Optimal reliability-based design of bulk water supply infrastructure-incorporating pumping systems

Michael Papathanasiou

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

Supervisor: Professor Jakobus E. van Zyl

Department of Civil Engineering
University of Cape Town

August 2015

Abstract

The optimal design of a bulk water supply system is centered on two major objectives: cost efficiency and the formation of a design solution that is appropriate for the conditions in which the system is to be implemented. The currently employed CSIR (2000) design guidelines utilise deterministic measures to size system components. The efficiency of following a deterministic approach to bulk water system design, involving pumping systems, was investigated. This was seen as necessary owing to the vast spectrum of influences and the interrelation of parameters that constitute a bulk water supply system.

A model developed by Chang & van Zyl (2012) sought to address this inefficiency by optimizing a bulk water supply system, with the core objectives of cost and reliability. The determination of these objectives was achieved by using a capital cost model for cost determination and a stochastic model developed by Van Zyl *et al.* (2008) for reliability. While this produced workable results, the application was relatively limited, and applied only to non-pumped, gravity-fed flow. As such, the failure mechanisms of the supply system did not include the effects of pump failure, an important influence on overall system reliability. In addition, the costing system was based solely on capital cost and did not take into account the life-cycle cost involved with the implementation of a bulk water supply system.

The investigation sought to expand the applicability of the model through the incorporation of pumping systems and life-cycle costing. It was further intended to compare the expanded model to both the model developed by Chang & van Zyl (2012) and the CSIR (2000) guidelines. A sensitivity analysis would also be performed. The expanded model that was subsequently developed was able to retain its integrity as a stochastic, optimisation model, inclusive of power transmission failure modelling, pumped flow and life-cycle costing through comparison to the Chang model. The major cost components were found to be pump infrastructure and pipeline costs (on the baseline system), in equal measure. Sensitivity to physical parameters such as static head and pipe length were found to be significant, while variation of stochastic parameters, other than pump failure frequency, was found to have little effect. The base system was found to require a higher supply ratio than stipulated by CSIR (2000), and a lower reservoir capacity. It was found that applying the design criteria

guidelines could produce vastly different reliability results, indicating that the CSIR (2000) guidelines, conventionally thought to be conservative are not necessarily conservative when the sensitivity to supply ratio is taken into account. It was recommended that a rule of thumb approach would be to design to a peak supply ratio of 2.3 AADD and 52 hours AADD reservoir capacity, as opposed to the 1.5 AADD supply ratio and 48 hours AADD reservoir capacity stipulated by CSIR (2000). The proposed approach would produce a far more robust and conservative, yet potentially inefficient system. It was found, however, that the CSIR (2000) design guidelines were not completely inefficient and that, if sensitivity to supply is taken into account, its use can produce results of acceptable reliability and with reasonable cost efficiency. This was not found by Chang & van Zyl (2012) as the range of pipe sizes and, by extension, supply ratios tested was limited and the model could therefore not optimise completely. This resulted in a higher than necessary supply ratio and lower than optimal reservoir capacity for the solutions meeting the desired failure frequency of 1 failure every 10 years, under seasonal peak demand conditions.

Reliability under load shedding conditions was found to decrease by a factor of 100 when continuous COCT (2015) Stage 2 load shedding was simulated. The reliability further decreased toward complete bulk water system failure under continuous Stage 3B/4 load shedding conditions. The converse of this finding is the substantial increase in cost required to either construct or maintain an optimally designed bulk water system at the desired reliability level of 1 failure every 10 years under seasonal peak conditions, as proposed by van Zyl *et al.* (2008), under load shedding conditions.

In addition, it was shown through the application of both USADOE (2014) - NERC (2014) and Nel (2009)-based power failure models that the developed base model has substantial applicability to systems found in real world bulk water systems. In this case, referring to the North American and South African systems, respectively, as it allows the designer to stay within his/her local legislative standards and still provide an optimal-reliability system (with the acceptance of a degree of inefficiency). The findings and developments made contribute toward the advancement of the optimal-reliability based model and the approach of optimal-reliability based design of bulk water infrastructure, including pumping systems.

Table of Contents

List of Acronyms

AADD:	Annual Average Daily Demand
CSIR:	Council for Scientific and Industrial Research
FPA:	Failures Per Annum
GA:	Genetic Algorithm
H:	Hours
H-W:	Hazen-Williams
ID:	Inside Diameter
MTTF:	Mean Time To Failure
NERC:	North American Electric Reliability Corporation
NPV:	Net Present Value
NSGA:	Non-Domination Sorted Genetic Algorithm
RC:	Reservoir Capacity
SAIDI:	System Average Interruption Duration Index
SAIFI:	System Average Interruption Frequency Index
SPD:	Seasonal Peak Demand
SR:	Supply Ratio
TTF:	Time To Failure
USADOE:	United States of America Department of Energy

List of Figures

List of Tables

Acknowledgements

To my family, for understanding the amount of time that studying part-time and working full-time takes, and for their understanding and support.

To my supervisor, Professor Kobus van Zyl, for his knowledge and encouragement as well as his understanding toward my circumstances and the frustrations of part-time study.

To Shameez, for her unwavering support and encouragement.

Finally, to Eskom, for allowing sporadic intervals of continuous electrical supply in order for the simulations to reach completion.

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

Signature Removed

Michael Papathanasiou

August 2015

1. Introduction

1.1. Background and motivation for research

Efficiency and reliability in the provision of water service is an important development objective, both globally and with specific relevance to the South African context. In general there tends to be a more significant focus on the optimisation of water distribution networks situated between the service reservoir and the consumer population. However, when one takes into consideration the ultimate goal of a successful water distribution system: to provide a reliable supply of potable water, the bulk water distribution component is an even more critical link, as its reliability determines the reliability of the entire system.

It is in the development and remediation of bulk water supply infrastructure that there is a need and an opportunity for the integration of methodologies to improve design efficiency. The current, local design procedure involves the use of deterministic guidelines to size service reservoirs and their supply pipelines. What is recommended by CSIR (2000) is to size the service reservoir to contain enough water to supply the population for 48 hours, under annual average daily demand (AADD) conditions, with additional allowance for fire-fighting capacity. The supply ratio into the service reservoir is recommended as being 1.5 times the AADD. This is conventionally seen as being a conservative methodology, as it involves the superposition of capacity required to withstand the excess demand from fire events, and interruption in supply from supply pipe breaks. In addition, the reservoir capacity acts as a buffer against peak demand from the population and interruptions in bulk supply to the reservoir. The result is a system that is expected to be resilient against a set of severe simultaneous circumstances (pipe break and fire event under high consumer demand).

This uncertainty leaves room for the development of an approach and methodology that is able to accurately simulate the conditions in which a bulk water supply system operates, and optimise the cost and reliability accordingly. A stochastic model for determining system reliability, incorporating both deterministic and probabilistic components was proposed by Nel & Haarhoff (1996).

This was further developed by Haarhoff & van Zyl (2002), Van Zyl & Haarhoff (2007) and van Zyl *et al*.(2008). The model and approach simulated a bulk water supply system under seasonal peak conditions, experiencing demand from the population, pipe failure events and fire flow demands. The reliability of the system was determined by calculating the average annual reservoir failures after simulating a sufficient number of reservoir failures to characterise the system.

The stochastic model was improved upon in terms of its computational efficiency by Chang & van Zyl (2010), through the incorporation of a compression heuristic. With the capability to accurately simulate the behavioural characteristics of a particular population and its associated phenomena (fire outbreaks, pipe breaks) in an efficient manner, what remained was to incorporate the stochastic model into an optimisation model that could provide a range of solutions of varying reliability and cost. The stochastic reliability model was incorporated by Chang & van Zyl (2012) into a multi-objective genetic optimisation algorithm, forming an optimisation model. The model was used to optimise a simple system consisting of a source, a service reservoir, a pipe network joining the source to the reservoir and a consumer population. The developed model was able to generate optimal, discrete solution sets that would allow a designer to choose a solution for implementation, based on the desired reliability level and available budget.

The model, however, was based on a simple, gravity fed supply system. The simplicity of the model allowed for reasonable assumptions to be made regarding the hydraulics, components and topology of the system. These assumptions, although simplifying, resulted in the model being restricted in its applicability to real-world bulk water supply systems involving pumped flow and capital cost.

The current advancement of the optimal reliability model is in the effort of expanding the stochastic and optimisation models to include both pumped flow and the reliability of pumping systems, and life-cycle costing. The envisaged result is an optimisation model that is applicable to a wider spectrum of real-world bulk water supply applications that can be compared to the model developed by Chang & van Zyl (2012) and the CSIR (2000)/Red Book guidelines.

1.2. Limitations and scope of investigation

This investigation was structured to expand the optimisation model developed by Chang & van Zyl (2012) by increasing the scope of its application. The investigation was directed in a manner such that the desired expansion could be validated against the Chang model, and the results compared.

The inclusion of components, such as pumps, was expected to involve the addition of an associated stochastic model to the original stochastic models (consumer demand, fire event and pipe failure event). The original stochastic models and associated assumptions, such as the seasonal peak factor (assumed to be 1.5) and demand patterns were unaltered, along with the assumed desired design criterion reliability of one failure every 10 years under seasonal peak demand as proposed by van Zyl *et al.* (2008). The definitions used in the original model for core system performance such as reliability and reservoir failure mode were left unchanged in line with the objective of developing an expanded model that is comparable to the original model. In line with this approach, the optimisation objectives of cost and reliability were left unchanged

1.3. Objectives

The primary objective of this investigation was to develop an expanded model, based on the model developed by Chang & van Zyl (2012) that would be more versatile in its application to the diverse range of conditions in which bulk water supply systems are designed, constructed and operated. The secondary objective is to investigate its feasibility as a design tool through comparison to both the original model and the CSIR (2000) design guidelines currently in effect.

These objectives are to be accomplished in the following stages:

- Compile a literature review in order that the theoretical structure behind the model is fully understood.
- Detail the operative mechanics of the original model in order that a full understanding of its functions, mechanisms and nuances can be had.

- Develop and incorporate a link to an external, dedicated hydraulic solver into the model, in order that pumped supply systems can be easily modeled and that the model can be capacitated for future development.
- Develop a comparative model that adheres to the original, gravity-fed approach, while allowing for a wider range of possible solutions through the incorporation of the hydraulic solver.
- Compare the results of the comparative model to the original, uniform pipe model to validate the accuracy of the expanded model and determine if any improvement was made.
- Develop an expanded model that includes the following expansions:
    - Network topology that removes the assumption of gravity fed flow and replaces it with pump fed flow against a static head.
    - A stochastic unit model for the pump system in terms of its susceptibility to failure and subsequent failure duration.
    - A cost model that replaces the capital cost estimation with lifecycle cost, based on the net present value method. The model is to incorporate all costs involved with a pump-fed bulk water supply system.
    - Capability to optimize within the CSIR (2000) specified design guidelines (current) and other deterministic guidelines (future research).
- Analyse the results to determine if the outcome is in line with what was expected.
- Compare the results from the expanded model against the original model developed by Chang & van Zyl (2012) and the CSIR (2000) design guidelines.
- Perform a sensitivity analysis of various physical and statistical parameters pertaining to the model, as well as example of its ability, outside of its normal application.
- Draw conclusions regarding the results obtained and the implication on the advancement of the applicability of the model.
- Provide recommendations for future investigation and further development of the model and approach.

1.4. Plan of development

A literature review is presented on topics pertinent to the development of both the original and pumped optimisation models. It covers the theoretical structure behind the stochastically modelled, genetic optimisation approach to optimal design and analysis of bulk water supply systems. It also covers the statistical backing to the stochastic backing to the pump model to be included.

The methodology is presented in detailed form, outlining the basis and processes involved in the development of the original and expanded optimisation models. The expanded optimisation model, with newly integrated functions and capabilities is adapted to mimic the original model parameters in order to verify its consistency and accuracy against the original optimisation model.

The relevant results from the expanded model are presented with analysis given in-chapter and compared to CSIR (2000) guidelines. Comparison is made between the expanded model results and results from the CSIR (2000) guidelines-optimised model. A sensitivity analysis is performed by altering the statistical and physical characteristics and parameters of the modelled system, following which the conclusions are drawn and recommendations are made for future research.

2. Literature Review

This investigation is centered around the expansion and application of the cost-reliability optimisation model proposed by Chang & van Zyl (2012). This literature review is therefore intended to document the process of acquiring a sufficient understanding of the theoretical background to the investigation at hand and a summary of the current state-of-knowledge of all pertinent contributing research topics. In addition, the literature review will show at each point the relevance, in terms of practical application, to the investigation at hand.

2.1. Design of bulk water supply and storage systems

The design of bulk water supply systems and design within the broader context of civil engineering as a whole is understandable dominated by design guidelines pertaining to the specific locality and discipline. Within a South African context, the design guidelines are almost wholly based on deterministic measures, with the majority founded on the results of empirical studies. The result is a set of rules that require the input of parameters that cannot effectively describe the context of application in sufficient detail as to provide an efficient solution. The South African design guidelines pertaining to bulk water supply design, stipulate the required storage provision for demand side parameters such as: consumer demands, storage balancing between inflows and outflows, fire demand and emergency parameters, (van Zyl *et al* , 2008). The consumer demand storage stipulated is 48 hours of Annual Average Daily Demand (AADD), plus fire-fighting reserves (CSIR, 2000). The reservoir inflow requirement is 1.5 times AADD (CSIR, 2000).

The result is a model that is expected to be conservative and is static with regards to the conditions of operation and environmental context of bulk water supply infrastructure, many parameters of which are seen to be peripheral to the operation of the bulk system. The approach of erring on the side of caution has the potential for substantial cost implications from the construction of overdesigned system components (pumps, pipes, reservoirs). However the potential exists for the construction of a system according to the guidelines, but is inappropriate to its design environment. The situation calls for an approach that is considerate of real world conditions and a method of analysis and optimisation that results in accurate results and a means of determining and comparing solutions for efficiency.

2.2. Objectives of bulk water supply infrastructure

The successful design of bulk water supply infrastructure should attend to all of the required objectives within the design context and provide a solution that is appropriate and efficient. While there are numerous objectives one may associate with a bulk water supply system, such as: water quality, water retention times, chlorine concentration, material usage and minimization of ecological disturbance and energy efficiency, amongst others. The two objectives that are given importance in orders of magnitude above the rest are reliability and ultimate cost to the prevailing water authority. The following gives a summary on the various definitions of reliability used in practice and theory and the models employed for estimating the cost of the system.

2.2.1. Defining and quantifying reliability

The overarching purpose of a bulk water supply system is to supply potable water to members of the serviced population. The challenge is thus to provide a reasonable definition of reliability that can be translated into a quantifiable measure with which to compare various systems. The current literature provides a variety of definitions for reliability. Quimpo and Shamsi (1987) define reliability as probabilistic and mission oriented, in the following manner: a system's reliability is defined as the probability that a system performs its mission under specified conditions during a specified time period. The mission or goal referred to is the previously stated purpose of supplying water to the consumer population. The conditions would be the physical and hydraulic context of the system, such as a system that is physically inclusive of all consumers in the population and supplies water that is of an adequate pressure head. The specified time would be the time frame under which the reliability is being tested.

A similar definition is given by Cullinane *et al* (1992), as cited in Chang & van Zyl (2012). The definition of reliability as applied to any purpose driven engineering system is the probability that a system performs its mission within specified limits for a given period of time within a specified environment. The influence of probability on design can thus be seen as being of significant importance, as will be emphasized further on in the literature review.
The above mentioned definitions pertain to engineering systems as a whole. When one considers the explicit purpose of a water distribution network, a proposition as outlined in Farmani *et al.* (2005) is that reliability is the capability of the distribution network to provide

the population with sufficient, high quality supply under normal and abnormal conditions. Normal being interpreted as supply under demand from the population only, and abnormal referring to supply under extenuating conditions (heightened demand and reduced supply). By extension, the reliability of a bulk water supply system has the same definition as it is a mechanism by which the distribution system is fed. Thus a failure of the bulk water supply system is a failure of the water supply and distribution network as a whole.

The reliability can thus be defined in terms of its susceptibility to failure or the lack thereof. What is required is a practicable definition of failure. The most applicable failure mode when considering the primary purpose of a bulk water distribution system to feed a water distribution network is hydraulic failure. Hydraulic failure is defined by Cullinane *et al* (1992) to occur when the system cannot supply the specified amount of water to the specified location, at the specified time, and at the specified pressure. This can be interpreted, in the context of bulk water systems as happening when the reservoir physically runs dry. This was emphasized and employed by Chang & van Zyl (2010). In addition the failure duration, determined as a property of the failure was defined as being the duration from which the reservoir is empty until a net inflow is experienced (Chang & van Zyl, 2010).

With the definition of reliability being accepted for this investigation as being in terms of a system's resilience to failure and with an applicable definition of failure, what remains is to quantify the tolerance for failure. Khomsi *et al.* (1996) states that repair time is important when quantifying reliability. It is seen as reasonable to translate that into the duration of failure, as both refer to a period of time where the consumer population experiences an interruption in supply. An acceptable degree of reliability is proposed by Van Zyl *et al.* (2008) as being one failure of the reservoir every 10 years under seasonal peak conditions. Further proposed were the results from a study that showed that consumers are less sensitive to the duration of failure as compared to the frequency of failure. It is on this basis that the failure duration will be investigated, but will not be considered as a means of quantifying reliability and comparing solutions. In the current investigation, reliability is therefore quantified as the inverse of failure frequency.

2.2.2. Costing and cost models

Bulk water system cost should not be viewed only as a requirement in terms of the budgetary constraint of the developer or prevailing water authority. When one considers that every possible solution system for a given application has to adhere to the same cost model, what can be seen is that cost is not just a measure of monetary requirement, it is a measure of system and operational efficiency. Thus relative cost efficiency between solutions can be considered as synonymous to relative system efficiency. This point is emphasized by Barta & Rowse (1998) and Swamee & Sharma (2008).

There are various methods that are currently utilized in order to provide an accurate cost estimate. Methods such as capitalization, annuity and Net Present Value (NPV) are employed in cost estimation. The appropriate method should be chosen based on the conditions and type of development and/or remediation required. When considering a gravity fed system devoid of significant energy costs, it can be seen to be appropriate to use the capitalization method or simply consider the current capital cost. The approach of limiting the scope to capital cost was employed by Chang & van Zyl (2012) and in certain real world applications such as the Hanoi network (Fujiwara & Khang, 1990, as cited in Chang & van Zyl, 2012) and the widely known problem of the New York tunnels (Schaake & lai,1969, as cited in Chang & van Zyl, 2012).

However, a substantial number of bulk water supply systems require the use of a pumping system to convey water from the source to the reservoir, owing to a difference in elevation and/or long distances. This added dimension renders the assumption that only current capital cost need be considered, unfeasible. The recurring energy and maintenance costs need to be evaluated and included in order to provide a realistic and comparable cost estimate. The cost needs to be calculated using an appropriate application of the interaction between inflation rate and interest rate. The NPV method takes into account the effect of discounting the nominal amount due each interval through the discount rate (Swamee & Sharma, 2008). The significance of this is highlighted when considering recurring energy costs, when comparing systems, an accurate effective cost is necessary to achieve cost efficiency. For example, a larger pump and smaller reservoir may be selected over the inverse arrangement when one considers the effect of the discount rate after each payment interval.

When comparing systems with relatively large energy costs, lifecycle costing is recommended in order to accurately conceptualise the total cost of the project as opposed to only considering the initial capital cost. This is effectively demonstrated in Figure 2.2.1.

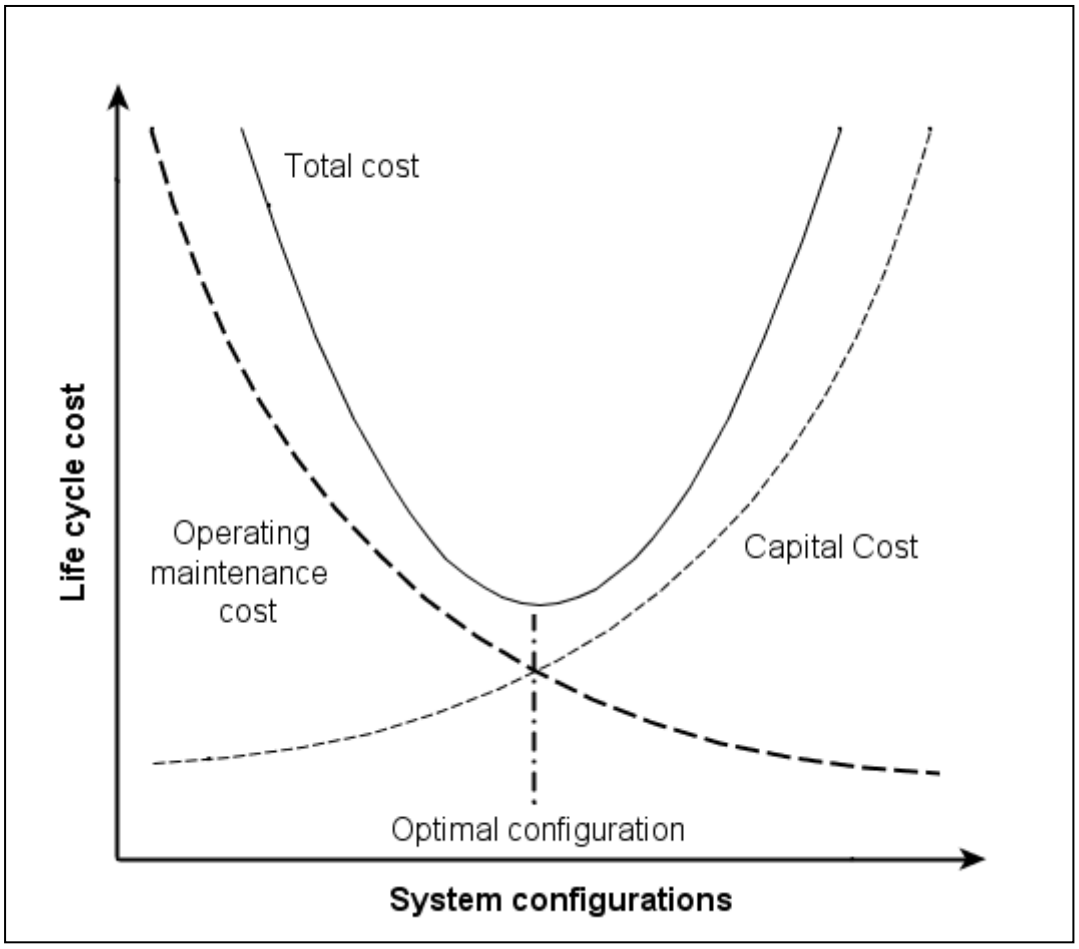

Figure 2.2.1: Variation of total cost with system configuration (Swamee& Sharma, 2008)

2.3. Stochastic analysis

The word stochastic is rooted in the Greek word 'stokhos', meaning an aim or a guess. This gives an indication as to the element of randomness in stochastic processes and stochastic analysis. Stochastic analysis is an analytic technique whereby the deterministic and probabilistic parameters of a system are simulated to model the system's behavior more accurately (Chang & van Zyl, 2010).

2.3.1. Stochastic analysis within science and engineering

Traditionally, stochastic analysis has been used in scientific disciplines that involve the modeling and analysis of exceptionally complex processes. Many of these complex processes involve influences from an incomputable number of sources and parameters. This forces the analyst to differentiate between that which is computationally determinate and that which requires a probabilistic/random approach.

It has thus found its way into disciplines such as advanced biochemistry and nuclear physics. In engineering it is prominent in the fields of hydrology, telecommunications and electrical distribution systems due to the assumed elements of randomness in meteorological and human behaviour.

One of the most widely known and fundamental applications of stochastic process is Brownian motion. It was studied by Albert Einstein in 1905 in the context of a theory describing the irregular movement of pollen immersed in water. The irregular motion was first noticed by a botanist R. Brown in 1824. The mathematical theory was further developed by N. Wiener (1923) and P. Levy (1948). Their research inspired the majority of research into stochastic process and the process known as the *Wiener-Levy* process is still used in engineering, economics, finance and mathematical biology amongst others, (Karatzas, 1988).

2.3.2. Stochastic analysis in bulk water infrastructure

Within the context of the reliability analysis of bulk water infrastructure, the suitability of stochastic analysis is emphasized by Nel (2009), who states that many of the variables involved possess probabilistic characteristics. The failure mode and failure characteristic of a bulk water supply system are random in nature (Quimpo & Shamsi, 1987) thus classifying bulk water and water distribution systems as inherent stochastic systems.

Work on applying stochastic process to the analysis of bulk and water distribution systems was carried out by Yang *et al* (1996) and Ostfeld *et al* (2001) amongst other researchers. This work was further developed by Nel (1993), Haarhoff & van Zyl (2002) and Van Zyl *et al* (2008) with the addition of stochastic models describing consumer water demand, fire occurrences and fire flow demands and reservoir supply pipe failure events.

The results from stochastic modeling of bulk water systems are intended to give an accurate representation of real world influences on the various components within a system. In the absence of a relatively isolated real-world test population where regular and exhaustive field measurements can be taken in order to calibrate the model, the results are predictable only in a statistical sense (Van Zyl *et al* 2008).

2.3.3. Simulation and sampling techniques

The reliability analysis of bulk water systems is categorised by (Xu & Powell, 1991, as cited by Chang & van Zyl, 2010), to include analytical methods, heuristic methods and Monte Carlo simulation.

Monte Carlo simulation is derived from the area in the principality of Monaco, which is widely known for its casinos. The method which draws multiple random samplings within specified boundaries and analyses them based on a preconceived set of rules. One can envision how this could be applied to blackjack, testing the success rate of a particular strategy over millions of games.

The appropriateness of Monte Carlo simulation is in its random approach to a process involving random characteristics. It is particularly effective when the underlying

probabilities/laws are known, but the results cannot be analytically solved (Huber, 1997). Zio (2009), as cited in Chang & van Zyl (2010), postulated 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 analyses.

This is not to assume that Monte Carlo simulation is without its disadvantages. As previously mentioned, Monte Carlo simulation involves the drawing of multiple random samples over a large number of iterations. Depending on the scale of the analysis and the accuracy required, the number of samples drawn/iterations performed can be in the order of billions. This gives an indication as to the computational demand involved.

As previously mentioned, Van Zyl *et al* (2008) developed and demonstrated the application of the stochastic model consisting of various unit models used to describe components of the systems and its influencing factors. A more detailed description of the mechanics of this approach will be given further on in the literature review.

Other sampling techniques include: Latin Hypercube sampling, Descriptive sampling and importance sampling. Latin Hypercube sampling uses a technique that ensures that all parts of the input probability distributions (for example, the frequency of occurrence of a fire event) are sampled (McKay *et al.* 1979). The advantage is that the entire distribution range is represented, which should require fewer samples to be drawn in order to represent the modelled system (Chang & van Zyl, 2012). This method was investigated by Chang & van Zyl (2012) with the result being a marginal increase in computational effort on the chosen platform.

Descriptive sampling and Importance sampling as summarised by Chang & van Zyl (2012) are able to decrease the computational effort of stochastic analysis by being computationally more efficient than Monte Carlo method. Descriptive sampling uses a deterministic set with a random sequence of samples being drawn, while the Monte Carlo method uses a random set and random sequence. Importance sampling assigns an assumed importance rating to each probabilistic component. However both methods either lack theoretical development or require a large deal to be known about the system pre-analysis.

2.4. Application of the stochastic analysis process

As with any form of modeling, the purpose of applying stochastic process to a bulk water supply system is to create a bounded space in which all pertinent aspects are reasonably identical to real world conditions. This simulated world is observed over an extended period of time in order to draw conclusions as to the reliability of the imposed system, as proposed by Van Zyl *et al.* (2008).

The reliability of a reservoir in terms of failure occurrence was found by Van Zyl, *et al*. (2008) to be approximately 17 orders of magnitude more reliable during the lower demand season (winter/rainy season) Thus the decision was made to concentrate on the most critical period of the year, more specifically, the most critical week of the year, under seasonal peak conditions. To restate the design tolerance earlier mentioned, one failure every ten years during this critical period was seen as an acceptable design criterion reliability level.

2.4.1. Integration of stochastic unit models

Van Zyl *et al.* (2008) developed stochastic unit models in order to build a holistic model that would be representative of the dynamic conditions in which a bulk water supply system exists in reality and where a theoretical system could be imposed. As previously mentioned, these models were designed to represent consumer demand, fire demand and pipe failure within the holistic model.

The mathematical and statistical components of these models were summarized by Chang & van Zyl (2010) as follows:

Consumer Demand

The task of accurately representing both the quantity of water used and the manner in which it would be used by the serviced population is difficult as there are many influencing factors. Van Zyl *et al.* (2008) recognized the influence of conditional factors such as the socio-economic situation of the users, climatic factors such as rainfall patterns and structural factors, including water metering stringency and size of the serviced population.

The consumer demand unit model developed by Van Zyl *et al.* (2008) consists of four components: average demand, cyclic patterns, persistence and randomness.

Average demand refers to the average demand over the modeled period or possibly over a typical year in which the most critical period is being modeled (AADD). Cyclic patterns increase in resolution from seasonal, monthly, weekly and diurnal patterns all the way down to the fluctuations from hour to hour. The model under investigation (Van Zyl *et al.* 2008) uses daily and hourly demand factors, as the modeled period is the most critical week, implying no variation in any factor pertaining to or of higher order than the weekly factor.

Demand persistence pertains to the tendency of every unit time step, in this case an hourly or daily time step, to be influenced by the preceding time step and subsequently influence the following time step, but to a lesser degree. To give an example, if the serviced population experienced a heat wave, the water usage following the cessation of the heat wave would still be elevated above normal due to the lag in response to the now reduced-to-normal temperature. Pools would still be refilling, water demanding air-conditioning units still set to use more water. The same influence-lag effect is experienced on an hourly basis. This effect was modeled by Van Zyl *et al.* (2008) in the form of autocorrelation coefficients.

If one had to visualize a graph of the total water demand of a given population over a year and sequentially remove the patterns pertaining to the seasonal, monthly, weekly, diurnal, hourly and finally persistent patterns, what would be left would be a pattern that by observation appears completely random. This random component was modeled by Van Zyl *et al.* (2008) as a white noise function. This white noise component is modeled by a statistical distribution function with mean 0 and variance$\sigma^2{}_D$.

The model itself is divided into two major constituents, the daily demand functions and hourly demand function that are partially dependent on the daily demand parameters. These are described in the equations to follow.

As mentioned earlier, the statistical and mathematical components as developed by Van Zyl *et al.* (2008) were summarized by Chang & van Zyl (2010) as follows:

$$D_d = D_{ave} C_{dow} \upsilon_d \tag{1.1}$$

where D_d = average demand in day

D_{ave} = average demand for the period studied

C_{dow} = day-of-week demand factor

υ_d = daily demand residual function

The daily demand residual function, υ_d, is described by

$$\ln \upsilon_d = \sum_{i=1}^{m} \phi_i \ln \upsilon_{d-i} + \ln \varepsilon_d, \qquad \ln \varepsilon_d \sim IN(0, \sigma_D^2) \tag{1.2}$$

where i = lag counter

m = number of daily autocorrelation lags

ϕ_i = daily autoregression coefficient for lag i

$\ln \varepsilon_d$ = white-noise process

The notation $\ln \varepsilon_d \sim IN(0, \sigma_D^2)$ indicates that the natural logarithms of the residuals are normally and independently distributed with mean 0 and variance σ_D^2.

A similar model is used for the hourly demand variation:

$$D_h = D_d C_h \upsilon_h \tag{1.3}$$

where D_h = average demand for hour

D_d = average demand for the current day

C_h = hourly demand factor

υ_h = hourly demand residual function

The hourly demand residual function is described with a similar equation as the daily demand residuals, as shown in equation (1.3).

Fire Demand

The primary source of water used to extinguish fires is municipally supplied water from the water distribution mains. It is widely understood within the design of municipal water distribution systems that the critical design factor is the flow required under fire outbreak conditions. Thus the significance of accurately modelling the fire flow becomes evident as it is the condition under which the reservoir will experience the greatest demand stress.

The important characteristics of fire demand and fire events were recognised by van Zyl *et al* (2008), these were: fire event frequency, fire event duration and required fire demand flow. The statistical and mathematical mechanics as developed by van Zyl *et al.* (2008) were summarized by Chang & van Zyl (2010), as follows: the fire frequency or occurrence of fire events is modeled using Poisson process with rate parameter λ, is described by:

$$P\big[N(t+\tau) = N(t)\big) = k] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!},\ k = 0,1,\ldots \qquad (2.1)$$

where P = probability of occurrence

t = time

$N(t+\tau) - N(t)$ = number of events in time period τ

k = number of occurrences

e = base of the natural logarithm

The expected number of events in time interval $(t, t+\tau]$ is given by

$$E[N(t+\tau) - N(t)] = \lambda\tau \qquad (2.2)$$

Fire events were modelled by simulating times between successive fires, which follows an exponential distribution. The probability density function of an exponential distribution is given by

$$f(\Delta t;\lambda) = \lambda e^{-\lambda t},\ \Delta t \geq 0 \qquad (2.3)$$

where Δt = time between successive fire events

Fires occurring before the end of the current fire were ignored.

Once a fire has occurred, the fire duration and fire flow are both estimated using log-normal distributions. The probability distribution function for a log-normal distribution, for the fire duration, is given by

$$f(T;\mu,\sigma) = \frac{1}{T\sigma\sqrt{2\pi}} e^{-(\ln T-\mu)^2/2\sigma^2},\ T \geq 0 \qquad (2.4)$$

where T = duration of the fire event

μ = mean of the logs of the durations

σ = standard deviation of the logs of the durations

Pipe Failure

The success and reliability of a bulk water distribution system do not rely solely on the size of the service reservoir. To better visualise the relationship between service reservoir capacity, pipe diameter (equated to reservoir inflow capacity) and reliability, first picture a large capacity reservoir with a small diameter feeder pipe. What follows from this is a system that has a substantial ability to handle large scale fire and pipe break events for a short duration.

However, should the reservoir experience a prolonged fire event, it will experience a dramatic reduction in reservoir level and possibly run dry (fail). This arrangement is also vulnerable to high frequency failures as a low relative inflow implies a large lag time to refill to a safe level. If the inverse is considered, the reservoir is resilient to longer duration, milder intensity events but not to high intensity events due to the smaller capacity to buffer resulting from a smaller volume.

The above mentioned point is emphasized by van Zyl *et al* (2008), in the importance stressed upon the capability of the source to reservoir pipe to provide a constant and uninterrupted flow.

There are many factors that influence the reliability of a pipeline. Nel (2009) demonstrated and presented findings on the effect of material, pipe length, soil type and conditions on the susceptibility of a pipe to failure. A critical analysis of pipe conditions is necessary, particularly in dolomitic areas, where sinkhole formation is possible.

Van Zyl *et al* (2008) identified four components that constitute a set of statistical characteristics that can be used to model the behaviour of supply pipe systems under failure conditions. The four components are: occurrence, duration, clustering and severity. The same approach as that applied to fire events was used in that a Poisson process was employed to model the occurrence of the events. Duration was once again simulated through a log-normal distribution. Clustering was disregarded as it was seen as being of less that significant importance Van Zyl *et al* (2008). The severity of a pipeline failure was determined by its impact on the feeder flow into the reservoir.

Pump Failure

Simple, gravity-fed models only consist of the above mentioned 3 unit models, however when the source is not at a position of higher gravitational potential energy (higher elevation) in relation to the service reservoir, the employment of a pumping system becomes a necessity. Pumping systems are also necessary when long length, horizontal feeder pipelines are used.

Cullinane (1985), as cited in Nel (2009) identified various parameters that were seen as pertinent with respect to the overall reliability of a pump unit.

$$R_S = R_P * R_M * R_C * R_{PT} * (R_V)^2 \qquad (2.5)$$

Where: R_S = reliability of the pumping system

R_P = reliability of the pump

R_M = reliability of the motor

R_C = reliability of the control unit

R_{PT} = reliability of the power transmission

R_V = reliability of the valves (1 intake and 1 delivery valve)

Nel (2009) stresses the extremely high relative importance of power transmission/supply on the overall reliability of the pumping unit as a whole. The operative assumption is that it is likely that the pumping system is subjected to regular maintenance of the mechanical (pump, motor, valves) and electronic (control unit), and that system redundancies exist within the pumping system to mitigate the effect of a failure of any component contained within the pump station. However, complete backup power supply to the pump system is not a commonly found occurrence, and the majority of pump systems become inactive during a failure of the power transmission grid, this is confirmed by data produced by Rand Water (South Africa), and categorised by Mbula (2008) as comprising external trips. External trips are defined by Fredericks *et al* (2007), as cited by Nel (2009), as resulting in neither the duty or standby pumps being operational, which is interpreted as a complete failure of the pump stations pumping ability. The reliability of the power transmission component is therefore strongly emphasized, to the point where all other factors influencing reliability are assumed to be close to 1. The reliability of the pumping system is simplified to:

$$R_S = R_{PT} * \quad (2.5a)$$

Where: R_S = reliability of the pumping system

R_{PT} = reliability of the power transmission

The reliability of the bulk water systems to be investigated is defined, as outlined in 2.2.1, as the inverse of failure frequency. This definition is extended to the reliability of the power transmission to the pump station. The distribution of power supply failure occurrence and the associated power supply duration were found by Nel (2009) to follow a lognormal distribution and, based on the average failure rates can be used to model the power supply failure occurrences and power supply failure duration in a similar manner as that employed to calculate pipe failures and fire events. The time between power supply failure events is calculated using a Poisson process, and the duration is generated using the inverse of the cumulative distribution function and normally distributed, generated random numbers.

A cumulative distribution is used to model power failure duration. An example of such a distribution, determined by Nel (2009), is presented in Figure 2.4.1. The power supply failures described by the lognormal distribution as presented by Nel (2009) and again by Nel

& Haarhoff (2011) are based on the failure of supply to 7 Rand Water pump failure installations, as collated by Mbula (2008). The lognormal mean in Figure 2.4.1 is -0.61 with a standard deviation of 1.54. This translates to an average duration of 1.6 hours per failure.

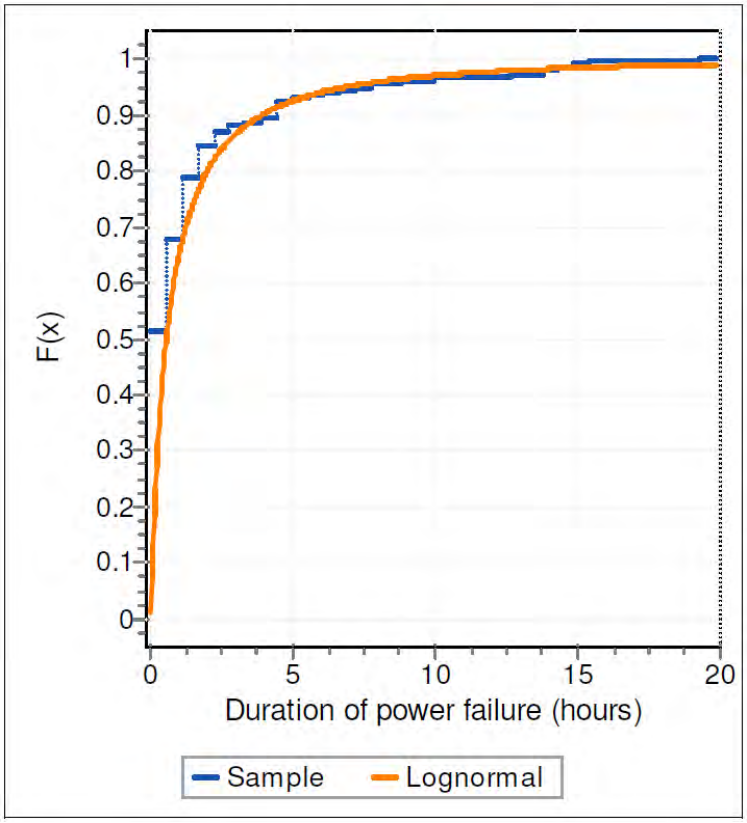

Figure 2.4.1: Cumulative distribution of power failure duration (Nel, 2009)

The distribution of annual failure frequency was also found to follow a lognormal distribution, and a cumulative distribution function was fitted by Nel (2009), in Figure 2.4.2. The lognormal mean in Figure 2.4.2 is 2.20 with a standard deviation of 0.70. The result is a failure frequency of 11.4 failures per annum.

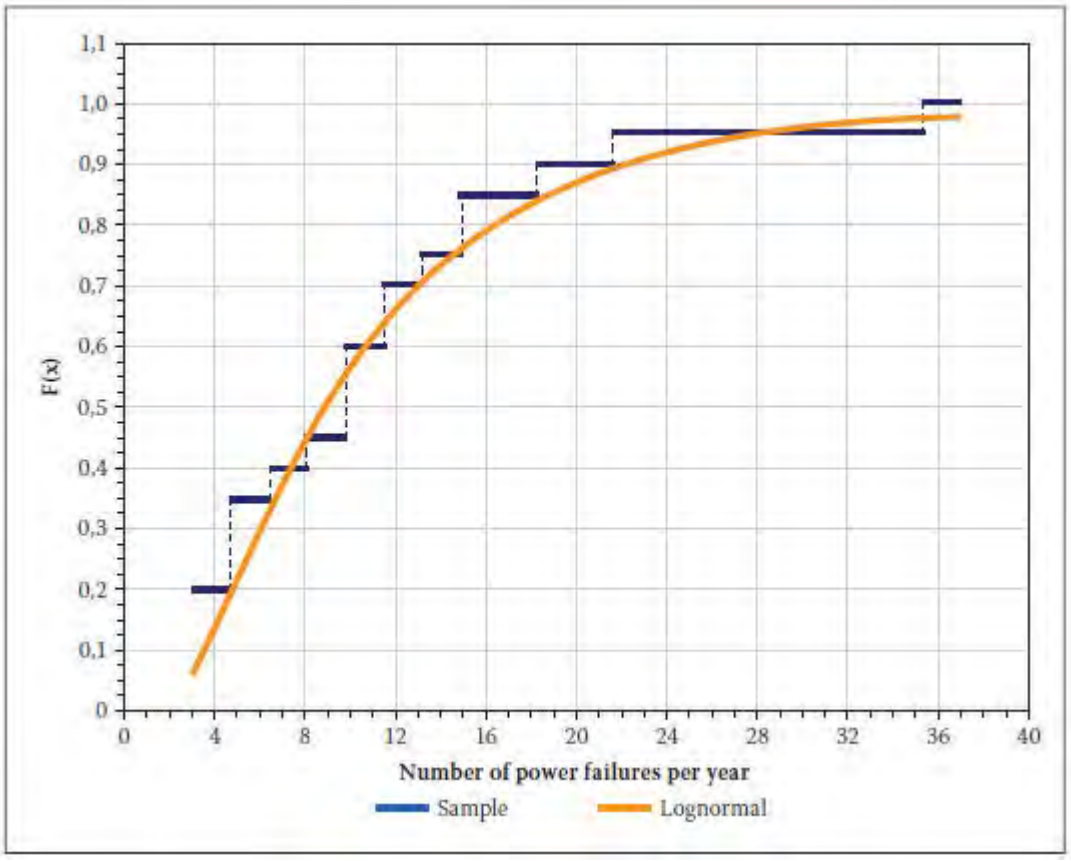

Figure 2.4.2: Cumulative distribution of power failure frequency (Nel, 2009)

The stochastic unit models outlined above allow for random numbers to be drawn and applied to statistical distributions to generate events that are representative of real world conditions.

As stated above, the distributions attributed to power supply failure frequency and duration can be used to stochastically model power supply failures within a South African power supply context. However, the distribution of power supply failure frequency as well as the associated duration of failure is seen to be vastly different between countries at varying stages of development.

This is shown in the System Average Interruption Duration Index (SAIDI) and System Average Interruption Frequency Index (SAIFI), presented by Ramakrishna *et al* (2009) and Eskom (2007), as cited by Nel & Haarhoff (2011). The SAIDA and SAIFA indices translate to the severity of occurrences and associated total duration of power supply failures over the course of a year for a given consumer; this is represented in equation 2.6 and 2.7.

$$SAIDI = \frac{\sum_{i=1}^{l} K_i D_i}{K} \quad (2.6)$$

$$SAIFI = \frac{\sum_{i=1}^{l} K_i}{K} \quad (2.7)$$

Where: i = an interruption event

D_i = restoration time for each interruption event

K_i = number of interrupted households for each sustained interruption event

K = total number of households served for the area considered

The stage of development of power transmission within a country is evident in its SAIDI and SAIFI indices. The basic stages of development of a country can therefore be associated with specific indices. The following index vales are presented in Table 2.4.1, along with the assumed stage of civil infrastructure development. A SAIDI of 120 minutes per annum, implies that the average consumer will experience 120 minutes of electrical supply downtime per year. A SAIFI of 1.26 implies that the average consumer will experience 1.26 electrical supply interruptions per year. A country with better developed infrastructure will suffer fewer electrical supply interruptions and for a shorter duration. Well developed electrical generation and transmission infrastructure will manifest in lower SAIDI and SAIFI indices.

Table 2.4.1: SAIDI - SAIFI Indices (after Ramakrishna *et al* (2009)

Country	Development Status	SAIDI (minutes/a)	SAIFI (number/a)
Netherlands	Highly Developed	20	0.23
USA	Developed	120	1.26
South Africa	Developing	3084	25.2
India	Underdeveloped	100 800	40

For this study, the South African and USA electrical supply condition were chosen to form part of the stochastic modelling of electrical supply failure. The USA conditions were used for the base model, to demonstrate the effectiveness of the application of optimal reliability-based design in a country that is considered as developed, but still experiences significant electrical transmission disturbances. The South African conditions were chosen owing to the

significant development still required, but where it was deemed possible that the approach could be implemented by local authorities.

For stochastic modelling purposes, the SAIDI and SAIFI indices cannot be used directly, as they do not represent the distribution of power failure occurrence and duration. With the lognormal distribution obtained from Nel (2009) and Nel & Haarhoff (2011), the statistical distribution for power supply failures in the USA remained.

This information is obtained from the Electrical Disturbance Event (EDE) summaries, compiled by the U.S Department of Energy (USADOE, 2014), for all recorded annual disturbances between 2003 and 2014. This information included the time and duration for each failure, as well as the number of affected customers and the North American Electric Reliability Corporation (NERC) region that each disturbance (failure) occurred in.

The NERC regions are as per Table 2.4.2 and represent defined zones within the United States of America that are monitored and analysed for reliability and function. These zones are used for area-specific infrastructure planning and strategic development.

Table 2.4.2: North American Electric Reliability Corporation Regions

FRCC	Florida Reliability Coordinating Council
MRO	Midwest Reliability Organisation
NPCC	Northeast Power Coordinating Council
PJM	Pennsylvania, Jersey, Maryland Power Pool
SERC	Southeast Electric Reliability Council
SPP	Southwest Power Pool
TRE	Texas Reliability Entity
WECC	Western Energy Coordination Council

The total number of failures per zone, per year for 2003 until 2014 (only 2008 - 2014 to be shown), were tallied and organised from the USADOE (2014) data, presented in Table 2.4.3.

Table 2.4.3: Power failures per year per NERC region

ZONE	2014	2013	2012	2011	2010	2009	2008
FRCC	0	3	1	3	1	1	8
MRO	14	8	3	4	6	1	3
NPCC	5	20	27	44	9	3	12
PJM	35	46	74	94	44	34	51
SERC	31	21	24	46	17	25	30
SPP	0	7	8	21	8	3	4
TRE	11	5	8	15	5	11	9
WECC	32	59	49	77	33	19	26

However, as the base assumption is that the power transmission feeding the bulk pumping system is affected by transmission failures that are experienced by the general population, the number of affected customers per zone and total customers per zone were determined as per USADOE (2014) and NERC (2014) data. The population figures were offset annually by the annual population growth factors published by the World Bank (2014).

The result is a matrix of annual supply failure probability per household over the duration of a year (where insufficient data was available, the values were removed to avoid skewing the distribution). This is presented in Table 2.4.4.

Table 2.4.4: Probability of power failure per year per NERC region

ZONE	2014	2013	2012	2011	2010	2009	2008
FRCC		0.00219		0.00336		0.00201	0.01191
MRO	0.02173	0.03271		0.00659	0.01341		0.03604
NPCC		0.00132	0.00397	0.00246	0.00431		0.00462
PJM	0.00289	0.00166	0.00340	0.00198	0.00227	0.00195	0.00294
SERC	0.00448	0.00311	0.00400	0.00368	0.00135	0.00279	0.00379
SPP		0.01097	0.01700	0.00666	0.00673		0.01140
TRE			0.00412	0.00581	0.01009	0.00906	0.03056
WECC		0.00052	0.00036	0.00094	0.00188	0.00198	0.00243

This was used in conjunction with the annual power failure statistics from Table 2.4.3, to generate the number of failures that could be experienced by a generic household, and by extension a bulk pumping installation, per annum. This is demonstrated in Table 2.4.5.

