h. Its low failure frequency is attributed to its high supply ratio. For the 10 km system, the cheapest solution meeting the desired reliability has quite a low supply ratio value of 1.34, which is compensated for by the reservoir capacity of 14.5 h. When the maximum reservoir capacity allowed was 16 h, the 100 km system required two parallel pipes to achieve the desired reliability, which resulted in an exorbitantly priced system of A\$ 44,150,300. If the range of the reservoir capacities was extended, then the acceptable solution would consist of a single 0.574 m pipeline (SR = 1.76) and an 18 h reservoir. This would be the more sensible and cost effective solution, as it would reduce the cost of the system to A\$ 29,348,258, producing a 34% saving in cost.

Table 7.13 Comparison of single solution for the different pipe lengths

Length (km)	Configuration (Pipes/Interconn.)	Diameter (m)	SR*	RC* (h)	Cost (A\$)	Failure frequency (failures/annum)
1	2/0	0.182	2.0673	4.0781	513710	0.0664
10	1/0	0.322	1.3367	14.5358	2356255	0.1048
100 ^a	2/0	0.428	1.6297	16	44150300	0.0006
100 ^b	1/0	0.574	1.7633	18	29348258	0.0954
100 ^b	1/0	0.574	1.7633	32	29658822	0.0007

* of seasonal peak demand

a – solution if maximum RC is 16 h

b – solution if maximum RC is greater than 16 h

From the comparison of the solutions in Table 7.13, it can be concluded that for systems with a short pipe length, it is cheaper to have parallel pipe configurations and small reservoir sizes. This is because the reservoir cost outweighs that of the pipe system. As the pipe length increases, however, the cost of the pipe forms a greater proportion of the total cost. Therefore, it becomes far more economical to design the system such that it has a single pipeline and a reservoir with a large capacity.

7.4.2 Effect of head

The system head has a wide range of possible values, as it varies depending on the location of the system, the pipe length, the pipe velocity, the pipe diameter and the pipe roughness. To find an estimated typical value, a design velocity of 1 m/s was assumed, with a Hazen-Williams roughness coefficient of 120. For a 0.2 m diameter pipeline that is 10 km long, an expected headloss of 63 m can be calculated. From this approximation, the typical value for the system head in the sensitivity analysis is 60 m. The low and high values are taken to be half and double the typical value respectively, i.e. 30 m and 120 m. Although 30 m may not actually be a low value, it was selected so that the range of values can give an indication of how the optimisation is affected by the system head, as demonstrated in Figure 7.11.

Similar to the pipe length, the system head will also directly influence with pipe diameters are selected for the optimisation procedure. Using the Hazen-Williams formula with a constant flow rate and a constant pipe length, the diameter decreases with an increasing head. Thus, a 120 m head requires smaller diameter pipes than a system with 60 m and 30 m head. Table 7.14 shows the allowed pipe diameters for each head value. Note that the selection of pipe diameters does not vary as much with the system head as it does with the system pipe length.

Table 7.14 Pipe diameter selection for each system head

H = 30 m	H = 60 m	H = 120 m
0.227 m	0.227 m	0.182 m
0.286 m	0.286 m	0.227 m
0.322 m	0.322 m	0.286 m
0.363 m	0.363 m	0.322 m
0.428 m		

Looking at the cheaper end of each curve in Figure 7.11, it is observed that the system with 30 m head has a steeper curve than those with a 60 m and 120 m head. This indicates that systems with smaller heads have a greater increase in failure frequency corresponding to an increase in cost. The curve for the 30 m head system has shifted to the right of the original curve, indicating that more expensive solutions are required because of the larger pipe diameters. The curve for the 120 m head system has shifted to the left, signifying that the solutions are cheaper because of the selection of smaller diameter pipes.

Both of the cheapest solutions for system heads of 30 m and 60 m have the same pipe configuration, diameter, and reservoir size, and therefore exactly the same cost. However, a comparison of their failure frequencies shows that the lower head of 30 m has fewer failures per annum; since for the same pipe diameter, the system with a lower head will have a higher flow rate, and higher supply ratio, therefore it is more reliable.

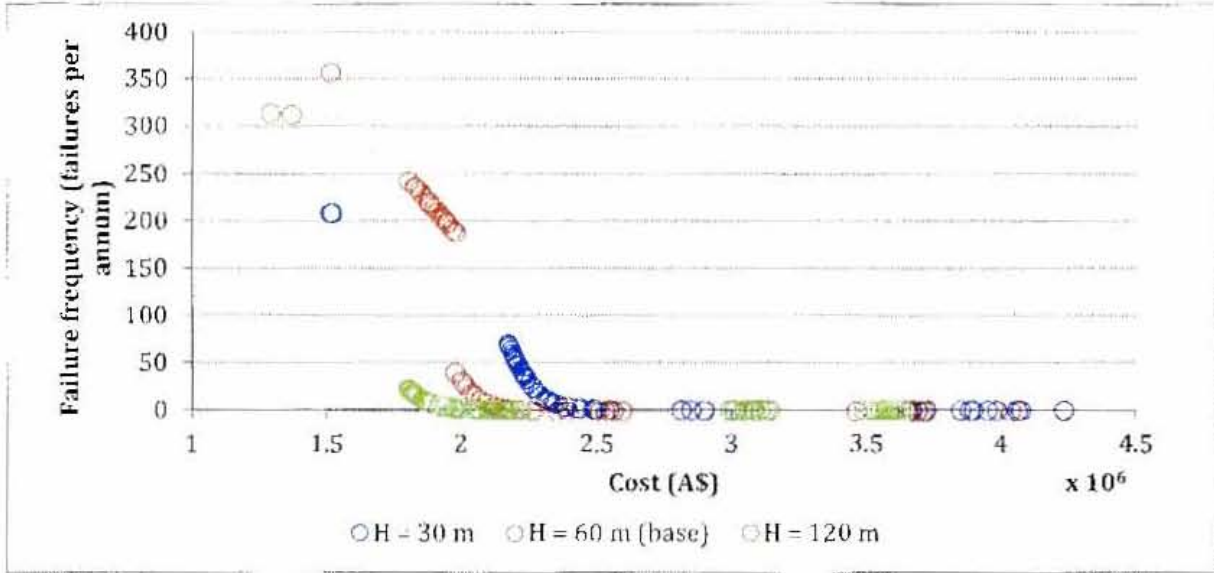


Figure 7.11 Pareto fronts for systems with different heads

For the three systems, the solutions that meet the design criterion of 0.1 failures/annum have a mixed selection of pipe configurations consisting of either one pipe or two parallel pipes with a varying number of interconnections. As before, the cheapest solutions closest to the design criterion for each system head were compared, and are summarised in Table 7.15. Each of these solutions requires only a single pipeline. Again, the diameter decreases with increasing head. With an increasing trend in the supply ratio, a smaller reservoir capacity is required. The solution with a 120 m system head has both a smaller diameter pipe and also a smaller reservoir; therefore it is cheaper than the other two solutions.

Table 7.15 Comparison of single solution for the different system heads

Head (m)	Configuration (Pipes/Interconnections)	Diameter (m)	SR*	RC*	Cost (A\$)	Failure frequency (failures/annum)
30	1/0	0.363	1.26	16	2,593,198	0.1030
60	1/0	0.322	1.3367	14.5358	2,356,255	0.1048
120	1/0	0.286	1.4229	13.7064	2,158,468	0.0829

* of seasonal peak demand

These results show that the system head does not have as great an impact on the optimal solutions as the length of the system. The effect that the system head has on the cost is not as significant as effect of the length as the costs of the solutions are within a similar order of magnitude. The system head mainly has an effect on the cost, rather than the reliability, of the system. The designer can control the system head by choosing to construct either an elevated or a ground level storage reservoir. If an elevated reservoir is chosen, it must be remembered that additional costs will be incurred as pumping of the water into the reservoir is required.

7.4.3 Effect of demand pattern

For a fixed average demand, the supply ratio could be selected as a parameter for sensitivity analysis; however, this is not necessary, as allowing the possibility of different pipe diameters to be selected will inherently include the effect of the supply ratio in the optimisation. Neither is the daily demand pattern varied as it is generally quite flat and there is little variation in demand from one day to the next. Only the hourly demand pattern is investigated in this sensitivity analysis. Different hourly demand patterns were investigated by selecting suitable low and high values for the peak hourly factor.

The low, typical and high values of the peak hourly factor for the demand pattern were chosen to be the same as those in a study by van Zyl *et al.* (2010) on user demand sensitivity parameters. The typical value was based on the measured demands of three small residential towns located in the Moselle area in the east of France, as described by van Zyl *et al.* (2008). It was found in the study of that area that the hourly factors follow a classical residential demand pattern and that the peak hourly factor was 1.49. Van Zyl *et al.* (2010) identified the low peak hourly factor as 1.25 and the high peak hourly factor as 1.75. Figure 7.12 shows the

hourly demand pattern for each peak hourly factor. Note that an average hourly factor of 1 was maintained for all hourly demand patterns.

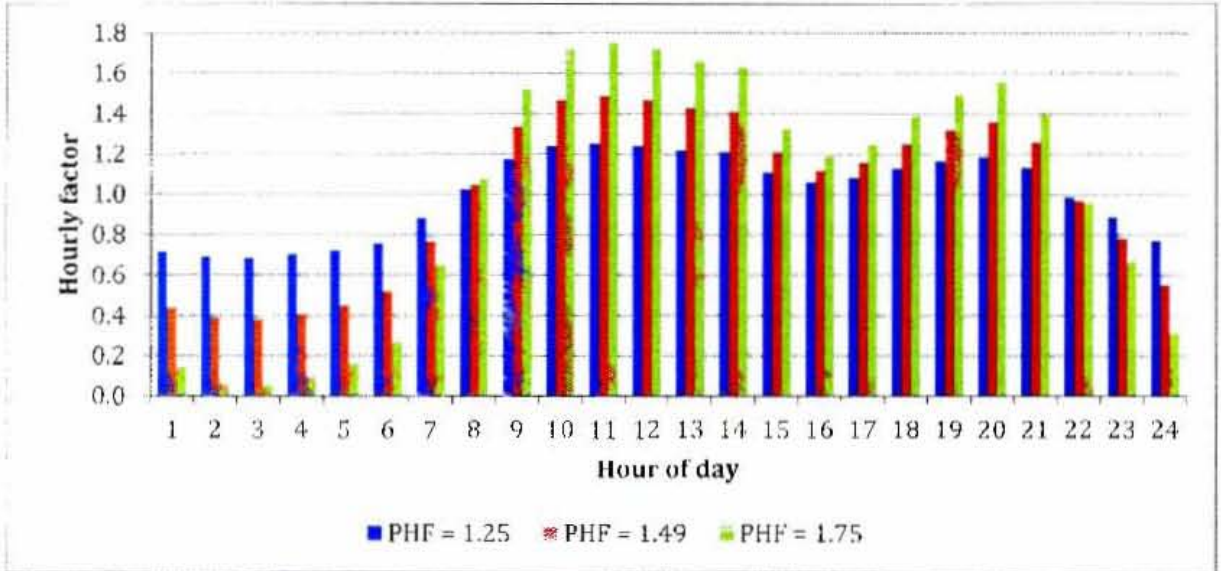


Figure 7.12 Hourly demand patterns

Changing the peak hourly factor has no effect on the system parameters, and therefore no effect on the practical diameters selected for the optimisation. It does, however, have an impact on the results for the pre-run and events run. In Figure 7.13, the effects of changing the peak hourly factor on the optimisation results are shown. For the high hourly peak factor of 1.75, a steeper curve is produced than for the typical and low hourly peak factors. This observation can be explained by the fact that a high peak hourly factor means that there is an elevated demand pattern. If the system experiences greater consumer demand, then the reservoir is more likely to run dry and more reservoir failures occur. Therefore, less reliable systems are obtained at the same cost. A similar explanation is provided for the flatter trade-off curve observed for the low peak hourly factor. The low value of 1.25 generates lower hourly consumer demand that would cause fewer reservoir failures. The system would thus have a lower failure frequency. Varying the peak hourly factor does not have a significant effect on the design values for the optimal solutions (see Table C.3, Appendix C). As a result, the solutions have similar costs.

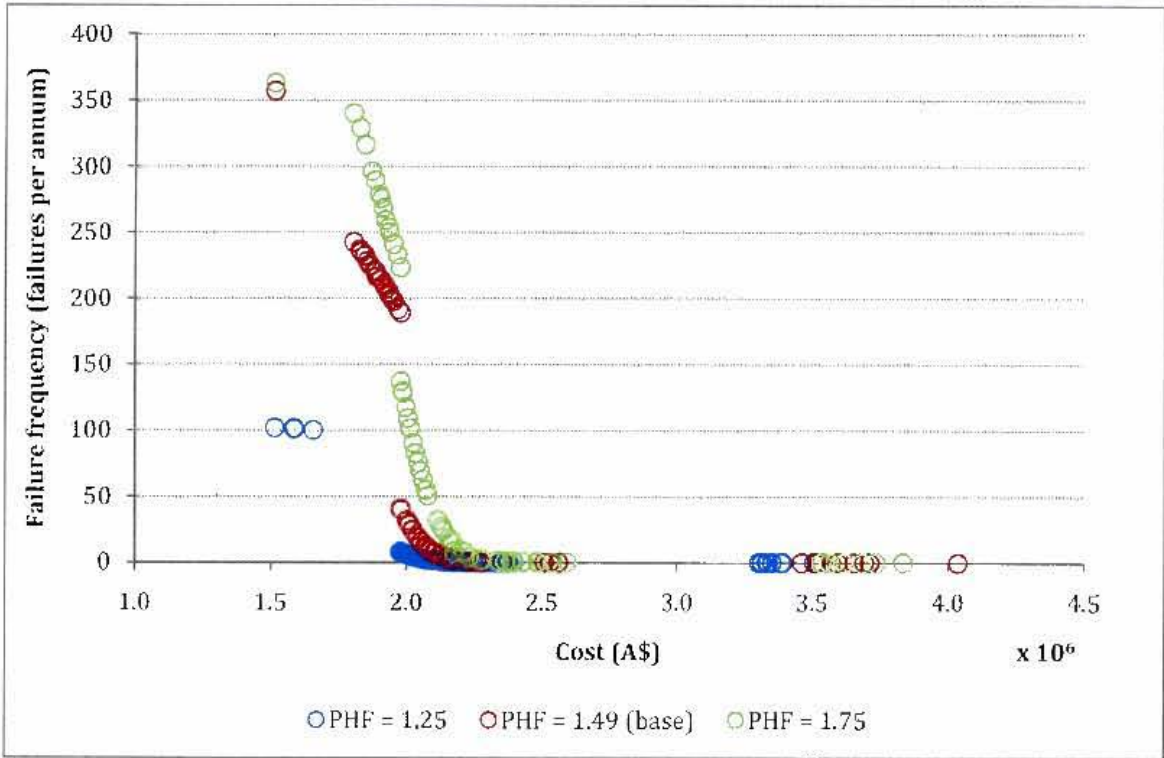


Figure 7.13 Pareto fronts for systems with different peak hourly factors

The main effect of varying the peak hourly factor is on the failure frequency for the system. This observation is illustrated by Table 7.16 below. The solutions in Table 7.16 are positioned where the vertical jump occurs in the trade-off curves. At this point in the curves, where the gap is, the optimisation produced exactly the same design solution for all three peak hourly factors. However, as the peak hourly factor increases, so does the system's failure frequency.

Table 7.16 Comparison the same solution with different peak hourly factors

PHF	Configuration (Pipes/Inter.)	Diameter (m)	SR*	RC* (h)	Cost (A\$)	Failure frequency (failures/annum)
1.25	1/0	0.322	1.3367	4	1977642	7.7699
1.49	1/0	0.322	1.3367	4	1977642	40.4481
1.75	1/0	0.322	1.3367	4	1977642	136.6199

* of seasonal peak demand

Comparing the acceptable solutions that meet the design criterion of 0.1 failures/annum, a trend is observed where as the peak hourly factor increases larger supply ratio values are required so as to maintain an acceptable reliability. A comparison of the cheapest solution with the closest adherence to the design criterion for each of the peak hourly factors is provided in Table 7.17. A comparison of the solutions with peak hourly factors of 1.25 and 1.49 indicates that an elevated hourly demand pattern requires the system to have a larger reservoir size so as to ensure that the reliability of the system is maintained. For an even

higher peak hourly factor of 1.75, an increased supply ratio was required and provided by a larger diameter of 0.363 m.

Table 7.17 Comparison of single solution for the different hourly demand patterns

PHF	Configuration (Pipes/Inter.)	Diameter (m)	SR*	RC* (h)	Cost (A\$)	Failure frequency (failures/annum)
1.25	1/0	0.322	1.3367	12.1437	2287628	0.0952
1.49	1/0	0.322	1.3367	14.5358	2356255	0.1048
1.75	1/0	0.363	1.8321	14.0259	2539344	0.0744

* of seasonal peak demand

The demand pattern is not a factor which can be directly controlled by the designer. As a result, the designer of the bulk supply system needs to be aware of what would happen in the event of a future elevated water demand. Certain measures must be taken to respond to an increased demand. These measures would include either increasing the reservoir capacity of the system, or augmenting the supply inflow to the reservoir by replacing the supply system with a larger diameter pipe or adding an additional pipe in parallel. From this analysis, it has been shown that the optimisation model can be used for scheduling the expansion of a bulk water supply system.

7.4.4 Effect of fire rate

The study by van Zyl *et al.* (2008) assumed that a fire brigade will be required to extinguish a fire once every two months for the low-density residential area analysed. This fire rate of 6 fires/annum is selected as the typical value. To examine the effect of no fires on the system, a low fire rate of zero fires/annum was used. While for a high value, it was assumed that two fires occur every month, i.e. at a rate of 24 fires/annum.

The fire rate has no effect on the selection of pipe diameters used in the optimisation process. It also has no impact on the demands failure frequency because fires are not simulated in the pre-run. Looking at Figure 7.14, it can be seen that the fire rate has very little effect on the Pareto-optimal solutions produced, as the trade-off curves for all three fire rates lie on the same front. All three systems have the same cheapest solution. The most expensive solution for each fire rate is different in terms of cost, but insignificantly so. It was also found that the solutions that met the desired design criterion for each system all have similar pipe configurations and diameters. For these solutions, the fire rate of 24 fires/annum generally required a larger reservoir capacity (8 to 16 h) compared to the fire rates of zero and 6 fires/annum (4 to 16 h) in order to provide the desired system reliability.

One solution that has a failure frequency close to 0.1 failures/annum for each fire rate is shown in Table 7.18. For the different fire rates, each of the solutions has the same configuration and diameter, giving the same supply ratio, and very similar reservoir capacities. Consequently, they have a similar reliability and cost. This result confirms the findings from a similar study by Vlok (2010), stating that a larger reservoir size is required

for a higher fire rate, however the increase is insubstantial. These results indicate that fires have very little effect on the failure behaviour of the system. A possible reason for this is because both the fire flow and fire duration are small. A fire event would thus have no significant effect unless it coincided with a high consumer demand or pipe failure event.

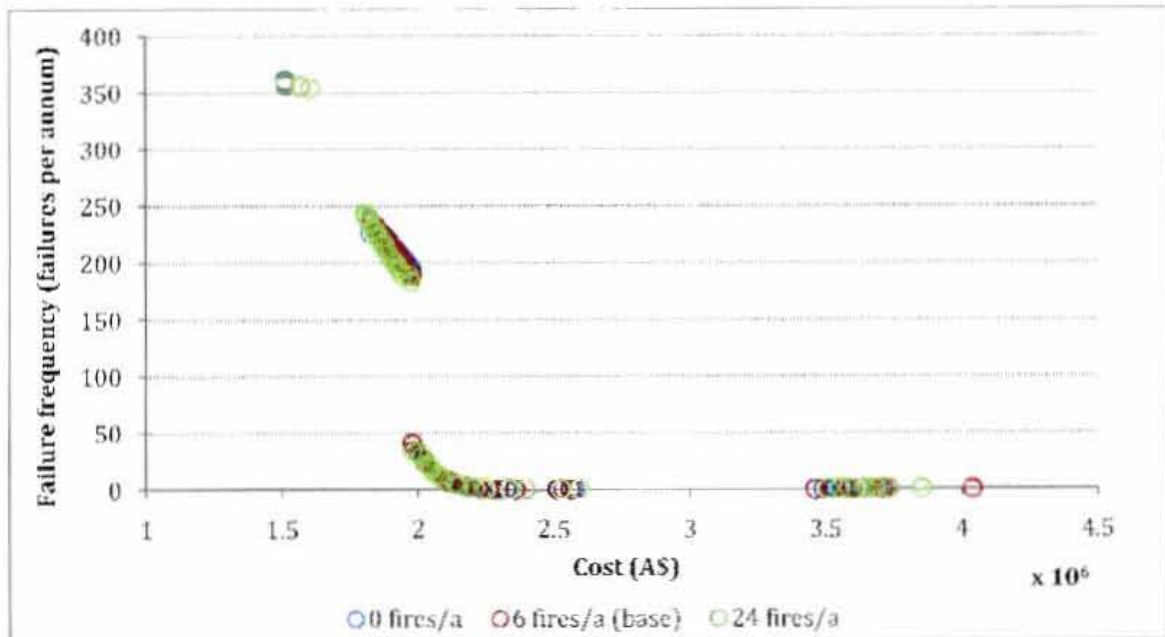


Figure 7.14 Pareto fronts for different fire rates

Table 7.18 Comparison of single solution for different fire rates

Fire rate (fires/a)	Configuration (Pipes/Inter.)	Diameter (m)	SR*	RC* (h)	Cost (A\$)	Failure frequency (failures/annum)
0	1/0	0.322	1.3367	14.4525	2353949	0.1056
6	1/0	0.322	1.3367	14.5358	2356255	0.1048
24	1/0	0.322	1.3367	14	2341327	0.1079

* of seasonal peak demand

7.4.5 Effect of pipe failure rate

The number of failures that the supply pipe will experience has a large impact on the reservoir reliability (van Zyl *et al.*, 2008). It is therefore important that the effect of different pipe failure rates is examined.

