Synthesis of a Model for Optimising a Potable Water Treatment Plant and Water Usage Analysis in the Ugu District

Submitted in fulfilment of the requirements for the degree of

Master of Engineering

in the Department of Industrial Engineering,

Faculty of Engineering and the Built Environment

At the Durban University of Technology

James Magombo

March 2017

Supervisor: Dr B Dzwairo…….....….Date.....

Co-Supervisor: Prof S Moyo ….....……Date.....

Co-Supervisor: Mr M Dewa…….....……Date.....

ABSTRACT

Access to clean and adequate water is a universal and basic human right that feeds into the 6th of the 17 Sustainable Development Goals (SDGs). This goal aims at ensuring availability and sustainable management of water and sanitation for all. Clean water is referred to as potable water, which is safe for human consumption and offer low risk of immediate or long term harm. Raw water undergoes rigorous processing which consists of coagulation, sedimentation, filtration, disinfection and storage, to produce potable water. Each module or stage consumes chemicals and energy resources and thus incurs costs.

To achieve the aim of the study, which was to synthesize an optimised potable water treatment network and a water usage analysis model, the Umzinto Water Treatment Plant (UWTP) and its distribution system was used as the study area. This treatment plant is located within Umdoni, a local municipality of the Ugu District Municipality in KwaZulu-Natal Province, South Africa.

This study's objectives were fourfold and the first objective was to identify and quantify key raw water quality parameters affecting treatment at the UWTP. The second objective was to design a genetic algorithm for the potable water treatment process control. The third objective was to evaluate the Umzinto Water Distribution System's Non-Revenue Water (NRW) while the fourth objective was to develop a model for water usage analysis.

For the first objective, data for water quality parameters for the water treatment from July 2006 to June 2013 were statistically analysed. This data were collected from the UWTP's historical records. To improve the data's integrity it was pre-processed using cubic hermite interpolation. After the pre-processing trend lines and box plots were used to determine the parameters' significance compared to the standard values stipulated in the South African National Standard (SANS 241). The trend lines were used to analyse the frequency of observations that were higher than the standard values according to SANS 241. The box plots were used to determine the minimum, median, maximum and mean of the data sets. The mean values for each parameter were compared to the SANS 241 value to determine their significance. The raw water quality parameters were then correlated to the chemical

dosages for lime, polymer, potassium permanganate and chlorine. The key parameters selected from the correlation analysis were algal count, manganese (Mn), iron (Fe), *Escherichia coli*, total coliforms, colour, odour, conductivity, turbidity, suspended solids (SS), pH, temperature, total organic carbon (TOC,) and Hardness.

A number of methods can be used to achieve such optimisation, including artificial neural networks, dynamic programming, linear and non-linear programming, and this study utilised a genetic algorithm as an optimisation tool to achieve the second objective of optimising water treatment at the UWTP. For the model development, data from the correlations obtained for objective 1 were used. The model was aimed at reducing the cost of chemical dosage and four chemical dosage prediction models were developed using genetic algorithms and these were then used to produce a combined chemical dosage cost prediction model. The programming interface utilised for these models was Matlab. In developing these models, the data were first pre-processed to remove outliers and fill in the blanks using a Microsoft Excel Add-in that was developed for this particular purpose. The next step involved a curve fitting exercise in Microsoft Excel 2013. Matlab was then used to code the genetic algorithm that combined and optimised the solutions obtained from the curve fittings. The results showed that genetic algorithms can be reliably used to predict the chemical dosages and hence reduce water treatment costs.

After treatment, water is pumped into the distribution system for consumption. It is therefore important to ensure that all the pumped out treated water reaches the consumer. The third objective therefore assessed the NRW for the Umzinto Water Distribution System for the period between July 2013 and June 2014. The data used for this objective was provided by the Ugu District Municipality. The method used combined the top-down approach and the component-based approach. This combined approach was modified to enable the calculation of all the components that are required in a standard South African Water Balance. The results showed that the distribution system had a high value of NRW, which was 27.9% of the System Input Volume. The major component of NRW was Real Losses, that is, losses that can be mitigated by improving maintenance.

The fourth objective was to develop a model for water usage analysis that would reduce the time taken to evaluate NRW and also improve the analysis of the NRW components

using Microsoft Visual Basics 2012 and Microsoft SQL Server 2012 development interfaces. The Visual Basics enabled the development of a graphic user interface that was user-friendly and minimised the time taken to learn the software. The software platform developed was able to import the data required to construct a standard International Water Asssociation (IWA) Water Balance, calculate all the components of NRW, store historical data for the water distribution systems and report on a rolling year basis. A model for water usage analysis was developed and made available for usage by practitioners in Ugu District. The model was developed for the specific study area and further studies would be required in order to validate it in a different setting.

The results obtained for the first objective led to the conclusion that, there was very high pollution emanating from communities and activities close to the raw water sources, especially the EJ Smith Dam. The results from the first objective were also used to determine parameters for the models developed in the second objective. From objective two it was concluded that genetic algorithms can be reliably used to predict chemical dosages and hence reduce water treatment costs. The third objective's results showed that 27.9% of treated water pumped into the distribution system is NRW. Which is a concern because 65% of this are real losses which have maintenance related problems. The fourth objective's results showed the practicality of designing model that could be used determine all the important components of NRW that would take time to evaluate manually. It would also store historical data for the water distribution system and report on a rolling year basis. Implementation of this software would help minimise the errors associated with manual calculation of NRW and improve the availability of data for research and analysis.

From the research findings, it is recommended that the treatment plant should change the way it is dosing chemicals in the balancing tank. The method currently being used is prone to error. The analysis of NRW showed that Real Losses were a major challenge in the Umzinto Distribution System. There is need to develop a maintenance program to cater for leakage. Communities also need to be educated on the importance of reporting leakage in the network.

DECLARATION

I hereby declare that this submission is my own work and to the best of my knowledge it neither contains material previously published or written by another person, nor material which to a major extent has been accepted for the award of any other degree at Durban University of Technology (DUT) or any other educational institution. I also declare that the intellectual content of this dissertation is the product of my own work. Any contribution made to the research by others especially in the use of equipment for sample analysis has been explicitly acknowledged in the dissertation.

_____

Mr. J Magombo

Student Number (21350339)

DEDICATION

I dedicate this dissertation to my wife, Judith, who has been a constant source of support and encouragement during the challenges of research and life. I am truly thankful for having her in my life. This work is also dedicated to my father, Wonder, whose work greatly inspired me to pursue engineering. I also dedicate this dissertation to my mother, Agnes, who has always loved me unconditionally and taught me to work hard for the things that I aspire to achieve.

ACKNOWLEDGEMENTS

Above all, I thank God for guiding me in all my endeavours. I would like to express my gratitude and respect to Dr B Dzwairo, Prof. S Moyo and Mr M Dewa for their endless support and guidance throughout the study.

I would like to acknowledge my academic friends for their support and encouragement. This helped me to overcome all the obstacles I faced and to complete this study.

I also would like to thank Durban University of Technology for the study environment and financial support provided in the form of a scholarship.

Finally, my sincere gratitude to Ugu District Municipality and Umgeni Water for offering their work environment as the study area.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

RESEARCH OUTPUTS

1. Magombo J, Dzwairo, B, Moyo S, Dewa M. 2014. Application of Genetic Algorithms Towards Sustenance of Water Resources Systems. 2014. Abstract proceedings. Journal of Green Economy & Development's (JGED) 1st Annual International Conference, Makaranga Lodge, Kloof, Durban, South Africa, 17 - 19 December 2014.
2. Magombo, J., Dzwairo, B., Moyo, S. and Dewa, M. 2015. Availability of Data in Small Drinking Water Treatment Plants. Poster presentation. 4th Young Water Professional-South Africa Biennial and 1st African YWP Conference - 16-18 November 2015.
3. Magombo, J., Dzwairo, B., Moyo, S. and Dewa, M. 2015. Inequality, Water Pollution and Access to Water in the Ugu District. Abstract proceedings. World Social Science Forum 2015. Durban International Convention Centre, Durban, from 13 – 16 September 2015.
4. Magombo, J., Dzwairo, B., Moyo, S. and Dewa, M. 2015. Data pre-processing for process optimization at a drinking water treatment plant in Ugu District Municipality, South Africa. *Environmental Economics*, 6 (1): 159-171.
5. Magombo, J., Dzwairo, B., Moyo, S. and Dewa, M. 2015. Developing a Non-Revenue Water Strategy for the Umzinto Water Supply Scheme: Towards Green Economy. Abstract proceedings. Journal of Green Economy & Development's (JGED) 2nd Annual International Conference, Makaranga Lodge, Kloof, Durban, South Africa, 08-10 July 2015.
6. Magombo, J., Dzwairo, B., Moyo, S. and Dewa, M. 2016. Coagulant Dose Prediction Using Curve Fitting Combined with Genetic Algorithms. Abstract proceedings. Water Institute of Southern Africa Conference, Durban International Biennial Conference Centre, Durban, South Africa, 15-19 May 2016.
7. Magombo, J., Dzwairo, B., Moyo, S. and Dewa, M. 2015. Non-Revenue Water Assessment Strategy for the Umzinto Water Supply Scheme: Towards Green Economy. *In press*. *Journal of Green Economy & Development*.

CHAPTER 1 – INTRODUCTION

1.1 Background

To produce potable water, raw water undergoes rigorous processing, which comprises of coagulation, sedimentation, filtration, disinfection and storage (Droste 1997; Crittenden *et al.* 2012; Honda 2014). At each of these modules or stages, chemicals and energy are used, which introduces a cost dimension to a water treatment system. This illustrates the complexity of the interconnected physical and chemical systems involved in the water treatment process. While it is not obvious how each individual module or stage affects the other, it is certain that failure to address a problem at one stage will quickly result in a much larger problem in one or more subsequent stages (Adgar, Cox and Böhme 2000; Liu *et al.* 2013).

Studies have indicated that the various stages could be optimised to produce high quality potable water that is compliant with governing standards and regulations on water quality. These studies have also focused on minimising the time taken to carry out a number of tasks, for example, quality assessments which have an impact on the effectiveness of a treatment network. While a number of methods such as non-linear programming could be used to achieve optimisation, due to the number of parameters involved and cost considerations, genetic algorithms are becoming popular (Chong *et al.* 2010; Simonovic 2012).

Genetic algorithms belong to a large class of evolutionary algorithms (EA), which provide answers for optimisation problems utilizing procedures propelled by common advancement or natural selection (Ashlock 2006). The problem at hand was that when planning, developing, managing, and operating water resource systems one can never be certain of the parameters. A change in a parameter will affect the value of another or other parameters in the water resources system. Seasonal variations in water quality parameters also bring about uncertainty in the system (Bartram and Ballance 1996; Loucks *et al.* 2005).

After treatment, water is pumped into the distribution system for consumption. Management of this resource beyond chlorination is critical because not all the water that

is produced in the treatment plant reaches the consumer (Lim and Park 2007). The difference between the amount of water produced at the treatment plant and the water that reaches the consumer is called Unaccounted-for-Water (Andey and Kelkar 2009). A model was developed to analyse such losses, some of which are the result of underground leakage, unauthorised usage, unavoidable network leakage, inaccurate/faulty meters, and other unusual causes (George *et al.* 2009; Kumar 2010). The model also incorporates an analysis of Non-Revenue Water (NRW) components.

This study developed a Genetic Algorithm model that optimised treatment at a small drinking water treatment plant. The design considered raw water pollution sources, and the various ways in which they affect the quality of treated water. The study also improved the analysis of water usage in the treatment plant's distribution system by incorporating a model that analysed the components of NRW. The water treatment plant and distribution system considered are situated in the KwaZulu-Natal province of South Africa, shown in Figure 1.1. The map was created using a software platform called ArcGIS for desktop student edition version 10.1 (ESRI 2013).