Table 2.4.5: Power failures per household per year per NERC region

ZONE	2014	2013	2012	2011	2010	2009	2008
FRCC	-	0.00658	-	0.01008	-	0.00201	0.09532
MRO	0.30426	0.26170	-	0.02638	0.08043	-	0.10812
NPCC	-	0.02634	0.10717	0.10808	0.03876	-	0.05547
PJM	0.10098	0.07633	0.25164	0.18575	0.10001	0.06631	0.14990
SERC	0.13892	0.06538	0.09611	0.16906	0.02299	0.06983	0.11381
SPP	-	0.07682	0.13599	0.13984	0.05388	-	0.04560
TRE	-	-	0.03293	0.08710	0.05046	0.09969	0.27507
WECC	-	0.03082	0.01779	0.07242	0.06219	0.03761	0.06314

These values were used to fit a lognormal cumulative distribution function for all of the NERC regions combined. Instead of drawing data from a single NERC region to be used as the base system, data from all NERC regions was used as power failures to be generated were to be representative of all regions of the USA. The lognormal cumulative distribution function was fitted to data representing power supply failures across the entire USA. The lognormal cumulative distribution function is presented in Figure 2.4.3, for 2003 to 2014.

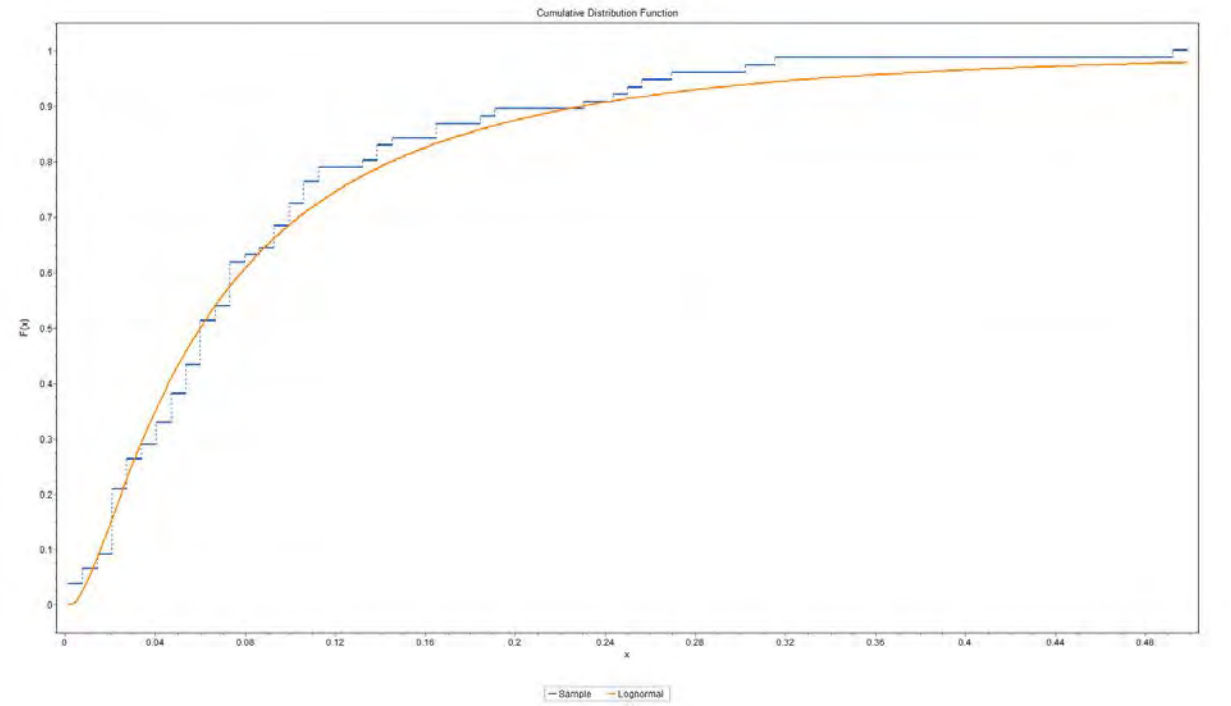

Figure 2.4.3: Cumulative distribution function of failure frequency - NERC regions

The result was a well-matched lognormal cumulative distribution function with $\mu = -2.81$ and $\sigma = 1.05$. As the power (and thus pump) failures are determined in a Time To Failure (TTF) manner, on an hourly basis, the mean value derived from this distribution is 0.104 failures per annum, or 1 failure every 9.6 years. This value is reduced to an hourly Poisson mean to be used in the stochastic model.

To generate the lognormal cumulative distribution function for the power failure duration, approximately 1500 failure event durations were used as input. The resultant cumulative distribution function and statistical parameters are presented in Figure 2.4.4 and Table 2.4.6.

The developed US statistical model for failure as defined by data from NERC (2014) and USADOE (2014) is to be used as the baseline model for simulations, with comparative simulations to be run for the Nel (2009) /Nel & Haarhoff (2011) statistical distributions. The cumulative distribution function for power failure duration is presented in Figure 2.4.4.

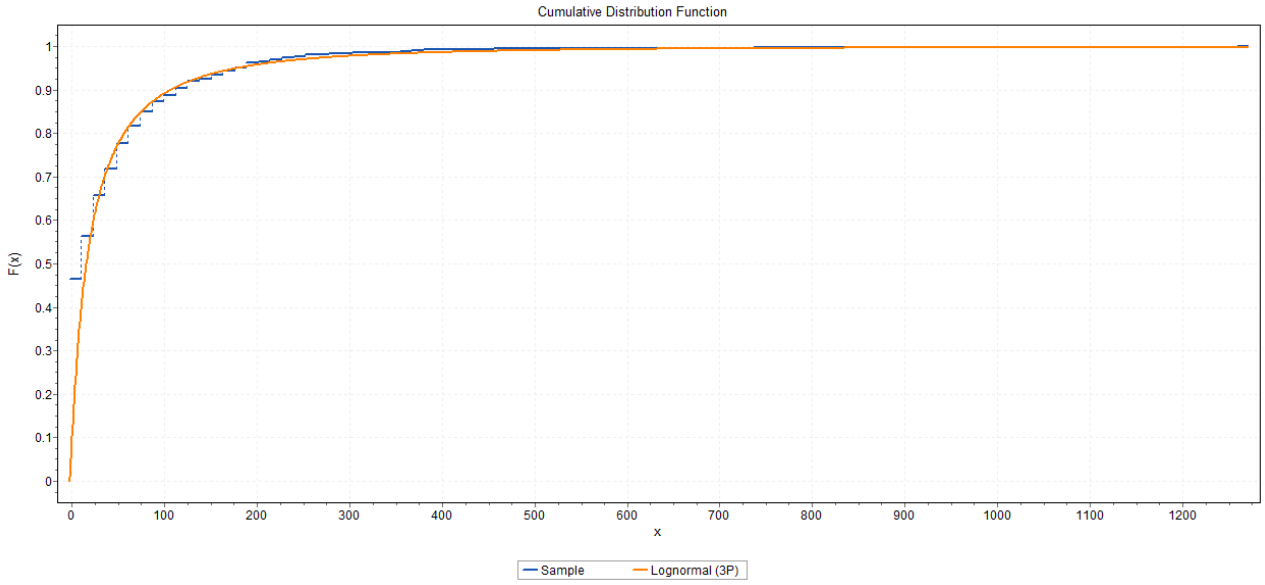

Figure 2.4.4: Cumulative distribution function of failure duration - NERC regions

The failure durations experienced were large, with $\mu = 2.92$ and $\sigma = 1.38$, giving an average of 48 hours per event. The resulting statistical distributions to be used in the base model and South African comparison are presented in Table 2.4.6.

Table 2.4.6: Power failure statistical distributions

		Parameter	Value
BASE MODEL	**NERC (2014) & USADOE (2014) US**	**Mean frequency (failures/annum)**	**0.104**
		Standard deviation (σ)	1.05
		Lognormal mean (μ)	2.92
		Mean duration (hours)	**48**
		Standard deviation (σ)	1.38
		Lognormal mean (μ)	2.92
SA COMPARISON	**NEL (2009) SA**	**Mean frequency (failures/annum)**	**11.4**
		Standard deviation (σ)	0.7
		Lognormal mean (μ)	2.2
		Mean duration (hours)	**1.6**
		Standard deviation (σ)	1.54
		Lognormal mean (μ)	-0.61

The difference in both frequency and duration between the models shows the contrast between developed and developing countries' power generation and transmission climates.

2.5. Optimisation

As mentioned previously, this investigation deals with the expansion of the optimisation model proposed and demonstrated by Chang & van Zyl (2012) to include pumping systems. This section will focus on a description and basic review of the optimisation method employed by Chang & van Zyl (2012). A more mechanistic approach, detailing the changes and additions made, will be given in the methodology.

2.5.1. Objectives of optimisation

Any optimisation technique has the goal of finding the most efficient solution within the specified constraints. As previously mentioned, the two most prominent factors, when designing a bulk water supply system, are the reliability of the system (represented as the inverse of failure frequency) and the total financial cost of the system. The pareto-optimal obejective is the minimisation of both cost and failure frequency.

The failure frequency tolerance proposed by van Zyl *et al.* (2008) of 1 failure in 10 years under seasonal peak conditions, suggests that using cost as the sole objective allows for the possibility that the system is cost efficient but not sufficiently reliable for the proposed development. This possibility demonstrates that multi-objective optimisation techniques are important in order to maximize the net benefit (in this case, in terms of reliability), defined as the benefit minus cost (Walski, 2001).

2.5.2. Genetic optimisation

When considering the various solutions for a given bulk water supply problem, the number of unique solutions available for a reasonably sized system are, by orders of magnitude, too large to realistically calculate the cost and reliability of each solution. This is especially true when considering the stochastic nature of determining reliability and the large associated computational load for a single solution. This necessitates an optimisation technique that is representative of the entire solution space and can sample and refine solutions without having to analyse a large number of unique solutions.

As the name suggests, a genetic algorithm is an optimisation technique that is based on the mechanics of natural selection and genetics (Holland, 1975; Goldberg, 1989 as cited by Chang & van Zyl, 2012). 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, as cited by Chang & van Zyl, 2012)

Genetic Algorithm: NSGA-II

As the genetic algorithm employed in the model suggested by Chang & van Zyl (2012) is a non-dominated sorting genetic algorithm (NSGA), the scope of the review will be limited as such. The basic operation of the NSGA-II algorithm is as follows:

The population size, decision variables, decision variable boundaries, number of generations, objectives and other parameters specific to the optimisation of the algorithm are inputted. A population is initialized by generating randomly within the solution space defined and bounded by the decision variables' limits. The decision variables in this application are the system components (In this instance, pump power, number of pipes, pipe size, reservoir size etc.). The population consists of a set of chromosomes representing a certain assortment of genes (decision variables).

The population is evaluated in terms of fitness; each chromosome is evaluated in terms of the objectives set (cost and reliability, in this instance). The fitness of each unique solution (chromosome) is translated into a probability for selection, with fitter solutions having a higher probability of being selected and recombined into the solution population.

Each solution set is sorted in the non-domination sorting algorithm. The algorithm is based on the NSGA definition of domination and involves ranking, sorting into domination fronts and assigning a crowding distance to each solution. The crowding distance value is based on the position in the front and is used to ensure that solutions do not coagulate around a particular solution subspace.

The sorting algorithm is defined below by Liong *et al*. (2004) and summarized by Chang & van Zyl (2012):

If one considers 2 unique solutions (chromosomes) within the initialized population: $\mathbf{x}_{(1)}$ & $\mathbf{x}_{(2)}$,

i. $\mathbf{x}_{(1)}$ 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;
ii. $\mathbf{x}_{(2)}$ 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;
iii. Otherwise, $\mathbf{x}_{(1)}$ & $\mathbf{x}_{(2)}$ are equally good.

This process is repeated until all chromosomes are ranked by domination. Crowding distance will be demonstrated later in the methodology.

The population is entered into the evolution process. The parents are selected based on the fitness probability and the reproduction process starts. The crossover process, which in NSGA-II is the Simulated Binary Crossover (SBX), ensures that two solutions with fit decision variables (genes) are combined to form a solution with a high probability to have a better fitness than the parents. To maintain diversity within the population, a certain probability exists with the possibility for a mutation to occur (±10%) of a decision variable that has a low probability for selection, i.e. an unfit chromosome.

Selection is performed on the population which now contains both the parents and the child solutions and the unfit individuals are replaced by the fit ones to maintain a constant population size (Seshadri, 2009). This process is iterated with the new chromosome population being entered back into the evolution process until the specified number of generations has been reached.

The output from the genetic algorithm (GA) is a Pareto-optimal front that represents a trade off curve between the objectives set at the initial stages. The real advantage of multi-objective GAs is seen at this point is that the front provides flexibility to the engineer upon which he can apply his/her engineering insight to provide a practicable solution to the client (Prasad & Park, 2004). This is discussed in more detail in the methodology.

3. Methodology

3.1. Application of the Monte Carlo and optimisation methods

The Monte Carlo based stochastic algorithm, developed by Van Zyl *et al.* (2008) in the C++ programming language, was studied by Chang & van Zyl (2010). The essential coding and core algorithms were re-coded using MATLAB.

Owing to the iterative nature of stochastic modeling, the computation effort involved is significant. For a stochastic model simulating a duration of 10 million hours, the model adapted by Chang & van Zyl (2010) had a run time of 60 minutes. When considering that the stochastic model is intended for optimisation, multiple systems must be simulated. For a single set of model parameters (head, pipe length, etc), 2500 solution systems are tested (50 solutions over 50 generations). This would result in the optimisation model running for an estimated 105+ days for a single set of model parameters. The run time is dependent on the computing system utilised. However, the intention is to develop an application that can be used on typical, commercially available desktop platforms.

Chang & van Zyl (2010) attempted to address the significant simulation time through the development of a compression heuristic. The coding of the stochastic model was adapted to allow the model to skip over the computational periods where no fire or pipe failure events take place. This was done by developing an algorithm that generated statistical distributions of individual solution systems under population demand-only (non-failure or fire demand) conditions. This is described in greater detail in section 3.5.5.

The optimisation model developed by Chang & van Zyl (2012), encompassed the stochastic model developed by Van Zyl *et al* (2008), the compression heuristic and the NSGA-II optimisation algorithm. The optimisation model was used to generate pareto-optimal solution curves for various model parameters based on a gravity-fed system. The model produced adequate results and demonstrated the sensitivity to varying model parameters. There were however, a number of limitations that have been addressed in the revised model proposed in this document.

3.2. Computational Platform

As mentioned in Section 3.1 above, the programming platform and language was MATLAB, a fourth generation programming language and numerical computing environment. It is considered a fourth generation programming language by Mathworks, owing to its high-level structure and user interface. The user input is more intuitive and natural in comparison to lower-level languages such as C or C++.

It was chosen as the computational platform as it was used to code the original optimisation model by Chang & van Zyl (2012), and owing to the numerous, integrated mathematical functions. Adopting the same platform ensured that the possibility of latent errors when re-coding was minimised, and that the comparison between original and revised models was more accurate. In addition, the design and structure of MATLAB™ is accommodating to the applied sciences, where the level of abstraction between user input and compiler/machine code is useful in reducing the coding and error rectification effort.

The benefits of using this particular platform are also its disadvantage in the application of stochastic process. The abstraction between user input and machine code implies marginally greater computational effort. When considering the vast number of iterations involved, a marginal difference becomes significant. The use of a lower level platform would be encouraged should this optimisation application be developed into a standalone program. However, for an academic endeavour, the difference can be considered acceptable.

3.3. Revised optimisation model base code

The original code as based on models developed by Van Zyl *et al.* (2008) and as integrated into an optimisation model by Chang & van Zyl (2012) was provided under full permission from all parties involved and subsequently used as a base upon which the broadened model could be developed. This was done in order to build upon existing research and provide a reliable basis for comparison.

3.4. Algorithm description

A macroscopic view of the model structure and mechanics of both the optimisation model developed by Chang & van Zyl (2012), and the currently proposed model, will be provided in the order of progression of the model. The amendments and subsequent appendments made will be detailed in order to gain an understanding of the process involved in revision of the original model to include lifecycle costing and pumping.

3.5. Mechanics of the Chang model

The optimisation algorithm as proposed and developed by Chang & van Zyl (2012), to be referred to as the Chang model, consists primarily of user inputs, a basic hydraulic calculation, a pre-run to determine reliability under normal water demand conditions only, the optimisation model, informed by cost and stochastic reliability model and the output of a solution population. This is presented in a simplified, topographical layout, in Figure 3.5.1.

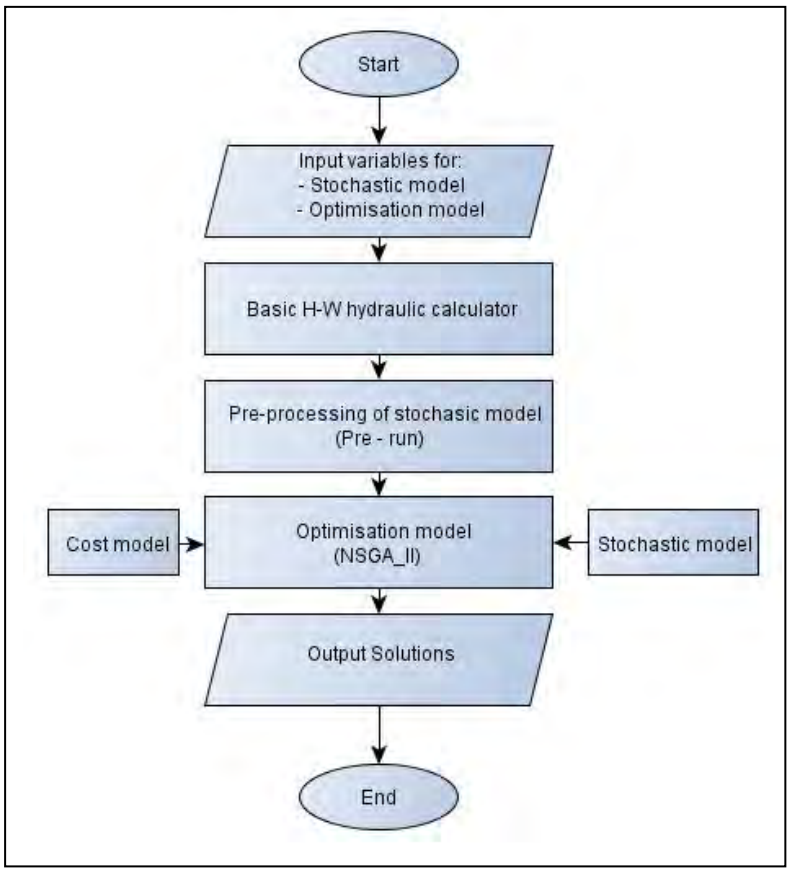

Figure 3.5.1: Algorithm of optimisation model (Adapted from Chang & van Zyl, 2012)

The model topology, upon which the original algorithm, above, is performed, is presented in Figure 3.5.2, showing water source, feeder pipes and reservoir (variables to be optimised) and demanding population.

3.5.1. Macroscopic model topology

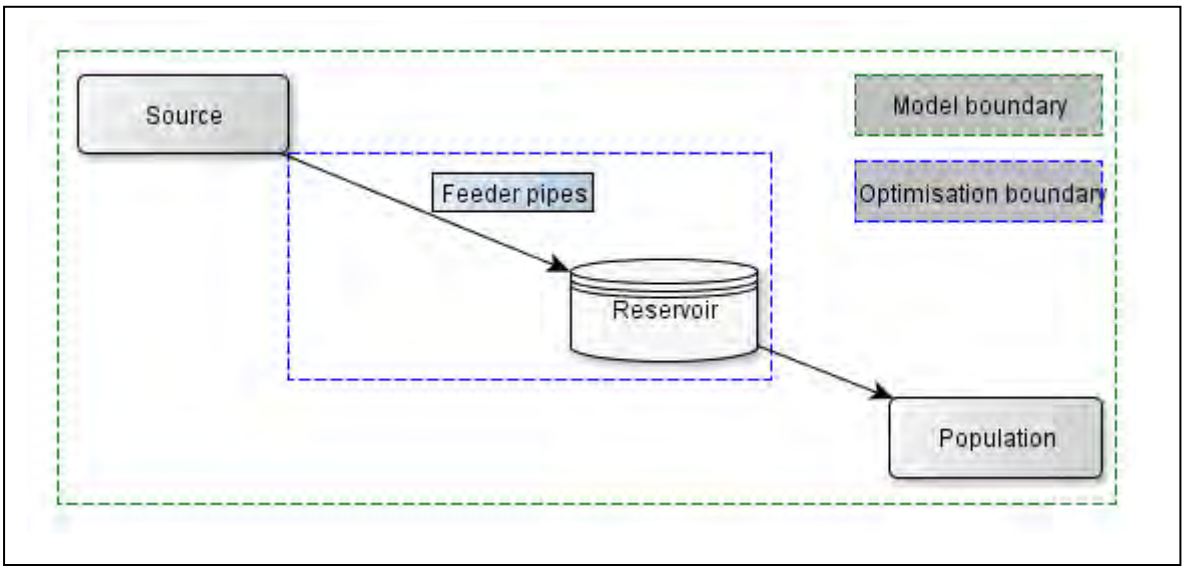

Figure 3.5.2: Layout of model used in Chang & van Zyl (2012)

The model is bounded by the supply from the source and the demand from the population.

User input and variable initialization

The following parameters were required as user inputs:

- Seasonal Peak Demand
- The total system head (elevation difference between source and reservoir)
- The total length of pipe between the source and the reservoir
- The Hazen-Williams coefficient of the pipe
- The minimum tank capacity
- The maximum tank capacity
- The resolution for tank capacity (step size)
- The simulation duration
- The seed number for the random number generators
- Population demand parameters (Van Zyl *et al*, 2008)
- Fire event frequency
- Pipe failure frequency.

The number and variation of pipe sizes and their configuration are based on commercially available quantities arranged in preconceived configurations.

The range of discrete pipe sizes as contained within the code consists of 26 available pipe sizes, as shown in Table 3.5.1:

Table 3.5.1: Commercially available pipe sizes (mm - ID)

127	145	182	227	286	322	363
428	479	530	574	626	675	726
777	828	878	929	976	1074	1176
1366	1568	1773	1970	2174		

The feeder pipe configuration is limited to the number of pipes and number of perpendicular connections between the pipes, in parallel. The pipe configurations used were as per Table 3.5.2.

Table 3.5.2: Possible pipe configurations

Configuration ID (#)	Parallel Pipes (#)	Perpendicular connections (#)
1	1	0
2	2	0
3	3	0
4	2	1
5	3	1
6	2	2
7	3	2

The basis behind employing perpendicular pipes between the source and reservoir is to increase system redundancy and resilience against pipe breaks. As previously mentioned and as defined by van Zyl *et al.* (2008), the severity of pipe failure is determined by the effect of the pipe failure on flow into the reservoir. Thus the more redundancies built into the network, the lower the severity of a particular pipe break. This is demonstrated in Figure 3.5.3.

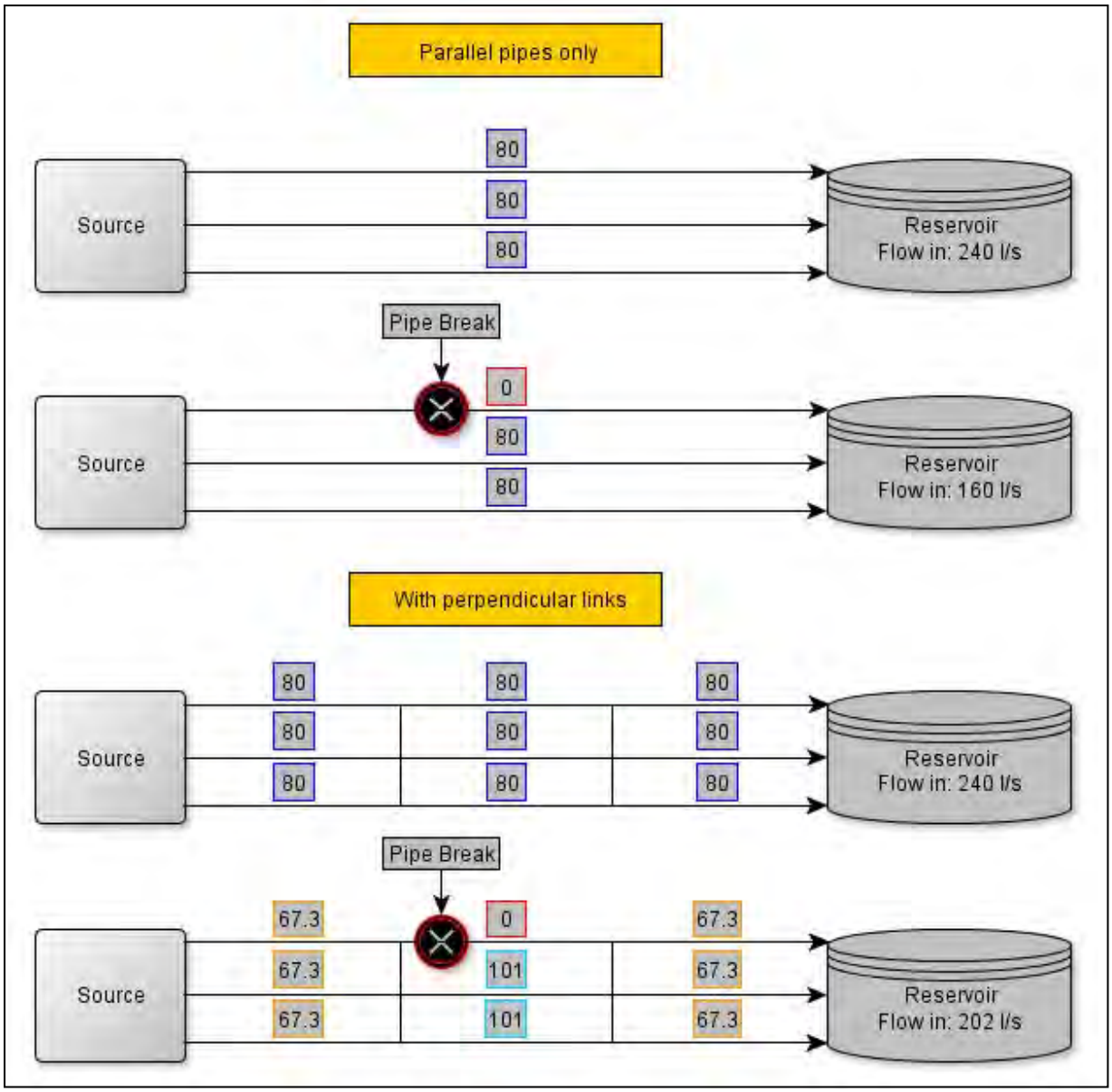

Figure 3.5.3: Effect of pipe failure with Ch = 120, L = 333m. D = 0.207 for each pipe

As can be seen from the above diagram, the increased redundancy through the use of perpendicular links between the main supply pipes results in a decrease in pipe failure severity of 26% with a flow remaining of 84.3% (Chang & van Zyl, 2012)

For this model, the ratio of flow remaining is defined as follows:

$$Q_{remaining} = \frac{Q_{pipe\ failure}}{Q_{operating}} \quad (3.1)$$

This is synonymous with the inverse of pipe failure severity.

3.5.2. Basic hydraulic calculation

During operating flow conditions, theoretically, the perpendicular pipes should not experience any flow, owing to the equal hydraulic energy on either end. The operating flow can thus be calculated in a determinate manner by employing the Hazen-Williams head loss formula:

$$V = 0.354 C_H D^{0.63} \left(\frac{h_f}{L}\right)^{0.54} \quad (3.2)$$

Where :
- V = velocity (m/s)
- C_h = H-W head loss coefficient
- D = Inner pipe diameter (m)
- h_f = Friction loss (m/m)

The velocity is used to calculate the feeder flow rate into the reservoir through:

$$Q_{inflow} = V \times A_{pipe} \quad (3.3)$$

Where:
- V = velocity (m/s)
- A_{pipe} = the cross-sectional area of the pipe (m^2).

By the previously mentioned assumption that no flow is experienced in the perpendicular pipes under operating flow conditions, the flow rate for each of the 3 parallel pipe configurations is iterated, with each iteration moving to the next pipe diameter until all 26 pipe diameters have been calculated for each parallel pipe configuration.

Thus, $Q_{530mm\ 2\ Parallel\ Pipes} = 2 \times Q_{530mm\ 1\ Pipe}$

$Q_{530mm\ 3\ Parallel\ Pipes} = 3 \times Q_{530mm\ 1\ Pipe}$

3.5.3. Supply ratio filtering

Following completion of the hydraulic calculation for each discrete pipe diameter, the algorithm sorts the pipe diameters based on the supply ratio calculated. The supply ratio as defined previously is:

$$Supply\ ratio = \frac{Q_{inflow}}{Average\ \ demand} \tag{3.4}$$

The average demand is modeled under seasonal peak condition and is constant for each run of the model. To ensure that the pipe diameters gave reasonable supply ratios for the given system restrictions were placed on the potential pipe diameters based on configuration. For a 1 pipe system, the supply ratio could not be greater than 2, and for a 3 pipe system, the supply ratio could not be less than 1. This is shown in Figure 3.5.4:

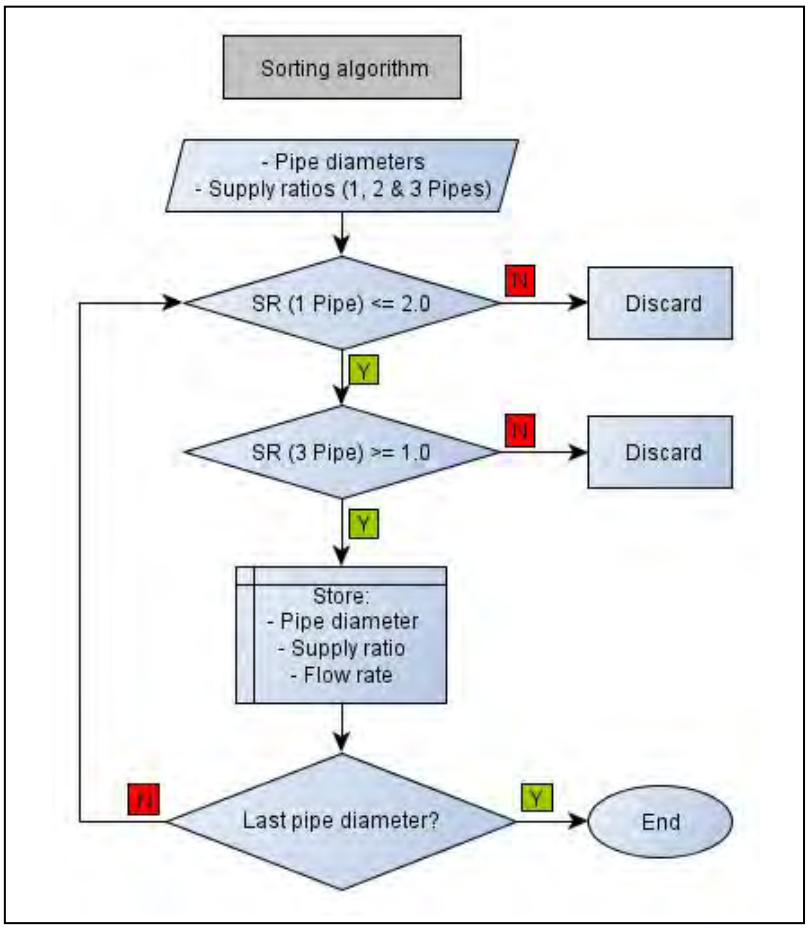

Figure 3.5.4: Supply ratio sorting algorithm

This ensures that the pre-run to be detailed later on does not waste computational time determining the reliability of systems that are either too expensive or too unreliable. It also filters out systems that would suffer from inefficiencies either in price or reliability. For example: a 3 pipe system with a supply ratio less than 1.0 would imply that an unreasonably small diameter pipe has been used. The consequence of using 3 small pipe diameters is a large construction cost with a small gain in supply ratio and thus a system that is unreliable and inefficient. An unreliable, expensive system would likely be dominated on the first evolutionary cycle, owing to this inefficiency, and thus the processing time is wasted.

3.5.4. NSGA II model parameters

The population size and number of evolutionary generations to be initialized and performed by the genetic algorithm is preset by the user. The number of generations used for this model was 50 with a population size of 50 chromosomes. This was seen to give a reasonable spread of solutions without resulting in an excessive computational time (Chang & van Zyl, 2012).

3.5.5. Demands pre-run

The demands-only pre-run is a component of the compression heuristic developed by Chang & van Zyl (2010) and later updated by Chang & van Zyl (2012). The purpose as mentioned previously is to reduce the computational load required by the full Monte Carlo method. The fundamental premise of the pre-run is to determine the reliability of the potential solution systems under consumer demand only and apply that reliability baseline to the model component that simulates pipe failures and fire events. The method employed is to test each feasible supply ratio with each tank capacity, against the flow demanded by the model population.

The demands failures are recorded along with associated lookup tables that are based on the reservoir level at the beginning of each week, on Sunday at 04H00. The purpose of this will be explained in greater detail in the events run section. This is also the time at which the reservoir is most likely to be full.

Each supply ratio associated with a particular number of pipes and pipe diameter is tested with the total number of loops through the pre-run algorithm being:

$$Total\ Loop\ Count = Length(SR) \times Length(RC) \quad (3.5)$$

Where: Length(SR) = the length of the vector storing each supply ratio
Length(RC) = the length of the vector storing each tank capacity

The pre-run includes the stochastic unit model developed by Van Zyl *et al.* (2008) for modeling consumer demands. The output variables are stored in an **n** x **m** matrix with **n** = supply ratio and **m** = reservoir capacity as is demonstrated in Table 3.5.3:

Table 3.5.3: Demands-only pre-run data storage matrix (Chang & van Zyl, 2012)

		Reservoir Capacity (RC)			
		4 h	6 h	8 h	10 h
Supply Ratio (SR)	0.52	• Number of failures • Simulated hours • Failure frequency • Convergence ratio • Full reservoir fraction			
	1.23				
	2.76				
	3.35				
	4.94				

Each output is stored in a separate two dimensional matrix. The simulated hour count is equal to the final iteration count at the point where the pre-run algorithm ends. This is due to the iterative step being in hourly units.

As mentioned previously, the measure of reliability is the frequency of failure or the inverse thereof. Therefore, after each failure event has occurred and subsequently ended, the time to failure is recorded as the duration (in years) between failures. The mean time to failure (MTTF) is also recorded. In order to determine the reliability of each supply ratio-reservoir capacity combination the following measure is used:

$$Failures\ per\ annum\ (FPA) = \frac{1}{MTTF} \quad (3.6)$$

The convergence ratio is based on the parameters used to terminate the demands-only pre-run. In order to obtain results that are realistic and accurate compared to their ultimate values, the stochastic process should proceed until the given system has experienced at least 2000 failures, and results are within 5% of their ultimate values (Van Zyl *et al.*, 2008). The problem, as identified by Chang & van Zyl (2012), is that when dealing with systems with a very high reliability, the run time required can become unrealistically large. Thus a termination criterion of confidence limits was proposed by Chang & van Zyl (2012). (Snedecor & Cochran, 1989; Ott & Longnecker, 2008, as cited by Chang & van Zyl , 2012) stated that the following equation can be used to determine an appropriate sample size:

$$n = \left(\frac{z \cdot S}{E}\right)^2 \quad (3.7)$$

Where: n = sample size
z = statistical z-value
S = standard deviation of sample
E = allowable error

Chang & van Zyl (2012) proposed an allowable error of 5% within a 95% confidence interval. When one takes into account the sample mean ($\bar{y}$) recorded at each reservoir failure, the range of tolerance can be described as follows:

$$\bar{y} \pm \frac{1.96S}{\sqrt{n}} \quad (3.8)$$

This equation was reformulated in order to develop a specific termination criterion, in that the pre run terminates when the mean time to failure falls within 5% of its ultimate value. This is expressed as follows:

$$\frac{1.96S}{\sqrt{n} \cdot \bar{y}} \leq 5\% \quad (3.9)$$

This termination criterion, as identified by Chang & van Zyl (2012) is effective for systems that have a low inherent reliability. This implies a large number of reservoir failures, which results in the mean time to failure quickly converging to within the 5% error tolerance.

However, what was discovered is that for systems with large supply ratios and/or reservoir capacities, the number of failures under demands-only conditions was minimal. The resulting time to convergence was not practical in terms of the computational time needed to converge. Thus another termination criterion had to be developed:

$$\bar{y} + \frac{1.96S}{\sqrt{n}} > \max\ Time\ To\ Failure \tag{3.10}$$

This allowed systems that had very long durations between failures to exit the pre-run loop when the upper confidence limit exceeded the maximum time to failure set by the user. This gave 2 termination criteria that covered systems within a wide spectrum of reliabilities. The final termination criterion was a hard limit set by the user in terms of a maximum number of hours to be simulated. In summary, the termination criteria developed by Chang & van Zyl (2012) are as follows:

i. Mean time to failure is within 5% of its ultimate value;
ii. The upper confidence limit of time to failure is above the maximum time to failure set by the user;
iii. Simulation duration exceeds the maximum duration set by the user.

The full reservoir fraction as shown in the two dimensional matrix previously is the fraction of the total number of passes through the beginning of the week that the reservoir has been completely full. Two three-dimensional matrices store the top-of-bin values for the reservoir lookup Table and relative cumulative frequency respectively. That is in an **n** x **m** x **z** matrix, **n** corresponds to the supply ratio values, **m** to the reservoir capacity values and **z** to the upper boundary of each of the ten divisions of the reservoir capacity values. The relative cumulative frequency matrix is identical with the exception that the z values now hold the values corresponding to the frequency count for each bin.

This is used to generate a lookup Table that incorporates the frequency of each reservoir level and full fraction so that a probabilistic estimate of the reservoir level at the beginning of each week (Sunday 04H00) can be made during the events run to be detailed in Section 3.5.8 and 3.6.6.

The final output of the demands-only pre-run is to fit a curve to the demands failure frequency for each supply ratio. MATLAB's curve fitting functions are used to fit a curve to the scatter-plot of reservoir capacity against the failure frequency (failures per annum). For example, a supply ratio of 1.2 may have 16 or more different reservoir capacities which it has been tested with, allowing a non-linear curve to be fitted.

The purpose of this is to formulate a continuous relationship between failure frequency and reservoir capacity, allowing the genetic operator to generate supply ratio and reservoir capacity values that may fall between the discretely tested values. In the original model, the supply ratios tested were definite and pre-determined. As will be explained in 3.6, in the proposed model, the supply ratios are determined in an ad-hoc manner which makes full use of this capability.

The graphical layout for the demands-only pre-run is demonstrated in Figure 3.5.5, showing the iterative nature by which the demands-only pre-run tests each supply ratio-reservoir capacity combination and records the results. The consumer demand model is integrated into the demands-only pre-run, and is further outlined in Figure 3.5.6.

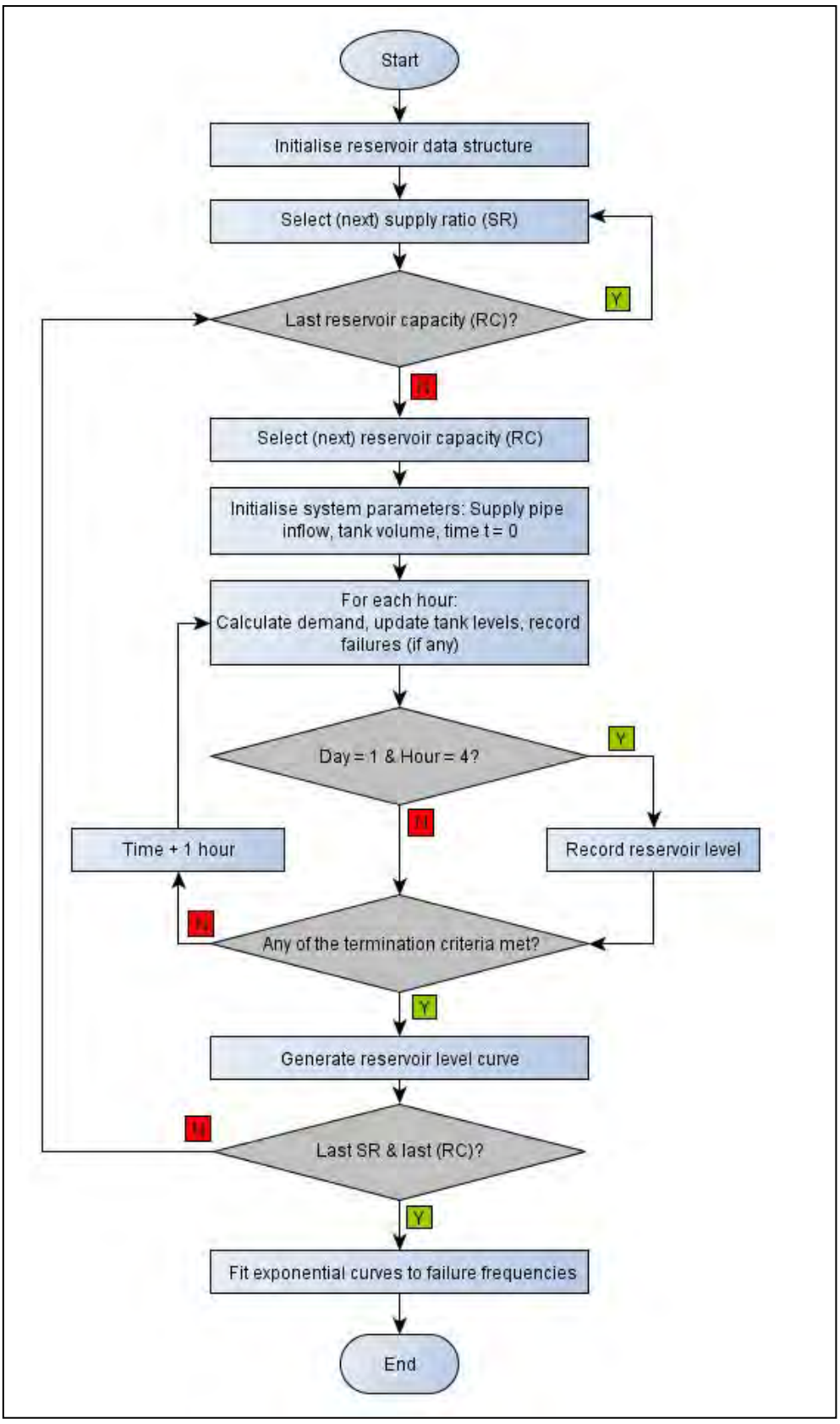

Figure 3.5.5: Demands-only pre-run algorithm, adapted from (Chang & van Zyl, 2012)

The consumer demand is calculated as described in the algorithm in Figure 3.5.5, and demonstrated in greater detail in Figure 3.5.6.

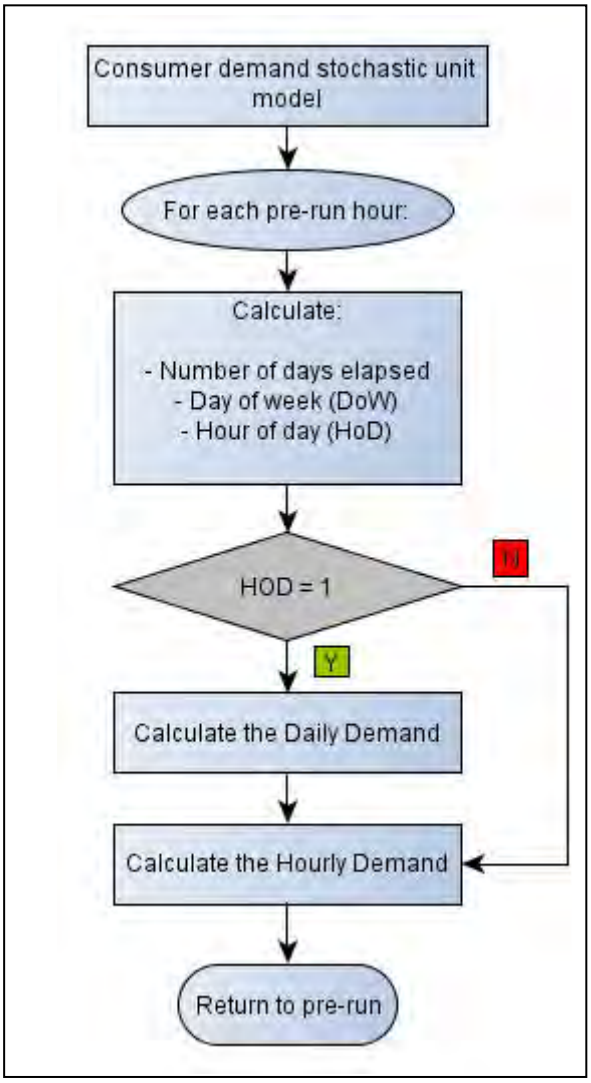

Figure 3.5.6: Consumer demand unit model

The hourly consumer demand is calculated according to the process described in the literature review. The consumer demand model is the only stochastic model utilised in the demands-only pre-run. The current tank volume is updated at the end of each hour in the following manner:

$$Current\ tank\ volume = previous\ tank\ volume + reservoir\ inflow - consumer\ demand$$

The termination criteria are checked and once the simulation reaches the last supply ratio and reservoir capacity the pre-run ends and returns all the required data called by the main function as previously shown. This concludes the pre-run, the parameters and data established are now collated and the model moves over to the optimisation process.

3.5.6. NSGA-II optimisation process

The data and statistical distributions generated by the demands-only pre-run are used as input into the optimisation function. The NSGA-II algorithm incorporates the following processes, defined in terms of their basic operations:

i. Generation of the initial population;
ii. Evaluation of the objectives (Cost and reliability);
iii. Determination of how closely spaced the chromosomes (solutions) are;
iv. Ranking of each chromosome in the population;
v. Comparison of each solution in the population to determine preference;
vi. Evolutionary processes

The theory behind the above processes is described in 2.5.2. The mechanics of these processes are described in the following chapters. The evolutionary process is iterative, and repeats until the required number of iterations (generations) has been reached. The outcome of the optimisation process, as detailed in 2.5.2, is a pareto-optimal front, a range of solutions that can reasonably be assumed to contain the best possible solutions for a particular cost or reliability. This allows the designer/user of the optimisation model to narrow down the potential solutions for implementation. This process is represented graphically in Figure 3.5.7.

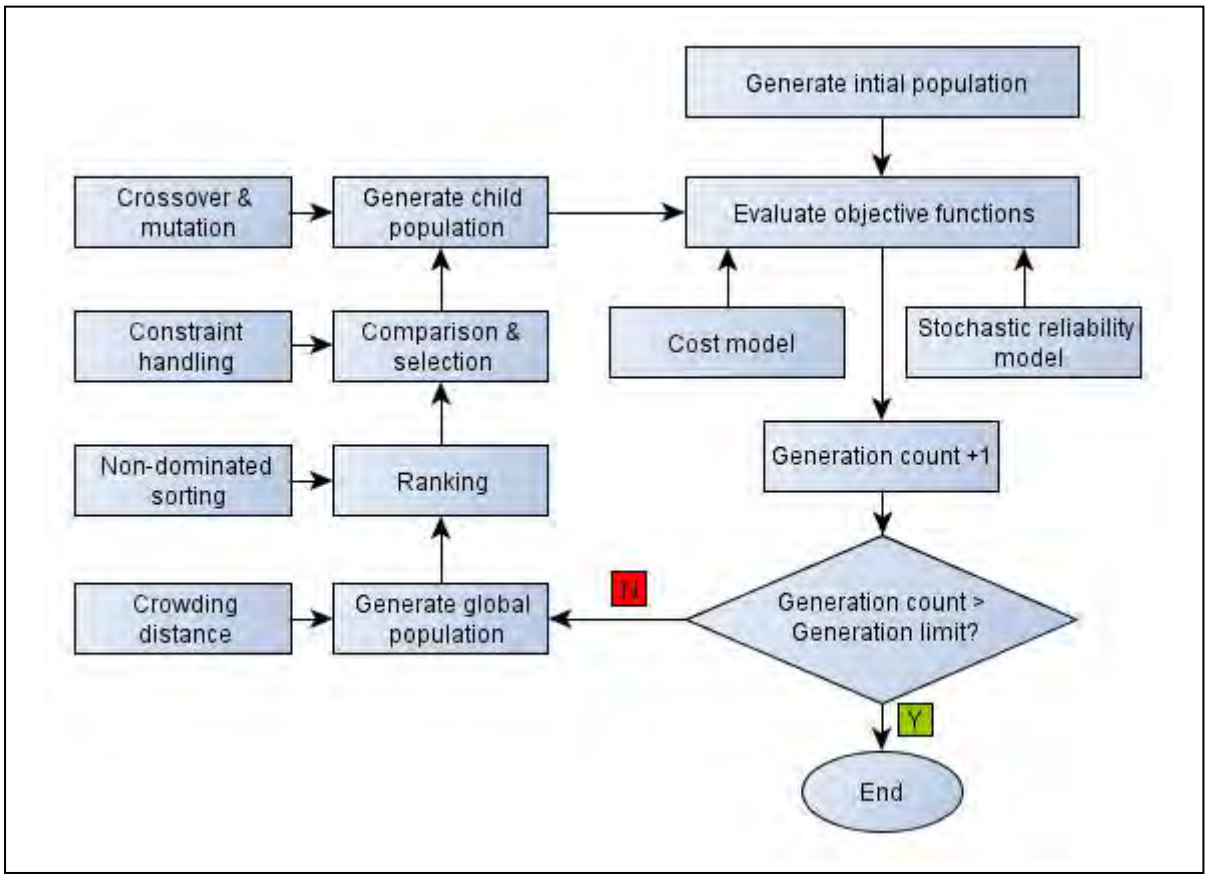

Figure 3.5.7: NSGA-II Genetic optimisation algorithm after Wu *et al* (2010)

The optimisation model operates within the model boundary as described at the start of section 3.5.1. That is, it operates to optimise the pipe configuration, pipe size and reservoir capacity of the model to achieve a set of pareto-optimal solutions (optimum reliability for a given cost or vice-versa).