The most commonly reported factors included in pipe failure prediction models are pipe material, pipe diameter and pipe age (Nel, 2009). The impact of these factors on the pipe failure rate and these factors in the pipe failure rate and the system reliability was not investigated, since the aim of this section is not to predict the pipe failure rate, but rather to get an idea of representative values. It was assumed that different pipe diameters have the same failure frequency, as do different pipe materials.

The pipe failure rate is clearly linked to the pipe length as it is usually measured in terms of number of failures per kilometre per annum. An extract of a table from Nel (2009) shows the average value failure rate for different pipeline materials (Table 7.19). Taking an average of these values gives a typical pipe failure rate of 0.2 failures/km/annum. This value is the same as the pipe failure rate of 2 failures/annum in the study by van Zyl *et al.* (2008), which is presumably the failure rate for a 10 km supply pipe. The low pipe failure rate is set at half of the typical value, i.e. 0.1 failures/km/annum, while the high value is 0.5 failures/km/annum.

Once again, note that the pipe failure rate is a function of the length of the pipe. Hence, if a single pipeline has a failure rate of 0.2 failures/km/annum, then the pipe failure rate for a bulk system with two pipes in parallel has a failure rate of 0.4 failures/km/annum. A system with multiple pipes in parallel will experience more failures.

Table 7.19 Average failure rates for different pipeline materials (Nel, 2009)

Pipe material	Average value (no./km/annum)
Steel	0.192
Cast iron	0.353
Ductile iron	0.071
Cement and concrete	0.322
Plastic	0.089

As with the fire rate, the pipe failure rate has no influence on the selection of pipe diameters used in the optimisation; and so for each of the failure rates investigated, the same pipe sizes with diameters 0.227 m, 0.286 m, 0.322 m, and 0.363 m are permitted. The pipe failure rate also does not affect the pre-run as no pipe failures are simulated. Pipe failures only affect the events run.

Figure 7.15 shows the Pareto-optimal solutions in trade-off curves for each of the different pipe failure rates. The pipe failure rate appears to have little effect on the Pareto front obtained, as the trade-off curves for the high and low pipe failure rates do not shift away from the curve for the typical value of 0.2 failures/km/annum, and there is a similar spread of solutions. All three pipe failure rates produce acceptable solutions that adhere to the design criterion of 0.1 failures/annum. These solutions either have single pipe configurations or configurations with two pipes in parallel and a varying number of interconnections. The solutions for each failure rate also have reservoir sizes which vary within the range selected. There appears to be no distinguishable trend.

Table 7.20 shows one solution for each of the different pipe failure rates, whose failure frequency gives the closest value to one failure in ten years under seasonal peak conditions. The pipe failure rates of 0.1 and 0.2 failures/km/annum have solutions with the same pipe diameter and configuration, while the reservoir capacity for the higher failure rate is larger than the lower failure rate. This corresponds with Vlok's (2010) finding that a larger reservoir

size is required for a higher pipe failure rate. For a higher failure rate of 0.5 failures/km/annum, the reservoir experiences more failures caused by the greater frequency of pipe breakages. A larger reservoir size could be used, but in this case, a higher supply ratio is needed to counteract this break in supply, which means that a larger diameter is required to supply an increased flow into the system. The larger pipe diameter is responsible for the increased cost of the system.

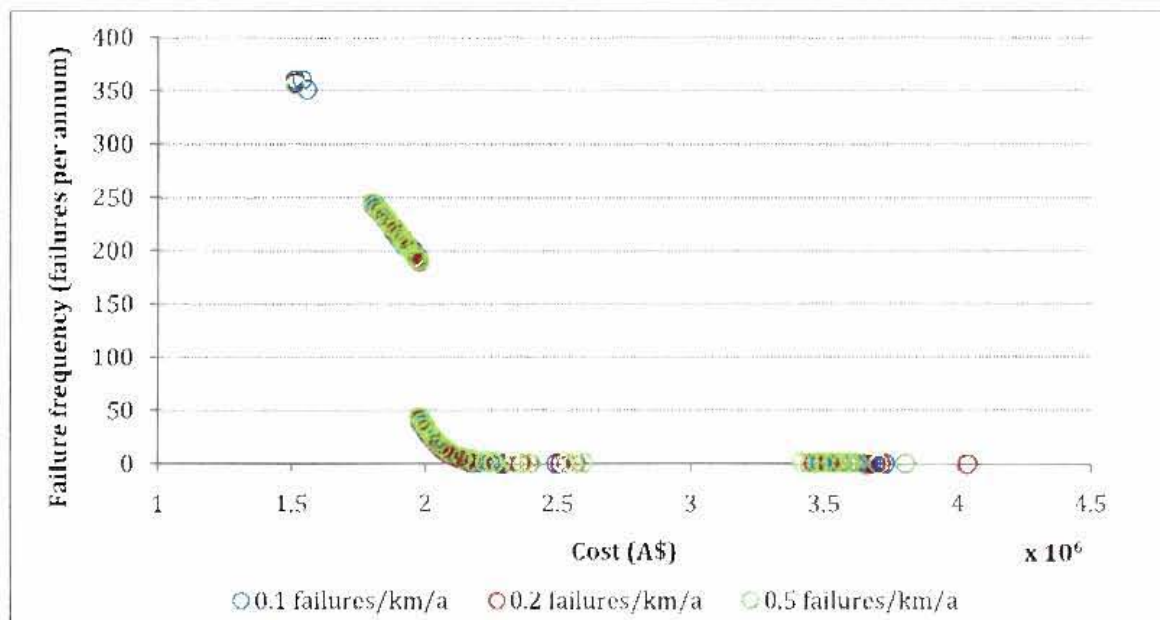


Figure 7.15 Pareto fronts for different pipe failure rates

Table 7.20 Comparison of single solution for different pipe failure rates

Failure rate [no./km/a]	Configuration (Pipes/Inter.)	Diameter [m]	SR*	RC*	Cost (A\$)	Failure frequency [failures/annum]
0.1	1/0	0.322	1.3367	13.6364	2331059	0.1068
0.2	1/0	0.322	1.3367	14.5358	2356255	0.1048
0.5	1/0	0.363	1.8321	13.509	2524720	0.1026

* of seasonal peak demand

The analysis of the results obtained shows that the pipe failure rate, like the fire rate, does not greatly affect the system’s failure behaviour. A similar Pareto-optimal set of solutions is obtained for each pipe failure rate. An explanation for this observation could be that when a pipe failure occurs, it lasts for a fairly short period of time (average failure duration of 4.5 hours), so it has very little impact on the reserve of water already stored in the reservoir. Taking this explanation into account, it was decided that a sensitivity analysis on the average failure duration for the pipe failure model should be performed.

7.4.6 Effect of failure duration

The study by van Zyl *et al.* (2008) estimated the logarithmic value for the average pipe failure duration as 1.49, which corresponds to an approximate pipe outage time of 4.5 h. This value was taken as the typical average failure duration. Low and high average failure durations of 3 h and 9 h were selected, corresponding to logarithmic values of 1.1 and 2.2 respectively.

As mentioned in the sensitivity analysis for the pipe failure frequency, no pipe failures are modelled in the pre-run, and therefore the demands failure frequency remains the same for the three average failure durations investigated. Only the events run is affected by the pipe failure duration. The selection of available pipe diameters for the different failure durations is also the same: 0.227 m, 0.286 m, 0.322 m and 0.363 m.

Like the graphs for the fire rate and the pipe failure rate, the three average failure durations produce very similar Pareto fronts, as shown in Figure 7.16. The only noticeable difference between the trade-off curves is the slight shifting of the individual solutions along the fronts. As it is difficult to see the solutions clearly, a close-up of the acceptable solutions that meet the design criterion of 0.1 failures/annum is provided in Figure 7.17. The low and typical failure durations provide very similar results in this region, where the low failure duration of 3 h has solutions which are shifted slightly to the left of the solutions for the typical failure duration of 4.5 h. This is expected as a more expensive solution is required to achieve the same system reliability when the average failure duration is longer. There is a more pronounced difference shift of solutions to the right for the high failure duration of 9 h, which is double the typical value of 4.5 h. These solutions represent systems that have parallel pipe configurations, which explain the increased cost of the systems.

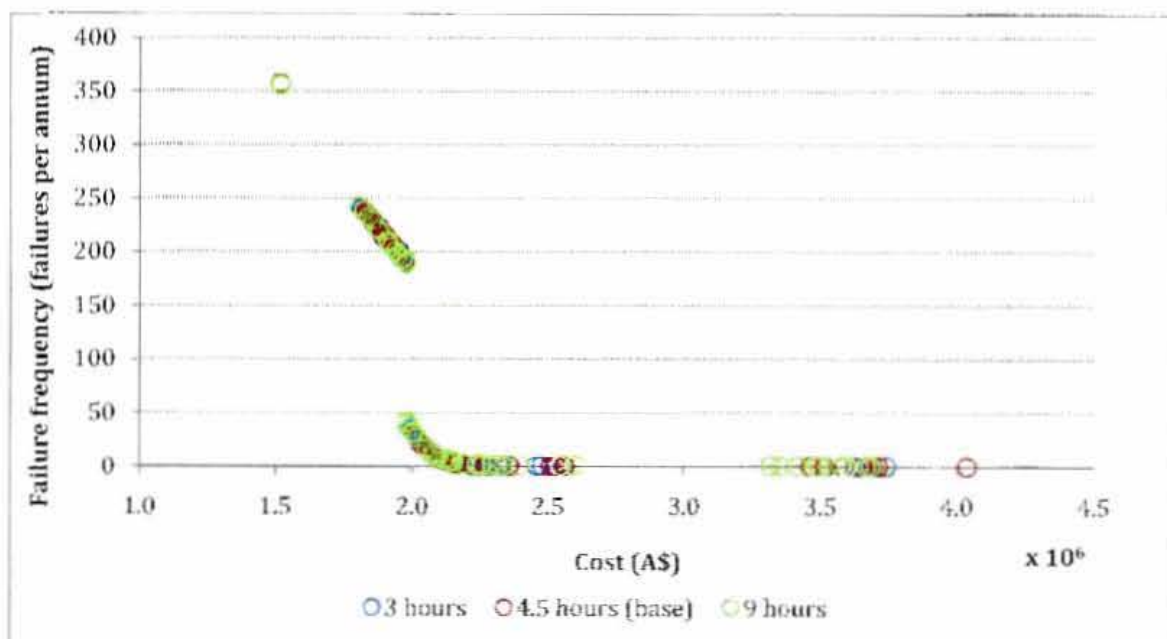


Figure 7.16 Pareto fronts for different pipe failure durations

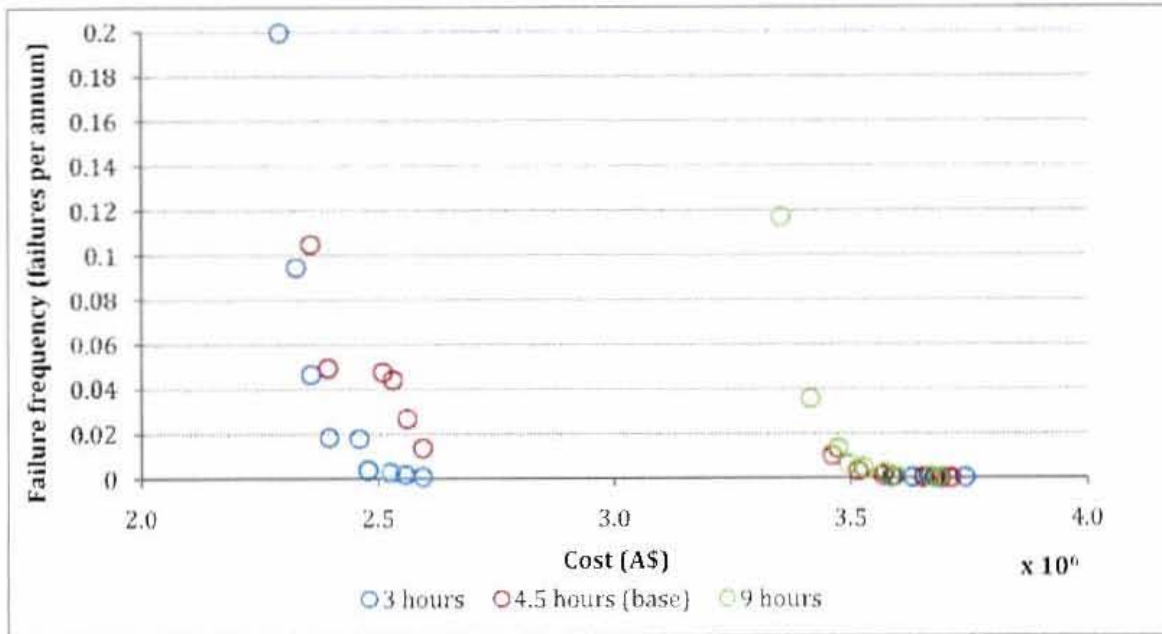


Figure 7.17 Solutions meeting the design criterion for the different pipe failure durations

The individual solutions for each average pipe failure duration that satisfy the design criterion of 0.1 failures/annum at minimal cost are shown in Table 7.21. Comparing the solutions for the low and the typical values, which are not very different, the typical system makes up for the larger failure duration by increasing the system's reservoir capacity, thereby causing a slight increase in cost. Examining the high average failure duration shows that its solution requires a parallel pipe system to be implemented so as to be able to meet the design criterion. Consequently, it has a much higher cost of A\$ 3,527,742, even though its reservoir capacity is 4.69 h. If the range of allowed reservoir sizes was increased such that the solution has a reservoir size of 21.5 h and a single 0.363 m diameter pipe, then a significant cost saving of A\$ 622,663 could be achieved (18.6%).

Table 7.21 Comparison of single solution for different pipe failure durations

Fail duration (h)	Configuration (Pipes/Inter.)	Diameter (m)	SR*	RC* (h)	Cost (A\$)	Failure frequency (failures/annum)
3	1/0	0.322	1.3367	13.4723	2326386	0.0945
4.5	1/0	0.322	1.3367	14.5358	2356255	0.1048
9 ^a	2/2	0.286	1.9573	4.6945	3352742	0.1168
9 ^b	1/0	0.363	1.8321	21.5	2730079	0.0980

* of seasonal peak demand

a - original solution in Pareto front

b - additional solution investigated

The investigation of the average failure duration of a pipe demonstrates that it has a greater effect on the optimisation results than the pipe failure frequency. This is illustrated by the cheapest solutions satisfying the desired reliability for each of the sensitivity parameters. For example, the reservoir capacity for the high average failure duration value is 21.5 h of storage

(Table 7.21), which is 8 h larger than the 13.5 h storage capacity for the high pipe failure frequency value (Table 7.20). This observation indicates that if the supply pipe system experiences short but frequent pipe outages, then the reservoir may not actually run dry because it may have a large enough reserve of water stored. In contrast, if the pipe failures are long but less frequent, the reservoir is able to store only a limited amount of water, which may not be able to last for the duration of the pipe outage. The storage reservoir is therefore more likely to fail when the average failure duration is longer than when the pipe failures are more frequent.

7.4.7 Comparison of sensitivity results

The above sensitivity analysis identified six parameters that have impacts on the optimal solutions selected:

- Length of the feeder pipe system;
- System head;
- Peak hourly factor;
- Fire occurrence rate;
- Pipe failure rate;
- Average pipe failure duration.

In this study, a design criterion of one failure in ten years under seasonal peak conditions was used. For each of the sensitivity parameters analysed, the optimal solution that provided the closest failure frequency to the design criterion at minimum cost was identified. These solutions are summarised in Table 7.22. It should be noted that each failure frequency is not exactly 0.1 failures/annum. This is because of the discrete nature of the supply ratio, and the inherent variability of the stochastic results. The impact of these factors on the system cost is shown in Figure 7.18.

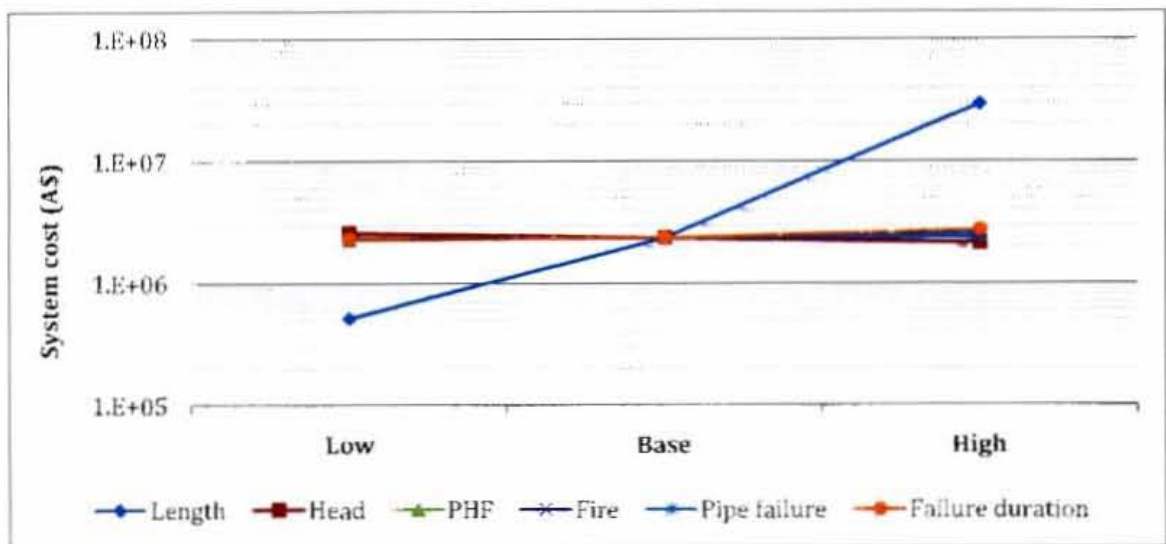


Figure 7.18 Impact of sensitivity parameters on system cost

Table 7.22 Near optimal solutions satisfying design criterion for each sensitivity parameter

Description	Configuration	Diameter (m)	SR	RC (h)	Cost (A\$)	Failure frequency (failures/annum) #	Comment
Base	1/0	0.322	1.34	14.54	2,356,255	0.1048	Typical system and stochastic parameters
Length - low	2/0	0.182	2.07	4.08	513,710	0.0664	L = 1 km
Length - high	1/0	0.574	1.76	18	29,347,258	0.0954	L = 100 km*
Head - low	1/0	0.363	1.26	16	2,593,198	0.1030	H = 30 m
Head - high	1/0	0.286	1.42	13.71	2,158,468	0.0829	H = 120 m
Demand - low	1/0	0.322	1.34	12.14	2,287,628	0.0952	PHF = 1.25
Demand - high	1/0	0.363	1.83	14.03	2,539,344	0.0744	PHF = 1.75
Fire rate - low	1/0	0.322	1.34	14.45	2,353,949	0.1056	Fire rate = 0 fires/annum
Fire rate - high	1/0	0.322	1.34	14	2,395,910	0.1079	Fire rate = 24 fires/annum
Pipe failure rate - low	1/0	0.322	1.34	13.64	2,331,059	0.1068	Failure rate = 0.1 failures/km/annum
Pipe failure rate - high	1/0	0.363	1.83	13.51	2,524,720	0.1026	Failure rate = 0.5 failures/km/annum
Failure duration - low	1/0	0.322	1.34	13.47	2,326,386	0.0945	Ave. failure duration = 3 h
Failure duration - high	1/0	0.363	1.83	21.5	2,730,079	0.0980	Ave. failure duration = 9 h*
Red Book	1/0	0.286	0.9787	32	2,583,671	140.0671	SR \approx 1.5 x AADD, RC = 48 h AADD

* allowed RC greater than 16 h

the solutions listed are those with reliability closest to one failure in 10 years

It is clear that the length is the dominant factor and has a large impact on the Pareto-optimal solutions produced. The length directly affects the range of practical pipe diameters selected for use in the optimisation, which in turn affects the pre-run and events run. A longer pipe length results in a higher system cost.