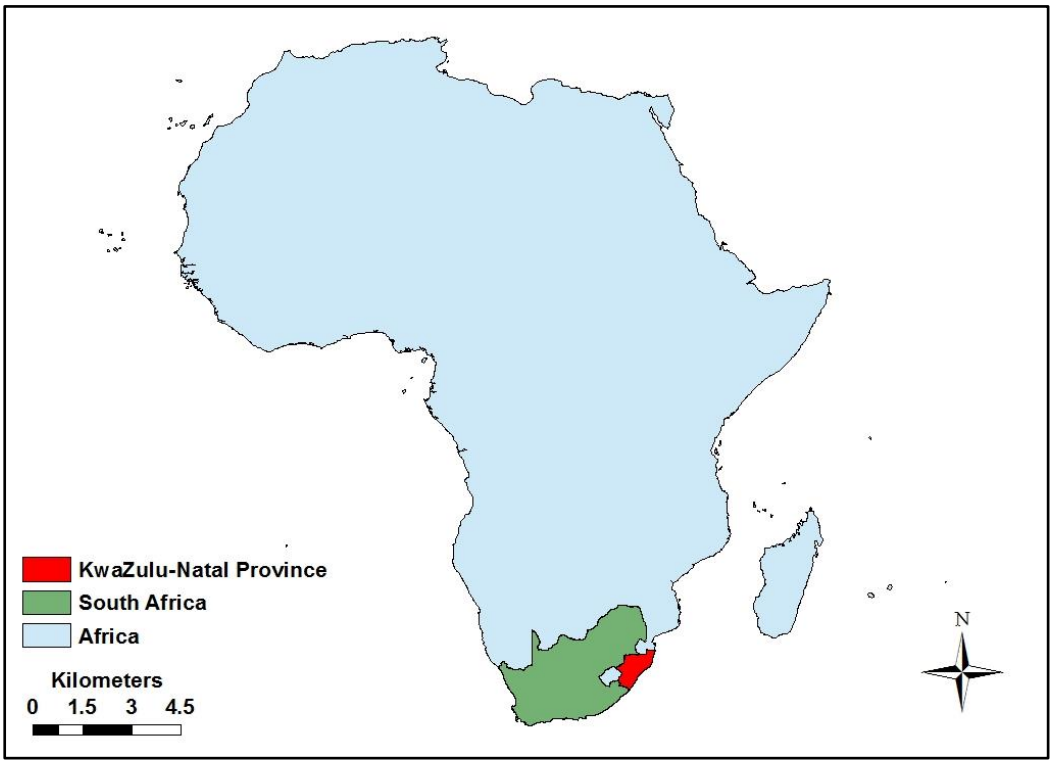


Figure 1.1: Location of the KwaZulu-Natal Province. Source of shape files: Internet web pages. Software platform: Environmental Systems Research Institute (ESRI) ArcGIS 10.1

The treatment plant and its distribution system fall under the Umdoni Local Municipality in the Ugu District. The treatment plant gets its name, Umzinto, from its major supply areas which are Umzinto town and rural. The map in Figure 1.2 shows the location of Umdoni Local Municipality in the district.

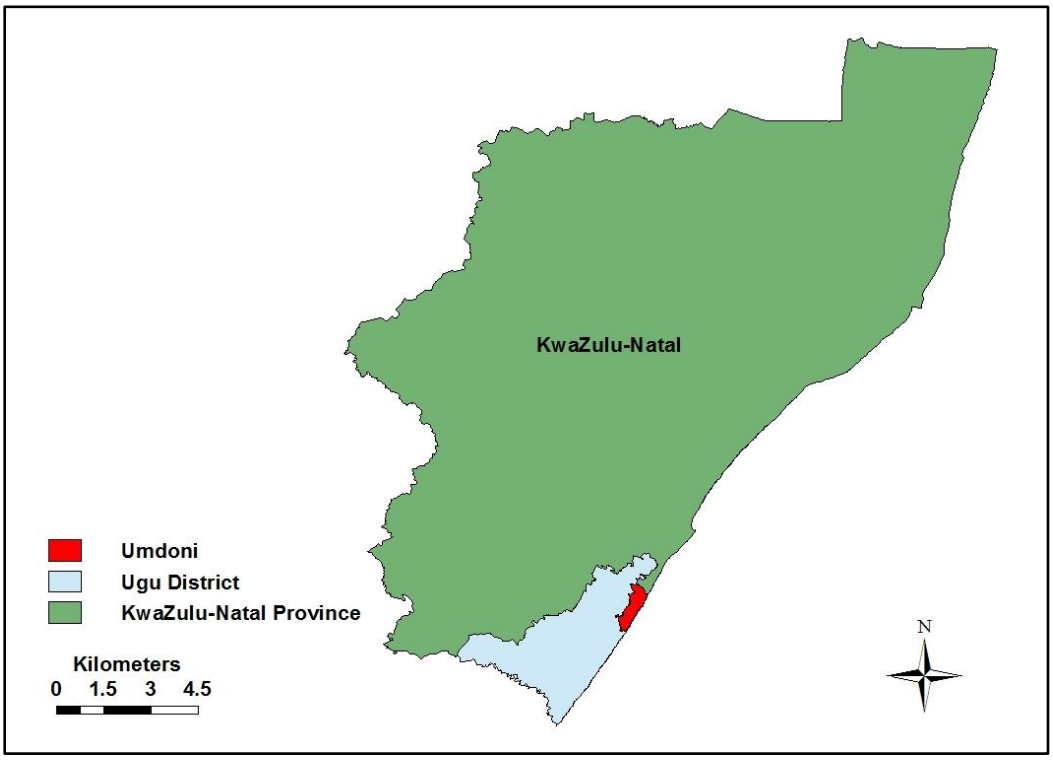


Figure 1.2: Location of Umdoni Local Municipality. Source of shape files: Internet web pages and Ugu District Municipality. Software platform: ESRI ArcGIS 10.1

The 2009 Blue Drop Report (DWA, 2009), indicated that Umzinto Water Treatment Plant (UWTP) was performing positively in terms of regular submission of drinking water quality results to Department of Water Affairs and Forestry (DWAF), complying with the South African National Standard (SANS241) and Drinking Water Quality Failure Response Management. These indicators among others mentioned in the report showed that the plant could attain Blue Drop status (DWA, 2009). Although the report is a generally useful tool for measuring compliance with drinking water quality standards, it however, does not give a full picture of the challenges faced by a treatment plant. For example, problems such as efficiency and effectiveness for relatively small treatment plants are not addressed in the report; hence the need for optimisation. These problems are seen to be a great challenge with the increasing water demand in such sectors of the society.

Further, some demographic results from a community survey, which was carried out in 2007 for Ugu District Municipality, indicated that the population of the district was predicted to continuously increase up to year 2030 (Ugu District Municipality 2007). Table 1.1 illustrates the population forecast under the three growth scenarios considered in the report. This could potentially exert pressure on the treatment plant's capacity to meet demand for potable water.

Table 1.1: Population and demographics for Umzinto Water Supply areas Source: Ugu District Municipality (2007)

Growth Scenarios		**2008**	**2010**	**2015**	**2020**	**2025**	**2030**
High growth scenario	Head Count	33 875	36 640	44 578	53 587	62 122	70 285
	Rate of Growth		4%	4%	4%	3%	2%
Median growth scenario	Head Count	33 875	35 590	40 267	45 006	49 205	53 008
	Rate of Growth		2%	2%	2%	2%	1%
Low growth scenario	Head Count	33 875	34 556	36 319	37 702	38 846	39 827
	Rate of Growth		1%	1%	1%	1%	0%

As the population increases, raw water abstraction also increases as well as the demand for adequately treated water. Table 1.2 shows the estimated water consumption and demand patterns from 2005 to 2008, for Ugu District Municipality. These trends are generally similar to those shown in the districts Non-Revenue Water (NRW) report (Ugu District Municipality 2011). The increase in water produced results in an increase in real water losses if the volume of water being pumped out of the treatment plant is increased without improving the infrastructure. This is discussed further in chapter 2.

In a 2011 report, the Ugu District Municipality noted that a preliminary assessment was conducted in the area supplied by the UWTP. The assessment showed that total system losses were high. Non-Revenue Water in the supply area was estimated at "approximately 38% of the water abstracted" (Ugu District Municipality 2011). It was concluded that water losses were mainly due to operational inefficiency and deteriorating machinery.

Thus without a well-defined water demand management plan the conditions were expected to worsen unless interventions were put in place to improve operational efficiency in both water treatment and distribution.

Table 1.2: Estimated water consumption for the Umzinto Water Distribution System. Source: Ugu District Municipality (2007)

Year	2005	2006	2007	2008
Raw water abstraction (million m^3/a)	2.92	3.07	3.37	3.57
Raw water abstraction (ML/d)	8.00	8.40	9.24	9.79
Required treated water production (million m^3/a)	2.67	2.81	3.10	3.36
Required treated water production (ML/d)	7.32	7.69	8.50	9.21

1.2 Problem Statement

Population growth poses an inevitable challenge to most nations (Buhaug and Urdal 2013), including South Africa. The experiences of other nations have shown that water scarcity is one of the main reasons for environmental conflict (Böhmelt *et al.* 2013). Beside the conflict brought about by the scarcity of water, another equally alarming problem is the environmental health hazard it poses. Also fuelled by the increase in population, the emanating environmental health hazards can be dangerous as illustrated by the situation experienced in Bulawayo, Zimbabwe (Nyemba *et al.* 2010). A long-term solution can be achieved by implementing a sound technical approach that includes all the factors affecting the availability of potable water, i.e., treatment and distribution.

Ugu District Municipality is no exception to the situation described above. The population in the district has been continuously increasing putting pressure on the water supply (City Brinkoff 2012). This has pushed the municipality to ration water supply, which in many cases has resulted in protests by consumers who do not get enough water. The increase in population in the district has also fuelled an increase in water pollution, which, has a great bearing on the cost of treating the water and the tariff structuring for the consumers, a cost-chain for water services (Dzwairo, Otieno and Ochieng 2010). If the treatment plants are not properly operated and controlled, the increased water pollution will subsequently result in serious environmental health problems. In order to minimise these health

challenges, the water treatment plant has to dose an optimum amount of chemicals to make the water safe for drinking. The process of determining the correct amount of chemicals to dose is complex as it involves many interrelated parameters, of which some are difficult or impossible to measure at regular intervals. Currently (2016), the water treatment plant has selected three major parameters that it is able to measure and use for estimating dosage (Umgeni 2015). The major challenge with this method is that it often results in underestimation or overestimation of the dose. This study therefore sought to develop a model that optimises chemical dosage at a water treatment plant. The study also assessed the levels of water loss in the distribution system. This was to done to measure the percentage of treated water that reaches the consumer.

1.3 Justification

When treating water, the primary goal is to produce what is known as potable water. This is water that is biologically and chemically safe, while anything less is unacceptable (Ndwambi 2009). In South Africa and in particular Ugu District Municipality and its environments (Umdoni and other local municipalities), there is intolerance of failure to meet the primary goal when compared to the national standard (SABS 2011). This calls for a need to develop an integrated system that incorporates skilled operators and an optimally designed plant. As much as there is a need for this integration, the water source and the integrity of the distribution system need to be considered equally important (Bartram and Ballance 1996). The primary goal is thus achievable by adhering to a reference design and standard regulations. The other goals of treating water include, producing water that is appealing to the consumer (Ezine 2012) and reaches the consumer at the right quantity and quality (Jeffrey Yang *et al.* 2008).

Research that has been done to improve the efficiency of the Umzinto Water Treatment plant sort to increase the correctness of the current highly manual tests done to determine the dosages. Freese *et al.* (2000), developed a modified jar test which helped to determine the amount of activated carbon to be dosed during water treatment. Although the methods helped to achieve the goals of treating water, the outputs of this study proposed a more automated and faster method of determining chemical dosages. This study further designed a model for analysing the water distribution system.

1.4 Aim

The aim of this study was to synthesize an optimised potable water treatment network and develop a water usage analysis model for the UWTP.

1.5 Specific Objectives

The objectives of the research were to:

1) Identify and quantify the key raw water quality parameters affecting treatment at the Umzinto Water Treatment Plant.
2) Design a genetic algorithm for the potable water treatment process control.
3) Analyse the Non-Revenue Water of the Umzinto Water Distribution System.
4) Develop a model for water usage analysis for assessing Non-Revenue Water as part of the water loss control.

1.6 Scope of the Research

While many models could be created for water quality management, this research centred on the design of a genetic algorithm for optimising dosage at a water treatment plant and development of a model for water usage analysis. Due to financial and time constraints, the research was used two sets of historical data. One set was for the UWTP, which was for the period between July 2006 to June 2013 and the other set was for the distribution system which was for the period between July 2013 to June 2014.

1.7 Methodology

In carrying out the research, the following activities were undertaken:

- **Literature Review**

  An extensive literature review was conducted on the optimisation of water treatment plants and the management of Unaccounted-for Water in both the local and international context. While aiming to learn from applicable methods in the international context, the focus was on local conditions. The literature review revealed the extent, achievements, gaps, issues and current norms of water treatment practices which were used to assess and address present and future needs.

- **Data collection**

  The Ugu District Municipality and Umgeni Water were consulted to enable acquisition of water treatment and usage data. Water treatment data were collected from the UWTP's historical records and water usage data were collected from Ugu District Municipality.

- **Data analysis**

  The data were analysed by trending water quality for algal count, manganese (Mn), iron (Fe), Geosmin, *Escherichia coli*, 2-methylisoborneol, total coliforms, colour, conductivity, turbidity, suspended solids (SS), pH, temperature, total organic carbon (TOC) and alkalinity. Graphs were plotted for a thorough analysis of the trends. This was done to determine the level of complexity of the model to be designed and also to allow for well-informed recommendations. Water usage data were also analysed to enable the formulation of a suitable method for assessing NRW.

- **Model development**

  The model development stage involved selection of key raw water quality parameters and relating them to chemical dosage data. From the relationships, a genetic algorithm model was designed in order to optimise water treatment using the Matlab programming interface (MathWorks 2014). Microsoft Visual Basics 2012 and Microsoft SQL Server 2012 development interfaces were then used to develop the water usage analysis model (Microsoft Corporation 2012a, 2012b).

