SIZING OF BULK WATER SUPPLY SYSTEMS WITH A PROBABILISTIC METHOD

J Haarhoff • JE van Zyl

WRC Report No. 985/1/02

Water Research Commission

SIZING OF BULK WATER SUPPLY SYSTEMS WITH A PROBABILISTIC METHOD

by

J. HAARHOFF AND J.E. VAN ZYL

Department of Civil and Urban Engineering
Rand Afrikaans University
Johannesburg

Final Report to the Water Research Commission on the Project

"K5/985 - Development of a stochastic technique for the optimisation of pipe and reservoir systems"

Project Leader	:	Professor J. Haarhoff
WRC Report No	:	985/1/02
ISBN No	:	1 86845 837 7

This study emanated in two outputs:

1. This report, entitled "Development of a Stochastic Technique for the Optimisation of Pipe and Reservoir Systems" WRC 985/1/02

2. The associated software package, Mocasim. Mocasim is exclusively available for download from the WRC website at http://www.wrc.org.za/wrcsoftware/mocasim.htm. It is recommended that the report and software be used together.

ACKNOWLEDGMENTS

The research in this report emanated from a project funded by the Water Research Commission and entitled :

"K5/985 - THE DEVELOPMENT OF A STOCHASTIC TECHNIQUE FOR THE OPTIMISATION OF PIPE AND RESERVOIR SYSTEMS "

The Steering Committee responsible for this project, consisted of the following persons :

Mr JN Bhagwan	Water Research Commission (Chairman)
Mr DS van der Merwe	Water Research Commission
Mrs CM Smit	Water Research Commission (Secretary)
Prof D Stephenson	University of the Witwatersrand
Mr NJ Manson	University of the Witwatersrand
Mr A Paton	Umgeni Water
Mr R Hill	Department of Water Affairs and Forestry
Ms MK Milstein	Department of Water Affairs and Forestry
Mr E van Huyssteen	City Council of Pretoria
Mr JJ van der Walt	Magalies Water
Mr BJ de Klerk	Bloem Water
Mr A Cross	Johannesburg Metropolitan Council

The financing of the project by the Water Research Commission and the contribution of the members of the Steering Committee are acknowledged.

This project was possible due to the co-operation of many individuals and institutions. The authors therefore wish to record their gratitude to the following :

Johannesburg Metropolitan Council	Water consumption data
Magalies Water	Water consumption data
City Council of Pretoria	Water consumption data
City Council of Windhoek	Water consumption data

EXECUTIVE SUMMARY

Background and motivation

The demography and water resources of South Africa are such that the vast majority of its inhabitants live a substantial distance away from the closest water source. This necessitates a vast, expensive and often complex network of *bulk water supply systems* (consisting mainly of pipes and storage tanks) to transport water from its source to the immediate neighbourhood of the consumers. With the current national emphasis on bringing water to the previous unserved communities (often those which are furthest away from existing sources), huge capital investments are being made towards new additions to the bulk supply systems. It is of obvious importance to ensure that the systems are optimally designed to provide acceptable service at an affordable cost.

Current design standards for bulk water supply systems do not allow much design flexibility. Firstly, they generally do not allow the designer to differentiate meaningfully between urban and rural systems, and secondly they do not allow the designer the freedom to assume different levels of reliability. A small rural community, which had been limping along for many years with their own rudimentary water supply system, may opt for a cheaper system, albeit at lower reliability, since they have an alternative during the short periods of non-supply. Likewise, a large industry which can ill afford to shut down due to a water shortage, may opt for higher reliability, even if at higher cost.

For these reasons, a methodology was deemed necessary to allow the designer to couple *reliability* with *system capacity*. Whereas similar methods are commonplace in many other fields of civil engineering, e.g. hydrology, no such tools are generally available to designers of bulk water supply systems. This report aims to establish a solid theoretical framework for such a *probabilistic* method, and to stimulate the use and further refinement of this approach in engineering practice.

Results and conclusions

Having developed a comprehensive framework for the probabilistic analysis of water supply systems, the first part of the analysis dealt with the sensitivity of system reliability with respect to the different input parameters. This sensitivity analysis brought forward a number of useful points:

1. The provision of water for fire fighting, even if the fire demand is grossly exaggerated, does not warrant the provision of additional storage as is suggested by current design guidelines.
2. The system reliability is greatly dependent on a parameter X, which was defined during the course of this study as the maximum weekly demand divided by the supply capacity between the source and the reservoir. As this ratio approaches unity, system reliability rapidly drops.

3. The daily demand peaks have a negligible effect, but the hourly demand peaks are somewhat more important.
4. The stochastic demand parameters (serial correlation and white noise) have important effects on system reliability.
5. The pipe break frequency has a large effect, but the actual repair times and the standard deviation thereof much less so.

The second part of the analysis applied the method to three different reservoir configurations, all commonly used in Southern Africa. The three case studies presented demonstrated how probabilistic analysis can bring a fresh insight into the behaviour of water supply systems. In the cases presented, the following was shown:

6. Estimates were derived for a typical urban bulk water supply system. The number of supply interruptions are typically between 0.1 and 1 interruptions per year, and the total annual interruption duration ranges between 1 and 10 hours per year.
7. The approach can be extended to a typical branched configuration as found in many cities, i.e. where water is supplied to a primary reservoir, from where it is further distributed to secondary reservoirs. The analysis of the Windhoek reservoir system, as an example, showed that about 8 hours of the AADD of the combined secondary reservoirs should be allocated to the primary tank if the primary tank has only a "transfer" function.
8. The analysis of the Mabeskraal system showed how the method is applied to a long, linear "backbone" system with small side reservoirs. In such systems, tank sizes should increase with distance from the source. Generally, the tanks can be significantly reduced in size (to between 12h and 24h of AADD) without a compromise in supply security.
9. The Mabeskraal analysis also pointed to some important differences between urban and rural systems. The demand peaks are much more attenuated, due to on-site storage and the use of standpipes. There is also stronger serial correlation between consecutive days than in urban areas. Pipe breaks are less frequent in rural areas, and repair times generally longer.

Research needs

The research conducted for this project pointed to the following areas for further research and development:

10. The interrelationship between reservoir size, supply capacity and weekly demand peaks was evident. This study introduced a parameter X (the maximum weekly demand divided by the supply capacity) which turned out to be of critical importance for system reliability. This specific issue should be investigated further, both from a theoretical perspective and by collecting data from a number of cases where failures had been reported.
11. The stage is now set for the more complex analysis of systems where reservoirs are used in multiple stages, i.e. where primary reservoirs lead to secondary and tertiary reservoirs. When such concatenated systems are

encountered, the *operational* issues turn out to be crucial. Often, the individual reservoirs in such systems are being controlled by different authorities. These systems, because of their practical relevance, deserve closer scrutiny from a theoretical perspective to derive optimal operational strategies.

12. The important differences between urban and rural systems have been clearly indicated. Traditional design guidelines for urban systems are not very relevant to rural systems. Further research should be directed at the development of rational, probabilistic design guidelines to be used for rural systems.

MOCASIM software package

This study generated a stand-alone stochastic analysis software package called MOCASIM. This package is available for download from the WRC web site http://www.wrc.org.za/software.

TABLE OF CONTENTS

GLOSSARY

Continuous variable. A random variable X is said to be *continuous* if it can take an infinite number of possible values associated with intervals of real numbers, and there is a function $f(x)$, called the *probability density function*, such that (Scheaffer & McClave, 1986):

- $f(x) \geq 0$, for all x.
- $\int_{-\infty}^{\infty} f(x)dx = 1$.
- $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$.

Deterministic model. A model that can predict a specific outcome for an experiment to an acceptable level of accuracy.

Discrete variable. A random variable X is said to be *discrete* if it can take on only a finite number, or a countable infinity, of possible values x. In this case (Scheaffer & McClave, 1986):

- $P(X = x) = p(x) \geq 0$.
- $\sum_{x} P(X = x) = 1$, where the sum is over all possible values of x.

The function $p(x)$ is called the *probability function* of X.

Distribution function. The distribution function $F(b)$ for a random variable X is defined as $F(b) = P(X \leq b)$. If X is discrete, $F(b) = \sum_{x=-\infty}^{b} p(x)$, where $p(x)$ is the probability function. If X is continuous, $F(b) = \int_{-\infty}^{b} f(x)dx$, where $f(x)$ is the probability density function (Scheaffer & McClave, 1986).

Experiment. The process of making an observation (Scheaffer & McClave, 1986).

Expected value. The *expected value* of a discrete random variable X having probability distribution $p(x)$ is given by $E(X) = \mu = \sum_{x} x\, p(x)$. (The sum over all values of x for which $p(x) > 0$.) (Scheaffer & McClave, 1986).

The *expected value* of a continuous random variable X having probability density function $f(x)$ is given by $E(X) = \int_{-\infty}^{\infty} x f(x)dx$.

Fail. A system or system element is said to fail when it ceases to perform its intended function.

Fail frequency. The failure rate, f, is defined as the number of failures in a time period divided by the total operational and repair time.

Failure rate. The failure rate, μ, is defined as the number of failures in a time period divided by the total operational time.

Failure state. The state in which a system or a system element is partially or fully in a condition of failure.

Operational state. The state in which a system or system element operates normally.

Probabilistic model. A model normally arising from statistical investigations, characterised by the fact that although specific outcomes of an experiment cannot be predicted with certainty, relative frequencies for various possible outcomes are predictable (Scheaffer & McClave, 1986).

Probability. Suppose that an experiment has associated with it a (finite) sample space S. A probability is a numerically valued function that assigns to every event A in S a real number $P(A)$ so that the following axioms hold (Scheaffer & McClave, 1986):

- $P(A) \geq 0$.
- $P(S) = 1$.
- If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$.

Probability distribution. A mathematical, table, graphical or other representation of the probability- or probability density function.

Reliability. The probability that a system or system element will not fail under some specified set of circumstances.

Repair rate. The repair rate, η, is defined as the number of repairs in a time period divided by the total repair time.

Risk. The potential (probability) for the realisation of unwanted consequences from impending events (Bowles, 1987).

Standard deviation. The square root of the *variance*, or $\sigma = \sqrt{\sigma^2}$ (Scheaffer & McClave, 1986).

State of a stochastic process. A state in a stochastic process is a description of the demands in terms of flow rate at all demand nodes in the system, the operational condition of all elements in the system, and a description of the level of supply in terms of flow rate in all links in the system (Duan & Mays, 1990).

Stochastic process. A stochastic process is a collection of random variables $X(t)$ referred to as the state of the process at time t (Duan and Mays, 1990).

Variance. The *variance* of a random variable X with *expected value* μ is given by $V(X) = \sigma^2 = E(X - \mu)^2$ (Scheaffer & McClave, 1986).

LIST OF SYMBOLS

Symbol	Description	Units
E	Expected value	varies
f	Failure frequency	s
N	Number of	-
p, P	Probability	-
t	Time	s
T	Time to failure	s
x, X	An event	-
V	Variance	varies
μ	Expected value	varies
μ	Failure rate	s^{-1}
η	Repair rate	s^{-1}
σ	Standard deviation	varies

CHAPTER 1 - INTRODUCTION

1.1 Background and Motivation

The demography and water resources of South Africa are such that the vast majority of its inhabitants live a substantial distance away from the closest water source. This necessitates a vast, expensive and often complex network of *bulk water supply systems* (consisting mainly of pipes and storage tanks) to transport water from its source to the immediate neighbourhood of the consumers. With the current national emphasis on bringing water to the previous unserved communities (often those which are furthest away from existing sources), huge capital investments are being made towards new additions to the bulk supply systems. Ensuring that the systems are optimally designed to provide acceptable service at an affordable cost, is of obvious importance.

Current design standards for bulk water supply systems do not allow much design flexibility. Firstly, they generally do not allow the designer to differentiate meaningfully between urban and rural systems, and secondly they do not allow the designer the freedom to assume different levels of reliability. A small rural community, which had been limping along for many years with their own rudimentary water supply system, may opt for a cheaper system, albeit at lower reliability, since they have an alternative during the short periods of non-supply. Likewise, a large industry which can ill afford to shut down due to a water shortage, may opt for higher reliability, even if at higher cost.

For these reasons, a methodology was deemed necessary to allow the designer to couple *reliability* with *system capacity*. Whereas similar methods are commonplace in many other fields of civil engineering, e.g. hydrology, no such tools are generally available to designers of bulk water supply systems. This report aims to establish a solid theoretical framework for such a *probabilistic* method, and to stimulate the use and further refinement of this approach in engineering practice.

It should be emphasised that this report deals with *bulk supply systems* only, which only covers the system from the source to the bulk storage tank. It excludes the *distribution network*, which is the part of the system from the outlet of the bulk storage tank to the consumer itself. The reliability of the network is another matter altogether which needs to be estimated in other ways, and is therefore excluded from this project.

1.2 Previous Work

Research into the probabilistic analysis of bulk water supply systems started a decade ago at the Water Research Group of the Rand Afrikaans University. The initial work of Danie Nel (1993) established the conceptual framework for estimating system reliability and demonstrated the feasibility of the method for urban supply systems. That study also provided a first set of estimated input parameters for the probabilistic analysis. It was later summarised and published by Nel and Haarhoff

(1996). Since then, related research has been ongoing, mainly to improve and verify the estimates of the numerous input parameters. The work of Kobus van Zyl (1993) provided stochastic information on fire probability, fire flow and fire duration, based on an analysis of 11 years of fire flow data in Johannesburg. The work of Angela van der Mey (1995) provided stochastic information on the probability of pipe breaks and repair times, based on an analysis of more than 3000 pipe breaks over three years in Johannesburg. The work of Lize Fourie (1996) provided information on the serial correlation of residential daily water demand, based on an analysis of five different supply areas in Johannesburg. Cobus Booysen (1995) made some early improvements to the simulation software, and Bernard Groenewald (1995) automated some procedures for extracting statistical parameters from actual water demand data.

This earlier work provided a solid platform for the work done for this project. Many more data sets were captured and analysed and the simulation software was completely redesigned, rewritten and significantly improved, allowing more sophisticated analyses than before. In addition, this report is also, for the first time, a comprehensive synthesis and summary of all the progress made at the Rand Afrikaans University in the area of the probabilistic analysis of bulk water supply systems.

1.3 Project Aims

The original research proposal for this project set out the following aims:

- To compile a comprehensive report on the state of knowledge regarding stochastic analyses of water supply systems and storage tank optimisation techniques.
- To develop an algorithm to use stochastic analysis techniques for the optimal design of pipe and storage tank systems. It is specifically aimed at optimising the bulk pipeline(s) from the water source to the storage tank.
- To test and verify the algorithm.
- To make the technique accessible to engineers as computer software and through publications.

1.4 Structure of the Report

Chapter 2 contains the literature study. Although there has been no coherent framework published for this type of analysis, a number of publications were found which contained parts which were relevant to the project.

In Chapter 3, the theoretical framework with its associated mathematics and statistics is developed.

Chapter 4 uses the method to demonstrate the sensitivity of system reliability to various input parameters. Most of the required input parameters are not generally used by civil engineers and the sensitivity analysis is deemed to be important to develop an intuitive feel for these parameters. Different model parameters are changed in turn to illustrate their individual effects on overall system reliability.

Chapter 5 deals with three typical case studies to illustrate how the method was applied in practice, and what was learned from these applications.

Chapter 6 summarises the main findings and shortcomings of this study, and also suggests some avenues for further research.

In addition, two software programs had to be developed. The first software application MOCASIM is used to perform a Monte Carlo-type simulation once the model parameters have been determined for a particular application. The second software application PAREXTRA was used to perform the extremely tedious analyses of historical data from water meters. While PAREXTRA will not be available as part of this project (it being continuously improved), its procedures for extracting trends, periodicity, peak factors, serial correlation and white noise from noisy and incomplete real-world data, may be useful to others working in this area.

Annexure A therefore contains the background to the Monte Carlo Simulation (MOCASIM) software. The software can be downloaded from the WRC web site http://www.wrc.org.za/software.

Annexure B summarises the procedures inherent to the Parameter Extraction (PAREXTRA) software which was used for extracting the required model parameters from actual data sets.

CHAPTER 2 - LITERATURE STUDY

2.1 Introduction

The literature survey shows that probabilistic modelling is an established technique in a wide range of fields in engineering and science. It is frequently used in the analysis of complex systems where risk and uncertainty play important roles, for instance nuclear power plants, environmental impact assessments and in industry.

In the water engineering field, probabilistic analysis is a popular technique used in the study of the hydrology, water quality of natural drainage systems and the design and operation of hydraulic structures. From the literature surveyed, however, it appears as though the application of probabilistic analysis in water supply systems, as proposed in this report, is unique. Whilst there are no other studies which set out a comprehensive framework, many studies were found which illuminated at least some individual aspects of the proposed method.

2.2 Terminology and Concepts

It is essential for any application of probability theory to have clear and unambiguous definitions for the terms and concepts used. The most important concepts used in this study are first enumerated within the context of water supply system analysis of bulk water supply systems.

Reliability: In the broadest sense, reliability is associated with dependability, successful operation, and the absence of breakdowns or failures. For quantitative engineering analysis, however, reliability is more stringently defined as the probability that a system will perform its intended function for a specified period of time under a given set of conditions. A system or system element is said to fail when it ceases to perform its intended function (Lewis, 1996). Duan & Mays (1990) defined reliability as the probability of remaining in an operational state as a function of the time given that the system started in the operational state at time $t = 0$.

Availability: Duan & Mays (1990) defined availability as the probability of being found in an operational state as a function of the time given that the system started in the operational state at time $t = 0$. The fundamental difference between reliability and availability is that the former requires that the system be in the operational state continuously, while the latter does not. Availability considers the operational-failure-repair process of repairable systems, while reliability only considers the operational-failure process.

Important Note: For bulk water supply systems, there is a significant difference between reliability and availability. Reliability would be a measure of a system in perfect working order, while availability would be a measure of consumers having water available in the storage tank. If the supply pipe to a tank would break, but be repaired before the water in the tank runs out, the reliability would be affected, but the availability not.

Risk: The concept of risk involves two aspects, namely uncertainty and some kind of loss or damage that are associated with an event (Kaplan & Garrick, 1981; Bowles, 1987). Symbolically, it may be written as:

$$\text{risk} = \text{uncertainty} + \text{damage}$$

Acceptable Risk: The definition for the notion of "acceptable risk" has been debated for a very long time. Kaplan & Garrick (1981) pointed out that there are two difficulties with the notion of acceptable risk:

- The first difficulty is that the notion implies that risk is linearly comparable. In other words, one can state that the risk of one scenario is greater than the risk of another scenario. This is often not the case. One scenario may have a smaller probability than a second scenario, but a greater loss associated with it. A good way to represent these scenarios would be to draw a graph showing frequency as a function of effect. There are ways of reducing these functions to a single scalar number, but a lot of information is lost in the process, which may make the numbers nonsensical as basis for certain decisions.
- The second difficulty with the notion of acceptable risk is that risk cannot be seen in isolation, but should be viewed in combination with the cost and benefits associated with it. Considered in isolation, no risk is acceptable, but if the returns on taking the risk are high enough, it may be worth taking.

2.3 Probability Distributions

In stochastic analysis it is often necessary to assign a probability of occurrence to different possible values of a variable. We may, for instance, know from experience that the minimum time required to repair a burst on a particular pipe is 3 hours, the maximum 36 hours and the median 12 hours. By fitting a suitable function to these values, a probability distribution function may be defined which can then be used in a stochastic analysis to simulate the pipe's failure duration behaviour.

Various probability distribution functions have been developed and are available in standard statistics text books. The functions most commonly used in stochastic analysis of water supply systems are summarised below.

2.3.1 Discrete Probability Distributions

The formal definitions and properties of the discrete probability distributions discussed are listed in Table 2.1.

Bernoulli Distribution: In a Bernoulli distribution, the outcome of an experiment has one of only two possible states, e.g. "success" or "failure". An example of a Bernoulli distribution is a pipe with a failure rate of 5 bursts/km/year. In any period, say an hour, the pipe will either be available (without bursts) or unavailable (having a burst).

The probability that a 10 km long pipe with this failure probability distribution will fail in a one hour period can thus be calculated as 0.57 %, using the formula in Table 2.1.

Binomial Distribution: If an experiment consisting of n identical, but independent trials is conducted, with each trial outcome following a Bernoulli distribution with Y the number of successes in the n trials, Y is said to possess a binomial distribution.

An example of a binomial distribution is a system consisting of two reservoirs interconnected by five independent parallel pipes, each 10 km long. We may apply the binomial distribution to calculate the probability that one, two or more of these pipes will fail simultaneously. To be able to apply the binomial theory the system must, however, meet the following conditions:

- the pipes must be identical.
- each pipe can only have one of two possible states, i.e. without a burst ("successful") or with a burst ("failure").
- the probability of failure is identical for each pipe, say 5 bursts/km/year or 0.57 % in any one hour.
- the pipes are independent of each other, meaning that if one pipe fails, this would not affect the other pipes' probability of failure.

The probability that, say, two pipes in the above case will fail in any one hour may now be calculated using the binomial theory to be 0.032 %, using the formula in Table 2.1.

Poisson Distribution: There are events which occur at random points in time or space, where we are only interested in the occurrence of the event, and not its non-occurrences. We may, for instance, want to know the number of pipe failures in a given time interval, but there is no sense in investigating the number of pipes not failing. To apply the Poisson distribution, the following conditions must be met:

- the probability of events occurring must be exactly the same
- events must be independent
- there must exist a constant and positive λ, so that the probability of exactly one event occurring in interval t is approximately proportional to the interval, that is λt.

Table 2.1 Properties of some discrete probability distributions

Name	p(x) =	x =	for p	Expected value (μ)	Variance (σ^2)
Bernoulli	$p^x(1-p)^{1-x}$	1	$0 \le p \le 1$	p	$p(1-p)$
Binomial	$\binom{n}{y} p^y (1-p)^{n-y}$	0,1,....,n	$0 \le p \le 1$	np	$np(1-p)$
Poisson	$e^{-\mu}\frac{(\mu)^x}{x!}$	0,1,...	$0 \le p \le 1$	λt	λt

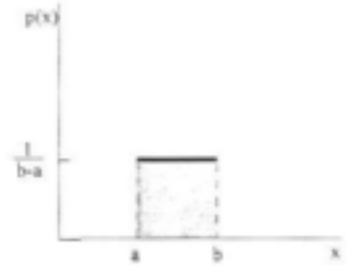


Figure 2.1 The continuous uniform distribution

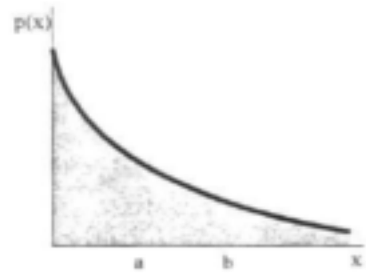


Figure 2.2 The continuous exponential distribution

2.3.2 Continuous Probability Distributions

The formal definition and properties of the continuous probability distributions discussed are summarised in Table 2.2.

Uniform Distribution: If the probability of an event x is constant over a given interval from a to b, then it is has a uniform distribution. See Figure 2.1.

Exponential Distribution: The exponential probability distribution may be thought of as being generated by a Poisson process. If, for instance, the occurrences of pipe failure events follow a Poisson distribution, then the distribution of the length of time between successive arrivals follow an exponential distribution. See Figure 2.2.

Normal distribution: The most frequently used probability distribution is the normal distribution. It may be presented in standard form, in which the mean is zero and variance (and therefore the standard deviation) is unity. Other forms may be constructed from the standard form. See Figure 2.3.

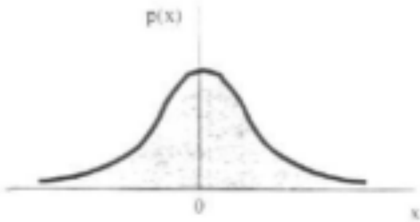


Figure 2.3 The continuous normal distribution

Table 2.2 Properties of some continuous probability distributions

Name	$p(x)$ =	x =	for p	Expected value (μ)	Variance (σ^2)
Uniform	$\frac{1}{b-a}$	$a < x < b$	$0 \leq p \leq 1$	$\frac{a+b}{2}$	$\frac{(a-b)^2}{12}$
Exponential	$\mu e^{-\mu x}$	$0 \leq x \leq \infty$	$0 \leq p \leq 1$	$\frac{1}{\mu}$	$\frac{1}{\mu^2}$
Normal (standard form)	$\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$	$\infty \leq x \leq \infty$	$0 \leq p \leq 1$	0	1

2.4 Monte Carlo Simulation

The performance of a typical engineering system depends on a combination of external or input parameters. If the input parameters are constant (gravitational acceleration and dead weight of structures are examples from structural engineering), the performance of the system will also be constant. More often than not, however, the input parameters are not constant but variable (live and wind loadings are examples). The performance of such a system will itself be a variable, depending of the combination of input parameters.