The other factors have less of an impact on the optimal solutions than the pipe length. A comparison of these factors is shown in Figure 7.19. It can be seen that the effect that they have on the system cost is in the same order of magnitude. Of these factors, the most significant ones are the head and the average pipe failure duration, while the fire rate, pipe failure rate and peak hourly factor have insignificant impacts. The system head, like the length, also affects the pipe diameters selected for optimisation but the range of diameters is not as drastic. From the analysis of the pipe failure duration, it was found that it is a more important factor than the pipe failure rate. Long, infrequent pipe outages cause more reservoir failures than short, frequent outages; thus, more reliable systems are required to maintain a desired system performance, at the price of a higher system cost.

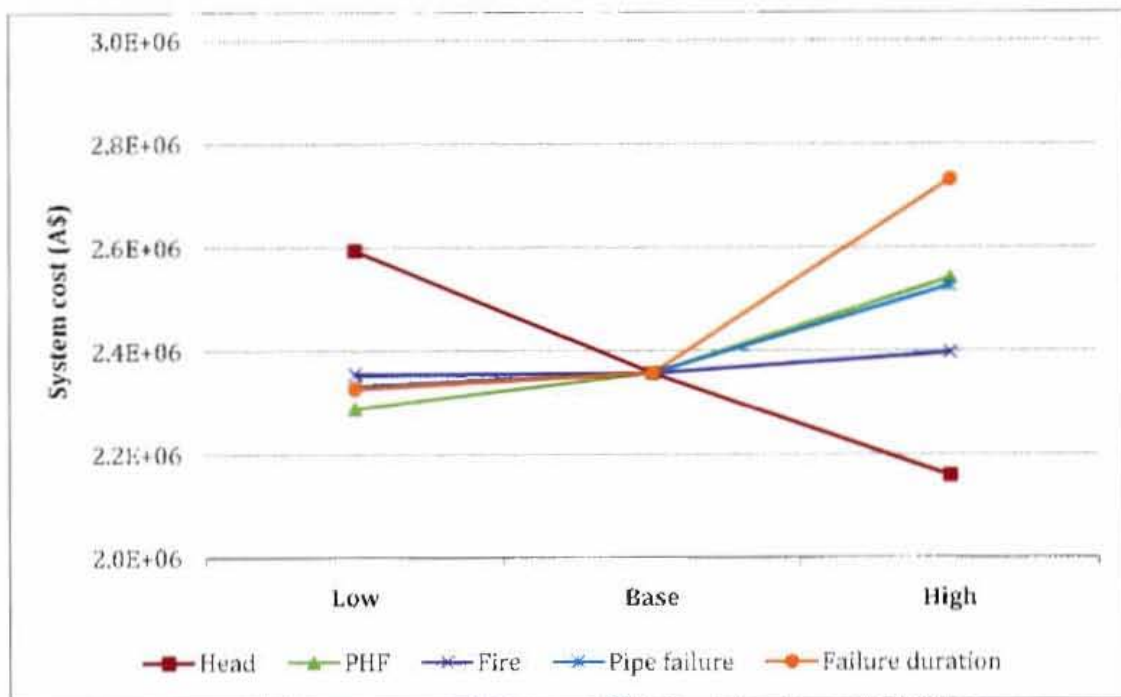


Figure 7.19 Impact of sensitivity parameters on system cost (excluding length)

Figure 7.20 provides a qualitative comparison of the optimal solutions summarised in Table 7.22 by plotting the combination of supply ratio and reservoir capacity for each optimal solution. The single pipe solution with supply ratio of approximately 1.5 times AADD and reservoir size of 48 h of AADD, as designed according to the Red Book, is marked on the figure. It can be seen that the optimised solutions that meet the design criterion all have a larger supply ratio and smaller reservoir capacity than the Red Book solution. This comparison once again emphasises that it is preferable to have a larger supply ratio than reservoir size.

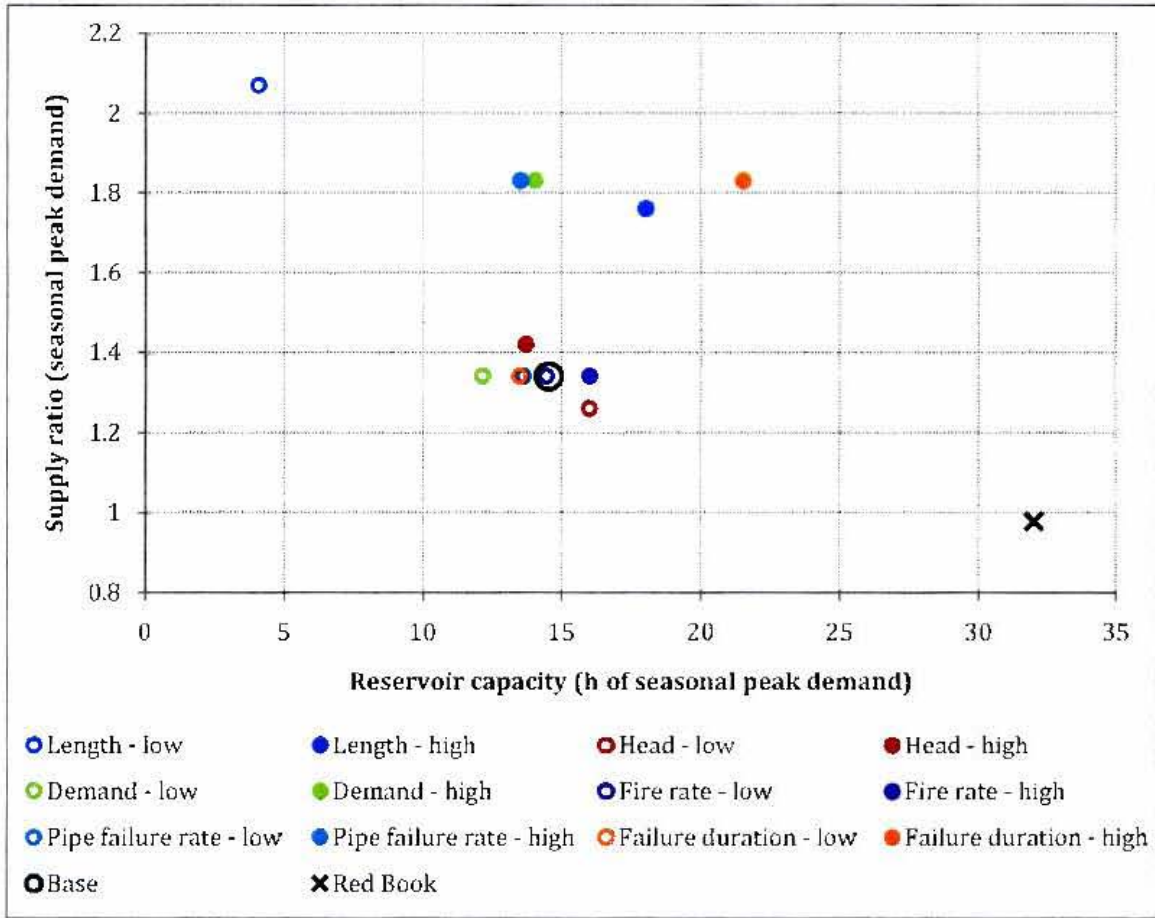


Figure 7.20 Trade-off of supply ratio against reservoir capacity for optimal solutions

The solution with a short pipe length consists of a small reservoir size and a very high supply ratio because it has a parallel pipe configuration. This solution reiterates the finding that a parallel pipe system with small reservoir size is preferred for short pipe lengths, whereas a single pipeline with large reservoir size is more economical for a long pipe length.

The system head produces a similar pattern in the trade-off between supply ratio and reservoir capacity as the length, although the differences between the low and high values are not as pronounced.

A change in the peak hourly factor, from a low to high value, results in a large increase in supply ratio from 1.34 to 1.83 times seasonal peak demand and minor increase in reservoir size from 12.14 h to 14.03 h of seasonal peak demand. This illustrates that a high peak hourly factor places more stress on the reliability of a system therefore a larger supply ratio and reservoir size is required to meet the desired design criterion.

The fire rate has minimal impact on the optimal solutions. The solutions for the low and high values have the same configuration. There is only a minor variation in the reservoir size. In contrast, different pipe failure rates result in solutions which have a similar reservoir capacity,

but have an increase in supply ratio for a high failure rate. The high number of failures caused by the high pipe failure rate is counteracted by increasing the supply ratio.

It was observed that the optimal solutions chosen are sensitive to the duration of the pipe failures. As the average failure duration increases, a larger reservoir capacity is required to store a greater amount of water when there is no flow into the reservoir, and a larger supply ratio is required to fill up the reservoir more quickly.

Figure 7.21 illustrates the proportions contributed by the individual components to the total cost for each optimal solution. Changing the system and stochastic parameters does not have a large impact on the relative costs. The pipe system consistently makes up approximately 70% of the total cost and the remainder of the cost is attributed to the storage reservoir. The only exception is for the two solutions where the pipe length is varied. For a short pipe length, the reservoir cost (62% of the total) outweighs the pipe cost (38%); while the cost of the pipe system forms nearly all of the total system cost (97%). This indicates that pipe length is the most significant factor when considering the cost of the system.

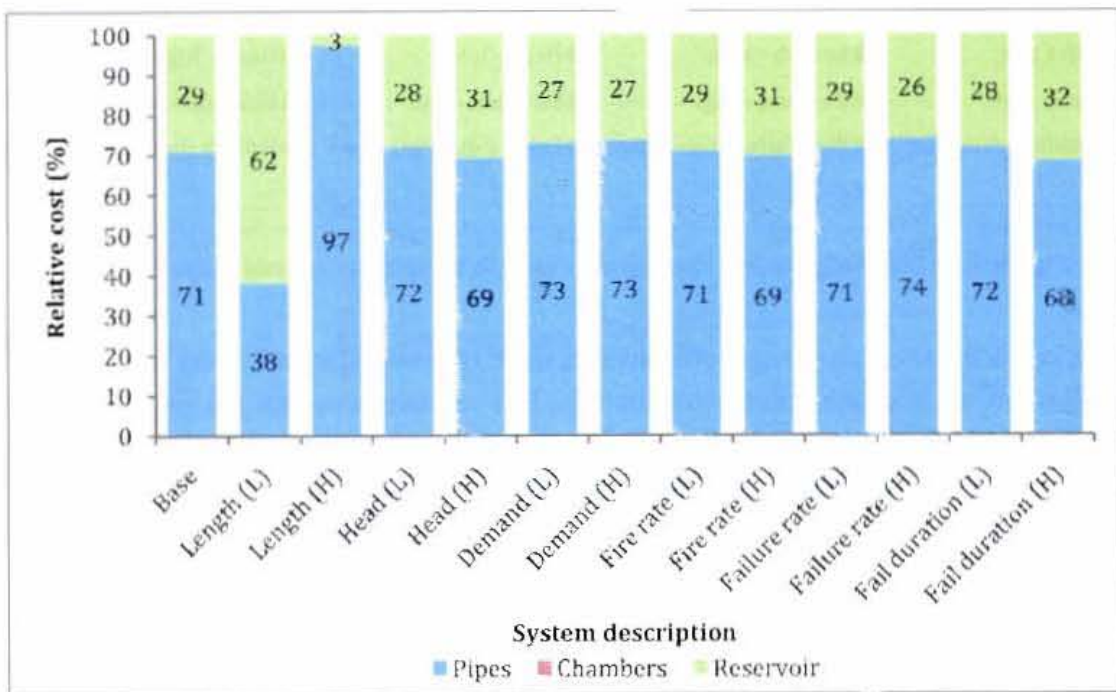


Figure 7.21 Relative costs of individual components for each optimal solution

A similar study on risk-based cost analysis of bulk water supply systems was performed by Vlok (2010). Where possible, the results obtained in this section were compared to the findings from Vlok (2010). This proved to be slightly difficult as the supply ratio was kept constant in that study, while in this model, both the supply ratio and reservoir capacity could be varied in order to produce a reliable yet cost-effective solution. For example, Vlok (2010) found that the capital cost of a system with a single pipe configuration was always the least expensive. However, in this study it was found that this is not true for short pipe lengths,

where a cheaper system can be obtained by selecting a parallel pipe configuration. This difference in the results can be attributed to the fact that the diameter of the pipe can be varied in the optimisation model, so that a larger supply ratio can be provided, leading to a more reliable solution at a cheaper cost.

7.4.8 Computational time savings

Simulation times for the design optimisation model will vary according to the system and stochastic parameters selected. For the thirteen systems analysed in the sensitivity analysis, the approximate simulation times for the optimisation of 50 solutions over 50 generations ranged from 20 h to 100 h.

The time taken by the model to complete the optimisation and produce a set of results depends on a number of factors, such as the reliability of the reservoir and the number of event generated. The more failures that are generated for a particular system, the faster the stochastic simulation will converge. If only a few failures are generated, then the stochastic model might need to run for the full simulation duration specified by the designer. On top of that if the system has a higher fire / pipe failure rate or longer pipe failure duration, more or longer critical periods must be fully simulated in the events simulation, resulting in a longer simulation time.

The optimisation of the typical system took an estimated time of 40.5 hours using the compression heuristic. If the optimisation was performed using the full Monte Carlo method instead, then it would take approximately 262 hours, as estimated using the trend shown in Figure 7.5 (Section 7.2). This shows the compression heuristic is capable of providing an 85% savings in computation time for the optimisation procedure. All the other systems analysed will be likely to produce computational savings in the same order of magnitude.

8 Conclusions

This project dealt with the design optimisation of bulk water supply systems using a well-known multi-objective genetic algorithm. The optimisation was aimed at minimising the cost and minimising the failure frequency (maximising the reliability) of the bulk systems, by selecting the optimal combinations of supply pipe configuration, diameter and reservoir size. In this chapter, a summary of the dissertation contents and methodology used in the research is provided. This chapter presents the conclusions drawn from the analysis and discussion of the results, the main contributions of this research, as well as proposals for future work.

8.1 Summary of work

A fair amount of work has been done in the field of water distribution system design optimisation. Through a literature survey it was found that the amount of research on optimisation of bulk water supply systems pales in comparison. Recognising the niche for work in this area, this study aimed to develop a model that can find a range of Pareto-optimal solutions for a bulk system, which has as its design variables: pipe configuration, pipe diameter and reservoir size.

Cost and reliability were chosen as the objectives for the multi-objective optimisation problem. Existing unit cost models that were simple and easy to incorporate into the optimisation were used to evaluate the system cost. It was a slightly more difficult task to measure the reliability of the system. From the literature, it was established that while there is no universally accepted definition for the reliability of a water distribution system, the reliability (or lack thereof) of a storage reservoir can be defined in terms of its failure. In this study, the reliability of the storage reservoir is used synonymously to measure the reliability of the bulk water supply system. The reliability of the reservoir is measured specifically in terms of its failure frequency, which is evaluated using a stochastic analysis model based on the work by van Zyl *et al.* (2008).

Stochastic analysis, although able to provide a realistic measure of reliability and account for the randomness of variables, is computationally intensive. It was thus necessary to use a compression heuristic (Chang & van Zyl, 2010) to increase the computational speed of the simulation. The aim of the compression heuristic is to produce statistically similar results as the full stochastic model using Monte Carlo sampling, but to do so over a shorter time span. The premise leading to the development of the compression heuristic was that most of the time steps in the full stochastic model simulate only consumer demand, while a minor proportion of the time steps experience a fire or pipe failure or both. It was deduced that time could be saved by fully simulating only these events in an events run. However, it would be foolish to ignore the time steps between these critical periods, thus the intermediate periods were simulated in a demands-only metaheuristic, referred to as the pre-run.

To test the practicality of using the compression heuristic, its performance was compared to the full stochastic model. Once satisfied with compression heuristic's ability to produce accurate and reproducible results, in a shorter period of time, the reliability-based optimisation model was applied to a typical test system and its results were analysed with reference to deterministic design guidelines. Following this application, a sensitivity analysis was performed to determine which parameters have the greatest effect on the design solutions produced. The results from the model application to the test system and the sensitivity analysis were discussed. Conclusions were made from these discussions and these are presented in the following section.

8.2 Main findings

Several conclusions can be drawn from the results obtained in this study. This section presents the conclusions based on the development of the compression heuristic, followed by a summary of the main findings from the application of the optimisation model to a typical system as well as the sensitivity analysis performed.

8.2.1 Compression heuristic

From a comparison of the compression heuristic with the full stochastic model, it was concluded that the compression heuristic is capable of accurately and reliably reproducing the failure results obtained by the full stochastic simulation. For the systems tested with ten different random seeds, the failure frequency results for the compression heuristic consistently fell within 5% of the average of the failure frequency produced by the full Monte Carlo method. In most cases, the compression heuristic provided more conservative estimates of the failure frequency with a lower variability.

The compression heuristic was implemented with the purpose of improving the computational speed of the stochastic analysis. It was found that this was not possible for a single simulation, because two runs (pre-run and events run) instead of one run (full stochastic simulation) are required to achieve convergence. The convergence criterion depends on the number of failures generated in the simulation. The compression heuristic takes longer to converge because fewer failures are generated in each run when the fire and pipe failure events are simulated separately in the events run.

For a larger number of systems, the true potential of the compression heuristic to reduce the simulation time was displayed. The required simulation time of the pre-run remains constant, even as the number of reliability evaluations increases. Thus, the greater the number of reliability evaluations performed by the compression heuristic, the greater the potential time savings. As an example, for a pre-run simulation with grid values $SR = \{1.0; 1.2; 1.4; 1.6; 1.8; 2.0\}$ and $RC = \{4; 8; 12; 16; 20; 24\}$ h, the compression heuristic has a similar simulation time as the full stochastic method after approximately 30 simulations. For more than 30 simulations, the compression heuristic becomes noticeably faster than the full stochastic

method. For the typical optimisation problem in this study, 2550 reliability evaluations are required. By using the compression heuristic, the total simulation time required for these evaluations can be reduced to 85% of the time for the full stochastic model.

8.2.2 Optimisation model

The optimisation model developed in this study successfully generated a set of Pareto-optimal solutions for a number of systems tested. The solution set ranges from extremely unreliable, but cheap systems to highly reliable, but expensive systems. The model allows the engineer to choose the most economical design for a bulk water supply system while conforming to a pre-selected level of reliability. Of course the solution chosen is not guaranteed to be the best one, as all optimisation models are inevitably limited by the uncertainties of variables and assumptions made during the model development. Nevertheless, the model is able to narrow down the innumerable selection of possible design solutions to just a handful. These solutions are then to be used in conjunction with the experience and judgment of the engineer to design the final system.

The application of the optimisation model to a typical system illustrated that it is capable of finding cheaper solutions than conservative deterministic design guidelines, while satisfying the design criterion of one failure in ten years under seasonal peak conditions. South African design guidelines specify that the bulk system should have a reservoir capacity of 48 h of AADD and a supply mains capacity of at least 1.5 times AADD (CSIR, 2000). From the optimisation, the cheapest system that met the design criterion had a reservoir capacity of 21.8 h AADD, and a supply capacity of 2 times AADD. A comparison of the costs of both systems showed that the cost could be reduced by at least 9%. From this comparison, it was concluded that it is more economical to have a larger capacity for the supply system rather than a larger reservoir size for the system tested. Further evaluation of the reliability of a guidelines-based design solution showed that it was in fact unable to meet the specified design criterion, and therefore a review of the design guidelines specified in the Red Book was suggested.

The sensitivity analysis involved an investigation of the effects of varying the pipe length, the system head, the peak hourly factor, the fire and pipe failure rates, and the average pipe failure duration. It was found that the length of the supply pipe system had the greatest impact on the design optimisation of the system, followed by the system head and pipe failure duration. The pipe length affects the reliability of the system because a longer pipe experiences more failures per annum; in addition to the cost of the system since the pipe cost is a function of its length. It was interesting to note that for a system with a short pipe length, two parallel pipes are preferred over a single pipeline; while for longer pipes, a single pipeline is always cheaper.

The peak hourly factor dictates the impact of the hourly consumer demand on the model. A higher peak factor results in an elevated demand pattern which increases the number of

failures experienced by a system, compared to when a lower peak factor is used. The fire and pipe failure rates have comparatively little effect on the trade-off curves compared to the peak hourly factor, length, and head of the system. The relative lack of impact of the increased fire rate was attributed to the fact that fire volumes are small compared to consumer demand volumes. The pipe failure rate also seemed to have insignificant effect compared to the average pipe failure duration. It was found that long, infrequent pipe outages tended to cause more storage reservoir failures than short, frequent pipe outages.

In conclusion, the optimisation model is useful for identifying the impacts of changing certain system and stochastic parameters, in addition to narrowing the search region for a reliable and cost-effective bulk water supply system design.

8.3 Contributions to research

An effort is made to define the contributions made to the field of bulk water supply system design and optimisation through the development of the reliability-based optimisation methodology developed in this study. This section first describes some of the research contributions made by the refinement of the stochastic analysis model. It then presents the contributions of the optimisation model.

Practical stopping rules were formulated for the termination of a stochastic model, specifically the compression heuristic. These stopping rules included a convergence criterion for the pre-run and events run, which ensure that the simulation stops when there is a 95% confidence that the failure result falls within 5% of its true value. The benefit of using this convergence criterion is that for low reliability reservoirs, no failures are simulated more than necessary. In addition, highly reliable reservoirs are not simulated for an indefinite period of time if not enough reservoir failures are generated.