  After the genetic algorithm model was developed, it was calibrated and validated. The predicted data were compared with actual data from the water treatment plant. Obtaining a Pearson's correlation coefficient close to one showed that the values predicted by the model were close to the actual. Additionally, the root mean square error between the predicted and actual chemical dosage data was used to assess the model's fit. This was expected to be less than 5% of the maximum actual value and less than 20% of the minimum actual value within a given dataset (MathWorks 2011).

For the water usage analysis model, a comparison was done between the results which were obtained using manual calculation of the water balance components and those from the developed model.

1.8 Significance of the Research

The significance of this research includes:

1. a genetic algorithm model that could be used to calculate chemical dosage in water treatment at the Umzinto Water Treatment Plant.
2. Water usage analysis results obtained for the Umzinto Water Distribution System.
3. Model developed to analyse water usage.
4. By designing an optimal water dosage model and efficient water usage analysis model, the end user would benefit from access to clean and safe water.

1.9 Structure of the Dissertation

The dissertation is structured as follows:

Chapter 1: Introduction

This chapter outlines the background and justification for the study. It also sets out the problem statement, the study's aim and objectives, and briefly discusses the methodology adopted.

Chapter 2: Literature Review

The literature review was based on the study's four specific objectives. The chapter begins by reviewing global water quality standards as well as those for South Africa. This is followed by a literature review on optimisation of water treatment plants, applications of genetic algorithms, water usage and non-revenue water.

Chapter 3: Study Area

This chapter describes the UWTP and its operations as well as its distribution system.

Chapter 4 Methodology

This chapter describes the methods which were employed to achieve the research objectives. Details of the methods used to assess performance of the models are also provided.

Chapter 5: Results and Discussions

Findings of the research are discussed in relation to the study's four specific objectives.

Chapter 6: Conclusion and Recommendations

This chapter concludes the dissertation and makes recommendations for future research.

1.10 Summary

This chapter outlined the background to the study, the problem statement, and the study's aim, specific objectives, scope, methodology and significance. The background and justification for the study were in a South African context and the aim of the study was to develop a model for optimising a water treatment plant and for water usage analysis. The research was limited to UWTP. The methodology which was followed embraced a comprehensive literature review, data collection, data analysis, model development, and model validation. The following chapter reviews the literature on water treatment, genetic algorithms and water usage.

CHAPTER 2 - LITERATURE REVIEW

2.1 Introduction

This chapter reviews the literature on water treatment and the analysis of its usage. Section 2.2 covers water treatment in general. It first compares the standards and country commitments to producing water that is safe for drinking. It then goes on to examine the challenges faced by South African water treatment plants. The last part of this section describes developments in water treatment optimisation. The literature reviewed in section 2.2 together with section 2.3, which discussed raw water quality parameters was essential for achieving objective 1. Section 2.4 focuses on genetic algorithms and their principles. In section 2.5, their application in water resource systems is reviewed. This was essential for development of the model using genetic algorithms in Objective 2. Section 2.6 reviews the literature on water usage analysis, focusing on the variation of NRW in water distribution systems. This was essential for the development of the methods used to achieve objective 3. Objective 4 utilised the literature reviewed in Section 2.6.

2.2 Water Treatment

There have been many changes in water management methods and skills over the years and global developments continue. Most of the changes are a result of lessons learnt from incidents such as outbreak of disease and conflict caused by poor water management practices. In developing countries, poor water treatment and distribution are among factors impeding development; they threaten many lives and cause political unrest (WHO 2006).

Research notes that the quality of drinking water is one of the major health determinants (Rajasekar, Philominathan and Chinnathambi 2006). The water quality itself is dependent on industrial, agricultural and other anthropogenic activities (Singh *et al.* 2009). While these may be necessary activities, they pollute raw water sources, mainly rivers, which are dynamic. To address this problem, municipalities have to abstract water and put it through treatment processes that make it clean, usable, pleasant and safe to use and drink.

The treatment process is a complex and costly activity. DeZuane (1997) noted that the cost of treating or obtaining good quality water is increasing. This was confirmed by the analysis

of global water tariffs between June 2010 and June 2011 in Zetland (2011), which estimated that water and wastewater tariffs increased by 6.8% during this period. Tariffs depend on labour costs, electricity overheads and the treatment technology used to abstract water (Dzwairo, Otieno and Ochieng 2012). In order to address this situation, standard regulations for raw and potable water are required. In addition to this, sound technical methods to treat and distribute water should be formulated.

The role played by water quality standards and guidelines is crucial in deciding on the treatment methods. Research on analysing the effectiveness of the South African drinking water standards compared them to the water quality standards from the Netherlands (NL), together with the guidelines from World Health Organisation (WHO) and the European Union (EU) (Mamba, Rietveld and Verberk 2008). As shown in Tables 2.1 and 2.2, they compared the standards' maximum limit for health related chemical parameters.

Table 2.1: Health related water quality chemical parameters (organic). Source: Mamba, Rietveld and Verberk (2008)

Determinant	**Unit**	**WHO max limit**	**EU max limit**	**NL max**	**SA max limit**
Polycyclic	µg/ℓ		0.1	0.10	Not
Trihalomethanes	µg/ℓ	Not mentioned	110	25	200
Polychlorinated	µg/ℓ	Not mentioned	Not mentioned	0.10	Not
PCBs (sum)	µg/ℓ	Not mentioned	Not mentioned	0.50	Not
Pesticides	µg/ℓ	Not	0.1		
Pesticides (sum)	µg/ℓ	Not mentioned	0.5	0.50	Not mentioned
Tetra- and tri-	µg/ℓ	40	20	10	Not
Vinyl chloride	µg/ℓ	0.3	0.5	0.50	Not
Dissolved organic carbon (DOC)	mg/ℓ	Not mentioned	*	*	10

Table 2.1 shows that, South Africa does not have maximum limits for most organic parameters such as Polychlorinated biphenyls, while the Netherlands, EU and WHO have limits for these parameters. This lenience in standards is also observed in Table 2.2, which shows inorganic parameters. For some inorganic chemicals, South Africa grants an allowance by which the maximum limit can be exceeded. This poses a danger if not properly monitored. Companies can easily abuse the allowance. Mamba, Rietveld and

Verberk (2008), suggested that such differences in strictness might explain why the Dutch are nicknamed the “water people”.

Table 2.2: Health related water quality chemical parameters (inorganic). Source: Mamba, Rietveld and Verberk (2008)

Determinant	Unit	WHO max limit	EU max limit	NL max limit	SA max limit
Aluminium	µg/ℓ	200	200	200	300
Ammonia	µg/ℓ	No guideline	500	200	1000
Antimony	µg/ℓ	5	5	5	10
Arsenic	µg/ℓ	10	10	10	10
Bromate	µg/ℓ	Not mentioned	10	1*	Not mentioned
Chromium	µg/ℓ	50	50	50	100
Copper	mg/ℓ	2	2	2	1
Iron	µg/ℓ	300	200	200	200
Lead	µg/ℓ	10	10	10	20
Manganese	µg/ℓ	500	50	50	100
Mercury	µg/ℓ	1	1	1	1
Nickel	µg/ℓ	20	20	20	150
Sodium	mg/ℓ	200	200	150	200
Zinc	mg/ℓ	3	Not mentioned	3	5
Chloride	mg/ℓ	250	250	150	200
Cyanide	µg/ℓ	70	50	50	50
Fluoride	mg/ℓ	1.5	1.5	1.1	1
Sulphate	mg/ℓ	500	250	150	400
Selenium	µg/ℓ	10	10	10	20
Nitrate	mg/ℓ	50 (as total N)	50	50	10 (as total N)
Nitrite	mg/ℓ	50 (as total N)	0.5	0.1	10 (as total N)

Momba, Tyafa and Makala (2004) reported that rural water treatment plants in the Eastern Cape of South Africa were failing to provide potable water to their consumers. Five sites in the Alice water distribution system, Ntselamanzi, Davidson Primary School, Victoria East Clinic, Hillcrest Primary School and the University of Fort Hare, were used to analyse both the chemical and microbiological quality of municipal water. The chemical quality results of the water from the treatment plants were generally within acceptable levels, except the values of turbidity which ranged from 1.6 to 4.7 NTU. The recommended limits

for turbidity in drinking water range from 0 to 1 NTU; thus, the high levels observed posed a great risk.

High turbidity levels result in low residual chlorine levels, creating an environment that favours the growth of microbiological contaminants. This is supported in a paper by Obi *et al.* (2008). The paper focussed on the interplay of factors involving chlorine dose, turbidity flow capacity and pH on the microbial quality of drinking water in small water treatment plants. Fifty five rural and peri-urban small water treatment plants and distribution systems in the Limpopo and Mpumalanga provinces were studied. The results showed that current flow capacity is a major determinant of final microbial water quality compared to chlorine dosage which is limited in a number of ways (Obi *et al.* 2008). One of the factors contributing to the ineffectiveness of chlorine is insufficient knowledge of its dosage and the mechanisms promoting coliform bacterial survival (Rand, Gagnon and Knowles 2014). The operators need to be well versed with the terms free, total and residual chlorine. These have been used by various researchers. According to (Casey *et al.* 2012), chlorine demand describes the amount of chlorine that is used up by metals and organic materials when chlorine is fist added to water. This chlorine will not be available for disinfection. When the demand chlorine has been accounted for, total chlorine is what is left. This is further divided into combined chlorine and free chlorine. Combined chlorine being the chlorine that reacts with nitrates and is not available for disinfection. Free chlorine is the one that inactivates the disease causing organisms (Luo *et al.* 2011).

A more extensive survey consisting of 181 water treatment plants was conducted across seven provinces in South Africa (Momba, Obi and Thompson 2009). The survey examined the ownership of the plants, design capacity, the type of raw water sources, the treatment processes and maintenance of the equipment. Microbial and Physiochemical tests were also carried out at the treatment and supply points using standard methods (Momba, Obi and Thompson 2009). The majority of the plants were operating under their design capacity and the questionnaires used in the study confirmed that they were poorly maintained. The low level of skills among the majority of the treatment plant operators resulted in improper dosage of coagulants and disinfectants in the treatment process (Momba, Obi and Thompson 2009). Most plants lacked basic monitoring and control equipment, resulting in

difficulties in managing residual chlorine levels, turbidity, pH and flow rate, all of which are critical for an optimised treatment process.

From these studies it was noted that the observed water treatment plants produced almost similar results and conclusions. This study develops a more optimum algorithm that takes into account the variations that occur along the treatment train.

2.2.1 Blue Drop Certification

This is a legislative requirement established in September 2008 by the Department of Water and Sanitation in South Africa (DWA, 2009). It enables municipalities and other water service institutions to develop suitable monitoring programs. Rivett *et al.* (2014) noted that it is an incentive program that encourages institutions to take a holistic approach to management. The implementation of this regulation has been a success in the water industry (DWA, 2009). It was initiated with the following main objectives:

- "Introducing incentive-based regulation of the drinking water quality management function
- Introducing key requirements for effective and efficient management of drinking water quality by water services institutions;
- Initiating transparency on the actual drinking water quality management performance of water services institutions;
- Provide information to the public on DWQ performance per water supply system (to prevent generalization), and;
- Facilitating closer working relationships between Water Services Authorities and Water Services Providers (where relevant)" (Rivett *et al.* 2014).

In implementing the regulations, Water Services Authorities assess and assign scores or penalties according to specific criteria that ensure that preventive approaches have been taken in providing safe and clean drinking water. The main concerns during evaluation are: water safety plans, process control, efficiency of monitoring programs, credibility of sample analysis, submission of results to the DWA, compliance with SANS 241 (SABS, 2005), failure response management, responsible publication of performance and efficacy of basic DWQ asset management (Rivett *et al.* 2014).