As seen above, the algorithm will continue to loop until it has reached the user-specified generation limit. Each generation represents a refinement of the solution space. The individual processes will be detailed further in the following sections.

3.5.7. Generation of initial population

The initial population is generated through the utilisation of pseudo-random number generators that generate values for each of the decision variables within a chromosome, and constrained to the boundaries set by the user. The algorithm for this process is as shown in Figure 3.5.8: The pipe diameter (Table 3.5.1) and pipe configuration (Table 3.5.2) are discrete while the tank capacity is selected from a continuous range. The chromosome vector

contains the genetic information for each chromosome. The genetic information is shown by the 3 green cells. Each chromosome is evaluated according to the objective function models to be described in 3.5.8. The evaluated cost and reliability are appended as shown in the orange cells.

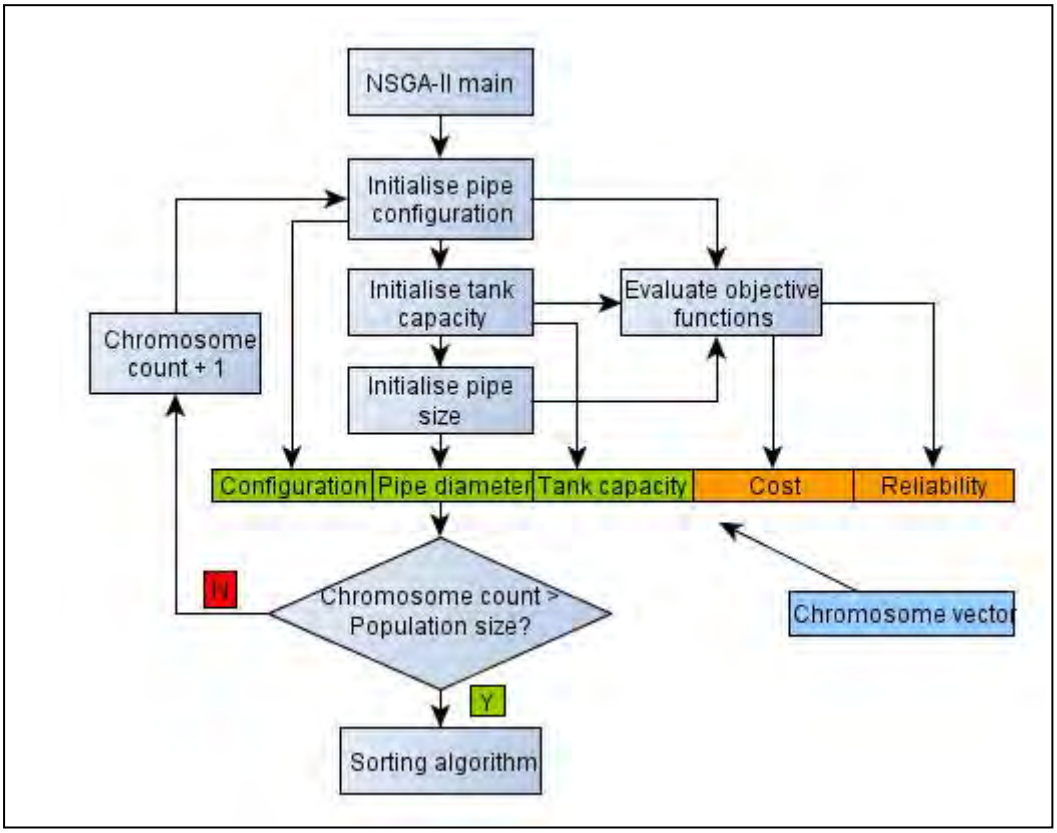

Figure 3.5.8: Initialisation of decision variables

At this stage the population has not been sorted or ranked. Following the completion of the non-domination sorting algorithm, the rank and crowding distance are further appended to the chromosome vector (not shown in Figure 3.5.8).

3.5.8. Evaluation of cost and reliability

The two objective functions in the original and subsequent algorithms are system cost and reliability. The cost is evaluated taking into account only the capital cost of the components that constitute the system; these components are evaluated as follows:

Pipeline cost is determined using the following equation:

$$C_m = k_m L D^m \tag{3.11}$$

Where: k_m and $_m$ are financial coefficients and exponents respectively.

The coefficient and exponent used in this model results in a pipeline cost equation as follows:

$$C_m = 480 L D^{0.935} \tag{3.12}$$

Storage reservoir cost is determined using the following equation:

$$C_R = \frac{290\ V_R}{[1+(\frac{V_R}{1100})^{5.6}\ ]^{0.075}} \tag{3.13}$$

Where: C_R = reservoir cost ($)

V_R = volume of reservoir (m^3)

Interconnecting chamber cost is determined in the following manner:

2 Pipe interconnections:

$$C_C = 0.0457D^2 + 16.519D + 2624.1 \tag{3.14}$$

3 Pipe interconnections:

$$C_C = 0.07D^2 + 24.898D + 2790.2 \tag{3.15}$$

The total cost is described by:

$$C_T = C_m + C_R + C_C \tag{3.16}$$

The overall reliability of the chromosome is evaluated through the events-run component of the stochastic algorithm. This is described as follows:

The overall reliability of the chromosome (solution system) in question is determined by performing a stochastic simulation of the periods where the system experiences an event. The events considered, as described earlier, are fire events, where the system experiences a heightened water demand from the supplied population, owing to fire-fighting considerations. The second type of event considered is a pipe failure event, where supply is interrupted or impeded owing to the closure of valves necessary to perform repair work on a damaged feeder pipe.

The fire and pipe (supply) failure event times are initialised from the mean and associated Poisson distribution for each, as determined by van Zyl *et al.* (2008). The duration of each event is calculated in the same manner. It sorts the events in order of occurrence and checks for any overlap. If there is, the end of the event is extended in order to concatenate the events into one extended-duration event.

The model jumps to the beginning of the week in which the event is to occur and begins the stochastic analysis. The reservoir level at the start of that week is obtained from the reservoir lookup Table generated during the demands-only pre-run described previously. The events model runs through each hour of the week, simulating consumer demands and the supply and demand conditions resulting from a fire outbreak or pipe failure, or both. The tank level is updated and any failures of the reservoir are recorded. The inflow during a pipe failure is pre-calculated as a ratio between operating inflow and flow during failure. This is static and assumes that failure will always occur at a particular point in the pipe layout for each configuration.

The event ends when the fire or failure has ended and the reservoir has returned to its maximum level. The terminating conditions for the event run are identical to those described for the demands-only pre run. The final reliability for the chromosome is determined by combining the results for the pre run and the events run in the manner described by equation 3.17.

$$Total\ MTTF = \left(\frac{1}{failures\ \ per\ \ annum}\right) = \frac{Demands\ \ fraction}{Demands\ \ MTTF} + \frac{Events\ \ fraction}{Events\ \ MTTF} \qquad (3.17)$$

Where: MTTF = Mean time to failure (years)

The fractions referred to are the fractions of the total time simulated as split between demands and events fractions. The events-run component of the model simulates the bulk water system during all pipe failure and fire events that the particular chromosome (solution system) may be subjected to, as follows:

The time of the first fire and pipe failure event are initialised from the stochastic models describing their frequency, outline in 2.4.1. The events are chronologically ordered to determine the first event. The full-fraction for the reservoir size and reservoir level distribution curve are used to generate the reservoir level at 04H00 on the Sunday (Day 1 Hour 4) preceding the event. The model proceeds hourly until the event begins. Once the event has begun the inflow into the reservoir is adjusted according the pipe failure conditions and pipe configuration (in the case of a pipe failure event). The outflow from the reservoir is adjusted according to the fire flow conditions (in the case of a fire event). For each hourly time step, the reservoir level is updated as described in 3.5.5, and any tank failures are recorded.

At the end of the current event, the time until the next event is calculated and the overall failure frequency is calculated from the demands-only pre-run and the current events run. If any of the termination criteria, described in 3.5.5 have been reached, the model terminates and returns the overall failure frequency, to be appended to the vector of the chromosome that has been tested. If none of the criteria have been reached, the next event is calculated and the model repeats the process until a termination criterion has been reached.

This process is represented graphically in Figure 3.5.9.

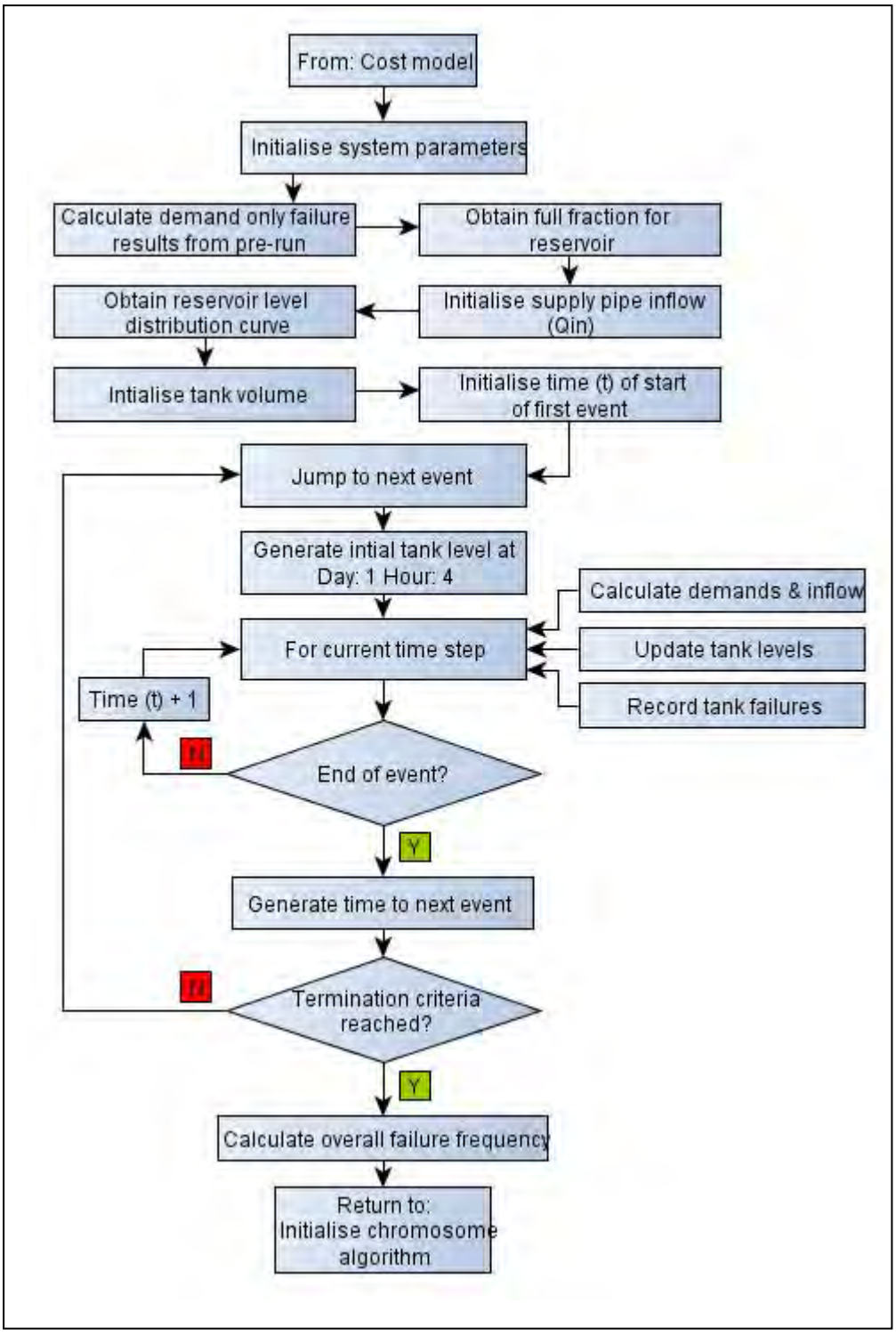

Figure 3.5.9: Events-run algorithm

3.5.9. Non-dominated sorting and crowding distance

Once the cost and reliability for each of the chromosomes have been determined, the population (set of 50 chromosomes), must be ranked, and the crowding distance determined. The genetic algorithm does this by forming domination fronts, using the definition of multiple-objective domination (cost and reliability, in this instance), as presented previously in 2.5.2.

2 unique solutions (chromosomes) **p** and **q** within the initialized population: **p** & **q**,

i. **p** is better than **q** if **p** is no worse than **q** in all objectives & at least one objective is better;
ii. **q** is better than **p** if **q** is no worse than **p** in all objectives & at least one objective is better;
iii. Otherwise, **p** & **q** are equally good.

This process is repeated until all chromosomes are ranked by domination.

A domination set (Sp) and a domination counter (Np) are initialised for each chromosome (solution). For each chromosome that is dominated by another solution, the domination counter is incremented by 1 and that chromosome is added to the domination set. The value Q represents the ranking front that is one rank below the front under consideration (i). The chromosomes are arranged into non-domination fronts, where the domination counter of each chromosome is equal, and criterion 3 in the above list is true. The non-domination front is appended to the chromosome vector and the result is the stratification of all chromosomes within a population in order for the evolution process to pick the fittest solutions.

This process is represented graphically in Figure 3.5.10.

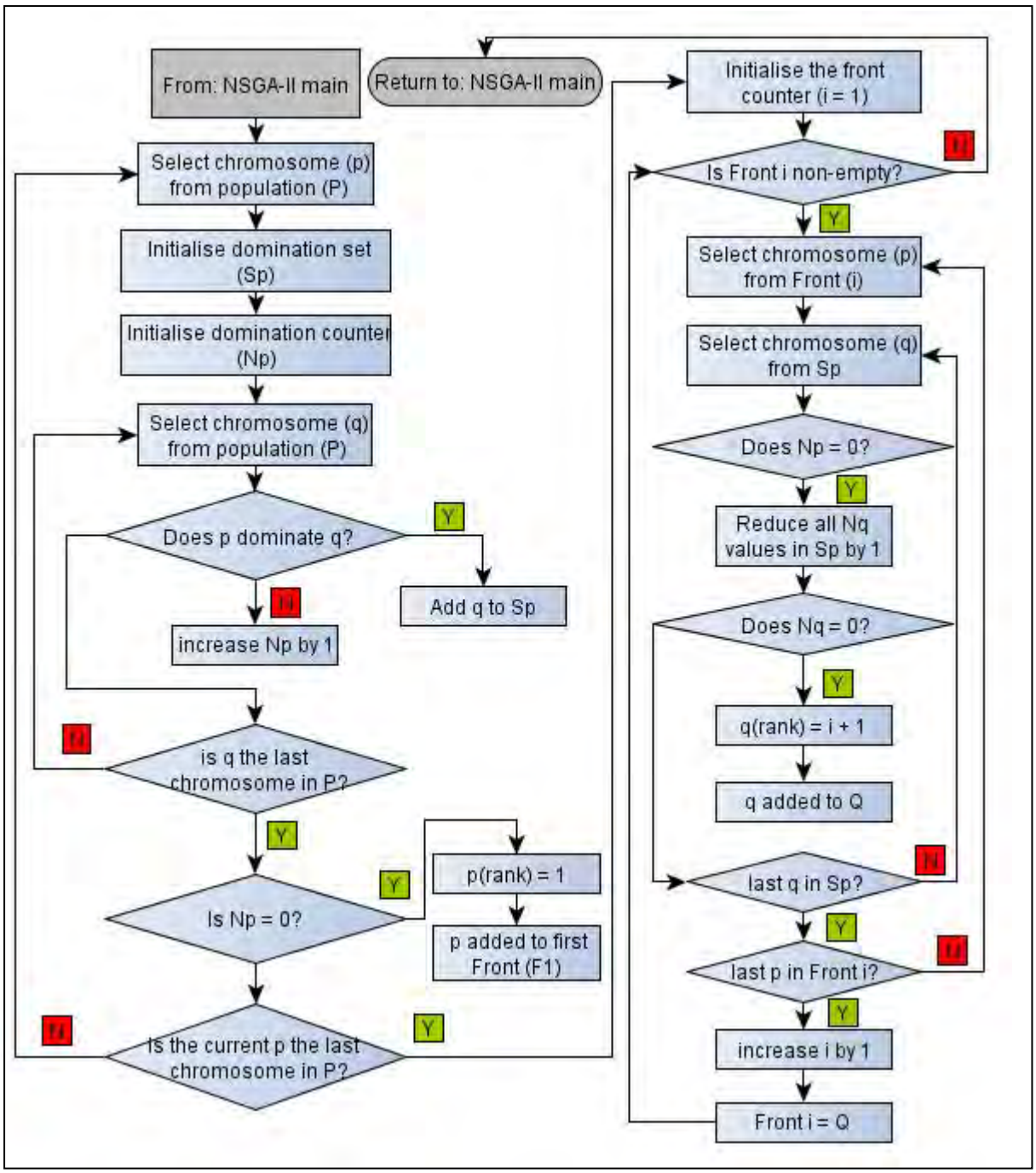

Figure 3.5.10: Non-dominated sorting algorithm

In order for the population to avoid crowding around a particular point, the genetic algorithm must take into consideration the closeness of a particular chromosome to the nearest chromosome. The crowding distance of each chromosome must be determined in order to discriminate against those that are too closely spaced.

The crowding distance of a particular chromosome is a measure of its diversity in relation to peer solutions of its front. The crowding distance algorithm when used in conjunction with the evolutionary process ensures that the solution front remains diverse (Seshadri, 2009).

The crowding distance is represented graphically in Figure 3.5.11.

The black circles represent chromosomes from the same front. The solutions for each front are sorted in descending order and the boundary values are assigned an infinite distance value while the intermediate values are assigned distance values based on the perimeter of the cuboid for each objective function. (Deb et al, 2002 as cited by Chang & van Zyl, 2012), the rank and crowding distance are appended to the chromosome vector, as per Figure 3.5.12

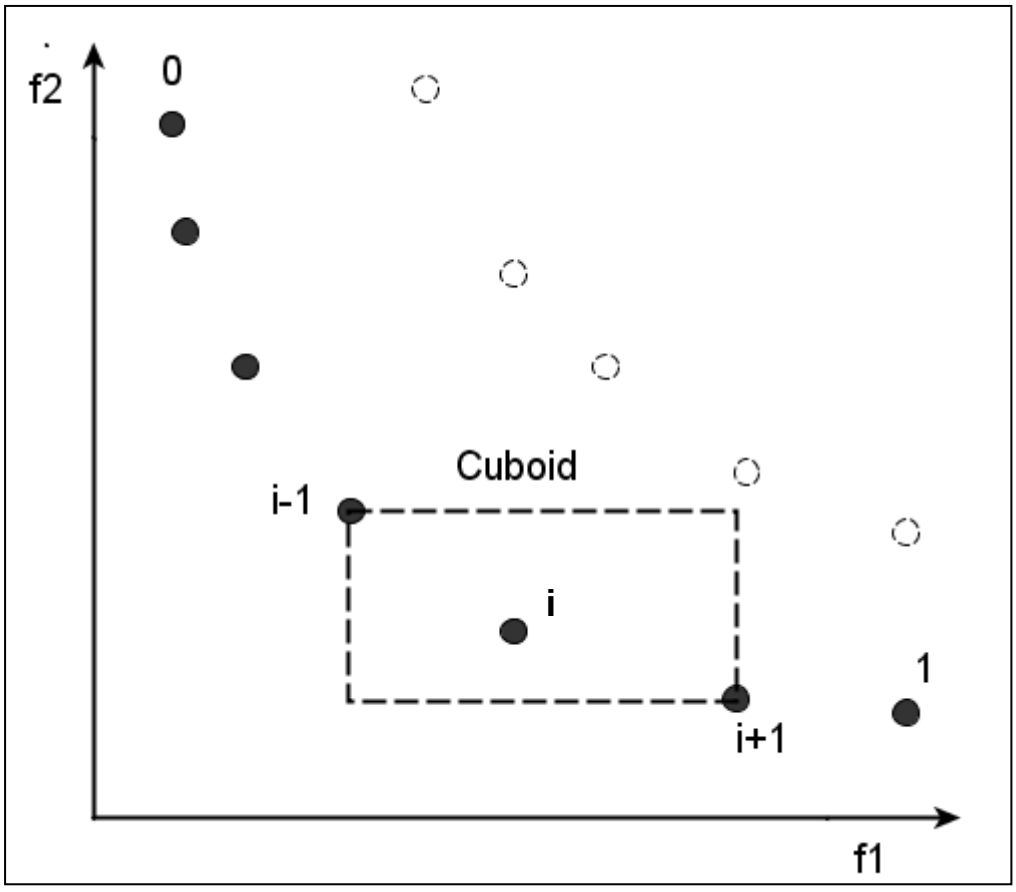

Figure 3.5.11: Crowding distance of solution (i) - Deb *et al (2002)*

.

Figure 3.5.12: Chromosome vector with appended optimisation parameters (blue)

The chromosome vector is appended with the rank and crowding distance forming a complete, sorted and ranked chromosome that has all the necessary parameters associated with it to be entered into the evolutionary process.

3.5.10. The evolutionary process

The evolutionary process employed in the NSGA-II genetic algorithm is mathematically complex and thus for the purpose of this investigation a qualitative summary is provided instead. For more information on the mathematical structure behind the NSGA-II evolutionary process, refer to Deb *et al* (2002), Deb & Agrawal (1995) and Seshadri (2009).

The evolutionary process consists of a series of phases through which the population is evaluated and manipulated over multiple generations in order to refine the pareto-optimal solution set. The individual sub processes that make up the evolutionary process are summarised as follows:

i. **Tournament selection:**

A selection policy is applied to the generation population, to choose chromosomes fit enough to form the mating pool, a sub-grouping of preferred solutions that will be used to form child solutions through the crossover and mutation processes. The selection of fit solutions is determined using a comparison operator, which factors in the non-domination rank and crowding distance explained earlier. The crowded-comparison operator is represented by the following symbol ($\prec_n$).To find the preferred solution, the following comparative test is used:

Consider two chromosomes of the initialised population (i ; j). For the outcome where: chromosome $i \prec_n$ j (i preferred over j)

$$i_{rank} < j_{rank}$$

The rank of chromosome i must be lesser (superior) to the rank of j.

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

If the solutions being compared are elements of the same non domination front then the preferred solution is that with the greater crowding distance (higher diversity). This method is used to select the parent chromosomes upon which the crossover and mutation processes will be performed. The crossover process consists of multiple mathematical and probabilistic mechanisms with a high level of complexity.

ii. Simulated binary crossover:

Based on the natural biological mechanism by which genetic information of the parent is transferred to the child, the crossover process is a probabilistically backed alteration to the genetic information. In the NSGA-II algorithm, this information is the value of each of the decision variables in the chromosome vector. That is, the value of each decision variable (e.g. pipe diameter) is changed slightly. The process is titled simulated binary crossover or SBX. This allows for both discrete and continuous variables as the possibility exists for forcing the changes to adhere to discrete values (the complexities associated with discrete-forcing are discussed in further chapters. The crossover has a 90% chance to occur for each round of the genetic operator. The parents are randomly selected from the mating pool formed during the tournament selection process and the alterations are made to each of the decision variables in the parent chromosomes. The altered versions are stored as offspring chromosomes following a check that determines whether the changes are permissible with respect to the boundary conditions imposed by the user.

iii. Mutation

The 10% probability that exists is the probability that a mutation of one of the members of the tournament selection pool will occur. This is carried out in much the same way as the crossover; however there is the added influence of perturbance factors that are randomly distributed and determine the degree of mutation. The

mutated chromosome represents a third child chromosome that is entered into the child population.

iv. Recombination and selection

The parent and child populations are recombined into a single intermediate population that is non-domination sorted in the same manner as the initialised population. Following the sorting, the evolved population is compiled by filling the fronts in ascending order until the population size, as determined by the user, is reached. That is, the weaker solutions in the higher (inferior) ranks are pushed out and replaced by the fitter solutions, i.e. they are simply crowded out of the solution space.

The recombination of parent and child solutions is to avoid the complete replacement of the previous or initialised population by the child population as elements of the child chromosome population may be less fit that of the parent population. The NSGA-II evolutionary process is repeated until the number of generations reaches that set by the user.

3.6. Development of the base model

The model, to be referred to as the Chang model, as developed by Chang & van Zyl (2012) described in the previous section while being intricately and comprehensively compiled had a key number of limitations. The model, deals primarily with discrete decision variables, in the form of pipe configuration and pipe diameter. This is explained by Chang & van Zyl (2012) to address the problem of inefficiencies caused by the rounding of solutions that are not economically or practically feasible in terms of what is commercially produced. This is valid as it provides solutions that can be taken directly from the Pareto optimal curve and applied in reality.

The greatest limitation of the model is a lack of applicability to real-world bulk water supply systems involving pumped flow. The complexity of simulated systems is also limited as the hydraulic capability of the model is limited to determinate configurations. These limiting factors were seen as a starting point for the development of a pumped flow model.

3.6.1. Integration of EPANET

To include a hydraulic solver was seen as a necessary step in order to expand the model's hydraulic calculation capabilities. To be able to calculate the hydraulic properties of any simple or complex, pumped or gravity fed system, in an ad-hoc fashion, during the optimisation process, was seen as necessary. This is to avoid having to calculate every hydraulic possibility before the optimisation process as in the Chang model, which is not feasible when using a continuous variable such as pump power (as opposed to a defined range of discrete pipe sizes and configurations).

The full potential of its application to large variations in the topology of the system is not explored in this investigation. EPANET was seen as an appropriate solver to use due to its input method, public domain license and comprehensive technical user support. It is also widely used in industry and in the public sector. This widespread use implies a large number of potential users who are familiar with the software, this in turn results in a significant benefit to the ease of use and further development of the model.

A baseline algorithm for the incorporation of EPANET into MATLAB had been developed by Jonkergouw (2007). This was able to return hydraulic values (e.g. flow and head) from the EPANET program. This algorithm was extended by Eliades (2009) who added functionality to the link that allowed MATLAB to call EPANET functions through integrated MATLAB C++ library call functions. This was integrated into the model and allowed for full hydraulic calculation of the hydraulic network at any point in the model. The Hazen-Williams head-loss equations, pipe material and other hydraulic constants were kept constant in order to keep the comparison between the different applications of stochastic process and genetic optimisation on the same basis.

3.6.2. Application of demands-only pre-run

As the exact supply ratio values to be used in the optimisation process cannot be pre-calculated, as with the Chang model, a new approach was necessary. The supply ratio for each chromosome is dependent on the reservoir inflow of each solution system, which in turn is determined by the pump size, pipe size and number of pipes (elevation head remains constant for each run of the model). To address this inability to pre-calculate, the demands-only failure results are calculated over a set range of supply ratios and reservoir capacities (as per 3.5.5). The solution system is composed of randomly generated values for each of the decision variables, as per 3.5.7. The supply ratio is calculated after initialisation, this is detailed in 3.6.4 and Figure 3.6.2.

The discrete results from the demands-only pre-run are interpolated between, to estimate the statistical distributions for the supply ratio and reservoir capacity of each randomly generated chromosome. The advantage of using this method is that the demands-only pre-run and curve fitting process only have to be calculated once, following that the results and coefficients for the curves are stored and recalled, as it is unaffected by any permutation of the user-input parameters. This saves a large amount of computational time as the interpolation process negates the need to calculate the characteristics of each, unique system under demand-only conditions.

3.6.3. Expanded base model topology

The topology of the base model needed to be expanded to include pumps and the effect of power source reliability. The model consisted of the same arrangement seen in the Chang model, with the potential for a maximum of 3 pipes in parallel, but with the possibility for 3 interconnections, as opposed to 2, and the capability of having 3 different pipe sizes, as opposed to 3 of the same pipe sizes. The implication is that instead of 7 configurations and 26 pipe diameters, resulting in 182 permutations (as per the Chang model), there are now a proposed 416 permutations for the pipe arrangement. When combined with pump power, the number of significant permutations for supply ratio is substantially larger, owing to the continuous nature of the pump power variable. The base model topology is shown in Figure 3.6.1.

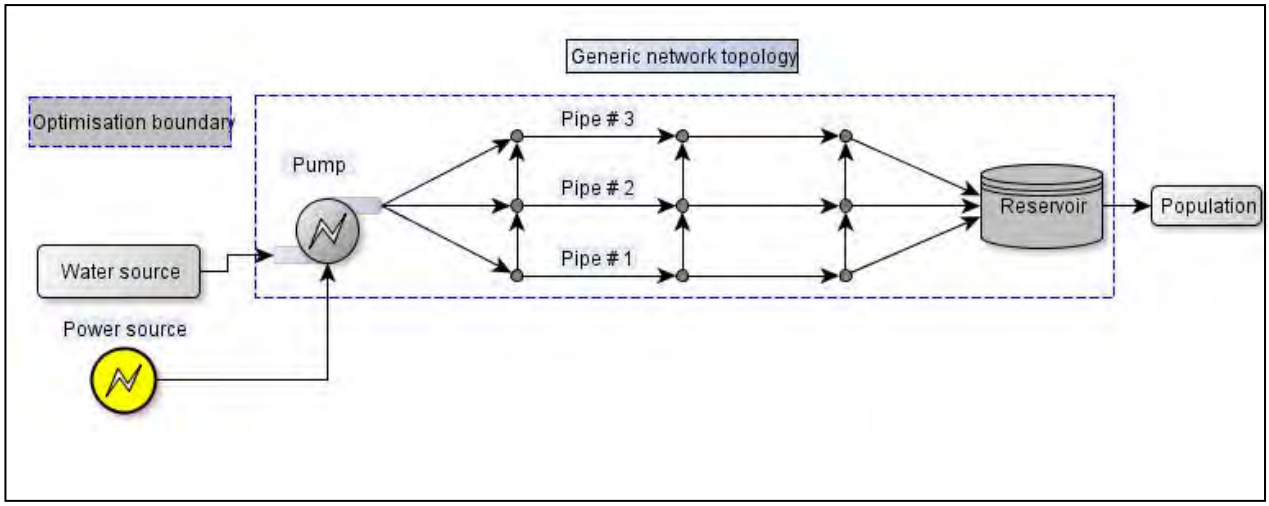

Figure 3.6.1: Generic network topology of base model

The findings of Chang & van Zyl (2012), demonstrated the increasing inefficiency of multiple pipe solutions. However both 2 and 3 pipe systems were included for comparability to the original model as well as for the possibility that the inclusion of a pump and the associated reliability model may have an influence on this inefficiency.

3.6.4. Decision variables and initialisation of population

Pipe Layout

As mentioned in 3.6.3, in an attempt to increase the level variability within the model, the single decision variable for pipe layout was replaced by a potential 5 decision variables: the number of pipes, the number of interconnecting pipes (if the number of pipes > 1), and the diameter of each pipe. During the crossover and mutation phase of the evolution process, the values of each of the variables are altered, the number of decision variables was thus structured to ensure that a higher value equals a more capacitated and failure-resilient hydraulic system. The logic is that, should systems composed of more than 1 pipe in parallel become an efficient solution, compared to a single pipe system, the crossover process will be able to find a finer, more efficient solution set by varying a larger number of decision variables (individual pipe diameters).

Pump model

There are a large number of parameters by which pumping systems are sized, along with the pump power, layout and system redundancies. There are also a significant number of complex considerations associated with the selection process of each system component. In order to focus on the relationship between pumping system capacity and the rest of the system components, the pumping system in its entirety is defined by its input power demand (in kW). The pump power is to be utilised as a continuous variable, owing to the multiple options that exist when setting up a bulk water pumping system, such as intelligent power control systems and pump layouts. The integrated hydraulic solver calculates the reservoir input from the pump power, pipe system parameters and overall system parameters (elevation head, length from source to reservoir, etc.). Sizing the pump system by pump power allows for easier optimisation of the energy cost of the system.

The set of chromosomes comprising the initial population is formed by assigning values to each of the decision variables. The decision variables are assigned as follows:

i.	Pump power	(kW, continuous)
ii.	Number of pipes	(#, discrete,)
iii.	Number of interconnecting pipes	(#, discrete) (if >1 pipe system)
iv.	Pipe 1 diameter	(m, discrete, preset diameters)
v.	Pipe 2 diameter (if iii = 2)	(m, discrete, preset diameters)
vi.	Pipe 3 diameter (if iii = 3)	(m, discrete, preset diameters)
vii.	Reservoir capacity	(h, continuous)

Following the initialisation of decision variables i through vi, the hydraulic solver is called to determine if the supply ratio of the chromosome falls within the boundary [0.5,5]. If it falls outside of the boundary, the decision variables are re-initialised and re-checked until the supply ratio falls within the bounds. It is this iterative algorithm that allows for a significant amount of diversity in the solution space. It is also possible to manipulate the solution space using a closed loop approach. The model can be forced to only accept initial solutions within a more specific supply ratio and/or reservoir capacity range in order to focus the solutions produced. This is investigated in more detail in 4.3.11.

Finally the reservoir size is randomly selected from within the range. Following the evaluation stage, the algorithm initialises the next chromosome until the population is at the size specified by the user, which in the standard case, is 50 chromosomes. The initialisation algorithm is detailed graphically in Figure 3.6.2.

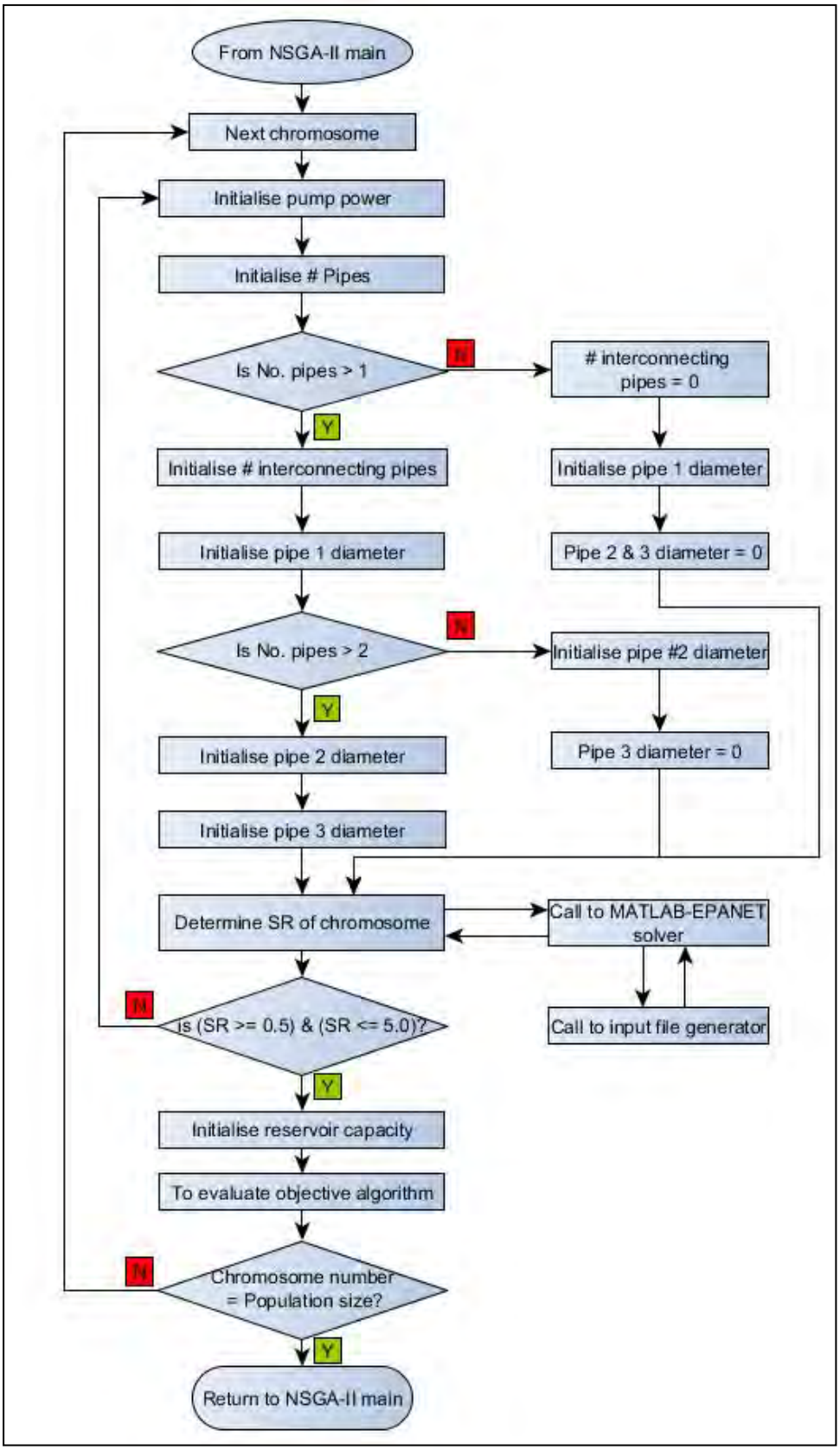

Figure 3.6.2: Algorithm for initialisation of decision variables within chromosome

3.6.5. Evaluation of cost and reliability

Following the initialisation of each of the decision variables, the objective functions evaluate the cost and reliability of the chromosome. This is repeated for each chromosome in the initial population. The basic structure is similar to the Chang model. However the cost and events run functions have been expanded to include life-cycle cost, an additional pump failure model and the integrated hydraulic solver.

Expansion of cost model

The Chang model considered only the capital cost of the system under consideration. As explained in the literature review, this is not appropriate when considering the costs associated with bulk pumping systems, such as operation and maintenance, and energy costs. The cost model was expanded as follows:

Net present value lifecycle costing

The net present value (NPV) method was employed in the cost estimation process to provide a more accurate financial estimate of total, life-cycle cost of each chromosome.

Net present cost is presented by Swamee & Sharma, (2008), as follows:

$$P_{NC} = C_0(1+r)^{-T} \tag{3.18}$$

Where : P_{NC} = Net present capital cost
C_0 = Current capital cost
r = Discount rate
T = Analysis period (years)

The assumption made is that the future capital cost is equal to the current capital cost (C_0) due to uncertainties in the projection of future cost. This is used to adjust the net present capital cost of a system component by factoring in the effect of the chosen discount rate.

The salvage value is represented by the following equation, and carries a negative cost:

$$P_{NS} = \alpha C_0 (1 + r)^{-T} \quad (3.19)$$

Where: P_{NS} = Net present salvage cost
C_0 = Current capital cost
r = Discount rate
T = Analysis period (years)
α = Salvage factor (of capital)

The operation and maintenance cost is represented by the following equation:

$$P_{NA} = \beta C_0 \frac{(1+r)^{-T}-1}{r(1+r)^{-T}} \quad (3.20)$$

Where: P_{NA} = Net present operation and maintenance cost
C_0 = Current capital cost
r = Discount rate
T = Analysis period (years)
β = Annual operation and maintenance factor (of capital)

The NPV of a generic pipe or reservoir system consisting of capital cost, maintenance and salvage, is represented as a combination of the above equations as follows:

$$P_N = C_0 [1 - \alpha(1 + r)^{-T} + \beta \frac{(1+r)^{-T}-1}{r(1+r)^{-T}}] \quad (3.21)$$

Where: P_N = Net present salvage cost
C_0 = Current capital cost
r = discount rate
T = analysis period (years)
β = operation and maintenance factor (of capital)

The energy cost of a generic pump system is represented by the following equation:

$$E_{NA} = E_a \frac{(1+r)^{-T}-1}{r(1+r)^{-T}} \quad (3.22)$$

Where: E_{NA} = Net present energy cost

E_a = Annual energy cost

r = Discount rate

T = Analysis period (years)

Cost and system parameters

The design life constants, maintenance and salvage factors utilised are as proposed by Swamee & Sharma (2008):

Discount rate: (r) = 0.04 (4%)

The discount rate, as explained by Swamee & Sharma (2008), should be roughly equal to the difference between the prevailing interest rate and inflation. It should also be roughly equal to the lower interest rate offered by government authorities to presiding water management authorities. A discount rate of 4% was chosen keeping in mind the current drive for greener civil services, and the drive for ecological and environmental awareness through reduced energy consumption. A lower discount rate biases investment toward capital cost by ensuring that the operation and maintenance cost is not disproportionately represented by being too far offset by interest accrued annually. This discount rate is applied to each component of the solution system.

Total system design life: (T_{system}) = 50 years

The design life of the solution system and individual components were set to achieve a realistic total system design life. The overall system design life (T_{system}) was chosen as 50 years, owing to 50 years being the approximate design life for protected uPVC, as well as the pump house. It is also the approximate amount of time in which 4 pumping systems with an

average design life of 12.5 years would have been installed owing to wear. It was also seen as reasonable that after 50 years, the service reservoir would still be structurally sound, but the capacity would need to be reviewed and revised according to corrected, actual demand and total system capacity. As such the design life of the individual system components, with a design life > 50 years, have been capped to 50 years.

Pipeline Costing

As the pipeline is constructed at the outset of the capital venture, all capital cost equations are kept as per Chang & van Zyl (2012) for comparability. The financial factors used for costing the pipe system are as follows:

Pipeline design life: ($T_{pipeline}$) = 50 years
Pipeline maintenance: ($\beta_{pipeline}$) = 0.04 (4%)
Pipeline salvage: ($\alpha_{pipeline}$) = 0

The pipeline design life for uPVC ($T_{pipeline}$) is listed by Swamee & Sharma (2008) as 60 years. However for this exercise it was limited to 50 years in line with the total system design life chosen. This value is seen as reasonable owing to the inherence variance in construction and material quality.

The operation and maintenance of the pipeline is seen as being a relatively minor overall cost. As such the percentage of capital cost attributed to maintenance ($\beta_{pipeline}$) is equally minor. In reality, the operation and maintenance cost would be dependent on the reliability (failure rate) of the pipeline, which in turn is a factor of length. However for simplicity and comparison's sake, the operation and maintenance cost for the pipeline and all other system components is kept static.

The salvage factor (α) was assumed to be 0 as uPVC is seen as unsalvageable at the end of its design life.

The capital cost of the pipeline was determined as follows:

$$Cc_{Pipe} = 480 \times L_{Pipe} \times (D_{Pipe})^{0.935} \quad (3.23)$$

Where: Cc_{Pipe} = Pipeline capital cost ($)
L_{Pipe} = Length of pipe (m)
D_{Pipe} = Diameter of pipe (m)

This is calculated for each pipe in the system, up to a maximum of 3 pipes, and added together to get the total pipeline capital cost. As the pipeline is to be installed at the outset of the capital project, the cost is naturally at present value.

The operation and maintenance cost of the pipeline(Com_{Pipe}) was determined using the financial factors listed above and equation 3.20. The capital cost of the pipe chambers used to connect the main pipelines to the interconnecting pipes was determined as follows, in line with:

$$Cc_{Chamber} = 0.0457 \times (D_{Pipe})^2 + 16.519 \times D_{Pipe} + 2624.1 \quad (3.24)$$

Where: $Cc_{Chamber}$ = Chamber capital cost ($)
D_{Pipe} = Diameter of pipe (m)

The cost of each of the chambers is based on the individual pipe size of the connecting pipes and the number of interconnections. No operation and maintenance or salvage cost is considered for the pipe chambers owing to the relatively minor influence of the chamber cost on the overall cost of the standard system.

The lifecycle cost of the pipeline components is as follows:

Pipeline: $C_{NPV-Pipe} = Cc_{Pipe} + Com_{Pipe}$

Chamber: $C_{NPV-Chamber} = Cc_{Chamber}$

Pump System Costing

The overarching pump system simulation was based on the pump's assumed design life and, as mentioned in the beginning of section 3.6.6, that future value is assumed to be equal to present value. Further base assumptions are that the pumping system power will not change between the first and last pump installations. The financial factors used are presented below:

Pump system design life: (T_{Pump}) = 12.5 years
Pump system maintenance: (β_{Pump}) = 0.03 (3%)
Pump system salvage: (α_{Pump}) = 0.2

Swamee& Sharma (2008) propose a design life of between 12 and 15 years for a pumping system. 12.5 years was chosen as a conservative amount that allowed for 4 salvage and replacement operations within the overall system design life of 50 years. The operation and maintenance, and salvage values were chosen from the reference values proposed by Swamee& Sharma (2008), as it likely that a percentage of the pumping system can be salvaged for sale or reuse.

The cost of the pump constructed at the outset of the capital project is calculated as follows:

$$Cc_{Pump} = 5560 \times (P_{Pump})^{0.7231} \quad (3.25)$$

Where: Cc_{Pump} = Pump capital cost ($)
P_{Pump} = Pump input power (kW)

The present capital cost of each subsequent pump is calculated as follows:

$$Cc_{PumpNPV}(x) = Cc_{Pump} \times (1 + 0.04)^{(x) \times 12.5} \quad (3.26)$$

Where: $Cc_{PumpNPV\ (x)}$ = Net present pump cost ($)
(x) = Pump replacement number (#, 1, 2, 3)

The total net present capital cost is therefore:

$$Cc_{PumpNPV} = Cc_{Pump} + Cc_{PumpNPV}(1) + Cc_{PumpNPV}(2) + Cc_{PumpNPV}(3)$$

The pump lifecycles were as follows:

- Initial pump: 0 years - 12.5 years
- First replacement pump: 12.5 years - 25.0 years
- Second replacement pump: 25.0 years - 37.5 years
- Third replacement pump 37.5 years - 50.0 years (system termination)

The salvage value is calculated in the same manner, as described by equation 3.19, but with an associated negative cost.

$$Cs_{Pump} = 5560 \times (P_{Pump})^{0.7231} \quad (3.27)$$

Where: Cc_{Pump} = Pump capital cost ($)

P_{Pump} = Pump input power (kW)

The present value of each subsequent pump is calculated as follows:

$$Cs_{PumpNPV}(x) = 0.2 \times Cc_{Pump} \times (1 + 0.04)^{(x) \times 12.5} \quad (3.28)$$

Where: $Cs_{PumpNPV\ (x)}$ = Net present pump salvage (-$)

(x) = Pump salvage number (#, 1, 2, 3,4)

The total net present salvage value is therefore:

$$Cs_{PumpNPV} = Cs_{PumpNPV}(1) + Cs_{PumpNPV}(2) + Cs_{PumpNPV}(3) + Cs_{PumpNPV}(4)$$

The pump salvage timeline was applied as follows:

- Initial pump salvage: 12.5 years
- First replacement pump salvage: 25.0 years
- Second replacement pump salvage: 37.5 years
- Third replacement pump salvage 50.0 years (system termination)

The operation and maintenance cost of the pump system($Com_{PumpNPV}$) was determined using the financial factors listed above and equation 3.20.

The energy cost of the system was calculated in the following manner:

The direct energy cost was acquired from Eskom's estimate for the year 2014/2015 and although it is understood that the energy cost varies between seasons and between peak and off-peak times, for simplification, an energy cost averaged across the year was used. The rationale is that the pump system is being tested under the most critical period of the year. The pump is thus run continuously to determine the system's critical capacity. However, the pumping installation will function throughout the year and thus the cost should be representative of this.

The pump is assumed to run 24 hours a day as the constant operation of the pump represents the optimum reliability state of a pumped bulk water supply system. Pump scheduling was not considered in this investigation

The annual energy cost of the pump is represented by the following equations:

$$Ce_{Pump} = P_{Pump} \times 24\ hours \times 365\ days \times PE_R \quad (3.29)$$

Where:

Ce_{PumpAR}	= Annual pump energy cost	($)
P_{Pump}	= Pump input power	(kW)
PE_R	= Pump energy rate	(R/kWh)

The total present energy cost is calculated as follows:

$$Ce_{PumpNPV} = Ce_{PumpAnn} \times RD_e \times \frac{(1+0.04)^{-50}-1}{0.04(1+0.04)^{-50}} \quad (3.30)$$

Where: $Ce_{PumpNPV}$ = Total pump energy cost ($)
RD_e = Rand-Dollar exchange rate (kW)
PE_R = Pump energy rate (R/kWh)

The system design life of 50 years was used, as the pump is assumed to run for the full life of the system. The energy rate used was the industrial usage rate of $\mathbf{0.54}\ R/kWh$. The exchange rate used was the average Rand/Dollar exchange rate for the past 10 years of **0.11** \$A/R. The total, lifecycle cost of the pumping system was calculated as follows:

$$C_{NPV-Pump} = Cc_{PumpNPV} + Cs_{PumpNPV} + Com_{PumpNPV} + Ce_{PumpNPV}$$

Reservoir Costing

The reservoir costing is in line with that implemented by Chang & van Zyl (2012), however with the inclusion of operation and maintenance cost.