The original compression heuristic (Chang & van Zyl, 2010) considered only one pipe configuration with a single pipeline. It was further developed in this study so that it was expanded to include reliability analyses of different pipe configurations, such as multiple pipes in parallel, both with and without interconnections.

Traditionally engineers use engineering judgment and follow deterministic guidelines to design bulk systems. These deterministic guidelines often contribute to the design of conservative and oversized systems. Although numerous studies on multi-objective optimisation have been carried out on water distribution systems, there is a lack of research in the design optimisation of bulk water supply systems. The optimisation model developed in this study can be used to assist engineers with selecting the design variables of a bulk water supply system such that they are optimised based on both cost and reliability considerations. This method allows engineers to weigh up different design options, while taking into account the cost and failure frequency of the system.

The reliability-based optimisation model also provides engineers with a method of investigating the effects of changing various system parameters and stochastic variables, which may or may not be within the engineer's control.

8.4 Recommendations for future work

There are a number of possible avenues for future work on both the stochastic analysis model and proposed optimisation method. First of all, the stochastic model can be improved by further reduction of the computational effort required by the compression heuristic. The compression heuristic can also be expanded to include other water supply system components. With respect to the optimisation model, its capabilities can be investigated and explored further, while it can be expanded to include other objectives.

8.4.1 Speeding up the stochastic model

Part of this project involved further development of the compression heuristic with the aim of speeding up the original stochastic analysis model by van Zyl *et al.* (2008). Elements of the compression heuristic that could be improved to achieve further reduction in computational effort were identified.

To begin with, the pre-run could be set to run for small reservoirs only, seeing as it takes less time for the simulations to converge than for large reservoir capacities. To obtain the demands failure frequency for large reservoirs, the results from the pre-run would just be extrapolated. This method of extrapolation would be adequate for the demands failure frequency as the results follow an exponential equation. However, it would be problematic to extrapolate the reservoir level distribution curve, as it does not follow any observable trend. Nevertheless, this idea is worth exploring further.

In the pre-run, each pair of values for the supply ratio and reservoir capacity is simulated one at a time. Another approach that could be explored is to run the pre-run simulation for each supply ratio, and the whole set of reservoir capacities at the same time. This means that for each supply ratio, the stochastic simulation would be performed and the failures and reservoir levels would be logged for the range of reservoir capacity values all at once. The smaller reservoir capacities would converge first. As the results for each reservoir capacity converge, the logging of the results is discontinued. The stochastic simulation would only be stopped for the current supply ratio, when results for all the reservoir sizes converge.

It was shown earlier in the sensitivity analysis that the fire demand has very little effect on the reliability of the storage reservoir. It is reasonable to consider simulating the fire demand along with the consumer demand in the pre-run simulation. This would make a marginal difference to the performance time of the pre-run, seeing as it already simulates every time step. It is possible that by combining the two demands in the pre-run, more failures would be generated within a shorter span of time, and the pre-run simulation would converge faster.

The computational speed of the compression heuristic would also be increased as only the pipe failure events would be simulated as critical periods in the events run.

Looking at the events run, the possibility of jumping to the start of the day of each event instead of the beginning of the week is considered. This approach would require a reservoir level distribution curve for each day of the week, which would result in a longer pre-run simulation time. The feasibility of this approach must be investigated.

Finally, while the use of Latin Hypercube sampling has been investigated and disregarded for use in the stochastic simulation, it is worthwhile investigating other sampling techniques in more detail. Importance sampling and descriptive sampling have been shown to be effective variation reduction techniques.

8.4.2 Expansion of the stochastic model

The stochastic model currently considers seven different pipe configurations in this project. It could be expanded to include other pipe configurations, such as an increased number of interconnections, or transitions from single pipelines to pipes in parallel.

There are a number of factors that affect the failure rate of pipes. To make the model more realistic, the effect of the pipe diameter and pipe material, etc. on the pipe failure rate can be included.

In addition, the stochastic model should be expanded to include more complex systems. These could include systems with more than one storage reservoir, or a system with a pipe network. Pumps and pumping stations were omitted from the stochastic analysis in this study. These should be built into the stochastic model so that the associated energy costs and reliability considerations of the pumps can be modelled and analysed.

8.4.3 Additional optimisation model capabilities

Thus far, the optimisation model has only been tested for a base system and several sensitivity analyses. The solutions in the trade-off curves range from very cheap, highly unreliable to very expensive, highly reliable systems. Most of these solutions would immediately be discarded by the designer, and only a very small selection of solutions would be acceptable. At this point, the optimisation model could be used again to narrow the search to the feasible region of solutions and explore them in more detail. This could be done through a hybrid optimisation technique where a local search is used in addition to the Monte Carlo simulation. Basically, once the optimisation has detected the possible pipe configuration and size that the designer will be interested in, the local search will be used to optimise the reservoir capacity.

The reliability-based optimisation model also provides engineers with a method of investigating the effects of changing various system parameters and stochastic variables, which may or may not be within the engineer's control.

8.4 Recommendations for future work

There are a number of possible avenues for future work on both the stochastic analysis model and proposed optimisation method. First of all, the stochastic model can be improved by further reduction of the computational effort required by the compression heuristic. The compression heuristic can also be expanded to include other water supply system components. With respect to the optimisation model, its capabilities can be investigated and explored further, while it can be expanded to include other objectives.

8.4.1 Speeding up the stochastic model

Part of this project involved further development of the compression heuristic with the aim of speeding up the original stochastic analysis model by van Zyl *et al.* (2008). Elements of the compression heuristic that could be improved to achieve further reduction in computational effort were identified.

To begin with, the pre-run could be set to run for small reservoirs only, seeing as it takes less time for the simulations to converge than for large reservoir capacities. To obtain the demands failure frequency for large reservoirs, the results from the pre-run would just be extrapolated. This method of extrapolation would be adequate for the demands failure frequency as the results follow an exponential equation. However, it would be problematic to extrapolate the reservoir level distribution curve, as it does not follow any observable trend. Nevertheless, this idea is worth exploring further.

In the pre-run, each pair of values for the supply ratio and reservoir capacity is simulated one at a time. Another approach that could be explored is to run the pre-run simulation for each supply ratio, and the whole set of reservoir capacities at the same time. This means that for each supply ratio, the stochastic simulation would be performed and the failures and reservoir levels would be logged for the range of reservoir capacity values all at once. The smaller reservoir capacities would converge first. As the results for each reservoir capacity converge, the logging of the results is discontinued. The stochastic simulation would only be stopped for the current supply ratio, when results for all the reservoir sizes converge.

It was shown earlier in the sensitivity analysis that the fire demand has very little effect on the reliability of the storage reservoir. It is reasonable to consider simulating the fire demand along with the consumer demand in the pre-run simulation. This would make a marginal difference to the performance time of the pre-run, seeing as it already simulates every time step. It is possible that by combining the two demands in the pre-run, more failures would be generated within a shorter span of time, and the pre-run simulation would converge faster.

It was assumed earlier in this study that pipes in parallel are laid in separate trenches, thus the cost of two parallel pipes is double the cost of a single pipeline. The optimisation model could be modified to include the cost impact of installing parallel pipes in the same trench. The model also gives rise to the possibility of scheduling the design implementation of a bulk water supply system over an extended length of time. Future researchers might want to consider phasing the construction so that capital is spent later in the project, thereby saving costs. For example, if the designer wishes to design for a scenario fifty years in the future, it would be more economical to only construct part of the system now, and add an additional pipe in parallel later on. The proposed method could be used in this instance.

8.4.4 Additional objectives for the optimisation problem

Currently the optimisation model deals with two objectives: cost and reliability. There are of course a myriad of other objectives which could be included in this multi-objective optimisation problem, such as water quality and system security.

It is a logical step to include water quality as an additional objective in the optimisation model. This is because a larger reservoir size, while providing greater system reliability, will have a longer water retention time, which may have a negative impact on the water quality of the system. If the water age and water mixing are not managed properly, then poor water quality will result in the depletion of disinfectant residual, biological problems, as well as problems with taste and odour of the water (National Research Council, 2006). Reliability and water quality are thus clearly conflicting objectives.

To include water quality as an objective in the optimisation model, an investigation into which measure would best represent the water quality is required. These measures could include the water age and the disinfectant residual. It would be prudent to consult the literature to find out what work on water quality modelling has been done already.

References

- Agrawal, R. 1995. *Simulated binary crossover for real-coded genetic algorithms*. Kanpur, India: Indian Institute of Technology. (Master's thesis).
- Alperovits, E. & Shamir, U. 1977. Design of optimal water distribution systems. *Water Resources Research*, 13(6): 885-900.
- Atiquzzaman, M., Liang, S.-Y. & Yu, X. 2006. Alternative decision making in water distribution network with NSGA-II. *Journal of Water Resources Planning and Management*, 132(2): 122-126.
- Awumah, K., Goulter, I. C. & Bhatt, S. K. 1990. Assessment of reliability in water distribution networks using entropy based measures. *Stochastic Hydrology and Hydraulics*, 4: 309-320.
- Awumah, K., Goulter, I. C. & Bhatt, S. K. 1991. Entropy-based redundancy measures in water-distribution networks. *Journal of Hydraulic Engineering*, 117(5): 595-614.
- Balakrishnan, P. V. & Jacob, V. S. 1996. Genetic algorithms for product design. *Management Science*, 42(8): 1105-1117.
- Bao, Y. & Mays, L. W. 1990. Model for water distribution system reliability. *Journal of Hydraulic Engineering*, 116(9): 1119-1137.
- Barta, B. & Rowse, N. 1998. Capital cost optimisation of pumping and reservoir system design. *WRC Report No 757/2/98*. Pretoria: Water Research Commission.
- Bhave, P. R. 1985. Optimal expansion of water distribution systems. *Journal of Environmental Engineering*, 111(2): 177-197.
- Blokker, E. J. M., Vreeburg, J. H. G. & van Dijk, J. C. 2010. Simulating residential water demand with a stochastic end-use model. *Journal of Water Resources Planning and Management*, 136(1): 19-26.
- Chang, C. & van Zyl, J.E. 2010. Speeding up the stochastic analysis of water distribution systems using a compression heuristic. *WDSA 2010 Conference*, Tucson, AZ, USA, 12-15 September 2010.
- Chilton, C. H. 1949. Cost data correlated. *Chemical Engineering*, 56:6-97.
- Clark, R. M. & Males, R. M. 1985. Simulating cost and quality in water distribution. *Journal of Water Resources Planning and Management*, 111(4): 454-466.
- Clark, R. M. & Dorsey, P. 1982. A model of costs for treating drinking water. *Journal American Water Works Association*, 74(12): 618-627.
- CSIR. 2000. *Guidelines for human settlement planning and management: Volume 2*. Pretoria: CSIR.
- Cullinane, M. J., Lansey, K. E. & Mays, L. W. 1992. Optimization-availability-based design of water-distribution networks. *Journal of Hydraulic Engineering*, 118(3): 420-441.
- Cui, L. & Kuczera, G. 2003. Optimizing urban water supply headworks using probabilistic search methods. *Journal of Water Resources Planning and Management*, 129(5): 380-387.

- Cui, L. & Kuczera, G. 2009. Assessment of the replicate compression heuristic to improve efficiency of urban water supply headworks optimization. *Journal of Water Resources Planning and Management*, 135(6): 451-457.
- Dandy, G. C., Simpson, A. R. & 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. (2006). Multi-objective trade-offs between cost and reliability in the replacement of water mains. *J. Water Resour. Plann. Manage.*, 132(2), 79-88.
- Deb, K. & Agrawal, R. B. 1995. Simulated binary crossover for continuous search space, *Complex Systems*, 9: 115-148.
- Deb, K. & Goyal, M. 1996. A combined genetic adaptive search (GeneAS) for engineering design. *Computer Science and Informatics*, 26(4): 30-45.
- Deb, K. & Goyal, M. 1997. Optimizing engineering designs using a combined genetic search. In T. Bäck (ed.), *7th International Conference on Genetic Algorithms*, 521-528. San Francisco: Morgan Kaufmann Publishers.
- Deb, K. & Gulati, S. 2001. Design of truss-structures for minimum weight using genetic algorithms. *Finite Elements in Analysis and Design*, 37: 447-465.
- Deb, K. & Kumar, A. 1995. Real-coded genetic algorithms with simulated binary crossover: Studies on multimodal and multiobjective problems. *Complex Systems*, 9: 431-454.
- Deb, K. Pratap, A., Agarwal, S. & Meyarivan, T. 2002. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2): 182-197.
- Dickson, R. D. 1972. Estimating water system costs. *Water treatment plant design*, R. L. Sanks, ed., Ann Arbor Science Publishers, Ann Arbor, Michigan, 763-799.
- Duan, N., Mays, L. W. & Lansey, K. E. 1990. Optimal reliability-based design of pumping and distribution systems. *Journal of Water Resources Planning and Management*, 116(2): 249-268.
- Farmani, R., Savic, D. A. & Walters, G. A. 2004. The simultaneous multi-objective optimization of Anytown pipe rehabilitation, tank sizing, tank siting, and pump operation schedules. *ASCE EWRI Conference*, Salt Lake City.
- Farmani, R., Walters, G. A. & Savic, D. A. 2005. Trade-off between total cost and reliability for Anytown water distribution network. *Journal of Water Resources Planning and Management*, 131(3): 161-171.
- Fonseca, C. M. & Fleming, P. J. 1993. Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. *Proceedings of the Fifth International Conference on Genetic Algorithms*, Forrest, S., Ed. San Mateo, CA: Morgan Kauffman, 416-423.
- Fowler, T. G., Manson, N. J. & Stephenson, D. 1997. Cost model of water supply in the rand water distribution area, Report to the Water Research Commission on the project "Optimisation of the Rand Water Board's Distribution System". *WRC Report No. 488/1/97*. Pretoria: Water Research Commission.
- Fujiwara, O. & De Silva, A. U. 1990. Algorithm for reliability-based optimal design of water networks. *Journal of Water Resources Planning and Management*, 116(3): 575-587.

- Fujiwara, O. & Khang, D. B. 1990. A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resources Research*, 26(4): 539-549.
- Giunta, A. A., Wojtkiewicz, S. F. & Eldred, M. S. 2003. Overview of modern design of experiments methods for computational simulations. In *41st AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, AIAA, AIAA-2003-0649.
- Goldberg, D. E. 1989. *Genetic algorithms in search, optimization and machine learning*. Reading, Massachusetts: Addison-Wesley Publishing Co.
- Goulter, I. C. 1992. Systems analysis in water distribution network design: From theory to practice. *Journal of Water Resources Planning and Management*, 118(3): 238-248.
- Goulter, I. 1995. Analytical and simulation models for reliability analysis in water distribution systems. *Improving efficiency and reliability in water distribution systems*, E. Cabrera & A. F. Vela, eds., Kluwer Academic, London, 235-266.
- Haarhoff, J. & van Zyl, J. E. 2002. Sizing of bulk water supply systems with a probabilistic method. *WRC Report No. 985/1/02*. Pretoria: Water Research Commission.
- Haestad Methods. 2003. *Advanced water distribution modelling and management*. Waterbury, CT: Haestad Press.
- Halhal, D., Walters, G. A., Ouazar, D. & Savic, D. A. 1997. Water network rehabilitation with structured messy genetic algorithm. *Journal of Water Resources Planning and Management*, 123(3): 137-146.
- Holland, J. H. 1975. *Adaptation in natural and artificial systems*. Cambridge, Massachusetts: MIT Press.
- Jacobs, P. & Goulter, I. C. 1988. Evaluation of methods for decomposition of water distribution networks for reliability analysis. *Civil Engineering Systems*, 5(2): 58-64.
- Kadu, M. S., Gupta, R. & 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.
- Kessler, A. 1988. Optimal design of water distribution networks using graph theory techniques (in Hebrew). Doctoral thesis in civil engineering, 142 pp., Technion, Israel Inst. Of Technol., Haifa.
- Khan, A. A., Lye, L. & Husain, T. 2008. Latin Hypercube sampling for uncertainty analysis in multiphase modelling. *Journal of Environmental Engineering and Science*, 7: 617-626.
- Khu, S. T. & Keedwell, E. 2004. Using multi-objective genetic algorithm to achieve design flexibility for water distribution systems. *Proceedings of the 6th International Conference on Hydroinformatics*, World Scientific, Singapore, 1011-1018.
- Kim, H. Y. & Clark, R. M. 1987. Input substitution and demand in the water supply production process. *Water Resources Research*, 23(3): 239-244.
- Kim, H. Y. & Clark, R. M. 1988. Economics of scale and scope in water supply. *Reg. Sci. Urban Econ.*, 18(7), 479-502.
- Knowles, J. & Corne, D. 1999. The Pareto archived evolution strategy: A new baseline algorithm for multiobjective optimization. *Proceedings of the 1999 Congress on Evolutionary Computation*. Piscataway, NJ: IEEE Press, 98-105.

- Kwietniewski, M. & Roman, M. 1997. Establishing performance criteria of water supply systems reliability. *Aqua*, 46(3): 181-184.
- Lansley, K. E., & Mays, L. W. 1989. Optimization model for water distribution system design. *Journal of Water Resources Planning and Management*, 115(10): 1401-1418.
- Lian, Y. & Yen, B. C. 2003. Comparison of risk calculation methods for a culvert. *Journal of Water Resources Planning and Management*, 129(2): 140-152.
- Lindsay, A. K. & Walski, T. M. 1982. Verification of cost estimating procedures for maps computer program. *Technical Report EL-82-3, U.S. Army Engineers Waterways Experiment Station, Vicksburg, Mississippi*.
- Liong, S. Y., Atiquzzaman, M. & Yu, X. 2004. Multi-objective algorithms to enhance decision making process in water distribution network problems. *Proceedings of the 2nd APHF Conference*, 138-146.
- Lobo, F. G. & Lima, C. F. 2005. A review of adaptive population sizing schemes in genetic algorithms. In *Proceedings of the 2005 Workshop on Parameter Settings in Genetic and Evolutionary Algorithms (PSGEA 2005), part of GECCO 2005*.
- Lu, Z. & Zhang, D. 2003. On importance sampling Monte Carlo approach to uncertainty analysis for flow and transport in porous media. *Advances in Water Resources*, 26: 1177-1188.
- Manache, G. & Melching, C. S. 2004. Sensitivity analysis of a water-quality model using Latin Hypercube sampling. *Journal of Water Resources Planning and Management*, 130(3): 232-242.
- MathWorks. 2011. *Pitfalls in fitting nonlinear models by transforming to linearity*. [Online]. Available:<http://www.mathworks.com/products/statistics/demos.html?file=/products/demos/shipping/stats/xform2lineardemo.html> [3 May 2011]
- Mays, L. W. & Tung, Y.-K. 1992. *Hydrosystems Engineering & Management*. Singapore: McGraw-Hill.
- McKay, M. D., Beckman, R. J. & Conover, W. J. 1979. A comparison of three methods for selecting input variables in the analysis of output from a computer code. *Technometrics*, 21(2): 239-245.
- Murphy, L. J. & Simpson, A. R. 1992. Genetic algorithms in pipe network optimization. *Research report no. R93*. University of Adelaide.
- National Research Council. 2006. *Drinking Water Distribution Systems: Assessing and Reducing Risks*. Washington DC: National Academies Press.
- Nel, D. T. 1993. *Bepaling van die optimal stoorkapasiteit van twee Johannesburgse Munisipale Diensreservoirs*. Johannesburg: Rand Afrikaans University. (M. Eng dissertation).
- Nel, D. & Haarhoff, J. 1996. Sizing municipal water storage tanks with Monte Carlo simulation. *Aqua*, 45(4): 203-212.
- Nel, D. T. 2009. *Factors that may compromise bulk water distribution reliability*. Johannesburg: University of Johannesburg. (PhD Thesis).
- Ostfeld, A. 2001. Reliability analysis of regional water distribution systems. *Urban Water*, 3: 253-360.