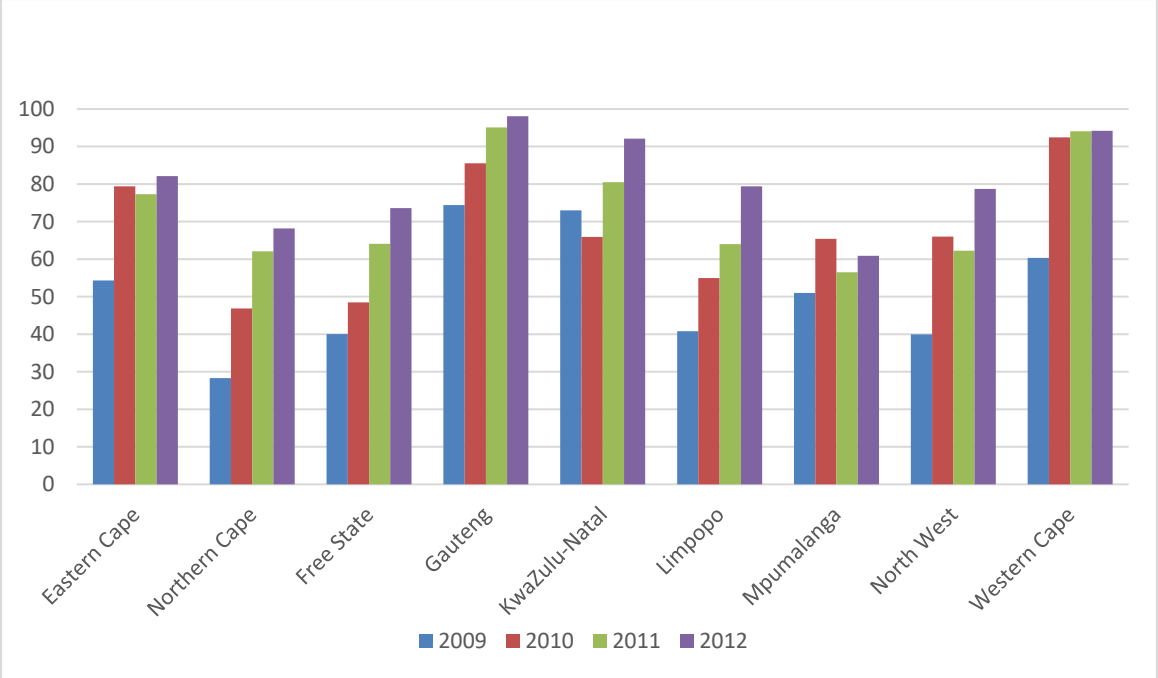


Figure 2.1: Provincial Blue Drop Scores from 2009 to 2012. Source: Rivett et al. (2014)

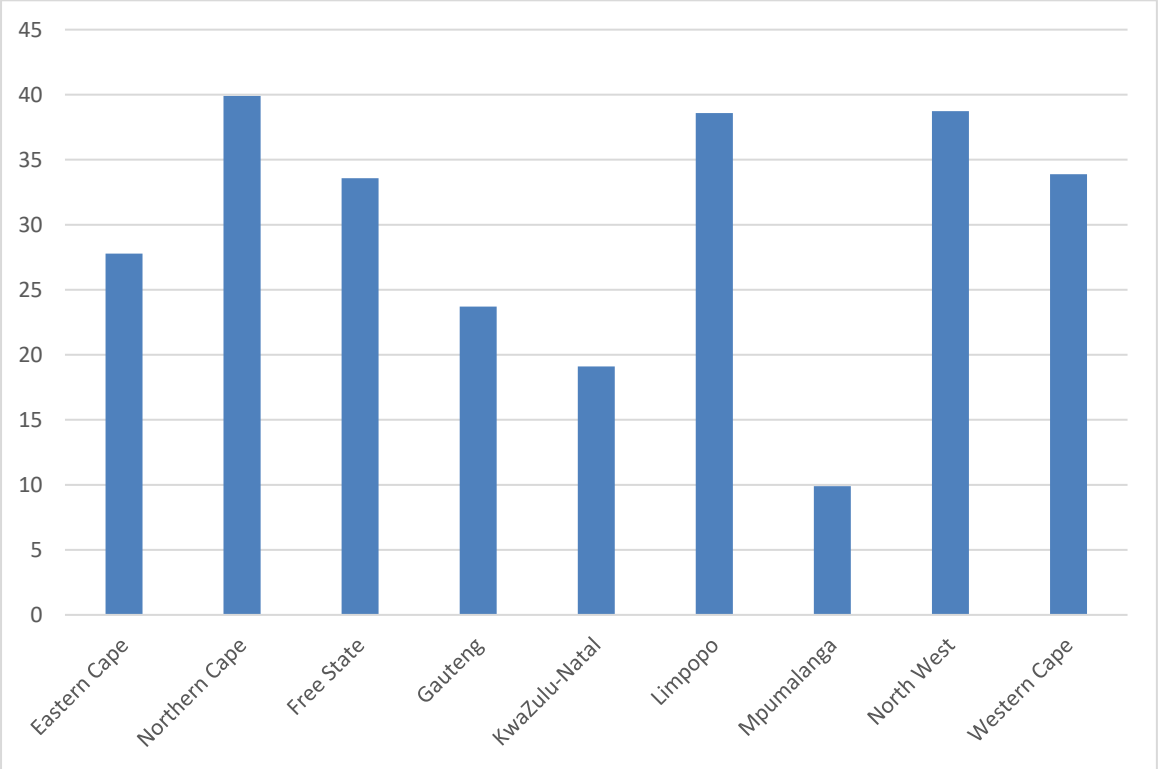


Figure 2.2: Percentage Improvement in Blue Drop Compliance. Source: Rivett et al. (2014)

The 2012 Blue Drop report notes that there has been a general increase in compliance with the regulations across all provinces. Rivett *et al.* (2014) attributed the sharp improvement

in 2010 shown in Figure 2.1 to infrastructure development for the 2010 World Cup. Figure 2.2 shows the percentage improvements.

2.2.2 Optimisation of water treatment

In water treatment, the three main objectives are the production of safe water, producing water that is appealing to the consumer and developing designs and concepts that require low operational costs. These goals are not once-off targets but should be met consistently for as long as a water treatment plant is functional. Achieving these goals requires the consideration of a range of variables ranging from cost, to chemistry, and the operation and operability of the plant; hence the issue of optimisation is essential and should be put to good use. Optimisation goes far beyond producing water that complies with quality standards because other dimensions are involved in the treatment process. For example, if turbidity is within the recommended limits, this does not necessarily mean that the system is optimised. Besides conformity, it is essential to determine the time and cost required to complete a process or to conform. The optimisation process will involve finding the right values of all the dimensions involved. Table 2.3 shows the general expectations when the various modules of water treatment are optimised. The expectations are those set out in the Area Wide Optimization Program (AWOP) (Dowbiggin 2012).

Table 2.3: General expectations for an optimised water treatment design. Source: Dowbiggin (2012)

Process	**General expectations when optimised**
Sedimentation	• When the daily average raw water turbidity is less than 10.0 NTU settled water turbidity should be less than 1.0 NTU 95 percent of the time. • When the daily average raw water turbidity is greater than 10.0 NTU, the settled water turbidity should be less than 2.0 NTU. 95 percent of the time.
Filtration	• An optimised filtration stage should be able to achieve water turbidity of less than 0.1 NTU 95 percent of the time based on the maximum values recorded during 2-hour time increments. • A tolerance of up to 0.2 NTU can be given for any filtered water measurement.
Monitoring Requirements	• It is crucial to continuously monitor and evaluate the daily turbidity of the raw water, settled water and the filtered water • Most utilities do theses checks at 2 hour increments from each filter • One filter backwash turbidity profile is performed each month for each filter
Recommended Instrumentation	• Each filter effluent is prepared such that turbidity is constantly checked and recorded • The pH of raw and filtered water is continuously checked and recorded

Process	General expectations when optimised
	• Plant is equipped with an adequately sized computer for recording and electronically transmitting raw, settled and filtered water data, and for generating turbidity vs. time graphs

Darr and Loflin (2011) reported on the benefits achieved by the Orange Water and Sewer Authority's Jones Ferry Road Water Treatment Plant through the implementation of optimization. These included:

- Water quality improvements
- Operational improvements
- Workforce improvements
- Facility improvements.

A popular approach to the optimisation of water treatment plants involves the optimisation of the individual modules of the treatment process and targeting the correct quality parameters involved at each module (Piri, Homayoonnezhad and Amirian 2010). Rietveld *et al.* (2010) noted that water quality indicators at each module of the water treatment plant need to be determined for good operation. In their research, a 10-step model development and evaluation method was used for numerical modelling of a drinking water treatment plant. The results show that, for each stage, the models developed gave more detailed information about the treatment process. An example is the model employed for the ozonation process. The model showed the changes in water quality values; this gave rise to the recommendation that a CT-value sufficient for disinfection be maintained. Thus, instead of decreasing ozone dosages as in the traditional process, the dosage should be increased (Rietveld *et al.* 2010).

Breese (2006) identified the factors that affect optimisation of coagulation and flocculation in drinking water treatment. The analysis of historical chemical dosing trends was considered the first step in optimisation. These trends assist in understanding how chemical dosage impacts the plant's performance (Piri, Homayoonnezhad and Amirian 2010). The analysis usually involves comparing the mass of coagulant added to a value of water quality indicator, e.g., the relation between raw water turbidity versus coagulant dose. Coagulation is a complex process to optimise as the chemical reactions which are meant to be precise take place in a few seconds. This chemistry is very important and involves the type of

coagulant used, its amount and effect and the overall alkalinity of the system. Further research showed that raw water natural organic matter, coagulant dose and the coagulation pH affected the dewatering properties of the sludge produced in drinking water treatment (Verrelli, Dixon and Scales 2009).

Research has considered optimisation of the sedimentation process as the Dutch's secret on how to produce drinking water without chlorine (Smeets, Medema and Van Dijk 2009). The five-stage method established by the research encouraged optimisation of the physical process of drinking water treatment without relying on chlorination. This technique has improved the efficiency of research and has encouraged joint research in the drinking water sector. Consumers in the Netherlands trust the water treatment methods, with at least 95% of the population drinking tap water (Smeets, Medema and Van Dijk 2009).

Filtration is the key to disinfection. If the treatment is not effective, problems arise with the quality of the potable water regardless of the quantity of disinfectants added. In most treatment plants, filtration is done to remove almost all particles from the water. Optimisation of filtration mainly focuses on the design parameters of the filter with consideration of the processes involved, such as back wash. Breese (2006) incorporated the issue of filter run time in optimisation of filtration. The filter indices representing a product of filtration rate and filtration run were presented.

In recent research, the optimisation of water treatment has increasingly shifted to producing fully automated plants (Piri, Homayoonnezhad and Amirian 2010). In taking the automation approach to optimisation, popular techniques include linear programming (Bagajewicz and Savelski 2001; Castro, Teles and Novais 2009; Platikanov, Martín and Tauler 2012), non-linear programming (Chachuat, Roche and Latifi 2001; Dellana and West 2009), fuzzy logic programming (Pual *et al.* 2001; Carrasco *et al.* 2004; Altunkaynak, Özger and Çakmakci 2005; Joon-Hong and Ju-Jang 2009), artificial neural networks (Adgar, Cox and Böhme 2000; Hyeon *et al.* 2004; Shariff 2004; Shariff, Cudrak and Stanley 2004) and genetic algorithm (Gupta and Shrivastava 2010). In other studies stochastic modelling has been used to incorporate variability in optimisation (Boccelli, Small and Diwekar 2007).

Bagajewicz and Savelski (2001) used a linear programming method for optimisation. The methodology applied is only applicable to processes that contain a single contaminant. Application of other optimisation techniques, such as Pinch Analysis can be very beneficial in optimising systems with multiple contaminants (Jeżowski, Bochenek and Poplewski 2007). In water treatment there is always more than one contaminant involved and the chemicals used need to be properly selected to achieve acceptable parameter values. Castro, Teles and Novais (2009) also applied linear programming in their research but unlike the method employed by Bagajewicz and Savelski (2001), theirs' was able to optimise systems with multiple contaminants. The main drawback in their research was that it demanded much more computationally (Castro, Teles and Novais 2009).

2.3 Water Quality Parameters

Water quality parameters mainly depend on the use of the water. For example, the parameters of concern when monitoring the parameters for drinking water are different from those for monitoring wastewater. Knowledge of water quality parameters is important for two essential reasons. It assists in understanding their effect on the water treatment process and their effect on health.

2.3.1 Algal count

Algae are minute life forms that are generally considered as simple aquatic plants which do not have roots, stems or leaves and have primitive methods of reproduction. This is on the grounds that they utilize photosynthesis to deliver sustenance. Algae live in a wide range of aquatic environments and are a natural component of most aquatic ecosystems. Algae require warmth, sunlight, and nutrients to grow and reproduce.

Some algae have a financial significance since they are a source of carotene, glycerol, and alginates and can be converted into a food source for aquaculture. Algae is beneficial in the simplest manner because algae release oxygen as part of their metabolism, they serve to oxygenate water. Algae also play an important part in the aquatic food chain, as they are the main food source for zooplankton and small fish, which in turn serve as food for larger fish and other wildlife. More algae in the water means that more carbon dioxide is used from the atmosphere and that more oxygen is released into the atmosphere.

Algae are good indicators of the atrophic status of a water body, that is, the degree of pollution and nutrients in that water. When the algae levels are high, the first concern is its impact on the aesthetic properties of water. It typically brings about offensive colour and smell, that is, the water will be aesthetically unacceptable. Literature notes that high levels of algae often result in increased use of powdered activated carbon, a chemical used to treat the toxins produced by some algae species (Dixon *et al.* 2011; Sevilla *et al.* 2012). This directly affects treatment costs. When chlorine is used, it reacts with the algae to produce dis-infection by-products, which poses a risk to cancer to the population that drinks the water (Chowdhury, Rodriguez and Sadiq 2011; Richardson and Postigo 2012; Henry 2013). Another concern is that this reaction tends to increase the amount of chlorine used.