Reservoir design life: (T_{Res}) = 50.0 years
Reservoir maintenance: (β_{Res}) = 0.015 (1.5%)
Reservoir salvage: (α_{Res}) = 0.0

The design life of a typical municipal service reservoir is proposed by Swamee& Sharma (2008) as being between 100 - 120 years. As mentioned previously, this value has been capped to 50 years, as it can reasonably be assumed that after a period of 50 years, the bulk supply system will be re-evaluated and the reservoir demolished and rebuilt or an additional reservoir added. Both of the aforementioned changes constitute a major alteration to the system which is outside of the optimisation and model boundaries.

The capital cost of the reservoir was calculated as per the below:

$$Cc_{ResNPV} = 290 \times \left(\frac{V_{Res}}{(1+{}^{V_{Res}}/_{1100})^{5.6}}\right)^{0.075} \quad (3.31)$$

Where: Cc_{ResNPV} = Annual pump energy cost ($)

V_{Res} = Reservoir volume (m^3)

The operation and maintenance cost of the reservoir(Com_{ResNPV}) was determined using the financial factors listed above and equation 3.20.

The total reservoir cost was determined as follows

$$C_{NPV-Reservoir} = Cc_{ResNPV} + Com_{ResNPV}$$

Total System Lifecycle Cost

The total lifecycle cost of each of the chromosomes was calculated as follows:

$$C_{NPV-Chrom} = C_{NPV-Pipeline} + C_{NPV-Chamber} + C_{NPV-Pump} + C_{NPV-Reservoir}$$

The four cost components of the system are combined and assigned to the chromosome being evaluated. This value gives a reasonable estimation of the life-cycle cost of the chromosome and forms one of the two objectives of the evolutionary process.

Expansion of events-run stochastic model

Following the successful completion of the cost evaluation, the chromosome is entered into the second objective's evaluation process, the stochastic determination of its reliability.

The events-run algorithm was appended to include the pump failure unit model, and the updated application of the hydraulic solver. The sorting algorithms within the events-run had to be amended to include a 3-way comparison in order to determine the order of fire events, pipe and pump failures.

The hydraulic model include two modes of operation: operating flow and failure flow. The operating flow mode assumes flow under normal operating conditions, while the failure flow mode determines the flow remaining during a supply pipe failure (burst pipe and subsequent valve closure). At the start of the algorithm, the operating flow and supply ratio are determined by the hydraulic solver, as well as all possible pipe failure flows (based on where the failure occurs in the system). The supply ratio is kept within the stipulated model boundary by verifying supply ratio at each point in the optimisation model where a variation to any of the decision variables occurs. This prevents a supply ratio being modeled that does not have an associated demands-only pre-run failure characteristic (out of bounds).

When a failure is generated by stochastic, pipe failure model, the position of the failure in the hydraulic network is generated randomly. The failure mechanism and hydraulic network modeling capacity of the model is utilized to a limited extent in the model, but requires more complex supply systems with extensive network redundancies to be fully utilised. As the assumed failure mode of the pumps is of a binary nature, owing to the assumed reliance on the electric grid and the binary characteristic of electrical supply. The pump is either fully functioning or not at all. It cannot operate during a power failure and thus there is no remaining flow for the duration of a power failure. The stochastic fire unit model is, as mentioned previously, part of the population component of the model, which affects demand only, and is thus outside of the optimisation space. This component remains unchanged between the proposed base and Chang models. The pump failure unit model is similar in nature to the pipe failure unit model in that the occurrence is based on a Poisson process and the duration on a lognormal distribution (calculated from the inverse cumulative distribution

function). As detailed previously in 2.4.1, The rate parameter used was an average of 0.104 power failures per annum supplied bulk power, as determined from the United States of America - Department of Energy, (USADOE, 2014) and the North American Energy Reliability Council (NERC, 2014) grid interruption data, demonstrating an average failure duration of 48 hours. This rate is representative of power failure events in North America, a developed country, where the optimal reliability-based approach could feasibly be adopted. A second power failure model is also tested, as developed by Nel (2009) and Nel & Haarhoff (2011) that represents power failure of pumping station in South Africa.

These values as presented in Table 3.6.1, are understood to vary significantly from country to country, and are sensitive to environmental factors such as system age, peripheral connections to the distribution network and potential load shedding (refer to 4.3.4).

Table 3.6.1: Power failure statistical distributions

	Parameter	Baseline
NERC (2014) & USADOE (2014) US	**Mean frequency (failures/annum)**	**0.104**
	Standard deviation (σ)	1.05
	Lognormal mean (μ)	2.92
	Mean duration (hours)	**48**
	Standard deviation (σ)	1.38
	Lognormal mean (μ)	2.92
NEL (2009) SA	**Mean frequency (failures/annum)**	**11.4**
	Standard deviation (σ)	0.7
	Lognormal mean (μ)	2.2
	Mean duration (hours)	**1.6**
	Standard deviation (σ)	1.54
	Lognormal mean (μ)	-0.61

The algorithm used to determine the reservoir failures during system events, is detailed in Figure 3.6.3 and follows a similar process to that described in 3.5.8 but with the addition of pump failures, sorting of the order of fire events and pump and pipe failures (also outline in Figure 3.6.4) and the integrated hydraulic solver. The hourly iteration during the week of the failure is also demonstrated, with the tank level being updated and any failures recorded.

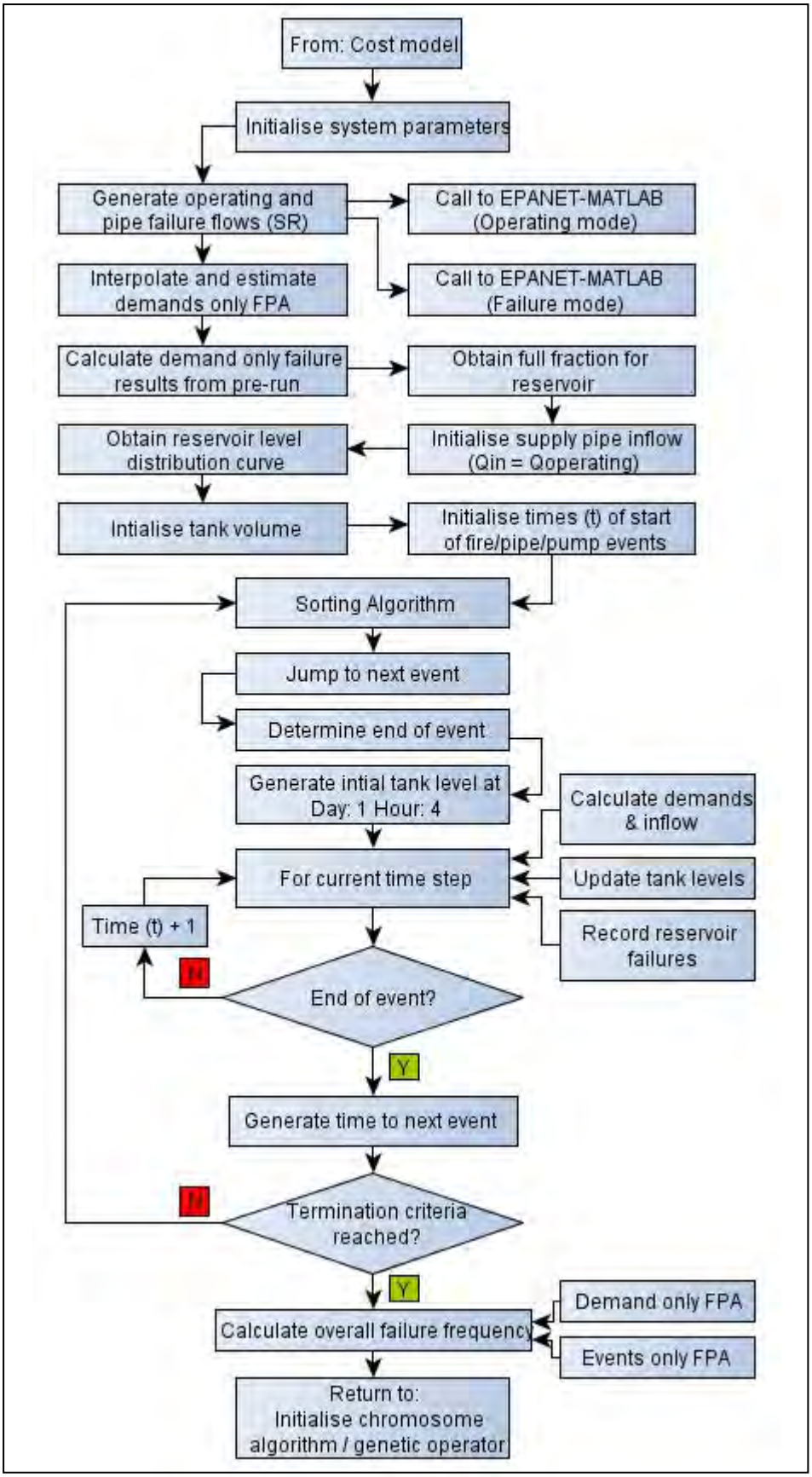

Figure 3.6.3: Expanded events run algorithm, including pump failures

The sorting algorithm was appended to include a 3 way comparison as shown in Figure 3.6.4.

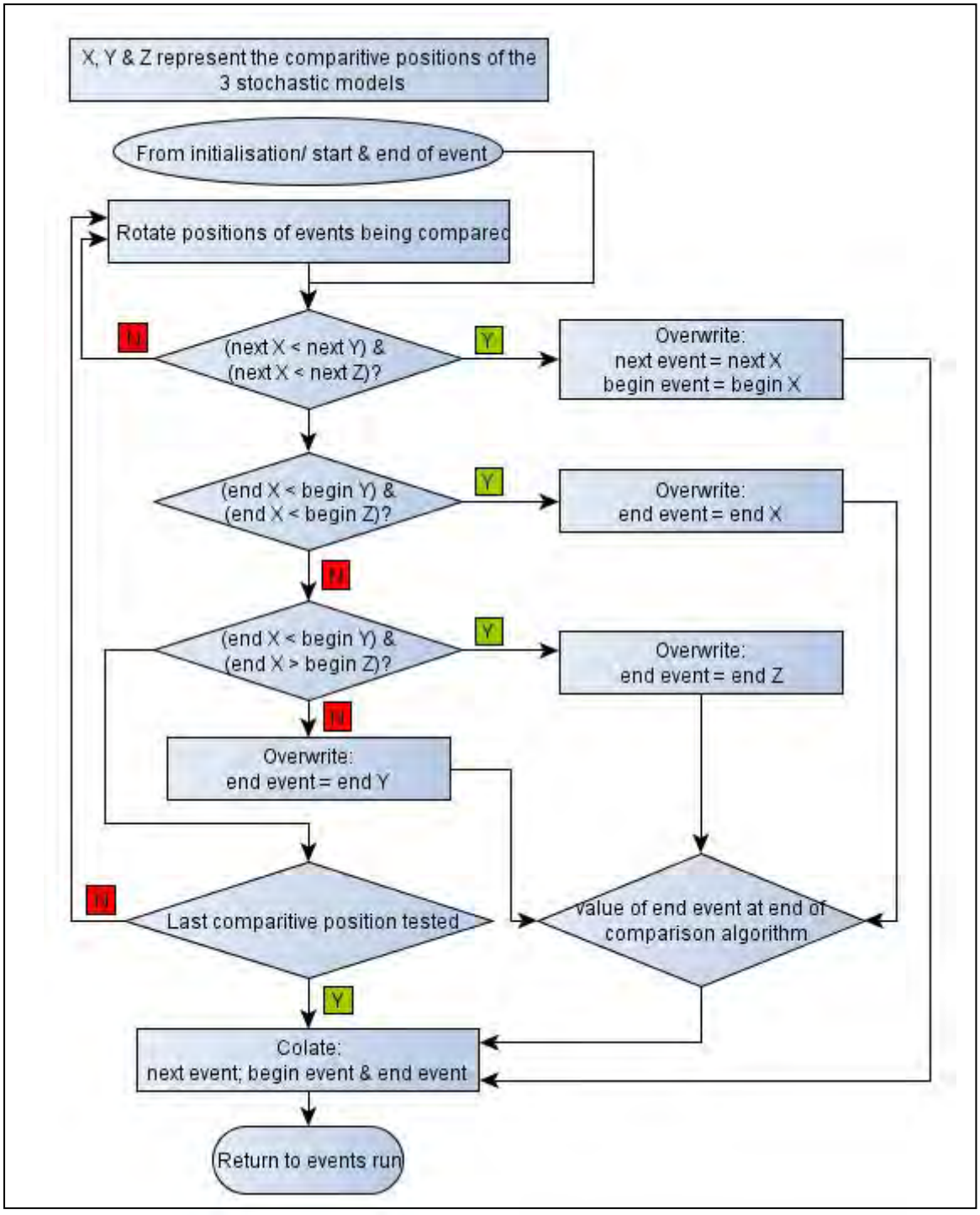

Figure 3.6.4: Event sorting algorithm

The sorting algorithm ensures that the events proceed in an ordered manner and that the duration of the event is appended if another event occurs before the end of the current event.

3.6.6. Genetic operator

The last significant amendment is to the genetic operator, the algorithm responsible for the simulated binary crossover of the decision variables of each parent chromosome and for polynomial mutation. As mentioned previously, the demands-only pre-run approach dictates that the product of the crossover and mutation process (child chromosome) must have a supply ratio that is within the boundary set [0.5, 5.0]. This is in addition to the controls set in place to ensure that the values produced by the crossover and mutation are appropriate with respect to the pipe and pump boundaries.

3.6.7. Verification of model and sensitivity analysis

A permutation of the base model was adapted to mimic the Chang model in its topology and application. This was done by removing the effect of the pump on the hydraulic system, the effect of pump failures on the evaluation of the reliability, and the life-cycle cost of the system. This was done in order to verify the integrity of the model against the Chang model, after the introduction of the hydraulic solver, updated pipe failure mechanism, ad-hoc calculation of supply ratio, increased number of decision variables as well as the inclusion of the pump failure mechanism on the sorting of events and other minor adaptations that had to be made to the base model. This is described in further detail in 4.1.

A sensitivity analysis is performed after the initial testing of the model, using, where possible, the same variations in physical and stochastic parameter used by Chang & van Zyl (2012) for comparability. The newly introduced stochastic distributions (pump failure frequency and duration) were varied by adding and subtracting one standard deviation from the lognormal mean. The adjusted normal mean was used for the Poisson based frequency process (to determine high and low times to failure) and the adjusted high and low duration cumulative distribution function was used to determine the high and low duration models respectively.

A load-shedding model was also developed based on the City of Cape Town (2015) load shedding schedule for an area with a population approximately the size of the generic population used to model the water demand (3000 - 5000 stands). The severity of load shedding was, as per the schedule, based solely on the frequency of load shedding with constant duration.

3.6.8. Methodology summary

The methods proposed above form an investigative platform upon which to test the synthesis of optimisation and stochastic analysis structure developed by Chang & van Zyl (2012) with the hydraulic solver and the inclusion of a pumped system and life-cycle costing.

This is done in order to investigate and draw conclusions regarding the applicability of the model to real world engineering problems and to form a foundation for future research. The developed method has shown how the structure of a stochastic analysis model coupled with a genetic algorithm can be expanded to include unit models that allow the model to more accurately represent real world conditions.

The inclusion of a hydraulic solver engine capable of solving complex hydraulic networks and pumps in to the network topology is a stepping stone in the creation of a research and design tool that can be used to analyse and design bulk water supply systems in a dynamic and holistic manner. The planned sensitivity analysis will provide insight into the areas of criticality, where the designer should focus attention on defining the stochastic parameters (based on demand patterns, failure frequencies, etc).

4. Results

4.1. Verification of base model

In order to verify the integrity of the base model, following the additions and revisions made to the structure of the Chang model, it must to be compared to the Chang model as developed by Chang & van Zyl (2012).

The model configuration was kept to the baseline configuration as determined by Chang & van Zyl (2012), to be representative of the typical system.

Table 4.1.1: Parameters used for verification of model

Parameter	Value	Unit
Physical Parameters		
Static system head	-60	m
Source to reservoir distance	10	km
Range of reservoir capacity	4 - 32	hours (SPD)
Optimisation Parameters		
Solution population size	50	Chromosomes
Number of evolutionary generations	50	Generations
Population Demand Parameters		
Seasonal peak demand	80	ℓ/s
Daily autoregression coefficient	0.12	-
Hourly autoregression coefficient	0.70	-
Daily standard deviation	0.068	-
Hourly standard deviation	0.13	-
Stochastic Parameters		
Fire rate	6	Fires/annum
Fire duration lognormal mean	-0.393	-
Fire duration standard deviation	0.66	-
Pipe failure rate	0.2	Failures/km/annum
Pipe failure duration lognormal mean	1.49	-
Pipe failure duration standard deviation	0.48	-

All parameters in accordance with those specified by van Zyl *et al.* (2008) and employed on the Chang mode with the exception of the maximum reservoir capacity, which was increased from 16 to 32 hours of seasonal peak demand. This was done in order to determine if a more cost efficient solution could be found by allowing for larger reservoir capacities.

The use of discrete pipe diameters was necessary in order to keep uniformity in the topology of the models being compared. The pipe diameters to be considered are those of a typical system, as determined by Chang & van Zyl (2012), summarised in Table 4.1.2.

Table 4.1.2: Typical, commercially available pipe sizes employed (mm ID)

127	145	182	227	286	322	363
428	479	530	574	626	675	726
777	828	878	929	976	1074	1176
1366	1568	1773	1970	2174		

4.1.1. Adaptation of base model for comparison

The base model was developed primarily to include the capability to process system configurations where the reservoir is at a greater altitude than the source. For an accurate comparison, the model was constrained to use the range of decision variables shown in Table 4.1.3.

Table 4.1.3: Adapted base model parameters

Decision Variable	Range/Value	Variable Type
Number of pipes	[1,3]	Discrete
Number of interconnecting pipes	[1,3]	Discrete
Pipe #1 diameter	[127,2174]	Discrete, Set
Pipe #2 diameter	[127,2174]	Discrete, Set
Pipe #3 diameter	[127,2174]	Discrete, Set
Reservoir capacity	[4,32]	Continuous

The reservoir capacity is continuous as there is no imperative to design a reservoir to a capacity defined in terms of a discrete number of hours of AADD (or seasonal peak demand in this case). It can be designed according to an exact need. This implies that it can be incorporated into both models, without changing the way in which the models work, and can be compared as such.

4.1.2. Incorporation of demands-only pre-run data

The demands-only pre-run output data needed to be used as input to both the Chang model as well as the adapted, base model to ensure a consistent base as both models were run as part of this investigation.

The demands-only pre-run, as described previously in section 3.5.5 and 3.6.2, simulates a population that draws from a reservoir that has assumedly perfect supply, i.e. the reservoir will never experience decreased inflow as a result of pump, pipe or any other supply related failure. The system configurations tested for reservoir failure frequency amongst other parameters, are as follows:

i. Reservoir capacity (4 h – 32 h, 1 h steps) – 29 capacities tested
ii. Supply ratio (0.5 – 5.0, 0.045 steps) – 101 supply ratios tested.

The total number of systems tested is thus 29 x 101 = 2 929. To test 2 929 systems of varying reliability is an extremely resource and processor intensive task. The demands run for this system took 170 hours to reach completion (detailed in section 3.6.2). This is owing to the high reliability systems (<0.01 failures per annum, under seasonal peak conditions) requiring an extended amount of time in order to achieve a statistically acceptable number of failures in order to be reliably classified).

Table 4.1.4 details the demands-only failure frequency under seasonal peak conditions. Not all the results are shown owing to the size of the dataset and only the results up until 32 hours of seasonal peak demand are relevant the verification exercise.

Table 4.1.4: Demands-only pre-run failure frequency (failures per annum)

	Reservoir capacity (h) (Seasonal Peak Demand)													
SR	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**28**	**32**	**48**
0.5	367.479	369.597	377.741	357.981	362.801	364.374	373.902	354.204	359.688	360.335	345.376	372.133	368.203	362.181
0.545	372.766	376.196	373.333	374.473	375.994	375.711	360.549	363.291	360.549	362.336	360.000	362.056	357.193	355.291
0.59	373.902	362.171	373.333	371.636	373.333	370.514	382.302	370.887	360.145	369.716	368.904	363.406	358.820	366.927
0.635	373.333	376.196	372.766	377.741	380.107	368.510	359.453	356.823	369.676	374.259	357.766	357.766	363.540	360.633
0.68	371.186	374.664	371.636	372.087	371.186	359.453	363.653	372.766	363.177	374.509	357.551	358.624	375.994	361.117
0.725	382.324	383.381	389.333	377.261	372.563	378.583	369.120	373.774	380.299	371.563	389.843	363.851	363.851	362.335
0.77	376.319	385.920	366.363	375.429	379.038	374.877	360.055	379.277	371.770	373.720	366.142	364.087	367.309	373.902
0.815	372.658	358.313	362.387	371.538	370.080	377.396	360.188	377.526	358.762	367.901	370.408	374.989	353.315	371.839
0.86	360.043	353.992	345.577	343.908	341.396	351.719	337.726	343.683	341.702	333.790	326.861	344.732	349.236	342.240
0.905	319.799	294.816	291.624	289.835	298.990	293.855	285.089	295.553	290.658	294.459	286.959	289.138	284.160	284.897
0.95	280.301	240.887	217.336	215.385	212.835	206.955	201.845	207.771	207.983	209.888	209.295	205.282	207.374	201.936
0.995	238.304	175.512	142.778	126.393	116.132	105.876	100.848	97.844	92.526	94.074	90.959	91.433	89.029	85.493
1.04	203.294	125.776	91.580	66.245	48.339	34.448	26.240	18.909	15.255	11.407	8.494	5.799	3.610	0.936
1.085	161.134	86.960	52.820	29.218	17.929	9.356	5.016	2.617	1.406	0.694	0.366	0.107	0.019	0.000
1.13	132.111	60.841	27.847	14.327	6.453	2.638	1.157	0.453	0.179	0.063	0.015	0.002	0.000	0.000
1.175	107.676	40.814	16.817	6.881	2.478	0.918	0.301	0.105	0.030	0.005	0.010	0.000	0.000	0.000
1.22	84.247	28.088	9.947	3.222	1.004	0.276	0.084	0.013	0.010	0.002	0.000	0.000	0.000	0.000
1.265	62.655	19.293	5.697	1.554	0.443	0.101	0.034	0.018	0.001	0.000	0.000	0.000	0.000	0.000
1.31	48.991	13.157	3.548	0.848	0.198	0.042	0.022	0.002	0.000	0.000	0.000	0.000	0.000	0.000
1.355	34.788	9.019	2.207	0.490	0.112	0.022	0.013	0.001	0.000	0.000	0.000	0.000	0.000	0.000
1.4	24.024	6.184	1.339	0.282	0.060	0.020	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000

	Reservoir capacity (h)(Seasonal Peak Demand) - Continued													
SR	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**28**	**32**	**48**
1.445	18.089	4.172	0.858	0.162	0.012	0.003	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.49	12.834	2.665	0.528	0.093	0.010	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.535	8.908	1.689	0.287	0.055	0.009	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.58	6.131	1.038	0.187	0.029	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.625	4.078	0.667	0.117	0.013	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.715	1.787	0.270	0.019	0.007	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.76	1.154	0.178	0.028	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.805	0.772	0.091	0.008	0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.85	0.503	0.058	0.007	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.895	0.309	0.036	0.007	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.94	0.206	0.016	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.985	0.142	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.03	0.075	0.013	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.075	0.053	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.12	0.048	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.165	0.033	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.21	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.255	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.3	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.345	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.39	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.435	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Out of 2929 tested configurations, 2 235 or 76.3% had failure frequencies less than 1 failure in 1000 years. A degree of variance exists within the very low failure frequency region as it results in the model reaching the user set duration limit before convergence to within 5% of its ultimate value, as outlined in 3.5.4. This degree of reliability (as the inverse of failure frequency) will not likely be considered when optimizing the system as it is substantially greater than any feasible system would be built. It therefore does not influence the relevance and accuracy of the model.

4.1.3. Verification model simulation results

Both the Chang and adapted base model were capacitated to record the full extent of each chromosome vector, including all decision variables, supply ratio, cost and reliability. The initial tests performed were those of the accuracy of and consistency between the models' hydraulic, stochastic and heuristic algorithms. The model and system parameters were consistent with the base model presented in Table 4.1.1 and Table 4.1.3. The same, single pipe, chromosome was used as input for both models, as per Table 4.1.5. All failure frequencies mentioned refer to hydraulic failure (running dry) frequency of the reservoir.

Table 4.1.5: Stochastic Model Output Comparison

Model	Pipe Diameter (m)	Tank Capacity (h)	Failure Frequency	Total Cost	Pipe Cost	Tank Cost
Base	0.322	15.00	0.0806	$2 274 755.70	$1 861 032.40	$413 723.31
Chang			0.0791	$2 274 755.70	$1 861 032.40	$413 723.31
Difference			0.0015	$ 0.00	$ 0.00	$ 0.00
Variance (%)			1.90	0.00	0.00	0.00

The 1.90 % variance between the models is within the 5% variance found by Chang & van Zyl (2010) when varying the seed for the pseudo-random generators in the stochastic model and can therefore be considered acceptable as the same variance is inherent in the base model. The stochastic components and associated algorithms of the adapted, pump-inclusive model were considered to have retained their integrity and comparability to the Chang model.

The Pareto optimal solution sets, of both models, after the 50th generation, are presented in Figure 4.1.1, showing the complete scope of failure frequencies.

Figure 4.1.2 shows the range of solutions situated around the desired failure frequency of 1 failure every 10 years under seasonal peak demand conditions. The failure frequency range presented spans from a failure more than once per day to a failure less than once every 10 000 years. The graphs show the different failure frequency zones. Figure 4.1.1 shows the entire range of failure frequencies spanning from 1 failure per day to less than 1 failure every 1 000 years.

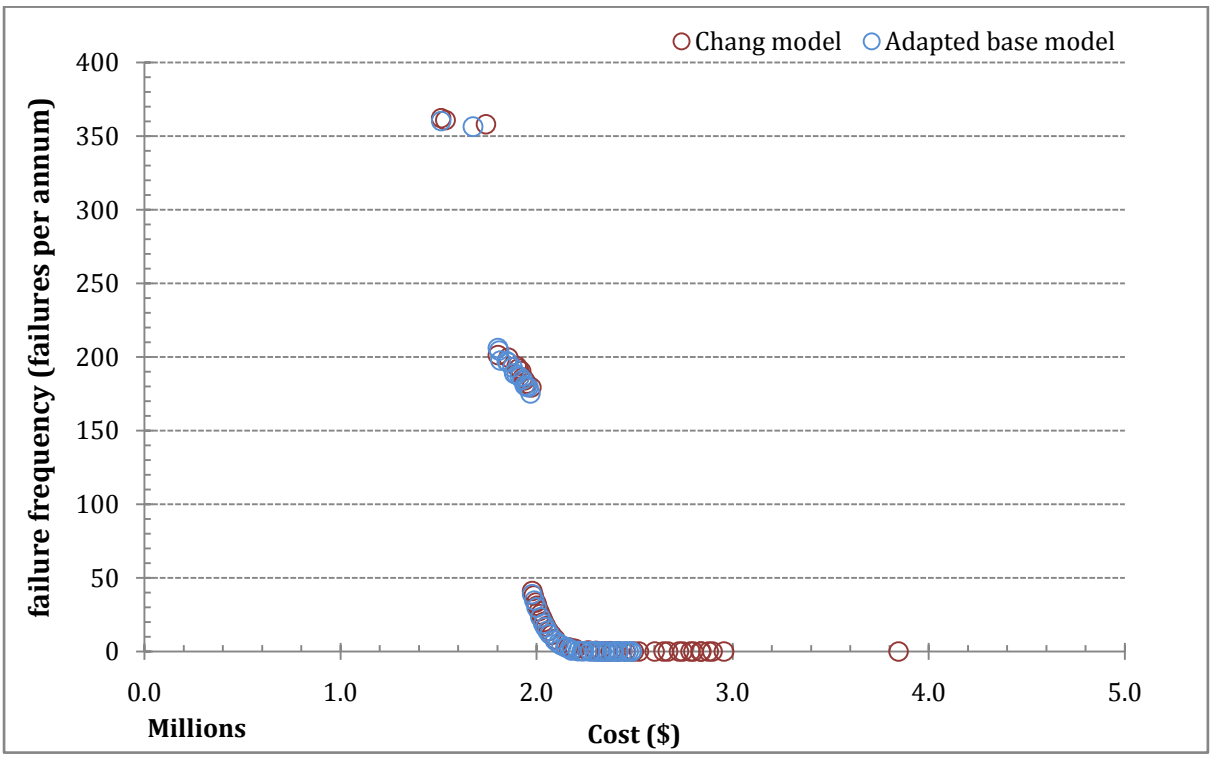

Figure 4.1.1: Pareto-optimal solution front – model comparison

The adapted base model closely matches the Chang model, exhibiting the same pattern in the distribution and spacing of the solutions. This distribution is addressed further in the comments and observations in 4.1.4. The majority of the solutions are unfeasible for implementation as the failure frequency is too high. The range of solutions around the design criterion is of use in this regard. This is shown in Figure 4.1.2.

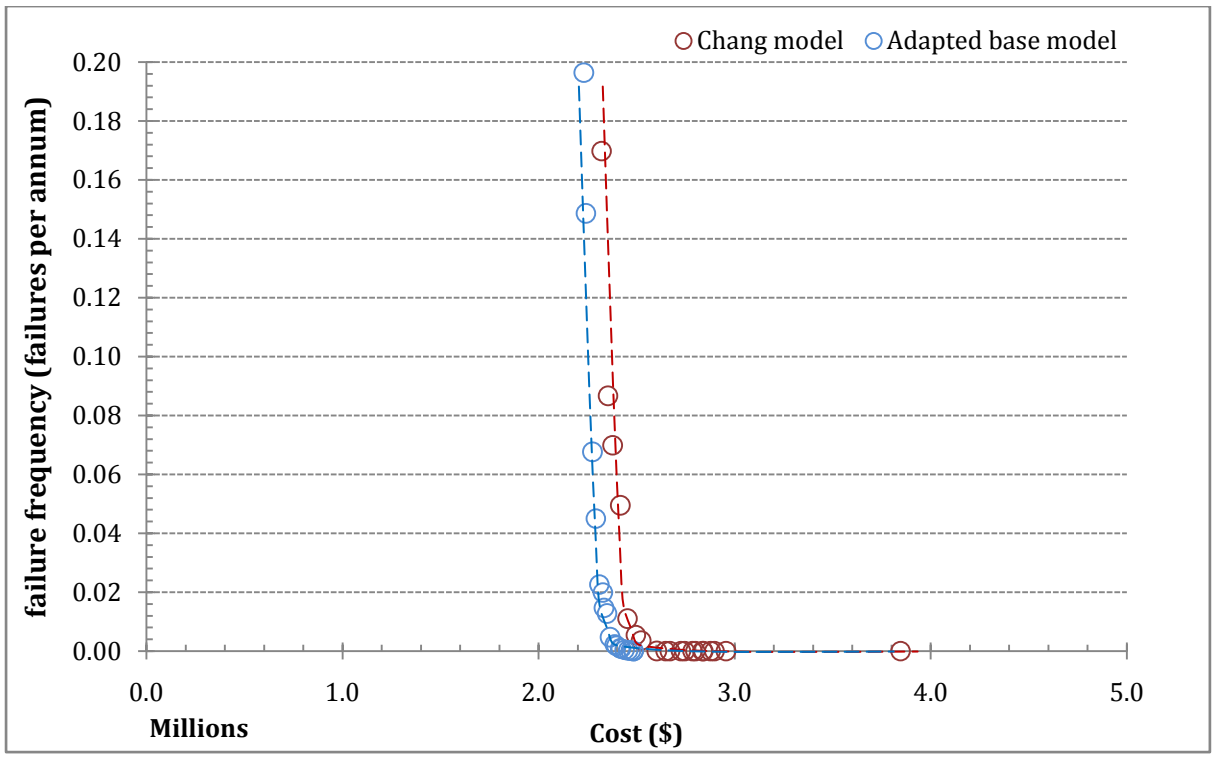

Figure 4.1.2: Pareto-optimal solution front - comparative systems (design-criterion reliability)

The adapted model produced a solution front that closely matches the Chang model but is inwardly offset. The implication of this is that, all else considered constant, the adapted base model is able to produce a solution set of marginally more efficient solutions. The reasoning behind this is discussed in 4.1.4.

The solutions fronts displayed in Figure 4.1.1 and Figure 4.1.2 are listed in tables 4.1.6 and 4.1.7, for adapted base model and Chang model solutions respectively. Different shades of blue are used to indicate solutions with the same supply ratio. The number of pipes, internal pipe diameter, tank capacity in hours of seasonal peak demand, supply ratio (seasonal peak demand), capital cost and failure frequency is presented for each model.

Table 4.1.6:50th Generation solution set - adapted base model

No. Pipes	Pipe ID (m)	Tank Capacity (h - SPD)	Supply Ratio	Capital Cost	Failures per Annum
1	0.43	4.000	2.83	$2,484,805.10	< 0.01
1	0.43	4.000	2.83	$2,484,805.10	< 0.01
1	0.36	11.678	1.84	$2,470,920.50	< 0.01
1	0.36	11.215	1.84	$2,456,729.20	< 0.01
1	0.36	10.650	1.84	$2,439,121.40	< 0.01
1	0.36	10.047	1.84	$2,419,845.30	< 0.01
1	0.36	9.935	1.84	$2,416,225.40	< 0.01
1	0.36	9.334	1.84	$2,396,398.60	< 0.01
1	0.36	9.099	1.84	$2,388,499.20	< 0.01
1	0.36	8.406	1.84	$2,364,664.60	< 0.01
1	0.36	7.981	1.84	$2,349,562.60	0.01
1	0.36	7.540	1.84	$2,333,506.70	0.01
1	0.36	7.400	1.84	$2,328,339.50	0.02
1	0.36	6.924	1.84	$2,310,289.70	0.02
1	0.36	6.457	1.84	$2,291,929.50	0.04
1	0.36	6.037	1.84	$2,274,755.70	0.07
1	0.36	5.272	1.84	$2,241,231.80	0.15
1	0.36	5.052	1.84	$2,230,921.40	0.20
1	0.36	4.636	1.84	$2,210,253.60	0.30
1	0.36	4.055	1.84	$2,178,213.10	0.56
1	0.32	8.427	1.34	$2,168,108.70	1.96
1	0.32	7.993	1.34	$2,152,719.40	2.65
1	0.32	7.835	1.34	$2,147,029.90	2.94
1	0.32	7.307	1.34	$2,127,556.20	4.19
1	0.32	6.895	1.34	$2,111,877.40	5.52
1	0.32	6.413	1.34	$2,092,864.20	7.63
1	0.32	5.785	1.34	$2,066,760.30	11.70
1	0.32	5.533	1.34	$2,055,770.00	13.95
1	0.32	5.300	1.34	$2,045,223.50	16.35
1	0.32	5.083	1.34	$2,035,100.50	18.74
1	0.32	4.777	1.34	$2,020,174.50	23.14
1	0.32	4.400	1.34	$2,000,473.30	29.83

No. Pipes	Pipe ID (m)	Tank Capacity (h - SPD)	Supply Ratio	Capital Cost	Failures per Annum
1	0.32	4.184	1.34	$1,988,402.70	34.55
1	0.32	4.000	1.34	$1,977,642.20	38.82
1	0.29	7.731	0.98	$1,968,664.40	175.17
1	0.29	7.462	0.98	$1,958,756.70	179.35
1	0.29	7.212	0.98	$1,949,430.30	179.83
1	0.29	6.966	0.98	$1,940,036.50	180.64
1	0.29	6.678	0.98	$1,928,852.50	184.50
1	0.29	6.328	0.98	$1,914,858.30	186.53
1	0.29	5.966	0.98	$1,899,871.60	187.79
1	0.29	5.643	0.98	$1,886,015.00	188.78
1	0.29	5.445	0.98	$1,877,227.90	192.95
1	0.29	4.983	0.98	$1,855,744.30	196.31
1	0.29	4.833	0.98	$1,848,355.90	197.06
1	0.29	4.226	0.98	$1,816,230.70	197.56
1	0.29	4.102	0.98	$1,809,098.00	204.39
1	0.29	4.000	0.98	$1,803,065.60	206.06
1	0.23	7.636	0.53	$1,675,871.00	356.39
1	0.23	4.000	0.53	$1,513,744.10	360.19

Table 4.1.7 :50th generation solution set - Chang model

No. Pipes	Pipe ID (m)	Tank Capacity (h - SPD)	Supply Ratio	Capital Cost	Failures per Annum
2	0.322	8.83	2.67	$3,845,764.10	<0.01
1	0.363	32.00	1.83	$2,955,536.30	<0.01
1	0.363	29.09	1.83	$2,896,649.80	<0.01
1	0.363	28.16	1.83	$2,877,257.10	<0.01
1	0.363	26.45	1.83	$2,841,014.50	<0.01
1	0.363	26.24	1.83	$2,836,467.30	<0.01
1	0.363	24.58	1.83	$2,800,160.40	<0.01
1	0.363	23.91	1.83	$2,785,314.50	<0.01

No. Pipes	Pipe ID (m)	Tank Capacity (h - SPD)	Supply Ratio	Capital Cost	Failures per Annum
1	0.363	22.13	1.83	$2,744,874.00	<0.01
1	0.322	30.42	1.34	$2,726,554.70	<0.01
1	0.322	27.67	1.34	$2,669,834.80	<0.01
1	0.322	26.62	1.34	$2,647,497.10	<0.01
1	0.322	24.57	1.34	$2,602,721.00	<0.01
1	0.322	21.07	1.34	$2,522,689.40	<0.01
1	0.322	19.95	1.34	$2,495,959.80	0.01
1	0.322	18.24	1.34	$2,453,851.10	0.01
1	0.322	16.79	1.34	$2,416,670.60	0.05
1	0.322	15.31	1.34	$2,377,500.60	0.07
1	0.322	14.45	1.34	$2,354,005.30	0.09
1	0.322	13.31	1.34	$2,321,821.70	0.17
1	0.322	12.67	1.34	$2,303,042.50	0.25
1	0.322	11.55	1.34	$2,269,865.50	0.43
1	0.322	11.26	1.34	$2,260,846.00	0.50
1	0.322	9.28	1.34	$2,197,236.40	1.61
1	0.322	8.99	1.34	$2,187,436.20	1.81
1	0.322	8.31	1.34	$2,163,970.20	2.73
1	0.322	6.50	1.34	$2,096,503.20	8.35
1	0.322	6.01	1.34	$2,076,280.40	11.40
1	0.322	5.69	1.34	$2,062,442.20	13.81
1	0.322	5.39	1.34	$2,049,148.50	17.00
1	0.322	5.12	1.34	$2,036,666.80	20.00
1	0.322	4.94	1.34	$2,028,215.60	22.58
1	0.322	4.77	1.34	$2,020,011.20	25.31
1	0.322	4.62	1.34	$2,012,320.80	27.38
1	0.322	4.44	1.34	$2,002,390.60	31.09
1	0.322	4.34	1.34	$1,997,397.80	33.16
1	0.322	4.13	1.34	$1,985,137.30	37.99
1	0.322	4.00	1.34	$1,977,642.20	40.98
1	0.286	7.86	0.98	$1,973,450.80	179.28
1	0.286	7.26	0.98	$1,951,395.10	181.88
1	0.286	6.93	0.98	$1,938,640.10	184.18

No. Pipes	Pipe ID (m)	Tank Capacity (h - SPD)	Supply Ratio	Capital Cost	Failures per Annum
1	0.286	6.70	0.98	$1,929,738.40	185.92
1	0.286	6.51	0.98	$1,922,028.70	190.35
1	0.286	6.16	0.98	$1,907,833.70	192.07
1	0.286	5.94	0.98	$1,898,908.50	193.75
1	0.286	4.99	0.98	$1,855,905.50	199.58
1	0.286	4.00	0.98	$1,803,065.60	201.19
1	0.227	9.54	0.53	$1,742,189.70	358.01
1	0.227	4.38	0.53	$1,535,740.00	360.77
1	0.227	4.00	0.53	$1,513,744.10	362.07

4.1.4. Comments and observations

General Absence of solutions containing more than one pipe

Both models had the capability to initialise and genetically alter systems containing up to 3 parallel pipes with multiple interconnections. In the case of the adapted base model, each parallel pipe could have a different size, as opposed to the Chang model that consisted of systems with uniform pipe sizes. This was done in an attempt to increase the variability of the model, through increasing the number of supply ratios that fall within the boundary condition [0.5, 5.0].

However, after the 1st generation in the adapted base model and the 15th generation in the original model, all (except one in the Chang model) pipe systems with more than one pipe had been dominated and eliminated from the solution population.

This is seen to be owing to a multitude of factors. The interdependent relationship between pipe size and tank capacity, and the associated cost of each affects the selection process and by extension the evolutionary process as a whole. The trade-off between selected pipe size and the number of pipes also exhibits a significant influence. The supply ratios falling into the acceptable range [0.5; 5.0], and the associated pipe sizes and number of parallel pipes is displayed in Table 4.1.8.

Table 4.1.8: Acceptable supply ratios

Pipe Sizes	Supply Ratios		
	1 Pipe	2 Pipe	3 Pipe
0.2270	0.533	1.066	1.60
0.2860	0.979	1.96	2.94
0.3220	1.34	2.67	4.01
0.3630	1.83	3.66	5.00

The closest supply ratios from systems varying by a single pipe size and number of parallel pipes are represented by bands of the same colour. The optimisation algorithm discriminates, through the crossover/mutation and recombination/selection processes, between modulating the supply ratio and modulating the tank size to achieve a superior solution. By extension, the trade-off within the supply ratio value is between the pipe size and the number of pipes in the solution system (chromosome). The capital cost of the pipes is shown in Table 4.1.9.

Table 4.1.9: Associated pipe costs

Pipe Sizes	Pipe Cost		
	1 Pipe	2 Pipe	3 Pipe
0.2270	$1 199 845.74	$2 399 691.48	$3 599 537.22
0.2860	$1 489 167.26	$2 978 334.51	$4 467 501.77
0.3220	$1 663 743.90	$3 327 487.81	$4 991 231.71
0.3630	$1 861 032.39	$3 722 064.78	$5 583 097.17

The cost of each pipe solution is presented above, with colours representing supply ratios with the closest value from systems varying by a single pipe size and number of parallel pipes, as per Table 4.1.8. It can be observed that to modulate the supply ratio through horizontal (pipe number) selection results in a 100% increase in pipe cost, when considering capital cost only. A selection movement (crossover/mutation) in this direction places the solution system in a far more reliable and far more expensive solution group. A selection movement vertically is a significantly more efficient move, as the increase in supply ratio is on average 52% while the increase in cost is on average 16%. This variation in cost efficiency as a result of the crossover process is what ultimately influences the evolutionary trend. The benefit of having multiple pipes is partial inflow to the reservoir under pipe

failure conditions (as opposed to no inflow), assuming interconnections are present. However in the Chang and adapted base model, the probability of pipe failure is a function of pipe length. This implies that with long pipelines, where the failure rate is high and the benefit of partial flow is great, the cost of having multiple pipe systems is even higher. In short pipe length systems, where the cost may be feasible, the benefit is less owing to fewer pipe failure events. As such, solutions consisting of more than a single pipe are dominated and crowded out by the more efficient single pipe systems. Owing to the limited extent of the mutation, once a multiple pipe system is genetically altered into a single pipe system, it is unlikely that it will be altered back into a multiple pipe system again due to the cost inefficiency of doing so.

Model Comparison and Confirmation of Adapted, Pump Inclusive Model Accuracy

The base model, adapted to model a gravity fed system has produced results with a high degree of correlation between the solution set. However, to compare the consistency of the adapted, base model to the Chang model, the feasible portion of the solution set ([0.01, 0.1] failures per annum) must be compared. All solutions presented below consist of a single pipe system after 50 generations/iterations of optimisation. These results are presented in tables 4.1.10 and 4.1.11. The two most closely spaced solutions are highlighted in green.

Table 4.1.10: Feasible solutions - adapted based model

Pipe Diameter (m)	Tank Capacity (h - SPD)	Supply Ratio (SPD)	Total Cost	Failure Frequency	Pipe Cost	Tank Cost
0.36	6.037	1.84	$2 274 755.70	0.0676	$1 861 032.40	$413 723.31
0.36	6.457	1.84	$2 291 929.50	0.0450	$1 861 032.40	$430 897.15
0.36	6.924	1.84	$2 310 289.70	0.0226	$1 861 032.40	$449 257.28
0.36	7.400	1.84	$2 328 339.50	0.0199	$1 861 032.40	$467 307.12
0.36	7.540	1.84	$2 333 506.70	0.0147	$1 861 032.40	$472 474.35
0.36	7.981	1.84	$2 349 562.60	0.0128	$1 861 032.40	$488 530.19
Mean						
0.36	7.057	1.84	$2 314 730.62	0.0304	$1 861 032.40	$453 698.23

Table 4.1.11: Feasible Solutions - Chang model

Pipe Diameter (m)	Tank Capacity (h - SPD)	Supply Ratio (SPD)	Total Cost	Failure Frequency	Pipe Cost	Reservoir C
0.322	14.45	1.34	$2 354 005.30	0.0867	$1 663 743.90	$690 261.41
0.322	15.31	1.34	$2 377 500.60	0.0699	$1 663 743.90	$713 756.68
0.322	16.79	1.34	$2 416 670.60	0.0495	$1 663 743.90	$752 926.67
0.322	18.24	1.34	$2 453 851.10	0.0110	$1 663 743.90	$790 107.17
Mean						
0.322	16.20	1.34	$2 400 506.90	0.0543	$1 663 743.90	$736 762.98

The solutions that meet the desired design criterion of 0.1 failure per annum (1 failure every 10 years) under seasonal peak demand as per tables 4.1.10 and 4.1.11, are presented in Figure 4.1.3. The trend of solutions for both models, with increasing cost, is linearly decreasing failure frequency.

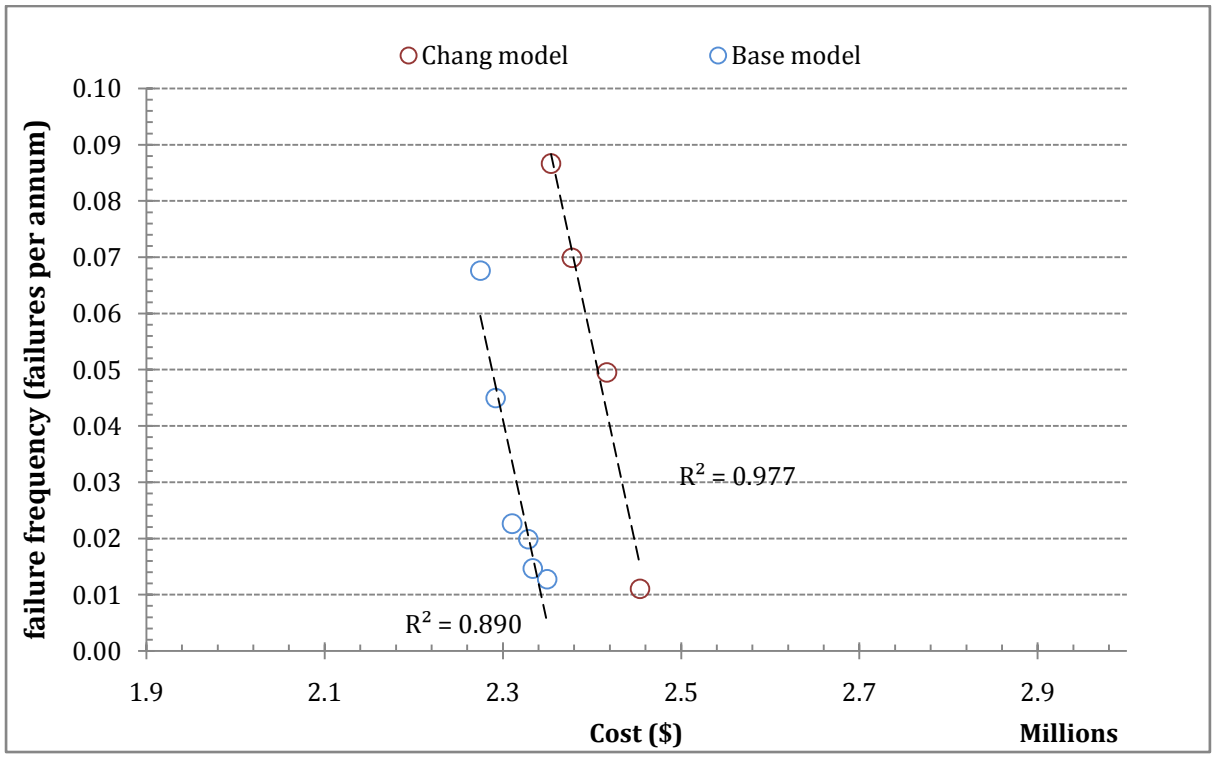

Figure 4.1.3:Pareto-optimal solution front for feasible solution region

What can be observed from the solutions presented above is a reasonable degree of correlation between models. It cannot be expected for the models to produce identical solutions as there is an inherent variance owing to the random functions employed.
There were also slight adaptations made to the various functions employed in order to enable the model to simulate continuous and discrete variable pumping systems (adapted in this instance for gravity fed flow). The hydraulic calculator employed also produces very slightly different flow readings under the H-W flow estimation. The rounding function employed in the genetic operator (crossover/mutation) to ensure the decision variables remain integers was also reviewed and implemented into the adapted model.