Monte Carlo simulation entails the repeated calculation of the system performance, each time with a different combination of input parameters. The input parameters are randomly selected from probability distribution functions that describe the variability of each input parameter. Each calculated system performance is compared against some performance criterion: if the performance criterion is satisfied, that event is recorded as successful. If not, the event is recorded as a failure. After a large number of iterations, the probability of system failure is estimated as the number of failures divided by the total number of iterations.

Monte Carlo simulation is a well established tool for the probabilistic evaluation of engineering systems. Examples include structural analysis, hydrology, scour at bridge piers and counter-current air stripping. It is also ideally suited for the specific problem at hand, namely the probabilistic evaluation of bulk water supply systems. The input parameters are those stochastic variables which affect the system; consumer demand, fire-fighting demand and emergency storage requirements. Once estimates of the probability distributions of these components are available, Monte Carlo simulation may be used to generate stochastic combinations of these variables for any required number of iterations, each iteration covering a time step of arbitrary length.

The performance criterion used for bulk water supply systems is whether the amount of water in storage at the beginning of a time step, plus the inflow during the step, is more than the water leaving the tank. If so, the supply to the customers is not interrupted and the system is successful. If not, it implies that the tank will be empty for at least a part of the time step and a failure is recorded. The process is repeated for each successive time step, with the tank level adjusted for the nett in/outflow after each step.

The question may be asked: "Why not use some of the more conventional, analytical methods for the analysis of this problem?" Such methods are desirable for their ability to provide exact solutions to the problems they are applied to. These analytical methods, however, are often associated with severe restrictions on the complexity of the problems they can solve. Text book problems and solutions often look deceptively simple as they are deliberately chosen to make calculations amenable to relatively simple algebraic manipulations (Chou, 1963).

Beim & Hobbs (1988) investigated the use of various analytical methods in the field of water supply system analysis. These methods were shown to have better computational efficiency, but lack the detail and accuracy of Monte Carlo or stochastic models. This is due to the many assumptions required in the application of analytical models. When models become complex and realistic, the level of mathematics required increases so that they may not be amenable to the application of analytical techniques at all, even by competent mathematicians. In such cases, a simulation approach becomes the best tactic for solving the problem.

Monte Carlo analysis does have some disadvantages. It requires fairly sophisticated computer resources and cannot produce precise estimates of reliability without long runs (Beim & Hobbs, 1988). These restrictions, however, are of less significance today due to the powerful computational capacities of modern microcomputer systems.

Various methods may be used in conjunction with Monte Carlo methods, including Bayesian analysis (Bernier, 1987), Queuing theory (Chou, 1964), the Modified Frequency-Duration model and Markov Chain Approach model (Beim & Hobbs, 1988). Bayesian analysis, for example, is often applied to the study of risk situations induced by geophysical extreme events such as earthquakes, waves and storm surges (Bernier, 1987).

2.5 Probabilistic Analysis in Related Areas

The underlying method of probabilistic analysis applied in this study is well established and a great number of publications are available on its application in related fields. A few of these publications are reviewed to show how the method is applied and the type of results that may be extracted from such analyses.

2.5.1 Study 1: Flood Frequency Analysis (Schultz, 1987)

One of the oldest and simplest stochastic hydrological techniques is found in flood frequency analysis. The application of this method usually results in information on the relationship between peak discharge values and corresponding probabilities of exceedence. The study may also include benefit/cost ratio estimation for different scenarios.

Schultz presented an example of such a study in which a water authority in Germany (Ennepe Water Authority) had to decide whether or not an existing multi-purpose dam should be raised. The purposes of the dam were drinking water supply, flood protection and low stream flow augmentation downstream of the dam. These purposes were all fulfilled at different levels of reliability which could all be improved by raising the dam wall. Decision criteria, however, also included economic considerations. For this reason the benefits of proposed schemes were compared with the respective costs on the basis of a benefit/cost ratio.

Considering the benefits relating to flood damage costs, the benefits of raising the dam wall may be calculated as the difference in damage cost for the raised wall case and present conditions. The valley downstream of the Ennepe dam is highly industrialised and flood damage costs were thus linked to the peak flood levels. These levels, however, represent a stochastic process, thus rendering damage cost stochastic. It was thus only possible to compute expected values of damage costs on the basis of the probability distributions of peak floods. The project lifetime was assumed to be 50 years and the present value of the reduced flood damage cost could thus be considered.

Hydraulic data for the stream flow into the dam made it possible to construct probability functions of reservoir inflow peaks and representative unit inflow hydrographs. These floods were then routed through both the present dam and a raised dam using standard flood routing procedures. Historic flood damage figures were used to estimate the flood damage due to different sizes of floods and the present value of benefits due to raising of the dam wall could thus be calculated and used as rational basis for decision making.

2.5.2 Study 2: Reliability of a Water Resource System (Schultz, 1987)

A stochastic technique that is often used in water resources systems is queuing theory. The author discussed an example of a water supply project in Ghana. The aim of the study was to investigate the reliability of the extraction from a dam for the projected future demand scenario. The existing dam was feared to be too small for the growing water demand load on it. The task was to decide whether or not the water demand could be met by raising the existing dam or whether a second dam needed to be built. Water demand is a function of population size and was thus expected to increase in the long term. Seasonal demand variations were also taken into consideration.

This type of problem is usually solved with the aid of simulation techniques: If the dam's capacity and control rules are known, stochastic analysis may be used to determine its probability of failure. Queuing theory, however, allowed a more direct calculation of the probability of failure of the reservoir and the probabilities of its storage contents, given the reservoir inflow probability distribution and demand functions. This technique may be applied if the operating rule can be expressed as a function of the storage content of the dam. This would typically consist of the release from the dam as a function of its content. A probability transition matrix is formulated and by simultaneous solution of the corresponding systems of equations, the probability distributions of the reservoir contents and dam releases are calculated. This results in a reliability distribution for the dam which may be used to make planning decisions. Queuing theory is, however, unable to take serial correlation of inflow into account.

For the Ghana dam analysis, queuing theory was applied on the basis of monthly time increments and a suitable discretisation of reservoir capacity. Computations were carried out for various reservoir capacities at different times in the planning horizon. The results of the analysis showed that although the present dam was adequate for current demand, it would have to be raised for future scenarios.

2.5.3 Study 3: Reliability of a Water Distribution System (Yang *et al*, 1996)

The authors applied stochastic simulation to estimate the reliability of a water distribution network. They used an optimisation model as the simulation process. An optimisation model is advantageous over a simulation model in that it converges to a solution, thus avoiding the repetition of trial and error. The reliability of the network was judged by a reliability index called *performance reliability* which is an indicator of the ability of a network to meet its demands. Performance failure was defined as a failure of a water distribution network to provide a satisfactory level of service at specified critical locations in the system.

Demand and source supply were modelled deterministically, while system component failures were modelled using appropriate failure probability distributions. A large number of possible "crippled" network configurations were generated and simulated over a sufficiently long period to capture variations in supplies and demands. The system reliability was then calculated from the results. Pipe failures were assumed to be independent and a simple exponential failure probability distribution model and constant repair time were used. For each link, a sequence of states (fail/not fail) was generated. A system scenario was then generated by combining the sequences for all the links in the system. A sufficiently large number of system scenarios were generated to provide a statistically significant reliability value.

The authors applied this methodology to investigate the regional water distribution network of Metropolitan Water District of Southern California. The system was modelled using a prototype of 214 links, four source nodes, nine surface water reservoir nodes, 18 ground-water recharge basin nodes, two spill nodes and 38 demand nodes. The planning period was 25 years and the simulation done in monthly intervals. Only failures of major links were considered for the analysis of which only 12 pipes were of significance (with normal failure durations longer than a month) with a mean time to failure of 20 years and a constant repair time of six months. Useful deductions could then be made for use in future planning. For example, if it is required that the shortage in supply should not be exceeded 20 % of demand for at least 88 % of the time, then the probability of achieving this would be (say) 95 %.

2.5.4 Study 4: Reliability of a Water Distribution System (Wagner *et al*, 1988)

The authors determined the reliability of water distribution systems subject to pipe and pump failures by using stochastic analysis. The simulation program used consisted of two parts. The first part generated failure and repair events according to specified probability distributions and the second part calculated the flow rates and pressures in the system.

Simulations were performed over an extended time period. When a link failed, it was removed from the system. The new pressures the demand nodes in the reduced network were then calculated leaving the demands unchanged. A set of demand nodes were selected for the reliability analysis and simulated as pressure-dependent demands. Once a link failed, a random repair time was generated before it was returned to full operation. Pipes were assumed to fail independently. They used exponential functions for pipe and pump failure times, a uniform pipe failure duration function from 3 to 72 hours, and a log normal distribution for pump repair duration. Most failure time and repair time data had been estimated according to what the authors saw as reasonable, as data for this was not readily available. Pipe failure rates were based on published figures.

2.5.5 Study 5: Reliability Analysis of a Pumping System (Damelin *et al*, 1972)

Damelin *et al* (1972) conducted a Monte Carlo analysis on the Nahal Oren bulk water supply system near Haifa, Israel, consisting of eight pumping wells, a booster pumping station with two pumps, 10 km length of pipeline and a reservoir with 4 Ml capacity.

The water demand was modelled deterministically and included seasonal, weekly and hourly variations. The pumping equipment was modelled with an exponential distribution function for the time periods between failures. The repair duration was modelled using a log normal distribution with a mean repair time varying between 50 and 52 hours and a distribution with 4 % of repair durations less than one hour, and 1 % greater than 700 hours. Pipe, water source and electricity failures were not considered.

The authors found that a simulation period of 20 years was sufficient to generate a representative sample of annual reliability factors. The sensitivity of the model to a number of input parameters was determined and is shown in Table 2.3.

Table 2.3 Results of seven Monte Carlo analyses on the Nahal Oren System, Haifa, Israel (Damelin et al, 1972)

No	Cumulative changes in input parameters from description in text	Full simulation		Worst year in simulation			
		No of failures	Ave failure duration (h/year)	No of failures	Ave failure duration (h/year)	Longest failure (h)	Second longest failure (h)
1	None	29	9.9	30	19.3	17	16
2	Mean pump fail duration (51 h) increased to 67 h. Max 4 % of fail durations (700 h) increased to 1000 h	18	7.0	28	19.8	83	43
3	Another well and pump similar to Pump 2 added	3	0.6	7	2.5	13	12
4	Reservoir capacity increased by 50 % to 6 Ml	6	1.8	20	10.2	17	17
5	6 hour lag added to repair period	7	2.7	20	9.7	14	14
6	12 hour lag added to repair period	13	4.9	21	11.9	32	15
7	Demand pattern increased by 10 %	33	15.4	60	34.9	20	17

2.6 Probabilistic Description of Pipe Reliability

Two measures are required for a comprehensive description of pipe reliability, namely the *probability* of failure and the typical *duration* of the supply interruption. Pipe failures occur relatively infrequently (Clark *et al*, 1982) and records, if they are kept at all, are rarely published. When they are published, comparison with other similar studies is normally difficult or even impossible.

2.6.1 Probability of Supply Interruption

Researchers agree that various independent factors may be responsible for pipe failures, including (Clark *et al*, 1982; O'Day, 1983; Andreou & Marks, 1986):

- Quality and material strength of pipe and fittings
- Environment: corrosiveness of soil, frost, heaving soils, external loads
- Construction practice and quality of workmanship
- Service conditions; notably internal working pressure and water hammer
- Length of time facility has been in operation
- Design practice when built
- Degree of land development

Various studies on pipe failure behaviour are available in the literature. These studies are complicated by the large number of factors influencing failures, and the limited amount of data available for analysis. As a result the conclusions of the studies vary significantly and are, in some cases, contradictory:

- Goulter & Coals (1986) reported that it is questionable whether pipe failures in distribution systems are independent events. While corrosion is often the root cause of breaks, they found that the events that trigger breaks are very cold weather, water hammer, expansive soils, earthquakes, and construction activity. Such events normally impact on large portions of a distribution system simultaneously, triggering clusters of failures.
- Good quality pipes are often more expensive than the poorer quality alternatives, although this is seldom a primary concern in the design of pipe systems. Andreou & Marks (1986) found that pipes installed during different time periods experienced significant differences in break-rates. Pipes installed in the most recent times (50's and later) mostly performed worse than older pipes.
- The probability of a pipe failing will increase once the pipe has failed for the first time. Clark *et al* (1982) did a study on pipe failures on a large (6275 km pipes) and a small (580 km of pipes) water supply system in the USA, both consisting mainly of cast iron pipes. They found that over a period of 40 years 52 % of the pipes had no failure events at all. A minority of the pipes were responsible for a majority of the maintenance events. They also found that after a pipe has failed, the time interval between successive failures became increasingly shorter.

- Goulter & Coals (1986) found that both spatial and temporal clustering were significant features of the breakage pattern. They found that more than 42 % of all failures occurred within 20 m of a previous failure of similar type. More than 13 % of all failures occurred at the very same location as a previous failure of the same type. A high fraction of these failures (65 %) occurred within one day of a previous failure, pointing to inadequate repair practices.

- Goulter *et al* (1993) developed a method to predict pipe failures, based on historic failure data and taking into consideration both temporal and spatial clustering of failures. They used the non-homogeneous Poisson distribution to model the failure rate, as this distribution can accommodate local variations in the failure rate over intervals of both time and space.
- Another study (Fitzgerald, 1968) reported on findings made on break data for cast-iron pipes in Antonio, Texas where the cause of breaks were identified and separated into corrosion and other causes. Corrosion-induced failures showed an exponential growth, while failures caused by other reasons showed a linear increase.
- Andreou & Marks (1986) found that an early second break was a very important predictor of future failures.
- Various conflicting results have been published regarding the effect of age on pipe failure rates. O'Day *et al* (1983) and Ciottoni (1983) concluded after studies on pipe failure rates of the New York and Philadelphia systems, that age is not a very important factor in pipe failure rates. On the other hand, other studies (Kettler & Goulter, 1985; Walski & Pelliccia, 1982) concluded that there is a definite trend of increasing pipe failure rate with age. Andreou & Marks (1986) concluded that failure rates initially decrease following installation, after which they start increasing with age. It seems that age is at best a poor, and at worst a misleading indication of failure rate (O'Day, 1983).
- There is general agreement among all the studies that pipe diameter is a significant factor in failure rate with an increase in failure rate with decreasing diameter. (O'Day, 1983; Ciottony, 1983; Sullivan, 1982; Kettler & Goulter, 1985).

Table 2.4 provides information on the average failure rates in various cities in the USA from a study by O'Day (1982) to give a typical range of rates for mixed diameters and mixed pipe materials within an urban environment. It is clear from the table that the rates are highly variable, with an average of 0,177 and a mean failure rate of 0,072/km/year.

It is much more meaningful if one could separate the failures according to some of the most important parameters. Such an analysis is shown in Table 2.5, where some international values are compared with the South African values compiled by the RAU Water Research Group during the past ten years.

Table 2.4 Published water main break rates (O'Day, 1982)

Location	Period studied	Pipe failure rate (#/km/year)
Boston	1969-78	0.027
Chicago	1973	0.033
Cincinnati	1969-78	0.13
Denver	1973	0.1
Houston	1973	0.8
Indianapolis	1969-78	0.067
Los Angeles	1973-74	0.027
Louisville, Ky	1964-76	0.077
Milwaukee	1973	0.145
New Orleans	1969-78	0.809
New York City	1976	0.047
San Francisco	1973	0.067
St. Louis	1973	0.048
Washington, DC	1969-78	0.11

Author's Note: Variations in reporting conventions for failures by water utilities limit the usefulness of this comparison. It is likely that the differences between cities are both definitional and real.

Table 2.5 Pipe failure rates from different sources

Study	Location	Material	Diameter mm	Area	Rate #/km/yr
Kettler et al (1985)	Winnipeg	Cast Iron	150	Urban	1.06
Kettler et al (1985)	Manhattan	Cast Iron	150	Urban	0.34
Kettler et al (1985)	Philadelphia	Cast Iron	150	Urban	0.32
Kettler et al (1985)	Manhattan	Cast Iron	300	Urban	0.11
Kettler et al (1985)	Winnipeg	Cast Iron	300	Urban	0.07
Kettler et al (1985)	Philadelphia	Cast Iron	300	Urban	0.05
Haarhoff et al (1999)	Magalies Water	Fibre Cement	150-300	Rural	0.074
GLS (1995)	Stellenbosch	Fibre Cement	50	Urban	1.63
GLS (1995)	Stellenbosch	Fibre Cement	75-100	Urban	1.43
GLS (1995)	Stellenbosch	Fibre Cement	150	Urban	0.48
GLS (1995)	Stellenbosch	Fibre Cement	200	Urban	0.10
GLS (1995)	Stellenbosch	Fibre Cement	250	Urban	0.05
Haarhoff et al (1999)	Magalies Water	Polyvinylchloride	90-160	Rural	0.041
Van der Mey (1995)	Johannesburg	Steel	100-199	Urban	0.58
Drusani (1995)	Northern Italy	Steel	100	Mixed	0.32
Van der Mey (1995)	Johannesburg	Steel	200-299	Urban	0.23
Drusani (1995)	Northern Italy	Steel	150-250	Mixed	0.16
Van der Mey (1995)	Johannesburg	Steel	300-449	Urban	0.12
Van der Mey (1995)	Johannesburg	Steel	450-599	Urban	0.12
Van der Mey (1995)	Johannesburg	Steel	600+	Urban	0.10
Nel (1993)	Rand Water (70's)	Steel	600+	Mixed	0.054
Nel (1993)	Rand Water (80's)	Steel	600+	Mixed	0.042
Drusani (1995)	Northern Italy	Steel	300-600	Mixed	0.04

2.6.2 Duration of Supply Interruptions

There are three common reasons for the supply to a storage tank to be interrupted, namely the failure of the raw water source or water treatment plant, the failure of pumping installations (e.g. through power cuts), and the failure of the pipe system feeding the storage tank. In this project, only pipe failures are considered, as they are the most common cause of prolonged supply interruptions in practice.

The total interruption time associated with a pipe break is assumed to be the sum of:

- reporting time from the instant of break until the maintenance centre is notified,
- mobilisation time to assemble crew, spares, equipment and vehicles,
- travel time from maintenance centre to point of break,
- preparation time, *i.e.* shutting valves, draining pipeline, and excavation to the level of the pipe break,
- repair time, and
- commissioning time, *i.e.* disinfection, air bleed, and controlled restoration of full pipeline pressure.

It is notable that the total interruption time for rural and urban will be dominated by completely different factors. In urban situations, the reporting and travel times will be minimal, as any break will be immediately visible and the maintenance centre is usually situated close by within the urban area. One could, however, expect that the excavation and repair time in a congested area, shared by numerous other services, will have to be slow and careful. For rural areas, the total interruption time is dominated by reporting and travel times due to remote location, absence of the people, and poor roads and communications. Excavation and repair time, however, will be shorter due to unrestricted room and mostly smaller pipeline diameters.

Once the average interruption time per break is determined (which should be estimated for each individual set of conditions), it is also necessary to estimate the *variability* of the interruption time, as all breaks obviously do not take exactly the same time to repair. It is reasonable to assume that the the total interruption time will follow a log-normal distribution, as the total interruption time can obviously not be less than zero.

Two cases have been analysed by the RAU Water Research Group during previous investigations. In the first case, total interruption time was estimated for the CBD of the City of Johannesburg, based on the analysis of the maintenance records over a period of 36 months. The job cards (which indicated the time of start and end of each job) were assumed to represent the total interruption time. The reporting and mobilisation times were therefore ignored for reasons mentioned earlier. In the second case, the total interruption time was estimated for the extensive rural system supplying a number of small villages along the northern edge of the Pilanesberg in the North-West Province. It will be called the Mabeskraal system, as Mabeskraal is the largest village at the far end of the system. Here the total interruption time for all the breaks over a period of 19 months was estimated by plotting each break and calculating the travel time based on an average travelling speed. Fixed times were assumed for reporting, mobilisation, preparation, repair and commissioning. Despite the shortcomings that are evident in both cases, the results obtained are considered to be valuable pointers until more exhaustive surveys and estimates are made. The results are summarised in Table 2.6.

Table 2.6 Duration of supply interruptions

Location	Mabeskraal	Johannesburg	Johannesburg	Johannesburg
System type	rural	urban	urban	urban
Diameter	90 - 350 mm	100 - 199 mm	200 - 299 mm	450 - 449 mm
Material	Mixed	Steel	Steel	Steel
Average duration	16.7 h	3.5 h	5.1 h	6.6 h
Distribution	Normal	Log-normal	Log-normal	Log-normal
Coefficient of variation	24%	14%	14%	14%

2.7 Probabilistic Description of Pump Reliability

Although this study is limited to simple bulk systems comprising of pipes and storage tanks only, it is relevant to include a brief description of pumping system reliability, as this would eventually also be included as the method is refined in future.

A pumping station consists of one or more pumping units or systems supported by appropriate electrical, piping and structural components. A pumping unit, in turn, consists of five subsystems: (Cullinane, 1985):

- pump,
- motor,
- power supply,
- valves, and
- controls.

The reliability of each of the subsystems may be determined and multiplied to obtain the pumping component reliability. Cullinane used exponential failure distributions, a constant failure rate and a constant repair rate for the pumping unit subsystems. The actual repair duration has a lower limit below which repairs are too minor to be recorded. The upper limit of the repair time can be taken as the time required for a complete overhaul of the equipment, or the time needed to obtain and install a replacement (Damelin *et al*, 1972).

Damelin *et al* (1972) reported on the reliability of 10 pumps which are part of the Nahal Oren Project supplying the city of Haifa in Israel. The mean repair time for the pumps varied between 50 and 52 hours, with 4 % of repair durations less than one hour, and 1 % greater than 700 hours. The mean time between repair events varied more as shown in Figure 2.4. Most of the times between failures are in a band between 800 and 1 400 hours, with no relation to the pump capacity evident.

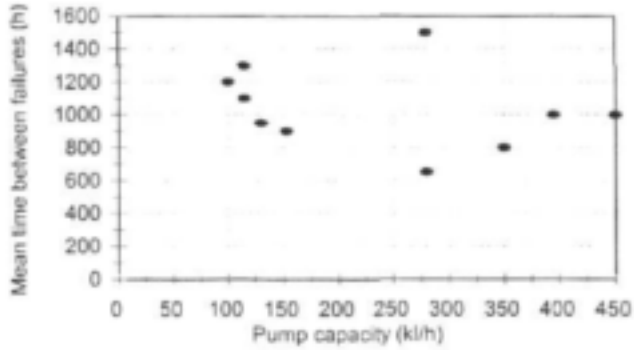


Figure 2.4 Mean time between failures for the Nahal Oren Project, Israel (Damelin et al, 1972)

Duan & Mays (1990) suggested a modified frequency duration analysis method for pumping systems which includes both mechanical and hydraulic failure, a method not reviewed here.

2.8 Probabilistic Description of Consumer Demand

A demand prediction system usually consists of a mathematical model predicting future water demand based on past measured data. The classical approach is one of time series analysis. The demand represents a physical process which posses both deterministic and probabilistic components (Brdys & Ulanicki, 1994). Various factors may influence the demand including weather, social behavioural patterns, emergencies, etc. A number of models have been suggested for modelling water demand. These include:

- Stochastic modelling with both periodic variations and trend. (Brdys & Ulanicki, 1994) The demand at a given time instance is calculated as the sum of a deterministic and probabilistic component. The deterministic component consists of a trend and periodic component, while the probabilistic component consist of auto-correlation, correlation (relationship with influencing factors such as temperature and rainfall) and a purely random component.
- Multiple regression is a method which is concerned with minimising squared error terms in an equation used to predict future demand (Billings & Agthe, 1998). The Auto Regressive Integrated Moving Average Model (ARIMA) is especially useful in modelling time series where the data reflect social, cultural, climatic, etc. behaviour (Cembrowicz, 1995; Brdys & Ulanicki, 1994).
- Exponential smoothing is a method that is simple and easy to implement, but is less accurate than the more sophisticated methods (Brdys & Ulanicki, 1994). In this method the demand data is first screened to eliminate errors and exceptionally low or high demands. A smoothing process is then applied to remove random fluctuations from the data by expanding the time series into its harmonic components using Fast-Fourier-Transform.