2.3.2 Alkalinity and pH

Alkalinity and pH are words that are often used interchangeably (Kramer 2012). The two do not mean the same nor are they opposites. The pH is an indication for the acidity of a substance. It is determined by the number of free hydrogen ions (H^+) in a substance. Acidity is one of the most imperative properties of water on the grounds that specific concoction procedures can just occur when water has a specific pH. For instance, chlorine reactions only take place when the pH has a value of between 6,5 and 8 (WHO 2006). Water is made up of H+ ions and the hydroxide (OH-) ions. It is considered neutral because the number of H+ ions equals the number of OH- ions. Its pH value is 7. The pH scale ranges from 0 to14, with basic or alkaline substance have pH above 7 and acidic substances have pH below 7. The further the pH lies above or below 7, the more basic or acid a solution is (Larson and Henley 1955).

The South African standards stipulate a pH range between 5 and 9.7, which means that the two raw water sources are within that range (SABS, 2011). This should be a major benefit in the control of lime dosage and disinfection by chlorine. The WHO (2006) stipulates that disinfection using chlorine is more effective at pH values below 8. In the raw water guidelines the stipulated target is between 6.0 and 9.0. At this pH there is No significant effects on health due to toxicity of dissolved metal ions and protonated species, or on taste (DWAF 1996).

Alkalinity is the water's ability to remain at a specific pH when a substance that would make it more acidic or alkaline is added to it. This is commonly known as the water's "buffering capacity." Alkalinity can also be referred to as the capability of water to neutralize acid. Alkalinity is often related to hardness because the main source of alkalinity is usually from carbonate rocks (limestone) which are mostly $CaCO_3$. If $CaCO_3$ actually accounts for most of the alkalinity, hardness in $CaCO_3$ is equal to alkalinity. Since hard water contains metal carbonates (mostly $CaCO_3$) it is high in alkalinity. Conversely, unless carbonate is associated with sodium or potassium which don't contribute to hardness, soft water usually has low alkalinity and little buffering capacity (Wesanekar *et al.* 2014; Wing 2014). So, generally, soft water is much more susceptible to fluctuations in pH from acid rains or acid contamination. In South Africa rivers that have been studied have shown the water to be highly alkaline (Gose *et al.* 2013; Dabrowski *et al.* 2014).This means that the raw water has significant capacity to buffer the pH and would result in low lime dosages.

2.3.3 Colour

Being colourless is one of the most widely recognized qualities of water. This depicts its capacity to permit light from the sun or whatever other sources to go through (Hongve, Riise and Kristiansen 2004). This transparency nature is greatly related to the consumer's perception on the purity of water. In systems where there is a high nutrient concentrations, large algal populations are promoted and these reduce water clarity. With algae being at the top or upper layer of the water, light will not reach the algae at the bottom of the water. This hinders primary production from algal photosynthesis (Shittu, Olaitan and Amusa 2008). Colour can also originate from metallic ions such as iron and manganese, humus, plankton, and industrial wastes. When water contains metallic ions of manganese and iron, it is shown by the reddish – brown colour.

It should, however, be noted that while colour in itself has no adverse effect on human health (DWAF 1996), due to stipulations by the South African Bureau of Standards (SABS, 2011), the colour of portable water should be less than 15^0H in order to be classified as Class I. Colour can impact the suitability of water bodies for recreational activities. Because they reduce water clarity, potentially dangerous obstructions such as rocks, stumps, sand bars, and logs may not be visible to boaters and swimmers.

2.3.4 Conductivity

Conductivity is a measurement of the ability of an aqueous solution to carry an electrical current. Water conducts electricity when it has ions dissolved in it, for example, table salt known as sodium chloride consists of sodium ions (Na+) and chloride ions (Cl-) creates a solution that conducts elsctricity when dissolved in water. This water parameter is used to estimates the total amount of solids dissolved in water -TDS, which stands for Total Dissolved Solids. TDS is measured in ppm (parts per million) or in mg/L.

When conductivity is very high (exceeds 370 mS/m) it causes a significant number of health problems. It can disturb the salt and water balance in children. Heart and renal patients will also face problems with their blood pressure. Besides the health problems electrical conductivity may also cause aesthetic problems such as making the water taste salty and making the water lose its quenching effect.

South African guidelines for raw water (i.e., <170 mS/m) (DWAF 1996). This suggests that conductivity should not greatly influence the treatment costs at the UWTP.

2.3.5 Total Coliforms and *Escherichia coli*

In the raw water quality guidelines total coliforms refer to all bacteria which produce colonies with a typical metallic sheen within 20 - 24 hours of incubation at 35 °C on m-Endo agar (DWAF 1996). Total coliform bacteria are not likely to cause illness, but their presence indicates that your water supply may be vulnerable to contamination by more harmful microorganisms. *Escherichia coli* (*E.coli*) is a member of the total coliform group of bacteria that is found only in the intestines of mammals, including humans. The presence of *E.coli* in water indicates recent faecal contamination and may indicate the possible presence of disease-causing pathogens, such as bacteria, viruses, and parasites. Although most strains of *E.coli* bacteria are harmless, certain strains, such as *E.coli* 0157:H7, may cause illness.

Total coliforms and *E.coli* are used as indicators to measure the degree of pollution and sanitary quality of water, because testing for all known pathogens is a complicated and expensive process. Total Coliform do not necessarily indicate recent water contamination by faecal waste, however the presence or absence of these bacteria in treated water is often

used to determine whether water disinfection is working properly. The presence of Faecal Coliform in well water may indicate recent contamination of the groundwater by human sewage or animal droppings which could contain other bacteria, viruses, or disease causing organisms. high levels of *Escherichia coli* suggest that treatment should be much more effective prior to disinfection otherwise this will result in a high demand for chlorine which in turn increases the treatment costs (DWAF 1996).

2.3.6 Manganese and Iron

Iron and manganese originate from rocks. This mean raw water can only have these metals when it has had contact with rocks containing them for a long time. In most drinking water supplies both iron and manganese are common and are the main causes of metallic taste and the staining. When water is brought out from wells or springs that have high manganese or iron content, staining in not readily apparent. Staining will only occur when the water has been oxidised by the oxygene in the atmosphere.

In the South African guidelines for raw water, the Department of Water and Sanitation (DWAF 1996) stipulates a limit of less than < 2 mg/L for iron concentration. Iron pose risk of a condition called haemochromatosis. This condition is commonly caused by eating acid foodstuffs cooked in kitchenware made of iron for too long. Naturally, water does not contain excessive high concentrations of iron, hence iron poisoning is very rare (DWAF 1996).

The South African standard for drinking water has two limits for the manganese content. One is for the aesthetic properties of the water and the other is for health purposes. The permissible aesthetic limit for manganese is <0.1 mg/L and the one for health is <0.5 mg/L (SABS, 2011). The drinking water guidelines stipulate a target range between 0 and 0.05 to mitigate against both the anaesthetic and health effects of manganese (DWAF 1996).

2.3.7 Total Hardness

The term hardness of water originates from the behaviour of water that has contains a significant level of Ca^{2+}, Mg^{2+} , and HCO_3^- ions. Hardness was originally defined as the capacity of water to precipitate soap. This result from the suds formed when water with Ca^{2+} or Mg^{2+} gets in contact with water. Hardness is measured as calcium hardness or magnesium hardness using an Ion Selective Electrode. The sum of calcium and magnesium

hardness is then known as Total hardness. When classifying water according to Total hardness, it can be described as soft when the total hardness is between 0 and 30 $CaCO_3$ mg/L, moderately soft between 30 and 60 $CaCO_3$ mg/L, moderately hard between 60 and 120 $CaCO_3$ mg/L, hard between 120 and 180 $CaCO_3$ mg/L, and very hard when the total hardness is greater than 180 $CaCO_3$ mg/L.

While South African standards do not set a limit for total hardness, it should be noted that this parameter is significantly linked to the taste of the water and consumer perceptions. Total hardness can also influence domestic uses of water such as laundry and bathing because it facilitates the formation of insoluble salts of long-chain fatty acids, the chief component of soaps. This makes total hardness important to control. This is also noted in the drinking water guidelines and a target range of 50 to 100 mg$CaCO_3$/L is stipulated (DWAF 1996).

2.3.8 Suspended solids

Suspended solids (SS) include all particles suspended in water which will not pass through a filter. High levels of SS decreases the ability of a water body to support a different species aquatic life. This happens because SS absorb heat from sunlight increasing water temperature and subsequently decreasing levels of dissolved oxygen (Lai *et al.* 2013). Photosynthesis also decreases, since less light penetrates the water. As less oxygen is produced by plants and algae, there is a further drop in dissolved oxygen levels. SS can also destroy fish habitat because suspended solids settle to the bottom and can eventually blanket the river bed.

The South African standards and guidelines do not set limits for suspended solids for raw and drinking water (DWAF, 1996; SABS,2011). Therefore, to control suspended solids, it is important to consider its correlation with other drinking water quality parameters (Shroff *et al.* 2015). For nonpoint sources, control measures should be implemented to reduce loading of suspended solids to streams, rivers and lakes. Farming practices such as no-till can minimize soil erosion and help protect water quality. For construction sites, controls such as silt fences and sedimentation basins are designed to prevent eroding soils from reaching surface waters. In urban areas, storm water retention ponds or a regular schedule

of street sweeping may be effective in reducing the quantity of suspended solids in storm water run-off.

2.3.9 Temperature

As expected, raw water temperature tends to show seasonal variations, with highs in summer and lows in winter. There is not much that can be done to control raw water temperature but the effects that it can have on coagulation and disinfection should be taken into account. Studies have shown that low raw water temperatures such as those in Figure 5.26, slow down the process of coagulation and flocculation. This has a major influence on the efficiency of reducing turbidity. Furthermore, studies by Hanson and Cleasby (1990) and Rossini, Garrido and Galluzzo (1999) showed that low temperatures result in the formation of flocs that can easily break up when subjected to shear force.

Veenstra and Schnoor (1980) noted that chlorine can be lost during summer due to radiation. Larger dosages of chlorine are thus required to compensate for radiation loss. High temperatures can also promote algae blooms.

2.3.10 Total organic carbon (TOC)

TOC is a required measurement in municipal water and wastewater systems. It is defined as the amount of carbon found in an organic compound and is often used as a non-specific indicator of water quality. Organic carbon readily binds with other elements in the environment, producing any of a large number of compounds, many of which are harmful to the environment or threaten the public health. In a number of researches it has been seen as a reliable indicator of precursors of disinfection by-products. The South African standard stipulates a limit of 10 mg/L (SABS, 2011). The box and whisker plot, and the temporal analysis show that most of the observations were within this limit. Those that were above would fall within the maximum allowable limit for a given period. Mamba *et al.* (2009) noted that this standard limit is between 10 and 70 mg/L for treated water, which would make this water safe according to this parameter. The raw water guideline states that total organic carbon is the sum of both the dissolved and suspended organic carbon, however, for drinking water purposes the dissolved organic carbon component is of more concern as it has a direct relationship with organic material in water (DWAF 1996). This type of material can pass through coagulation, flocculation, sedimentation and filtration, to the

disinfection stage. If the disinfectant is chlorine, disinfection by-products may be formed and pose a great risk to health (Hrudey 2009; Richardson and Postigo 2012).

2.3.11 Turbidity

Turbidity is greatly related to the colour and suspended solids in water therefore greatly describes the clarity of water. The South African standard limit for drinking water turbidity is 0 to 1 NTU (SABS, 2011). The calculated means and medians therefore show that the water from the two abstraction points is unsafe before treatment. In the water quality guidelines consumption of highly turbid water is shown to noted have no direct health effects (DWAF 1996). However, a major concern is that, high levels of turbidity will reduce the effectiveness of disinfection. They also increase the chances of formation of disinfection by-products, such as trihalomethanes, halogenic acetic acids, chlorine hydrates, chloropicrin, halofuranones and bromohydrins, when using chlorine. These pose a risk of cancer to the consumer. Another concern is the fact that high turbidity inhibits the action of disinfectant on pathogenic bacteria (DWAF 1996).

2.4 Genetic Algorithms

Uncertainty is always present when arranging, creating, overseeing, and working with water resource systems. It emerges because of the quantity of parameters that influence the performance of water resource systems that are not and cannot be known with certainty when a system is planned, designed, built, managed and operated (Loucks *et al.* 2005). Since various parameters are involved, solving or unravelling ecological processes such as the analysis of water usage, and treatment network design requires a multi-objective approach (Dzwairo, Otieno and Ochieng 2012). This calls for the use of bio-inspired methods such as evolution algorithms (EA) (Michalewicz and Schoenauer 1996).

The two main forms of EA are genetic algorithms and evolutionary strategies (Lawrence 1991). The most common are genetic algorithms which were pioneered by Holland and a group of students (Goldberg and Deb 1991). Genetic algorithms emphasize the use of genotypes which should be decoded and evaluated (Whitley 2001). An example is the theories of natural selection and the law of nature, “survival of the fittest” (Jagielska, Matthews and Whitfort 1999).

Significant research has been conducted on water treatment network design using nonlinear programming (Franceschi *et al.* 2002) and mass integration (Huang *et al.* 1999). While these studies have made important contributions to water treatment, they also have shortcomings. These include:

- Nonlinear programming models require the use of an initial minimum value which is almost impossible to find in a real system that contains multiple variables affecting water usage and optimization of the treatment process (Wang and Smith 1995).
- The more manual approach used in the nonlinear models means that the results depend on the user's experience and expertise.
- A large number of factors characterize the quality of water and not all are affected by the same pollutants (Zinkus, Byers and Doerr 1998).
- The available models for mass exchangers or waste interceptors cannot be used to describe all water-treatment processes(Bartram and Ballance 1996).
- All water-using units cannot be modelled as sinks or generators in the mass integration framework.