The models have produced a similar set of results; however the system composition is significantly different. The adapted model has produced what is observed to be a more efficient solution set, by using larger pipe sizes (and consequently larger supply ratios) and smaller capacity reservoirs.

This becomes clear when comparing two solutions that achieved a very similar reliability, as presented in Table 4.1.12.

Table 4.1.12: Comparison of similar reliability systems

Pipe ID (m)	Tank Capacity (h - SPD)	Supply Ratio	Cost	Failure frequency	Pipe Cost	Reservoir Cost
Adapted base model (Generation 45)						
0.36	6.037	1.84	$2 274 755.70	0.0676	$1 861 032.40	$413 723.31
Chang model (Generation 45)						
0.322	15.31	1.34	$2 377 500.60	0.0699	$1 663 743.90	$713 756.68
Difference			-$102 744.90	-0.00227	$197 288.50	-$300 033.37

It can be observed from the above Table that the reliability of the systems being compared is within 3.4% and the cost within 4.5%, showing a reasonable correlation between the models.

It can further be observed that the solution system for the adapted model, employing a stronger crossover trend toward increasing supply ratio over reservoir capacity, produces marginally more efficient solutions than the Chang model, in this instance.

This is, with a reasonable degree of certainty, attributed to the rounding mechanism employed during the operation of the genetic operator algorithm. The genetic operation and mutation is performed as a percentage of the decision variable (e.g. pipe size) value. This works well on continuous variables, as the genetic operation produces unrounded child solutions from the parents.

However when a system is composed of a number of discrete, bounded variables, the variables must be rounded off to an integer in order for the system to be representative of reality. The way it is rounded (i.e. near round, floor round, ceiling round and boundary limits affects whether solution systems become incorrectly biased toward a particular variable and/or upward or downward sticky.

The adapted, pump inclusive model has a genetic operator rounding function that does not round off the position of the pipe sizes in the allowable pipe sizes vector, but rather allows the genetic operator to act on the pipe size itself, following which, the solution is adjusted to the nearest pipe size. This allows the operator to perform the crossover and mutation functions on the decision variables actual value, removing any upward or downward bias.

4.1.5. Summary

It can be concluded, with a reasonable degree of certainty, that the base model, adapted in this instance to simulate a non-pumped, gravity fed system, has proven the accuracy of its integrated stochastic models. It has also proven the accuracy of its integrated genetic optimisation model, as was demonstrated by the closeness of its 50^{th} generation solution set to the Chang model

4.2. Base model

The base model, incorporating pumping systems, was developed to investigate the optimisation of a pumped bulk water supply system, utilizing life-cycle costing and stochastic analysis to determine cost and reliability and to compare the optimisation results to the results of the Chang model. The parameters utilised in the base model were, as far as was possible, aligned with the base model employed by Chang & van Zyl (2012) for comparability, where possible, between gravity fed and pumped systems.

4.2.1. Base model simulation results

The parameters for analysis were implemented as per Table 4.2.1.

Table 4.2.1: Parameters for base model against Chang model

Parameter	Chang Model	Base Model	Unit/Type
Physical Parameters			
Static system head	-60	+60	m
Source to reservoir distance	10	10	km
Range of reservoir capacity	4 - 32	4 - 48	hours (SPD)
Range of supply ratios	0.5 - 5.0	0.5 - 5.0	-
Range of pump powers	N/A	15 - 1500	kW
Pump efficiency	N/A	75	%
No. of parallel pipes	0 - 3	0 - 3	-
No. of interconnections	0 - 2	0 - 3	-
Range of pipe sizes	[127, 2174]	[127,2174]	mm- ID (Table 4.1.2)
System design life	N/A	50	years
Optimisation Parameters			
Solution population size	50	50	Chromosomes
Evolutionary generations	50	50	Generations
Population Demand Parameters			
Seasonal peak demand	80	80	ℓ/s
Daily autoregression coefficient	0.12	0.12	-
Hourly autoregression coefficient	0.70	0.70	-
Daily standard deviation	0.068	0.068	-
Hourly standard deviation	0.13	0.13	-

Parameter (continued)	Chang Model	Base Model	Unit/Type
Stochastic Parameters			
Fire rate	6	6	Fires/annum
Average fire duration	0.839	0.839	hours
Fire duration (μ)	-0.393	-0.393	Lognormal
Fire duration (σ)	0.66	0.66	Lognormal
Pipe failure rate	0.2	0.2	Failures/km/annum
Average pipe failure duration	4.5	4.5	hours
Pipe failure duration (μ)	1.49	1.49	Lognormal
Pipe failure duration (σ)	0.48	0.48	Lognormal
Pump (Power) failure rate	N/A	0.104	Pump failures/annum
Average pump failure duration	N/A	48	hours
Pump failure duration (μ)	N/A	2.92	Lognormal
Pump failure duration (σ)	N/A	1.38	Lognormal

The above parameters form the simulation baseline to be compared against results from the Chang model, the power failure distribution proposed by Nel (2009) for South African power failure affecting bulk water pumping systems. It is also this basis upon which the sensitivity analysis is to be performed.

The demands only failures per annum were re-calculated using the demands-only pre-run, extended to 48 hours of seasonal peak demand (72 hours of annual average daily demand) from 32 hours of seasonal peak demand (48 hours of annual average daily demand) used in the base model verification. This was done as an attempt to ensure that an optimal solution set could be produced without constraints imposed by limitation of the decision variable parameters. Increasing the maximum reservoir capacity past 48 hours of seasonal peak demand was not seen as necessary as the 48 hour limit is already 50% greater than that recommended by the CSIR (2000), fire flow capacity not considered.

The resulting pareto-optimal curve for the 50th generation is displayed in Figure 4.2.1, in comparison to the 50th generation pareto-optimal curve generated by the Chang model.

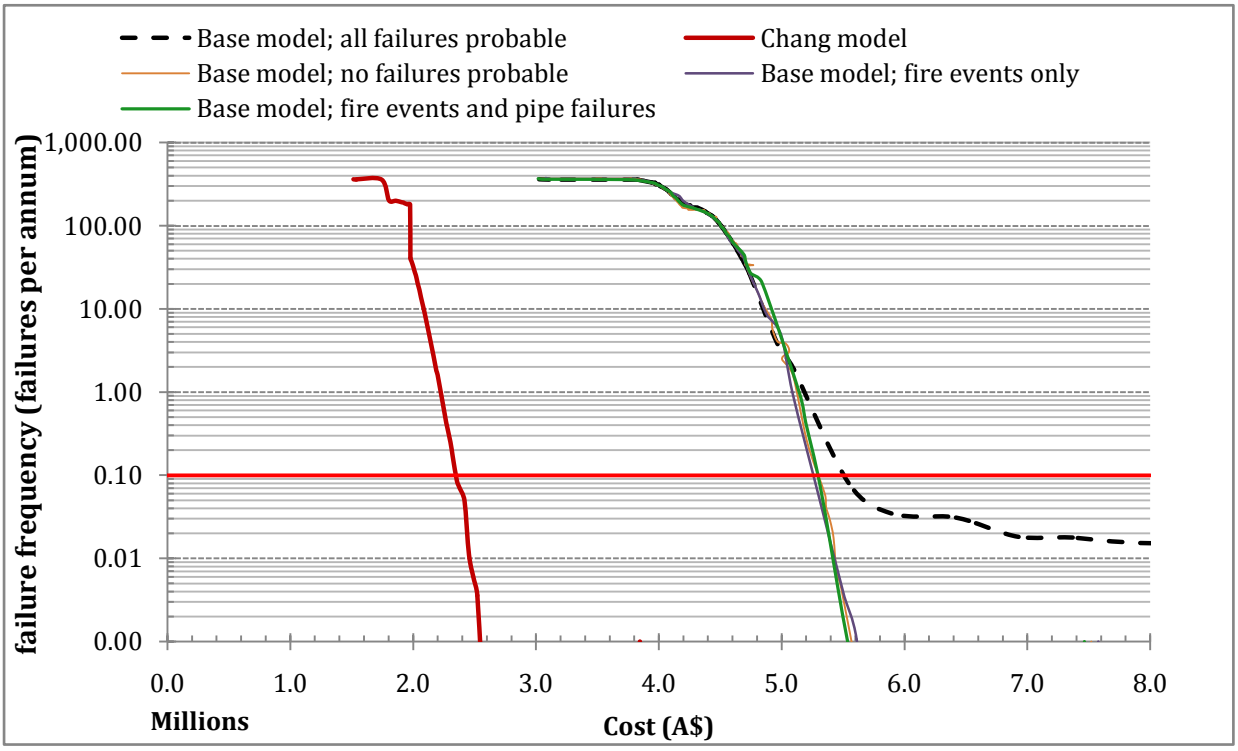

Figure 4.2.1: Pareto-optimal solution front - Chang and base models

The average cost difference is substantial, with a typical solution meeting the 0.1 failure per annum design criterion costing an additional $3.2 million under the pumped, base model. This is to be expected owing to the inclusion of operation, maintenance, repeated capital outlay for replacement of pumps and energy cost over the lifecycle of the system, as well as an elevated head as compared with the negative elevation of a gravity-fed system. The reliability of the pumping system has a significant effect on the failure frequency seen in the above Figure. Owing to the influence of pump system failures, the base model reaches a failure frequency asymptote at around 0.02 failures per annum, with cost increasing exponentially to this point. In contrast to this, the Chang model is observed to produce solutions that extend into the zone of extremely low failure frequency, around 1E-3 failures per annum. In this zone, the storage capacity of the reservoir is sufficient to sustain supply during all but the most extreme pipe failure events and/or fire events.

Removing the effect of pump failures has a substantial effect on failure frequency and in particular the position of the asymptote. Without the effect of pump failures on the system, the system is capable of much lower failure frequencies, equal to the Chang model. Further removing the effects of pipe failure yields far less of a response from the model, indicating

that the failure frequency characteristics of the system are dominated by the occurrence and duration of pump failure events. This is to be explored further in the sensitivity analysis.

Ideally, for a solution system to be as efficient as possible according to the adopted failure criterion, it should experience a failure frequency of as close to 0.1 failures per annum as possible. A solution with a failure frequency too far below exhibits an inefficiently high cost, and too far above, an impractically high failure frequency. When designing optimally for the desired failure frequency criterion, the effect of pump failures on cost results in an increase of approximately 5%. This can be extended to pump, pipe and failure events, owing to the domination exhibited by pump failure events. Chang & van Zyl (2014) demonstrated a linear (when plotted on log-scale) relationship for failure frequency against reservoir capacity for a typically found supply ratio. The supply ratios were constrained consecutively, and the results were compared to the relationship presented by Chang & van Zyl (2014), for a gravity fed system, in Figure 4.2.2.

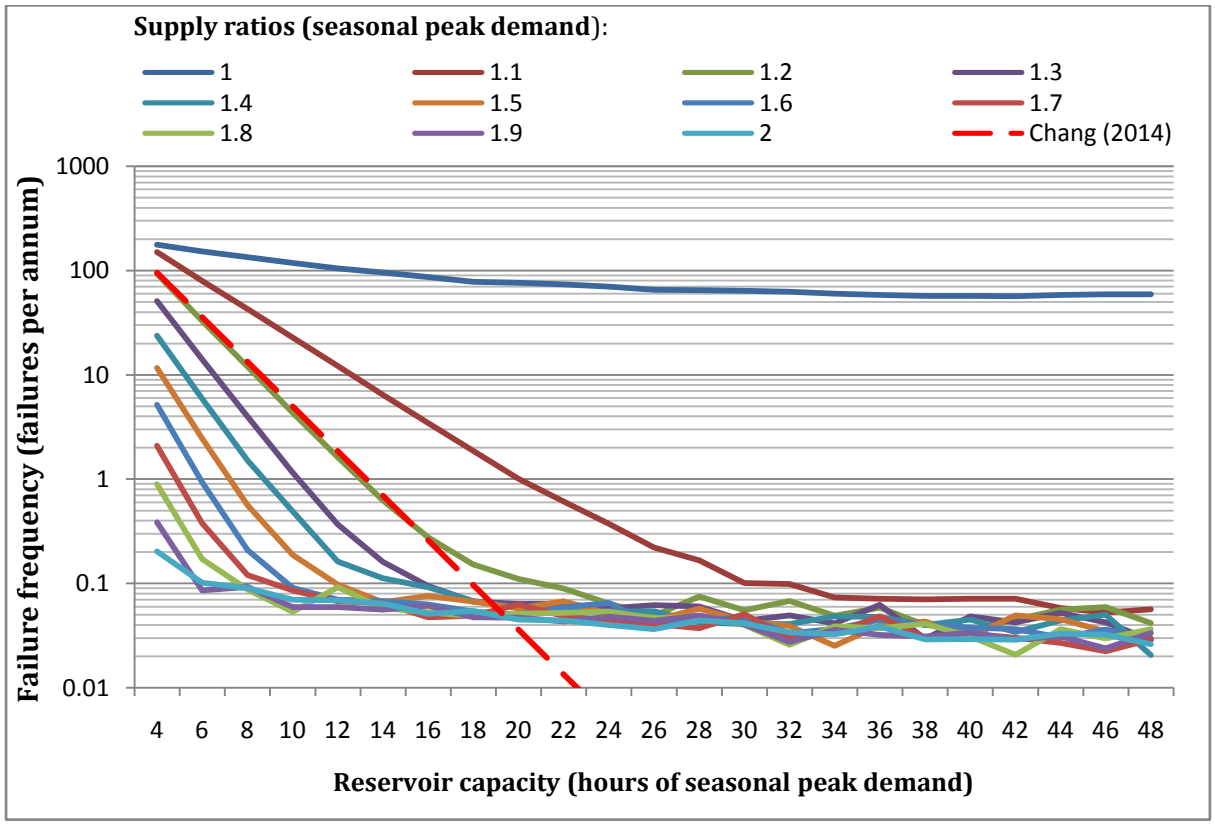

Figure 4.2.2: Failure frequency vs. reservoir capacity for varying supply ratio

There is a substantial gap observed between the supply ratios of 1.0 and 1.1, with decreasing effect toward a supply ratio of 2.0. This indicates the highly supply-sensitive nature of failure frequency in the region around 1.1 times seasonal peak demand. Achieving the desired failure criterion with a supply ratio less than 1.1 is not possible. Achieving the desired failure criterion with a supply ratio of 1.1 is possible, but with reservoir capacity greater than 30 hours of seasonal peak demand. The trade-off is determined by the cost function, as increasing the supply ratio to achieve a solution with lower failure frequency, increases all associated pump costs. This cost increase must result in a greater decrease in failure frequency in order to dominate solutions with increased reservoir capacity. In addition, when a pump failure occurs, the supply ratio immediately falls to 0. With an average pump failure duration of 48 hours, the efficiency of increasing supply ratio over reservoir capacity is likely to be low. This will be addressed further in the sensitivity analysis.

The failure frequency asymptote is also observed around 0.02 failures per annum. The same linear relationship and slope is observed between the Chang & van Zyl (2014) results, indicating the existence of a zone where the slope and by extension, efficiency of the increase in reservoir capacity is reasonably similar. This zone spans from 365 failures per annum to approximately 0.2 failures per annum and indicates that increasing supply ratio is at peak efficiency in this zone. For failure frequencies lower than 0.2 failures per annum in the base system, the utility of increasing supply ratio rapidly declines as the long duration (yet infrequent) failures (> 48 hours) result in almost certain reservoir failure, as there is no inflow, but sustained demand, and potentially fire events. This decrease in utility is not present in the Chang & van Zyl (2014) model as the duration of pipe failure is less by a factor of 10 and the utility of increasing reservoir capacity remains constant.

The trade-off that must take place when aiming to design a bulk water supply system that conforms to the design criterion dictates that, should the model allow for it, as reservoir capacity increases, supply ratio must decrease to maintain the failure frequency of 1 failure every 10 years under seasonal peak demand. With reference to Figure 4.2.3, the base model demonstrates this trade-off clearly, showing a decrease in supply ratio to an asymptote of 1.0. For a given system the Chang model manifests a step-wise relationship, with supply ratio staying constant over a range of reservoir capacities

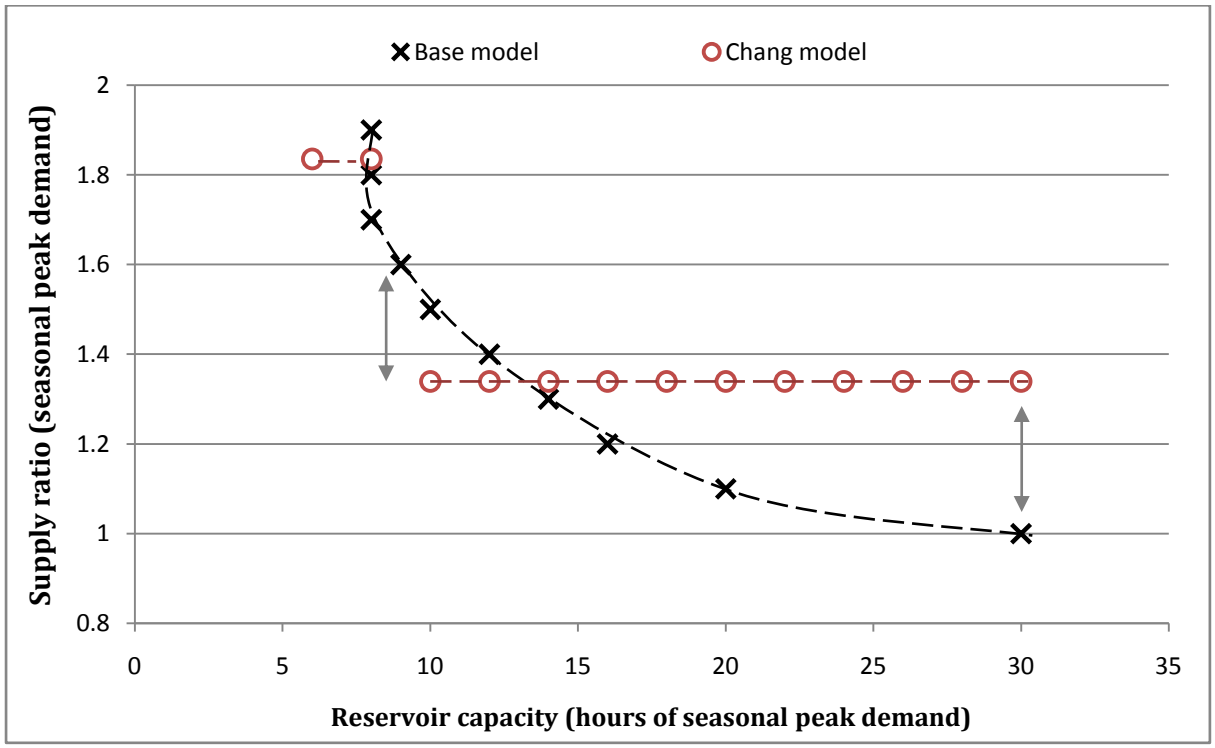

Figure 4.2.3: Trade-off at design criterion reliability

This is owing to the low resolutions of available supply ratios, which in turn is owing to the low resolution of available pipe diameters for single pipe systems. Table 4.2.2 demonstrates the supply ratio, as in the Chang model, for a given pipe diameter. As has been demonstrated previously, a supply ratio less than 1.0 (0.98 in this case) will result in a solution that experiences complete system failure (more than 100 reservoir failures per annum regardless of reservoir capacity).

A supply ratio greater than 1.83 requires a reservoir capacity of 6 hours or less, of seasonal peak demand, in order to remain as a non-dominated solution achieving approximately 1 failure every 10 years, which is impractical. The result is a single , acceptable supply ratio of 1.34 times seasonal peak demand that is utilised in conjunction with reservoir capacities ranging from 4 hours to 30 hours. This limitation results in a large inefficiency in the optimisation of solutions with reservoir capacities either more or less than approximately 14 hours of seasonal peak demand. The gaps indicated in Figure 4.2.3, represent cost inefficiency, thus creating pressure from the high probability of domination of solutions with reservoir capacity either more or less than approximately 14 hours of seasonal peak demand.

Table 4.2.2 Acceptable supply ratios for design criterion solutions (Chang model)

Internal Pipe Diameter (mm)	Supply Ratio (seasonal peak demand)
227	0.53
286	0.98
322	**1.34**
363	1.83
428	2.83
479	3.81
530	4.97

The 50th generation population is displayed in Table 4.2.3, with the system component costs detailed. The first 8 chromosome, delineated by the horizontal line, show those solutions that meet the design criterion. The sharp increase in pump power, pipe size and reservoir capacity is noted as the solutions progress into the range of solutions with failure frequency below the design criterion. In line with observations made in 4.1.4, the population consists of single pipe systems only.

Table 4.2.3: Base Population (50 Chromosomes) - Generation 50

Pump Power	Pipe Diam.	Reservoir Capacity	Supply Ratio	Failure Frequency		Total Cost	Total Pipe	Total Pump	Pump Capital	Pump Salvage	Pump O&M	Pump Energy	Total Tank
kW	*m*	*hour*	*#*	*(1:Years)*	*Failures/annum*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*-Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*
213.0	1.568	48.00	3.39	66.87	0.014954	13,809,755	8,094,660	3,884,221	2,589,629	845,638	172,924	1,967,305	1,830,874
160.3	0.574	47.68	2.22	65.51	0.015264	8,028,935	3,163,244	3,041,965	2,108,872	688,647	140,822	1,480,919	1,823,726
187.3	0.363	48.00	1.64	56.23	0.017786	7,369,758	2,060,928	3,477,957	2,360,449	770,799	157,621	1,730,686	1,830,874
160.0	0.363	48.00	1.51	54.89	0.018217	6,928,765	2,060,928	3,036,964	2,105,949	687,693	140,626	1,478,081	1,830,874
139.8	0.363	44.81	1.39	35.36	0.028277	6,525,331	2,060,928	2,705,015	1,909,902	623,674	127,535	1,291,252	1,759,389
125.2	0.363	48.00	1.30	31.40	0.031846	6,354,641	2,060,928	2,462,840	1,764,129	576,072	117,801	1,156,982	1,830,874
123.8	0.363	30.01	1.30	28.75	0.034784	5,894,644	2,060,928	2,439,493	1,749,944	571,440	116,854	1,144,136	1,394,223
103.3	0.363	29.98	1.16	12.46	0.080250	5,543,976	2,060,928	2,089,599	1,534,293	501,020	102,454	953,872	1,393,449
113.5	0.363	10.83	1.23	0.53	1.879096	5,098,073	2,060,928	2,265,416	1,643,392	536,646	109,739	1,048,931	771,729
100.8	0.363	13.00	1.14	0.26	3.777982	4,965,840	2,060,928	2,046,868	1,507,538	492,283	100,667	930,946	858,044
96.2	0.363	9.34	1.10	0.04	26.790578	4,736,151	2,060,928	1,967,298	1,457,455	475,929	97,323	888,449	707,925
95.2	0.363	8.44	1.10	0.03	39.351936	4,678,529	2,060,928	1,950,060	1,446,559	472,371	96,595	879,277	667,541
95.0	0.363	7.70	1.09	0.02	49.786174	4,640,756	2,060,928	1,947,501	1,444,940	471,842	96,487	877,916	632,327
95.3	0.363	7.28	1.10	0.02	55.512418	4,625,125	2,060,928	1,952,433	1,448,060	472,861	96,695	880,538	611,765
94.2	0.363	7.28	1.09	0.02	60.308342	4,605,732	2,060,928	1,932,984	1,435,749	468,841	95,873	870,203	611,820
92.6	0.363	7.33	1.08	0.01	69.869668	4,581,195	2,060,928	1,905,743	1,418,469	463,198	94,719	855,752	614,525
91.6	0.363	6.95	1.07	0.01	82.699032	4,544,977	2,060,928	1,888,492	1,407,505	459,618	93,987	846,618	595,557
89.4	0.363	7.01	1.05	0.01	98.845075	4,508,661	2,060,928	1,849,511	1,382,664	451,506	92,328	826,024	598,223

Pump Power	Pipe Diam.	Reservoir Capacity	Supply Ratio	Failure Frequency		Total Cost	Total Pipe	Total Pump	Pump Capital	Pump Salvage	Pump O&M	Pump Energy	Total Tank
kW	*mm*	*hour*	*#*	*(1:Years)*	*Failures/annum*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*-Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*
88.5	0.363	6.87	1.04	0.01	109.019010	4,485,885	2,060,928	1,833,843	1,372,654	448,237	91,660	817,766	591,114
86.2	0.363	7.05	1.03	0.01	122.730180	4,455,289	2,060,928	1,793,794	1,347,001	439,860	89,947	796,706	600,568
85.3	0.363	6.89	1.02	0.01	132.197080	4,431,096	2,060,928	1,777,995	1,336,854	436,547	89,269	788,419	592,173
84.5	0.363	6.35	1.01	0.01	144.780840	4,388,263	2,060,928	1,763,539	1,327,555	433,510	88,649	780,845	563,797
83.2	0.363	6.50	1.00	0.01	149.289550	4,373,185	2,060,928	1,740,334	1,312,603	428,628	87,650	768,709	571,923
85.5	0.363	4.78	1.02	0.01	164.595990	4,313,905	2,060,928	1,781,590	1,339,164	437,301	89,424	790,303	471,387
82.4	0.363	4.93	0.99	0.01	166.659510	4,268,113	2,060,928	1,725,881	1,303,273	425,581	87,027	761,162	481,304
81.9	0.363	4.86	0.99	0.01	174.734270	4,254,464	2,060,928	1,716,527	1,297,227	423,606	86,623	756,283	477,009
83.9	0.363	4.00	1.01	0.01	178.457100	4,228,790	2,060,928	1,752,815	1,320,650	431,255	88,187	775,233	415,047
81.3	0.363	4.00	0.99	0.01	185.600270	4,183,322	2,060,928	1,707,348	1,291,289	421,667	86,227	751,500	415,047
80.9	0.363	4.00	0.98	0.01	191.743860	4,174,894	2,060,928	1,698,919	1,285,831	419,885	85,862	747,111	415,047
79.6	0.363	4.27	0.97	0.00	204.328160	4,173,442	2,060,928	1,676,794	1,271,483	415,200	84,904	735,607	435,720
79.4	0.363	4.00	0.97	0.00	207.008860	4,149,219	2,060,928	1,673,245	1,269,178	414,447	84,750	733,764	415,047
78.4	0.363	4.00	0.96	0.00	218.627140	4,131,267	2,060,928	1,655,293	1,257,509	410,637	83,971	724,450	415,047
78.1	0.363	4.00	0.96	0.00	225.084110	4,125,746	2,060,928	1,649,771	1,253,915	409,463	83,731	721,588	415,047
77.9	0.363	4.00	0.96	0.00	231.129020	4,121,347	2,060,928	1,645,373	1,251,051	408,528	83,540	719,310	415,047
76.9	0.363	4.00	0.95	0.00	239.482710	4,103,579	2,060,928	1,627,605	1,239,468	404,745	82,766	710,116	415,047
76.3	0.363	4.00	0.94	0.00	246.387150	4,094,185	2,060,928	1,618,211	1,233,335	402,743	82,357	705,262	415,047
75.7	0.363	4.01	0.94	0.00	256.197640	4,082,936	2,060,928	1,605,868	1,225,268	400,109	81,818	698,890	416,141

Pump Power	Pipe Diam.	Reservoir Capacity	Supply Ratio	Failure Frequency		Total Cost	Total Pipe	Total Pump	Pump Capital	Pump Salvage	Pump O&M	Pump Energy	Total Tank
kW	*mm*	*hour*	*#*	*(1:Years)*	*Failures/annum*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*-Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*
74.5	0.363	4.00	0.93	0.00	272.153080	4,060,747	2,060,928	1,584,773	1,211,457	395,598	80,896	688,019	415,047
73.7	0.363	4.00	0.92	0.00	279.406280	4,047,285	2,060,928	1,571,310	1,202,626	392,715	80,306	681,093	415,047
84.1	0.322	4.00	0.90	0.00	299.401470	4,014,270	1,842,448	1,756,775	1,323,201	432,088	88,358	777,305	415,047
83.7	0.322	4.00	0.90	0.00	302.636810	4,007,163	1,842,448	1,749,668	1,318,622	430,593	88,052	773,588	415,047
70.8	0.363	4.08	0.89	0.00	312.213600	4,002,000	2,060,928	1,519,399	1,168,454	381,556	78,024	654,476	421,674
81.7	0.322	4.00	0.89	0.00	316.767260	3,971,243	1,842,448	1,713,748	1,295,430	423,020	86,503	754,835	415,047
69.2	0.363	4.07	0.88	0.00	330.074330	3,970,433	2,060,928	1,489,162	1,148,461	375,027	76,689	639,040	420,343
80.0	0.322	4.00	0.88	0.00	333.326350	3,941,810	1,842,448	1,684,315	1,276,364	416,794	85,230	739,515	415,047
76.4	0.322	4.00	0.85	0.00	347.469050	3,876,269	1,842,448	1,618,507	1,233,529	402,806	82,370	705,415	415,314
74.2	0.322	4.00	0.83	0.00	355.384170	3,837,583	1,842,448	1,579,852	1,208,230	394,545	80,680	685,486	415,282
73.4	0.322	4.00	0.83	0.00	358.149250	3,823,477	1,842,448	1,565,832	1,199,029	391,540	80,066	678,277	415,197
45.2	0.286	4.00	0.54	0.00	359.600280	3,106,895	1,649,120	1,042,728	844,421	275,744	56,387	417,664	415,047
40.9	0.286	4.00	0.50	0.00	362.416290	3,023,755	1,649,120	959,588	785,665	256,557	52,463	378,017	415,047

Note: The cost constituents for the pumping system have been detailed in the above Table. The cost breakdown for the pipeline and reservoir were recorded individually, but have not been displayed.

4.2.2. Comments and observations

Continuity in pareto optimal front

As mentioned under Section 4.2.1, the most noticeable characteristic of the solution front is that it forms a mostly continuous succession of solutions. The reason for this is evident from the widespread of supply ratios demonstrated in the 50th generation population. There exists the potential for 50 unique supply ratios within the [0.5, 5.0] boundary, as opposed to the 5 unique supply ratios observed in the Chang model. This is as a result on the continuous nature of the pump power decision variable. As described previously, the pump system input power does not necessarily refer to the input power of a single pump, but rather the input power of the pumping system, composed of any number of compatible pumps in an arrangement exhibiting the same flow and pressure head characteristics as a single larger pump. On this basis the pump power decision variable was kept as continuous as opposed to employing a discrete set of pump sizes. The more continuous spread is seen when the pareto-optimal solution front is viewed on a natural scale, as in Figure 4.2.4.

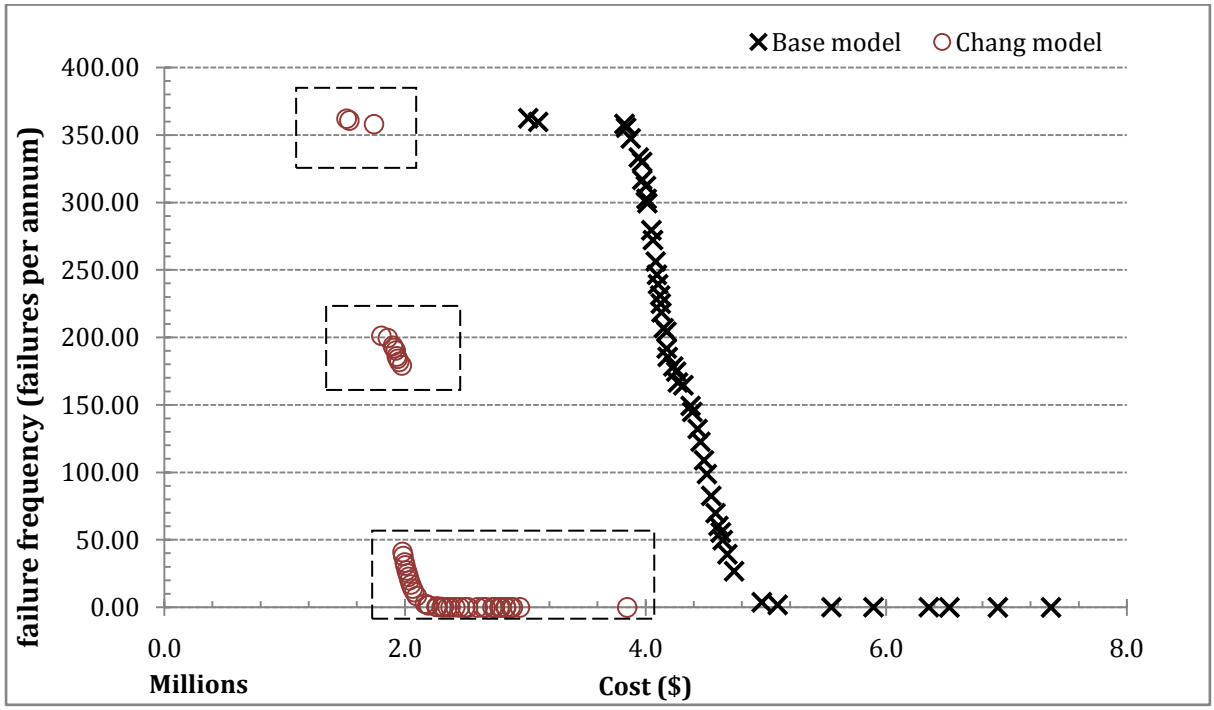

Figure 4.2.4: Pareto-optimal solution front - original and base models (Natural Scale)

The number of decision variables also plays a part, as each chromosome in the expanded model has 6 decision variables, as opposed to 3 decision variables in the original optimisation model. In this particular solution population however, all the chromosomes are single pipe systems. This reduces the number of relevant decision variables to 4, as the only significant crossover/mutation that is occurring is for the pump power, pipe diameter and reservoir capacity. In the original model, the reduction in solution diversity to only single pipe solutions reduces the effective decision variables from 3 to 2 (pipe size and reservoir capacity).

Lower number of solutions that meet the design criterion

As can be seen, the number of solutions that meet the design criterion is significantly less than the original optimisation model, this can be explained in that the NSGA-II algorithm uses the crowding distance to evenly space solutions within a given population, as the solutions now form a more continuous front, there are fewer solutions that meet the design criterion and more that are now spread into the zones that were once empty. As will be discussed in following sections, this is remedied in the application of this methodology by analysing the results and constraining either a single or multiple decision variables to achieve the desired range of solutions.

Supply Ratio - Reservoir Capacity Relationship vs. CSIR (2000) guidelines

The model responds to variation in supply ratio and reservoir capacity differently based on the system characteristics. It can be observed in Figure 4.2.4 that there is a reliability at which the supply ratio becomes critical, and a small variation in supply ratio results in large variation in reliability. It must be kept in mind that solutions on the inside of the pareto optimal front are not physically possible, as they represent a reliability that cannot be obtained at a given cost. The following figures, where possible, will exhibit the CSIR (2000) guidelines in the form of a blue line at the guideline stipulations of 48 hours annual average daily demand (32 hours seasonal peak demand) reservoir storage and 1.5 times annual average daily demand (1 times seasonal peak demand) supply ratio. The desired design criterion of 0.1 failure per annum or 1 failure every 10 years under seasonal peak conditions, as proposed by van Zyl *et al.* (2008) will be shown as a red line in the following figures

The optimal supply ratio and reservoir capacity are determined by the genetic algorithm for each chromosome as the algorithm spreads chromosomes across the range of possible failure frequencies. The supply ratios are presented in Figure 4.2.5.

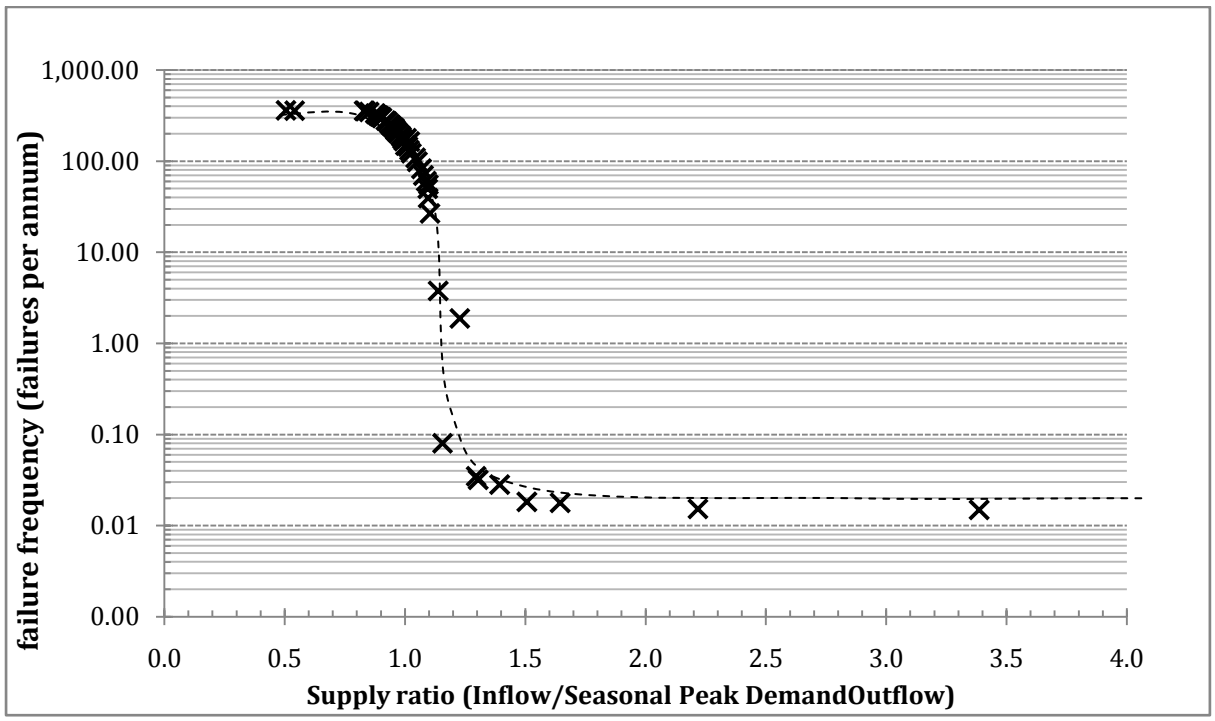

Figure 4.2.5: Reliability - supply ratio trend (Base model)

The supply ratio values between 1.0 and 1.2 can be considered critical. The gradient of influence is exponential/near vertical toward the design criterion of 0.1 failures per annum. Supply ratio values less than 1, produce a consistantly high failure rate, and values more than 1.7 produce an acceptable, constant failure rate. A supply ratio of 1.5 appears to be an acceptable point at which the effectiveness of supply ratio on failure frequency is diminished For the solutions that are crowded toward the low failure frequency end of the spectrum (around 0.02 failures per annum), the average pump failure duration manifests as a significant influence, negating the utility of increased supply owing to insufficient storage capacity of the reservoir for long duration failures. When considering the CSIR (2000) supply ratio design criterion of 1.5 times AADD (1.0 times seasonal peak demand), it can be seen that following this approach would be inadequate in this application. This value falls within the critical zone and although the reservoir capacity plays a significant role, having a supply ratio of 1.0 SPD would not produce an acceptable solution according to the design criterion, in this modeled environment (< 0.1 failure per annum under seasonal peak demand conditions).

The reservoir capacity of pareto-optimal chromosomes, corresponding to the supply ratios presented in Figure 4.2.5 is presented in Figure 4.2.6.

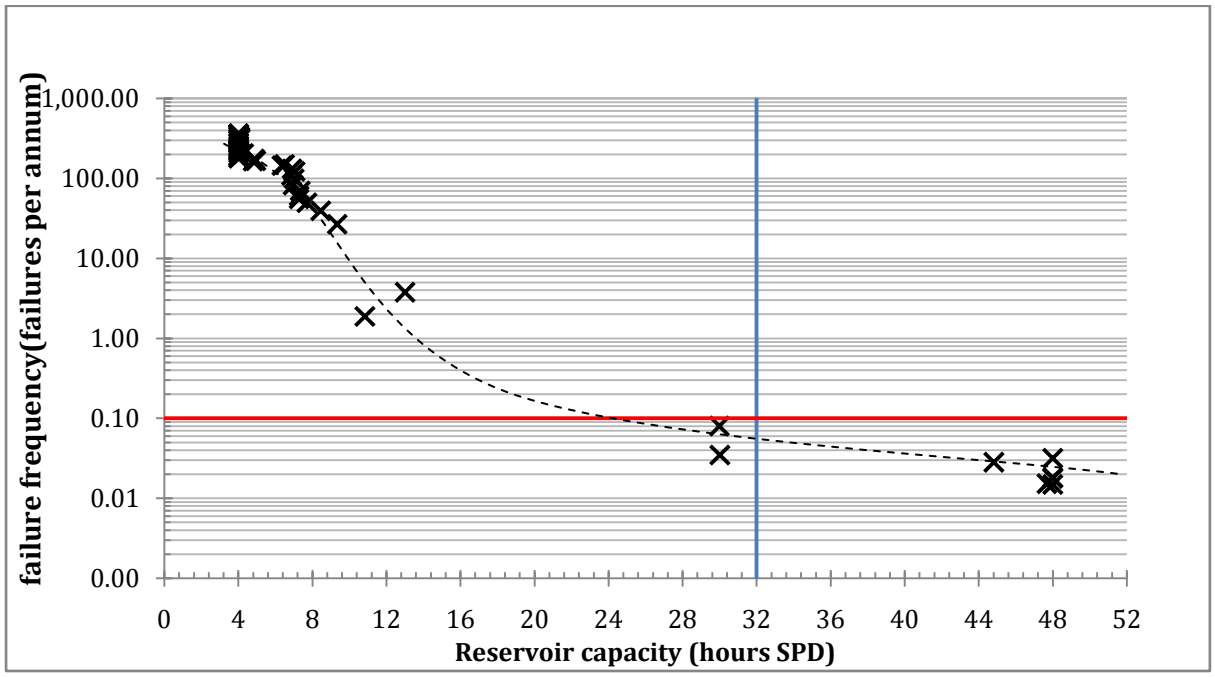

Figure 4.2.6: Reliability - reservoir capacity trend (Base model)

The primary observation from Figure 4.2.6 is that around and below the design criterion failure frequency, the cost efficiency of reservoir capacity increases, and results in the optimal solutions consisting of large reservoir capacities and increasing sharply below the design criterion. The crowding around the boundary conditions indicates that a more efficient solution could possibly be obtained in the very high reliability (< 1 in 50 year failure) zone. However as this zone is largely more reliable than deemed necessary or efficient according to the design criterion, the model boundary is deemed to be sufficient.

The relationship between reservoir capacity and supply ratio is complex and highly dependent on the system characteristics, such as the frequency and duration of supply failure as well as the frequency and duration of fire events, amongst other parameters. In a pumped system the system supply is subject to both pump and pipeline failure, which makes the trade-off between supply and capacity especially critical.

Should design be considered for this particular modeled system - according to the desired design criterion (0.1 failure per annum), and without constraining any of the decision

variables, what is shown again and demonstrated by Figure 4.2.4 is that for the base system a supply ratio of less than 1.0 will result in an unacceptably high failure frequency when designing optimally. Similarly, a reservoir capacity of less than 24 hours SPD will result in an unacceptable solution when designing optimally. The model could, at this point, be focused to exclude systems with a supply ratio of less than 1.10 and with a reservoir capacity of less than 24 hours of seasonal peak demand should the user want to restrict solutions to the feasible range.

The solutions that meet or exceed both of the CSIR (2000) stipulations (more than 32 hours of peak seasonal demand and reservoir capacity and more than 1.0 times seasonal peak demand supply ratio, in this case), do not necessarily correlate to meeting the design criterion. As such, the CSIR (2000) guidelines exclude 2 chromosomes that meet the design criterion failure frequency, but do not meet the CSIR (2000) minimum reservoir capacity. This is shown in Figure 4.2.7, with supply ratio plotted on the x - axis.

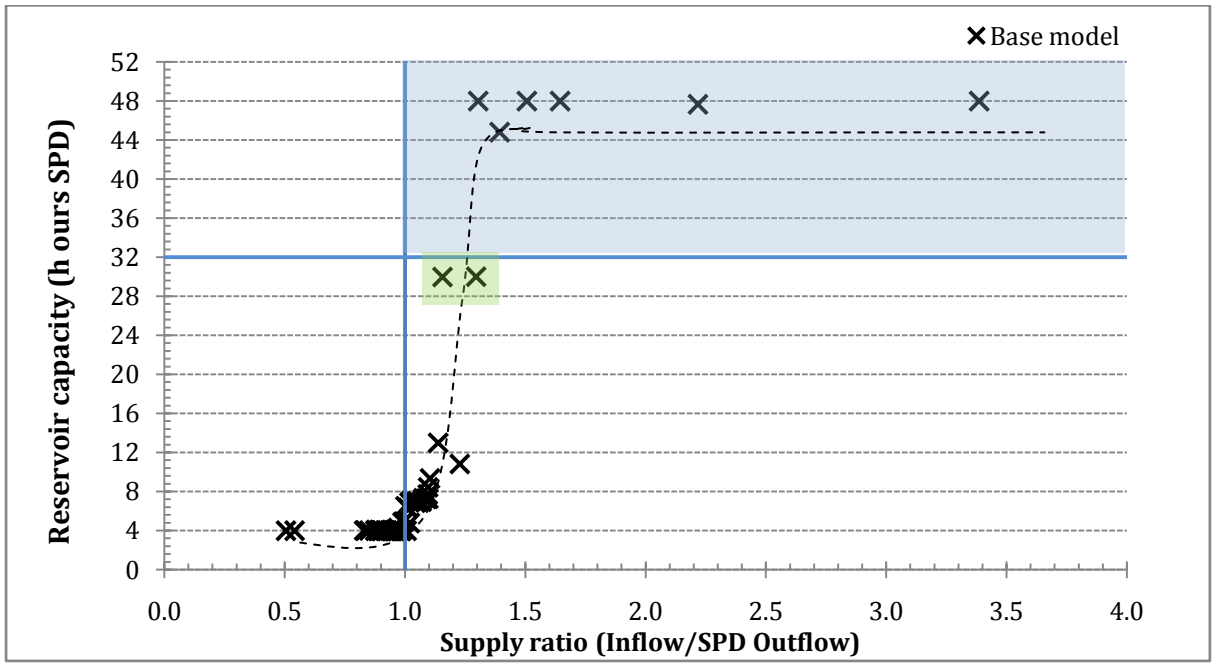

Figure 4.2.7: Reservoir capacity vs. supply ratio (CSIR (2000) stipulations in blue)

The shaded, blue area above represents the range of solutions that meet and exceed the CSIR (2000) guidelines. The shaded, green area represents the realistic area of optimisation based on the design criterion. The two solutions outlined above correlate to failure averages of 0.08 and 0.03 failures per annum, or 1 failure every 12 and 29 years respectively.

Following on from the observations stated previously, there is a zone where variation in both supply ratio and reservoir capacity have a substantial effect on the change in reliability, when variation occurs outside of this zone, the resultant change in reliability is significantly biased toward either supply or reservoir capacity. As can be seen in Figure 4.2.7 for failure frequency to decrease efficiently to an acceptable level less than 1 failure every 10 years), the reservoir capacity must increase substantially. This substantial increase in reservoir capacity between the supply ratios of 1.0 and 1.5 is indicative of the need for consideration of sensitivity in component sizing. The plateau to the right of the critical zone is a result of the solutions finding the 48 hour seasonal peak demand ceiling and the crowding function forcing an increase in supply ratio to achieve lower failure frequencies. Solutions that have reached the limit of any of the decision variable cannot be considered as efficient and should be disregarded for implementation.

Cost distribution

The cost of each of the system components varies over the reliability range based on the efficiency zones for the modulation of each component, as discussed above. By comparing the shift in cost of each component as presented in Figure 4.2.8, observations can be drawn about system and cost efficiency.

As has been mentioned previously, system component cost is analogous to system component sizing. One of the more significant observations that can be made is the constancy of the pipe system cost (and size). The pipe cost stays constant from 280 failures per annum to 0.018 failures per annum with a pipe size of 363 mm internal diameter and average velocities of 0.71 and 1.26 m/s, respectively for supply ratio between 0.92 and 1.64. The unchanging pipe diameter is an indication of the greater cost efficiency of the modulation of pump power and reservoir capacity, as demonstrated in section 4.1.4.

In the zone of failure frequency lower than the failure criterion, the effect of pump and pipe failure on system reliability increases and the pump size and reservoir capacity increases substantially to compensate. At a failure rate of less than 0.018 failures per annum (< 1:55 years) the reservoir capacity is at the maximum of 72 hours of annual average daily demand storage.