- Ostfeld, A., Kogan, D. & Shamir, U. 2001. Reliability simulation of water distribution systems – single and multiquality. *Urban Water*, 4: 53-61.
- Ostfeld, A. 2004. Reliability analysis of water distribution systems. *Journal of Hydroinformatics*, 6(4): 281-294.
- Ott, L. & Longnecker, M. 2008. *An Introduction to Statistical Methods and Data Analysis*. 6th ed. Belmont, CA: Brooks/Cole Cengage Learning.
- Pidd, M. 1998. *Computer simulation in management science*. 4th ed. Chichester: John Wiley & Sons.
- Prasad, T. D. & Park, N. 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.
- Quimpo, R. G. & Shamsi, U. M. 1987. Network analysis for water supply reliability determination. In Ragan, R. M. (ed.), *Proceedings of the American Society of Civil Engineers National Conferences on Hydraulic Engineering*. New York: American Society of Civil Engineers, 716-721.
- Quimpo, R. G. & Shamsi, U. M. 1991. Reliability-based distribution system maintenance. *Journal of Water Resources Planning and Management*, 117(2): 321-339.
- Quindry, G. E., Brill, E. D. & Liebman, J. C. 1981. Optimization of looped water distribution systems. *Journal of Environmental Engineering*, 107(4): 665-679.
- Raad, D. N. 2010. *Multi-objective optimisation of water distribution systems design using metaheuristics*. Stellenbosch: University of Stellenbosch. (PhD Dissertation).
- Renze, J. & Weisstein, E.W. 2011. *Law of Large Numbers – Wolfram MathWorld*. [Online]. Available: <http://mathworld.wolfram.com/LawofLargeNumbers.html> [4 June 2011]
- Saliby, E. 1990. Descriptive sampling: A better approach to Monte Carlo simulation. *Journal of the Operational Research Society*, 41(12): 1133-1142.
- Saliby, E. 1997. Descriptive sampling: An improvement over Latin Hypercube sampling. *Proceedings of the Winter Simulation Conference*, Rio de Janeiro, 230-233.
- Samra, S. & Essery, C. 2003. *NSW reference rates manual for valuation of water supply, sewerage and stormwater assets*. NSW, Australia: Ministry of Energy and Utilities.
- Savic, D. A. & 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.
- Schaake, J. C. & Lai, D. 1969. Linear programming and dynamic programming applications to water distribution network design. *Report. 116*, Hydrodynamics Laboratory, Department of Civil Engineering, Cambridge, Massachusetts.
- Schuyler, J. R. 1997. *Monte Carlo stopping rule*. [Online]. Available: <http://www.maxvalue.com/tip025.htm> [20 July 2010]
- Seshadri, A. 2009. *NSGA-II: A multi-objective optimisation algorithm*. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/10429-nsga-ii-a-multi-objective-optimization-algorithm>. [8 July 2011]

- Simpson, A. R., Dandy, G. C. & Murphy, L. J. 1994. Genetic algorithms compared to other techniques for pipe optimization. *Journal of Water Resources Planning and Management*, 120(4): 423-443.
- Snedecor, G. W. & Cochran, W. G. 1989. *Statistical Methods*. 8th ed. Iowa: Blackwell Publishing.
- Srinivas, N. & Deb, K. 1995. Multiobjective function optimization using nondominated sorting genetic algorithms. *Evol. Comput.*, 2(3): 221-248.
- Su, Y., Mays, L. W., Duan, N. & Lansey, K. E. 1987. Reliability-based optimisation model for water distribution systems. *Journal of Hydraulic Engineering*, 114(12): 1539-1556.
- Swamee, P. K. & Sharma, A. K. 2008. *Design of water supply pipe networks*. New Jersey: John Wiley & Sons, Inc.
- Tanyimboh, T. T. & Templeman, A. B. 1993. Maximum entropy flows for single-source networks. *Engineering Optimization*, 22: 49-63.
- Tanyimboh, T. T. & Templeman, A. B. 2000. A quantified assessment of the relationship between reliability and entropy of water distribution systems. *Engineering Optimization*, 33: 179-199.
- Tari, M. & Dahmani, A. 2005. Refined descriptive sampling: A better approach to Monte Carlo simulation. *Simulation Modelling Practice and Theory*, 14: 143-160.
- Thompson, K. D., Stedinger, J. R. & Heath, D. C. (1997). Evaluation and Presentation of Dam Failure and Flood Risks. *Journal of Water Resources Planning and Management*, 123(4): 216-227.
- Todini, E. 2000. Looped water distribution networks design using a resilience index based heuristic approach. *Urban Water*, 2(3): 115-122.
- Tung, Y.-K., Yen, B.-C. & Melching, C. S. 2006. *Hydrosystems engineering reliability assessment and risk analysis*. New York: McGraw-Hill.
- Vamvakeridou-Lyroudia, L. S., Savic, D. A. & Walters, G. A. 2007. Tank simulation for the optimization of water distribution networks. *Journal of Hydraulic Engineering*, 133(6): 625-636.
- van Zyl, J. E. & Haarhoff, J. 1999. The effect of feeder pipe configuration on the reliability of bulk water supply systems. *Water Industry Systems: Modelling and Optimization Applications*, Vol 2, Savic, D.A., Walters, G.A. (eds.), ISBN 0-86380-249-4, Research Studies Press, Baldock, England, 205-215.
- van Zyl, J. E., Piller, O. & Le Gat, Y. 2008. Sizing municipal storage tanks based on reliability criteria. *Journal of Water Resources Planning and Management*, 134(6): 548-555.
- van Zyl, J.E., Piller, O., le Gat, Y. & Walski, T.M. 2010. The risk of a municipal storage tank running dry due to user demands. *12th Annual Conference on Water Distribution Systems Analysis WDSA 2010*, 12-15 September, Tucson, Arizona.
- Vlok, G. 2010. *Optimal risk-based design of bulk water supply systems*. Cape Town: University of Cape Town. (MSc-Thesis).
- Wagner, M. J., Shamir, U. & Marks, D. H. 1988a. Water distribution reliability: analytical methods. *Journal of Water Resources Planning and Management*, 114(3): 253-275.

- Wagner, M. J., Shamir, U. & Marks, D. H. 1988b. Water distribution reliability: simulation methods. *Journal of Water Resources Planning and Management*, 114(3): 276-294.
- Walski, T. M. 1995. Optimization and pipe-sizing decisions. *Journal of Water Resources Planning and Management*, 121(4): 340-433.
- Walski, T. M., Brill, E. D. Jr., Gessler, J., Goulter, I. C., Jeppson, R. M., Lansey, K., Lee, H. L., Liebman, J. C., Mays, L. Morgan, D. R. & Ormsbee, L. 1987. Battle of the network models: Epilogue. *Journal of Water Resources Planning and Management*, 113(2): 191-203.
- Walski, T. M. 2001. The wrong paradigm – Why water distribution optimization doesn't work. *Journal of Water Resources Planning and Management*, 127(4), 203-205.
- Walters, G. A., Halhal, D., Savic, D. & Ouazar, D. 1999. Improved design of “Anytown” distribution network using structured messy genetic algorithms. *Urban Water*, 1: 23-38.
- Wu, W., Simpson, A. R. & Maier, H. R. 2010. Multiobjective genetic algorithm optimization of water distribution systems. *Journal of Water Resources Planning and Management*, 136(2): 146-155.
- Xu, C. & Goulter, I. C. 1998. Probabilistic model for water distribution reliability. *Journal of Water Resources Planning and Management*, 124(4): 218-228.
- Xu, C. & Powell, R. S. 1991. Water supply system reliability: Concepts and measures. *Civil Engineering Systems*, 8:191-195.
- Yang, S., Hsu, N., Louie, P. W. F. & Yeh, W. 1996. Water distribution network reliability: Stochastic simulation. *Journal of Infrastructure Systems*, 2(2):65-72.
- Zhang, M., Olbrich, M., Kinzelbach, H., Seider, D. & Barke, E. 2006. A fast and accurate Monte Carlo method for interconnect variation. In *ICICDT 2006*, January, IEEE, 207-210.
- Ziha, K. 1995. Descriptive sampling in structural safety. *Structural Safety*, 17: 33-41.
- Zio, E. 2009. *Computational methods for reliability and risk analysis*. Singapore: World Scientific.
- Zitzler, E. & Thiele, L. 1998. Multiobjective optimization using evolutionary algorithms – A comparative case study. *Parallel Problem Solving From Nature, V*, A. E. Eiben, T. Back, M Schoenaur, & H.-P. Schwefel, Eds. Berlin, Germany: Springer-Verlag, 292-301.

Appendix A: Comparison of reservoir level distributions

Table A.1 Reservoir level distribution for pre-run when (1) both failure and reservoir level results converge and (2) only failure results converge for $SR = 0.8$

$SR^* = 0.8$											
RC* = 2 h			RC* = 4 h			RC* = 6 h			RC* = 8 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0.00	0.0086	0	0.00	0.0058	0	0.00	0.0616	0.1111	44.55
0	0	0.00	0.0776	0.0833	6.84	0.3468	0.4074	14.87	0.7773	0.7407	4.94
0	0	0.00	0.3707	0.4167	11.04	0.9306	0.9259	0.51	0.9905	0.963	2.86
0.0581	0.1176	50.60	0.7672	0.75	2.29	0.9884	0.963	2.64	0.9953	0.963	3.35
0.1628	0.2941	44.64	0.9828	1	1.72	0.9942	0.963	3.24	0.9953	0.963	3.35
0.3488	0.4118	15.30	0.9914	1	0.86	0.9942	0.963	3.24	0.9953	0.963	3.35
0.5349	0.5882	9.06	1	1	0.00	0.9942	0.963	3.24	0.9953	0.963	3.35
0.7674	0.8235	6.81	1	1	0.00	0.9942	0.963	3.24	0.9953	0.963	3.35
0.9419	0.9412	0.07	1	1	0.00	1	1	0.00	1	1	0.00
1	1	0.00	1	1	0.00	1	1	0.00	1	1	0.00
RC* = 10 h			RC* = 12 h			RC* = 14 h			RC* = 16 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0.1481	0.2593	42.88	0.3297	0.4074	19.07	0.5288	0.5556	4.82	0.7507	0.7407	1.35
0.9794	0.963	1.70	0.9928	0.963	3.09	0.9968	0.963	3.51	0.9971	0.963	3.54
0.9959	0.963	3.42	0.9964	0.963	3.47	0.9968	0.963	3.51	0.9971	0.963	3.54
0.9959	0.963	3.42	0.9964	0.963	3.47	0.9968	0.963	3.51	0.9971	0.963	3.54
0.9959	0.963	3.42	0.9964	0.963	3.47	0.9968	0.963	3.51	0.9971	0.963	3.54
0.9959	0.963	3.42	0.9964	0.963	3.47	0.9968	0.963	3.51	0.9971	0.963	3.54
0.9959	0.963	3.42	0.9964	0.963	3.47	0.9968	0.963	3.51	0.9971	0.963	3.54
0.9959	0.963	3.42	1	1	0.00	1	1	0.00	1	1	0.00
1	1	0.00	1	1	0.00	1	1	0.00	1	1	0.00
1	1	0.00	1	1	0.00	1	1	0.00	1	1	0.00

* of seasonal peak demand

Table A.2 Reservoir level distribution for pre-run when (1) both failure and reservoir level results converge and (2) only failure results converge for $SR = 0.9$

SR* = 0.9											
RC* = 2 h			RC* = 4 h			RC* = 6 h			RC* = 8 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0.00	0	0	0.00	0.0073	0.0076	3.95	0.0059	0.0067	11.94
0	0	0.00	0	0	0.00	0.0803	0.0833	3.60	0.3077	0.3154	2.44
0	0	0.00	0.0814	0.0833	2.28	0.5109	0.5152	0.83	0.9349	0.9396	0.50
0	0	0.00	0.314	0.3333	5.79	0.9343	0.9318	0.27	0.9704	0.9799	0.97
0	0	0.00	0.686	0.6806	0.79	0.9781	0.9773	0.08	0.9941	0.9933	0.08
0.1224	0.15	18.40	0.9419	0.9444	0.26	0.9927	0.9924	0.03	0.9941	0.9933	0.08
0.2449	0.35	30.03	0.9884	0.9861	0.23	0.9927	0.9924	0.03	0.9941	0.9933	0.08
0.4898	0.65	24.65	0.9884	0.9861	0.23	0.9927	0.9924	0.03	0.9941	0.9933	0.08
0.6735	0.75	10.20	1	1	0.00	0.9927	0.9924	0.03	1	1	0.00
1	1	0.00	1	1	0.00	1	1	0.00	1	1	0.00
RC* = 10 h			RC* = 12 h			RC* = 14 h			RC* = 16 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0.0102	0.0066	54.55	0.0652	0.0728	10.44	0.1542	0.1419	8.67	0.2971	0.3082	3.60
0.6531	0.6821	4.25	0.9217	0.9404	1.99	0.9605	0.9742	1.41	0.971	0.9686	0.25
0.9643	0.9801	1.61	0.9826	0.9934	1.09	0.9921	0.9935	0.14	0.9964	0.9937	0.27
0.9949	0.9934	0.15	0.9957	0.9934	0.23	0.996	0.9935	0.25	0.9964	0.9937	0.27
0.9949	0.9934	0.15	0.9957	0.9934	0.23	0.996	0.9935	0.25	0.9964	0.9937	0.27
0.9949	0.9934	0.15	0.9957	0.9934	0.23	0.996	0.9935	0.25	0.9964	0.9937	0.27
0.9949	0.9934	0.15	0.9957	0.9934	0.23	0.996	0.9935	0.25	0.9964	0.9937	0.27
0.9949	0.9934	0.15	0.9957	0.9934	0.23	1	1	0.00	1	1	0.00
1	1	0.00	1	1	0.00	1	1	0.00	1	1	0.00
1	1	0.00	1	1	0.00	1	1	0.00	1	1	0.00

* of seasonal peak demand

Table A.3 Reservoir level distribution for pre-run when (1) both failure and reservoir level results converge and (2) only failure results converge for $SR = 1.0$

SR* = 1.0											
RC* = 2 h			RC* = 4 h			RC* = 6 h			RC* = 8 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0	0	0	0	0	0	0	0.0015	0.0015	0
0	0	0	0.0057	0.0057	0	0.0024	0.0024	0	0.0314	0.0314	0
0	0	0	0.0057	0.0057	0	0.1059	0.1059	0	0.4237	0.4237	0
0	0	0	0.0571	0.0571	0	0.4824	0.4824	0	0.7305	0.7305	0
0	0	0	0.2686	0.2686	0	0.7647	0.7647	0	0.8308	0.8308	0
0	0	0	0.56	0.56	0	0.8518	0.8518	0	0.8937	0.8937	0
0.05	0.05	0	0.8457	0.8457	0	0.9153	0.9153	0	0.9281	0.9281	0
0.15	0.15	0	0.9371	0.9371	0	0.9624	0.9624	0	0.9641	0.9641	0
0.55	0.55	0	0.9771	0.9771	0	0.9694	0.9694	0	0.982	0.982	0
1	1	0	1	1	0	1	1	0	1	1	0
RC* = 10 h			RC* = 12 h			RC* = 14 h			RC* = 16 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0.001	0.001	0	0.0013	0.0013	0	0.0078	0.0078	0	0.0192	0.0192	0
0.1527	0.1527	0	0.3227	0.3227	0	0.454	0.454	0	0.4994	0.4994	0
0.6381	0.6381	0	0.6258	0.6258	0	0.6303	0.6303	0	0.6239	0.6239	0
0.7617	0.7617	0	0.7275	0.7275	0	0.7243	0.7243	0	0.7193	0.7193	0
0.8385	0.8385	0	0.8279	0.8279	0	0.8027	0.8027	0	0.7946	0.7946	0
0.8988	0.8988	0	0.8846	0.8846	0	0.8816	0.8816	0	0.8726	0.8726	0
0.9397	0.9397	0	0.9355	0.9355	0	0.9323	0.9323	0	0.9271	0.9271	0
0.9767	0.9767	0	0.9694	0.9694	0	0.9669	0.9669	0	0.9644	0.9644	0
0.9883	0.9883	0	0.9909	0.9909	0	0.9903	0.9903	0	0.9912	0.9912	0
1	1	0	1	1	0	1	1	0	1	1	0

* of seasonal peak demand

Table A.4 Reservoir level distribution for pre-run when (1) both failure and reservoir level results converge and (2) only failure results converge for $SR = 1.1$

SR* = 1.1											
RC* = 2 h			RC* = 4 h			RC* = 6 h			RC* = 8 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0.0009	0.0009	0
0	0	0	0.0036	0.0036	0	0.0131	0.0131	0	0.0684	0.0684	0
0	0	0	0.0036	0.0036	0	0.1062	0.1062	0	0.3056	0.3056	0
0	0	0	0.0468	0.0468	0	0.3592	0.3592	0	0.4509	0.4509	0
0.0526	0.1667	68.45	0.1906	0.1906	0	0.5883	0.5883	0	0.5815	0.5815	0
0.0526	0.1667	68.45	0.4496	0.4496	0	0.6957	0.6957	0	0.7025	0.7025	0
0.0526	0.1667	68.45	0.759	0.759	0	0.8067	0.8067	0	0.8083	0.8083	0
0.2632	0.3333	21.03	0.9209	0.9209	0	0.9105	0.9105	0	0.9122	0.9122	0
1	1	0	1	1	0	1	1	0	1	1	0
RC* = 10 h			RC* = 12 h			RC* = 14 h			RC* = 16 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0	0	0	0	0.0001	0.0001	0	0.0003	0.0003	0
0.0105	0.0105	0	0.0287	0.0287	0	0.0349	0.0349	0	0.0276	0.0276	0
0.1638	0.1638	0	0.1476	0.1476	0	0.0972	0.0972	0	0.0589	0.0589	0
0.2924	0.2924	0	0.2214	0.2214	0	0.1624	0.1624	0	0.1081	0.1081	0
0.4014	0.4014	0	0.3193	0.3193	0	0.2466	0.2466	0	0.1845	0.1845	0
0.5196	0.5196	0	0.4378	0.4378	0	0.3638	0.3638	0	0.2942	0.2942	0
0.6454	0.6454	0	0.5774	0.5774	0	0.5064	0.5064	0	0.4417	0.4417	0
0.7718	0.7718	0	0.724	0.724	0	0.6696	0.6696	0	0.6243	0.6243	0
0.896	0.896	0	0.8709	0.8709	0	0.8423	0.8423	0	0.822	0.822	0
1	1	0	1	1	0	1	1	0	1	1	0

* of seasonal peak demand

Table A.5 Reservoir level distribution for pre-run when (1) both failure and reservoir level results converge and (2) only failure results converge for $SR = 1.2$

SR* = 1.2											
RC* = 2 h			RC* = 4 h			RC* = 6 h			RC* = 8 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0.0088	0.0088	0
0	0	0	0.0033	0.0033	0	0.0152	0.0152	0	0.0872	0.0872	0
0	0	0	0.0033	0.0033	0	0.1154	0.1154	0	0.2028	0.2028	0
0	0	0	0.0266	0.0266	0	0.305	0.305	0	0.3018	0.3018	0
0	0	0	0.1495	0.1495	0	0.5018	0.5018	0	0.4294	0.4294	0
0.1111	1	88.89	0.4086	0.4086	0	0.6458	0.6458	0	0.5997	0.5997	0
0.2222	1	77.78	0.7442	0.7442	0	0.8345	0.8345	0	0.7859	0.7859	0
1	1	0	1	1	0	1	1	0	1	1	0
RC* = 10 h			RC* = 12 h			RC* = 14 h			RC* = 16 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0	0	0	0	0	0	0	0	0	0
0.0004	0.0004	0	0.002	0.002	0	0.0027	0.0027	0	0.0022	0.0022	0
0.0233	0.0233	0	0.0274	0.0274	0	0.0159	0.0159	0	0.008	0.008	0
0.0901	0.0901	0	0.0564	0.0564	0	0.0328	0.0328	0	0.0186	0.0186	0
0.1553	0.1553	0	0.103	0.103	0	0.0654	0.0654	0	0.0408	0.0408	0
0.2417	0.2417	0	0.1789	0.1789	0	0.1262	0.1262	0	0.0874	0.0874	0
0.3668	0.3668	0	0.2961	0.2961	0	0.2323	0.2323	0	0.1806	0.1806	0
0.5387	0.5387	0	0.4698	0.4698	0	0.4037	0.4037	0	0.347	0.347	0
0.7537	0.7537	0	0.7083	0.7083	0	0.6629	0.6629	0	0.6204	0.6204	0
1	1	0	1	1	0	1	1	0	1	1	0

* of seasonal peak demand

Table A.6 Reservoir level distribution for pre-run when (1) both failure and reservoir level results converge and (2) only failure results converge for $SR = 1.3$

SR* = 1.3											
RC* = 2 h			RC* = 4 h			RC* = 6 h			RC* = 8 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	NaN	-	0	0	0	0	0	0	0	0	0
0	NaN	-	0	0	0	0	0	0	0	0	0
0	NaN	-	0	0	0	0	0	0	0.0006	0.0006	0
0	NaN	-	0	0	0	0.0016	0.0016	0	0.0198	0.0198	0
0	NaN	-	0	0	0	0.0317	0.0317	0	0.0898	0.0898	0
0	NaN	-	0	0	0	0.1496	0.1496	0	0.1734	0.1734	0
0	NaN	-	0.0576	0.0576	0	0.3496	0.3496	0	0.2761	0.2761	0
0	NaN	-	0.1975	0.1975	0	0.5398	0.5398	0	0.4361	0.4361	0
0	NaN	-	0.5226	0.5226	0	0.7407	0.7407	0	0.6758	0.6758	0
1	NaN	-	1	1	0	1	1	0	1	1	0
RC* = 10 h			RC* = 12 h			RC* = 14 h			RC* = 16 h		
Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference	Fail & level converge	Failure only converges	% difference
0	0	0	0	0	0	0	0	0	0	0	0
0.0001	0.0001	0	0.0001	0.0001	0	0.0002	0.0002	0	0.0002	0.0002	0
0.0039	0.0039	0	0.0046	0.0046	0	0.0028	0.0028	0	0.001	0.001	0
0.0265	0.0265	0	0.0159	0.0159	0	0.008	0.008	0	0.0036	0.0036	0
0.0602	0.0602	0	0.0347	0.0347	0	0.019	0.019	0	0.0105	0.0105	0
0.1153	0.1153	0	0.0737	0.0737	0	0.0455	0.0455	0	0.0278	0.0278	0
0.2097	0.2097	0	0.1514	0.1514	0	0.1076	0.1076	0	0.0752	0.0752	0
0.365	0.365	0	0.298	0.298	0	0.2408	0.2408	0	0.1928	0.1928	0
0.6207	0.6207	0	0.5644	0.5644	0	0.5099	0.5099	0	0.4599	0.4599	0
1	1	0	1	1	0	1	1	0	1	1	0