2.5 Application of Genetic Algorithms in Water Resource Systems

Genetic algorithms have become increasingly popular in different areas of optimisation, from job scheduling (Skinner *et al.* 2013) to control design (Messai *et al.* 2011), among other applications. However, this study focuses on their various applications in the water industry. Selected applications include:

- Surface Water Hydrology
- Water and Wastewater Treatment
- Water Distribution System.

2.5.1 Surface water hydrology

Surface water hydrology focuses on watershed planning and management and instream management of flows (Nicklow *et al.* 2009). These aspects are unique in surface-water flow and its interaction with the climate, topography and local soils.

Genetic algorithms were first used in watershed-level planning by Yeh and Labadie (1997) in a paper that applied multi-objective genetic algorithms to provide a robust and efficient means of generating non-dominated solutions for system costs. Far, Yumei and Zeng-chuan (2001) used a decision-making genetic algorithm to solve a multi-objective problem in watershed planning. In a case study involving the City Lake watershed in North Carolina, Harrell and Ranjithan (2003) applied genetic algorithms to identify detention pond designs and land use allocations. The results of their study showed cost improvements when simultaneously considering the decisions for pond locations and sizes, and land management (Harrell and Ranjithan 2003). Perez-Pedini, Limbrunner and Vogel (2005) used a distributed hydrologic model combined with a genetic algorithm to determine the optimal location of infiltration-based best management practices for storm water management. A Pareto frontier that described the trade-off between the number of best management practises for project cost and watershed flooding was developed (Perez-Pedini, Limbrunner and Vogel 2005).

Muleta and Nicklow (2005) used genetic algorithms combined with other computational methodologies to control non-point source pollution from watersheds. In their paper, genetic algorithms are used for two specific purposes. Firstly, a genetic algorithm is combined with the U.S. Department of Agriculture's Soil and Water Assessment Tool (SWAT) for single objective evaluations and, secondly, for the calibration of the simulation model (Muleta and Nicklow 2005). The model developed by Muleta and Nicklow (2005) can be applied for both watershed-scale and field-scale analysis. In addition, the use of a trained ANN instead of SWAT to identify the final land use patterns resulted in an 84% reduction in CPU (Central Processing Unit) time (Muleta and Nicklow 2005). The mathematical formulation of the model included the transition constraints and the crop management constraints as shown in Equation 2.1.

$$\text{Maximize } Z = \frac{Y_2 - Y_1}{P_2 - P_1} \qquad \text{Equation 2.1}$$

Where

$$Y_{(1,2)} = f(H, C_s, X_s, M_s, t, s)$$ Equation 2.2

$$P_{(1,2)} = f(H, C_s, X_s, M_s, R, t, s)$$ Equation 2.3

$$g(C_s, X_s, M_s, t, s) \leq 0$$ Equation 2.4

In the equations, *Z* represented the function to be maximised; the average annual sediment yield at the outlet of the watershed over the past three-year decision period is represented by *Y, P* represents net average annual economic benefit over the watershed. It should be noted that *Y* and *P* are followed by subscripts 1 and 2 which correspond to values that result by covering the alternative solutions by the most environmentally favoured land use and management practices and options that generate the best net agricultural profits, respectively. *H* represents a set of hydrologic response units that were created for the conservation program. C_s represents the cropping pattern while M_s the landscape management practices for the target watershed implemented during season *s* of year *t*. X_s is the generic term that represents all other hydrologic and hydraulic factors that may affect the sediment rooting and crop yield during season *s* of year *t*. The average market price for crop *C* over the predefined decision period is represented by *R.*

Arabi, Govindaraju and Hantush (2006) also integrated computational methods to develop a methodology that could result in the economical allocation of landscape practices. They used SWAT, representation of best management practices, an economic component, and a genetic algorithm-based spatial search procedure. The implementation of this methodology at two small watersheds in Indiana showed that it was more effective to select and place best management practices by optimization than targeting pollutant loads (Arabi, Govindaraju and Hantush 2006).

In watershed planning, genetic algorithms can also be combined with geographic information systems. Park *et al.* (2006) paper focussed on the design of a water quality monitoring network in a large river system. When the methodology was applied in the Nakdong River system in Korea, the results showed that the area's monitoring network needed to be carefully examined.

Yandamuri, Srinivasan and Murty (2006) used a non-dominated sorting genetic algorithm, to develop a model for optimal waste load allocation. This multi-objective approach focussed on cost, waste discharge and the performance measure that reflects the dissolved oxygen violation characteristics. The methodology was practically implemented for the Willamette river system in the United States and its results led to the conclusion that the detailed information produced makes the model a powerful decision-making tool (Yandamuri, Srinivasan and Murty 2006).

In the past five years, research has shifted from a focus on decision-making tools to the simulation and prediction of watershed conditions. Wang *et al.* (2007) combined the time delay neural network with a genetic algorithm to simulate and predict run off in watersheds. Using records from Linshan Watershed, Sichuan Province, PRC, Wang *et al.* (2007) were able to train the neural networks for the predictions. The model was effective although the results contained a few under-predictions (Wang *et al.* 2007).

Reichold *et al.* (2009) used genetic algorithms to mimic or simulate the predevelopment flow regime. They used the indicator of hydrologic alteration coupled with the range of variability approach in their methodology (Reichold *et al.* 2009). A hybrid algorithm consisting of a combination of a genetic algorithm and the Nelder-Meade approaches was used (Reichold *et al.* 2009). Bekele and Knapp (2010) applied genetic algorithms combined with SWAT in the automatic calibration of a model to forecast the impact of climate change on water supply availability. This approach proved effective when applied on the Fox River, but it did not consider the impact of population growth on water supply availability (Bekele and Knapp 2010).

2.5.2 Water treatment

In water treatment, genetic algorithms have found applications in the design (Cisty 2010), operation (Dürrenmatt and Gujer 2012) and management (Arabi, Govindaraju and Hantush 2006) of water treatment plants. This is the result of intense research in the field of water to make the production of quality water more efficient. As early as 1968 researchers such as Lagvankar and Gemmell (1968) began to demonstrate the interactions of water quality parameters among each other. This non-conservative nature has thus supported the need to engage techniques that can handle such interrelated variability.

While early water treatment research mainly focussed on producing drinkable water, the situation changed in the 1980s when researchers began to examine the efficiency of the process rather than simply the output, with a focus on plant performance in the design and operation of water treatment. Lawler, O'Melia and Tobiason (1980) developed mathematical models to evaluate the performance of each module in the treatment train. The models for flocculation and sedimentation allowed a variety of particle sizes to be considered. The main restriction in this research was the filtration model which could only handle homogeneous suspensions. Despite this drawback, the results obtained resembled that of water treatment practise (Lawler, O'Melia and Tobiason 1980).

Ramaley *et al.* (1981) confronted the same problem relating to filtration but showed that design variables and the parameters of the particles can be used to predict the performance of treatment plants. The results showed a strong relation between filtration, flocculation and sedimentation. A change in the velocity gradient during flocculation, the detention time in sedimentation, and the media size in filtration would result in a significant change in filtrate concentration and head loss development. It was also shown that suspension of different particle compositions required different design parameters (Ramaley *et al.* 1981).

Further research, coupled with technological developments improved the use of mathematical models to optimise water treatment plants (Chang, Brill and Hopkins 1982; Burn and McBean 1985; Banga and Casares 1987). The most common mathematical models were linear programming, non-linear programming and dynamic programming (Izquierdo, Pérez and Iglesias 2004). Also of importance were the environmental decision support systems that were essential in simulating treatment processes (Cortes *et al.* 2000) . Rodriguez and Sérodes (1998) assessed the empirical linear and non-linear modelling of residual chlorine using two different cases of urban drinking water systems. The first case involved the use of a storage tank and the second was the main pipeline of a city's distribution system. The two models assessed were the linear autoregressive model with external inputs (ARX) and the non-linear artificial neural network (ANN) model. From the case 1 results, the ANN seemed to be a favourable model for forecasting. In case 2, the ANN did not show a significant advantage but the size of the dataset and errors in the data

could be the cause. This research paved the way for the development of forecasting models for the different steps in the treatment train (Rodriguez and Sérodes 1998).

The number of variables or parameters involved in treating water promoted research into techniques that are able to handle such complexity. As it became more evident that finding the Jacobian or the Hessian was difficult and in some cases impossible when linear and nonlinear programming was used, researchers started to use genetic algorithms (Tsompanakis, Lagaros and Papadrakakis 2002). Since they are based on matheuristics, genetic algorithms do not rely on the current point (the Jacobian or the Hessian) but use the value of an objective or parameter at a set point. They thus offer flexibility which allows for modification to closely represent real life systems (Gen and Cheng 2000).

A paper by Karr and Gentry (1993) inspired most of the earliest work on the application of genetic algorithms in water treatment (Alander 1993; Jang and Sun 1995; Sing and Postlethwaite 1997; Huang *et al.* 1999). The paper's main objective was to control pH using a combination of fuzzy logic and genetic algorithms. The pH systems commonly found in industry are nonlinear and rapidly changing. To investigate this, Karr and Gentry (1993) used a laboratory pH system that contained both non-linearities and the changing process dynamics. The general structure of the laboratory system is shown in Figure 2.3.

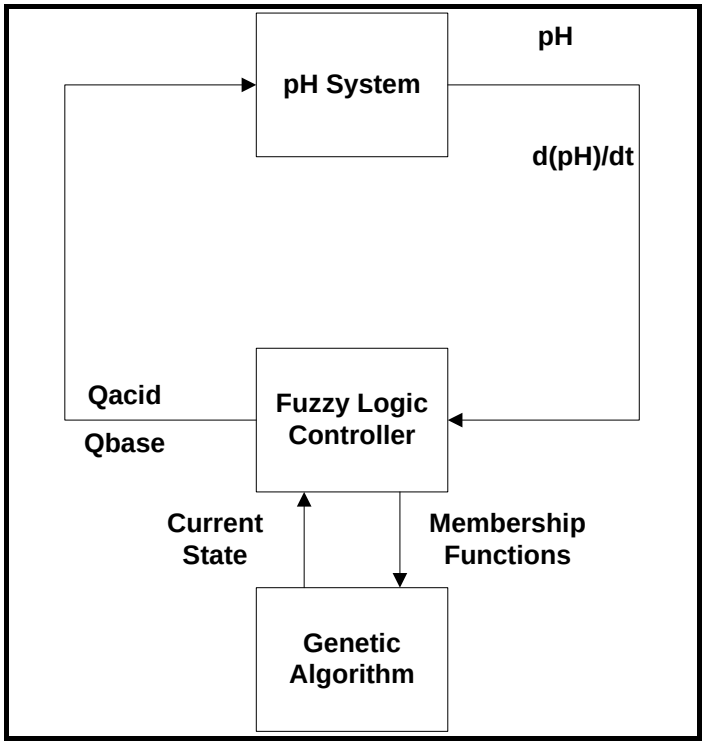


Figure 2.3: Computer Model for a Fuzzy Logic Genetic Algorithm pH system. Source: Karr and Gentry (1993)

Karr and Gentry (1993) found that fuzzy-genetic algorithms offer a powerful alternative to conventional process control techniques.

Following rapid development in multi-objective algorithms, researchers such as Chen and Chang (1998), introduced them in the field of water using biochemical oxygen demand and dissolved oxygen as the parameters for control. Tan *et al.* (2005) used genetic algorithms to model and control a pilot pH plant. The model by Tan *et al.* (2005) used genetic algorithms to establish the Pareto optimal sets (Warburton 1983) and estimated the Wiener-model for the feed forward control. Figure 2.4 shows a diagram of the pH plant. The genetic algorithms would be used to find the optimum parameter for the Wiener-model controller, controlling pumps, valves and flow sensors from Tank 1 and Tank 2.