- Expert systems which may be applied in combination with a time series analysis to handle additional information like weather or changes in the structure of the water network (Hartley & Powell, 1991).
- Neural networks may also be applied to demand forecasting by training them on historical data. They are especially useful when the demand prediction model is non-linear. (Brdys & Ulanicki, 1994).

The consumer demand forecasting model used in this project was developed by the RAU Water Research Group and will be described in detail in Chapter 3.

2.9 Probabilistic Description of Fire Demand

2.9.1 Estimation of Total Fire Volume

To properly estimate the effect of fire-fighting on the reliability of bulk water supply systems, the total *volume* of water used for each fire is required. In practice, however, fire volumes are rarely specified. More often, the maximum fire *duration* and well as the maximum fire *flow rate* is specified. The fire volume can then be calculated as the fire *duration* times the average fire *flow rate*.

Fire codes will provide the upper limits to both flow rate and duration, from which an upper limit of fire volume may be estimated. An international review of different fire codes by Van Zyl (1993), however, showed that wide discrepancies exist amongst these codes; in terms of their underlying philosophy as well as their numerical guidelines. Table 2.7 shows a selection of such values coming from different fire codes.

To get a true *probabilistic* estimate (not just a single maximum estimate), the fire duration as well as the fire flow rate have to be described in statistical terms. In one of the very few studies done in this regard (Van Zyl, 1993), the fire flow records of Johannesburg were analysed for 12 consecutive years. From this huge data base, the "large" fire events (those using more than 5000 litres of water) were selected and subjected to frequency analysis. Figures 2.5 and 2.6 summarise the results from this analysis. From data such as this, the mean, the appropriate statistical distribution and the standard deviation can be obtained.

2.9.2 Probability of a Fire Occurring

It is difficult to estimate the probability of a fire occurring in the supply area of a water supply system, even if historical records are available. Most fires require very little or no fire water from the water supply system, and these smaller fires are obviously of no consequence for the design of the supply system. What is really needed, is the probability of a *large* fire occurring; a fire that will exert a substantial demand from the supply system. This is especially important if the fire water volume is estimated from the local fire code, as this value indicates an extreme fire event.

Table 2.7 Comparison of Fire Codes (Van Zyl, 1993)

Parameter	Germany	Netherlands	USA	South Africa
Fire flow (l/min)				
high risk	3200	6000	17700	12000
moderate risk	1600	3000	11800	6000
low risk	800	1500	3800	900
Pressure (m)				
high risk	15	20	14	15
moderate risk	15	20	14	15
low risk	15	20	14	7
Fire duration (h)				
high risk	2	2	4	6
moderate risk	2	2	3	4
low risk	2	2	2	2
Code	DVGW-W405	KIWA #50 (1977)	AWWA M31 (1989)	SABS 090 (1972)

In the case of Johannesburg, all the large fires (refer to previous paragraph) amounted to 149 fire events. The total Johannesburg supply area was divided into approximately 30 zones, each with its own storage tank(s). This indicates a historic "large fire frequency" of 0,4 fires/year for each zone.

2.10 Philosophy of Designing for Risk

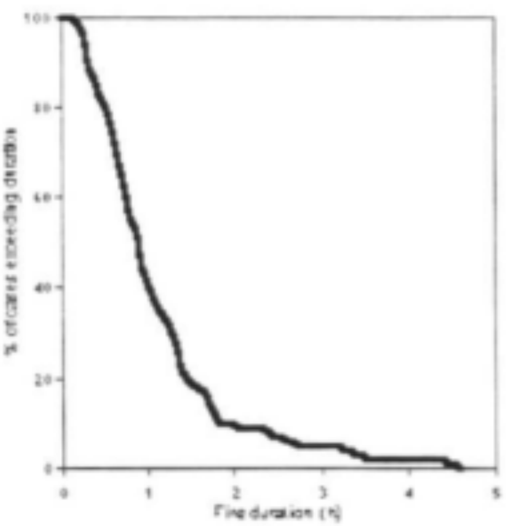


Figure 2.5 Johannesburg fire duration (Van Zyl, 1993)

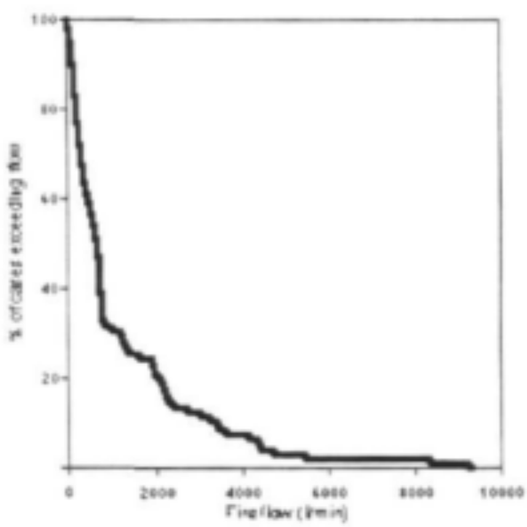


Figure 2.6 Johannesburg fire flow (Van Zyl, 1993)

The risk of failure of water supply systems is small compared to most other industrial systems; so small that its quantification had hardly been considered in the past. However, when tools become available for determining the probability of failure of a bulk water supply system, the next problem would be to specify an acceptable level of risk. In this section the classification of risk in various fields and published criteria for performance of water supply systems are discussed to provide some background.

The impact of a shortfall in supply in water supply systems will, of course, vary with the type of consumer. In municipal water supply, a minor shortfall will merely cause some inconvenience to the consumer, while a major shortfall can have disastrous results. In agriculture, where water is required for production, a shortfall may lead to a reduction in yield or even a total loss of the crop. In industrial water supply situations, a shortfall may lead to a temporary shutdown, which is not easy to evaluate economically (Damelin *et al*, 1972).

2.10.1 Example 1 : Water Infrastructure Risk Ranking and Filtering Method

One risk classification method for water infrastructure was developed by Dolezal *et al* (1994). It is called a Water Infrastructure Risk Ranking and Filtering (WIRRF) method and is a derivate of the Risk Ranking and Filtering (RRF) methodology initially developed by NASA as a decision tool to locate the twenty components that contributed most to the Shuttle Program's risk. The method is divided into a number of phases, namely 1) definition of the risk hierarchy, 2) quantification of the attributes and severities, 3) use of the telescopic filter, 4) use of the analytical hierarchy process in prioritisation and 5) application of the sensitivity analysis.

<u>Definition and Quantification of risk</u>. Four risk criteria for pipe networks were identified according to the WIRRF methodology. These risk criteria are used to generate a severity score between 1 (very low) and 5 (very high) for networks or network components.

- Prior risk information. This takes into account the previous operation of the network. Factors that characterise historical trends that affect pipe networks are identified. These include criteria that can be used to identify trends in pipe failure rates and critical areas. In this way critical areas, as well as areas of high performance, can be identified and prioritised for application of available resources. Important information in this category include the pipe age, type, number of previous breaks and trends in the data.
- Functionality. The pipe network's ability to operate in the presence of defects and anomalies, i.e. how robust the pipe network is in remaining operational. For example, a pipe network with a high degree of redundancy is considered less of a potential problem than a network with less redundancy. Important information in this category includes the level of redundancy, the number of sources and users and the part of the system that is pumped vs. gravity fed.
- Moderate-event risk characterises the most probable failure events. This may be partitioned into two parts, the likelihood of a failure event occurring and the

likely effect of the event. Important information in this category includes the number of people not served, fraction of the system affected, the economic loss of industries relying on a constant water supply and the fire loss due to the deficiency in pressure or flow available for fire fighting.

- Extreme-event risk goes further than moderate-event risk to provide the user with information on low probability / high consequence events or worse case scenarios. They have a large impact on public reaction and on the communities perception of risk. As with moderate-event risks, this category may be partitioned into two parts, the likelihood of an extreme event occurring and the likely effect of the event. Important information includes the number of people and the fraction of the system affected. Dolezal *et al* used a critical threshold above which an event is seen as critical.

Telescoping filter phase. This is an iterative procedure to retain only the systems or system components with the highest risk, or alternatively the systems or system components that require the most attention and resources. The severity scales defined and quantified in the previous phase are used as filters to compare systems with each other.

Prioritisation phase. The systems or system components retained in the previous step are now prioritised using mainly engineering judgement.

Sensitivity analysis. A sensitivity analysis may be conducted on the criteria used to determine the impact and importance of different weighting criteria.

2.10.2 Example 2 : ANCOLD Risk Classification System

Another example of a risk classification system used by the Australian National Committee on Large Dams (ANCOLD) is shown in Table 2.8 (Mackenzie, 1994). This table link the consequences of a dam failure event in terms of high, significant or low consequences in terms of loss of life, economic loss and repairs to the dam. This is linked to the probability of a flood event causing such damage.

Table 2.8 Incremental flood hazard categories, ANCOLD 1996

CONSEQUENCES	HIGH	SIGNIFICANT	LOW
Loss of life	Loss of identifiable life expected because of community or other significant developments downstream	No loss of life expected, but the possibility recognised. No urban development and no more than a small number of habitable structures downstream.	No loss of life expected
Economic loss	Excessive economic loss such as serious damage to communities, commercial or agricultural facilities, important utilities, the dam itself or other storage downstream.	Appreciable economic loss, such as damage to secondary roads, minor railways, relatively important public utilities, the dam itself or other storage downstream.	Minimal economic loss, such as farm buildings, limited damage to agricultural land, minor roads, etc.
Repairs to dam	Dam essential or services and repairs not practicable	Repairs to dam practicable or alternative sources of water / power supply available.	Repairs to dam practicable. Indirect losses not significant
DESIGN FLOOD ANNUAL EXCEEDANCE PROBABILITY	Probable maximum flood (PMF) to 1 in 10000	1 in 10 000 to 1 in 1 000	1 in 1 000 to 1 in 100

2.10.3 Reliability Performance Criteria

Kwietniewski & Roman (1997) formulated reliability requirements for water supply systems, based on considerations of

- the community's interests in general,
- the interests of the community supplied by the system, and
- interests of individual water users.

They took into consideration data on people's feelings on sanitary threats, comfort and good living conditions to identify three reliability criteria, namely fail frequency, mean repair time and proportion of time for which water is available. The recommendations of their analysis are given in Table 2.9.

Table 2.9 System reliability criteria (Kwietniewski & Roman, 1997)

No of system users	Consequence of failure (Fraction of users without water)	Max average fail frequency (failures per year)	Max mean repair time (h)	Average failure time (h/year)
> 50 000	0%	3	24	72
	30%	2	24	48
	100%	0.02	24	0.5
500 - 50 000	0%	6	24	not available
	30%	3	24	72
	100%	0.2	24	5
< 500	0%	12	24	288
	30%	6	24	143
	100%	1	24	24

The authors do not properly motivate the values suggested, except to say that it is based on "different water supply systems and the author's experience." It is, however, one of the few instances in which an author has been brave enough to suggest concrete reliability values.

Kwietniewski & Roman (1997) also conducted a survey on 100 families in an attempt to determine the acceptability of failure events. Table 2.10 gives the fraction of consumers that found such an event "tolerable".

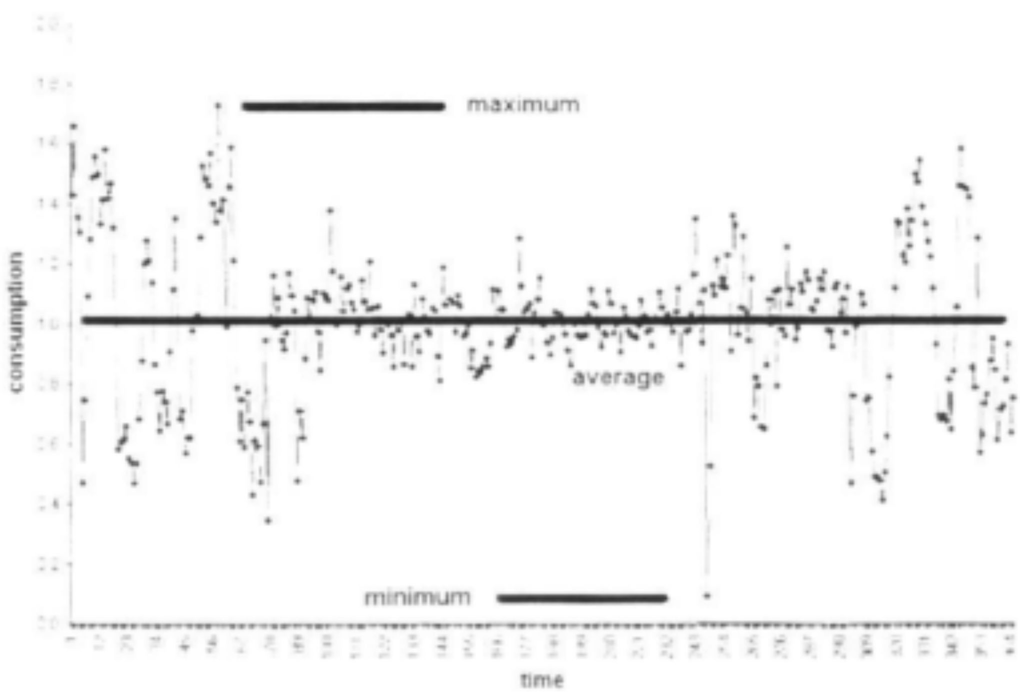
Table 2.10 Fraction of users describing different failure events "tolerable" (from Kwietniewski & Roman, 1997)

Scenario	Fail frequency (failures/year)	Failure duration (h)	Fraction of consumers finding it "tolerable"
1	1	24	74%
2	2	24	46%
3	4	12	33%
4	12	24	13%
5	12	2	55%

CHAPTER 3 - DEVELOPMENT OF A THEORETICAL FRAMEWORK FOR PROBABILISTIC ANALYSIS

3.1 Introduction

The traditional approach for the design of bulk water systems is to size the bulk supply pipeline and the storage tank independently. The bulk water supply pipeline capacity is usually sized for a constant fraction of the average daily demand rate (typically 1.5 times), and the storage tank to hold a constant fraction of the average daily demand (typically 24 to 48 hours of the average daily demand). Whilst these empirical design guidelines (which are still embodied in many design manuals) have served the industry and the public well, it is well recognized that there is an interplay between the sizes of the bulk supply pipeline and the storage tank. If a typical water demand pattern is analysed (Figure 3.1 shows such a trace), it is obvious that the storage tank size depends directly on the bulk water supply pipeline capacity. If the bulk water supply rate is more than Q_{max}, the required tank size is zero, as even the peak demand can be directly supplied from the pipeline. If the bulk water supply rate is less than Q_{avg}, then a tank of infinite volume is required, as the long-term demand outstrips the supply capacity. It intuitively follows that larger tanks can be used with smaller pipelines, and *vice versa*. This led to deterministic methods which could be used to determine the least-cost solution for the complete tank and pipeline system.



La **Figure 3.1 Typical water demand pattern** rger tanks and/or larger supply lines will obviously also reduce the probability of non-supply, an interrelationship schematically shown in Figure 3.2. Strangely, the probability of non-supply, which is the ultimate performance indicator for any water supply system, is not explicitly formulated or calculated for water supply systems. Traditional engineering, in other words, has only dealt with the two lower elements in Figure 2. The method proposed in this report could eventually eliminate this shortcoming.

In order to address the existing shortcomings, it is necessary to firstly identify those parameters which could contribute to system failure, and secondly to fingerprint the stochastic nature of these parameters. The rest of this chapter is devoted to the development of appropriate models to be used in the simulation software MOCASIM, which is described further on.

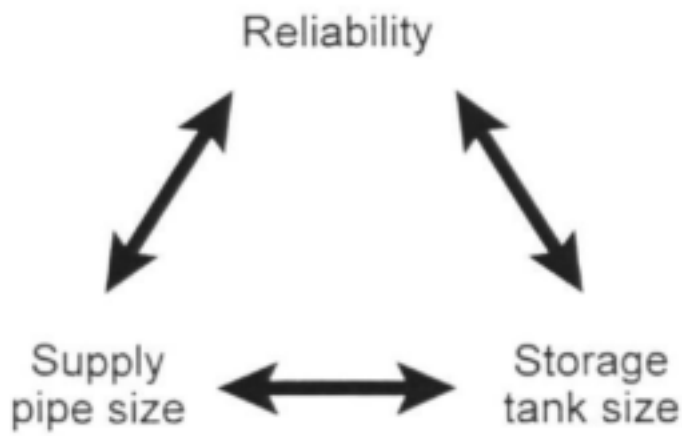


Figure 3.2 Interrelationship amongst bulk water supply variables

3.2 Storage Tank Requirements

An analysis of storage tanks requires the identification and definition of the diverse functions that have to be met. Very few storage tanks perform only one function; they have to provide more than one or all of the functions listed below:

- **Emergency storage**. Storage tank volume must be provided to guarantee the supply to the consumers, even when the supply to the tank is partially or completely discontinued. Such events may be scheduled for maintenance purposes, which is not a stochastic variable. More important for probabilistic analysis are the unscheduled events due to pipe, power or source failures. The volume required for these unscheduled events is stochastic, as neither the time of occurrence nor the duration of the interruption can be predicted.

- **Fire storage**. It is essential to have an adequate water supply available for fire-fighting. This is mostly supplied through the water reticulation system, and is thus also drawn from the storage tank. Most guidelines will specify an additional volume of water for which allowance must be made in the storage tank. This approach is based on the conservative assumption that the fire demand will coincide with a period of maximum consumer demand. The required fire storage is stochastic, as neither the time of occurrence nor the actual volume required can be predicted.

- **Demand storage**. Consumers draw water from a storage tank at a variable rate, whereas the supply line usually delivers water at a constant rate. The storage tank has to absorb the difference between inflow and outflow rates. The flow from the supply line into the tank is easily determined and usually well controlled. The consumer demand, on the other hand, is determined by the cumulative effect of a multitude of stochastic variables and is therefore itself a stochastic variable.

- **Operational requirements**. There could be additional requirements for storage tank volume, such as freeboard (dependent on the sophistication of level sensing and control equipment), bottom storage (dependent on the potential of air or sediment entrainment at the outlet), or a pump control band (required for automatic switching of pumps if water is being pumped to or from the tank). These components are all deterministic, *i.e.* they can be calculated once and simply added to the volume required for the stochastic components described above.

3.3 Model for Emergency Storage

Emergency storage is required when the supply to the storage tank is interrupted. The required volume of emergency storage for each interruption event is simply the total water demand during the time period of the supply interruption. The problem therefore reduces to two questions: a) how to estimate the *duration* of the supply interruptions, and b) how to estimate the *frequency* of the supply interruptions.

3.3.1 Frequency of Supply Interruptions

The available literature on interruptions due to pipe breakage was summarised in paragraph 2.6.1. To model the pipe breakages, it is only required to have the average pipe breakage rate. Knowing the length of pipe involved (L) and the average pipe breakage rate (b), the average number of breaks per year can be calculated (b.R). On any given day, the probability of a break is then p = (b.R)/365. It is now only necessary to generate a random number between 0 and 1 - if the number is below p, a break is assumed.

Once a break is assumed, the model will then proceed to estimate the duration of the interruption, to be described in the next paragraph.

3.3.2 Duration of Supply Interruptions

The nature of supply interruptions, as well the available literature was summarised in paragraph 2.6.2. To model the duration of these interruptions in a probabilistic way, it is necessary to have a) a statistical distribution that adequately fits the observed or anticipated interruptions, b) the mean interruption time, and c) the coefficient of variation describing the variability of the interruption time about the mean.

The literature, based on local studies, indicated that the log-normal distribution is appropriate when the interruption times are relatively short (as in urban areas), and that the normal distribution is appropriate when interruption times are longer (as in rural areas). These distributions, together with the locally measured coefficients of variation, will be used in the simulations that follow in later chapters.

3.4 Model for Fire Storage

3.4.1 Fire Duration and Fire Flow Rate

The available literature on fire durations and fire flow rates was summarised in paragraph 2.9.1. To get a true probabilistic estimate, the fire duration as well as the fire flow rate has to described in statistical terms. The literature review showed the results of such a study undertaken in Johannesburg. From data such as this, the mean, the appropriate statistical distribution and the standard deviation can be obtained.

Such data is rarely available in South Africa, and their retrieval and analysis are extremely tedious. Moreover, as large fires are rather infrequent, most of the sources will yield too few data points to allow meaningful statistical analysis. A much simpler, more direct approach is to get a deterministic estimate from the relevant fire water codes. These codes usually specify both the maximum fire flow rate and the fire duration as a function of the land use category, population density, or other parameters. By multiplying these values, an estimate can be obtained. Such values were reflected in Table 2.7.

In this project, fire water volume was estimated from a local fire code, for the following reasons:

- It is clear that actual fire water volumes measured in Johannesburg were very much less than the volumes estimated on the basis of the fire water codes. The fire water codes will thus give a conservative value.
- Earlier analyses showed that the fire demand, even if estimated by the conservative fire water codes, has almost no effect on the reliability of urban supply systems. In the case of rural supply systems, fire demand can usually be neglected due to the their low population density and lack of equipment.

If an initial analysis of a particular system would indicate that fire water does play a significant role, then it would be prudent to invest more effort in a stochastical description of the fire water parameters.

3.4.2 Probability of a Fire Occurring

It is difficult to estimate the probability of a fire occurring in the supply area of a water supply system, even if historical records would be available. Most fires require very little or no fire water from the water supply system, and these smaller fires are obviously are of no consequence for the design of the supply system. What is really needed, is the probability of a *large* fire occurring; a fire that will exert a substantial demand from the supply system. This is especially important if the fire water volume is estimated from the local fire code, as this value indicates an extreme fire event.

The very sparce literature on fire probability was earlier summarised in paragraph 2.9.2. Data from Johannesburg shows about 0.4 major fires per supply zone per year.

For modelling purposes, the fire probability will be expressed as 0, 1 or 2 fires per year for every system.

3.5 Model for Water Demand

An actual water demand pattern is the sum of two sets of factors:

- A number of cyclic factors, introduced by the seasons of the year or the days of the week.
- A random factor (termed white noise in this report), which is introduced by the cumulative effects of unaccounted-for variables such as temperature, rainfall, holidays, special events within the supply area, *etc.*

Note that the objective of the model described in this report is to generate a series of hypothetical water demand volumes which, over the long term, will have the same statistical fingerprints as an actually measured data set over a long term. The fact that an one-on-one comparison between measured and generated data will bear no correlation is not important; it is only important that the long-term cycles and data variability are similar.

3.5.1 Random Component

The estimation of a daily volume of water used starts out by selecting a random value X from a normal distribution with mean 1 and standard deviation σ:

$$X_{random} \sim N(1,\sigma)$$

3.5.2 Serial Correlation

Daily water demand data normally exhibit a strong degree of serial correlation. That means that each data point is to some extent determined by the preceding data point. In other words, if yesterday had a higher-than-average demand, today is also likely to have a higher-than-average demand. Behaviour of this kind is statistically described with a serial correlation coefficient ϕ. It is only necessary to take one preceding day into account (a lag-one model), as the analysis of numerous South African data sets have indicated that further serial correlation coefficients are not significant. The first correction (for serial correlation) is given by:

$$X_i \quad = \quad (1 - \phi_1) * X_{random} \quad + \quad \phi_1 * X_{i-1}$$

3.5.3 Systematic Factors

The second correction is due to the systematic day-to-day variation found in a typical week. The average daily volume for a particular day is calculated by multiplying by a peak factor $PF_{Mon..Sun}$:

$$= \quad PF_{Mon..Sun} * X_i$$

The third correction is due to the systematic annual seasonal variation. Some prefer to describe the variation in terms of 12 months, others in terms of 52 weeks. In this project, the seasonal pattern is described by 13 time units of exactly 28 days each. The average daily volume for each of these time units, is calculated by multiplying by an appropriate peak factor $PF_{1..13}$:

$$= \quad PF_{1..13} * PF_{Mon..Sun} * X_i$$

The fourth and final step is to dimensionalise the daily volume by multiplying with the Annual Average Daily Demand AADD:

$$= \quad AADD * PF_{1..13} * PF_{Mon..Sun} * X_i$$

The complete model is thus given by:

$$V_i \quad = \quad AADD * PF_{1..13} * PF_{Mon..Sun} * [(1 - \phi_1) * X_{random} \quad + \quad \phi_1 * X_{i-1}]$$

If the simulation is done with a time step of one day, the above model is complete. Experience has shown that a time step of one day is fairly coarse and does not provide adequate resolution at very high levels of reliability. In this project, a time step of one hour is used. In this case, the approach is to generate a daily volume as indicated above, and then to disassemble the daily volume into 24 hourly volumes according to a fixed hourly pattern. The fixed hourly pattern is defined by $PF_{1..24}$. The hourly volumes are thus obtained by:

$$H_{1..24} \quad = \quad PF_{1..24} * [V_i / 24]$$

3.6 Summary of Input Parameters

In each of the above sections, one or more model parameters were introduced. For the convenience of the reader, Table 3.1 below summarises and lists all the relevant parameters.