* of seasonal peak demand

Appendix B: Comparison of stochastic simulation times

Table B.1 List of simulation times for systems analysed with the full Monte Carlo method and the compression heuristic (events run only)

No.	SR*	RC* (h)	Time to converge		Hours simulated (h)		Total failure frequency	
			Full MC	CH (events)	Full MC	CH (events)	Full MC	CH
1	1.9526	23	22.84	35.36	997218	13245021	0.0176	0.0013
2	1.2696	8	59.78	28.37	2659917	10009269	6.3298	6.4042
3	1.6265	16	1074.80	130.03	48000000	48003148	0.0215	0.0202
4	1.5117	14	1054.12	129.19	48000000	48001636	0.0535	0.0536
5	1.8967	22	336.88	31.45	15350278	11918345	0.0011	0.0015
6	1.774	20	191.85	129.50	8738926	48000796	0.002	0.005
7	1.4836	13	1052.39	130.57	48000000	48002980	0.0858	0.088
8	1.0676	4	1.25	0.82	56755	261207	183.9824	168.6504
9	1.8303	21	112.05	129.69	5102996	48000292	0.0034	0.0035
10	1.4725	13	1052.38	129.92	48000000	48000001	0.0891	0.0943
11	1.6347	16	1052.83	129.69	48000000	48001468	0.0239	0.0242
12	1.8023	20	78.44	130.35	3572489	48002812	0.0049	0.0039
13	1.9257	23	225.45	18.49	10269477	6740851	0.0017	0.0026
14	1.7513	19	54.36	128.61	2475304	48000796	0.0071	0.0071
15	1.2175	7	20.23	11.16	920439	3797294	17.1119	18.1159
16	1.4354	12	1052.33	130.97	48000000	48000796	0.1668	0.1714
17	1.9387	23	354.35	67.71	16149713	24950122	0.0011	0.0021
18	1.9211	23	72.02	81.38	3280870	30387189	0.0053	0.0017
19	1.4398	12	1052.23	129.88	48000000	48000001	0.158	0.1621
20	1.899	22	335.11	129.20	15266414	48000001	0.0017	0.0037
21	1.105	5	2.65	1.89	119778	579065	104.6566	104.1553
22	1.3852	11	1053.67	131.86	47851916	48001636	0.3502	0.3773
23	1.8225	21	286.35	128.97	13046650	48001804	0.0013	0.0039
24	1.0594	4	1.06	0.87	48403	297061	190.9344	174.4875
25	1.1819	6	7.68	5.81	349721	1884615	41.4052	40.288
26	1.2426	8	40.24	23.40	1833035	7922007	8.2437	8.4106
27	1.2388	8	34.87	22.46	1588773	7836163	8.9211	8.7332
28	1.6236	16	1052.55	129.12	48000000	48000964	0.0241	0.0237
29	1.3086	9	159.50	57.85	7269836	20483926	2.3196	2.4669
30	1.2389	8	37.87	22.79	1724997	7719404	8.7295	8.7574
31	1.0645	4	1.16	0.91	52675	272683	186.9243	170.6061
32	1.7594	19	52.83	129.69	2404890	48000292	0.0109	0.0057
33	1.4728	13	1052.19	131.29	48000000	48001468	0.0947	0.091
34	1.9352	23	143.74	41.33	6548819	15229168	0.0027	0.0012
35	1.4927	13	1052.31	130.82	48000000	48000460	0.0872	0.0876
36	1.4477	12	1052.19	131.85	48000000	48001300	0.1518	0.158
37	1.8539	21	144.62	128.75	6588259	48000460	0.0027	0.0042
38	1.5489	15	19.28	131.38	878807	48001972	0.0199	0.0344
39	1.2425	8	41.19	22.90	1875524	7961831	8.2671	8.4181
40	1.6885	18	38.09	129.78	1735407	48001300	0.0101	0.0084

Table B.1 List of simulation times for systems analysed with the full Monte Carlo method and the compression heuristic (events run only) (continued)

No.	SR*	RC* (h)	Time to converge (s)		Hours simulated (h)		Total failure frequency	
			Full MC	CH (events)	Full MC	CH (events)	Full MC	CH
41	1.8462	21	92.84	15.52	4226709	5543808	0.0041	0.0032
42	1.0687	4	1.25	0.98	56757	318806	185.9823	166.7553
43	1.6972	18	27.08	129.16	1233398	48000796	0.0142	0.0068
44	1.4105	11	1052.28	131.07	48000000	48002140	0.3088	0.3258
45	1.8402	21	59.02	29.03	2686840	10785884	0.0065	0.0016
46	1.5277	14	1052.60	129.41	48000000	48000292	0.0509	0.0533
47	1.724	18	118.94	128.66	5416959	48000796	0.0129	0.0098
48	1.4574	13	1052.14	131.18	48000000	48001804	0.0933	0.104
49	1.3394	10	403.85	106.71	18419995	38201634	0.8846	1.0229
50	1.2302	7	21.15	12.93	962946	4373897	15.9654	16.1148
51	1.2338	8	37.59	21.83	1712244	7391320	9.2499	9.2781
52	1.6981	18	135.01	130.78	6151821	48001804	0.0142	0.0116
53	1.3376	10	365.69	118.77	16682062	42995206	0.9725	1.0185
54	1.5646	15	1052.50	129.97	48000000	48000796	0.0381	0.0371
55	1.1933	7	15.50	9.11	706245	3060980	23.0832	22.8123
56	1.713	18	75.97	129.19	3458565	48006676	0.0051	0.0112
57	1.4095	11	1051.98	132.10	48000000	48001300	0.3101	0.3236
58	1.867	22	301.25	128.90	13724088	48000964	0.0013	0.0029
59	1.861	21	224.13	128.51	10211731	48001972	0.0017	0.0028
60	1.6139	16	1052.70	128.35	48000000	48000460	0.0186	0.0235
61	1.5217	14	1052.63	128.35	48000000	48000001	0.0509	0.0553
62	1.9048	22	362.34	18.49	16511399	6767178	0.0011	0.0026
63	1.8305	21	174.76	130.81	7962762	48000001	0.0022	0.0035
64	1.6627	17	42.02	129.32	1911837	48000628	0.0092	0.0158
65	1.8271	21	116.91	128.66	5325341	48001132	0.0033	0.0034
66	1.6772	17	168.50	129.65	7675867	48000001	0.0148	0.0173
67	1.3749	11	956.65	133.09	43648076	48000628	0.3881	0.4405
68	1.3252	10	335.27	100.11	15290680	35772301	1.0753	1.1965
69	1.3741	11	819.93	132.50	37409876	48000964	0.4046	0.4261
70	1.5574	15	1052.69	130.71	48000000	48001300	0.0354	0.0317
71	1.7408	19	86.79	128.16	3949811	48001468	0.0111	0.0062
72	1.3438	10	397.19	100.63	18119109	36159960	0.922	0.9966
73	1.8466	21	103.51	129.30	4713561	48001804	0.0056	0.0029
74	1.5897	15	1052.69	130.21	48000000	48000001	0.0318	0.0335
75	1.4019	11	1052.03	130.91	48000000	48000460	0.3283	0.3379
76	1.7176	18	92.78	129.83	4225527	48001636	0.0041	0.0097
77	1.5692	15	1052.58	130.60	48000000	48000124	0.031	0.0329
78	1.4726	13	1052.29	131.23	48000000	48000001	0.0907	0.0914
79	1.7098	18	29.71	128.69	1353508	48000001	0.0129	0.0077
80	1.6402	17	70.17	129.40	3195693	48000292	0.0055	0.0103

Table B.1 List of simulation times for systems analysed with the full Monte Carlo method and the compression heuristic (events run only) (continued)

No.	SR*	RC* (h)	Time to converge (s)		Hours simulated (h)		Total failure frequency	
			Full MC	CH (events)	Full MC	CH (events)	Full MC	CH
81	1.8051	20	37.25	129.85	1693974	48000001	0.0103	0.0027
82	1.959	24	206.37	34.74	9400078	12810332	0.0019	0.0014
83	1.5465	14	1052.53	129.87	48000000	48000292	0.0557	0.0462
84	1.8861	22	203.76	94.51	9282019	35294686	0.0019	0.0007
85	1.2143	7	17.19	11.36	782683	3912909	18.859	18.319
86	1.9808	24	39.10	25.83	1779929	9427102	0.0098	0.0019
87	1.3079	9	144.46	57.87	6580364	20227862	2.3643	2.4798
88	1.2897	9	121.56	53.89	5540949	19176430	2.86	2.9696
89	1.882	22	324.42	38.24	14783082	13952565	0.0012	0.0019
90	1.7504	19	48.70	128.84	2218618	48000292	0.0079	0.0062
91	1.1797	6	7.20	5.62	327956	1884615	41.1882	40.7972
92	1.0612	4	1.06	0.95	48328	298957	191.2308	173.026
93	1.8992	22	106.78	19.84	4859231	7347787	0.0036	0.0024
94	1.2392	8	39.30	21.56	1789678	7438004	8.7028	8.7255
95	1.3338	10	388.01	101.06	17699465	36292291	0.9587	1.0839
96	1.6784	17	43.58	129.55	1984197	48000124	0.0088	0.013
97	1.3202	9	160.82	63.81	7328348	22938352	2.203	2.1856
98	1.4958	13	1052.42	130.27	48000000	48001972	0.0834	0.0893
99	1.1115	5	2.40	2.13	109216	673334	100.4205	100.7842
100	1.9889	24	663.17	25.95	30218606	9569755	0.0006	0.0018

* of seasonal peak demand

Appendix C: Sensitivity analysis results

Table C.1 Optimisation results for sensitivity analysis of supply pipe length

Length = 1 km						Length = 10 km						Length = 100 km					
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq
2/0	0.227	3.70	4.00	553867	0.00	2/0	0.363	3.66	4.00	4035963	0	2/1	0.530	2.86	5.64	53444689	0
2/0	0.227	3.70	4.00	553867	0.00	2/0	0.363	3.66	4.00	4035963	0	2/2	0.479	2.19	10.74	48860528	0.0002
2/1	0.182	2.07	4.65	552255	0.03	2/0	0.286	1.96	16.00	3710500	0	2/2	0.479	2.19	9.07	48805889	0.0006
2/1	0.182	2.07	4.54	546724	0.03	2/2	0.286	1.96	14.43	3690007	0	2/0	0.428	1.63	16.00	44150300	0.0006
2/1	0.182	2.07	4.48	543468	0.04	2/1	0.286	1.96	13.56	3654428	0.0002	2/2	0.428	1.63	7.80	43936236	0.1916
2/1	0.182	2.07	4.41	539814	0.04	2/1	0.286	1.96	11.45	3592311	0.0008	1/0	0.574	1.76	16.00	29296484	0.1937
2/1	0.182	2.07	4.33	535115	0.04	2/2	0.286	1.96	10.32	3568094	0.0011	1/0	0.574	1.76	14.82	29264565	0.3083
2/1	0.182	2.07	4.27	532082	0.05	2/1	0.286	1.96	9.06	3515449	0.0031	1/0	0.530	1.43	16.00	27243963	0.3279
2/1	0.182	2.07	4.18	526939	0.05	2/1	0.286	1.96	7.52	3461257	0.01	1/0	0.530	1.43	14.94	27215303	0.4992
2/0	0.182	2.07	4.24	523114	0.06	1/0	0.363	1.83	16.00	2593198	0.0136	1/0	0.530	1.43	7.37	26978118	7.3511
2/1	0.182	2.07	4.07	520244	0.06	1/0	0.363	1.83	14.78	2560323	0.0267	1/0	0.530	1.43	6.83	26957243	9.1548
2/1	0.182	2.07	4.02	517380	0.06	1/0	0.363	1.83	13.69	2529893	0.0441	1/0	0.479	1.10	16.00	24850924	9.2407
2/1	0.182	2.07	4.00	516224	0.06	1/0	0.363	1.83	12.95	2508717	0.0476	1/0	0.479	1.10	13.93	24794308	15.0872
2/0	0.182	2.07	4.08	513710	0.07	1/0	0.322	1.34	15.90	2393161	0.0494	1/0	0.479	1.10	12.75	24760472	19.9326
1/0	0.227	1.85	5.43	507386	0.21	1/0	0.322	1.34	14.54	2356255	0.1048	1/0	0.479	1.10	11.59	24725874	25.8099
1/0	0.227	1.85	5.36	504293	0.22	1/0	0.322	1.34	11.61	2271602	0.4311	1/0	0.479	1.10	10.54	24693266	32.3619
1/0	0.227	1.85	5.29	501157	0.23	1/0	0.363	1.83	11.20	2258879	0.5468	1/0	0.479	1.10	10.40	24688837	33.9954
1/0	0.227	1.85	5.21	497144	0.25	1/0	0.322	1.34	4.89	2223046	1.0748	1/0	0.479	1.10	9.79	24669337	40.0427
1/0	0.227	1.85	5.10	492293	0.26	1/0	0.322	1.34	8.54	2171954	2.4837	1/0	0.479	1.10	9.51	24659957	43.2509
1/0	0.227	1.85	5.04	489374	0.28	1/0	0.322	1.34	7.89	2149155	3.5715	1/0	0.479	1.10	9.46	24658205	45.4556
1/0	0.227	1.85	5.00	487416	0.28	1/0	0.322	1.34	7.16	2122044	5.5949	1/0	0.479	1.10	8.96	24641422	48.3808
1/0	0.227	1.85	4.95	485069	0.29	1/0	0.322	1.34	6.66	2102664	7.661	1/0	0.479	1.10	8.54	24627205	52.7429
1/0	0.227	1.85	4.89	482098	0.31	1/0	0.322	1.34	6.33	2089494	9.2225	1/0	0.479	1.10	8.16	24613703	57.4561
1/0	0.227	1.85	4.83	479272	0.32	1/0	0.322	1.34	5.99	2075292	11.5775	1/0	0.479	1.10	7.75	24599064	62.2649
1/0	0.227	1.85	4.72	473652	0.36	1/0	0.322	1.34	5.85	2069363	12.5632	1/0	0.479	1.10	7.38	24585215	68.1466
1/0	0.227	1.85	4.64	469190	0.37	1/0	0.322	1.34	5.48	2053515	15.8343	1/0	0.479	1.10	7.29	24582029	73.6386
1/0	0.227	1.85	4.61	467937	0.38	1/0	0.322	1.34	5.18	2039777	19.183	1/0	0.479	1.10	6.99	24570437	77.043
1/0	0.227	1.85	4.57	465705	0.38	1/0	0.322	1.34	4.80	2021150	24.5558	1/0	0.479	1.10	6.62	24556281	84.2388
1/0	0.227	1.85	4.51	462348	0.41	1/0	0.322	1.34	4.56	2009090	28.9622	1/0	0.479	1.10	6.50	24551409	87.4202
1/0	0.227	1.85	4.46	459679	0.42	1/0	0.322	1.34	4.41	2001219	31.6986	1/0	0.479	1.10	6.34	24545069	91.5503
1/0	0.227	1.85	4.39	456075	0.45	1/0	0.286	0.98	4.00	1977642	40.4481	1/0	0.479	1.10	6.13	24536491	95.7209
1/0	0.227	1.85	4.34	453302	0.47	1/0	0.286	0.98	7.98	1977576	188.9838	1/0	0.479	1.10	5.97	24529732	97.4365

Length = 1 km												Length = 10 km												Length = 100 km											
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq												
1/0	0.227	1.85	4.27	449790	0.47	1/0	0.286	0.98	7.68	1966836	191.6911	1/0	0.479	1.10	5.91	24526962	102.4689	1/0	0.227	1.85	4.00	383609	277.26												
1/0	0.227	1.85	4.25	448134	0.49	1/0	0.286	0.98	7.23	1950036	197.8197	1/0	0.479	1.10	5.65	24515801	106.4773	1/0	0.227	1.85	4.08	388551	272.21												
1/0	0.227	1.85	4.18	444398	0.51	1/0	0.286	0.98	7.07	1944116	199.638	1/0	0.479	1.10	5.57	24512400	112.6987	1/0	0.227	1.85	4.08	385458	274.88												
1/0	0.227	1.85	4.10	439508	0.56	1/0	0.286	0.98	6.83	1934813	201.9789	1/0	0.479	1.10	5.30	24500108	116.3712	1/0	0.227	1.85	4.08	388551	272.21												
1/0	0.227	1.85	4.05	436597	0.58	1/0	0.286	0.98	6.75	1931838	204.3048	1/0	0.479	1.10	5.25	24497792	123.1437	1/0	0.227	1.85	4.08	388551	272.21												
1/0	0.227	1.85	4.00	433883	0.59	1/0	0.286	0.98	6.54	1923215	207.1175	1/0	0.479	1.10	5.08	24489890	128.7346	1/0	0.227	1.85	4.08	388551	272.21												
1/0	0.182	1.03	4.37	432915	189.05	1/0	0.286	0.98	6.31	1914173	210.3233	1/0	0.479	1.10	4.98	24485064	132.0619	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.30	428877	191.22	1/0	0.286	0.98	6.00	1901304	214.1469	1/0	0.479	1.10	4.80	24476318	136.4796	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.26	426485	192.06	1/0	0.286	0.98	5.64	1885879	216.39	1/0	0.479	1.10	4.75	24473649	143.8229	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.18	421997	195.62	1/0	0.286	0.98	5.60	1884072	219.9988	1/0	0.479	1.10	4.75	24473649	143.8229	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.13	419278	196.47	1/0	0.286	0.98	5.47	1878391	220.9281	1/0	0.479	1.10	4.56	24463928	148.7341	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.11	418269	198.29	1/0	0.286	0.98	5.16	1864136	224.5219	1/0	0.479	1.10	4.42	24456488	152.0757	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.08	416281	199.31	1/0	0.286	0.98	4.92	1852651	228.3321	1/0	0.479	1.10	4.19	24443758	164.9193	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.04	413929	200.32	1/0	0.286	0.98	4.74	1843725	232.1735	1/0	0.479	1.10	4.16	24441781	169.9015	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.182	1.03	4.00	411489	201.14	1/0	0.286	0.98	4.49	1830484	235.1221	1/0	0.479	1.10	4.00	24432656	173.5365	1/0	0.182	1.03	4.08	388551	272.21												
1/0	0.127	0.40	4.00	383609	277.26	2/0	0.182	2.07	4.00	1513744	356.9697	1/0	0.322	0.39	4.00	16951337	250.5471	1/0	0.127	0.40	4.08	385458	274.88												
1/0	0.127	0.40	4.03	385458	274.88	1/0	0.227	0.53	4.00	1803066	242.4032	1/0	0.322	0.39	7.03	17090804	248.5537	1/0	0.127	0.40	4.08	385458	274.88												