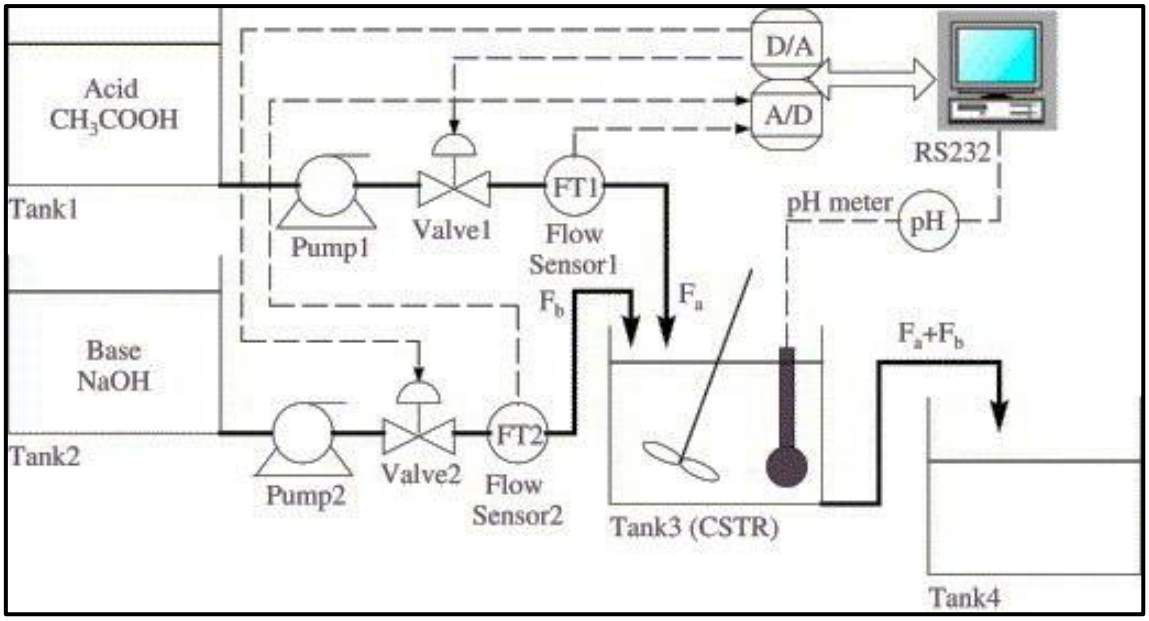


Figure 2.4: Schematic diagram of a pH plant. Source: Tan *et al.* (2005)

The methodology consisted of firstly identifying the inverse titration relationship using genetic algorithms. The result of this inverse neutralisation equation is taken as a component in the Wiener model that linearizes the pH process. This is used to cancel the nonlinearity in the process and then a multi-objective evolutionary algorithm is used to generate the Pareto optimal sets used by the PID controllers. The experimental results showed the effectiveness of the method (Tan *et al.* 2005).

Vasquez *et al.* (2000) presented a method that can be applied to achieve water quality reliability. A genetic algorithm is linked to a first order reliability method. The model was applied in a case study with the aim of estimating the reliability of the system under a given wasteload allocation. The genetic algorithm used an objective function that would minimise the treatment costs and the penalty for not achieving a targeted reliability standard. This method seemed to be more efficient than the reliability models combining genetic algorithms and Monte Carlo simulation developed in the same era (Vasquez *et al.* 2000).

Gupta and Shrivastava (2008) demonstrated the economic impact that parameter uncertainty can have on decision making. Most deterministic models for reliability overlook this aspect. The main objective of their paper was to develop an optimal design of a water treatment plant that would incorporate parameter uncertainty. To achieve this, a genetic algorithm was linked to a Monte Carlo Simulation. In its implementation the

genetic algorithm would first reduce the velocity gradient in the flocculate. The change in the gradient would result in an increase in the floc size, enabling its removal during sedimentation. This is an almost cost free improvement to reliability. The results showed that an approach using a genetic algorithm and Monte Carlo Simulation had great potential in developing optimal water treatment plants under uncertainty (Gupta and Shrivastava 2008).

A later study by Gupta and Shrivastava (2010) also combined a genetic algorithm and Monte Carlo Simulations. A model for reliability constrained optimisation was developed and run as a case study on a water treatment plant. The model worked by minimising the treatment costs subject to the design and performance constraints that are required to achieve a desired level of water quality. The results obtained showed that that reliability can be achieved by control of uncertain parameters (Gupta and Shrivastava 2010).

Other research on water and wastewater treatment includes two case studies by Béraud *et al.* (2007). A multi-objective approach was used to optimise a wastewater treatment plant. The paper concluded that genetic algorithms can be effectively used for desktop optimisation of wastewater treatment plants (Béraud *et al.* 2007).

2.5.3 Water distribution system

In examining the application of genetic algorithms in water distribution systems, the focus should be on the design, construction, rehabilitation, operation, effect of disinfection chemicals, and other cases such as network resilience (Prasad and Park 2004). Design and construction is all about determining where and how many pumps, valves, pipes, bends, etc., are required to ensure minimum costs and the most effective distribution. Over the years, linear and nonlinear programming models have been used for cost effective design of water distribution systems (Lansey and Mays 1989; Duan, Mays and Lansey 1990; Lansey and Basnet 1991). The number of variables and the need to obtain more accurate results has resulted in increased focus on applying genetic algorithms in distribution systems.

The application of genetic algorithms in water distribution was pioneered by Simpson, Dandy and Murphy (1994). Their research used case studies to compare genetic algorithms

to other techniques for pipe optimisation. First, they employed linear and nonlinear programming in a case study. Next, their study employed genetic algorithms and concluded that genetic algorithms would produce an optimal or near optimal solution in fewer evaluations (Simpson, Dandy and Murphy 1994). In 1996, Dandy, Simpson and Murphy (1996) proved the conclusions reached by Simpson, Dandy and Murphy (1994); they published an improved genetic algorithm that had an exponent in its fitness function and used gray coding for decision variables. This research was the solution to the New York tunnels' problem.

Huang *et al.* (1999) produced a mathematical model that could be used in the design of water treatment networks in industry. Two years later, Tsai and Chang (2001) improved the concept by using genetic algorithms. Their paper showed that genetic algorithms can be very beneficial in optimising a network design and usage analysis. The design involved the identification of water using units, treatment units and the mixers. These were taken as nodes and after their implementation in a case study, the benefits were clear. There was no need to have a good estimate for the initial value and fewer evaluations were required to reach the near optimal or optimal solution.

Recent research on the design of water distribution systems used genetic algorithms alongside other techniques to make the process more efficient. Cisty (2010) combined genetic algorithms and linear programming to develop an algorithm that was more stable in obtaining the optimal solution. This was achieved by using genetic algorithms to decompose looped networks into branched networks that could be handled by linear programming to obtain a more dependable global optimum. For almost the same reasons, Haghighi, Samani and Samani (2011) combined genetic algorithms with integer linear programming. Figure 2.5 illustrates the model used. The model starts by selecting the initial parameters for the genetic algorithm. The genetic algorithm will provide a set of pipe diameters for integer-linear programming (ILP). The ILP then reduces the search spaces and sends optimum pipe diameters back to the genetic algorithm. The evolution will continue depending on the values received from the ILP. The case studies in which the technique was used proved that the search time can be significantly reduced and that near optimal designs are achievable (Haghighi, Samani and Samani 2011).

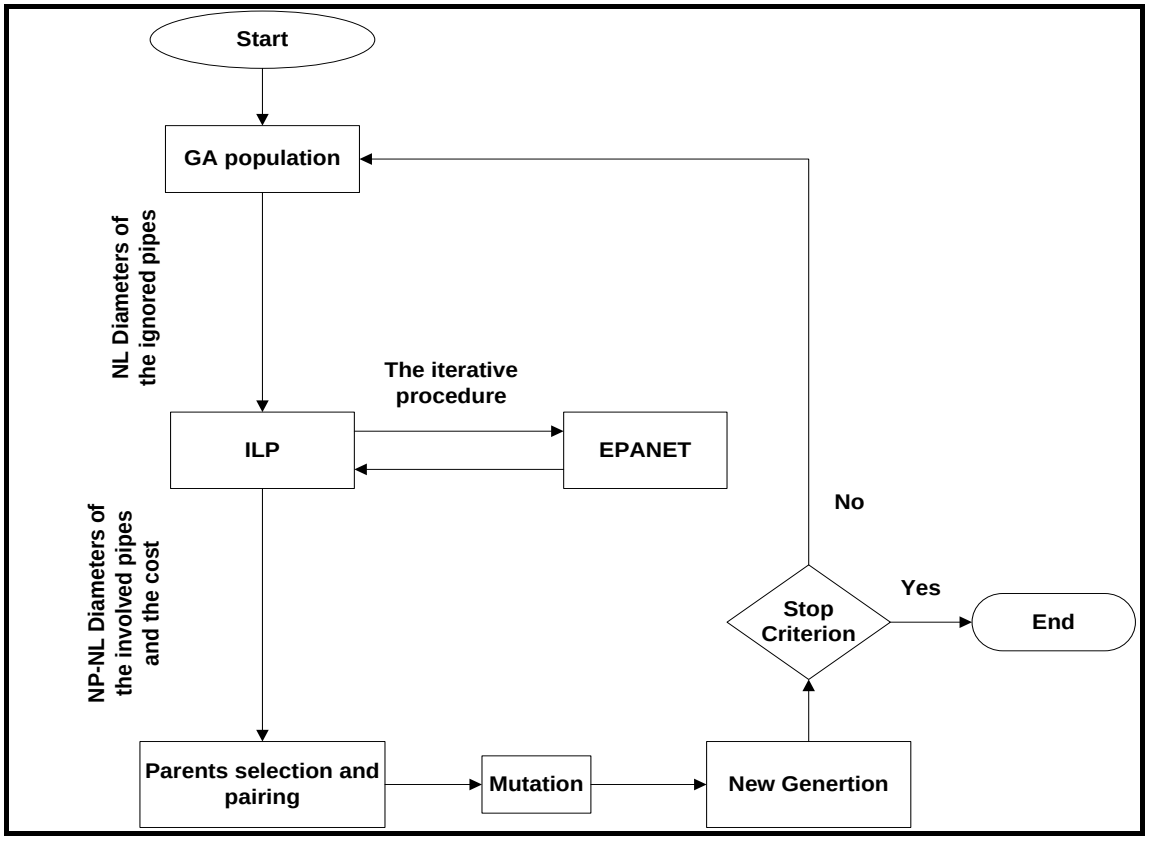


Figure 2.5: GA-ILP algorithm. Source: Haghighi, Samani and Samani (2011)

Wu *et al.* (2011) incorporated energy saving devices in a genetic algorithm model for a water distribution system. The results from the case study show that energy saving devices such as variable frequency drives and variable speed pumps can be well-combined with genetic algorithms. In the case study this resulted in significant cost savings and reduced greenhouse gas emissions in the distribution system (Wu *et al.* 2011).

Turning to the rehabilitation aspect of water distribution, components such as pipes, pumps, and valves will wear out and need to be replaced. The main purpose of using genetic algorithms in rehabilitation is to determine the most efficient sequence of replacing the worn out components and doing so the lowest possible cost. Goldberg and Kuo (1987) conducted some of the earliest research in this regard. They optimised the design of pipe networks to ensure ease of rehabilitation. A year later research focussed on the reliability of the water distribution system was presented with two measures of reliability (Wagner, Shamir and Marks 1988). Their research paved the way for further studies on rehabilitation.

In 1997 , Halhal *et al.* (1997) published a paper in which a messy genetic algorithm was used to address the issue of the optimal sequence for pipe replacement. This form of genetic algorithm is able to solve either a multi-objective problem or a problem with a few variables. With a limited budget, the algorithm was able to choose the best possible network improvements and performed far better than conventional methods of optimisation (Halhal *et al.* 1997).

Alvisi and Franchini (2006) focused on near-optimal rehabilitation scheduling of water distribution systems based on a multi-objective genetic algorithm. The variables were centred on minimising the costs incurred and unserved demand during the rehabilitation process. These were all subject to the funds available for a given period of time. The objective functions were derived for the hydraulic simulator which incorporated pipe failures and nodal demand scenarios. When the algorithm was used in a case study, it was concluded that the procedure has great potential to be a useful tool in minimising costs and unserved demand during rehabilitation. It also enabled the costs to be spread over a number of years and these costs remained within the confines of the available funds (Alvisi and Franchini 2006). In 2009, Alvisi and Franchini (2009) used a multi-objective approach to optimise rehabilitation and incorporated leakage detection in their procedure (Alvisi and Franchini 2009).

2.6 Model Calibration and Genetic Algorithms

Model calibration is a critical phase in the modelling process (Rajasekar, Philominathan and Chinnathambi 2006). This calls for a well-established and efficient calibration strategy. Model calibration involves an estimation of the various constants and parameters making up a model's general structure. Model coefficients and constants are evaluated by comprehension of the model mathematical statement for the parameters of interest subsequent to supplying observed values for both the indigent and autonomous variables. In this research, calibration focuses on developing a water treatment model. This is an environmental model and it hence contains a large number of parameters that are interrelated in a variety of ways.

Wang (1997) pointed out that genetic algorithms can be used as a global searching technique; hence, their potential to be used to solve optimisation problems in which the objective function solutions contain multiple optima and other anomalies. In their research they investigated the practicality of using genetic algorithms in calibrating environmental models in the context of calibrating rainfall-runoff models. The developed models were tested in four case studies and in each case, genetic algorithms proved very effective.

Wu and Sage (2006) used genetic algorithms to calibrate a water loss detection model. In this study, the model calibration is formulated as a nonlinear model that is solved using genetic algorithms. The methodology showcased a framework that can be used to integrate hydraulic simulation and optimisation modelling. The results from the case studies showed that the method can be effectively used to detect water loss as part of the hydraulic calibration of the network model. An observed challenge in this technique is that the accuracy of water loss detection depends mostly on the quality of the measured data (Wu and Sage 2006).