Table 3.1 Summary of all the model parameters listed in this chapter

CATEGORY	PARAMETER	COMMENT
Supply pipe	Capacity	The hydraulic capacity of the supply pipeline, expressed in terms of the annual average
Supply pipe	Break probability	The probability of a break in the pipe, expressed as breaks/year/km
Supply pipe	Break duration	The mean duration of a supply interruption
Supply pipe	Break duration variability	The standard deviation of the supply interruptions, expressed as a percentage of the mean
Demand pattern	Seasonal variability	The weekly or monthly demand pattern, expressed in terms of the annual average
Demand pattern	Daily variability	The day-of-the-week demand pattern, expressed in terms of the annual average
Demand pattern	Hourly variablity	The hour-of-the-day demand pattern, expressed in terms of the annual average
Demand pattern	Serial correlation	The serial lag-one coefficient for daily demand
Demand pattern	White noise	The standard deviation of the random component, expressed as a percentage of the mean
Fire	Fire probability	The probability of a fire in the supply area, expressed as fires/year
Fire	Volume required	The volume of fire water required per fire event, in this report conservatively based on fire codes

CHAPTER 4 - SENSITIVITY OF SYSTEM RELIABILITY TO SYSTEM CHARACTERISTICS

4.1 Introduction

The previous chapter indicated that a large number of parameters could influence the reliability of a water supply system. It is the purpose of this chapter to identify those parameters which have a significant influence on system reliability, and which parameters are less critical. This sensitivity analysis is not only useful for guidance towards the most profitable research areas in future, but may also be useful to readers to acquire some intuitive feel for system reliability - a concept not generally dealt with in the context of water supply systems.

The sensitivity analysis is based on a five-level discretisation of each variable. For each parameter, five ("very low", "low", "typical", "high" and "very high") values are assigned based on the experience of the project team, and on the international literature cited earlier. In some cases where very little variability is evident, only three ("low", "typical" and "high") values are assigned. The case where all parameters are "typical" is considered to be the "reference" case. All subsequent analyses are performed by only varying one parameter at a time, from the "very low" to the "very high" value.

The result of each analysis provides a curve relating three different performance indicators to the storage volume provided, of which two will be graphically shown in this chapter. The storage volume is expressed as hours of the AADD. In South Africa, the usual storage recommendations are between 36 and 48 hours of the AADD. For the analyses done here, a much broader range of storage volume is considered - from 12 hours to 96 hours.

4.2 Performance Indicators

Four performance indicators were included to assess system reliability:

- The *number* of supply interruptions per year.
- The *total number of hours* per year during which the supply was interrupted, or the *total annual interruption duration*.
- The *average duration* of a single supply interruption.

Different types of consumers will assign different weights to these indicators. Domestic consumers will mostly be irritated by the number of interruptions per year; industrial consumers with some on-site storage will only be concerned with interruptions longer duration than their own storage ability, *et cetera*.

4.3 The "Reference" Case

The ranges of variables used for the sensitivity analysis are shown in Table 4.1.

Table 4.1 Parameter values used for sensitivity analysis

Variable	Units	Very low	Typical	Very high
Feeder pipe				
Capacity of AADD (C)		1.3	1.5	1.7
Pipe break frequency	p.a.	1	5	30
Average failure duration	h	3	6	12
Failure standard deviation (log normal)		5%	14%	25%
Demand				
Weekly variability on AADD				
- Peak factor for maximum week (M)		1.2	1.5	1.8
- Peak factor for minimum week		0.9	0.7	0.5
- X = M / C		0.8	1.0	1.2
Daily variability on weekly average demand				
- Max factor		1.05	1.1	1.3
- Min factor		0.95	0.9	0.7
Hourly variability on daily average demand				
- Max factor		1.4	1.8	3.5
- Min factor		0.6	0.4	0.1
Lag-one daily serial correlation coefficient		0.1	0.4	0.7
Daily demand standard deviation		0.1	0.2	0.3
Fire				
Frequency	p.a.		0	52
Duration	h		2	4
Flow rate (% of AADD)			70%	140%

The following comments are made in relation to Table 4.1:

- As shown in Chapter 3, the pipe break probability is a product of the pipe length and the failure probability per unit length of pipe. In this sensitivity analysis, these parameters are not estimated separately, but combined into a single pipe break frequency. The "typical " value of 5 breaks per year could therefore result from a 10km pipe with failure probability of 0.5 breaks/km/year, or a 50km pipe with failure probability of 0.1 breaks/km/year, *et cetera*.
- As will become evident in the sensitivity analysis, a critical situation arises when the maximum weekly demand approaches the feeder pipe capacity. To bring this possibility into sharper focus, a secondary parameter X was defined as the ratio of the maximum weekly demand to the feeder pipe capacity, which will be discussed later on in this chapter.
- The "very high" fire frequency was taken as 1 fire per week, which is a dramatic exaggeration of what could reasonably be expected. Likewise, the fire flow rate is assumed to be an extremely high value in terms of the local fire codes.

The reference case was analysed first to get an idea of the performance indicators under "typical" conditions. These results are summarised in Table 4.2.

Table 4.2 Performance indicators for the "typical" reference case

Performance indicator	Volume = 24h AADD	Volume = 48h AADD
Number of supply interruptions	10 / year	4 / year
Total annual interruption duration	70 h / year	25 h / year
Average interruption duration	6.2 h	6.2 h

These results may indicate relatively poor performance, as supply systems in general suffer fewer interruptions and fewer interrupted hours than indicated in Table 4.2. It should be borne in mind that the case studied above is one which is utilised to the absolute limit with X = 1 (see the discussion on X under the following heading). In general, systems are expanded in discrete steps which results in inevitable under-utilisation, unlike the case analysed here. The general trends which will emerge from the sensitivity analysis, however, will remain valid as the attention will be upon the excursions relative from the "typical" reference case, and not the absolute values.

4.4 Effect of Supply Capacity and Weekly Water Demand Pattern

It is easy to see why the capacity of the supply pipe plays a pivotal role in system reliability. If the capacity is much larger than the demand, any shortage in the storage tank can be swiftly made up, but if the capacity barely exceeds the demand, shortages will take longer to be made up, leaving the system vulnerable for longer times. For this reason, the supply capacity is linked to the demand, commonly as 1,5 times the AADD.

Upon closer inspection, it becomes evident that this requirement should be defined more carefully. It is not the *average* demand that should be linked to the supply capacity, but the *maximum* demand that could be expected during the time that it typically would take to fill up the storage tank. A proposal was made (Vorster *et al*, 1995) that the supply capacity should therefore be related to the *maximum daily demand* (a supply capacity of 1,2 times maximum daily demand was proposed). In this report, we have assumed that the supply capacity should be correlated to the *maximum weekly demand*. It is thought that a storage tank of typical size will recover from nearly empty to full in about one to two days under *average* demand. When *peak* demand is experienced, the recovery time should be say two to three times longer, or about a week.

If it is accepted that the supply capacity should be correlated to the maximum weekly demand, how should it be correlated? To conceptualise this easier, we have defined a parameter X, which is the ratio of the maximum weekly demand to the supply capacity. If $X = 1$, then the supply capacity is exactly equal to the maximum weekly demand. It should be clear that a system will be subject to frequent and severe interruptions if the value of $X > 1$. Normally X should be less than 1. The "reference" case assumes a value of $X = 1$ (higher than would be expected).

Figures 4.1 shows the dramatic effect that X has on both the average number of supply interruptions per year, and Figure 4.2 the same on the total annual interruption duration. For a 36h-tank (typical in SA), both the number of interruptions as well the total annual interruption duration drop tenfold if the X-parameter is reduced from $X = 1.0$ to $X = 0.9$. If one wishes to keep the number of supply interruptions unchanged, consider the case where it is desired to keep the number of interruptions at 1 interruption per year. Figure 4.1 shows that this would require an 84h-tank if $X = 1.0$, but only a 30h-tank if $X = 0.9$. In other words, a 10% increase in supply capacity could reduce the tank volume by 64%! A similar conclusion can be drawn from Figure 4.2 if the average interruption time is limited to say 10h.

Figures 4.1 and 4.2 reinforce the previous argument that the region where $X > 1.0$ would be impractical, as the number of interruptions and interruption duration sharply jump into a completely unacceptable range.

4.5 Effect of Daily and Hourly Demand Pattern

The effects of daily and hourly demand patterns are examined next. Typical daily and hourly demand patterns were taken for the reference case. The amplitude of the peaks and dips were then increased to get higher variability, and decreased to get lower variability. Further details are shown in Table 4.1.

Figures 4.3 and 4.4 show the effects of the day-of-the-week variability. Here there is almost no difference to either the number of interruptions, or the total annual interruption duration. This indicates that supply interruptions are not primarily caused by high demands over a single day, but rather due to high demand over a longer period such as a week. This provides further motivation to link the supply capacity to the maximum weekly demand than to the maximum daily demand.

Figures 4.5 and 4.6 show the effects of the hour-of-the-day variability. Once again, the differences are not large, and the sensitivity therefore low. The number of interruptions are slightly higher when variability is higher, due to the fact that higher hourly variability causes greater fluctuations in tank level and therefore leaves the tank vulnerably empty more often. The total annual interruption duration is unchanged, as the hourly peaks are not sustained for the typical duration of effecting a pipe repair.

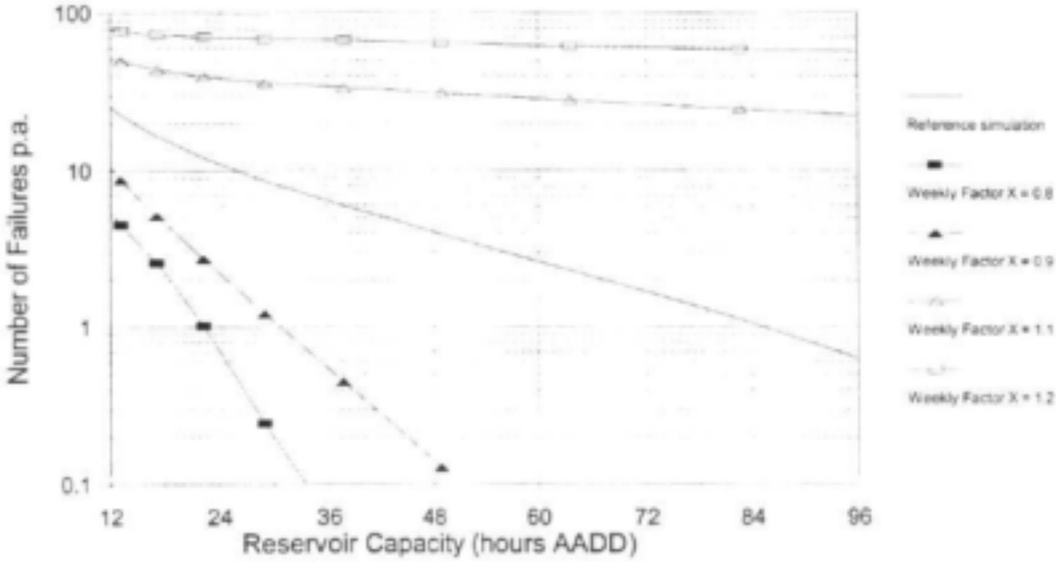


Figure 4.1 Effect of parameter X on the average number of interruptions per year.

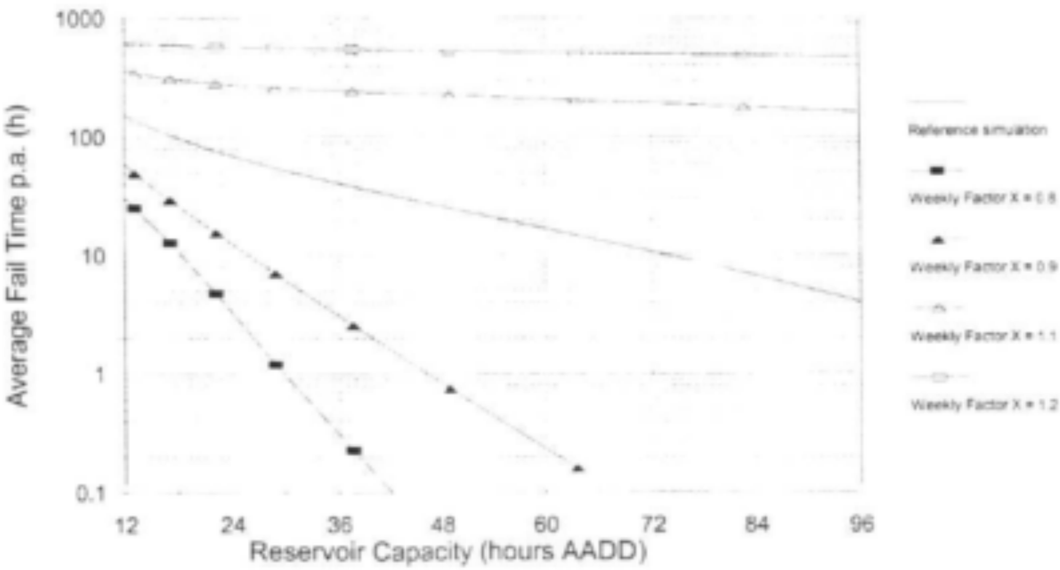


Figure 4.2 Effect of parameter X on the total annual interruption duration.

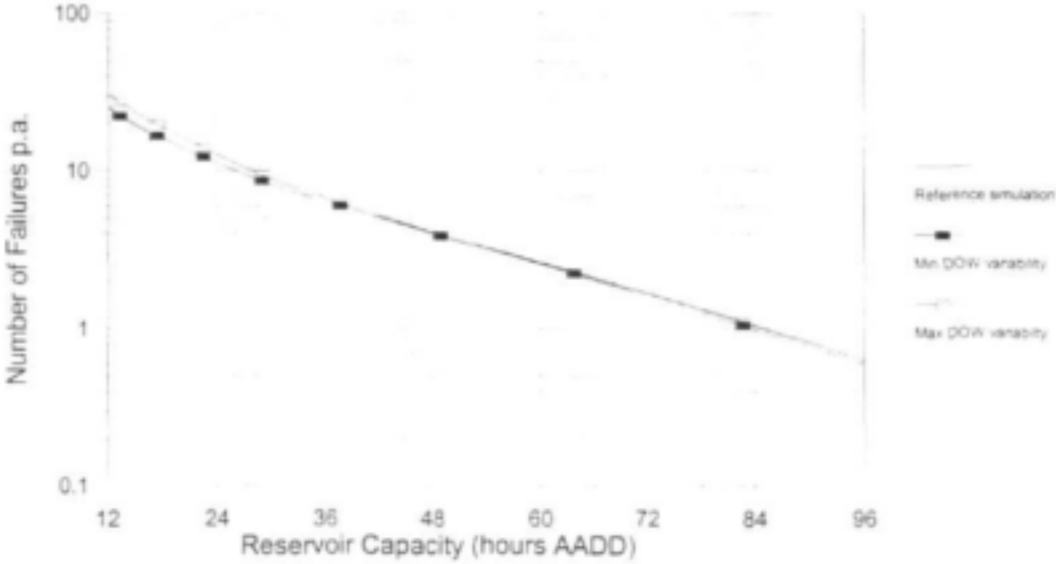


Figure 4.3 **Effect of day-of-the-week variability on the average number of interruptions per year.**

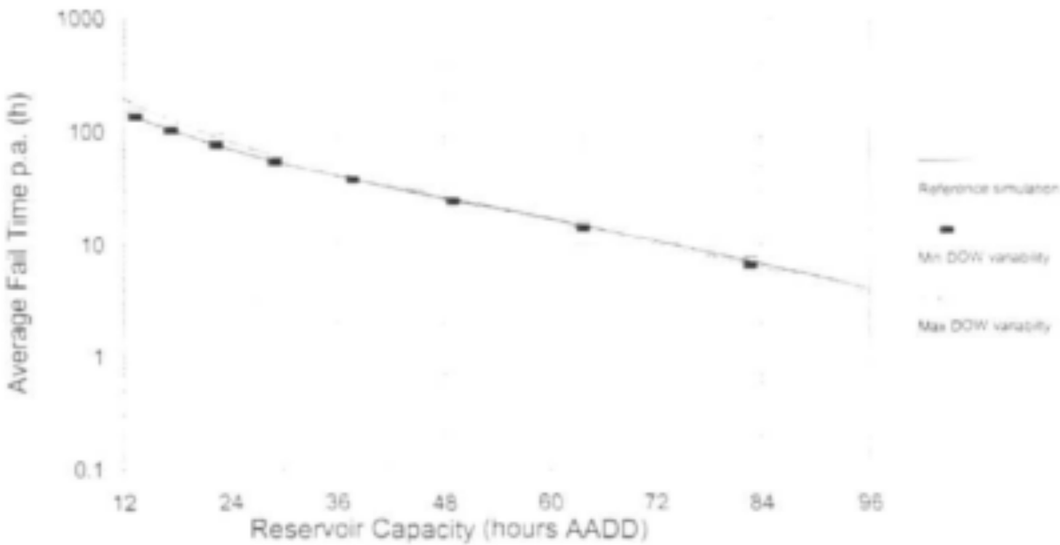


Figure 4.4 **Effect of day-of-the-week variability on the total annual interruption duration.**

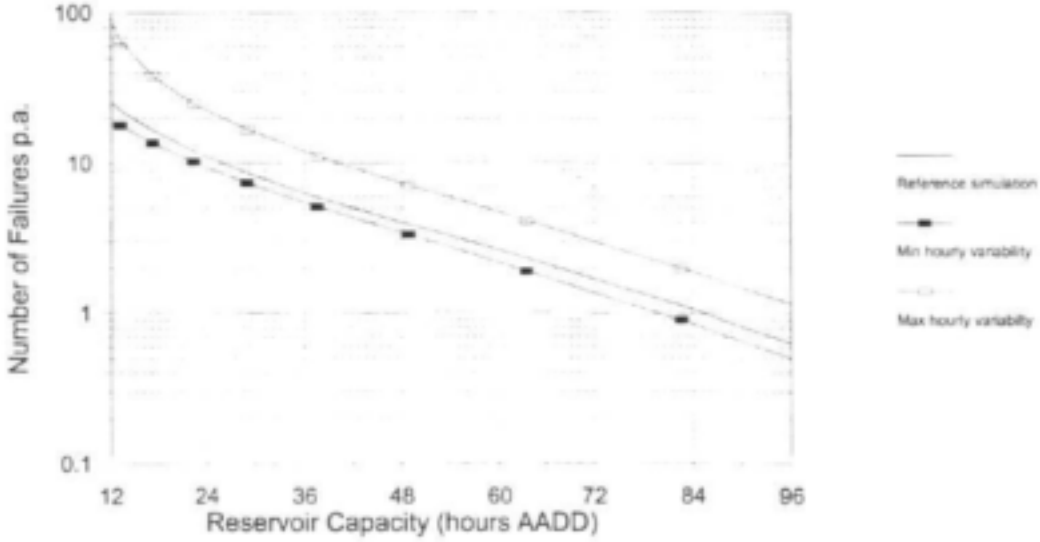


Figure 4.5 Effect of hour-of-the-day variability on the average number of interruptions per year.

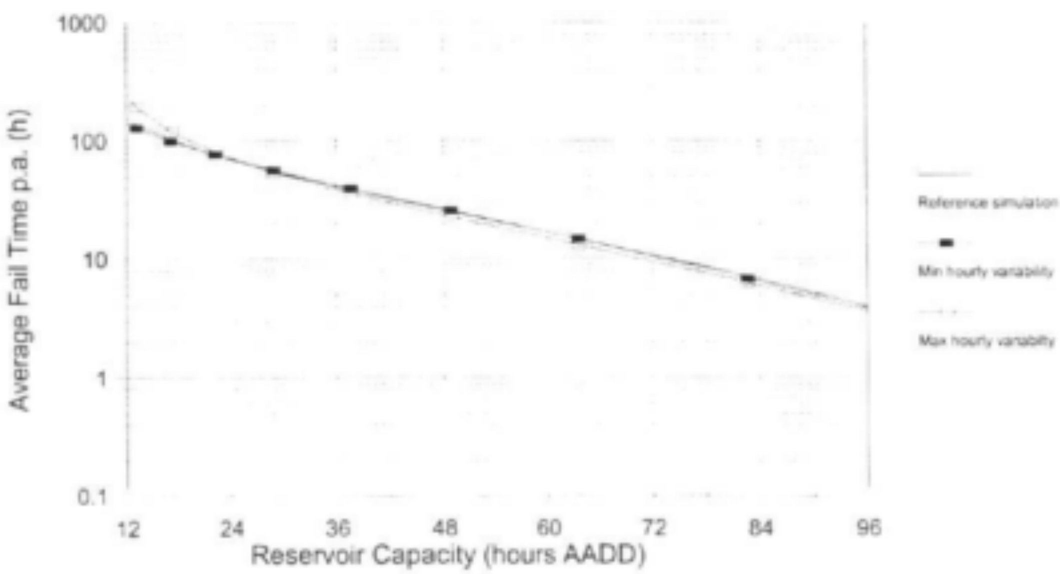


Figure 4.6 Effect of hour-of-the-day variability on the total annual interruption duration.

4.6 Effect of Serial Correlation and White Noise

The effect of the serial correlation of the water demand on consecutive days is shown in Figures 4.7 and 4.8. The range of variation is indicated in Table 4.1. When the serial correlation is high, the daily data tends to "clump" - once the daily demand is high, there is a tendency for a number of high demand days to follow. It is therefore understandable that the number of failures and the interruption duration increase with higher serial correlation.

The effect of the white noise of the water demand (the pure stochastic component of water demand after removal of serial correlation) is shown in Figures 4.9 and 4.10. The range of variation is indicated in Table 4.1. The white noise component determines the "bumpiness" of the data. With much white noise, the random variation of the demand about the deterministic pattern is large, more often leading to days with excessively high demand. The higher the white noise, the higher the number of supply interruptions and the total annual interruption duration.

4.7 Effect of Fire Flow

Previous unpublished studies at the RAU Water Research Group have shown the fire flow to have an almost imperceptible effect on system reliability. The "fire" parameters in Table 4.1 were therefore chosen to amplify the fire demand significantly above what would be reasonably expected to check whether the fire demand can be ignored even under extreme circumstances.

Once again, there is no discernible difference in any of the performance indicators between the reference case and the case with extreme fire flows. The resulting graphs are therefore not shown. This is an important finding on two counts:

- Fire flows can be safely ignored in the further calibration of the reliability model proposed in this report. This brings not only a mathematical simplification, but removes the difficult burden of collecting and collating the sparse and difficult-to-get fire water statistics, which would be required for the rigorous calibration and verification of the fire flow model proposed earlier.
- The South African fire water code calls for fire water storage over and above the storage required for balancing and emergency purposes. This analysis shows that this is overly conservative. Even at extraordinary fire water usage, the interruptions stay the same. Put in another way, the SA fire water code assumes that a serious fire will coincide exactly with a major pipe failure and high consumer demand. The analysis shows that such an event is extremely unlikely, to the point where it can be ignored. This has obvious consequences for future formulations of the local fire water code.