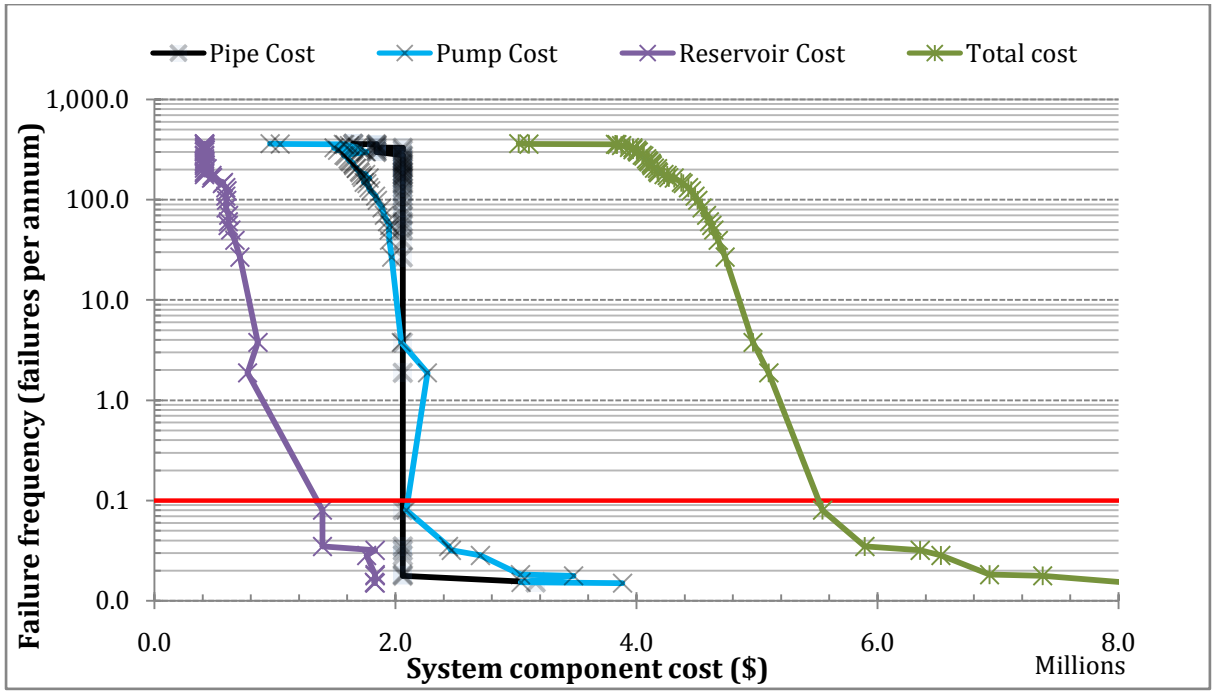

Figure 4.2.8: failure frequency versus system component cost (base model)

The split between capital and operational cost is shown in Figure 4.2.9. The capital and operational cost appear to follow a similar pattern with decreasing failure frequency, with capital cost increasing by a greater amount toward the failure criterion as the reservoir size increases. The split between capital and operational cost stays at approximately 70/30 for the range of failure frequencies considered.

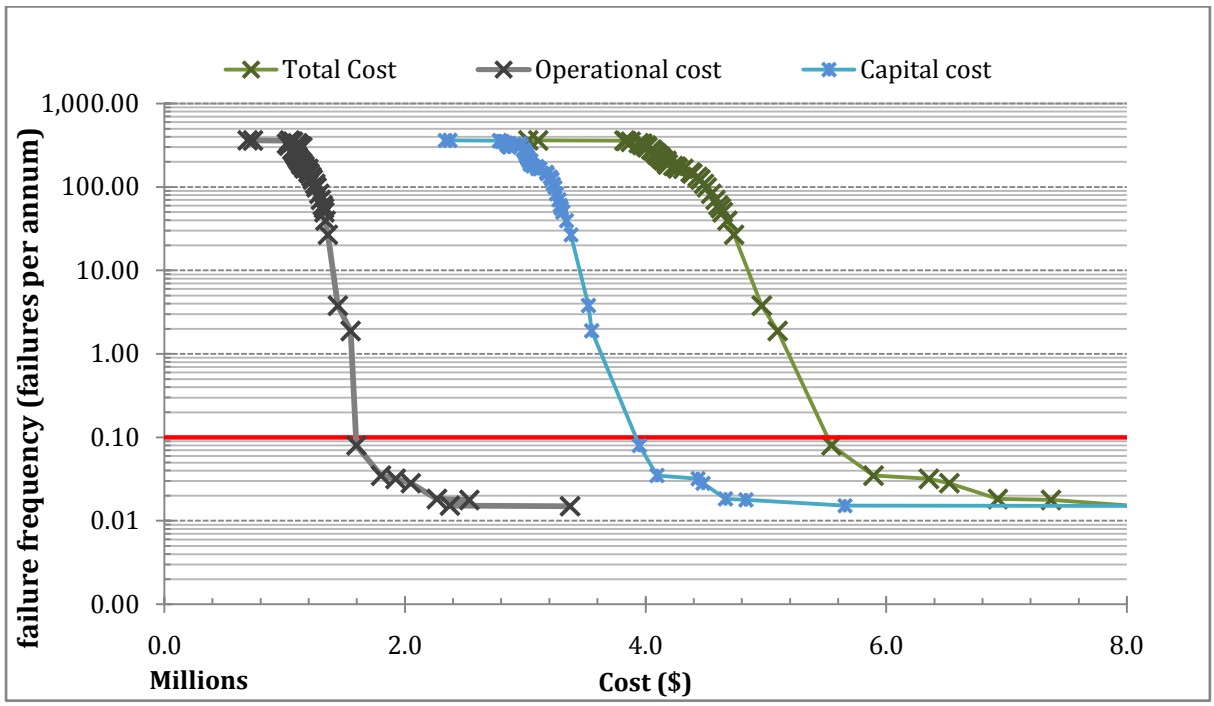

Figure 4.2.9: Capital cost - Operational cost split

A comparison to the gravity fed, Chang model presented earlier, and shown in Figure 4.2.10, gives an indication of the effect that pumping systems and their associated failures has on the cost of the pipe and reservoir capital cost. The supply failure probability is lower in the Chang model, owing to the supply system being unaffected by the failures of the electrical grid. In the Chang model this results in a shift toward increasing the pipe diameter and by association, increasing supply. The inverse is true for the base model, which as mentioned previously, requires a larger reservoir capacity to offset the increased failure duration.

Each increase in pipe diameter is accompanied by a drop in reservoir capacity, or vice versa, as the optimisation model finds the more efficient mechanism to vary the reliability, this is true of both models, as can be seen in Figure 4.2.7. The base model exhibits a larger pipe diameter than the Chang model over the majority of the failure frequency range. This can be attributed to the lower frictional losses and, by extension, lower energy cost, as well as the increase in flow capacity, from using a pipe with a larger internal diameter.

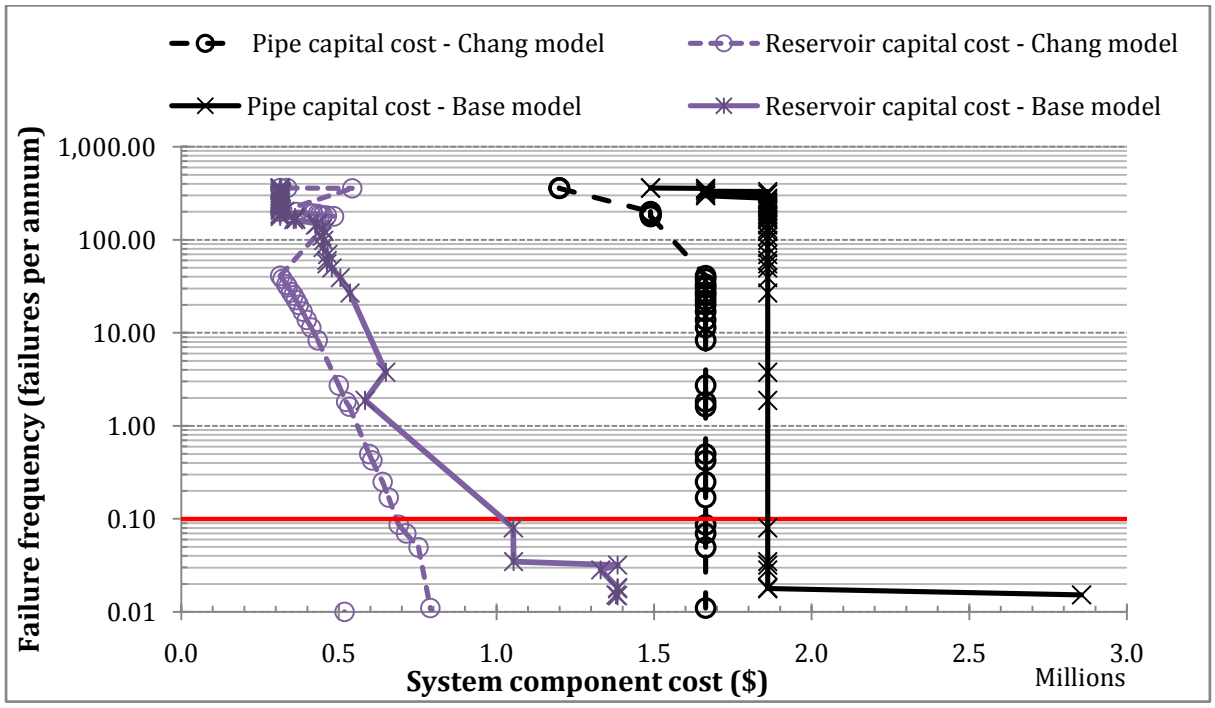

Figure 4.2.10: Reliability versus system cost (original model)

Cost convergence

The overall cost efficiency of each generation increases towards the final, optimised level with each generation, owing to the evolutionary process. The size of the range of system parameters influences how quickly the model converges to the pareto-optimal front. For this investigation the range of system parameters (as per Table 4.2.1) was large to allow the model to optimise without The result of allowing for a large range of parameters is the initial generation and a small number of generations following, are largely inefficient, as can be seen in Figure 4.2.11 and Table 4.2.4.

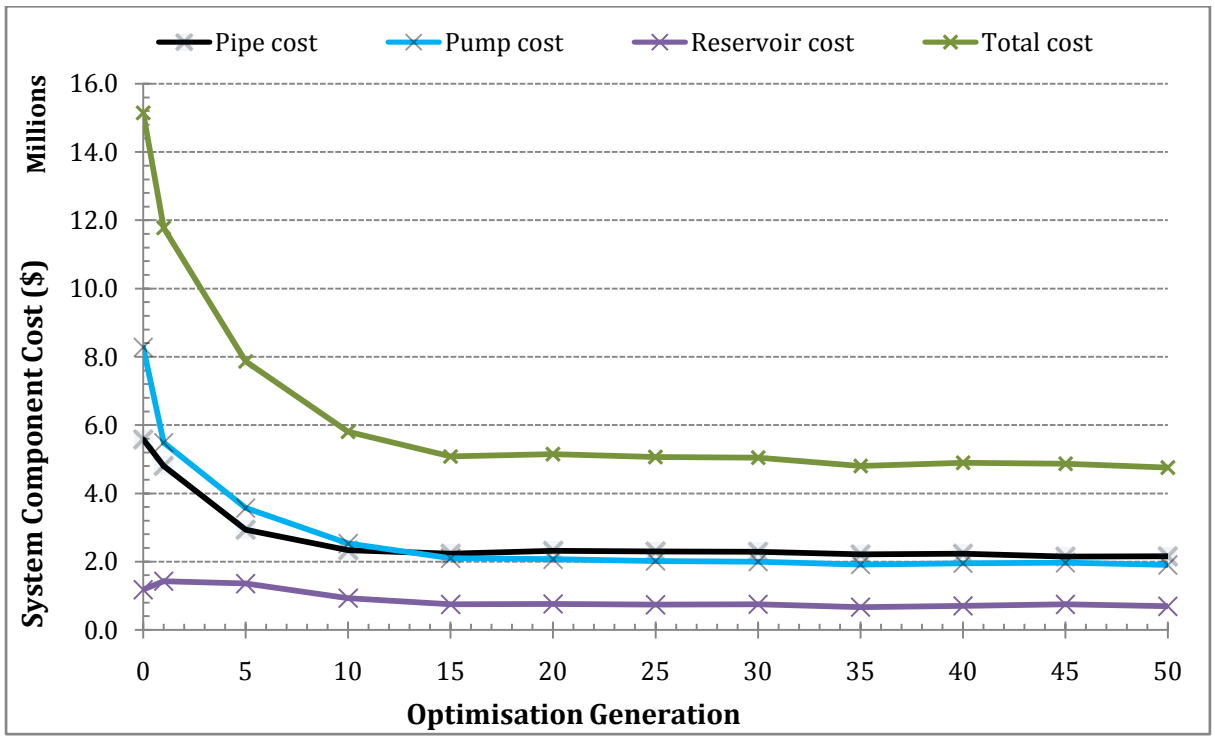

Figure 4.2.11: System cost components vs. optimisation generation (expanded model)

The pump cost (capital and operational) declines rapidly within the first 15 generations. This is owing to the wide range of pump powers allowed for during the initialisation stage, which results in a portion of the chromosomes consisting of large pump powers. The same is true for the pipe range to a lesser extent. The reservoir capacity does not exhibit such a large variance on the scale shown above as the reservoir cost is limited to between $0.4m and $1.8m.

Generation 0, in Table 4.2.4, represents the randomised, initial generation. The solutions are scattered randomly around the solution space and have not been through any of the selection processes. After a single optimisation cycle, the average system cost has dropped to 148%, after the 5th generation, that number is at 66%, a significant improvement. Within 15 generations, the average system cost falls to within 10% of the 50th generation cost. At this point the solutions produced can be considered pareto-optimal, but not uniformly crowded. After 35 generations the cost is within 5% of the 50th generation cost. After 25 generations, the solution efficiency tends to oscillate around the ultimate value as the crowding function manipulates already efficient solutions in order to space solutions evenly within the solution population.

Table 4.2.4: Average system cost per generation

Gen.	Total Cost	Total Pipe	Total Pump	Total Res.	% Ultimate
0	$15,151,005	$5,581,743	$8,281,322	$1,181,683	218.8
1	$11,782,850	$4,805,170	$5,491,383	$1,432,350	147.9
5	$7,873,963	$2,940,985	$3,573,082	$1,359,897	65.7
10	$5,803,364	$2,341,746	$2,532,072	$929,547	22.1
15	$5,088,751	$2,237,918	$2,100,870	$749,963	7.1
20	$5,151,067	$2,315,637	$2,073,240	$762,190	8.4
25	$5,062,431	$2,300,084	$2,021,101	$741,245	6.5
30	$5,049,496	$2,291,359	$2,004,134	$754,002	6.2
35	$4,801,337	$2,211,293	$1,918,946	$671,097	1.0
40	$4,897,271	$2,235,694	$1,954,263	$707,314	3.0
45	$4,870,996	$2,152,723	$1,968,929	$749,344	2.5
50	$4,753,254	$2,156,589	$1,902,495	$694,170	0.0

At this point the designer would constrain the number of generations to reduce the computational load. The number of generations, however, will be kept at 50 for the sensitivity analysis to allow for fluctuation. The progression of optimisation is shown in Figure 4.2.12.

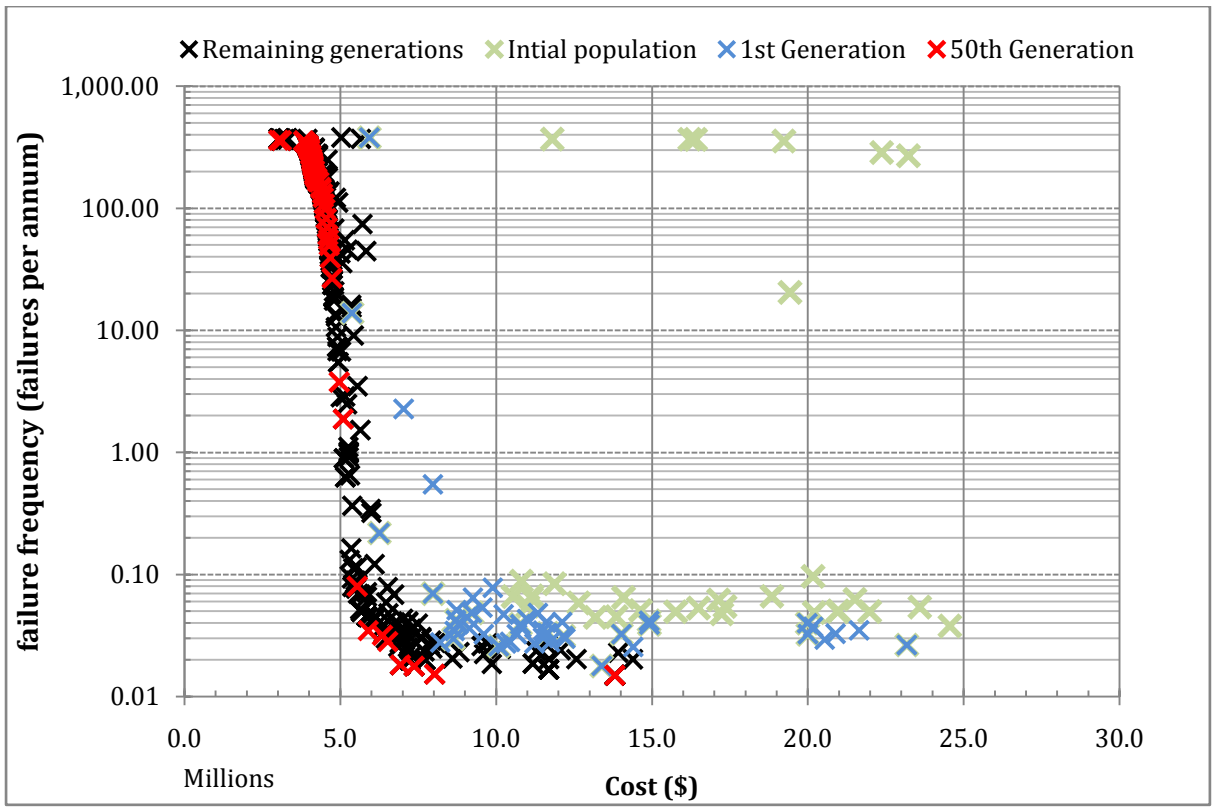

Figure 4.2.12: Failure frequency versus system cost (multiple generations)

The initial population (Generation 0) is observed to be scattered over the solution space. The first generation is substantially closer to the pareto-optimal front, but with a densely crowded population . The rest of the generations are distributed around the pareto-optimal front, as indicated by the 50th generation. As discussed above, the system converges to its optimal arrangement around the 35th generation.

The optimisation process has an effect on the cost balance between different components as the population goes through multiple generations. The optimal balance is achieved through crossover, recombination and selection to determine the fittest solutions. This distribution is presented in Figure 4.2.13. The initialised generation, as represented in Figure 4.2.11 and Figure 4.2.12 has a high percentage pump cost owing to large range of allowable pump sizes, mentioned above. As the population is passed through each subsequent generation, the pump and pipe (and by extension, supply) costs reduce and the reservoir cost becomes marginally more prominent. It is important to note the relatively minor average reservoir cost of 15%. The balance observed for optimal solution is 45/40/15% between pipe system, pump and reservoir is achieved after 25 Generations.

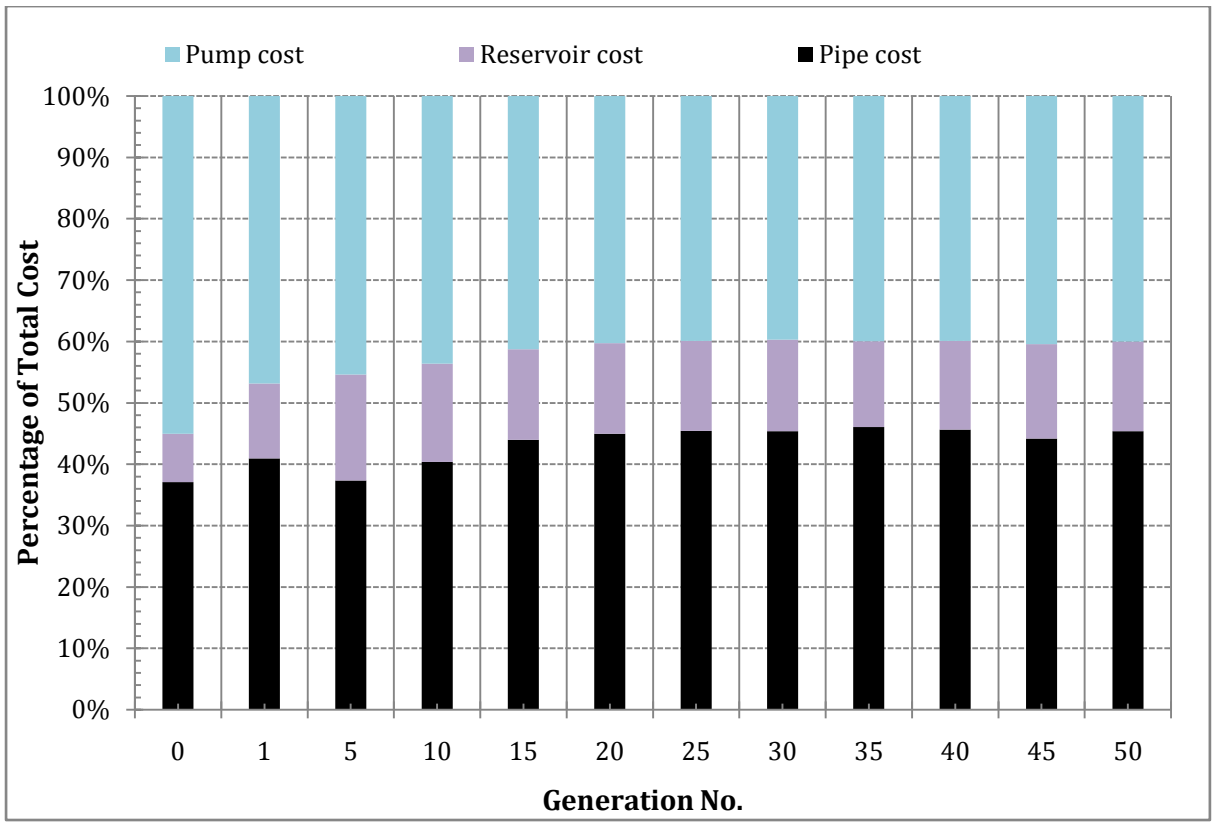

Figure 4.2.13: System Component Cost per Generation

The cost balance can also be traced within a single chromosome of 50 solutions. Figure 4.2.14 characterises the cost efficiency of each of the 50 solutions from the highest failure frequency to the lowest failure frequency chromosome, in the 50th generation population. The pump and pipe costs make up the majority of the cost of the high failure frequency chromosomes. The chromosome cost balance begins to shift more toward reservoir capacity and away from pipe size when considering chromosomes with failure frequency decreasing toward the design criterion

When considering the solutions with decreasing failure frequency, approaching the failure criterion (indicated as a vertical, red line) from the right-hand side, the reservoir portion of the total cost starts to increase more quickly, until it reaches a maximum value after the design criterion. The optimisation model forces an increase in pipe size in order to achieve the failure frequencies on the lower end of the failure frequency scale. The percentage shifts toward pipe cost as the unit cost to increase pipe size is substantial. The sharp spike seen in the lowest failure frequency range (furthest left) is as a result of the reservoir capacity reaching the 48 hour seasonal peak demand storage level.

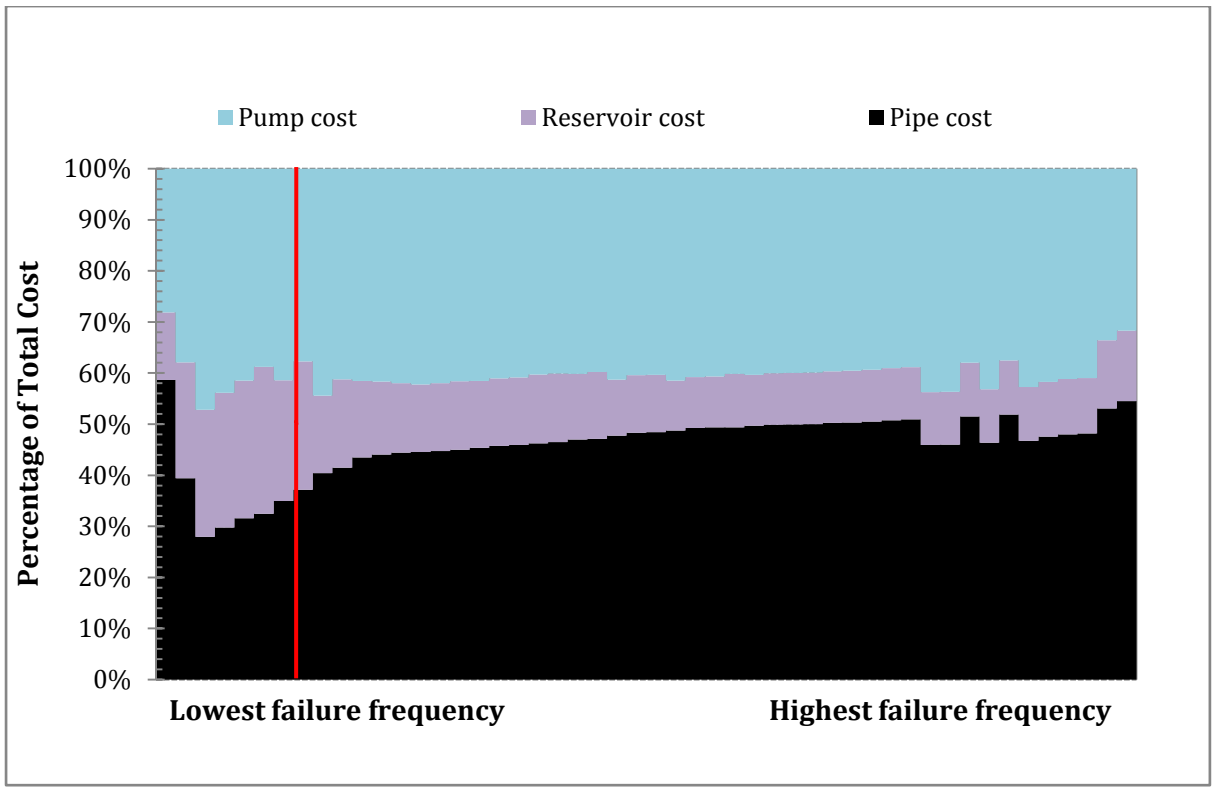

Figure 4.2.14: System component cost (50th generation, red line = design criterion)

4.2.3. Summary

Expanded optimisation model

The base model was able to provide solutions across the spectrum of reliability for systems that are inclusive of pumping systems, where power is drawn from the national, electrical distribution grid. The base model produced solutions within the acceptable design criterion of 1 failure in 10 years under seasonal peak demand. It gave insight into the influence of both supply ratio and reservoir capacity on the baseline system reliability.

It was shown that reservoir cost has a lesser influence on system reliability, around the design criterion, whereas pump power, and by extension supply into the reservoir has a substantially greater influence in the same region. This demonstrates the relative criticality and importance of consideration toward the sizing of the pump and desired supply ratio. This is in line with the observation from Figure 4.2.4, that optimisation through supply ratio variation; around the design criterion failure frequency is an efficient means of optimisation (large decrease in failure frequency for a small increase in supply). As only pump power is

varied in this range, while the pipe size remains unchanged by the evolutionary process, the modulation of pump size was, by extension, critical to achieving the desired failure frequency.

The genetic optimisation process was able to converge to a pareto-optimal front within 10% of the ultimate front within 15 Generations. The following generations can be considered as refining generations and allow the influence of the crowding distance to ensure that the solution front is relatively uniform. As each solution produced is valid (not always practicable), the overlay of pareto-optimal fronts from each generation provides reliability and cost results for 2500 solutions. From this wide range, the designer can narrow down the allowable system component sizes to achieve a desired range of costs and/or failure frequency.

4.3. Sensitivity analysis

In this section, the cost sensitivity of the converged generations (above 15th generation), solutions fronts, will be tested against variation in system parameters, such as pipe length, static pumping head, pump failure frequency, amongst others, and application to real-world environmental conditions (load shedding). The 50th generation will be presented graphically for comparison for each sensitivity parameter tested. However, all generations after the 45th generation, which have been shown to exhibit reasonable convergence, will be used to obtain the closest solution to the desired design criterion for summary analysis. The primary objective of the sensitivity analysis is to identify those system parameters that are cost critical, and draw observations to that effect. The sensitivity analyses were performed with parameter variations in alignment, where possible, with those performed by Chang & van Zyl (2012). The parameters to be tested are outlined in Table 4.3.1

Table 4.3.1: Sensitivity analysis parameters

Parameter	Low	Baseline	High	Unit
Pump failure frequency	0.036	**0.104**	0.297	Failures/annum (average)
Pump failure duration	12	**48**	190	Hours (average)
Power failure mechanism	Base model (NERC 2014)-USA vs. Nel (2009)-SA			
Load shedding severity	1	**2**	4	CoCT Stage
Pipe failure rate	1	**2**	5	Failures/annum
Pipe failure duration	3	**4.5**	9	Hours (average)
Fire event frequency	0	**6**	24	Failures/annum (average)
Fire event duration	0.44	**0.84**	1.62	Hours (average)
Pipe length	1	**10**	100	Km
Static head	30	**60**	120	m
Reliability focusing	N/A			

Each of the parameters listed in Table 4.3.1 were tested for 50 generations and the 50th generation pareto-optimal curve is represented graphically. A comparison of the systems that closely meet the design criterion is listed, and conclusions are drawn.

4.3.1. Sensitivity to pump failure frequency

The sensitivity of the cost of a bulk water supply system to supply failure frequency is an important consideration when designing a bulk supply system. Analysis of the failure frequency sensitivity allows for the designer to consider the cost efficiency of providing backup power (if necessary) or geographical placement of the pump station in relation to power reliability zones, as described in 2.4.1. The failure frequency values for the US (NERC, 2014) region reliability zones are as per Table 4.3.2.

Table 4.3.2: Base model failure frequency sensitivity values

Parameter	Low Value	Baseline	High Value
Mean frequency (failures/annum)	0.036	0.104	0.297
1 Failure :X years	27.5	9.6	3.37
Standard deviation (σ)	1.05	1.05	1.05
Lognormal mean (μ)	-3.86	-2.81	-1.76

The mean frequency is obtained as the natural mean of the lognormal cumulative distribution function as described by the standard deviation and lognormal mean. The standard deviation of the lognormal distribution is kept constant and the lognormal mean varied by one standard deviation either side of the mean. The time to failure or time between failures is determined from a Poisson process informed by this natural mean. The effect of this variation on the system cost of solutions produced is demonstrated in Figure 4.3.1

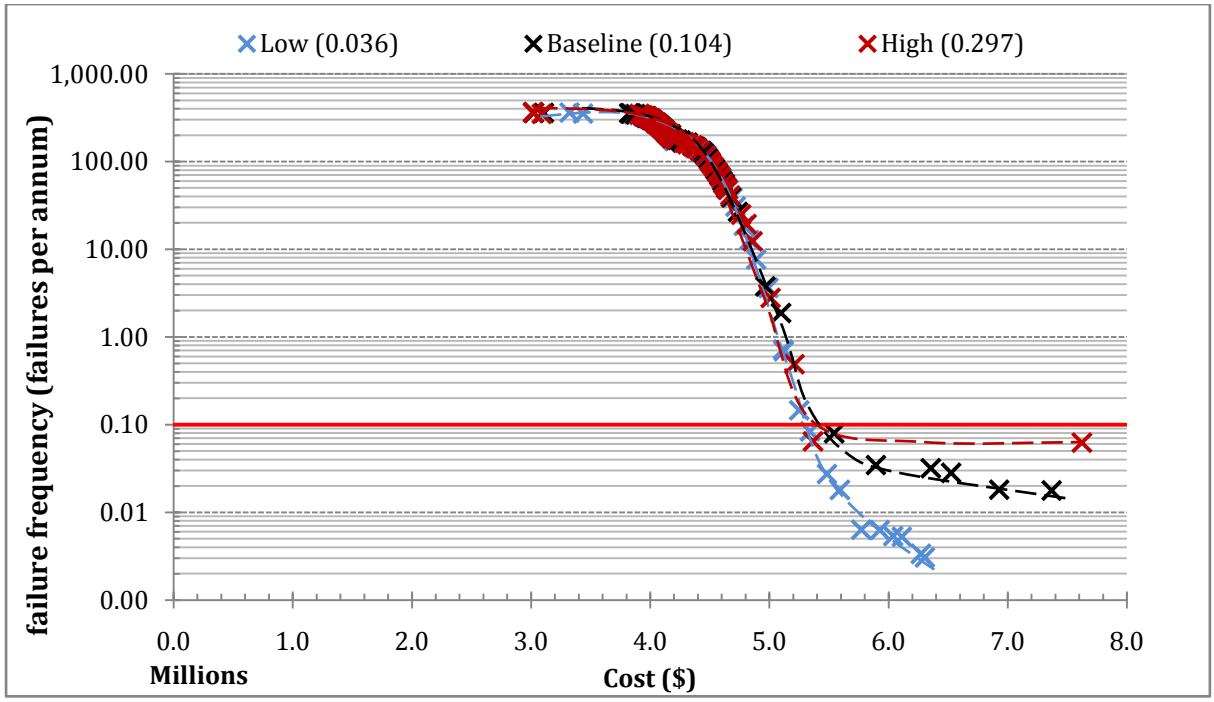

Figure 4.3.1: Pareto-optimal solution fronts for variable pump failure frequencies

The point at which the effect of the different pump failure frequency values, on system cost, manifests happens at the design criterion failure frequency. What can be taken from this is that there is sufficient room for optimisation within the system parameters to ensure that a system can be designed at and above the desired design criterion failure level of 1 failure every 10 years under seasonal peak demand, under varying power failure frequencies. This is significant as it shows that for a design criterion failure system, the consideration of power and pump failure frequency under US, developed world conditions shows significance but is not highly critical.

However, at reliability levels greater than this, each of the frequency levels exhibit substantially different behaviour. This is considered reasonable, as the mean pump failure duration of 48 hours ensures that any failure with duration equal to or greater than the mean will almost certainly result in a bulk system failure regardless of reservoir capacity. A lower frequency of power failures is therefore directly related to a lower reservoir failure frequency. The singular solution in Figure 4.3.1 where the high failure frequency model produces a more efficient solution, is the result of a crossover variation in the reservoir capacity, that increased cost, without increasing reliability by a significant amount (owing to pump failure duration). This is the result of a low resolution in the 0.2 to 0.01 failures per annum region. A focused

model, producing a higher resolution of acceptable reliability solutions would remedy this during application as the crowding distance would have less influence. For solutions with failure frequency more frequent than 1 failure every 10 years (0.1 failures per annum), the power failure frequency has little to no effect.

4.3.2. Sensitivity to pump failure duration

As with sensitivity to pump failure frequency, the pump failure duration has a direct influence on reservoir capacity buffer against failure events. The duration of power failure events is substantial. This is owing to the nature of the power failure events experienced, being distribution failures resulting from extreme weather, vandalism or malfunction. The repair and restoration times for these events are substantial. The power failure durations tested were as described in Table 4.3.3. The lognormal distribution parameters obtained from the USADOE (2014) and NERC (2014) were used with a variation of 1 standard deviation adjustment to the lognormal mean (mean of natural logarithm).

As mentioned in 2.4.1, the electrical reliability indices for countries with varying stages of development of electrical infrastructure were considered. The United States of America was chosen owing to the wide range of electrical interruptions experienced and the potential for implementation of the optimal reliability-based design approach, as the USA has sufficiently developed infrastructure to support pumping installations, but where the electrical supply interruptions are substantial enough to warrant their consideration in the evaluation of bulk water supply system reliability. The Nel (2009) failure duration lognormal distribution, exhibiting a substantially lower duration for pump failure duration is presented in 4.3.3. The results are demonstrated in Figure 4.3.2.

Table 4.3.3: USADOE/NERC(2014) Power failure duration sensitivity values

Parameter	Low Value	Baseline	High Value
Mean duration (hours)	**12**	**48**	**190**
Standard deviation (σ)	1.38	1.38	1.38
Lognormal mean (μ)	-1.54	2.92	4.30

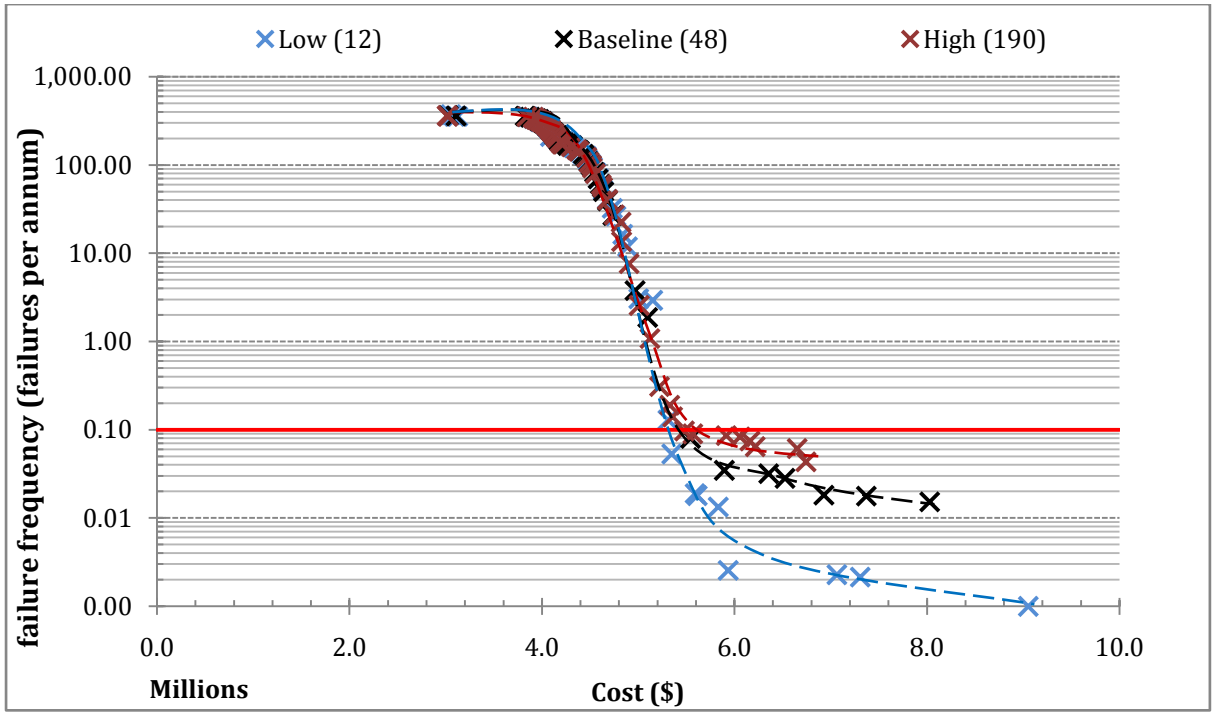

Figure 4.3.2: Pareto-optimal solution fronts for varying power failure duration values

Similar observations, as for power failure frequency can be drawn from Figure 4.3.2. At failure frequency level higher than 1 failure every 10 years, the power failure duration has little to no effect as at the level of failure frequency higher than 1 failure every 10 years, the system is failing owing to insufficient supply to the reservoir and/or insufficient reservoir capacity. All systems that do not meet the design criterion have either a supply ratio less than 1.1 or reservoir capacity less than 20 hours of seasonal peak demand or both. The result is that the duration of pump failure is inconsequential as the system becomes highly vulnerable to failure under any event of any duration (fire, pipe or pump failure) or even the result of a peak in normal population demand. Therefore a change in frequency or duration of any of the 3 types of events, mentioned above, on its own has little effect. This is observed throughout the sensitivity analysis.

The divergence point is also seen to be at the design criterion. This is likely attributed to the physical environment in which the system is constructed i.e. the characteristics of the demanding population. Again the conclusion is that when designing to the criterion of 1 failure every 10 years, the failure duration is not seen to be critical within the range of durations investigated for this particular system. Designing systems at failure frequency levels equal to and below this value, however, would require careful analysis of the bulk

systems environmental conditions i.e. demanding population and power supply/distribution reliability, as this would likely influence the divergence point seen in figures 4.3.1 and 4.3.2.

The low failure duration model is capable of producing optimal solutions that are more reliable than that produced by the low frequency model. At a mean power failure duration of 12 hours, there is a lesser chance of a system failure occurring, as a system with reservoir capacity greater than 12 hours, experiencing seasonal peak demand, will have sufficient storage to buffer against supply failure. Comparing low duration solutions against the base would see solutions in the same cost region at failure frequency levels 17 times less. This reinforces the assertion that low failure frequency systems (less than 1 failure every 10 years under seasonal peak demand) are highly sensitive to supply failure duration (and frequency, as seen in 4.3.1). The exact position that this sensitivity occurs is dependent on the characteristics of the population being supplied.

4.3.3. Sensitivity to power failure mechanism

The base power failure model, as mentioned in previous sections, is based on statistical distributions gained from power supply interruption data collected for the NERC regions over a number of years. This allows for stochastic modeling of North American power supply reliability. The extension is that this is used to represent power supply disruption in the developed world. Comparing developed world power supply to developing world power supply manifests a different mechanism altogether.

Developed world power supply is assumed to have sufficient generation capacity as a result of adequate funding and careful master planning. The failure events were reported as being the result major disruptions, such as fires, extreme weather and vandalism by USADOE (2014). Consequently the failure mechanism results in low frequency, long duration power failure events owing to the severity of the cause. When considering South Africa as a developing country, the failure events are commonly a result of insufficient generation capacity, poor maintenance, minor vandalism, cable theft and overloaded reticulation networks. These power-generation type failures result in higher frequency, shorter duration power failure events (external power trips) compared to the NERC (2014) data. The comparison of these failure mechanisms provides insight into the cost implication of each.

The lognormal distributions parameters presented in Table 4.3.4, in the case of failure frequency, are used to generate the mean frequency which is used as input to the Poisson process to calculate the time between failure from the average occurrence (as mentioned in 4.3.1). In the case of failure duration, the failure duration is calculated directly from the lognormal cumulative distribution function.

Table 4.3.4: Pump (power) failure lognormal distribution parameters

Parameter	Baseline - USA (NERC, 2014)	SA (Nel, 2009)
Mean frequency (failures/annum)	**0.104**	**11.4**
Standard deviation (σ)	1.05	0.7
Lognormal mean (μ)	2.92	2.2
Mean duration (hours)	**48**	**1.6**
Standard deviation (σ)	1.38	1.54
Lognormal mean (μ)	2.92	-0.61

The result of this variation in power failure mechanism is demonstrated in Figure 4.3.3, Taking into consideration the vast differences in power failure frequency and duration, the solution sets are remarkably similar. The Nel (2009) mechanism produces marginally more expensive options, at the near design criterion reliability level, however the differences are not as substantial as one would assume.

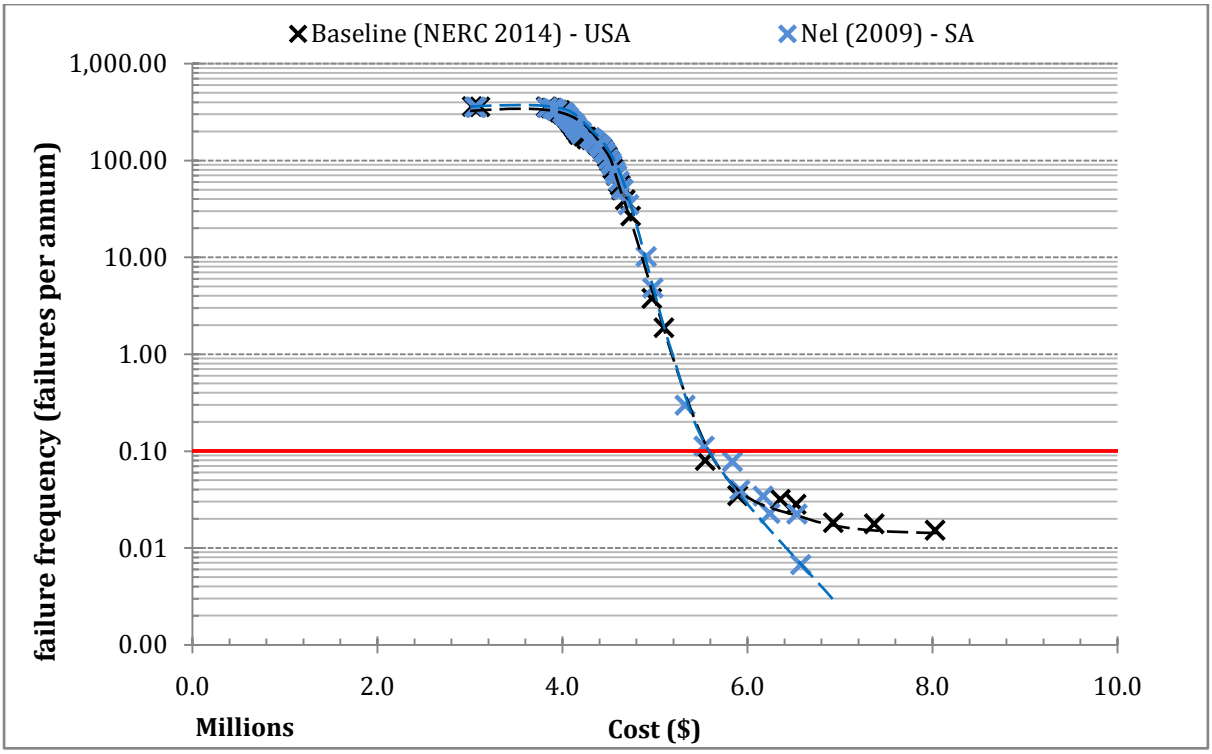

Figure 4.3.3: Pareto-optimal solution front for varying power failure mechanism

4.3.4. Sensitivity to Load shedding severity

Load shedding, also referred to as planned power outages or rolling blackouts are used as a means of reducing the demand on the electrical grid by switching power off to a specific area or sets of areas in succession. This, aside from placing an economic and social burden on the affected areas, has an effect on the delivery of civil services to the affected areas. Municipal/service reservoirs supplied with water by pumped systems, and with no backup power generation capacity, will be affected by power outages equally.

The severity of these outages can be measured either in outage time and/or outage frequency. The City of Cape Town produced a load shedding schedule in 2015, which outlined when load shedding would occur based on the day of the month and the affected zone(s). The zones are delineated by the City of Cape Town and the specific time-of-day allocation per zone was randomly selected. The scaling of the outages was done by increasing the frequency as opposed to increasing the duration of the outage. This is seen to be reasonable as there are many time dependent appliances that are affected by a longer duration outage, such as commercial and domestic refrigerators, battery backup power supplies, etc. The load shedding frequency is classed according to the frequency of power outages, as demonstrated in Table 4.3.5.