* of seasonal peak demand

Table C.2 Optimisation results for sensitivity analysis of system head																	
Head = 30 m				Head = 60 m				Head = 120 m									
Config	Diam (m)	SR*	RC† (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC† (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC† (h)	Cost (A\$)	Fail freq
2/2	0.363	2.52	7.75	4231564	0	2/0	0.363	3.66	4.00	4035963	0	2/0	0.322	3.89	4.00	3641386	0
2/1	0.322	1.84	16.00	4072335	0	2/0	0.363	3.66	4.00	4035963	0	2/0	0.322	3.89	4.00	3641386	0
2/0	0.322	1.84	16.00	4059654	0	2/0	0.286	1.96	16.00	3710500	0	2/0	0.286	2.85	12.76	3620397	0
2/2	0.322	1.84	12.39	3983925	0.0001	2/2	0.286	1.96	14.43	3690007	0	2/0	0.286	2.85	12.26	3605582	0
2/1	0.322	1.84	8.62	3851287	0.009	2/1	0.286	1.96	9.06	3515449	0.0031	2/0	0.286	2.85	10.38	3547960	0.0001
2/1	0.286	1.35	16.00	3721587	0.0094	2/1	0.286	1.96	7.52	3461257	0.01	2/0	0.286	2.85	9.86	3531177	0.0001
2/0	0.286	1.35	15.17	3688263	0.0202	1/0	1.83	16.00	2593198	0.0136	2/0	0.286	2.85	9.36	3514707	0.0001	
2/1	0.286	1.35	14.01	3667268	0.043	1/0	1.83	14.78	2560323	0.0267	2/0	0.286	2.85	8.62	3489136	0.0002	
2/0	0.286	1.35	13.62	3645214	0.0597	1/0	1.83	13.69	2529893	0.0441	2/1	1.55	16.00	3140586	0.0002		
1/0	0.428	1.94	15.79	2897517	0.0645	1/0	1.83	12.95	2508717	0.0476	2/1	1.55	15.51	3127422	0.0003		
1/0	0.428	1.94	15.75	2896433	0.065	1/0	1.34	15.90	2393161	0.0494	2/1	1.55	14.65	3104093	0.0007		
1/0	0.428	1.94	14.00	2848440	0.0837	1/0	1.34	14.54	2356255	0.1048	2/2	1.55	13.68	3085685	0.0022		
1/0	0.428	1.94	12.75	2812626	0.0945	1/0	1.34	11.61	2271602	0.4311	2/1	1.55	13.66	3076507	0.0024		
1/0	0.363	1.26	16.00	2593198	0.103	1/0	1.34	11.20	2258879	0.5468	2/1	1.55	13.08	3059714	0.0033		
1/0	0.363	1.26	14.18	2543677	0.2639	1/0	1.83	4.89	2223046	1.0748	2/1	1.55	12.52	3043559	0.0053		
1/0	0.363	1.26	12.52	2496068	0.597	1/0	1.34	8.54	2171954	2.4837	2/2	1.55	11.07	3008487	0.0185		
1/0	0.428	1.94	4.00	2484805	1.2465	1/0	1.34	7.89	2149155	3.5715	2/2	1.55	10.57	2992833	0.0284		
1/0	0.363	1.26	10.69	2440434	1.5688	1/0	1.34	7.16	2122044	5.5919	1/0	1.42	16.00	2221333	0.0289		
1/0	0.363	1.26	10.22	2425284	2.0333	1/0	1.34	6.66	2102664	7.661	1/0	1.42	15.79	2215789	0.0358		
1/0	0.363	1.26	9.09	2388146	3.7443	1/0	1.34	6.33	2089494	9.2225	1/0	1.42	15.02	2195062	0.0429		
1/0	0.363	1.26	8.86	2380505	4.2627	1/0	1.34	5.99	2075292	11.5775	1/0	1.42	14.38	2177414	0.0646		
1/0	0.363	1.26	8.06	2352340	6.7028	1/0	1.34	5.85	2069363	12.5632	1/0	1.42	13.71	2158468	0.0829		
1/0	0.363	1.26	7.76	2341471	7.8813	1/0	1.34	5.48	2053515	15.8343	1/0	1.42	12.79	2132116	0.1223		
1/0	0.363	1.26	7.39	2327967	9.7934	1/0	1.34	5.18	2039777	19.183	1/0	1.42	12.28	2117106	0.1604		
1/0	0.363	1.26	7.21	2321202	10.866	1/0	1.34	4.80	2021150	24.5558	1/0	1.42	11.95	2107131	0.1969		
1/0	0.363	1.26	6.99	2312990	12.3484	1/0	1.34	4.56	2009090	28.9622	1/0	1.42	11.59	2096306	0.2242		
1/0	0.363	1.26	6.44	2291050	14.7671	1/0	1.34	4.41	2001219	31.6986	1/0	1.42	10.67	2067948	0.3625		
1/0	0.363	1.26	6.30	2285698	18.119	1/0	1.34	4.00	1977642	40.4481	1/0	1.42	10.57	2064682	0.3743		
1/0	0.363	1.26	5.92	2269782	22.6314	1/0	0.286	0.98	1977576	188.9838	1/0	0.286	1.42	10.08	2049036	0.4853	

* of seasonal peak demand																	
Head = 30 m			Head = 60 m			Head = 120 m											
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq
1/0	0.363	1.26	5.83	2265901	23.9564	1/0	0.286	0.98	7.68	196836	191.6911	1/0	0.322	1.94	4.33	1996442	1.0876
1/0	0.363	1.26	5.56	2234074	28.152	1/0	0.286	0.98	7.23	1950036	197.8197	1/0	0.322	1.94	4.00	1977642	1.1888
1/0	0.363	1.26	5.31	2242907	32.3891	1/0	0.286	0.98	7.07	1944116	199.658	1/0	0.286	1.42	7.66	1966212	1.9517
1/0	0.363	1.26	5.23	2239461	33.9246	1/0	0.286	0.98	6.83	1934813	201.9789	1/0	0.286	1.42	7.15	1947035	2.647
1/0	0.363	1.26	5.11	2233440	36.5636	1/0	0.286	0.98	6.75	1931838	204.3048	1/0	0.286	1.42	6.90	1937512	3.1301
1/0	0.363	1.26	5.00	2228593	38.5695	1/0	0.286	0.98	6.54	1923215	207.1175	1/0	0.286	1.42	6.43	1918815	4.2068
1/0	0.363	1.26	4.93	2224886	40.6535	1/0	0.286	0.98	6.31	1914173	210.3233	1/0	0.286	1.42	5.74	1890372	6.7467
1/0	0.363	1.26	4.67	2212138	46.7704	1/0	0.286	0.98	6.00	1901304	214.1469	1/0	0.286	1.42	5.36	1873494	8.702
1/0	0.363	1.26	4.61	2208873	48.3743	1/0	0.286	0.98	5.64	1883879	216.39	1/0	0.286	1.42	5.10	1861251	10.4798
1/0	0.363	1.26	4.51	2203534	51.955	1/0	0.286	0.98	5.60	1884072	219.9988	1/0	0.286	1.42	4.65	1839156	14.2673
1/0	0.363	1.26	4.41	2198327	54.407	1/0	0.286	0.98	5.47	1878391	220.9281	1/0	0.286	1.42	4.48	1830183	16.0167
1/0	0.363	1.26	4.35	2195227	56.8615	1/0	0.286	0.98	5.16	1864136	224.5219	1/0	0.286	1.42	4.41	1826445	16.8354
1/0	0.363	1.26	4.27	2190667	59.4865	1/0	0.286	0.98	4.92	1852651	228.3321	1/0	0.286	1.42	4.30	1820125	18.1664
1/0	0.363	1.26	4.19	2185825	63.2237	1/0	0.286	0.98	4.74	1843725	232.1735	1/0	0.286	1.42	4.16	1812634	19.9394
1/0	0.363	1.26	4.07	2179078	67.0713	1/0	0.286	0.98	4.49	1830484	235.1221	1/0	0.286	1.42	4.05	1803940	21.6246
1/0	0.363	1.26	4.00	2174931	69.5443	1/0	0.286	0.98	4.38	1824976	236.7835	1/0	0.286	1.42	4.00	1803066	22.1381
1/0	0.227	0.37	4.10	1519574	208.1297	1/0	0.286	0.98	4.00	1803066	242.4032	1/0	0.182	0.43	5.33	1367900	312.0492
1/0	0.227	0.37	4.00	1513744	208.4193	1/0	0.227	0.53	4.00	1513744	356.9697	1/0	0.182	0.43	4.00	1289804	313.7986

Table C.3 Optimisation results for sensitivity analysis of peak hourly factor for hourly demand pattern

PHF = 1.25						PHF = 1.49						PHF = 1.75					
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq
2/2	0.286	1.96	8.46	3505947	< 0.0002	2/0	0.363	3.66	4.00	4035963	< 0.0002	2/1	0.322	2.67	8.12	3833554	< 0.0001
2/2	0.286	1.96	5.49	3390525	0.0002	2/0	0.363	3.66	4.00	4035963	< 0.0002	2/2	0.286	1.96	16.00	3732674	0.0001
2/2	0.286	1.96	5.38	3385612	0.0007	2/0	0.286	1.96	16.00	3710500	< 0.0002	2/2	0.286	1.96	14.20	3683773	0.0019
2/2	0.286	1.96	4.65	3350499	0.0007	2/2	0.286	1.96	14.43	3690007	< 0.0002	2/2	0.286	1.96	11.85	3615645	0.0037
2/2	0.286	1.96	4.37	3335829	0.0009	2/1	0.286	1.96	13.56	3654428	0.0002	2/1	0.286	1.96	10.51	3563224	0.0143
2/2	0.286	1.96	4.30	3331588	0.0012	2/1	0.286	1.96	11.45	3592311	0.0008	2/1	0.286	1.96	9.61	3533927	0.0286
2/2	0.286	1.96	4.00	3314406	0.0014	2/2	0.286	1.96	10.32	3568094	0.0011	1/0	0.363	1.83	16.00	2593198	0.0293
2/1	0.286	1.96	4.00	3303319	0.0026	2/1	0.286	1.96	9.06	3515449	0.0031	1/0	0.363	1.83	14.03	2539344	0.0744
1/0	0.322	1.34	16.00	2395910	0.0163	2/1	0.286	1.96	7.52	3461257	0.01	1/0	0.363	1.83	12.04	2481904	0.1369
1/0	0.322	1.34	15.29	2376885	0.0294	1/0	0.363	1.83	16.00	2593198	0.0136	1/0	0.363	1.83	10.32	2428573	0.2511
1/0	0.322	1.34	14.99	2360830	0.0304	1/0	0.363	1.83	14.78	2560323	0.0267	1/0	0.363	1.83	9.23	2392818	0.4093
1/0	0.322	1.34	14.37	2351654	0.0357	1/0	0.363	1.83	13.69	2529893	0.0441	1/0	0.363	1.83	8.38	2363740	0.5876
1/0	0.322	1.34	13.87	2337721	0.0403	1/0	0.363	1.83	12.95	2508717	0.0476	1/0	0.363	1.83	7.70	2339328	0.8428
1/0	0.322	1.34	13.38	2323814	0.0506	1/0	0.322	1.34	15.90	2393161	0.0494	1/0	0.363	1.83	6.10	2277245	2.0167
1/0	0.322	1.34	12.53	2299133	0.0803	1/0	0.322	1.34	14.54	2356255	0.1048	1/0	0.363	1.83	5.08	2232307	3.9555
1/0	0.322	1.34	12.14	2287628	0.0952	1/0	0.322	1.34	11.61	2271602	0.4311	1/0	0.322	1.34	9.72	2211800	8.0407
1/0	0.322	1.34	11.63	2272176	0.1146	1/0	0.322	1.34	11.20	2258879	0.5468	1/0	0.363	1.83	4.00	2174931	8.2387
1/0	0.322	1.34	10.93	2250501	0.1591	1/0	0.363	1.83	4.89	2223046	1.0748	1/0	0.322	1.34	8.47	2169487	14.9638
1/0	0.322	1.34	10.77	2245460	0.1753	1/0	0.322	1.34	8.54	2171954	2.4837	1/0	0.322	1.34	8.11	2156768	17.7867
1/0	0.322	1.34	10.42	2234605	0.2037	1/0	0.322	1.34	7.89	2149155	3.5715	1/0	0.322	1.34	7.54	2136361	24.0569
1/0	0.322	1.34	10.18	2226857	0.2235	1/0	0.322	1.34	7.16	2122044	5.5949	1/0	0.322	1.34	7.33	2128475	26.7132
1/0	0.322	1.34	9.58	2207303	0.291	1/0	0.322	1.34	6.66	2102664	7.661	1/0	0.322	1.34	6.99	2115417	31.4651
1/0	0.322	1.34	9.22	2195177	0.3571	1/0	0.322	1.34	6.33	2089494	9.2225	1/0	0.322	1.34	6.05	2077978	50.3982
1/0	0.322	1.34	8.85	2182732	0.4238	1/0	0.322	1.34	5.99	2075292	11.5775	1/0	0.322	1.34	5.89	2071421	54.7122
1/0	0.322	1.34	8.45	2169058	0.5037	1/0	0.322	1.34	5.85	2069363	12.5632	1/0	0.322	1.34	5.63	2060042	62.2673
1/0	0.322	1.34	8.08	2155882	0.6173	1/0	0.322	1.34	5.48	2053515	15.8343	1/0	0.322	1.34	5.40	2049731	69.3308
1/0	0.322	1.34	7.76	2144347	0.741	1/0	0.322	1.34	5.18	2039777	19.183	1/0	0.322	1.34	5.22	2041531	75.7582
1/0	0.322	1.34	7.25	2125553	0.9857	1/0	0.322	1.34	4.80	2021150	24.5558	1/0	0.322	1.34	5.03	2032768	82.3683
1/0	0.322	1.34	6.49	2096126	1.4985	1/0	0.322	1.34	4.56	2009090	28.9622	1/0	0.322	1.34	4.85	2023951	90.0335
1/0	0.322	1.34	6.19	2083737	1.757	1/0	0.322	1.34	4.41	2001219	31.6986	1/0	0.322	1.34	4.62	2011997	101.254
1/0	0.322	1.34	5.75	2065406	2.3297	1/0	0.322	1.34	4.00	1977642	40.4481	1/0	0.322	1.34	4.47	2004255	109.0901
1/0	0.322	1.34	5.71	2063693	2.4583	1/0	0.286	0.98	7.98	1977576	188.9838	1/0	0.322	1.34	4.31	1995784	117.2122

PHF = 1.25						PHF = 1.49						PHF = 1.75					
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq
1/0	0.322	1.34	5.44	2051610	2.93	1/0	0.286	0.98	7.68	1966836	191.6911	1/0	0.322	1.34	4.14	1985765	128.255
1/0	0.322	1.34	5.30	2045014	3.2083	1/0	0.286	0.98	7.23	1950036	197.8197	1/0	0.322	1.34	4.12	1984909	129.0422
1/0	0.322	1.34	5.14	2037773	3.5222	1/0	0.286	0.98	7.07	1944116	199.638	1/0	0.322	1.34	4.00	1977642	136.6199
1/0	0.322	1.34	5.00	2030951	3.9673	1/0	0.286	0.98	6.83	1934813	201.9789	1/0	0.286	0.98	7.96	1976981	222.966
1/0	0.322	1.34	4.85	2023735	4.3277	1/0	0.286	0.98	6.75	1931838	204.3048	1/0	0.286	0.98	7.63	1965013	232.4592
1/0	0.322	1.34	4.75	2018913	4.6424	1/0	0.286	0.98	6.54	1923215	207.1175	1/0	0.286	0.98	7.25	1950692	240.8446
1/0	0.322	1.34	4.67	2014954	4.8885	1/0	0.286	0.98	6.31	1914173	210.3233	1/0	0.286	0.98	6.85	1935661	250.018
1/0	0.322	1.34	4.55	2008372	5.3014	1/0	0.286	0.98	6.00	1901304	214.1469	1/0	0.286	0.98	6.76	1932048	253.2158
1/0	0.322	1.34	4.54	2008118	5.3448	1/0	0.286	0.98	5.64	1885879	216.39	1/0	0.286	0.98	6.50	1921772	259.0222
1/0	0.322	1.34	4.44	2002369	5.8517	1/0	0.286	0.98	5.60	1884072	219.9988	1/0	0.286	0.98	6.23	1910926	268.2834
1/0	0.322	1.34	4.31	1995757	6.3093	1/0	0.286	0.98	5.47	1878391	220.9281	1/0	0.286	0.98	6.08	1904461	275.3615
1/0	0.322	1.34	4.27	1993214	6.4631	1/0	0.286	0.98	5.16	1864136	224.5219	1/0	0.286	0.98	5.89	1896609	278.4651
1/0	0.322	1.34	4.16	1986745	6.9396	1/0	0.286	0.98	4.92	1852651	228.3321	1/0	0.286	0.98	5.55	1881819	289.2719
1/0	0.322	1.34	4.00	1977642	7.7699	1/0	0.286	0.98	4.74	1843725	232.1735	1/0	0.286	0.98	5.28	1869934	296.0936
1/0	0.227	0.53	7.12	1656643	100.0372	1/0	0.286	0.98	4.49	1830484	235.1221	1/0	0.286	0.98	4.77	1845083	316.2047
1/0	0.227	0.53	5.40	1585901	100.8692	1/0	0.286	0.98	4.38	1824976	236.7835	1/0	0.286	0.98	4.44	1828140	328.6777
1/0	0.227	0.53	5.39	1585622	101.692	1/0	0.286	0.98	4.00	1803066	242.4032	1/0	0.286	0.98	4.00	1803066	340.0393
1/0	0.227	0.53	4.00	1513744	101.6972	1/0	0.227	0.53	4.00	1513744	356.9697	1/0	0.227	0.53	4.00	1513744	362.9056

* of seasonal peak demand

Table C.4 Optimisation results for sensitivity analysis of fire rate

Fire rate = 0 fires/annum						Fire rate = 6 fires/annum						Fire rate = 24 fires/annum					
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq
2/1	0.322	2.67	4.73	3694446	< 0.0001	2/0	0.363	3.66	4.00	4035963	< 0.0002	2/0	0.322	2.67	9.01	3851810	< 0.0003
2/2	0.286	1.96	14.18	3683142	< 0.0001	2/0	0.363	3.66	4.00	4035963	< 0.0002	2/2	0.286	1.96	16.00	3732674	< 0.0003
2/2	0.286	1.96	13.17	3654415	0.0001	2/0	0.286	1.96	16.00	3710500	< 0.0002	2/2	0.286	1.96	14.27	3685679	0.0003
2/2	0.286	1.96	12.07	3622194	0.0002	2/2	0.286	1.96	14.43	3690007	< 0.0002	2/2	0.286	1.96	13.08	3651927	0.0003
2/2	0.286	1.96	10.19	3563977	0.0021	2/1	0.286	1.96	13.56	3654428	0.0002	2/2	0.286	1.96	12.47	3634110	0.0007
2/2	0.286	1.96	9.20	3531335	0.0022	2/1	0.286	1.96	11.45	3592311	0.0008	2/2	0.286	1.96	10.67	3579133	0.0008
2/2	0.286	1.96	8.77	3516851	0.0035	2/2	0.286	1.96	10.32	3568094	0.0011	2/1	0.286	1.96	10.07	3548938	0.0017
2/1	0.286	1.96	8.33	3490406	0.0067	2/1	0.286	1.96	9.06	3515449	0.0031	2/2	0.286	1.96	8.50	3507305	0.0036
2/2	0.286	1.96	7.18	3459460	0.0116	2/1	0.286	1.96	7.52	3461257	0.01	1/0	0.363	1.83	16.00	2593198	0.0052
1/0	0.363	1.83	16.00	2593198	0.0147	1/0	0.363	1.83	16.00	2593198	0.0136	1/0	0.363	1.83	13.52	2525082	0.0302
1/0	0.363	1.83	15.30	2574522	0.0211	1/0	0.363	1.83	14.78	2560323	0.0267	1/0	0.322	1.34	16.00	2395910	0.0449
1/0	0.363	1.83	14.69	2557906	0.0259	1/0	0.363	1.83	13.69	2529893	0.0441	1/0	0.322	1.34	13.81	2336089	0.1361
1/0	0.363	1.83	13.80	2532995	0.035	1/0	0.363	1.83	12.95	2508717	0.0476	1/0	0.322	1.34	10.88	2248940	0.6442
1/0	0.363	1.83	13.20	2515728	0.0454	1/0	0.322	1.34	15.90	2393161	0.0494	1/0	0.363	1.83	4.57	2206799	1.2308
1/0	0.322	1.34	16.00	2395889	0.0471	1/0	0.322	1.34	14.54	2356255	0.1048	1/0	0.363	1.83	4.00	2174931	1.6094
1/0	0.322	1.34	14.45	2353949	0.1056	1/0	0.322	1.34	11.61	2271602	0.4311	1/0	0.322	1.34	7.92	2150098	3.5181
1/0	0.322	1.34	13.35	2322820	0.1832	1/0	0.322	1.34	11.20	2258879	0.5468	1/0	0.322	1.34	7.48	2133981	4.626
1/0	0.322	1.34	12.47	2297403	0.2852	1/0	0.363	1.83	4.89	2223046	1.0748	1/0	0.322	1.34	6.31	2088817	9.262
1/0	0.322	1.34	11.28	2261334	0.532	1/0	0.322	1.34	8.54	2171954	2.4837	1/0	0.322	1.34	5.94	2073537	12.0064
1/0	0.322	1.34	9.91	2218091	1.1151	1/0	0.322	1.34	7.89	2149155	3.5715	1/0	0.322	1.34	5.67	2061925	13.8478
1/0	0.363	1.83	4.21	2186997	1.4737	1/0	0.322	1.34	7.16	2122044	5.5949	1/0	0.322	1.34	5.39	2049129	16.7476
1/0	0.322	1.34	8.25	2161852	2.9331	1/0	0.322	1.34	6.66	2102664	7.661	1/0	0.322	1.34	5.19	2039936	19.497
1/0	0.322	1.34	7.75	2143875	3.9416	1/0	0.322	1.34	6.33	2089494	9.2225	1/0	0.322	1.34	5.01	2031726	21.603
1/0	0.322	1.34	7.16	2121832	5.6452	1/0	0.322	1.34	5.99	2075292	11.5775	1/0	0.322	1.34	4.56	2009197	28.2359
1/0	0.322	1.34	6.76	2106719	7.1134	1/0	0.322	1.34	5.85	2069363	12.5632	1/0	0.322	1.34	4.37	1998998	32.3207
1/0	0.322	1.34	6.37	2091105	9.1217	1/0	0.322	1.34	5.48	2053515	15.8343	1/0	0.322	1.34	4.00	1977642	33.7668
1/0	0.322	1.34	5.99	2075532	11.5587	1/0	0.322	1.34	5.18	2039777	19.183	1/0	0.286	0.98	7.88	1974141	182.609
1/0	0.322	1.34	5.48	2053405	15.8917	1/0	0.322	1.34	4.80	2021150	24.5558	1/0	0.286	0.98	7.68	1966836	188.7659
1/0	0.322	1.34	5.09	2035214	20.4672	1/0	0.322	1.34	4.56	2009090	28.9622	1/0	0.286	0.98	7.03	1942648	188.7834
1/0	0.322	1.34	4.91	2026621	22.8	1/0	0.322	1.34	4.41	2001219	31.6986	1/0	0.286	0.98	6.87	1936454	192.8455
1/0	0.322	1.34	4.73	2017818	25.7403	1/0	0.322	1.34	4.00	1977642	40.4481	1/0	0.286	0.98	6.54	1923389	195.8879
1/0	0.322	1.34	4.55	2008410	28.71	1/0	0.286	0.98	7.98	1977576	188.9838	1/0	0.286	0.98	6.44	1919465	199.6324