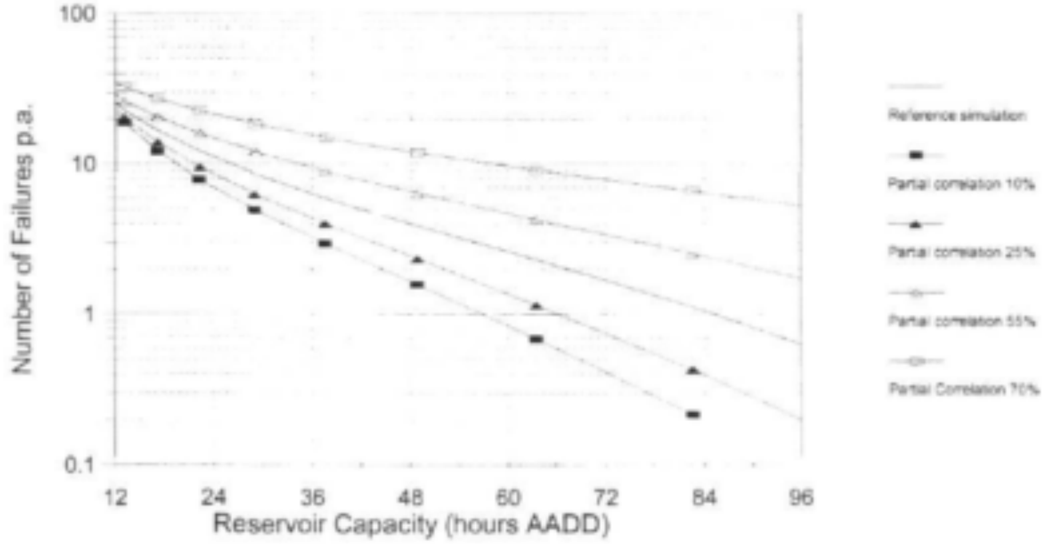


Figure 4.7 Effect of serial correlation between successive days on the average number of interruptions per year.

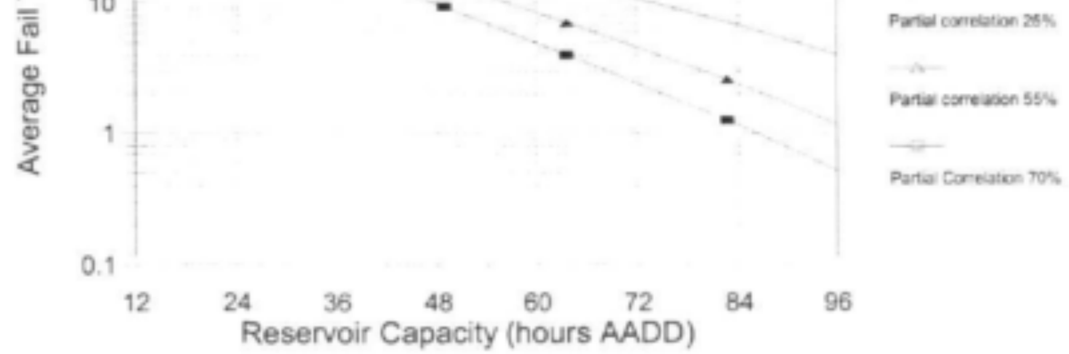


Figure 4.8 Effect of serial correlation on the total annual interruption duration.

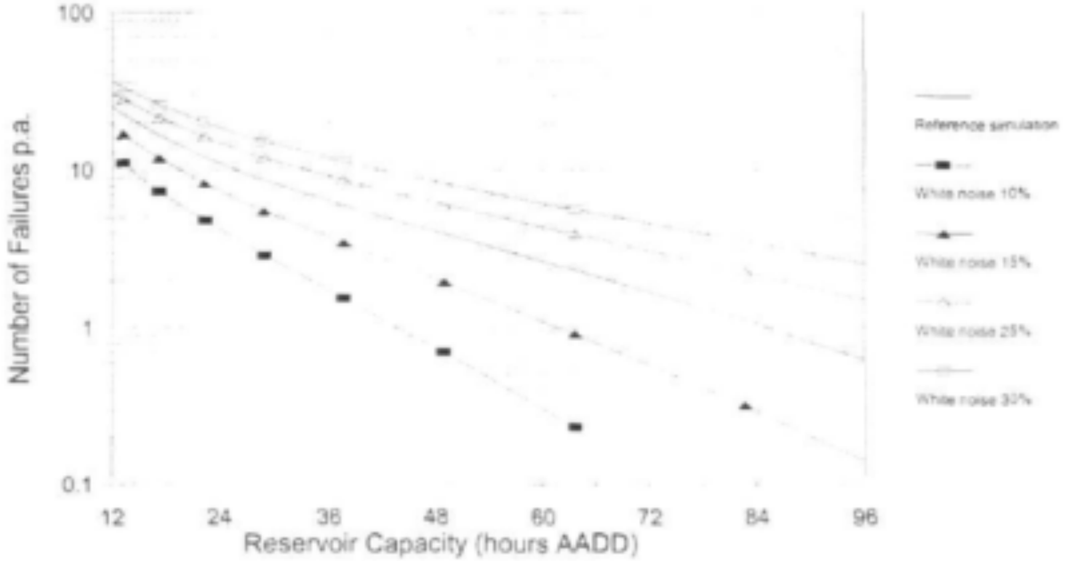


Figure 4.9 Effect of white noise on the average number of interruptions per year.

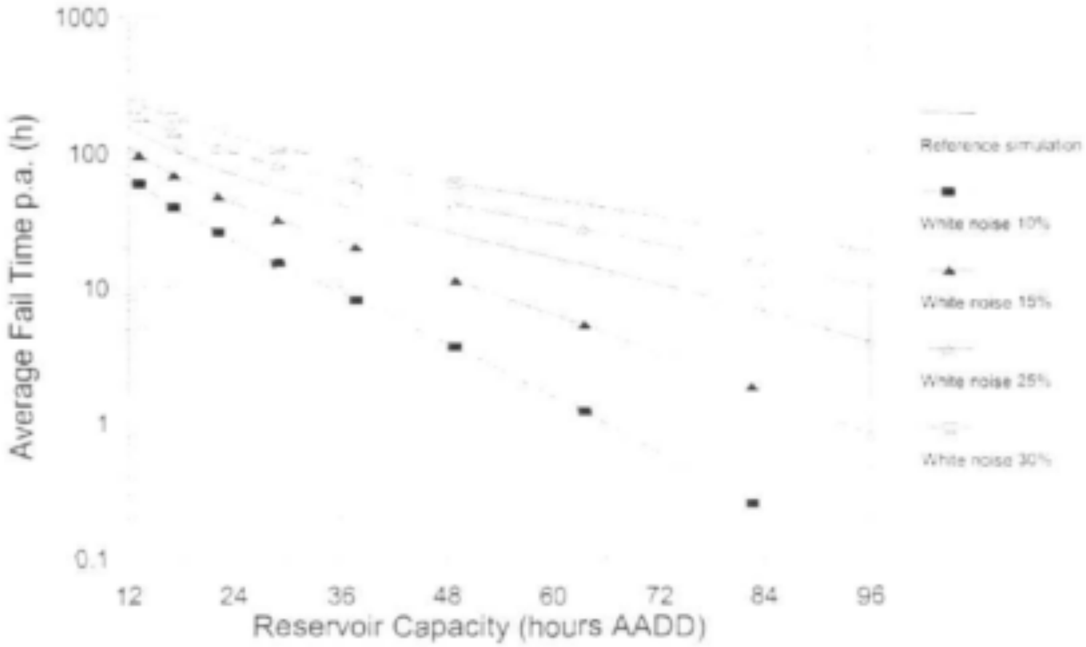


Figure 4.10 Effect of white noise on the total annual interruption duration.

4.8 Effect of Supply Pipe Failures

The frequency of failure of the supply pipe has an obvious effect on the number of supply interruptions. Every time the supply pipe fails, there is a chance of a supply interruption. Also, if the pipe fails frequently, two consecutive pipe failures could be so close together that there is insufficient time for complete system recovery. These effects are shown in Figures 4.11 and 4.12.

The effect of pipe failure frequency is perhaps not as large as one would intuitively expect. If the failure frequency increases from 5 to 30 breaks per year, the number of failures approximately doubles. If the pipe failure frequency is reduced from 5 to only 1 break per year, the number of interruptions drop by only about 10%.

4.9 Effect of Pipe Repair Procedures

Two parameters are used to characterise the time required to repair a pipe. The first is the *average* repair time per break. This parameter was allowed to change from 3 to 27 hours - a wide variation. Its effect is shown in Figures 4.13 and 4.14. There is a remarkably small variation in both the average number of supply interruptions and the total annual interruption duration.

The second pipe repair parameter is the standard deviation of the pipe repair time. If this is very high, it means that some breaks can take much longer to repair than the average time; conversely, a small standard deviation means that all breaks are repaired in approximately the same time. Its effect is shown in Figures 4.15 and 4.16. Here the effect is so small that there are no discernible differences amongst the lines.

4.10 Summary

In order to get an overall sense of which parameters are the most critical, the following procedure was followed:

- The three performance indicators (listed in paragrap 4.2) were applied to all the analyses described above. (The figures depicted only two of these parameters, namely the average number of supply interruptions per year and the total annual interruption duration).
- On each of these graphs, a vertical line corresponding to a 36h-tank was drawn, as this is the typical tank size in SA.
- The maximum and minimum values for each performance indicator were read from this vertical line.
- The range of each performance indicator was then calculated by subtracting the minimum value from the maximum value.
- The range of each performance indicator was divided by its value for the reference case.

These final values, therefore, provide single-point estimates for the sensitivity of the

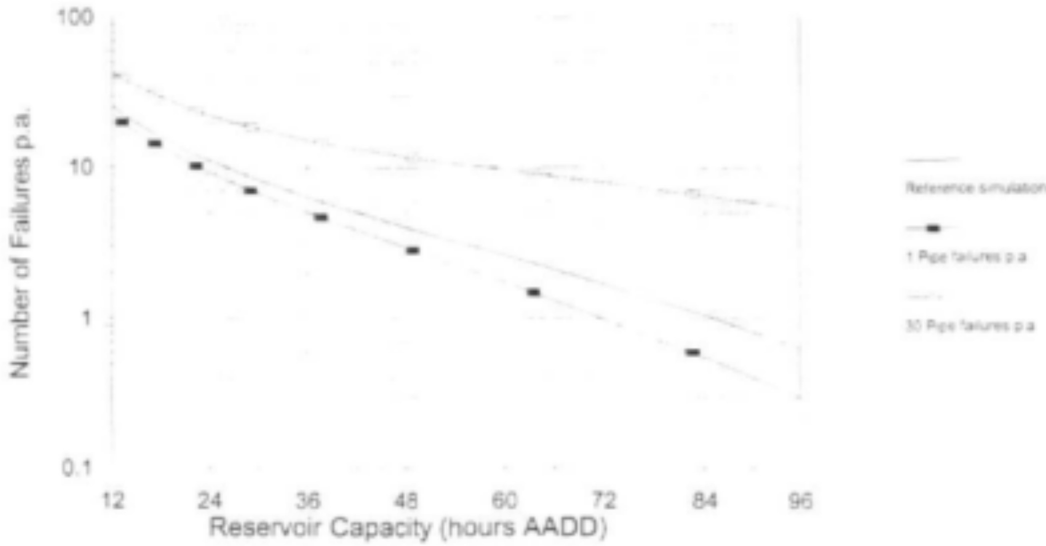


Figure 4.11 Effect of pipe failure frequency on the average number of interruptions per year.

Figure 4.12 Effect of pipe failure frequency on the total annual interruption duration.

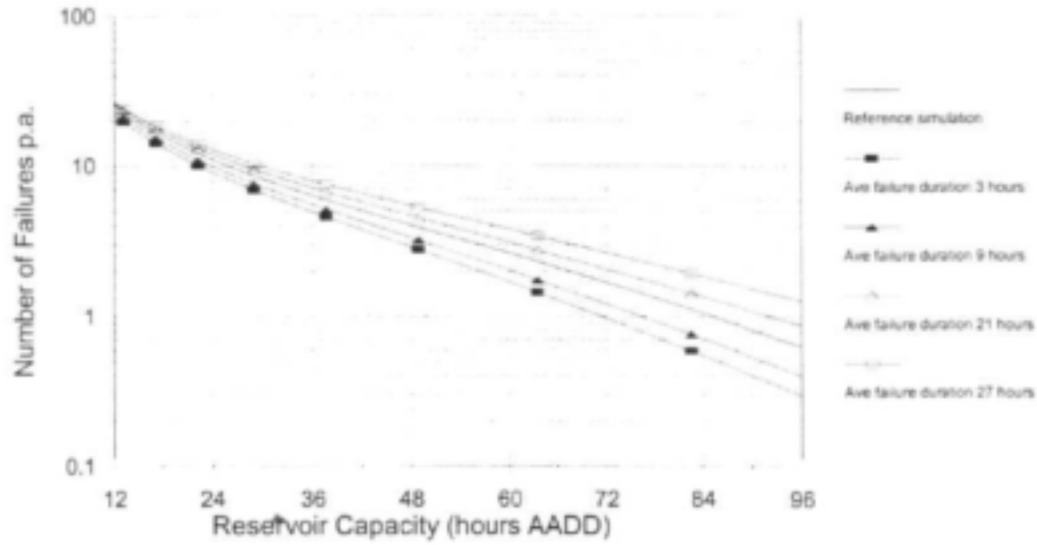


Figure 4.13 ***Effect of average pipe repair time on the average number of interruptions per year.***

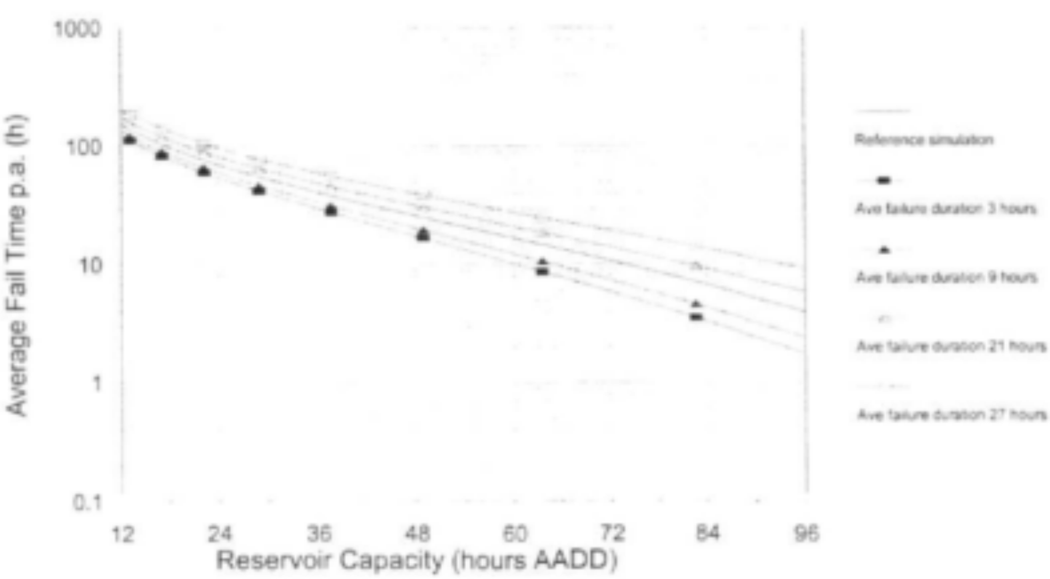


Figure 4.14 Effect of average pipe repair time on the total annual interruption duration.

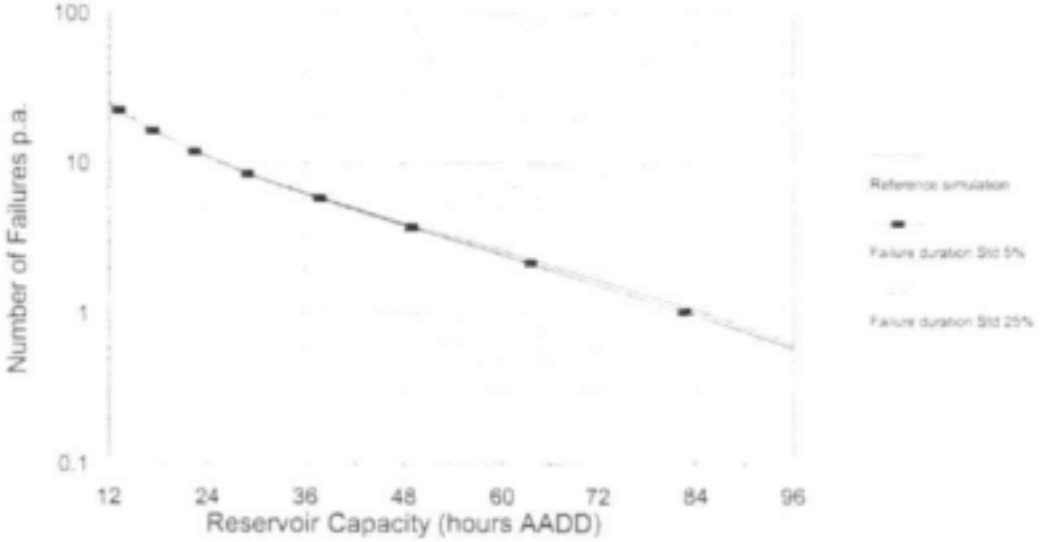


Figure 4.15 Effect of repair time standard deviation on the average number of interruptions per year.

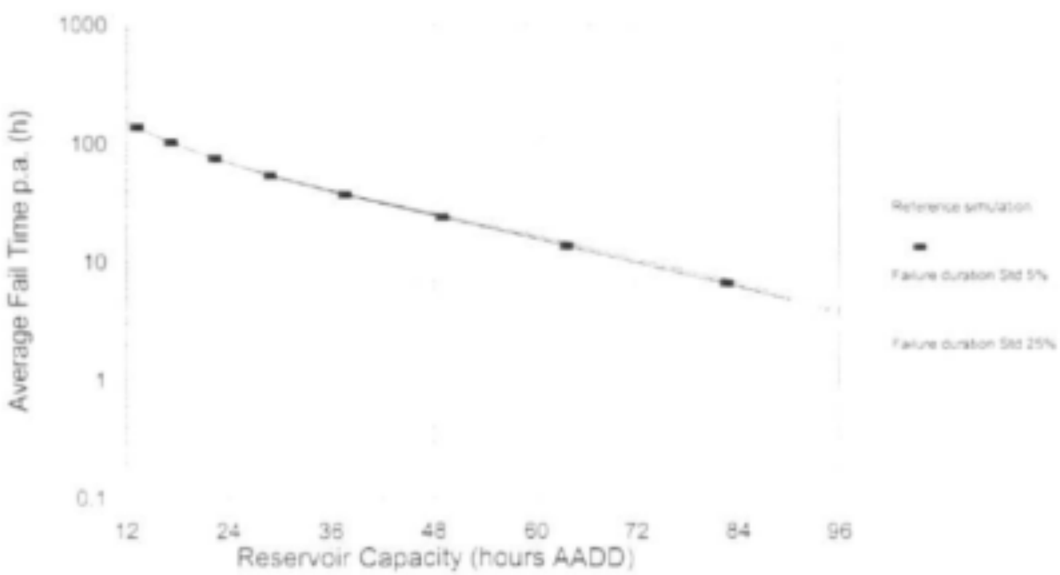


Figure 4.16 Effect of repair time standard deviation on the total annual interruption duration.

system to each input parameter over the ranges indicated in Table 4.1. The results from this analysis are summarised in Table 4.3. The four most important parameters for each performance indicator are marked as "important", while those with very little or no effect were marked as "negligible".

Table 4.3 Ranking of the input parameters in terms of their effects on different performance indicators

	Number of interruptions	Total interruption time	Average interruption time
Parameter X	important	important	important
Day-of-the-week variability	negligible	negligible	negligible
Hour-of-the-day variability	important		
Serial correlation		important	important
White noise	important	important	important
Pipe break frequency	important	important	important
Average repair time			
Repair time deviation			
Fire flow	negligible	negligible	negligible

The following conclusions follow from Table 4.3:

- Approximately the same input parameters prove to be important for all the performance indicators considered. This has a simplying effect, as an improvement in any one performance indicator will probably also cause an increase in the others.
- The effects of fire flow and day-of-the-week variability are negligible in all cases, which means that average values can be assumed for these parameters without introducing serious errors to the probabilistic analysis.
- Parameter X, which is determined from the supply capacity and the maximum weekly demand, is highly critical and holds a significant key to the reliability of bulk water supply systems.
- The pipe break frequency, not surprisingly, turns out to be important for all performance indicators. This is intuitively recognised by water supply engineers and much effort had been devoted to its reduction. What is needed, however, are better methods for monitoring and reporting of pipe failures to enable more accurate probabilistic assessments.
- The stochastic parameters describing daily water demand (serial correlation and white noise) play a meaningful role and further efforts are required to better describe and predict these variables.

CHAPTER 5 - APPLICATION TO TYPICAL SOUTHERN AFRICAN SUPPLY SYSTEMS

5.1 Introduction

Having gained a general grasp of the critical parameters for bulk water supply systems in the previous chapter, this chapter will apply the proposed probabilistic method to three typical Southern African systems. The examples were chosen not only to demonstrate how the method works, but also to gain some insight into the problems faced by different kinds of systems. The three cases presented, had all been studied earlier by the RAU Water Research Group, and the input parameters are therefore realistic and actually measured. The measurement and estimation of the input parameters will not be discussed - only their application.

5.2 A Large Metropolitan System - a Typical Johannesburg Reservoir

The following analysis of such a typical system is primarily aimed at determining how reliable urban systems generally are. At present, as indicated in the literature review, there are no benchmarks against which system reliability can be compared.

The City of Johannesburg obtains its drinking water in bulk from Rand Water, who operates an extensive pipe network throughout its supply area. There are multiple off-takes from the Rand Water network throughout the city. The agreement between Rand Water and its consumers is such that consumers have to attenuate their demand peaks using their own reservoirs. The typical Johannesburg "primary" reservoir (of which there are about 30) is therefore fed from the Rand Water network through a relatively short supply pipeline. There are often cases where "secondary" reservoirs are fed in turn from primary reservoirs; also by relatively short supply pipelines.

The analysis of the water demand from numerous Johannesburg reservoirs yielded a comprehensive database of input parameters, from which fairly robust estimates of typical input parameters could be made. The input parameters adopted for this analysis are summarised in Table 5.1. Using these input parameters, a probabilistic analysis was made to determine the main performance indicators (as before) as a function of storage tank size.

The results are shown in Figures 5.1, 5.2, 5.3 and 5.4. The following comments can be made:

- The average number of supply interruptions for a typical Johannesburg storage tank with storage of 24 to 36 AADD, lies between 0.1 and 1.0 interruptions per year. This number is significantly lower than the typical values reported in the sensitivity analysis in the previous chapter, which is due to the typical $X = 0.72$ used - much lower than the very conservative $X = 1.0$ used for the sensitivity analysis.

- The total interruption time per year is about between 1 and 10 hours per year.