Table 4.3.5: City of Cape Town (2015) load shedding schedule severity

	Load shedding Severity Level			
Parameter	**Stage 1**	**Stage 2**	**Stage 3A**	**Stage 3B**
Sensitivity Parameter	**Low**	**Baseline**	Not Used	**High**
Outage Frequency (TTF) (hours)	29.5	13.5	8.5	5.5
Outage Duration (hours)	2.5	2.5	2.5	2.5

The time to failure (TTF) was calculated by selecting a zone randomly with population equal to the generic model population (3000 - 5000 stands) and working out the time between power outages. The way that the load shedding schedule has been created ensures that each zone is equally affected by power outages. The acting assumption is that the load shedding stage does not change during the period when load shedding is in effect and that all load shedding events are in addition to the power failures experienced (as described by Nel, 2009). The effect that the implementation of load shedding has on the cost of designing a bulk supply system is demonstrated in Figure 4.3.4:

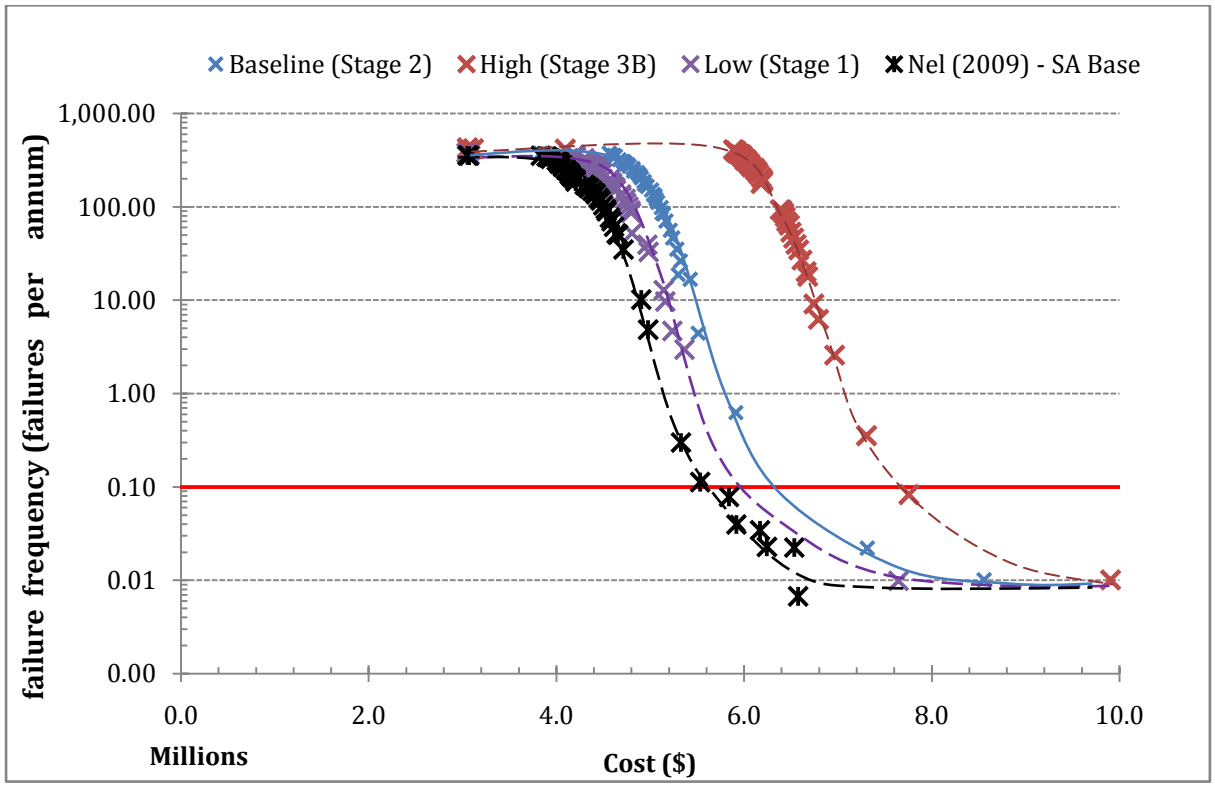

Figure 4.3.4: Pareto-optimal solution fronts for varying load shedding severity

As can be seen from Figure 4.3.4, the load shedding severity has a significant impact on the cost of bulk water supply system at an equal failure frequency level. The cost offset between different load shedding stages is seen to be constant throughout the failure frequency range. Between low and base (Stage 1 and Stage 2), the offset is approximately $400k, while the offset between base and high (Stage 2 and Stage 3B), is approximately $1.2m - $1.5m.

Comparing the results to the existing South African reliability model as defined by Nel (2009), gives an indication of the additional cost to the water authority should they choose to develop a bulk water supply system exhibiting design criterion failure frequency. At base-level load shedding (Stage 2), also under consideration of design criterion failure frequency, the difference is slightly more pronounced At a failure frequency lower than 1 failure every 10 years, the different results from different stages of load shedding begin to converge to the minimum failure frequency of approximately 0.01 failures per annum. It is proposed to be likely that at this failure frequency level the random power failures with substantially lower frequency, but higher duration (Nel, 2009) are more likely to result in failure compared to the lower duration, load shedding outages. This results in convergence to the minimum failure frequency as exhibited by the Nel (2009), South African model. This is, however, only observed in the low failure frequency zone (0.08 – 0.01 failure per annum).

The most significant observations are obtained when comparing the baseline, Nel (2009) model to the high-level, stage 3B load shedding results. The extra cost of a desired design criterion solution is approximately $2.2m or an equivalent 38% increase in cost. These results are summarised in Table 4.3.6.

Table 4.3.6: Load shedding cost and parameter sizing implication by severity level

System Average	**Nel (2009)**	**Stage 1**	**Stage 2**	**Stage 3B**
Sensitivity Parameter	**Baseline**	**Low**	**Base**	**High**
Pump Power (kW)	93	98	105	142
Modal Pipe Size (m)	0.36	0.36	0.43	0.48
Supply Ratio (/SPD)	1.08	1.17	1.27	1.78
Reservoir Capacity (h SPD)	13.23	15.5	15.5	11.6
Reservoir Failure Duration (h)	10.7	3.76	3.84	1.45
Design Criterion Cost ($m)	5.5	5.9	6.4	7.7

From the above Table, the influence of load shedding can be quantifiably compared. It can be seen that the efficient means of producing an optimally efficient set of solutions under high-level load shedding is to drastically increase supply. This, following on from prior observations would be in response to the decreased refill time between load shedding events and power failures. Another relevant observation is the reservoir failure duration, the time between when the reservoir empties, to when it receives a net inflow. It can be seen, in this instance, that the increased supply results in shorter reservoir failure duration.

Another, more concerning, observation is the consequence should a system be designed and constructed optimally as per the proposed method, to perform at design criterion failure frequency level and be subject to the implementation of load-shedding, not considered during design. In this case, the failure frequency under low-level load shedding increases by a factor of 10, to 1 failure per year. Under base-level load shedding, failure frequency increases by a factor of 100, from 1 failure every 10 years to 10 failures per annum. Under high-level load shedding, complete system failure occurs. It is proposed that this is not seen more prominently in South Africa, under 2015 load shedding conditions, as the implementation of load-shedding is not continuous and the existing bulk water systems are not designed optimally, implying the possibility of excess reservoir capacity and/or supply.. There are also times when load shedding is suspended completely. In addition, the occurrence of high-level (Stage 3B) load shedding is infrequent. However, should Stage 2 or Stage 3A/B load shedding be implemented in a continuous manner, without implementation of additional backup generation, it is likely that potable water service levels will decrease. The exact effect of this decrease will be dependent on the demanding population and system.

4.3.5. Sensitivity to pipe failure rate

Pipe failure, in a pumped supply system, results in lost or reduced supply capacity and an associated reduction in the effective supply ratio. In the case of power failure and resultant pump failure, the loss of supply is absolute. However, when considering a multiple pipe system, there is the possibility for the retention of partial supply. The verification results presented in section 4.1 and base results presented in section 4.2, for the expanded optimisation model manifest an inherent inefficiency in systems comprising more than a single pipe. The result of this inefficiency is the equivocation of the effect of pipe failure to

the effect of pump failure, a complete loss of supply. The sensitivity values presented in Table 4.3.7 were kept in line with those tested by Chang & van Zyl (2012).

Table 4.3.7: Pipe failure rate sensitivity values

Parameter	Low Value	Baseline	High Value
Pipe Failure Rate (failures/km/annum)	0.1	0.2	0.5
Effective Failure Rate (10km) (failures/annum)	1	2	5

The result of varying the pipe failure rate is mostly insignificant above the design criterion. At the design criterion failure frequency level, the effect of pipe failure frequency has little effect. This is proposed as owing to the domination of the pump failure duration influence over the model (discussed in 4.3.6) and the resultant increase in supply ratio and reservoir capacity. The high-failure model appears to produce results of a similar efficiency compared to the baseline. However, this is not the case, the variation is owing to the sparsity of very low failure frequency solutions and the effect that this has on optimisation in this range. The results from the sensitivity analysis are demonstrated in Figure 4.3.5.

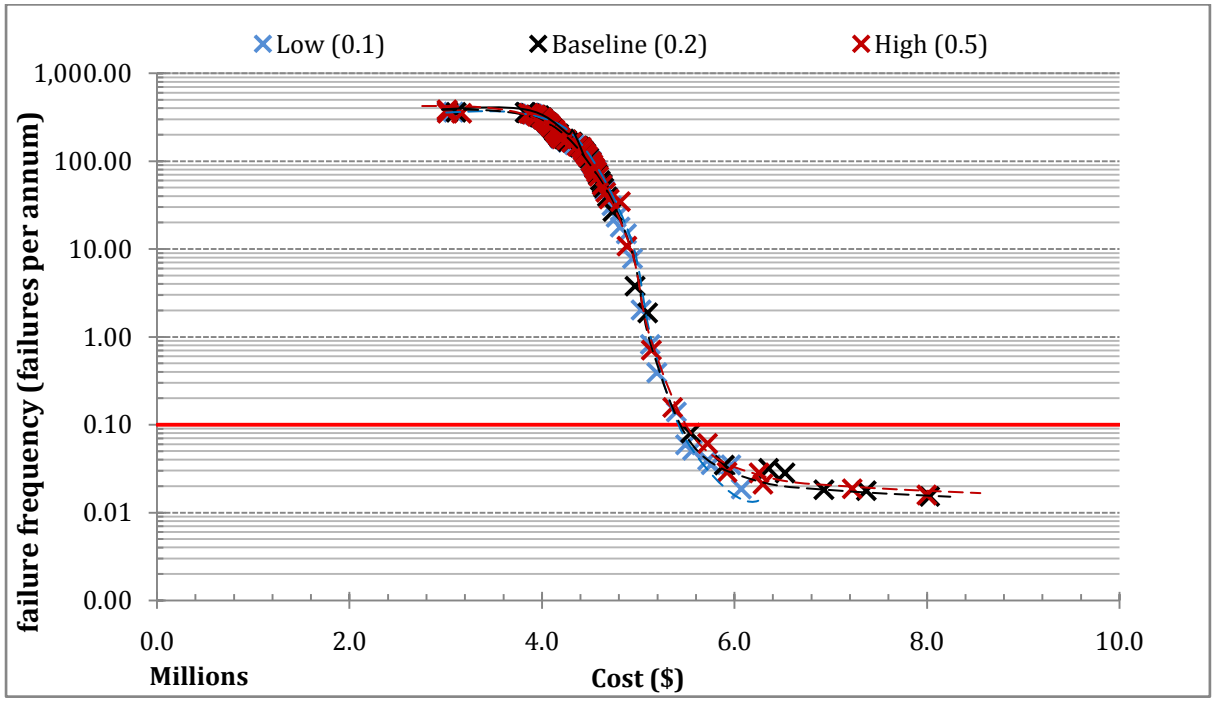

Figure 4.3.5: Pareto-optimal solution front for varying pipe failure frequency

4.3.6. Sensitivity to pipe failure duration

The values to be tested and results for cost sensitivity to pipe failure are presented in Table 4.3.8, and are the same as those tested by Chang & van Zyl (2012). These values represent the mean duration, which allows for the variation in repair duration (influenced by various environmental conditions).

Table 4.3.8: Pipe failure duration sensitivity values

Parameter	Low Value	Baseline	High Value
Pipe Failure Duration (h) – Poisson Mean	3	4.5	9

As with pipe failure frequency, the systems that achieve the design criterion of 1 failure every 10 years under seasonal peak demand, have large reservoir capacities as an adaption to the pump failure events. The order of reservoir capacities is in the order of 25 hours of seasonal peak demand (or 37.5 h annual average daily demand). This provides a very substantial buffer against both high frequency and long duration pipe failure events (4.5 hour average pipe failure duration, compared to 48 hour average pump failure duration). The results from the sensitivity analysis are demonstrated in Figure 4.3.6.

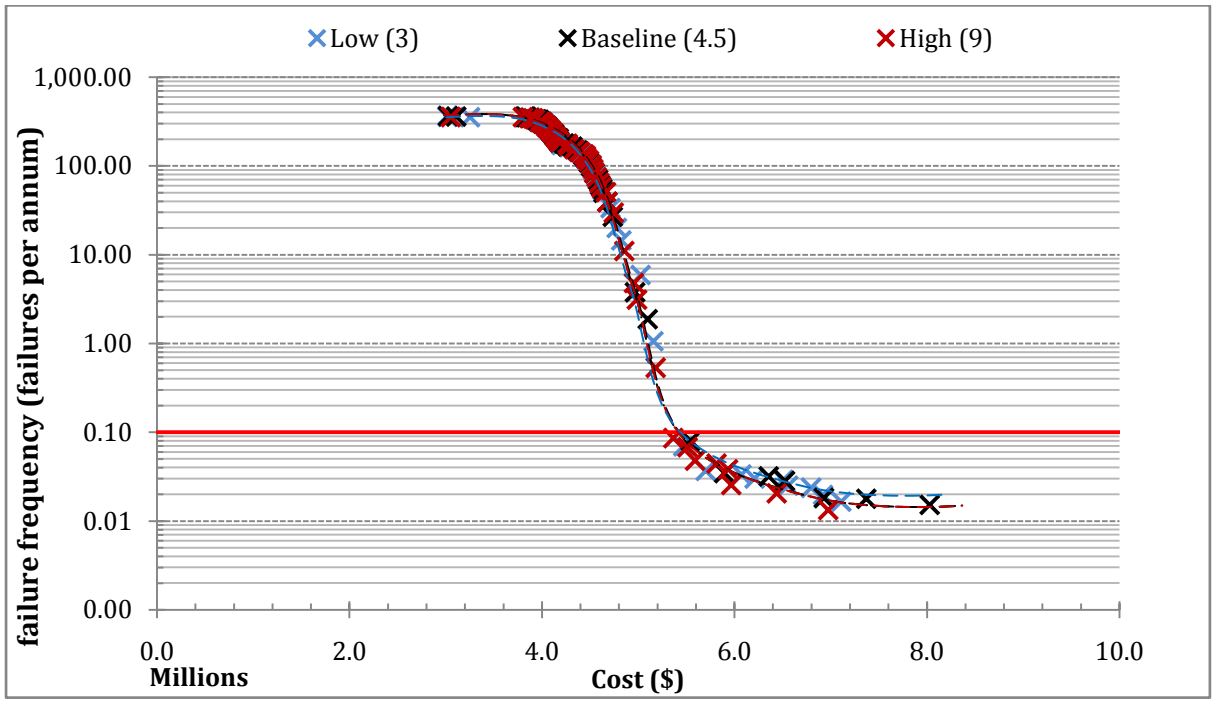

Figure 4.3.6: Pareto-optimal front for varying pipe failure duration

4.3.7. Sensitivity to fire event frequency

Fire flow, as a result of fire-fighting demand, places strain on the bulk water supply system. Fire events that happen during the seasonal peak result in the highest water demand condition experienced by the system. By varying the tested value, the sensitivity of the system to the frequency that peak demand occurs at, can be observed. The values to which the rate was varied are presented in Table 4.3.9, and are in line with the values tested by Chang & van Zyl (2012). As with pipe failure and pipe frequency, the effect of fire frequency on failure frequency is insignificant owing to chromosomes that are optimised to resist failure to long duration failures. The implication is a system is resilient to pipe failures and fire events.

Table 4.3.9: Fire frequency sensitivity values

Parameter	Low Value	Baseline	High Value
Fire event frequency (average fires/annum)	0	6	24

The large variation in fire durations tested is owing to the varying environmental conditions possible in various human settlements. From basic, isolated house fires, to large area of effect bush fires, found in dry and/or populous areas.

Only slight variation in system cost with varying fire frequency is observed. This is intuitively correct as the failure frequency level that is considered during the optimal design approach, the design criterion necessitates either (or both) a large supply ratio and a substantial reservoir capacity to buffer against pump and pipe failure. In addition, the fire events last 0.84 hours on average, if the fire event is not super positioned with a supply failure event, only systems with a small capacity reservoir and/or low supply ratio are likely to experience a significant decrease in reservoir level. The overall result is that the cost efficiency of a design criterion solution that is resilient to long duration supply failure events is insensitive to fire event frequency.

In the very low failure frequency range (+- 0.01 failures per annum), the results exhibit random variance owing to the small number of failures experienced during the simulation.

The results from the sensitivity analysis are demonstrated in Figure 4.3.7.

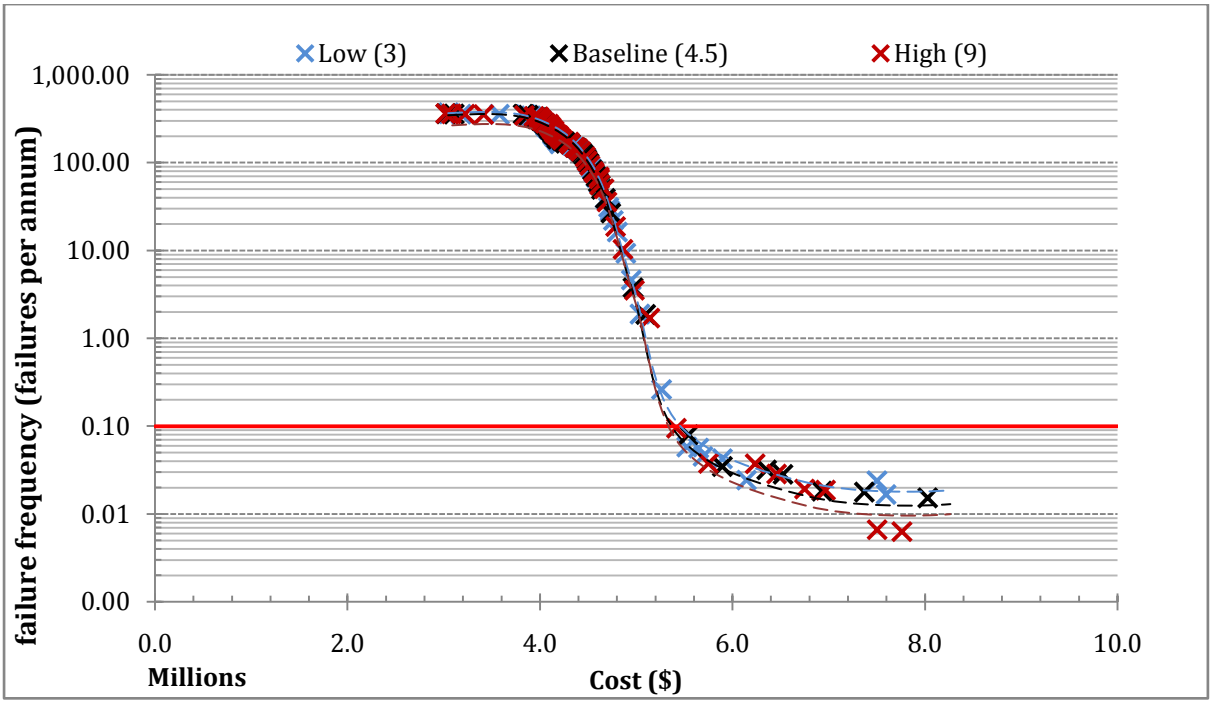

Figure 4.3.7: Pareto-optimal front for varying fire event frequency

4.3.8. Sensitivity to fire event duration

The longer the bulk supply system is subjected to a fire event, the longer the reservoir has to buffer against the reduced net inflow (or increased net outflow). The longer the fire event duration, the greater the possibility of consequent supply failure event occurring. The sensitivity of the supply system to these events needs to be tested in order to gauge the tolerance for variation. The values tested are listed in Table 4.3.10, in line with those tested by Chang & van Zyl (2012).

Table 4.3.10: Fire duration sensitivity values

Parameter	Low Value	Baseline	High Value
Fire event duration (Poisson mean (hours))	0.44	0.84	1.62

The values in Table 4.3.10 vary significantly with respect to each other, however when compared to a system that experiences pump failure durations with an average of 48 hours, the duration appears insignificant. The results of the sensitivity analysis are presented in Figure 4.3.8. In line with the observation regarding the fire event durations to be tested, the results appear to be largely unaffected by the change in fire event duration. The same

reasoning, as for the fire event frequency can be applied. The duration is not significant enough to cause an already largely reliable solution system, at or below design criterion, to fail more frequently, as the size of the components built into the system are resilient to long ,supply (pump failure) interruptions.

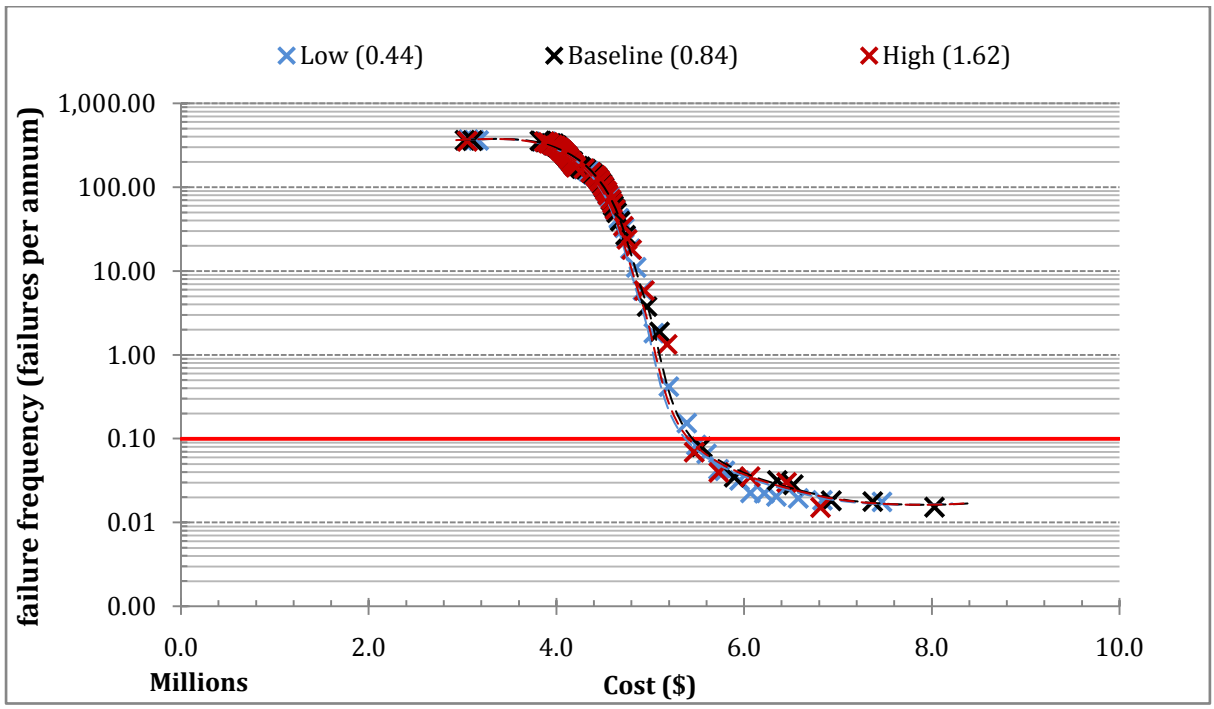

Figure 4.3.8: Pareto-optimal front for varying fire event duration

4.3.9. Sensitivity to pipe length

As was demonstrated in section 4.2.2 above, the pipe system cost makes up approximately 45% of the overall system cost. The pipe cost is influenced by the number of pipes, the length of the pipes and pipe diameter and is determined for each pipe from the equation presented below:

$$Cc_{Pipe} = 480 \times L_{Pipe} \times (D_{Pipe})^{0.935}$$

Where:

Cc_{Pipe}	= Pipeline capital cost	($)
L_{Pipe}	= Length of pipe	(m)
D_{Pipe}	= Diameter of pipe	(m)

From this equation, as presented previously in 3.6.6, the resultant increase in the cost of the pipeline system as a response to an increase in the pipe length is evident and easily calculated. However what is not as simply determined is the effect that the extra pipe length has on the pump and pipe sizing, and the impact of the sizing of these supply parameters on system configuration and cost. An increase in pipe length means an increase in primary, frictional losses. In order to respond to the pressure and flow reduction, the pump size must be increased, however the increase in flow results in a subsequent increase in frictional losses. This creates evolutionary pressure to increase the pipe diameter, and reduce pump power. The increased supply side cost also creates evolutionary pressure toward reservoir capacity as a more cost effective means of decreasing failure frequency; however increasing reservoir capacity is only effective at reducing failure frequency over a certain range of supply ratios. The lengths to be used in the variation of the pipe system are presented in Table 4.3.11, in line with the values tested by Chang & van Zyl (2012).

Table 4.3.11: Pipe system length sensitivity values

Parameter	Low Value	Baseline	High Value
Pipe System Length (km)	1	10	100

As can be seen in Figure 4.3.9, the difference in cost is highly substantial, with the high value pipe system cost a factor of 100 more than the low value pipe system cost. The total system cost increases at a higher rate in the high value system, owing to the additional cost to overcome frictional losses through increasing pump size. Equally, increasing pipe diameter has a far more substantial cost in a 100 km pipe system compared to a 10 km pipe system.. The observation however should be taken in the relative cost of solutions that meet the design criterion failure frequency.

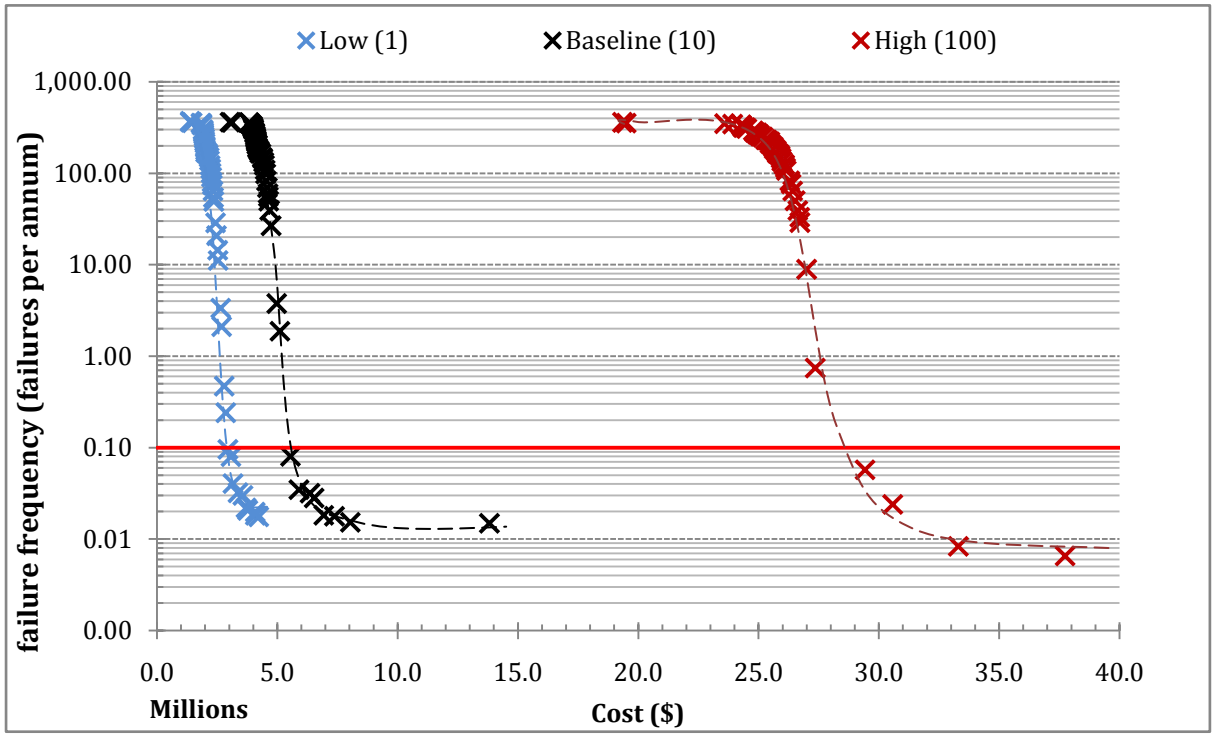

Figure 4.3.9: Pareto-optimal front for varying pipe length

The shift in balance between system components as the length of the pipe increases provides insight into the cost implication of system component design under varying system constraints. This variation is represented as component size, as a percentage relative to the baseline in Figure 4.3.10. This demonstrates the component sizing of the nearest solution that meets the design criterion. A significant pump power response to increase in length is observed as the reservoir needs to experience a certain supply ratio regardless of pipe length, to attain a certain failure frequency (in this case the design criterion). The increase in pump power is non-linear. This is due, in part, to the non-linear increase in frictional losses and the additional power required in overcoming them. The converse is true for the decrease in pump power for the low value system. The lesser increase in pipe diameter is as a result of the large increase in cost involved in increasing the size of a pipe that is 100 km in length. The reservoir increase is also relatively minor and is likely due to the inefficiency of increasing the reservoir capacity in the design criterion failure frequency zone. What is proposed is that increasing the supply ratio at the design criterion remains an efficient means of decreasing failure frequency for long length pipelines. The distinct lack of multiple pipe systems in the short distance model, that could be expected owing to the reduced cost associated with adding an extra pipe, is likely due to the reduced pipe failure rate (in failures/**km**/annum), and

the relative insensitivity to pipe failure frequency and duration, as demonstrated in 4.3.5 and 4.3.6. The result is the complete inefficiency of pipe systems consisting of more than a single pipe.

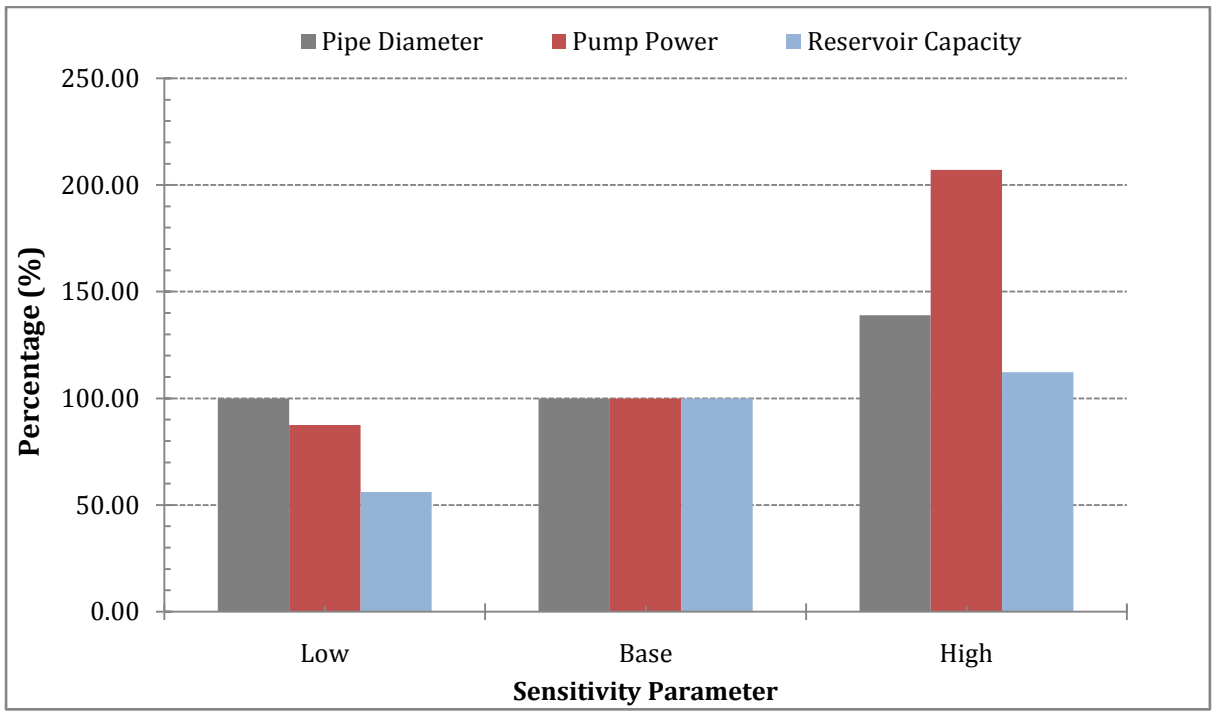

Figure 4.3.10: Relative Component Sizing for Varying Pipe Length

4.3.10.Sensitivity to static head

Following from the testing of pipe length sensitivity, the sensitivity of pumped systems to an increase in static head remains to be tested. The increase in static head necessitates an increase in pump power in order to both transport and lift the required reservoir supply. This is the case for most systems as the reservoir needs to supply a constant pressure to the demanding population, and it is done by making use of a difference in elevation.

The parameters to be tested are presented in Table 4.3.12, in line with Chang & van Zyl (2012)

Table 4.3.12: Static head sensitivity values

Parameter	Low Value	Baseline	High Value
Static Head (m)	30	60	120

The results of the static head sensitivity analysis are presented in Figure 4.3.11.

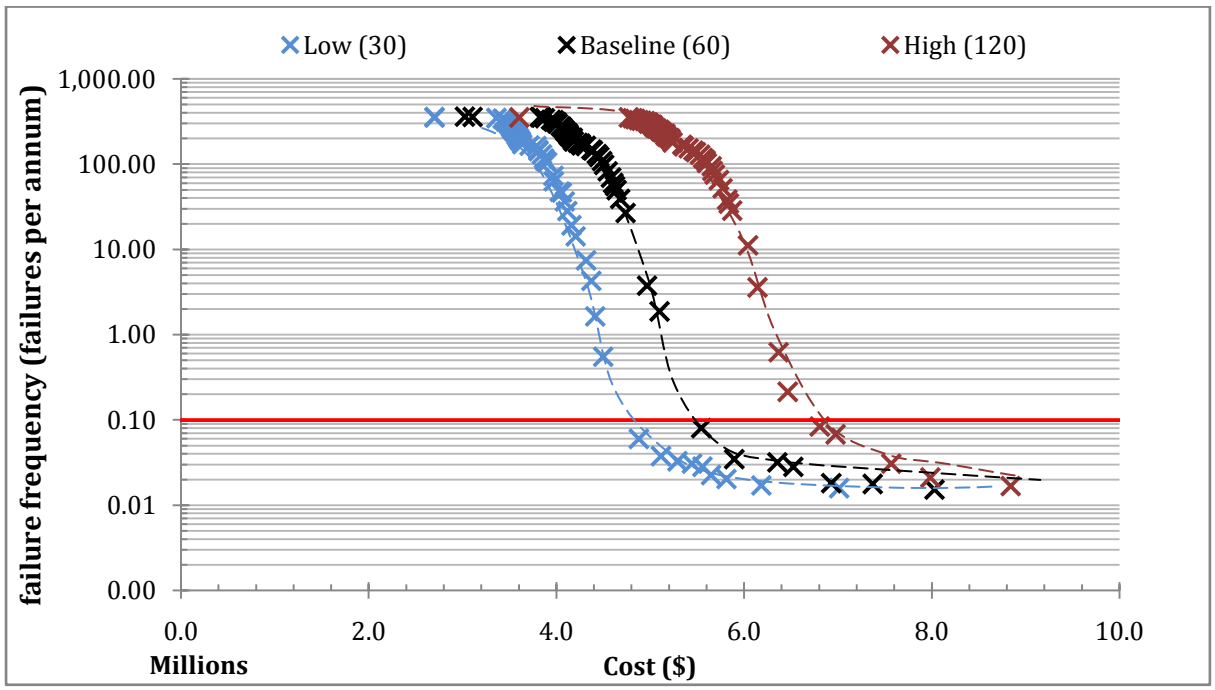

Figure 4.3.11: Pareto -optimal Front for Varying Static Head

The increase, and decrease, in static head reveal a similar cost offset for each possible solution, as for the varying pipe length. However, the magnitude of the cost offset is substantially lower, and the cost gradient for reducing failure frequency is relatively constant between different systems. This is likely due to the lesser frictional losses, as the length of the pipe remains unchanged.

The higher elevation, high value system exhibits a proportionally larger difference in cost, compared to the baseline system, than the lower elevation, low value system. This is likely owing to the secondary costs associate with an increased pump power (4 pump systems, energy cost, etc). What can also be observed is that the minimum failure frequency (+- 0.02 failures per annum) is unchanged as the physical (as opposed to statistical/stochastic) constraints do not affect the failure frequency of the system, only the cost of attaining a certain failure frequency.

The effect on component sizing, owing to a change in static head, is demonstrated in Figure 4.3.12, below:

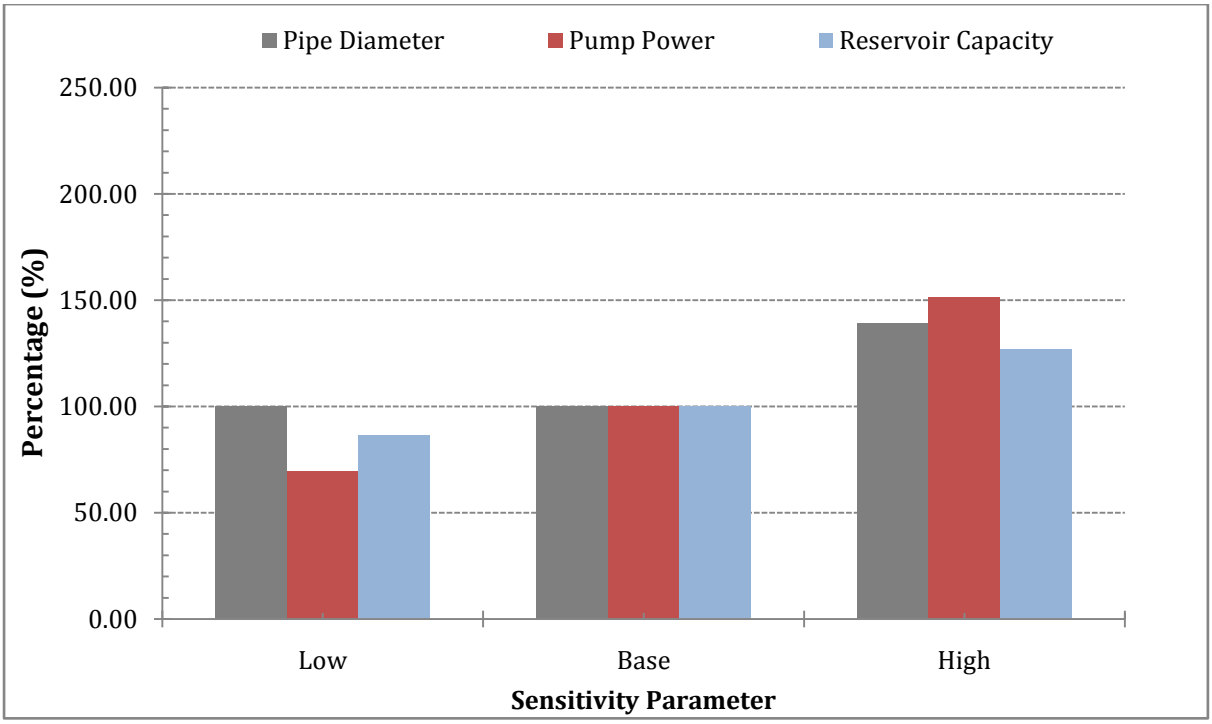

Figure 4.3.12: Relative component sizing for varying static head

The increase in pump power from the low static head to high static head system is near linear. The genetic optimisation mechanism of varying reservoir capacity to the point where it's utility in increasing (or decreasing) failure frequency becomes inefficient, is evident. When comparing the pump power increase in Figure 4.3.11 to Figure 4.3.9, the reservoir capacity increase is comparable. The decrease in reservoir capacity from the base system to the low static head system is minimal, as a larger relative decrease in pump power results in a more efficient system. The inverse is true for the varying pipe length model.

4.3.11. Sensitivity to focusing by parameter restriction

The NSGA-II genetic optimisation algorithm uses crowding distance as a means avoiding solutions that are too closely spaced. The result is that the model tends to space solutions over the entire failure frequency range (approximately 0.02 failures per annum to 360 failures per annum). As the vast majority of solutions generated fall outside of the range acceptable by the design criterion of 1 failure every 10 years under seasonal peak demand, what was looked into was a method of focusing the optimisation algorithm. This would be done in order to only generate and optimize solutions that were within the bounds of what could be considered feasible.

A workable solution found is to first run the model to determine the size of the system components of those systems that meet or are situated in the desired failure frequency criterion range. This exercise was performed for the base model and the following parameters, shown in Table 4.3.14, were identified as describing the bounding limits of the near design criterion failure frequency level:

Table 4.3.13: Design criterion failure frequency component sizes

Parameter	Baseline Value	Restricted Value
Pump power range (kW)	[15, 1500]	[95,115]
Pipe diameter range (m)	26 discrete diameters between [0.127 - 2.174]	3 discrete diameters: [0.363, 0.428, 479]
Range of reservoir capacity (hours SPD)	[4,48]	[27,48]

This limitation of the decision variables forces the optimisation algorithm to choose an initial population and perform crossover/mutation with a boundary defined by higher minimum and lower maximum values.

The advantage of this method is that all solutions in the range are valid, and representative of the true cost of the system. The disadvantage is in the uncertainty involved with selecting the appropriate values to set for the raised minimums, and the potential to introduce inefficiencies. The results of the focusing are presented in Figure 4.3.13.

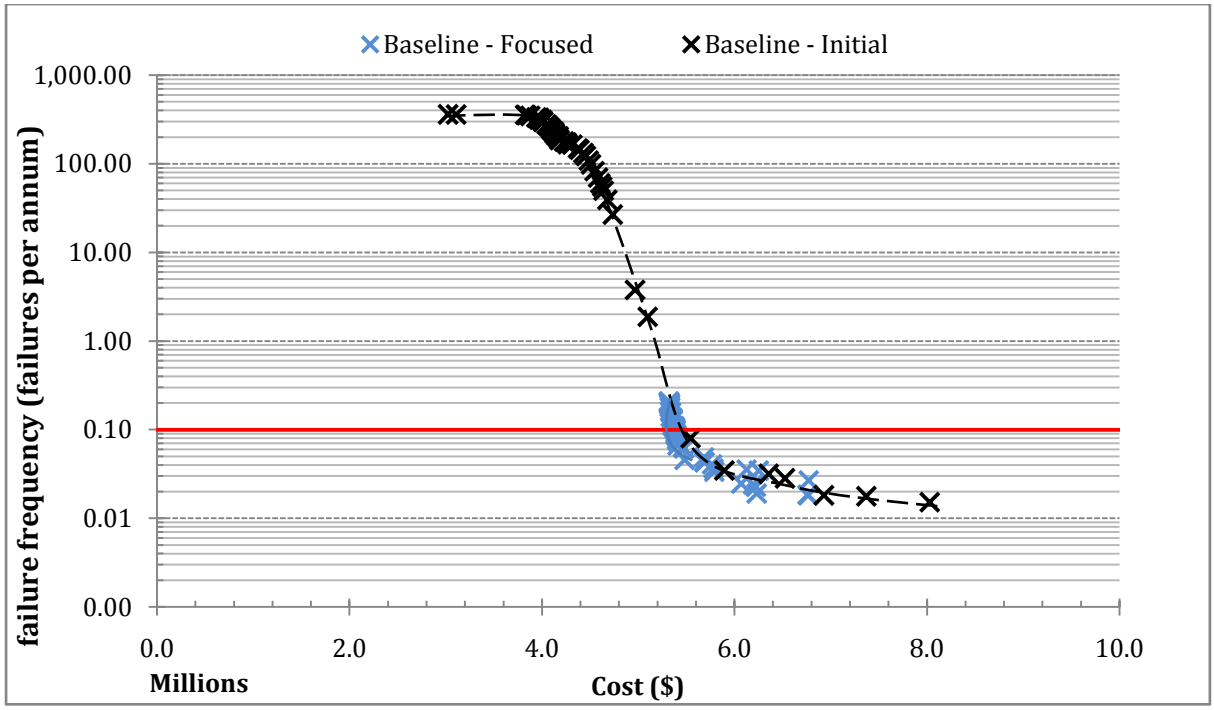

Figure 4.3.13: Pareto -optimal solution fronts for initial and focused models

As is observed from the above Figure, the focusing through implementing limitations on the decision variables has successfully produced results in the near-design criterion failure frequency range. The near design criterion failure frequency range of solutions is demonstrated in Figure 4.3.14.

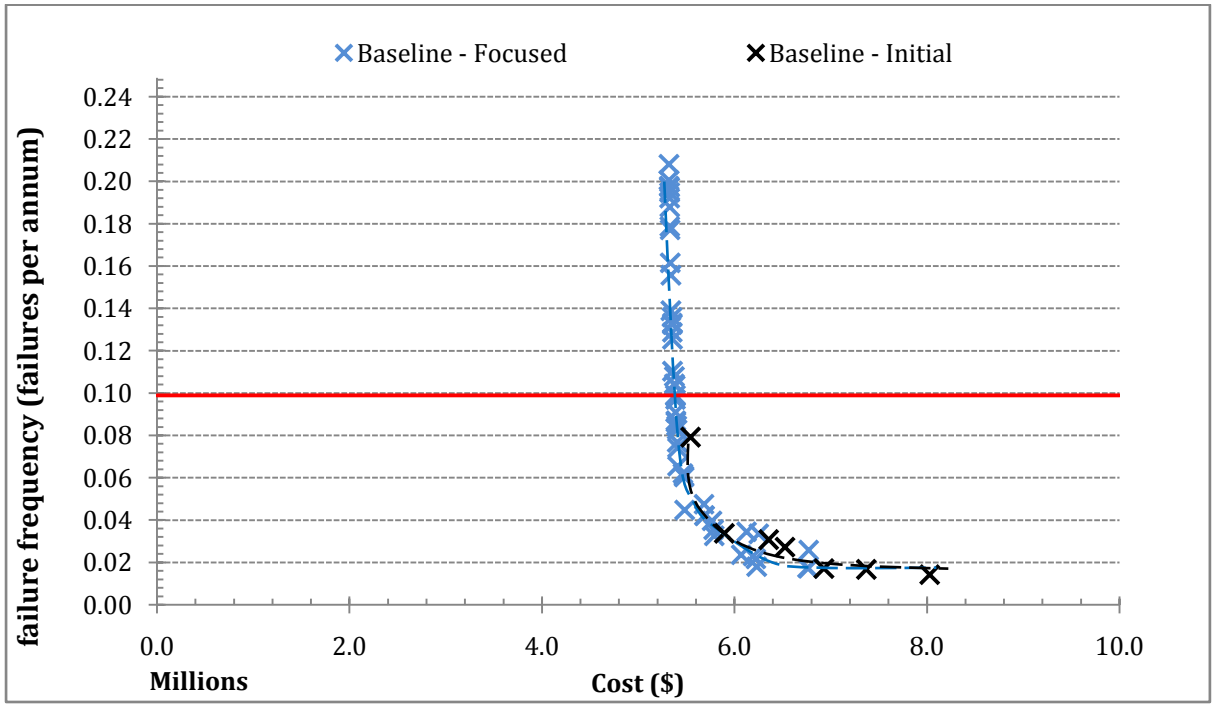

Figure 4.3.14: Pareto-optimal solution fronts for initial and focused models (failure criterion)

As can be seen in Figure 4.3.15, this method is effective in focusing the optimisation process to a range of acceptable solutions. By adjusting the decision variable limits, the lowest failure frequency was reduced from 360 failures per annum to 0.21 failures per annum.

The near-vertical slope of the solution curve is as a result of the high sensitivity of the model to slight variations in sizing of system components in the near design criterion failure frequency zone. The percentage deviation from the design criterion solution for each of the system component sizes is demonstrated in Figure 4.3.15. Pipe diameter remains constant over the range of solutions presented, and is therefore not shown.

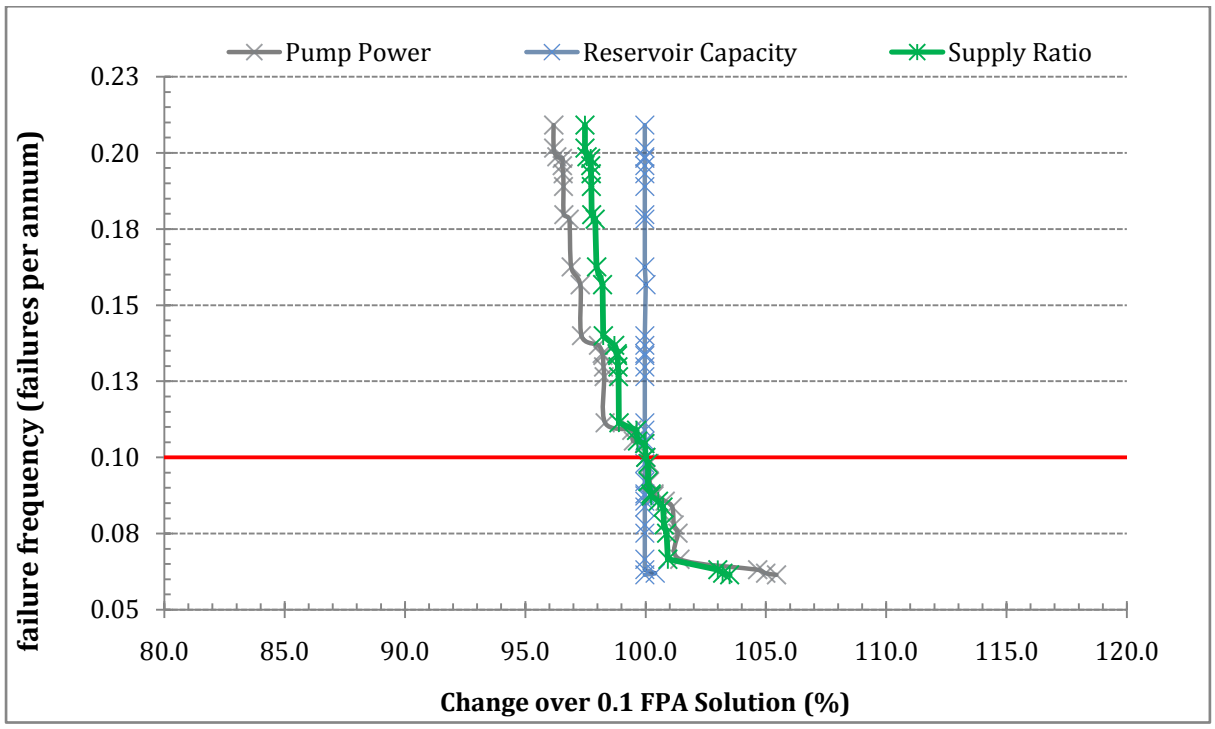

Figure 4.3.15: Relative component sizing as percentage of a design criterion solution

What is observed from the above Figure is that over a large range of solutions, near the desired failure frequency, the most efficient means of optimisation remains modulation of supply, through the modulation of pump power. This reiterates the importance and sensitivity of system failure frequency to supply in the near design criterion range of solutions. The difference in pump power between a 1 in 17 year failure frequency solution and a 1 in 5 year solution is 9% and the difference in supply ratio, only 6%. As mentioned previously, this method does have the disadvantage of requiring that the minimums be manually set by the designer. However, it is likely that the designer will have a reasonable

amount of experience and can thus set the decision variables to values that he/she deems to be reasonable, based on the initial, wide-spectrum simulation.