Table C.5 Optimisation results for sensitivity analysis of pipe failure rate

Pipe failure rate = 0.1 failures/km/annum						Pipe failure rate = 0.2 failures/km/annum						Pipe failure rate = 0.5 failures/km/annum					
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq
2/1	0.322	2.67	5.42	3726998	< 0.0001	2/0	0.363	3.66	4.00	4035963	< 0.0002	2/1	0.322	2.67	7.34	3805066	< 0.0004
2/2	0.286	1.96	14.01	3678404	< 0.0001	2/0	0.363	3.66	4.00	4035963	< 0.0002	2/2	0.286	1.96	14.87	3702168	< 0.0004
2/1	0.286	1.96	13.98	3666372	< 0.0001	2/0	0.286	1.96	16.00	3710500	< 0.0002	2/2	0.286	1.96	12.12	3623806	0.0004
2/2	0.286	1.96	12.32	3629605	0.0001	2/2	0.286	1.96	14.43	3690007	< 0.0002	2/2	0.286	1.96	11.29	3598495	0.0008
2/2	0.286	1.96	10.00	3557940	0.0013	2/1	0.286	1.96	13.56	3654428	0.0002	2/2	0.286	1.96	10.33	3568371	0.0016
2/1	0.286	1.96	9.10	3516953	0.0033	2/1	0.286	1.96	11.45	3592311	0.0008	2/2	0.286	1.96	9.48	3540599	0.0021
2/1	0.286	1.96	8.24	3487285	0.005	2/2	0.286	1.96	10.32	3568094	0.0011	2/1	0.286	1.96	9.45	3528587	0.0027
2/1	0.286	1.96	7.94	3476596	0.0062	2/1	0.286	1.96	9.06	3515449	0.0031	2/1	0.286	1.96	8.58	3498991	0.0044
1/0	0.363	1.83	16.00	2593198	0.0068	2/1	0.286	1.96	7.52	3461257	0.01	2/1	0.286	1.96	7.97	3477607	0.0116
1/0	0.363	1.83	14.92	2563972	0.0115	1/0	0.363	1.83	16.00	2593198	0.0136	2/1	0.286	1.96	7.13	3446591	0.0208
1/0	0.363	1.83	12.42	2493230	0.0268	1/0	0.363	1.83	14.78	2560323	0.0267	2/1	0.286	1.96	6.33	3415232	0.0352
1/0	0.322	1.34	16.00	2395910	0.0288	1/0	0.363	1.83	13.69	2529893	0.0441	1/0	0.363	1.83	16.00	2593198	0.041
1/0	0.322	1.34	13.64	2331059	0.1068	1/0	0.363	1.83	12.95	2508717	0.0476	1/0	0.363	1.83	14.71	2558301	0.0747
1/0	0.363	1.83	6.50	2293671	0.3168	1/0	0.322	1.34	15.90	2393161	0.0494	1/0	0.363	1.83	13.51	2524720	0.1026
1/0	0.363	1.83	5.82	2265391	0.4141	1/0	0.322	1.34	14.54	2356255	0.1048	1/0	0.322	1.34	16.00	2395910	0.108
1/0	0.363	1.83	5.37	2245618	0.53	1/0	0.322	1.34	11.61	2271602	0.4311	1/0	0.322	1.34	14.86	2365048	0.1883
1/0	0.363	1.83	4.84	2220555	0.6531	1/0	0.322	1.34	11.20	2258879	0.5468	1/0	0.322	1.34	13.69	2332532	0.2938
1/0	0.363	1.83	4.12	2182125	1.0581	1/0	0.363	1.83	4.89	2223046	1.0748	1/0	0.322	1.34	11.16	2257766	0.8876
1/0	0.363	1.83	4.00	2174931	1.1109	1/0	0.322	1.34	8.54	2171954	2.4837	1/0	0.322	1.34	10.14	2225668	1.4874
1/0	0.322	1.34	8.07	2155510	2.9292	1/0	0.322	1.34	7.89	2149155	3.5715	1/0	0.322	1.34	8.68	2176969	3.0424
1/0	0.322	1.34	7.24	2124876	4.9565	1/0	0.322	1.34	7.16	2122044	5.5949	1/0	0.322	1.34	7.57	2137321	5.34
1/0	0.322	1.34	6.52	2097216	7.8001	1/0	0.322	1.34	6.66	2102664	7.661	1/0	0.322	1.34	6.48	2095372	9.9045
1/0	0.322	1.34	6.15	2082332	9.9576	1/0	0.322	1.34	6.33	2089494	9.2225	1/0	0.322	1.34	6.06	2078562	12.5066
1/0	0.322	1.34	5.71	2063385	13.2614	1/0	0.322	1.34	5.99	2075292	11.5775	1/0	0.322	1.34	5.64	2060584	16.0878
1/0	0.322	1.34	4.94	2028385	21.5914	1/0	0.322	1.34	5.85	2069363	12.5632	1/0	0.322	1.34	5.26	2043303	20.2729
1/0	0.322	1.34	4.72	2017054	25.196	1/0	0.322	1.34	5.48	2053515	15.8343	1/0	0.322	1.34	5.02	2031948	23.3935
1/0	0.322	1.34	4.46	2003537	29.7871	1/0	0.322	1.34	5.18	2039777	19.183	1/0	0.322	1.34	4.90	2026158	24.9823
1/0	0.322	1.34	4.20	1989359	35.4103	1/0	0.322	1.34	4.80	2021150	24.5558	1/0	0.322	1.34	4.58	2009852	30.4318
1/0	0.322	1.34	4.00	1977642	39.9214	1/0	0.322	1.34	4.56	2009090	28.9622	1/0	0.322	1.34	4.41	2001235	33.4950
1/0	0.286	0.98	7.93	1975960	189.296	1/0	0.322	1.34	4.41	2001219	31.6986	1/0	0.322	1.34	4.23	1991049	37.5485
1/0	0.286	0.98	7.89	1974264	191.7301	1/0	0.322	1.34	4.00	1977642	40.4481	1/0	0.322	1.34	4.18	1988123	38.6265
1/0	0.286	0.98	7.43	1957456	197.1467	1/0	0.286	0.98	7.98	1977576	188.9838	1/0	0.322	1.34	4.00	1977642	42.3445

Pipe failure rate = 0.1 failures/km/annum												Pipe failure rate = 0.2 failures/km/annum												Pipe failure rate = 0.5 failures/km/annum											
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq												
1/0	0.286	0.98	7.33	1953872	199.1346	1/0	0.286	0.98	7.68	1966836	191.6911	1/0	0.286	0.98	7.96	1976889	189.9146	1/0	0.286	0.98	7.33	1953872	199.1346												
1/0	0.286	0.98	7.01	1941799	200.2031	1/0	0.286	0.98	7.23	1950036	197.8197	1/0	0.286	0.98	7.64	1965324	191.5965	1/0	0.286	0.98	7.01	1941799	200.2031												
1/0	0.286	0.98	6.52	1922463	203.8052	1/0	0.286	0.98	7.07	1944116	199.6338	1/0	0.286	0.98	7.38	1955558	195.8298	1/0	0.286	0.98	6.52	1922463	203.8052												
1/0	0.286	0.98	6.08	1904657	209.4592	1/0	0.286	0.98	6.83	1934813	201.9789	1/0	0.286	0.98	6.78	1933013	202.7113	1/0	0.286	0.98	6.08	1904657	209.4592												
1/0	0.286	0.98	5.89	1896690	213.704	1/0	0.286	0.98	6.75	1931838	204.3048	1/0	0.286	0.98	6.60	1925654	204.9008	1/0	0.286	0.98	5.89	1896690	213.704												
1/0	0.286	0.98	5.53	1880956	217.0825	1/0	0.286	0.98	6.54	1923215	207.1125	1/0	0.286	0.98	6.55	1923944	207.4004	1/0	0.286	0.98	5.53	1880956	217.0825												
1/0	0.286	0.98	5.37	1873790	221.1079	1/0	0.286	0.98	6.31	1914123	210.3233	1/0	0.286	0.98	6.24	1911146	208.9733	1/0	0.286	0.98	5.37	1873790	221.1079												
1/0	0.286	0.98	5.12	1862323	224.2252	1/0	0.286	0.98	6.00	1901304	214.1469	1/0	0.286	0.98	6.13	1906912	212.2765	1/0	0.286	0.98	5.12	1862323	224.2252												
1/0	0.286	0.98	4.97	1855015	227.9758	1/0	0.286	0.98	5.64	1885879	216.39	1/0	0.286	0.98	5.90	1896934	215.9851	1/0	0.286	0.98	4.97	1855015	227.9758												
1/0	0.286	0.98	4.74	1843737	230.9834	1/0	0.286	0.98	5.60	1884072	219.9988	1/0	0.286	0.98	5.55	1882077	219.3325	1/0	0.286	0.98	4.74	1843737	230.9834												
1/0	0.286	0.98	4.54	1833191	235.2129	1/0	0.286	0.98	5.47	1878391	220.9281	1/0	0.286	0.98	5.16	1864088	223.9391	1/0	0.286	0.98	4.54	1833191	235.2129												
1/0	0.286	0.98	4.44	1828135	236.4885	1/0	0.286	0.98	5.16	1864136	224.5219	1/0	0.286	0.98	5.02	1857675	228.4529	1/0	0.286	0.98	4.44	1828135	236.4885												
1/0	0.286	0.98	4.22	1816019	238.6279	1/0	0.286	0.98	4.92	1852651	228.3321	1/0	0.286	0.98	4.75	1843978	231.3065	1/0	0.286	0.98	4.22	1816019	238.6279												
1/0	0.286	0.98	4.14	1811503	242.7198	1/0	0.286	0.98	4.74	1843725	232.1735	1/0	0.286	0.98	4.64	1838762	234.8995	1/0	0.286	0.98	4.14	1811503	242.7198												
1/0	0.286	0.98	4.00	1803066	243.4496	1/0	0.286	0.98	4.49	1830484	235.1221	1/0	0.286	0.98	4.43	1827441	236.9673	1/0	0.286	0.98	4.00	1803066	243.4496												
1/0	0.286	0.98	4.83	1558951	350.5573	1/0	0.286	0.98	4.38	1824926	236.7835	1/0	0.286	0.98	4.14	1811156	241.087	1/0	0.286	0.98	4.83	1558951	350.5573												
1/0	0.227	0.53	4.46	1540034	359.8966	1/0	0.286	0.98	4.00	1803066	242.4032	1/0	0.286	0.98	4.00	1803066	243.7355	1/0	0.227	0.53	4.46	1540034	359.8966												
1/0	0.227	0.53	4.00	1513744	359.9548	1/0	0.227	0.53	4.00	1513744	356.9697	1/0	0.227	0.53	4.00	1513744	355.1802	1/0	0.227	0.53	4.00	1513744	355.1802												

of seasonal peak demand

Table C.6 Optimisation results for sensitivity analysis of average pipe failure duration																	
Ave failure duration = 3 h						Ave failure duration = 4.5 h						Ave failure duration = 9 h					
Config	Diam (m)	SR*	RC*(h)	Cost (\$)	Fail freq	Config	Diam (m)	SR*	RC*(h)	Cost (\$)	Fail freq	Config	Diam (m)	SR*	RC*(h)	Cost (\$)	Fail freq
2/1	0.322	2.67	5.72	3740528	0	2/0	0.363	3.66	4.00	4035963	0	2/2	0.322	3.66	4.30	3684246	0
2/1	0.286	1.96	14.79	3688882	0	2/0	0.363	3.66	4.00	4035963	0	2/2	0.286	1.96	14.79	3594747	0.0007
2/1	0.286	1.96	14.39	3677833	0	2/0	0.286	1.96	16.00	3710500	0	2/2	0.286	1.96	11.17	3594747	0.0007
2/1	0.286	1.96	12.70	3629679	-0.0001	2/2	0.286	1.96	14.43	3690007	0	2/2	0.286	1.96	10.47	3573019	0.0027
2/1	0.286	1.96	11.28	3586996	-0.0003	2/1	0.286	1.96	13.56	3654428	0.0002	2/2	0.286	1.96	9.06	3526715	0.0045
1/0	0.363	1.83	16.00	2593198	-0.0004	2/1	0.286	1.96	11.45	3592311	0.0008	2/2	0.286	1.96	8.12	3494195	0.006
1/0	0.363	1.83	14.62	2555955	-0.0015	2/2	0.286	1.96	10.32	3568094	0.0011	2/2	0.286	1.96	7.55	3473299	0.0132
1/0	0.363	1.83	13.50	2524520	-0.0026	2/1	0.286	1.96	9.06	3515449	0.0031	2/2	0.286	1.96	6.07	3415710	0.0354
1/0	0.363	1.83	11.90	2477579	-0.0036	2/1	0.286	1.96	7.52	3461257	0.01	2/2	0.286	1.96	4.69	3352742	0.1168
1/0	0.322	1.34	16.00	2395910	0.0183	1/0	0.363	1.83	16.00	2593198	0.0136	2/2	0.286	1.96	4.07	3318693	0.2106
1/0	0.322	1.34	14.60	2358071	0.0465	1/0	0.363	1.83	14.78	2560323	0.0267	2/2	0.286	1.96	4.00	3314406	0.2178
1/0	0.322	1.34	13.47	2326386	0.0945	1/0	0.363	1.83	13.69	2529893	0.0441	1/0	0.363	1.83	16.00	2593198	0.2919
1/0	0.322	1.34	12.28	2291817	0.1994	1/0	0.363	1.83	12.95	250817	0.0476	1/0	0.363	1.83	14.13	2542237	0.4478
1/0	0.322	1.34	11.94	2281642	0.2533	1/0	0.322	1.34	15.90	2393161	0.0494	1/0	0.322	1.34	16.00	2395910	0.4624
1/0	0.322	1.83	5.37	2245837	0.4822	1/0	0.322	1.34	11.61	2271602	0.4311	1/0	0.322	1.34	13.14	2316918	0.8814
1/0	0.363	1.83	4.82	2219616	0.6864	1/0	0.322	1.34	11.20	2258879	0.5468	1/0	0.322	1.34	12.32	2292878	1.0579
1/0	0.322	1.34	9.43	2202195	1.2115	1/0	0.363	1.83	4.89	2223046	1.0748	1/0	0.322	1.34	11.21	2259160	1.4052
1/0	0.322	1.34	8.33	2164570	2.4477	1/0	0.322	1.34	8.54	2171954	2.4837	1/0	0.363	1.83	5.19	2237475	1.7922
1/0	0.322	1.34	7.76	2144371	3.5365	1/0	0.322	1.34	7.89	2149155	3.5715	1/0	0.363	1.83	4.28	2190943	2.1893
1/0	0.322	1.34	7.09	2119239	5.4777	1/0	0.322	1.34	7.16	2122044	5.5949	1/0	0.322	1.34	7.93	2150388	4.5317
1/0	0.322	1.34	6.64	2101925	7.2767	1/0	0.322	1.34	6.66	2102664	7.661	1/0	0.322	1.34	7.70	2142057	5.0527
1/0	0.322	1.34	6.20	2084336	9.5897	1/0	0.322	1.34	6.33	2084994	9.2225	1/0	0.322	1.34	7.11	2120239	6.7189
1/0	0.322	1.34	5.85	2069706	12.1933	1/0	0.322	1.34	5.99	2075292	11.5775	1/0	0.322	1.34	6.53	2097584	9.1407
1/0	0.322	1.34	5.61	2059041	14.1774	1/0	0.322	1.34	5.85	2069363	12.5632	1/0	0.322	1.34	6.17	2082771	11.3408
1/0	0.322	1.34	5.28	2044140	17.5224	1/0	0.322	1.34	5.48	2053515	15.8343	1/0	0.322	1.34	5.95	2073609	12.7533
1/0	0.322	1.34	4.92	2027297	22.1052	1/0	0.322	1.34	5.18	2039777	19.183	1/0	0.322	1.34	5.48	2053208	16.6328
1/0	0.322	1.34	4.31	1995675	33.078	1/0	0.322	1.34	4.80	2021150	24.5558	1/0	0.322	1.34	5.45	2051982	17.1332
1/0	0.322	1.34	4.18	1988110	35.9405	1/0	0.322	1.34	4.56	2009090	28.9622	1/0	0.322	1.34	4.74	2018403	26.247
1/0	0.322	1.34	4.00	1977642	40.0959	1/0	0.322	1.34	4.41	2001219	31.6986	1/0	0.322	1.34	4.49	2005261	30.2871
1/0	0.286	0.98	7.92	1975398	190.5718	1/0	0.322	1.34	4.00	1977642	40.4481	1/0	0.322	1.34	4.42	2001449	32.0072
1/0	0.286	0.98	7.79	1970714	192.1441	1/0	0.286	0.98	7.98	1977576	188.9838	1/0	0.322	1.34	4.21	1989745	36.3306

Table C.6 Optimisation results for sensitivity analysis of average pipe failure duration

Ave failure duration = 3 h						Ave failure duration = 4.5 h						Ave failure duration = 9 h					
Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq	Config	Diam (m)	SR*	RC* (h)	Cost (A\$)	Fail freq
1/0	0.286	0.98	7.31	1953219	195.4029	1/0	0.286	0.98	7.68	1966836	191.6911	1/0	0.322	1.34	4.01	1977953	41.1142
1/0	0.286	0.98	7.30	1952642	200.1608	1/0	0.286	0.98	7.23	1950036	197.8197	1/0	0.322	1.34	4.00	1977642	41.1679
1/0	0.286	0.98	7.01	1941699	201.9636	1/0	0.286	0.98	7.07	1944116	199.638	1/0	0.286	0.98	7.97	1977133	188.9882
1/0	0.286	0.98	6.59	1925324	206.4184	1/0	0.286	0.98	6.83	1934813	201.9789	1/0	0.286	0.98	7.94	1976348	191.5903
1/0	0.286	0.98	6.19	1909282	208.8802	1/0	0.286	0.98	6.75	1931838	204.3048	1/0	0.286	0.98	7.56	1962257	194.5356
1/0	0.286	0.98	5.88	1896143	212.4143	1/0	0.286	0.98	6.54	1923215	207.1175	1/0	0.286	0.98	7.13	1946469	197.4397
1/0	0.286	0.98	5.76	1890942	215.268	1/0	0.286	0.98	6.31	1914173	210.3233	1/0	0.286	0.98	6.92	1938428	200.8565
1/0	0.286	0.98	5.58	1883093	219.0618	1/0	0.286	0.98	6.00	1901304	214.1469	1/0	0.286	0.98	6.50	1921866	204.517
1/0	0.286	0.98	5.44	1876919	223.1225	1/0	0.286	0.98	5.64	1885879	216.39	1/0	0.286	0.98	6.24	1911357	207.89
1/0	0.286	0.98	5.12	1862461	224.5388	1/0	0.286	0.98	5.60	1884072	219.9988	1/0	0.286	0.98	6.18	1908654	210.7793
1/0	0.286	0.98	5.01	1857089	227.3134	1/0	0.286	0.98	5.47	1878391	220.9281	1/0	0.286	0.98	5.49	1879260	220.1512
1/0	0.286	0.98	4.77	1845465	230.9211	1/0	0.286	0.98	5.16	1864136	224.5219	1/0	0.286	0.98	5.23	1867232	221.9744
1/0	0.286	0.98	4.67	1840124	233.3744	1/0	0.286	0.98	4.92	1852651	228.3321	1/0	0.286	0.98	4.97	1855051	230.0868
1/0	0.286	0.98	4.52	1832337	234.9185	1/0	0.286	0.98	4.74	1843725	232.1735	1/0	0.286	0.98	4.63	1837877	232.8917
1/0	0.286	0.98	4.35	1823422	237.7903	1/0	0.286	0.98	4.49	1830484	235.1221	1/0	0.286	0.98	4.34	1822659	238.7563
1/0	0.286	0.98	4.19	1813968	240.2913	1/0	0.286	0.98	4.38	1824976	236.7835	1/0	0.286	0.98	4.00	1803066	242.8684
1/0	0.286	0.98	4.00	1803066	243.0931	1/0	0.286	0.98	4.00	1803066	242.4032	1/0	0.227	0.53	4.09	1519328	354.8453
1/0	0.227	0.53	4.00	1513744	355.2153	1/0	0.227	0.53	4.00	1513744	356.9697	1/0	0.227	0.53	4.00	1513744	355.9989

* of seasonal peak demand