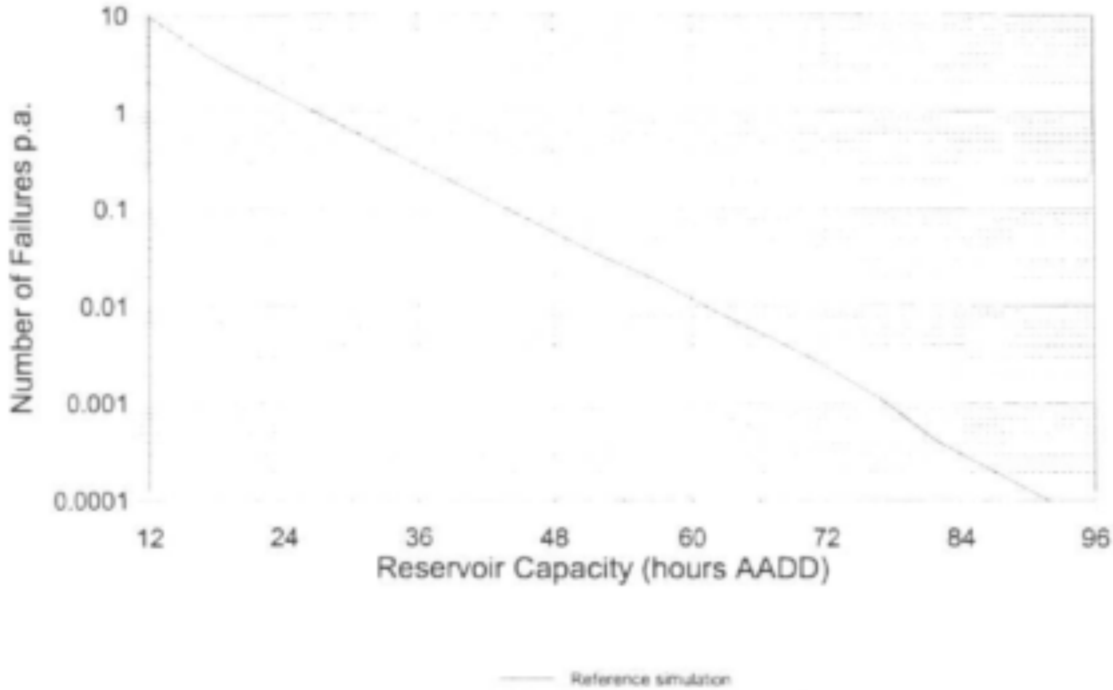


Figure 5.1 Average number of failures per year for a typical Johannesburg reservoir

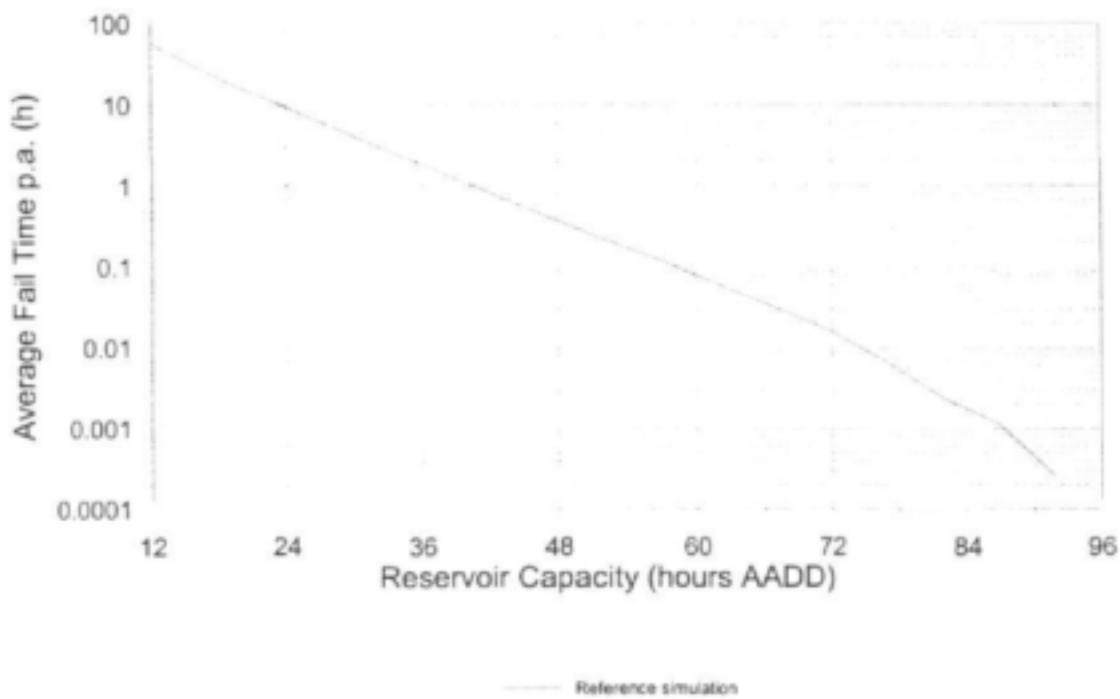


Figure 5.2 Total interruption duration for a typical Johannesburg reservoir

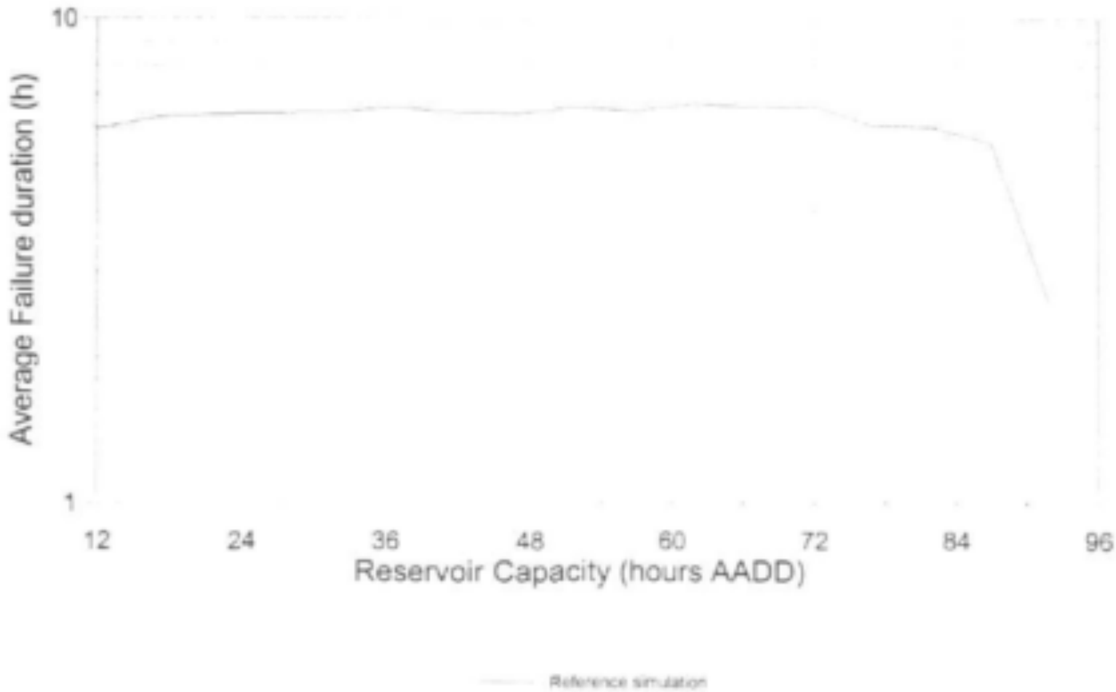


Figure 5.3 Average interruption duration for a typical Johannesburg reservoir

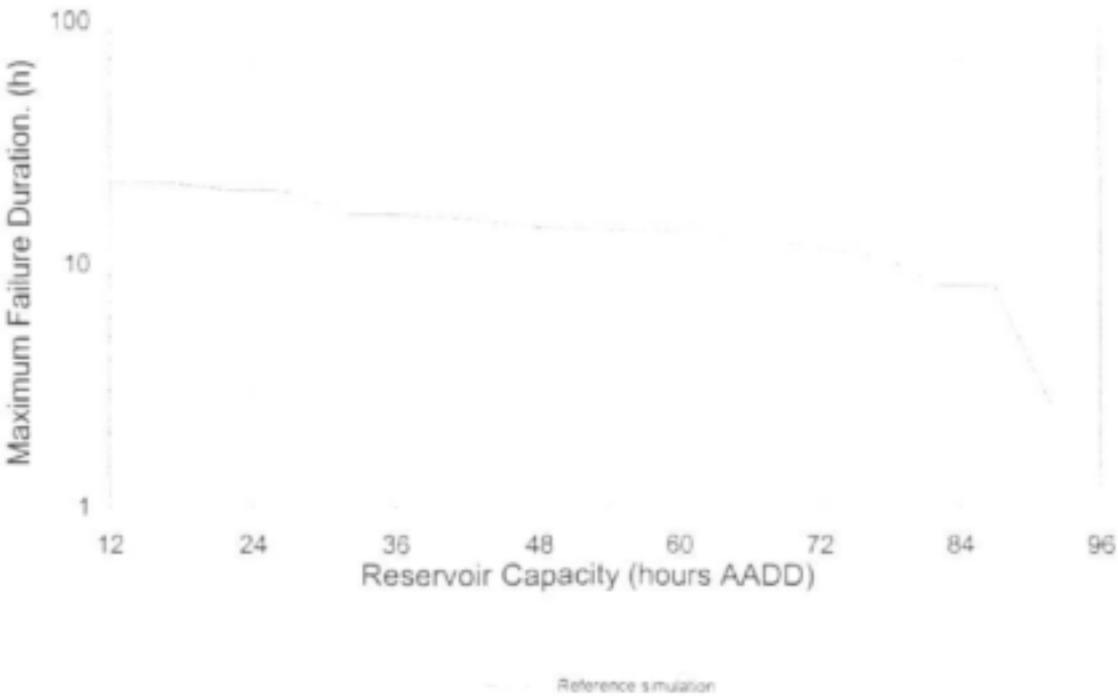


Figure 5.4 Maximum failure duration for a typical Johannesburg reservoir

- The average interruption duration is between 6 and 7 hours, regardless of the size of the storage tank.
- The maximum interruption duration during a typical year is 15 hours.

Table 5.1 Input parameters for a typical Johannesburg reservoir.

Supply pipe capacity	1.5 times AADD
Seasonal variability	Maximum weekly factor = 1.08 Minimum weekly factor = 0.94 Parameter X = 0.72
Daily variability	Maximum daily factor = 1.1 Minimum daily factor = 0.9
Hourly variability	Maximum hourly factor = 1.58 Minimum hourly factor = 0.44
Serial correlation between days	0.43
White noise on daily demand	0.32
Pipe break probability	0.12
Pipe break average duration	6 hours
Coefficient of variation of above	14%

5.3 A More Complex Urban Arrangement - Windhoek

The second example is based on an arrangement found in the City of Windhoek, which is typical of many such systems in Southern Africa. In this case, a central reservoir (the High Sam reservoir complex) is supplied from a distant source (the Von Bach Dam near Okahandja) through a long pipeline. From this central reservoir, there are four separate off-takes:

- Direct consumer demand, with an AADD of 6800 m^3/d.
- A pipeline feeding the secondary reservoir Aussablick, from which an AADD of 4800 m^3/d is supplied.
- A pipeline feeding the secondary reservoir Luipaardsvallei, from which an AADD of 800 m^3/d is supplied.
- A pipeline feeding the secondary reservoir Kleine Kuppe, from which an AADD of 7800 m^3/d is supplied.

The AADD of the combined supply areas is 20200 m^3/d. The direct demand from the first reservoir is therefore about one-third of the total AADD. This is an example of a very common arrangement of a primary reservoir having dual purposes, namely to balance out the fluctuations of its own consumers, as well as being a "transfer" tank which must balance out the imbalances between source and end reservoirs.

The emphasis is on the general behaviour of the entire four-reservoir system. The actual reservoir volumes were therefore treated as variables to see how they influence the reliability. Three cases were analysed:

- In the first case, the secondary reservoirs were fairly large, namely having volumes equal to 60h at the AADD.
- In the second case, the volumes of the secondary reservoirs were reduced to 48h at the AADD.
- The third case was similar to the second, with the exception that the direct demand from the primary reservoir was eliminated. In this case, the primary reservoir was thus a pure "transfer" tank.

The control system was such that the secondary reservoirs had first priority to the water in the primary reservoir. In other words, even if the primary reservoir was almost empty, water would be transferred to the secondary reservoirs if there were the slightest shortfall. The nett effect of this control philosophy is that consumers of the primary tank would suffer more interruptions than consumers from the secondary reservoirs.

The results of the analyses are shown in Figures 5.5 to 5.8. The following can be deduced from the results:

- The capacity of the secondary reservoirs can be increased from 48h to 60h by adding about 7000m^3 storage volume in total to the secondary reservoirs. This small increase has a very large effect on the total number of interruptions, as shown in Figure 5.5.
- To keep the total interruption time to about 1 hour per year, the primary reservoir has to be increased by about 8000m^3 if the secondary storage is reduced by 7000m^3 (in other words, reduced from 60h to 48h). Measured by this performance indicator, it does not seem to make a large difference whether extra storage is allocated to the primary or the secondary reservoirs.
- If the demand from the primary reservoir is ignored, the primary reservoir acts purely as a transfer tank. To keep the total interruption time to about 1 hour per year, the primary tank volume should be about 4000m^3, or about 8h of the AADD of the combined secondary reservoirs.

5.4 A Typical Rural System - Mabeskraal

The Mabeskraal bulk water supply system is a typical extended rural water supply system through an arid, sparsely populated area in the Northwest Province. The system is operated and maintained by Magalies Water, a statutory water board. Most of the numerous small villages are bedroom communities with women and children at home while men return home after hours or over weekends. The living standard is typically low and a minority of homes have a piped supply into the stand or home. The majority obtain their water from standpipes and haul the water home in containers. Houses are far apart and no provision is made to supply fire water through the system.

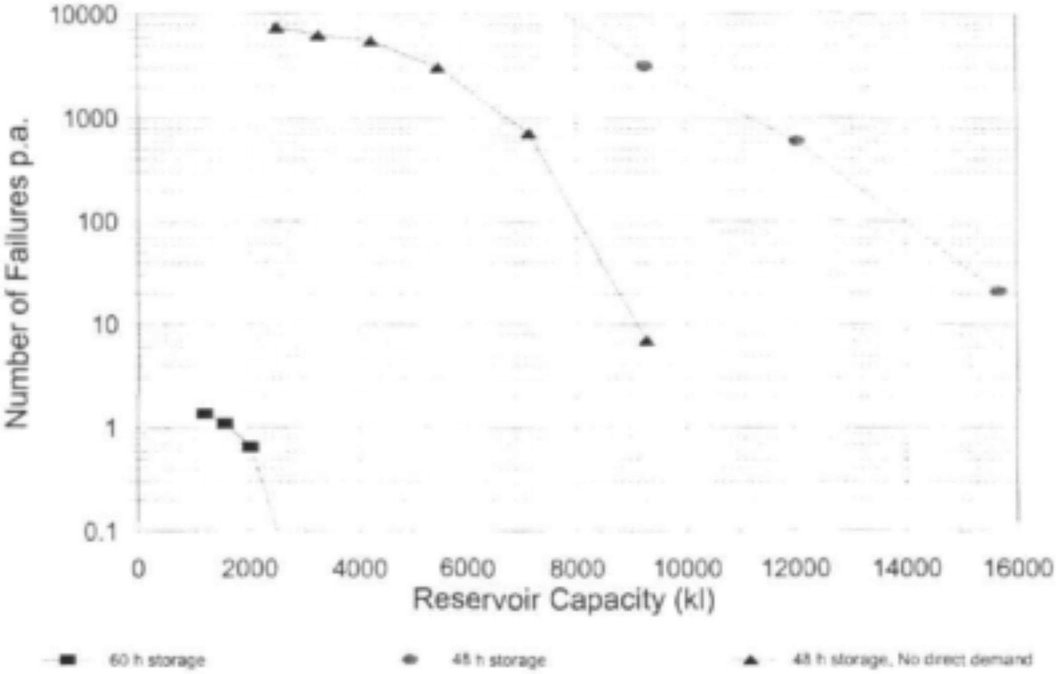


Figure 5.5 Average number of failures per year for the Windhoek reservoir system.

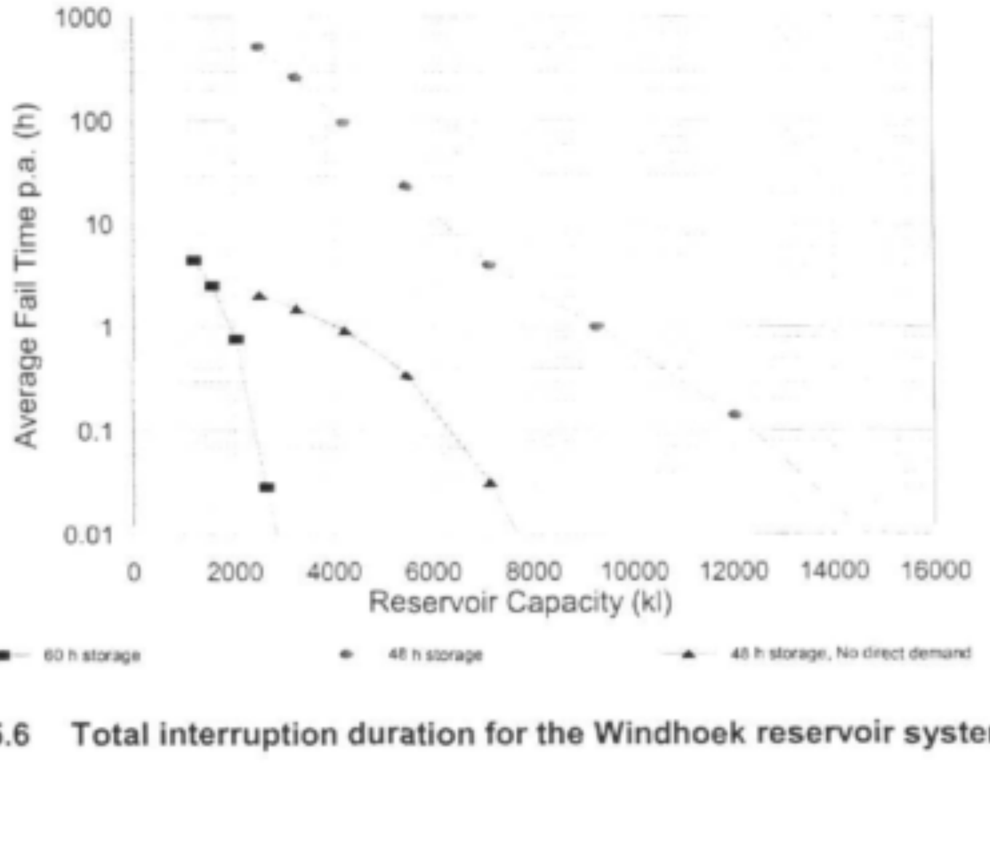


Figure 5.6 Total interruption duration for the Windhoek reservoir system

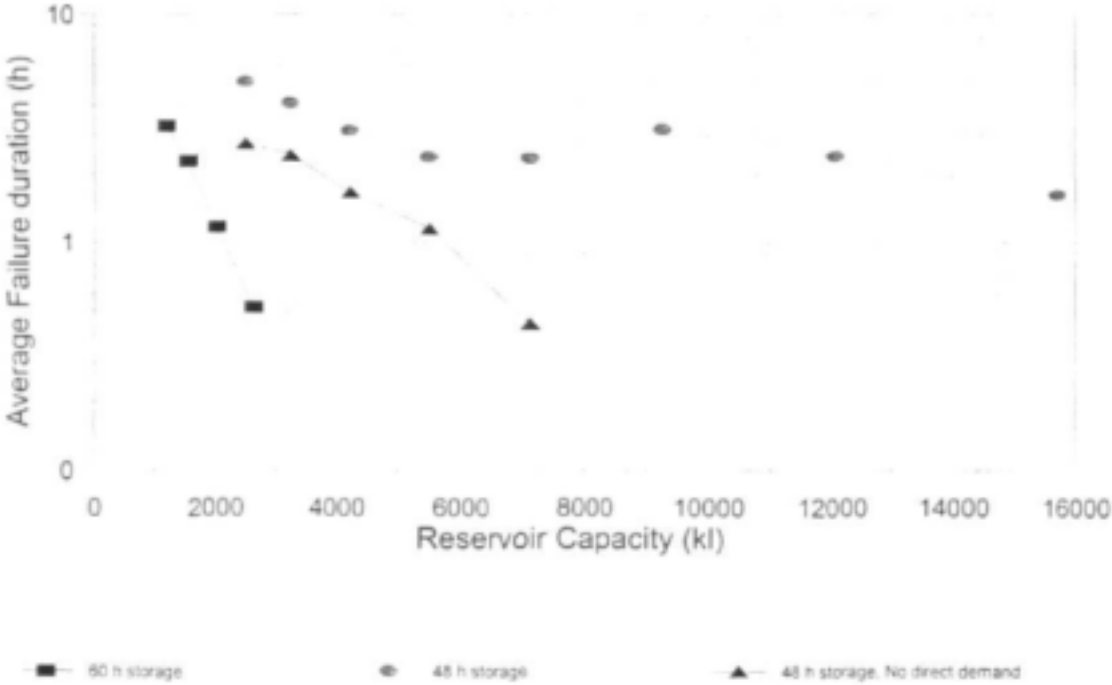


Figure 5.7 Average interruption duration for the Windhoek reservoir system.

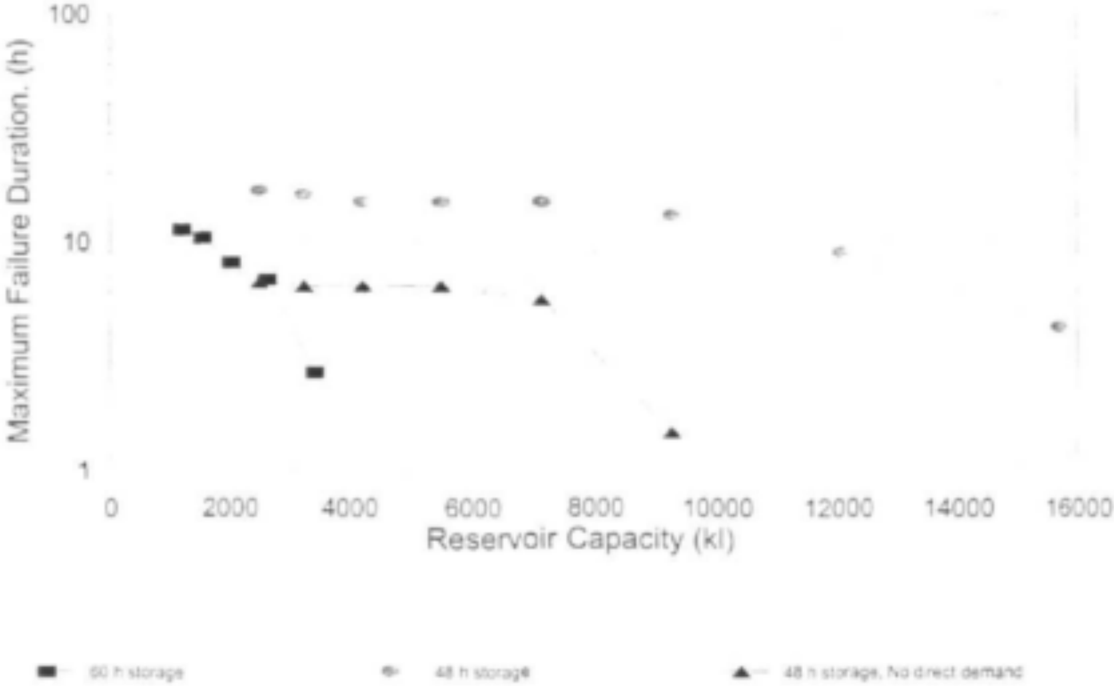


Figure 5.8 Maximum failure duration for the Windhoek reservoir system.

The water supply system consists of a main "back-bone" pipeline, fed from a large storage tank site which supplies treated water to three different systems. From the main pipeline, there are thirteen branch pipes at regular intervals, each leading into a storage tank. From these tanks, individual villages are supplied. Pipes enter the storage tanks at top level, which prevents back-flow from the tank should a pipe break occur. The main pipeline is mostly fabricated from fibre-cement, while the smaller branch pipelines are made of uPVC. The details of the main pipeline are shown in Table 5.2.

Water meters at the outlets of each of the storage tanks are read weekly. These weekly readings were analysed for the hydrological year October 1, 1997 to September 30, 1998. These results, as well as the pipe and tank details at each point of consumption are shown in Table 5.3. Additional flow data was obtained from continuous logging of the overall system at the source, as well as intermittent logging at individual consumers; all at 15 minute intervals.

Table 5.2 Details of Mabeskraal main pipeline

Main Pipeline Section	Distance	Pipe Details
A - B	3 750 m	350 FC + 200 FC
B - C	3 350 m	350 FC + 200 FC
C - D	1 900 m	350 FC + 200 FC
D - E	4 500 m	350 FC + 200 FC
E - F	1 790 m	350 FC + 150 FC
F - G	700 m	350 FC + 160 uPVC
G - H	2 850 m	350 FC + 125 uPVC
H - I	11 100 m	300 FC + 90 uPVC
I - J	8 800 m	300 FC

Note : FC = fibre-cement, uPVC = unplastized polyvinylchloride

There is no significant seasonal trend in the water consumption pattern over a year, which indicates minimum water use for domestic purposes, with little to spare for gardens. Similarly, the different days of the week showed very little variation. The daily demand for the entire system shows a surprisingly strong lag-one auto-correlation coefficient of 0.67 (compared to the 0.40 to 0.60 found for urban systems). The lag-two and further coefficients are statistically not significant.

The electronic data, taken at 15-minute intervals, was used to establish the hourly consumption patterns. The maximum hourly peak factors are much lower than for urban areas; only 1.10 for the entire system and 1.40 for a typical village, both occurring at about 10 o'clock in the morning. The low peaks are explained by the small number of communal taps compared to the much larger number of plumbing fixtures in urban areas, and also a high loss rate from the distribution network.

All maintenance and repair records for three supply systems in the area were analysed for the period February 1977 to August 1998. From this, break probabilities of 0.074 and 0.041 (rounded to 0.10 and 0.05) breaks/km/year were calculated for fibre-cement and polyvinylchloride respectively. The much lower break probability for the rural Mabeskraal system (see literature review for a comparison) is likely due to low construction activity in the area, with very few pipes broken by external forces.