4.3.12. Summary

Both the physical and stochastic parameters, as well as the optimisation model itself, were varied and tested. The statistical parameters describing the population demand (consumption) and demand patterns were not covered in this sensitivity analysis as the goal was to test the sensitivity of a model simulating a generic population. The adaptation of the model to a specific population to be serviced by the design system is beyond the scope of this investigation. The solution population size (> 50) was also not tested as the purpose of this sensitivity analysis and investigation was to test the capability of the optimisation model to generate both reasonable solutions and demonstrate the effect of the variances on the entire failure frequency spectrum. A finer resolution of solutions can be generated, but with a significant increase in computational time cost.

The model was most sensitive to variation in pipe length, with large variations in cost following an exponentially increasing trend owing to the large pipe cost, and increase in pump power necessitated by the increase in frictional losses. The variation in static head against which the integrated pump must operate has a less critical impact, and follows a linear increase with increasing head. The sensitivity of the optimal solution to variations in the stochastic model: pipe failure, pump (power) failure and fire events was investigated and the results compared to the results from the physical parameter sensitivity analyses, this is demonstrated in figures 4.3.15 and 4.3.16.

A large factor in ensuring that a model is applicable is ensuring that it can be adapted to generate solutions that are reasonable. It was proven with the focusing models, that it is possible to generate only a range of solutions that are of an acceptable failure frequency by restraining the system components to produce only reasonable values. The increase in efficiency of doing so is between 2-5%. This is an example of the synthesis between engineering experience and intuition, and the tool of an optimisation model.
However, as the intention was to compare not only those solutions that met the desired design criterion failure frequency level, but those across the full spectrum of failure frequencies, the sensitivity analyses, in general, were not focused.

The results of the sensitivity analyses for the base system are presented in Figure 4.3.16.

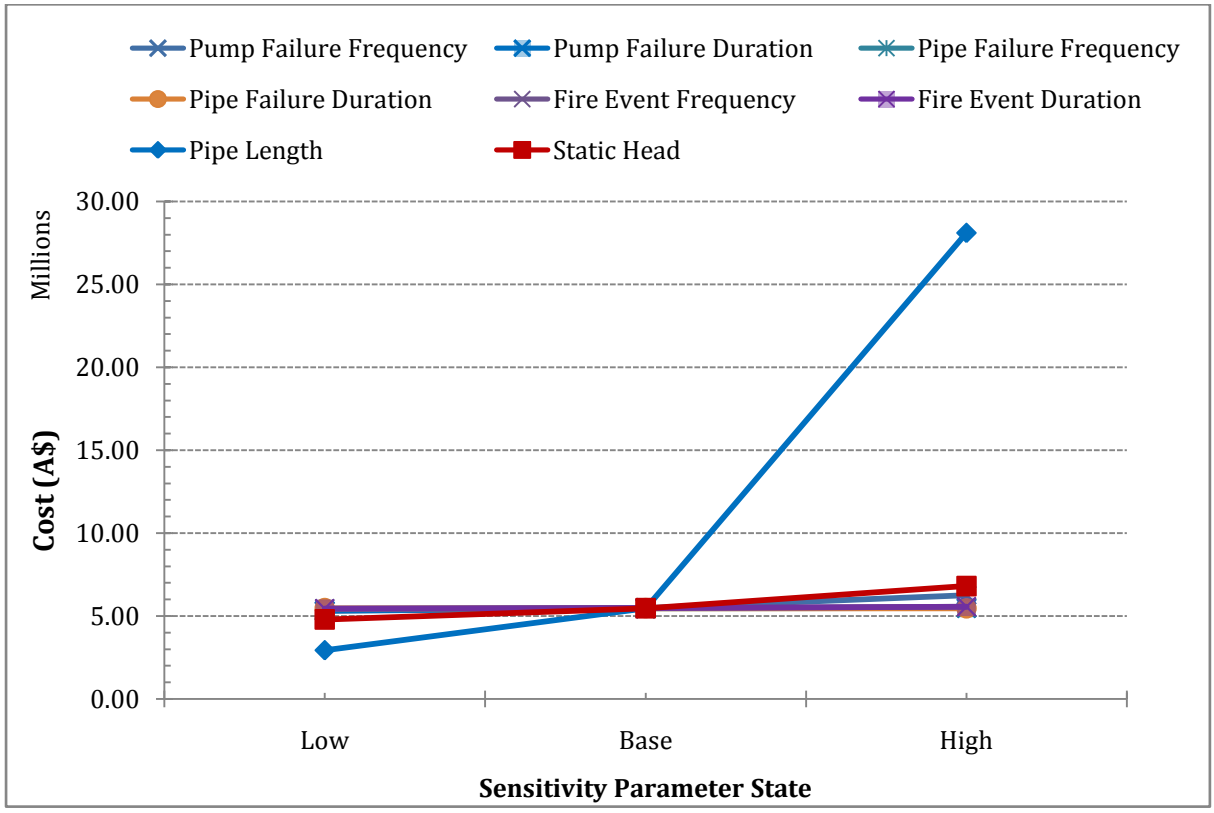

Figure 4.3.16: Comparison of sensitivity analysis parameters

The change in system cost with pipe length is highly prominent and greater than the variations in stochastic parameters by a factor of 5. This, however, provides a poor base for comparison between stochastic parameters. A comparison, excluding pipe length, is presented in Figure 4.3.17.

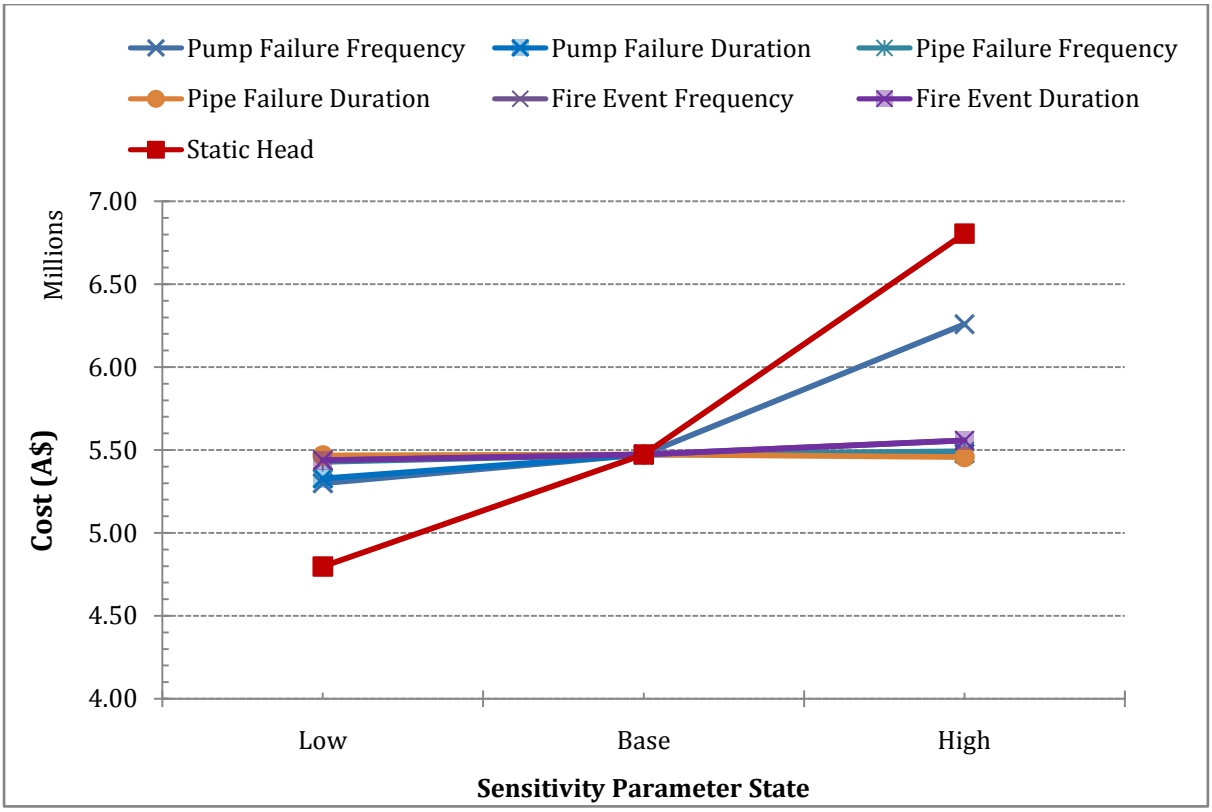

Figure 4.3.17: Comparison between sensitivity analysis parameters (excluding pipe length)

As expected, the increase in static head and associated increase in pump power has a significant effect on system cost. The pump failure frequency has a prominent effect on system cost when increased, as the cost to increase pump and reservoir capacity to achieve the desired design criterion failure frequency increases sharply as pump failure becomes more frequent. A decrease in pump failure frequency does not result in an equal reduction in system cost, which is evident of the non-linear relationship between desired failure frequency and cost and the criticality of pump failure frequency. Pump failure duration has a far lesser effect of system cost when increased, as the average duration of 48 hours is already so large that a failure that is above average has a very high probability of causing a system failure, the effect of increasing the duration further does not result in a significant number of additional failures. The inverse is not true, an average duration of 12 hours, allows for a reasonable probability of a system resisting failure and this is evident in the reduction in cost. The remaining stochastic model parameters have a largely insignificant effect in comparison to the pumping system parameters. All sensitivity parameters are presented in Table 4.3.14.

Table 4.3.14: Sensitivity analysis, solutions meeting design criterion failure frequency (Generation 15 - 50) - Seasonal peak demand

Description	Pump Power	Pipe Diam	Reservoir Capacity	Supply Ratio	Failure Frequency		Total Cost	Total Pipe	Total Pump	Total Tank
	kW	*m*	*Hour*	*#*	*(1:Years)*	*Fail/annum*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*
Base - USADOE/NERC(2014)	99.80	0.36	29.57	1.13	11.77	0.085	5,473,323	2,060,928	2,030,030	1,109,037
Pump Failure Frequency - Low	101.09	0.36	22.72	1.14	11.33	0.088	5,299,846	2,060,928	2,052,336	1,186,582
Pump Failure Frequency - high	120.20	0.36	47.52	1.27	11.00	0.091	6,259,136	2,060,928	2,378,005	1,820,203
Pump Failure Duration - Low	115.00	0.36	16.29	1.24	10.45	0.096	5,329,225	2,060,928	2,290,149	978,149
Pump Failure Duration - high	101.67	0.43	17.30	1.31	10.28	0.097	5,479,462	2,404,086	2,062,440	1,012,936
Base - Nel (2009)	138.90	0.32	33.83	1.21	10.92	0.092	6,027,453	1,842,448	2,690,487	1,494,518
Pump Failure Frequency - Nel (2009) - Low	106.82	0.36	23.78	1.18	10.38	0.096	5,429,970	2,060,928	2,150,839	1,218,203
Pump Failure Frequency - Nel (2009) - High	125.29	0.36	38.42	1.30	10.23	0.098	6,133,787	2,060,928	2,463,693	1,609,167
Pump Failure Duration - Nel (2009) - Low	125.02	0.36	13.85	1.30	10.59	0.094	5,410,336	2,060,928	2,459,185	890,223
Pump Failure Duration - Nel (2009) - High*	203.67	1.18	48.00	3.22	3.64*	0.275*	11,754,185	6,185,587	3,737,725	1,830,874
Load Shedding - Low (Stage 1)	108.53	0.36	26.73	1.19	10.74	0.093	5,544,794	2,060,928	2,180,056	1,303,811
Load Shedding - Base (Stage 2)	118.50	0.43	18.53	1.46	12.87	0.078	5,807,468	2,404,086	2,349,383	1,053,999
Load Shedding - High (Stage 3B/4)	139.15	0.57	33.99	1.98	11.40	0.088	7,356,521	3,163,244	2,694,634	1,498,644
Pipe Failure Frequency - Low	89.59	0.43	22.23	1.19	11.65	0.086	5,428,137	2,404,086	1,852,600	1,171,451
Pipe Failure Frequency - High	103.65	0.43	16.68	1.33	12.28	0.081	5,492,143	2,404,086	2,096,425	991,632
Pipe Failure Duration - Low	118.50	0.36	18.62	1.26	10.81	0.092	5,467,449	2,060,928	2,349,341	1,057,181
Pipe Failure Duration - High	104.72	0.36	25.98	1.17	10.53	0.095	5,458,087	2,060,928	2,114,783	1,282,376
Fire Event Frequency - Low	111.38	0.36	21.49	1.21	10.64	0.094	5,438,333	2,060,928	2,228,622	1,148,784
Fire Event Frequency - High	97.90	0.36	34.05	1.12	10.84	0.092	5,558,445	2,060,928	1,997,278	1,500,240

Description	Pump Power	Pipe Diam	Reservoir Capacity	Supply Ratio	Failure Frequency		Total Cost	Total Pipe	Total Pump	Total Tank
Continued	*kW*	*m*	*Hour ()*	*#*	*(1:Years)*	*Fail/annum*	*Cost ($)*	*Cost ($)*	*Cost ($)*	*Cost ($)*
Fire Event Duration - Low	101.40	0.36	25.79	1.14	10.22	0.098	5,395,605	2,060,928	2,057,732	1,276,945
Fire Event Duration - High	85.39	0.48	21.00	1.21	10.97	0.091	5,583,145	2,670,938	1,778,836	1,133,371
Pipe Length - Low	84.96	0.36	15.83	1.29	10.40	0.096	2,939,326	206,093	1,771,228	962,005
Pipe Length - High	367.76	0.36	25.13	1.14	11.31	0.088	28,110,306	20,609,276	6,243,096	1,257,934
Static Head - Low	67.57	0.36	25.77	1.16	10.23	0.098	4,797,547	2,060,928	1,460,103	1,276,517
Static Head - High	153.89	0.43	32.66	1.12	11.92	0.084	6,805,900	2,404,086	2,937,377	1,464,437
Focusing - Parameter Restriction	98.78	0.36	27.01	1.12	9.99	0.100	5,385,016	2,060,928	2,012,421	1,311,668

*Nel (2009), high duration system does not meet the desired design criterion, however it was demonstrated to show that a high-frequency (11.4 Power failure events per annum (average), high-duration, system cannot obtain the desired reliability without significantly increasing the resevoir capacity above 48 hours seasonal peak demand.

All solutions are shown in the above Table, including those not demonstrated in the sensitivity analysis section (owing to differing base systems). The wide spectrum of supply ratios and reservoir capacities gives credence to the assertion that the use of a deterministic guidelins that does not factor in the random influence has the potential to either result in an system that is not reliable enough, or is sufficiently reliable, but cost-inefficient. For the majority of solutions, the implementation of a 32 h seasonal peak demand reservoir capacity (48h AADD) reservoir would be in excess of what is stipulated and would not necessarily result in an appreciable decrease in failure frequency. Similarly, should a 1.0 seasonal peak demand, (1.5 AADD) supply ratio be implemented, the solution would not be able to achieve the desired reliability.

This crucial point is again demonstrated in Figure 4.3.18, where the supply ratio and reservoir capacity of the design criterion solution for each of the tested sensitivity vlauesis presented. Figure 4.3.19 demonstrates the equivalent relationship for the Chang & van Zyl (2012) results.

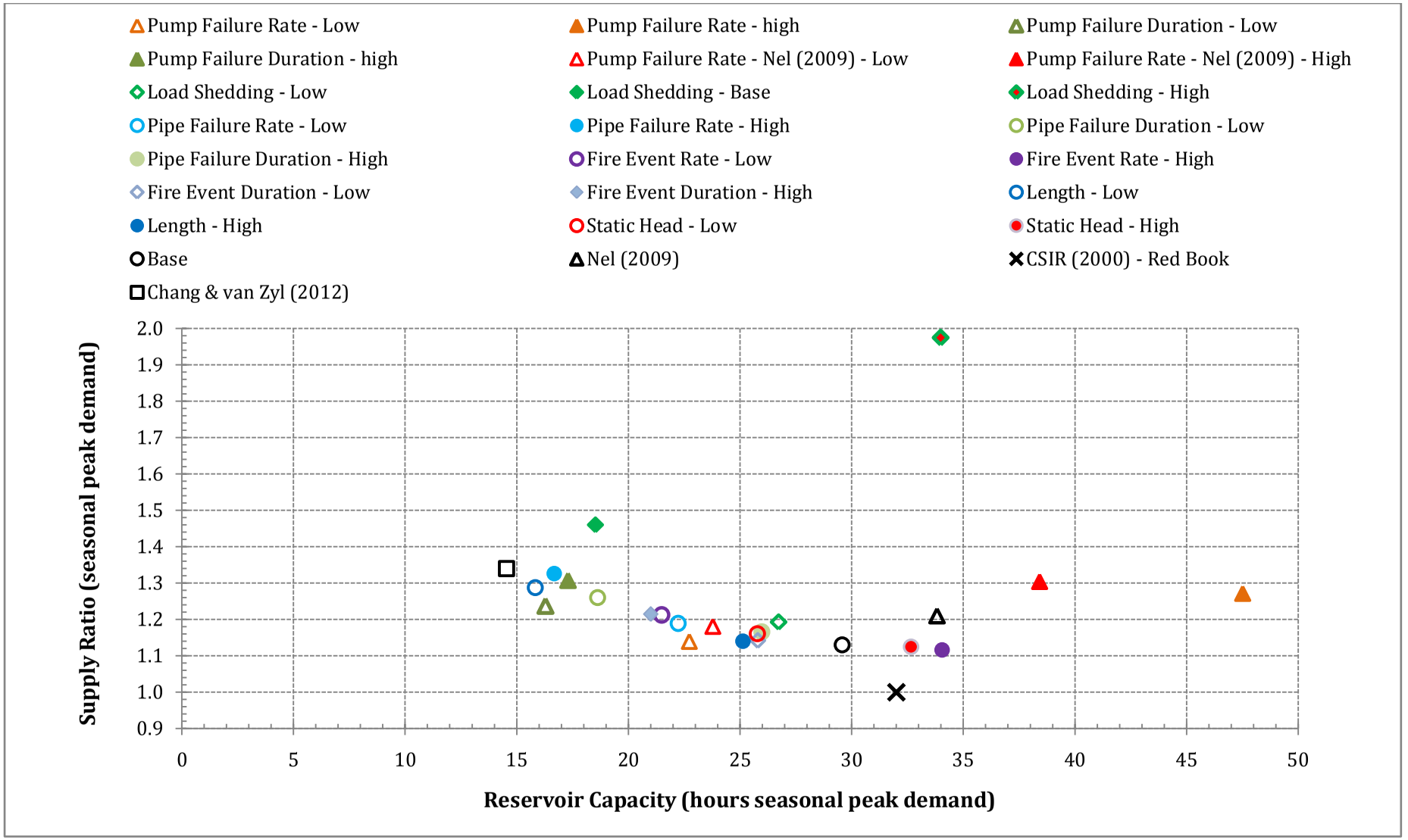

Figure 4.3.18: System trade-off (1 failure every 10 years) - Base model

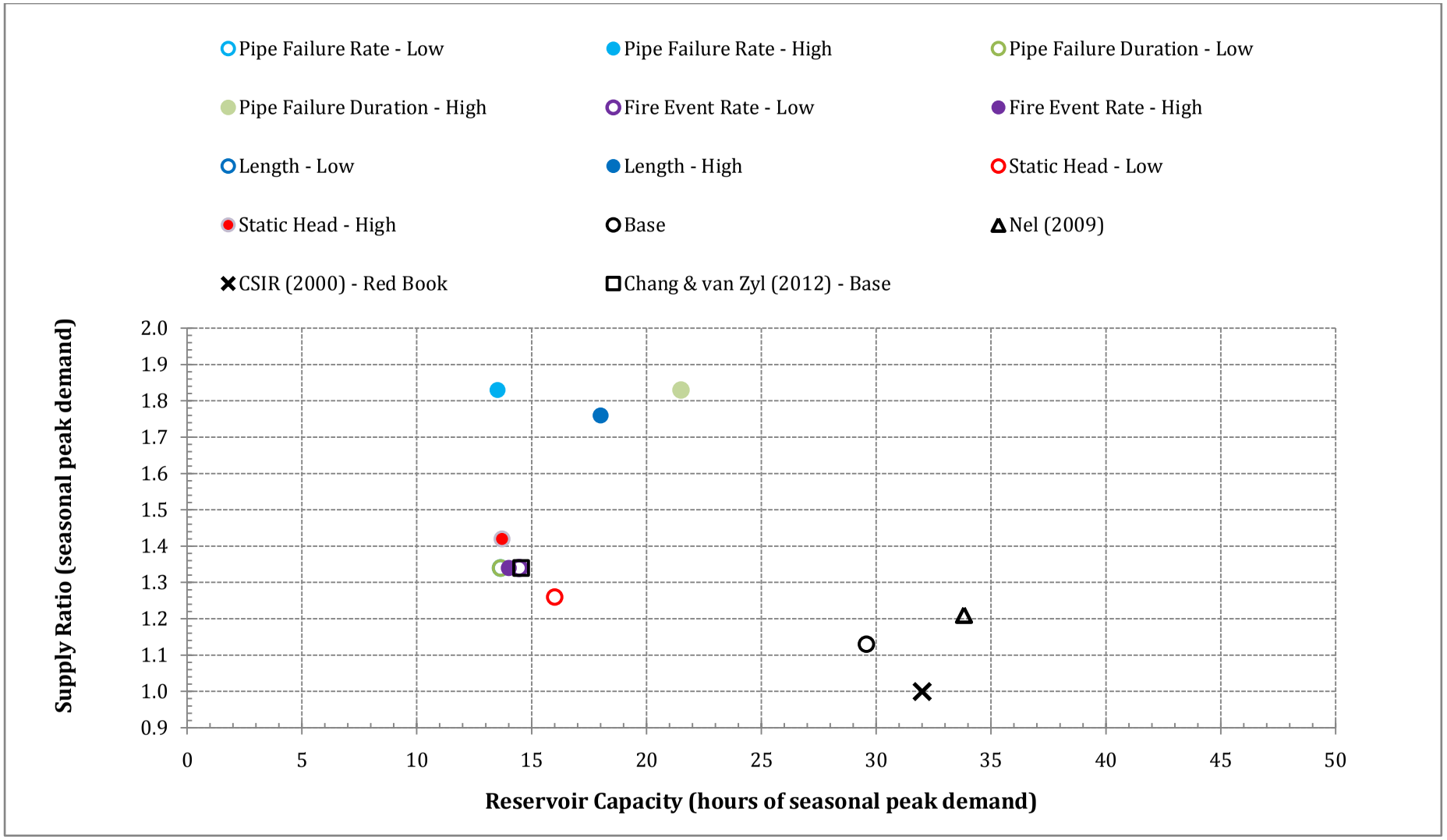

Figure 4.3.19: System trade-off (1 failure every 10 years) - Chang & van Zyl (2012)

Figure 4.3.17, shows that within the wide spectrum of solutions possible, there is no optimal solution that meets the desired design criterion with a supply ratio of less than 1. In addition, the majority of solutions, including the base solution do not require a reservoir capacity of 32 hours of seasonal peak demand (48 h annual average daily demand), with most solutions providing acceptable failure frequency with between 15 - 27 hours of reservoir capacity and between 1.1 and 1.5 supply ratio (1.65 and 2.25 annual average daily demand). The CSIR (2000)/Red Book solution stipulates a supply ratio of 1 that could be considered insufficient when compared to both the base and Chang & van Zyl (2012) base solutions.

As mentioned in 4.2.2, the Chang model produced solutions, in the near design criterion range, all with a supply ratio of 1.34. This was shown to be as a result of the limited number of pipe sizes available. It is proposed that if the model could utilise a finer resolution of pipe sizes, the pareto-optimal solution at the design criterion would have consisted of a lower supply ratio and higher reservoir capacity, resulting in a more efficient solution, that would be closer to both the base model and CSIR (2000) stipulation. The extension of this is that the sensitivity to supply ratio could not be fully explored. This is seen in Figure 4.3.19, with solutions crowding around the supply ratio of 1.34, with the exception being solutions of the high value model, that necessitate the next largest pipe size. The base model sensitivity results show a more diverse spread of solutions owing to the lesser restriction on supply ratio.

The closeness of both the base and Nel (2009) models to the CSIR (2000) stipulation show that consideration has been made in the CSIR (2000) guidelines for systems that are subject to a wide range of supply failures and demand variation events (fire events). The pump failure rate has a very pronounced influence on the reservoir sizing, for both the base and Nel (2009) models. The pipe failure rate and pipe failure duration all have a significant influence on supply ratio, with the lower pipe failure rate and duration showing a trend toward increased supply ratio and reduced reservoir capacity.Optimal results for load shedding results in an increased supply ratio and decreased reservoir capacity, to take advantage of refill times between failures, for the low (stage 1) and base (stage 2) levels. Stage 3B results in an inrease in both supply ratio and reservoir capacity in an effort to offset the effect of approximately 1600 load shedding events per annum. The cost buildip of each sensitivity test is presented in Figure 4.3.20, demonstrating the cost sensitivity of the model to variation in system parameters.

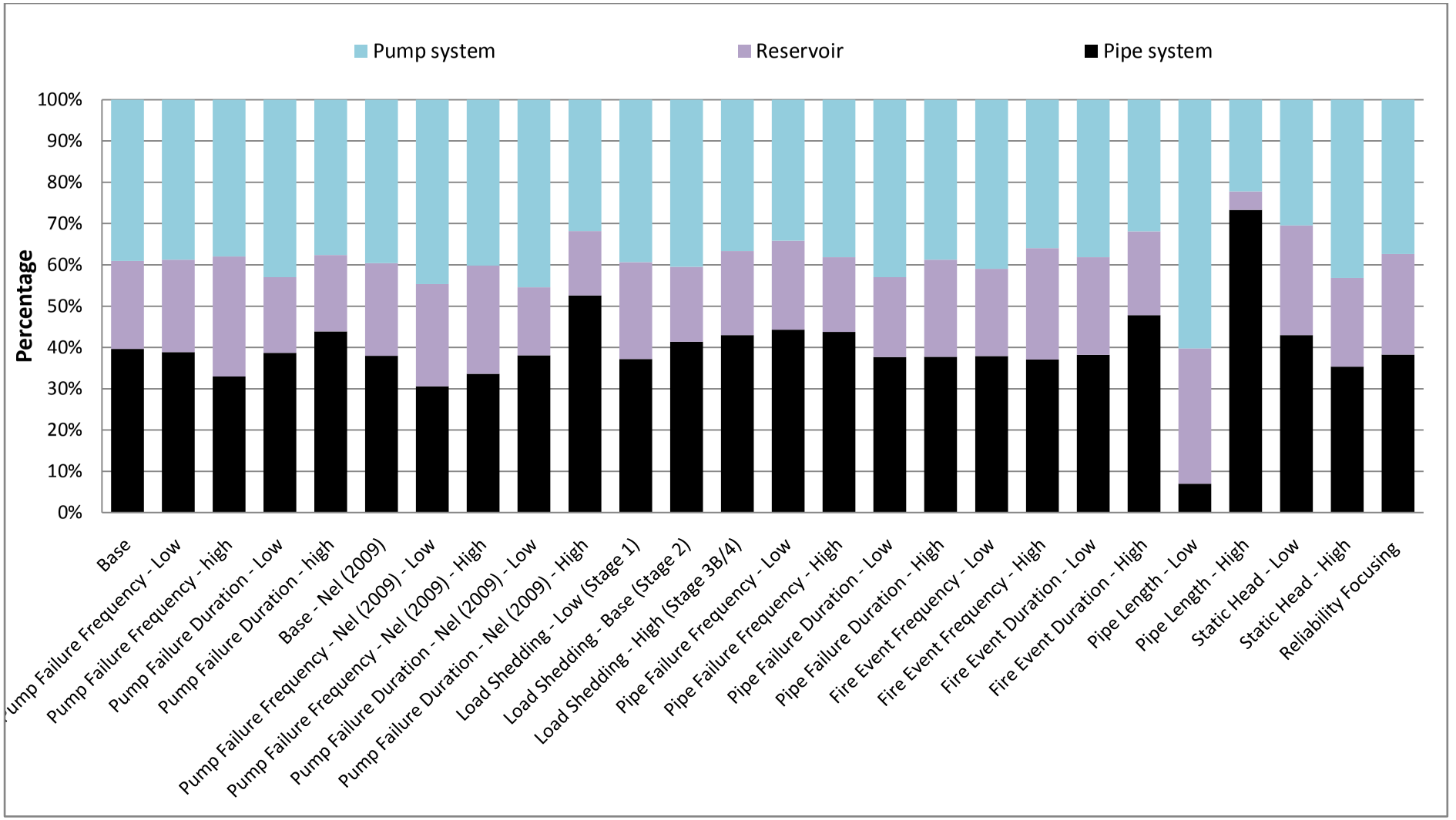

Figure 4.3.20: Relative system component cost - sensitivity parameters

The cost of the systems, as can be seen in Table 4.3.14 and Figure 4.3.20 are divided in an approximate proportion of 45/20/35, between pump/reservoir/pipe cost. This is intuitively so as the pumping system, inclusive of replacement, energy costs and all other life-cycle costs, over the design life is substantial. This cost is still dominant even compared to the significant capital cost of laying a pipe line. In the low and high length pipe system model, this balance is shifted toward pump and pipe cost respectively.

These results are in contrast to the findings of Chang & van Zyl (2012), which showed a 30/70 split between reservoir and pipe cost. As the original system did not include pumped flow, the 30/70 split can be compared to the 20/80 split between supply and storage component costs. In this regard, the supply-storage split cost of a gravity-fed system is comparable to that of a pumped system, with supply cost far outweighing storage cost. This is in line with the previous observation of the criticality and sensitivity of supply ratio around the desired design criterion failure frequency level. Should the design desire a far more reliable solution, in the order of 1 failure every 100 years, the likelihood is that this balance will shift toward a 50/50 split between supply and storage, or even further toward storage.

Pump failure frequency increases, increase reliance on reservoir capacity to provide a greater supply failure buffer and reduce the effect on reservoir level, from a supply failure. Pump failure duration increases, increase reliance on pipe size to provide greater inflow and recharge between failure events. Expectedly, the change in static head increases and decreases the pump cost, respectively.

The failure frequency focusing produces a slightly more efficient base solution by altering the cost split to 38/24/38. The overall observation is that aside from the major physical variations, the variance in cost split only varies between 5-8% for each of the 3 major cost components.

5. Conclusions

The investigation aimed to expand the existing optimisation model, developed by Chang & van Zyl (2012), by including pumping systems. This was done in order to increase the applicability of the model to real-world, bulk water supply systems using pumps. The electrical grid was modeled and integrated, as well as the addition of an external hydraulic solver. The shift in focus from pre-calculating supply ratios before optimisation, to the implementation of an enforced supply ratio range [0.5 - 5.0], allowed the investigation to make use of wide range of supply ratios during the optimisation process. This gave the model the ability to produce results that explore the effect of and relationship between supply ratio and reservoir capacity. Modeling of North American and South African power failures allowed for the testing of real world scenarios and sensitivities, in contrast to the desired design criterion of 1 failure every 10 years under seasonal peak conditions van Zyl *et al* (2008). A summary of the progression and development of the model as well as the mechanics of the model and results and sensitivities, are presented in the following sections.

5.1. Method summary

A literature review regarding the modeling, design and optimisation of bulk water supply systems was undertaken, with the intent to become familiar with the theoretical structure and context of the optimisation model under investigation. This included the basic definitions of the objectives of bulk water infrastructure design, stochastic process and its various applications. Optimisation and specifically genetic optimisation was detailed. In addition, the literature review aimed to provide theoretical backing to the stochastic pump system and cost models that needed to be developed.

The original model and its various mechanisms were detailed in depth for the purpose of adequately defining the base upon which the additions and amendments would be made. This was done in order that the effect of the alterations and additions on the mechanics of the model could be adequately placed and understood by the reader. The expanded, pump-inclusive model was developed with each of the components being outlined in detail. The explicit and in depth detail was for the purpose of obtaining transparency within the coding process.

The expanded model was constrained to model gravity-fed systems and the results were compared to the model produced by Chang & van Zyl (2012). This verified the integrity of the programming and accuracy of the model. The baseline system to be simulated was outlined, and the results of the optimisation were presented. A sensitivity analysis was performed in order to determine the effects of various physical and stochastic parameters on the balance between system component sizes and cost. The results were compared to those produced by Chang & van Zyl (2010). The method followed allowed for the completion of all objectives mentioned in 1.3.

5.2. Primary findings

The pertinent findings from the various application of the developed model are presented in the following sections.

5.2.1. Verification of the model

The comparison between the results of the base model against the model developed by Chang & van Zyl (2012) - the Chang model, confirmed that the base model had retained its stochastic integrity, by comparing a single chromosome analysed by both the Chang and base models. In addition, it was confirmed that the expanded model had retained its integrity as a genetic optimisation model by the closeness of the resulting 50th generation set to that produced by the Chang model.

It was found that systems consisting of multiple pipes in parallel were inefficient, and as such were crowded out of the solution space by more efficient, single-pipe systems. This observation is in line with the observation made by Chang & van Zyl (2012), with respect to gravity-fed systems.

5.2.2. Base model

The model was able achieve the primary goal of providing solutions across the spectrum of reliability (interpreted as the inverse of failure frequency) for bulk water supply systems making use of pumps, where power is drawn from the national, electrical grid. In addition, the model produced solutions within the acceptable design criterion of 1 failure per annum

under seasonal peak demand, with costing governed by a cost model describing the cost of each system over a design life of 50 years. The results obtained from the base model were indicative of the complex relationship between pump, pipe, reservoir size and cost for the standard system exposed to environmental conditions (component failures and fire events). The effect of supply system reliability on component sizing at reliability levels around the desired design criterion failure frequency was shown to be highly significant in comparison to the original, gravity-fed model. The cost gradient of supply failure frequency for pumped systems was shown to be steep in comparison to gravity-fed systems, owing to the influence of power grid reliability and the compound effect of replacement cost over the total system life-cycle.

The results obtained produced relatively few solutions in the acceptable range, owing to the crowding distance function of the NSGA-II optimisation model. This sparsely populated zone of solutions surrounding the desired design criterion can be overcome through focusing the model to produce more reliable solutions, by limiting the value of the decision variables (system components). This method was found to be effective, but in order to demonstrate the effect of various sensitivities over the full failure frequency spectrum; it was done only as an example. The focusing of solutions should only be attempted when the system is well defined (implementation stage, known constraints) and the desired reliability level is known, as the risk is that the efficiency and spectrum of solutions can be affected. The genetic optimisation process was able to converge to a pareto-optimal front within 10% of the ultimate front, within 15 Generations. The following generations can be considered as refining generations and allow the influence of the crowding distance to ensure that the solution front is relatively uniform.

5.2.3. Comparison to the Chang model and CSIR (2000) guidelines

The developed, base model produced solutions, around the desired design criterion of 1 failure every 10 years under seasonal peak conditions, with lower supply ratio and higher reservoir capacity compared to the Chang model. This was shown to be partially owing to the restriction in supply ratio experienced by the Chang model, and partially owing to the efficiencies involved in the trade-off between supply ratio and reservoir capacity. The restriction of the minimum feasible supply ratio of the Chang model to 1.34 shifted the balance of the optimal, design criterion solution toward a lower capacity reservoir, in order to

remain at the desired failure frequency level and stay pareto-optimal. The base model, owing to the additional influence of power failures, necessitated a larger reservoir capacity to buffer against the long power failure (and by extension pump failure) durations experienced. The base model's design criterion solution system consisted of a supply ratio of 1.13 and reservoir capacity of 29.6 hours of seasonal peak demand storage. This is in contrast to the findings of Chang & van Zyl (2012) that showed the design criterion solution to consist of a supply ratio of 1.34 and reservoir capacity of 14.54 hours of seasonal peak demand. This shows that it is possible to optimise a pumping system subjected to power failures and obtain results that are within the range of typically found reservoir capacities and supply ratios.

The base, Nel (2009) and Chang & van Zyl (2012) model results were compared to the CSIR (2000) / Red Book design stipulation. The finding was that the stipulations are relatively efficient, with reference to the difference in reservoir size and supply ratio, from the optimally produced solutions for the base and Nel (2009) models. The stipulated supply ratio, however, was found to be largely insufficient under seasonal peak conditions, taking into account the high sensitivity to supply ratio in the range of supply ratios between 1.0 and 1.2.

5.2.4. Sensitivity analysis

The wide spectrum of parameters varied in the sensitivity analysis, demonstrated the high sensitivity to physical variations (static head and pipe length), and the relative insensitivity of the system to variation of the majority of stochastic parameters. Pump failure rate and duration had significant influence on system cost, while fire and pipe failure parameters were seen to be largely insensitive. This was determined to be owing to evolutionary response of the system to pump failure events, resulting in a system resilient to long duration failures. As such, the short duration of pipe and fire events had relatively little impact on system failure frequency.

Two unrelated stochastic power-supply failure models were tested, being the NERC (2014) based model (the base model) and Nel (2009) model, relating to power failure distributions for North America and South Africa, respectively. The results of these models were compared and found to be surprisingly similar, producing solutions with an optimal cost within 10% of each other and equally spaced from the CSIR (2000)/Red Book design

solution. This was not anticipated as the failure mechanisms vary significantly. The NERC (2014) based model produced infrequent, long duration failures, while the Nel (2009) based model produced frequent, short duration failures. This observation as well as the relative insensitivity of the model to variation of the stochastic model parameters, is testament to the ability of the optimisation model to minimise both cost and failure frequency within a wide spectrum of conditions.

Owing to current power supply conditions experienced in South Africa, the effect of load-shedding on service reservoir reliability was also investigated, using the Nel (2009) stochastic model for baseline failure events (not related to load shedding). When comparing to this baseline, it was found that for the City of Cape Town (2015) stage 1 load-shedding schedule, there is an increase in failure frequency by a factor of 10 from 1 failure every 10 years to 1 failure per year. Stage 2 load shedding reduces the reliability (increases the failure frequency) of optimally designed reservoirs from 1 failure every 10 years to 10 failures per year. Stage 3B reduces reliability to unsustainable levels, resulting in complete system failure.

This may not be manifested as often in reality as the implementation of staged load shedding is not constant and the application of this observation is to optimally designed systems. Should systems be designed according to deterministic guidelines or with substantial safety factors, the effect is likely to be less severe. However, should load shedding be implemented continuously at Stage 2 or 3B, it is likely that reservoir failures and a reduction in potable water reticulation service levels will occur. This result emphasizes the need to accurately define the environmental condition and context before designing optimally.

5.2.5. Contributions to research

The developed, base model was shown to be stable and retained its integrity as a stochastic and optimisation model after altering and making additions to the Chang model. An external, hydraulic solver (MATLAB) was integrated, allowing for the model to be applied widely and with a large array of supply ratios. The expanded model was able to produce usable results when integrated with the Nel (2009)/Nel & Haarhoff (2011), and NERC (2014) and USADOE (2014) based power failure stochastic models. This constraint showed that following the guidelines and designing a system with exactly 1.5 times AADD supply ratio and 48 hours AADD reservoir storage would produce a system with impractical failure frequency. It was further found that limiting the model to the two minimum guideline conditions, mentioned above, produced systems that ranged from 100 failures per annum, to 1 failure every 100 years, based on very small adjustments to the supply ratio.

From this it is recommended firstly that deterministic design guidelines should preferably be used in conjunction with a stochastic, optimisation model, to ensure that the respective sensitivity to the design environment is well defined, particular for pumped systems reliant on national energy production. Secondly, if a stochastic process cannot be followed, that a rule of thumb approach be to design to a peak supply ratio of 2.3 times AADD and 52 hours AADD reservoir capacity, as opposed to the 1.5 times AADD and 48 hours AADD stipulated by the Red Book (CSIR,2000). This would ensure resilience against a wide spectrum of possible system conditions (power failures, pipe failures and demand side events), excluding load shedding. These figures are based on variations to the specific bulk water system used during the investigation, which is affected by environmental conditions, and should not be applied generically. When taking into account load shedding, the supply ratio should be increased accordingly. This approach will not produce optimal results and ideally should be only in the absence of access to a stochastic, optimisation modeling approach.

The results from the typically used deterministic guidelines (CSIR, 2000/Red book) produce results that can vary greatly. However, when the base solution reservoir capacity and supply ratio, at 1 failure every 10 years under seasonal peak conditions, is compared to the guideline stipulations, the difference is reasonably small. With careful consideration of supply ratio, the CSIR (2000) guidelines can produce solutions that are sufficiently reliable and reasonably efficient. . This gives the model applicability as it allows the designer to stay within his/her

local legislative standards and still provide an optimal system. All of these factors contribute toward the advancement of the model and the optimal reliability based-design of bulk water supply systems, including pumped flow.

5.3. Recommendations for future research

Although a substantial advancement of the optimisation model was made in the inclusion of pumping systems and life-cycle costing, there is still a significant opportunity to improve both the developed model and the approach to the optimisation of bulk water supply systems.

5.3.1. Development of the current model

There are many ways in which the developed model can be improved. One of the more technical ways is to increase the computational efficiency of the coding itself. The code was written and implemented on MATLAB, which as described previously, is computationally inefficient due to the number of overheads and the MATLAB to C++ link that is called every time the hydraulic engine is called. This would allow for shorter run times and increase the feasibility of the use of the model on less powerful computers. This was not included in this investigation as computational efficiency was not one of the objectives, as it was for Chang & van Zyl (2010).

The model can further be developed to include variable network topology. The topology used in the current model was limited to three parallel pipes with a maximum of 3 interconnections and varying pipe diameters. This was done to find the most efficient pipe system layout. However, the model may be used by a designer at a more advanced stage of design, where the pipe system has already been dictated by the supply side restraints. There is also the possibility for multiple reservoirs and pumping installations, which is hydraulically possible in the current model, but has not been coded into the stochastic models. The stochastic models used for the pipe failure can be updated to include further environmental factors such as pipe material and soil conditions that could help to increase the accuracy of the stochastic models, which further increases the holistic accuracy and applicability of the model.

The pumping system stochastic model was limited to consideration of the critical case of power failures under the assumption that a mechanical failure of the backup pump during maintenance of the operating pump (or vice-versa) is highly improbable. This model can be expanded to include the effects of mechanical failure of the pump mechanism; however this is not seen to be as critical as a power failure.

5.3.2. Research into the optimisation approach

The overall structure of the model could be developed to incorporate scheduled upgrades to the system. This would allow for the amount of wasted capacity to be minimized as the traditional long term design horizon would be replaced by multiple, incremental design horizons, staged to follow sequential upgrades.

6. References

Barta, B. &Rowse, N. 1998. Capital cost optimisation of pumping and reservoir system design. *WRC Report No 757/2/98*. Pretoria: Water Research Commission.

City of Cape Town, Areas 1 - 16 Monthly Load shedding Schedule https://www.capetown.gov.za/en/electricity/LS2015/EG2014_02_Schedule.pdf

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

Chang, C & van Zyl, J.E. 2012. Comparison of a bulk water supply optimisation technique to deterministic design guidelines

Chang, C & van Zyl, J.E. 2012. Optimal reliability-based design of bulk water supply systems - MSc(Eng) Dissertation, University of Cape Town.

Chang, C & van Zyl, J.E. 2014. Optimal reliability-based design of bulk water supply systems, *Journal of Water Resources Planning and Management (2014.140:32-39)*

Cullinane, M. J., Lansey, K. E. & Mays, L. W. 1992.Optimisation-availability-based design of water-distribution networks. Journal of Hydraulic Engineering, 118(3): 420-441.

CSIR. 2000. *Guidelines for human settlement planning and management: Volume 2.* Pretoria:CSIR.

Deb, K. & Agrawal, R. B. 1995.Simulated binary crossover for continuous search space,*Complex Systems*, 9: 115-148.

Deb, K. Pratap, A., Agarwal, S. &Meyarivan, T. 2002. A Fast and Elitist Multi objective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182-197.

Eliades. D. 2009. EPANET MATLAB Toolkit.http://www.mathworks.com/matlabcentral /fileexchange/25100-epanet-matlab-toolkit.

Eskom 2007. Annual Report - Directors' report: ensuring reliable electricity supply. http://www.eskom.co.za

Farmani, R., Walters, G. A. &Savic, D. A. 2005. Trade-off between total cost and reliability for Any town water distribution network. *Journal of Water Resources Planning and Management,* 131(2): 161-171.

Fredericks M, Khan F (2007), Report on plant reliability for Rand Water for 2006/2007, *Internal Rand Water report,* 9th July 2007, Rand Water Head Office, South Africa.

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.

Goldberg, D. E. 1989. *Genetic algorithms in search, optimisation and machine learning*. Reading, Massachusetts: Addison-Wesley Publishing Co.

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.

Holland, J. H. 1975. *Adaptation in natural and artificial systems*. Cambridge, Massachusetts: MIT Press.

Huber, T. 1997. Introduction to Monte Carlo Simulation For Summer 1997 Envision-It! Workshop. http://phsics.gac.edu/ ~huber/envision/instruct.MonteCar.html

Jonkergouw.P. 2007. Initial development of a EPANET MATLAB Toolkit. http://www.mathworks.com/matlabcentra/fileexchange/25100-epanet-matlab-toolkit

Karatzas, I. 1988. A tutorial introduction to stochastic analysis and its applications. *Department of Statistics Columbia University New York,* N.Y. 10027.

Khomsi, D., Walters, G. A., & Thorley, A. R. 1996.Reliability Tester for Water-Distribution Networks. *Journal of Computing in Civil Engineering.*10(1).10-19.

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.

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.

Mbula, V 2008. Rand Water pump stations external trip data. Maximo data sheets. Johannesburg: Rand Water Head Office.

Nel, D. T. 1993. *Bepaling van die optimal stoorkapasiteit van twee Johannesburg se Munisipale Diens reservoirs*. Johannesburg: Rand Afrikaans University. (M. Eng dissertation).

Nel, D.T & Haarhoff,J. 1996. Sizing municipal water storage tanks with Monte Carlo simulation. *Aqua*, 45(4): 203-212.

Nel, D T and Haarhoff, J. 2011, *Bulk Water Distribution Power Supply Failures. J. S. Afr. Inst. Civ. Eng.* (vol.53, n.1), pp. 55-60. ISSN 1021-2019.

Nel, D. T. 2009. *Factors that may compromise bulk water distribution reliability*. Johannesburg: University of Johannesburg. (PhD Thesis).

North American Electric Reliability Corporation, 2014. 2014 Long-Term Reliability Assessment, http://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/2014LTRA_ERATTA.pdf

Ostfeld, A., Kogan, D. & Shamir, U. 2001.Reliability simulation of water distribution systems – single and multi quality. *Urban Water*, 4: 53-61.

Ott, L. & Longnecker, M. 2008.*An Introduction to Statistical Methods and Data Analysis*. 6th ed. Belmont, CA: Brooks/Cole Cengage Learning.

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.

Quimpo. R. G. & Shamsi, U.M. 1987. Network analysis for water supply reliability determination. In Raga, 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.

Ramakrishna, S V & Gupta, S A 2009. Reliability of power supply. http://www.powermin.gov.in/distribution/apdrpbestprac/CEA-reliabiliy.pdf

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.

Seshadri, A. 2009 .*NSGA-II: A multi-objective optimisation algorithm*. [Online].Available:http://www.mathworks.com/matlabcentral/fileexchange/10429-nsga-ii-a-multiobjective- optimisation-algorithm.

Snedecor, G. W. & Cochran, W. G. 1989.*Statistical Methods*. 8th ed. Iowa: Blackwell Publishing.

Swamee, P. K. & Sharma, A. K. 2008.*Design of water supply pipe networks*. New Jersey: John Wiley & Sons, Inc.

The World Bank, 2014. [Online] Databank - Annual Population Growth Statistics, data.worldbank.org/indicator/SP.POP.GROW

United States Department of Energy, 2014, Electric Disturbance Events (OE-417) https://www.oe.netl.doe.gov/OE417_annual_summary.aspx

Van Zyl, J.E. &Haarhoff, J. (2007) Reliability analysis of municipal storage reservoirs using stochastic analysis. *Journal of the South African Institute of Civil Engineering*, 49(3), 27-32.

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.

VSA &Consultburo. 1996. Costing and planning for bulk water infrastructure. *VSA & Consultburo manual*.

Walski, T. M. 2001. The wrong paradigm – Why water distribution optimisation doesn'twork. *Journal of Water Resources Planning and Management*, 127(4), 203-205.

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.

Zio, E. 2009.*Computational methods for reliability and risk analysis*. Singapore: World Scientific