Table 5.3 Details of Mabeskraal draw-off points

Point	Source	Pipe	Tank	AADD	Hours
1	B	400 m 110 uPVC	600 m^3	341 m^3	42
2	C	2200 m uPVC	200 m^3	298 m^3	16
3	C	4900 m 110 uPVC	600 m^3	790 m^3	18
4	D	300 m 110 uPVC	600 m^3	950 m^3	15
5	E	2400 m 150 FC	600 m^3	527 m^3	27
6	F	1800 m 150 FC	800 m^3	455 m^3	42
7	G	700 m 150 FC	600 m^3	689 m^3	21
8	H	900 m 160 uPVC	600 m^3	137 m^3	105
9	I	200 m 200 ST	200 m^3	94 m^3	51
10	J	4400 m 90 uPVC	600 m^3	223 m^3	65
11	J	18100 m 250 FS	3500 m^3	1341 m^3	63

Note : Last column indicates the available storage in terms of hours at AADD

The average length of each non-supply event was 16.7h with standard deviation of 4.0h. This time includes call-out, drainage, repair and commissioning.

Analyses were performed at the outlet of each of the storage tanks, for a range of capacities, using the MOCASIM software. The capacities were "standardised" to hours of demand at AADD to be able to compare villages and tanks of different sizes. The results are shown in Figure 5.9.

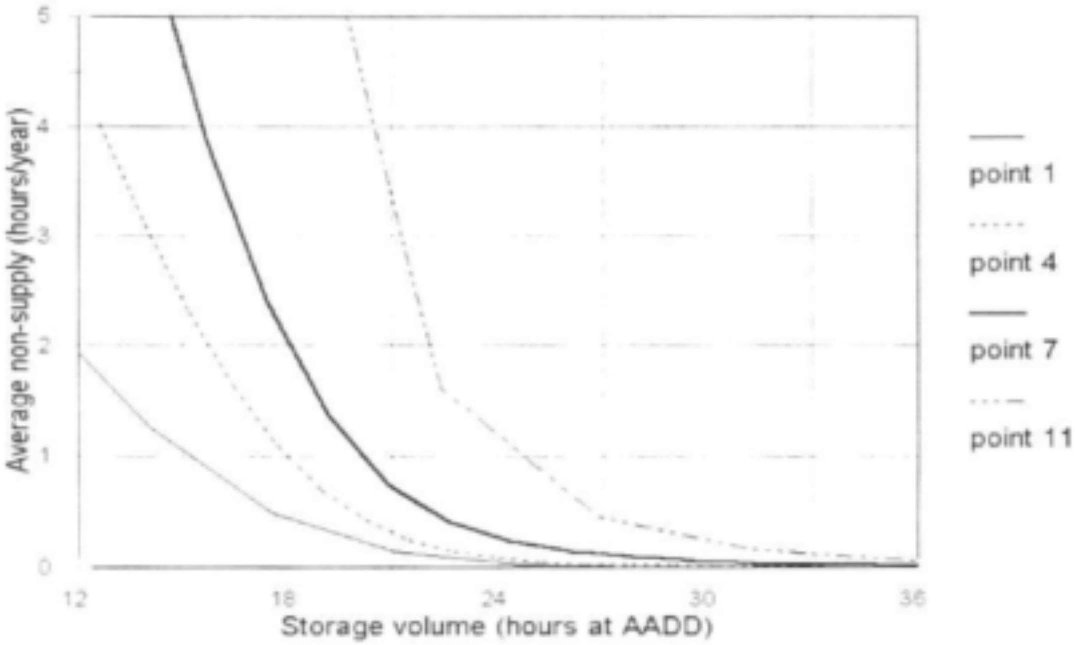


Figure 5.9 Total annual interruption duration per year at four selected points in the Mabeskraal system, as a function of storage tank size. Point 1 is the closest to the source; Point 4 the furthest.

From the analysis, the following points became evident:

- All the non-supply events were associated with pipe failures, and none with excessive demand only. Tanks remained very close to full during all times except for pipe breaks.
- The average duration of an interruption, and its standard deviation, is about the same for all points.
- From the installed storage tanks (refer to Table 5.3 for storage volumes in terms of AADD), the average time of non-supply is zero or very close to it. This seems excessively safe.
- If the average time of non-supply is limited to 1 hour, storage tank volumes of 12 to 24 hours of AADD would be acceptable (refer Figure 2).
- The average time of non-supply increases as one moves downstream, due to the dependence on a longer length of pipeline.

The Mabeskraal case study highlighted some important features of rural water supply systems:

- There is, by nature of the fact that most water is manually hauled to homes, almost no systematic seasonal or day-of-the-week variation in the demand.
- There is a strong lag-one serial correlation coefficient of 0.67 in daily demand, with no significant lag-two or further coefficients. This is in contrast with urban systems where serial correlation coefficients are weaker.
- Due to the hydraulic constraint of the limited number of communal standpipes, and the fact that the water is stored "off-line" in the homes, the hourly peak factors are much lower than in urban systems - 1.10 for the entire system and 1.40 for a typical village. A high leakage rate (not quantified) from the distribution network contributed to this observation.
- Small-diameter pipes (which have been found in urban systems to be broken mostly by external forces) have a significantly lower break probability than their urban counterparts due to the very low level of construction activity. Conservative values of 0.10 and 0.05 breaks/km/year were used for fibre-cement and polyvinylchloride respectively.
- Due to the long lengths of pipelines and small consumer demand fluctuations, storage tanks are predominantly required for "emergency storage" and much less for "balancing storage".
- The average time of non-supply per year for the Mabeskraal system (whose storage tanks were designed according to urban standards) was practically zero (with the exception of Point 4). Analysis of other systems and recommendations in the literature suggest that an average time of non-supply of about 1 h/year would be acceptable for all consumers - urban and rural.
- To limit the average time of non-supply to 1 h/year, storage tanks need to have capacities of between 12 and 24 hours of AADD - significantly less than currently installed.

5.5 Summary

The three case studies presented demonstrated how probabilistic analysis can bring a fresh insight into the behaviour of water supply systems. In the cases presented, the following was shown:

- Estimates were derived for a typical urban bulk water supply system. The number of supply interruptions are typically between 0.1 and 1 interruptions per year, and the total annual interruption duration ranges between 1 and 10 hours per year.
- The approach can be extended to a typical branched configuration as found in many cities, i.e. where water is supplied to a primary reservoir, from where it is further distributed to secondary reservoirs. The analysis of the Windhoek reservoir system, as an example, showed that about 8 hours of the AADD of the combined secondary reservoirs should be allocated to the primary tank if the primary tank has only a "transfer" function.

- The analysis of the Mabeskraal system showed how the method is applied to a long, linear "backbone" system with small side reservoirs. In such systems, tank sizes should increase with distance from the source. Generally, the tanks can be significantly reduced in size (to between 12h and 24h of AADD) without a compromise in supply security.

- The Mabeskraal analysis also pointed to some important differences between urban and rural systems. The demand peaks are much more attenuated, due to on-site storage and the use of standpipes. There is also stronger serial correlation between consecutive days than in urban areas. Pipe breaks are less frequent in rural areas, and repair times generally longer.

CHAPTER 6 - SUMMARY AND CONCLUSIONS

6.1 Summary

This report represents a milestone in a research initiative that had been ongoing at the RAU Water Research Group for the past decade. The contract with the Water Research Commission provided not only the opportunity to develop the computer model more comprehensively and to analyse three relevant case studies, but to record in a single volume the progress that had been made since the early 1990's.

The project originally set out the following research objectives:

- To compile a comprehensive report on the state of knowledge regarding stochastic analyses of water supply systems and storage tank optimisation techniques.
- To develop an algorithm to use stochastic analysis techniques for the optimal design of pipe and storage tank systems. It is specifically aimed at optimising the bulk pipeline(s) from the water source to the storage tank.
- To test and verify the algorithm.
- To make the technique accessible to engineers as computer software and through publications.

The project did meet all the objectives above and the report made the following specific contributions:

- A comprehensive summary of the world literature which is relevant to the probabilistic analysis of bulk water supply systems.
- A conceptual model of the different elements of a water supply systems, with a mathematical description of the factors that act upon each of these.
- An algorithm which allows Monte Carlo simulation on the entire system to obtain probabilistic estimates of system reliability.
- A computer program MOCASIM which can rapidly and conveniently perform such analyses for site-specific input parameters, thereby enabling engineers in practice to perform similar analyses.
- A comprehensive sensitivity analysis to determine the relative importance of the numerous input parameters, to improve the intuitive understanding of system behaviour and to identify the most critical parameters for further research.
- Three case studies which illustrate how the method can be applied to three different system configurations; a system with a single tank, a system with one primary and multiple secondary tanks, and a long linear "backbone" pipeline with multiple "side" reservoirs.
- An illustration of the significant differences between larger, more compact urban systems and smaller, drawn-out rural systems.

6.2 Research Needs

The research conducted for this project pointed to the following areas for further research and development:

- The interrelationship between reservoir size, supply capacity and weekly demand peaks was evident. This report introduced a parameter X (the maximum weekly demand divided by the supply capacity) which turned out to be of critical importance for system reliability. This specific issue should be investigated further, both from a theoretical perspective and by collecting data from a number of cases where failures had been reported.
- The stage is now set for the more complex analysis of systems where reservoirs are used in multiple stages, i.e. where primary reservoirs lead to secondary and tertiary reservoirs. When such concatenated systems are encountered, the *operational* issues turn out to be crucial. Often, the individual reservoirs in such systems are being controlled by different authorities. These systems, because of their practical relevance, deserve closer scrutiny from a theoretical perspective to derive optimal operational strategies.
- The important differences between urban and rural systems have been clearly indicated. Traditional design guidelines for urban systems are not very relevant to rural systems. Further research should be directed at the development of rational, probabilistic design guidelines to be used for rural systems.

REFERENCES

Andreou, S., and D. Marks. (1986) A new methodology for modelling water pipe breaks, Water Forum '86, ASCE, New York, 2,1726-1733

Beim, G., and B. Hobbs. (1988) Analytical simulation of water system capacity reliability. 2: Markov chain approach and verification of the models, Water Resources Research, **24** (9), 1445-1458

Bernier, J. (1987) Bayesian analysis: Further advances and applications, Engineering Reliability and Risk in Water Resources, Martinus Nijhoff Publishers, Dordrecht, Netherlands, 465-484

Billings, R.B., and D. E. Agthe. (1998) State-space versus multiple regression for forecasting urban water demand, Journal of Water Resources Planning and Management, **124** (2), 113-117

Booyens, J.D. (1995) Rekenaarprogram vir die Probabilistiese Bepaling van Reservoirvolumes. B.Eng project, RAU.

Bowles, D. (1987) A comparison of methods for risk assessment of dams, Engineering Reliability and Risk in Water Resources, Martinus Nijhoff Publishers, Dordrecht, Netherlands, 147-173

Brdys, M., and B. Ulanicki. (1994) Operational control of water systems: Structures, algorithms and applications, Prentice Hall, 364

Chou, Y-L. (1969) Statistical Analysis, Holt, Rinehart and Winston, New York, 794

Ciottony, A. (1983) Computerised data management in determining causes of water main failure, Proceedings of the 12th International Symposium on Urban Hydrology, Hydraulics and Sediment Control, Lexington, Kentucky, 323-329

Clark, R., C. Stafford and J. Goodrich. (1982) Water distribution systems: A spatial and cost evaluation, Journal of Water Resources Planning and Management, 108,WR3, 243-256

Cullinane, M.J. (1985) Reliability evaluation of water distribution system components, Hydraulics and hydrology in the small computer age. ACE, New York, **1**, 353-358

Damelin, E., U. Shamir and N. Arad. (1972) Engineering and economic evaluation of the reliability of water supply, Water Resources Research, **8** (4), 861-877

Dolezal, T., Y. Haimes and D. Li. (1994) Water infrastructure risk ranking and filtering method, Risk-Based Decision Making in Water Resources VI, ASCE, New York, USA, 56-72

Duan, N., and L. Mays. (1990) Reliability analysis of pumping systems, Journal of Hydraulic Engineering, **116** (2), 230-248

Fitzgerald, J. III. (1968) Corrosion as a primary cause of cast-iron main breaks, Journal AWWA, **60** (8)

Fourie, L. (1996) Probabilistiese Ontleding van Munisipale Reservoirs. B.Eng project, RAU.

GLS Consulting Civil Engineers. (1995) Stellenbosch Waternetwerk Pypbreukmodel. Report to the Town Engineer, Stellenbosch.

Goulter, I., and A. Coals. (1986) Quantitative approaches to reliability assessment in pipe networks, Journal of Transportation Engineering, **112** (3), 287-301

Goulter, I. (1986) Multi-objective optimisation of water distribution networks, Civil Engineering Systems, **3** (4), 222-231

Goulter, I., J. Davidson and D. Jacobs. (1993) Predicting water-main breakage rates, Journal of Water Resources Planning and Management, **119** (4), 419-436

Groenewald, B. (1995) Periodisiteit van Stedelike Waterverbruikspatrone. B.Eng project, RAU.

Haarhoff, J., Van Zyl, J.E., Nel, D., and Van der Walt, J.J. (1999) Application of Probabilistic Analysis to a Rural Water Supply System. IWSA , Buenos Aires, 20-24 September 1999.

Hartley, J. and R. Powell. (1991) The development of a combined water demand prediction system, Civil Engineering Systems, **8** (4), 231-236

Kaplan, S., and B. Garrick. (1981) On the quantitative definition of risk, Risk Analysis, **1**

Kettler, A., and I. Goulter. (1985) An analysis of pipe breakage in urban water distribution networks, Canadian Journal of Civil Engineering, **12**,286-293

Kwietniewski, M., and M. Roman. (1997) Reliability assessment of water supply systems, Aqua, **46** (5), 283-287

Kwietniewski, M., and M. Roman. (1997) Establishing performance criteria of water supply systems reliablility, Aqua, **46** (3), 181-184

Lewis, E. (1996) Introduction to Reliability Engineering, John Wiley & Sons, 435

Mackenzie, P. (1995) ANCOLD and its interest in risk assessment, Acceptable risks for major infrastructure, AA Balkema, Rotterdam, Nel, D.T. (1993) Bepaling van die Optimale Stoorkapasiteit van Twee Johannesburgse Munisipale Diensreservoirs. M.Eng dissertation, RAU.

Nel, D.T., Haarhoff, J., and Engelbrecht, R.J. (1995) A probabilistic technique for sizing municipal water storage tanks. In : Third International Conference on Computer Methods and Water Resources, September 25-28 1995, Beirut, Lebanon, 145-152. Computational Mechanics Publications, Southampton.

Nel, D., and Haarhoff J. (1996) Sizing Municipal Water Storage Tanks with Monte Carlo Simulation. Aqua **45** (4) 203-212.

O'Day, D. (1983) Analysing infrastructure condition - A practical approach, Civil Engineering, **53** (4), 39-42

O'Day, D. (1982) Organising and analysing leak and break data for making main replacement decisions, Journal AWWA, **74** (11), 589-594

Scheaffer, R., and J. McClave. (1986) Probability and Statistics for Engineers, Second Edition, Duxbury Press, Boston, 648

Schultz, G. (1987) Application of models for reliability assessment in reservoir operation, Engineering Reliability and Risk in Water Resources, Martinus Nijhoff Publishers, Dordrecht, Netherlands, 249-271

Sullivan, J. (1982) Maintaining aging systems - Boston's approach, Journal AWWA, **76** (11), 554-559

Van der Mey, A. (1995) The Failure Probability of Underground Pipelines. B.Eng project, RAU.

Van Zyl, J.E. (1993) Oorsig en evaluasie van standaarde vir brandwatervoorsiening met gepaardgaande implikasies vir netwerkkostes. M.Eng dissertation, RAU.

Van Zyl, J.E., and Haarhoff, J. (1997) South African Fire Water Guidelines and their Impact on Water Supply System Cost. Journal of the South African Institution of Civil Engineers, **39** (1), p16-22.

Van Zyl, J.E. (1998) SOH computer software and manual. Unpublished.

Van Zyl, J.E., and Haarhoff, J. (1999) The Effect of Feeder Pipe Configuration on the Reliability of Bulk Water Supply Systems. Computing and Control for the Water Industry CCWI99, Exeter, 13-15 September 1999.

Venter, M.J. (1999) Ekstraksie van Spitsfaktore vanaf Waterverbruiksdata. B.Eng project, RAU.

Vorster, J., Geustyn, L., Loubser, E., Tanner, A., and Wall, K. (1995) A Strategy and Master Plan for Water Supply, Storage and Distribution in the East Rand

Region. Journal of the South African Institution of Civil Engineers, **37** (2), 1-5.

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

Walski, T., and A. Pelliccia. (1982) Economic analysis of water main breaks, Journal AWWA, **74**,140-147

Yang, S., N.-S. Hsu, P. Louie and W.-G. Yeh. (1996) Water distribution network reliability: Stochastic simulation, Journal of Infrastructure Systems, **2** (2), 65-72

ANNEXURE A

DEVELOPMENT OF MONTE CARLO SIMULATION SOFTWARE (MOCASIM)

A.1 Introduction

A software simulation model called Monte Carlo Simulation Software (MOCASIM) was developed as part of the project. The software does stochastic analyses on bulk water supply systems and allows the user to generate capacity / reliability curves for a reservoir at any node in the system. The software was used to generate all the results published in this report and will allow engineers to do similar analyses on a wide range of bulk water supply systems. The program has been thoroughly tested for correct modelling of its various components.

MOCASIM is available for download from the WRC website at http://www.wrc.co.za/software.

A.2 Design

The software design was done using the object-oriented model of subdividing the problem into a large number of self-contained smaller problems. The smaller problems are programmed into "classes" which interact with each other through predefined methods. There are various advantages attached to this way of programming, including a logical software structure, minimum repetitive coding and ease of testing, maintenance and upgrading.The main program structure is shown in Figure 1.

The programming language used is Java version 1.1. Java is a modern high level programming language which has gained much popularity due to a number of advantages it has over other programing languages. These include that a Java program is platform independent (can be run on most types of operating systems including Windows, Unix, Apple Mac, Sun SPARC and Linux), it is a full object-oriented language, and it is geared toward applications running on the Internet. It is widely regarded as one of the programming languages of the future.

A.3 User interface

The user interface is a typical window based system with a menu and various editing windows. Commands and options are grouped logically for simple and intuitive data entry and editing. The input and result files is in ASCII format which allows for easy interface with other programs for entering or displaying data. For installation the program is packaged in a self-extracting installer that guides the user through the installation process.

Figure A.1: Program structure of MOCASIM showing the main class categories and the number of classes in each category in brackets

A.4 System Elements

The main elements of the program are links, nodes, demand patterns, fail models and fire demand models.

Links

Links or pipes in the system have properties relevant to the hydraulic and failure behavior of the system. The pipe failure behavior in the simulation is modelled according to a fail model that may be selected for it. Pipe failures may be logged in a data file for detailed analysis.

Nodes

Nodes in the system may be sources, reservoirs, junctions or demands. Sources have a fixed water level, while the level in reservoirs vary according to the mass balance that is performed between simulation time steps. Each demand node may be linked to a demand and a fire demand model. A reservoir reliability analysis may be done at each node, irrespective of its type.

Demand Patterns
Various different demand patterns may be defined in the model. Demand models simulate the various deterministic and stochastic elements of system demands.

Fail Models
Fail models simulate the failure behavior of links based on a failure rate and a failure duration distribution model. Both deterministic and stochastic parameters are modelled.

Fire models
Fire models model fire occurrences, demand and durations for demand nodes using user defined distribution functions.

A.5 Simulation runs

After a water supply system is entered into the model, a simulation may be run. The program allows the user to enter any time step length and simulation duration. Typical simulations are done using one hour time step lengths and a simulation duration of 10 000 years. Progress with the simulation is displayed on the screen. After completion of the simulation, various results are summarized and displayed. The user also has the option of having the program write a log of relevant system behavior like pipe and reservoir failures, fires, etc.

ANNEXURE B

DATA ANALYSIS PROCEDURES

B.1 Introduction

Modern water meters and data logging equipment are rapidly becoming state-of-the-art equipment for most water supply authorities. These systems provide data at very high resolution (at 1-hour or 15-minute intervals, or even less) which open a window on water demand patterns which was not previously available. High data resolution, however, translates into vast volumes of data, which are not amenable to simple spreadsheet analysis. It was therefore necessary to develop software for the rapid and standard analysis of data retrieved from data loggers.

Numerous alternative data analysis procedures were experimented with by using a number of typical data sets, before the approach used in this project was adopted. The purpose of this annexure is twofold:

- to demonstrate how the methods are applied to a practical data set, and
- to serve as a preliminary proposal for a standard data analysis procedure which could also be useful to others working in this field.

The test data is a data set provided by Mr. Ben van Merwe, former Water Engineer of the City of Windhoek. Water flow rate was continuously recorded by pen onto monthly paper strips at the outlet of the High Sam Reservoir Complex, one of several serving the City of Windhoek. The strips were manually digitised at hourly intervals by Bernard Groenewald, a former final-year civil engineering student at the Rand Afrikaans University.

B.2 Data filtering and patching

Data sets will rarely be perfectly complete, or perfectly accurate. Typical errors are null values during power or other logging interruptions, a zero followed by a very high value when the volume used during one period is not transmitted properly and then inadvertently added to the volume used during the next period, *etc.* The first step is therefore to screen the data for imperfections. The simplest method is to plot the complete data set, sorted from high to low. Excessively high values, as well as zero or very low values are then immediately and easily visible. By setting realistic upper and lower limits, the dubious points falling outside the acceptable range can be flagged.

The dubious points can either be completely stripped from the data set (leaving undefined gaps at those points), or they can be patched using some average substitute values. A stripped data set will have data discontinuities which require a significant complication in the method of analysis. For that reason, the general preference would be to rather patch a data set to enable simple and rapid analysis.

The choice between stripping and patching is, however, not only one of convenience. It is also determined by the type of analysis required. When extracting extreme peak factors for example (not covered in this project), patching would be in order, as an average substitute value could not possibly be an extreme value. When determining serial correlation, for example (as done in this project), it is essential to revert to stripping. When a data point is completely missing, the correlation with the preceding point simply cannot be calculated. The serial correlation for the entire set is thus only based on original values. If the set is to be patched with average values, the serial correlation for the entire set will be contaminated by correlation between real and patched values. For this project, stripping was therefore performed rather than patching.

B.3 Removal of annual trend

The raw daily volumes, expressed as peak factors, are shown in Figure B.1, after all the dubious data points have been stripped. (Consecutive data points are connected with a line - stripped points are thus present where there are breaks in the connecting line.) In this case, it is quite obvious that there is a consistent trend over the full year. To determine this trend, a linear regression line is fitted through all the data points, which is shown in bold in Figure B.1.

The annual trend is removed by dividing each daily value by the corresponding value on the linear trend line. The resulting data points are shown in Figure B.2.

B.4 Removal of seasonal trend

The seasonal trend is calculated by dividing the year into 13 blocks of 28 days each. (This accounts for 364 days - the remaining day is dropped for convenience with negligible error.) The average peak factor for every block is calculated and plotted in the centre of each block, with straight lines connecting these points. This results in the seasonal trend line, shown in bold in Figure B.2.

The seasonal trend is removed by dividing each daily value by the corresponding value on the seasonal trend line. The resulting data points are shown in Figure B.3.

B.5 Removal of weekly pattern

The average peak factors for each day of the week is calculated next. These average weekly factors are shown in Figure B.4. The weekly pattern is removed by dividing each daily value by the peak factor of the corresponding day of the week. The resulting data points are shown in Figure B.5.

B.6 Removal of serial correlation

The serial correlation is determined by plotting each value against the preceding value. (If there is no preceding value as a result of data stripping, nothing is done and the procedure moves forward by one step.) From this plot, a correlation coefficient is determined - the serial correlation coefficient.

The serial correlation can then be removed from each data point (see Chapter 3 in the report for the appropriate equations) to leave the remaining white noise. The white noise obtained in this case is shown in Figure B.6. Although the year has a maximum of 365 days and can therefore theoretically yield 364 points of white noise, the result of data stripping in this case reduced the number to 197 points of white noise. (One solitary missing data point will lead to the loss of two points of white noise.)

B.7 Characterisation of white noise

The available points of white noise are then sorted into a number of "bins" in order to plot a histogram of the white noise. The histogram is shown in Figure B.7. This is a typical result found in all the numerous cases analysed during this project. The shape of the histogram strongly indicates a normal distribution.

As a final step, the hypothesis (that the white noise follws a normal distribution) is tested. In this case, the hypothesis was shown to be confirmed at a confidence level of 99,9%. The white noise is finally characterised by a normal distribution with mean of 1,0 and standard deviation of 0,46 which is graphically shown in Fibure B.8.

B.8 Summary of parameters

The objective of the exercise described in this annexure is to reduce a typical data set to a number of numerical parameters which can be used for stochastical simulation. In the case of the data set used above, the equivalent parameters are given by:

Seasonal variation:

PF_1 = 1.0985
PF_2 = 0.7643
PF_3 = 1.2920
PF_4 = 1.0898
PF_5 = 1.0946
PF_6 = 0.9229
PF_7 = 0.8181
PF_8 = 0.7952
PF_9 = 0.9451
PF_{10} = 1.0157
PF_{11} = 1.1233
PF_{12} = 1.0920
PF_{13} = 1.0231

Weekly variation:

PF_{monday}	=	1.0222
$PF_{tuesday}$	=	0.9908
$PF_{wednesday}$	=	1.0522
$PF_{thursday}$	=	0.9530
PF_{friday}	=	1.0095
$PF_{saturday}$	=	1.0931
PF_{sunday}	=	0.8536

Serial correlation coefficient:

Lag-one coefficient = 0.526

White noise:

Average = 1.00
Standard deviation = 0.46

Note Stochastic simulation is done for a stationary situation, in other words where there is not a long-term growth trend. For this reason the linear trend is first removed from the data. In other words, the linear trend is only removed to convert the data to a stationary situation - not to use it for stochastic modelling purposes.

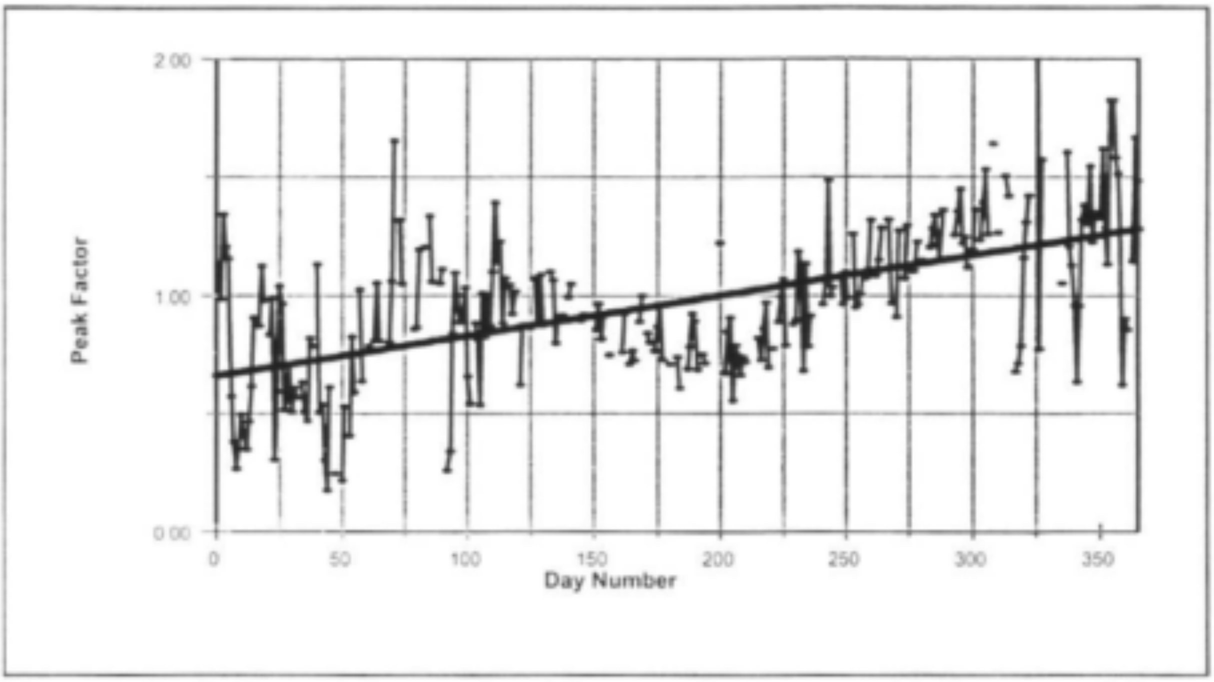


Figure B.1 Daily peak factors (raw) for sample data set with linear trend line.

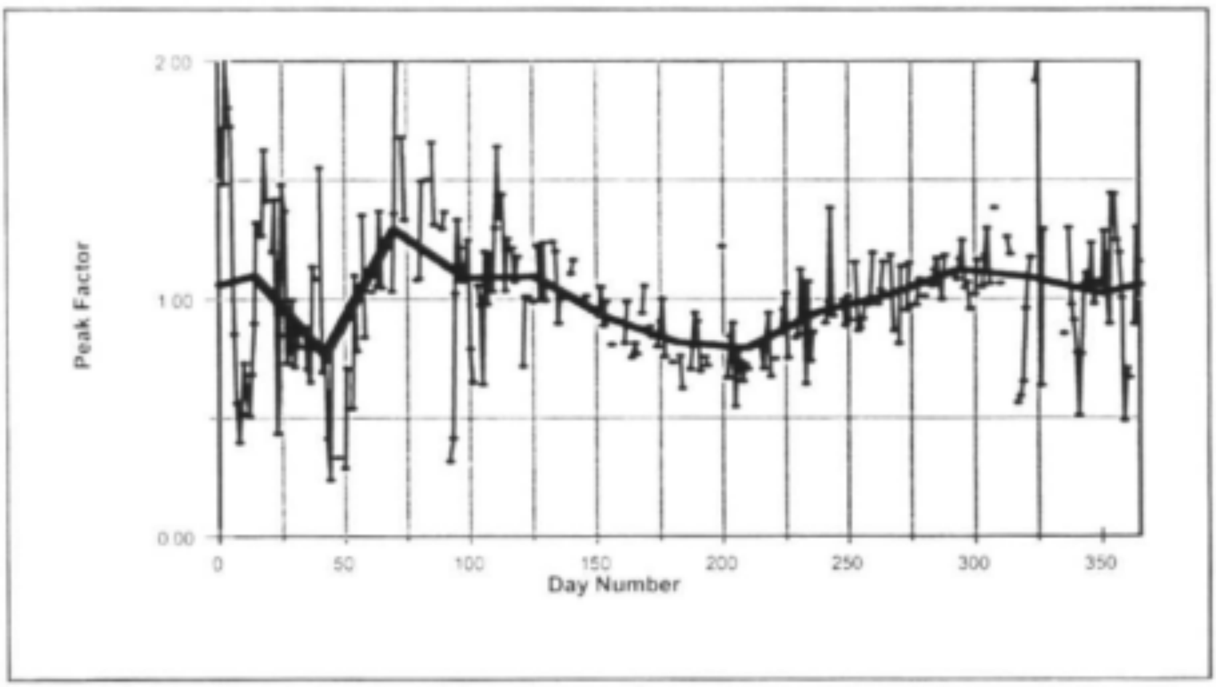


Figure B.2 Daily peak factors (after removal of linear trend) for sample data set with seasonal trend

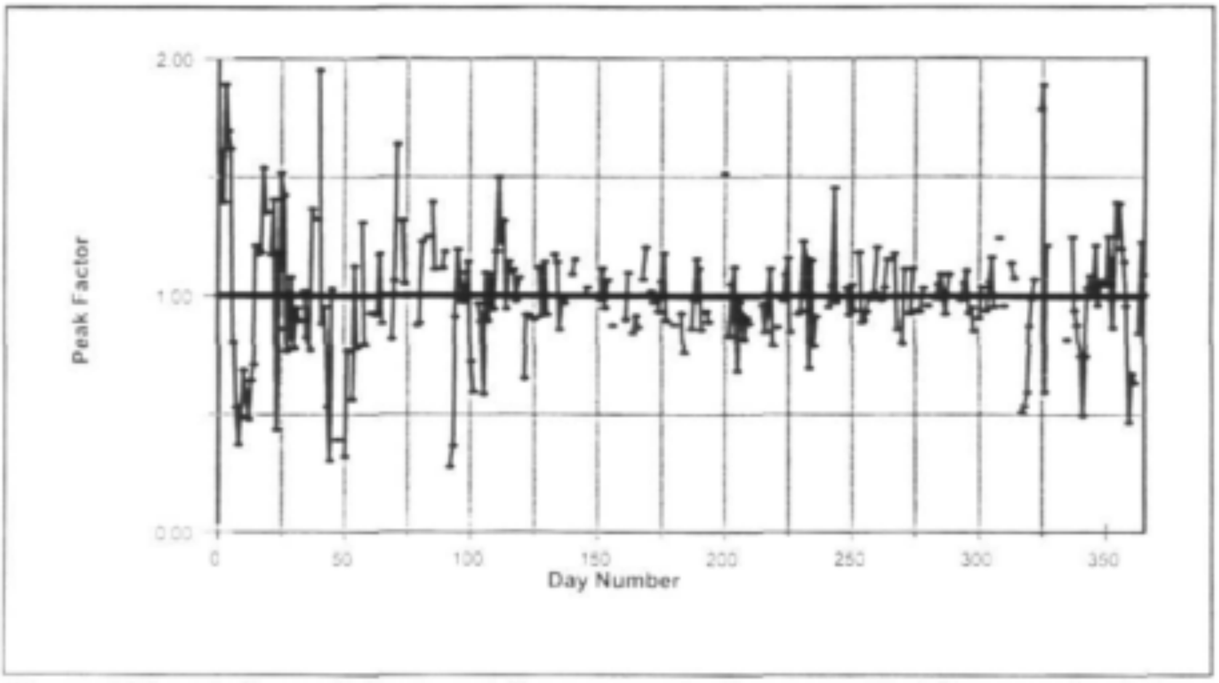


Figure B.3 Daily peak factors (after removal of seasonality) for sample data set.

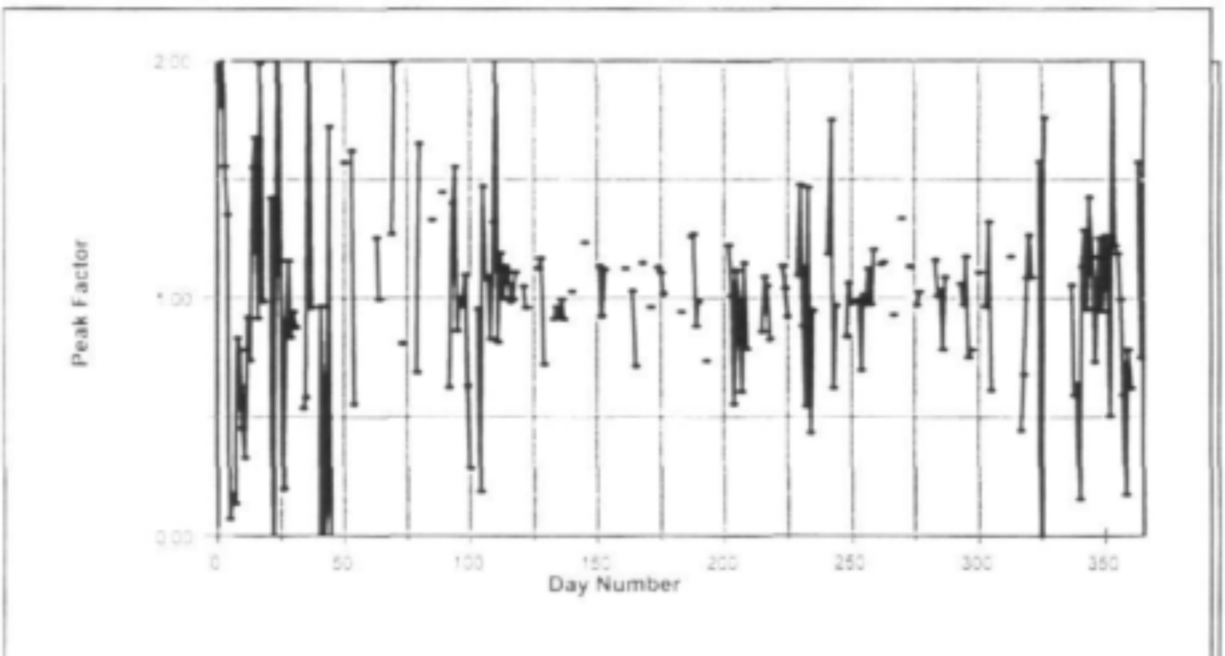


Figure B.6 Remaining white noise (after removal of serial correlation) for sample data set.

Figure B.4 Average peak factors for different days of the week.

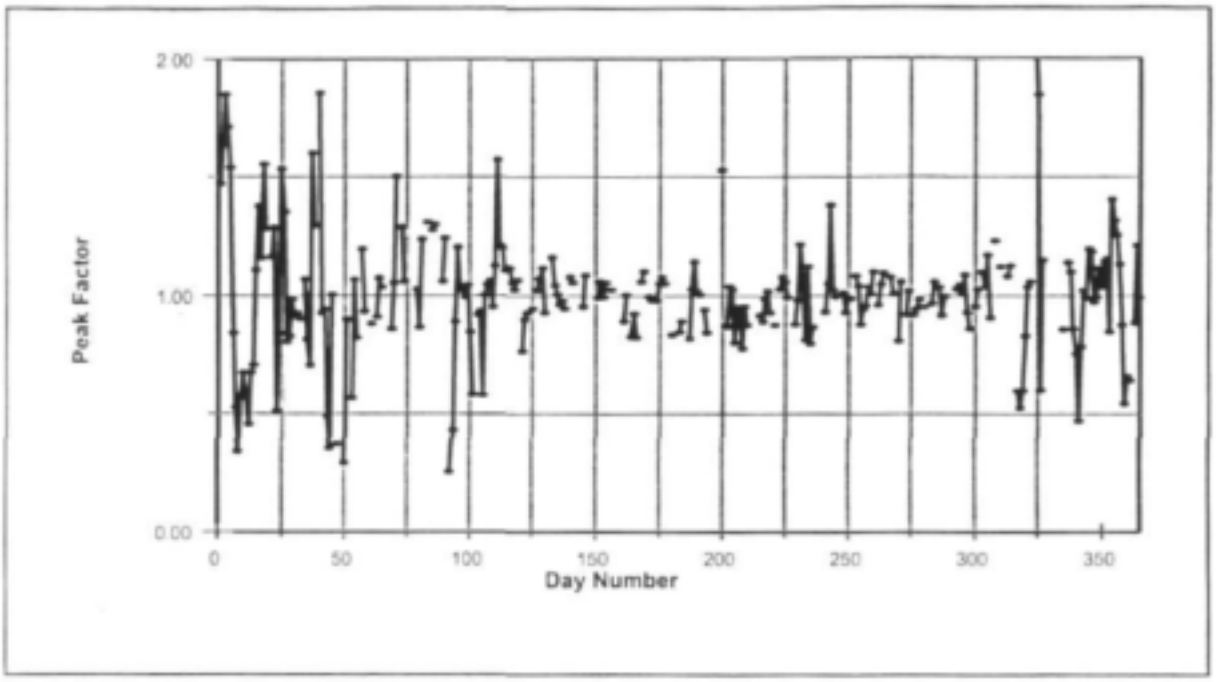


Figure B.5 Daily peak factors (after removal of weekly pattern) for sample data set.

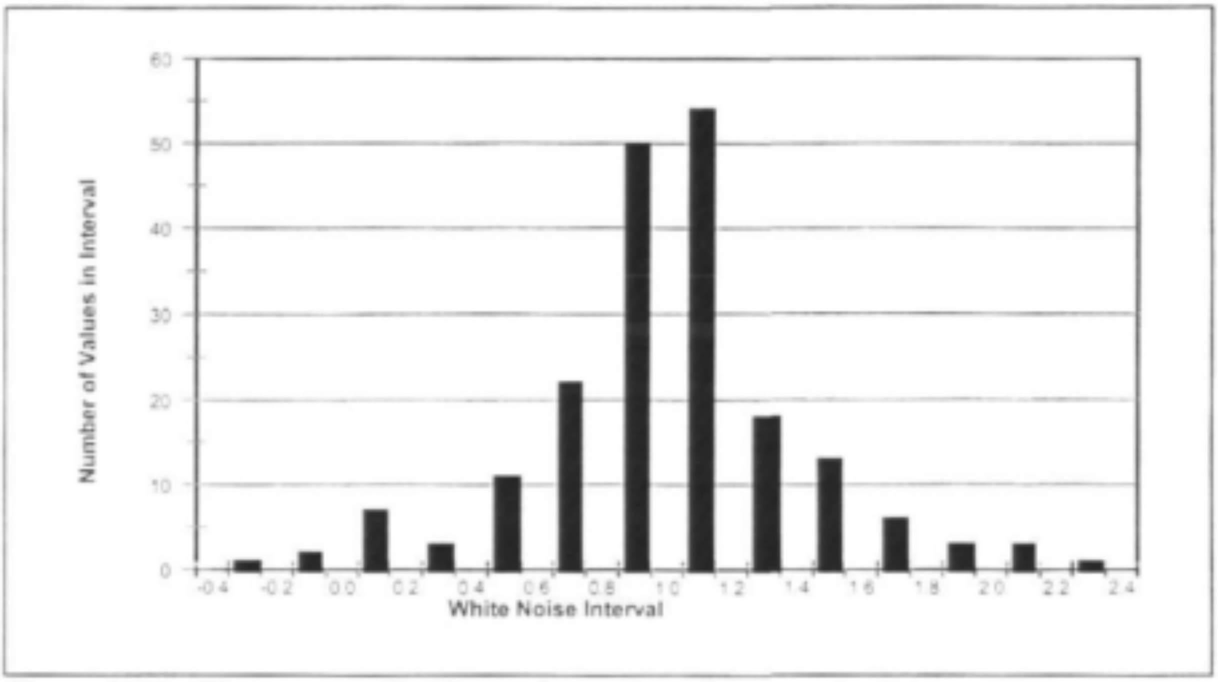


Figure B.7 Histogram of white noise values for sample data set.

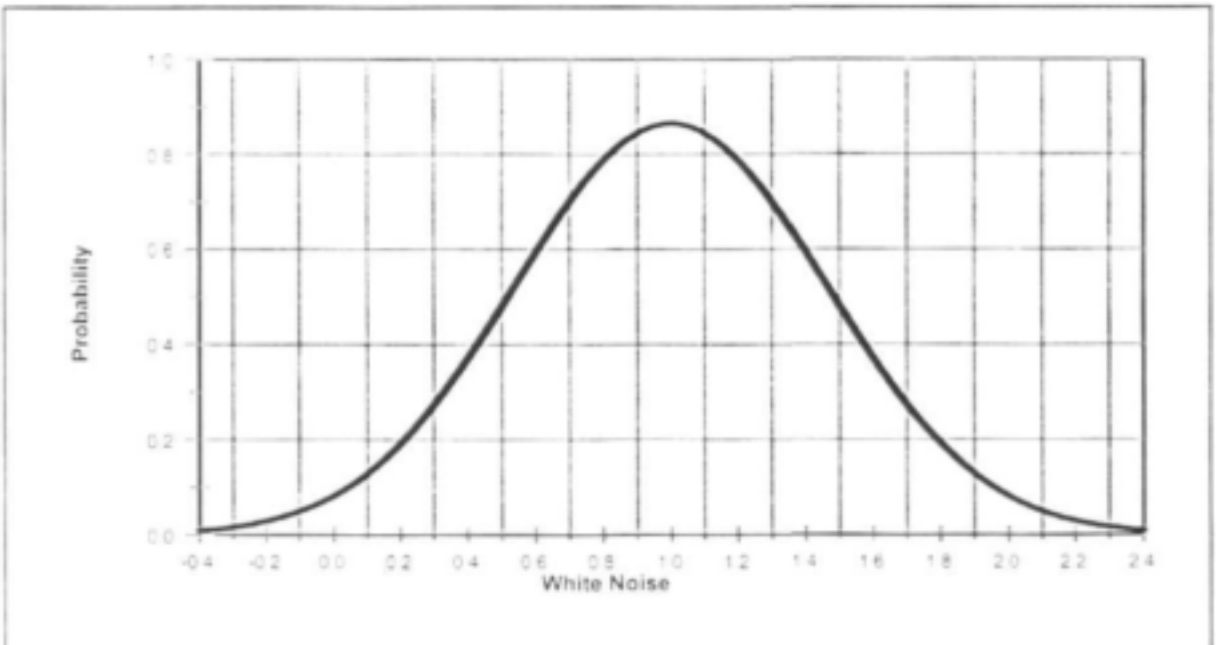
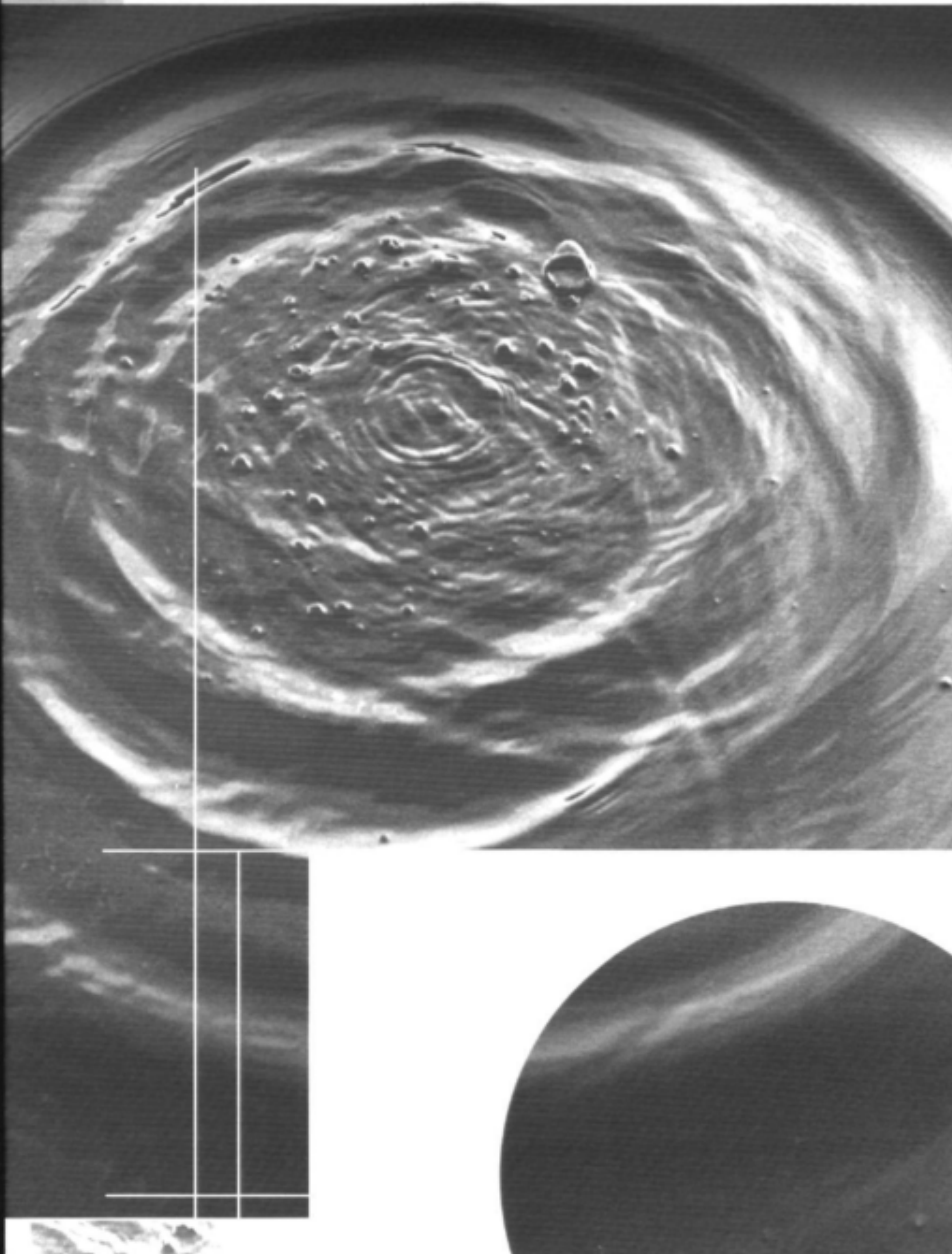


Figure B.8 Normal distribution curve describing white noise for sample data set.



1868458377

Water Research Commission

PO Box 824, Pretoria, 0001, South Africa

Tel: +27 12 330 0340, Fax: +27 12 331 2565

Web: http://www.wrc.org.za