Liu *et al.* (2007) confirmed the usefulness of genetic algorithms in the calibration of models with a large number of parameters as noted by Wang (1997). Liu *et al.* (2007) calibrated a model consisting of 78 parameters. Previous studies had indicated difficulty in full range calibration using all parameters in a diffuse pollution model. In this study, calibration was attained by defining the objective function as the minimisation of the modulus of the difference between predicted values and observed values. The results showed a good fitness and required reasonable computing time (Liu *et al.* 2007).

In more recent work on the use of genetic algorithms in calibration, genetic algorithms have been combined with other optimisation tools to improve the accuracy and computing time in modelling. Kaini, Artita and Nicklow (2012) research integrates genetic algorithms with the SWAT. The approach taken is similar to that of Muleta and Nicklow (2005) in applying genetic algorithms to watershed management. The SWAT tool is used to develop a general structure of the model and the genetic algorithm is then applied for the calibration.

Multi objective genetic algorithms can also be used for model calibration as evidenced by the calibration of a reservoir water quality model (Afshar, Shojaei and

Sagharjooghifarahani 2013). One of the models that has been calibrated is the two dimensional CE-QUAL-W2 water quality and hydrodynamic model. This model is an excellent illustration of models consisting of conflicting objectives. The results of the study showed that multi-objective swarm optimisation could effectively be used to calibrate models that can be employed to predict the physical, chemical, and biological properties of a reservoir (Afshar, Shojaei and Sagharjooghifarahani 2013).

A model for the prediction of chlorine concentration in various hydraulic conditions was developed for a pilot scale water distribution system (Kim, Kim and Koo 2014). The study was driven by the need to maintain a sustainable level of chlorine in drinking water. A genetic algorithm was used to calibrate the parameters of the various models and hydraulics. Regression analysis under turbulent conditions indicated that the fitted parameters from several chlorine models were significantly correlated with Reynolds numbers.

2.7 Water Usage Analysis

The efficiency of a water treatment plant can be described as the extent to which resources such as time, money and energy are utilised for an intended purpose. Mathematically, this can be described in a ratio relating output to input. Generally the higher the ratio-output: input is, the more efficient a system is said to be.

A potable water distribution treatment network is a complex system containing a great number of modules that comprise of pipes, pipe fittings, bends, and valves that must perform in the most efficient way. Most parameters affecting the efficiency of the system are dealt with during the design of the individual modules, mainly with the aim of ensuring that the plant produces the water required at a specific time at the lowest possible cost. This approach is becoming less popular as it lacks the conditions required for accountability and responsibility for what happens after the water leaves the treatment plant. Thus, there is a need to further analyse and develop ways to manage Unaccounted-for Water, the water that never reaches where it is supposed to.

2.7.1 Unaccounted-for-Water

Unaccounted-for-Water is the difference between the total billed water and the volume of water produced at the treatment plant (Motiee, McBean and Motiei 2007). Ideally, the volume of water produced at a treatment plant should match the volume of water that is billed as consumed. Due to a variety of forces in a distribution system and some local factors, different losses occur along the distribution system. This will affect the balance of water in the network (Farley 2001). The water balance in a water network system can be mathematically defined using Equation 2.5.

$$Q_P = Q_C + Q_L \qquad \text{Equation 2.5}$$

Where:

Q_p = Produced volume per time

Q_c = Consumption volume per time

Q_L = Total water losses volume per time

The total water loss can calculated from the relationship shown in Equation 2.6.

$$Total\ Water\ Loss\ (UFW) = WaterProduced - Water\ Billed \qquad \text{Equation 2.6}$$

Farley (2001) manual and Johnson (1996) research showed that Unaccounted-for Water has two main components that need to be managed: (a) physical losses and (b) non-physical losses for example, under-registration by water meters.

Physical losses consist of all types of leakage from the components that make up the water treatment and distribution system. These can be caused by a sudden rupture of a joint or broken pipe. A significant percentage of water can be lost through underground leakage. This is mainly caused by ageing pipes, soil conditions, traffic loading, pipe movement, poor installation practises and electrolysis (Johnson 1996).

Non-physical losses are caused by inaccurate meters, unauthorised use of water from hydrants and poor administrative systems (Abushams 2001). The general motto in water conservation is, ‘every drop counts’. This calls for consideration of other unusual losses

that occur in the system. An example is losses due to recirculation of water in the system, or intersystem connection. Losses due to leakage of reservoirs can also be included (Gomes, Marques and Sousa 2013).

The relationship can be summarised using Equation 2.7, 2.8, and 2.9.

$$Total\ Water\ Loss\ (UFW) = Physical\ Losses\ (real) + Non - Physical\ Losses\ (apparent) \quad \text{Equation 2.7}$$

$$Physical\ Losses = Large\ leakages\ upon\ rupture\ of\ joint + Small\ leakage\ on\ joints + Unavoidable\ leakages \quad \text{Equation 2.8}$$

$$Non - Physical\ Losses = Inaccurate\ meters\ (master, industrial, domestic) + Unauthorised\ use\ (unmetered\ use, illegal\ connections, bypass\ of\ meters) + Unusual\ causes \quad \text{Equation 2.9}$$

In Africa the level of Unaccounted-for Water is as high as 70% (Marunga, Hoko and Kaseke 2006), while Van der Zaag (2003) suggested an acceptable range of 15% to 25%. In the SADC region, Unaccounted-for Water levels are said to be 10% lower than for the rest of Africa (Gumbo 2004), around 11% to 60%. These figures are alarming considering the scarcity of water in Africa and the fact that demand for water increases by 4% to 5% in South Africa each year. This suggests that, if not controlled, there will be continuous water stress within the next half century.

2.7.2 Relationship between Non-Revenue Water and Unaccounted-For-Water

According to Farley *et al.* (2008) and (Bhagwan *et al.* 2014), the average percentage of global NRW is estimated at 35%. This is equivalent to an estimated 48.6 billion cubic metres (BCM) per year of which 32.7 BCM is real (physical) losses (Kingdom, Liemberger and Marin 2006). In addition to water wastage, this costs water utilities worldwide US$14 billion a year. Saving just half of this amount could provide water services to an additional 100 million people without any further investment (Kingdom, Liemberger and Marin 2006; Farley *et al.* 2008).

Liemberger *et al.* (2007) noted that the term Unaccounted-for-Water (UAW or UFW) is usually used to indicate the level of wastage in a water distribution system. While many water utilities use this term, it is open to subjective judgement. This is because it is very dependent on assumptions and poses a risk of manipulation of results. Numerous papers on this subject have been exhibited at research platforms around the globe and all plainly suggest that the term UAW is supplanted with the term NRW which can't be controlled to the same degree (Liemberger *et al.* 2007; Frauendorfer and Liemberger 2010). Figure 2.6 illustrates the difference between UAW and NRW as described in a conceptual framework by Park (2007). The major components within the framework are Total Water Produced, Non-Revenue Water, Water Loss , Unaccounted-for Water, Restoring rate of Water Loss, Identifying rate of Water Loss (Park 2007). Total water produced is a combination of pre-treated water acquired from other utilities and water that is produced by the water utility itself. Total water produced should be adjusted for known errors such as system input metering errors. The NRW is the difference between total water produced and billed water consumption (Revenue Water).

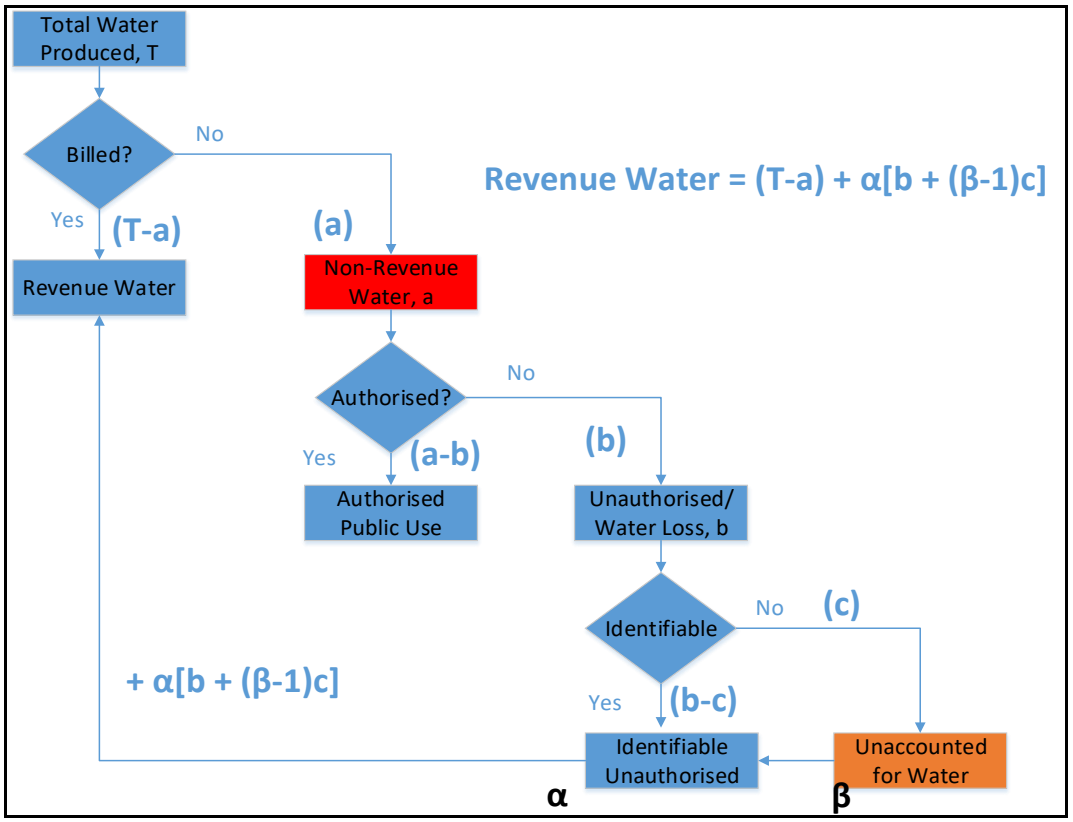


Figure 2.6: Difference between Non-Revenue Water and Unaccounted-for-Water Source:Park (2007)

T: Quantity of Total Water Produced
a: Quantity of Non-Revenue Water
b: Quantity of Water Loss
c: Quantity of Unaccounted-for Water
α: Restoring rate of Water Loss
β: Identifying rate of Water Loss

2.7.3 Water Balance

The water balance refers to the detailed picture of the water produced, imported and exported, consumed and lost. It can be described as a guide that provides information about the water lost as a result of meter inaccuracies, data handling errors, Apparent Loss and Real Loss. It also provides information about the amount of water lost and the reasons for that loss (AWWA 2009).

Traditional Water Balance

Until recently the standard approach to expressing leakage/losses in a system involved taking the water supplied to a system and subtracting authorised use in order to establish total losses which were often termed UAW. This would be represented in the form of a pie chart similar to that shown in Figure 2.7. The approach considers three main components: Authorised Metered, Authorised Unmetered and the remainder which represents all UAW and is often referred to as Real and Apparent Losses.

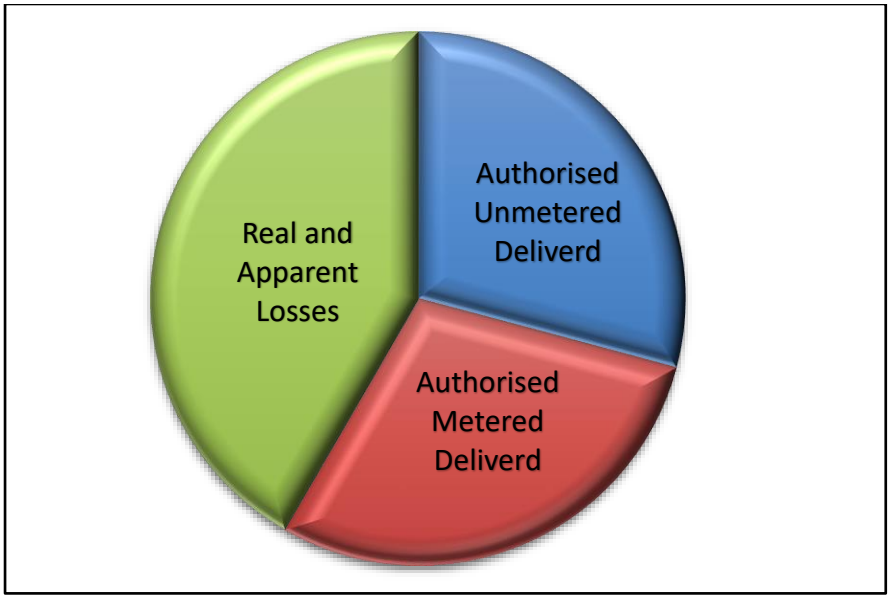


Figure 2.7: Traditional Water Balance

Circular Water Balance

Traditionally losses in the water balance were represented by a large portion consisting of the Real and Apparent Losses. This approach was changed by separating it into smaller segments that could either be measured or evaluated. This enabled a more in-depth understanding of the different components and also of their significance to the overall water balance. Figure 2.8 shows the components in a water balance.

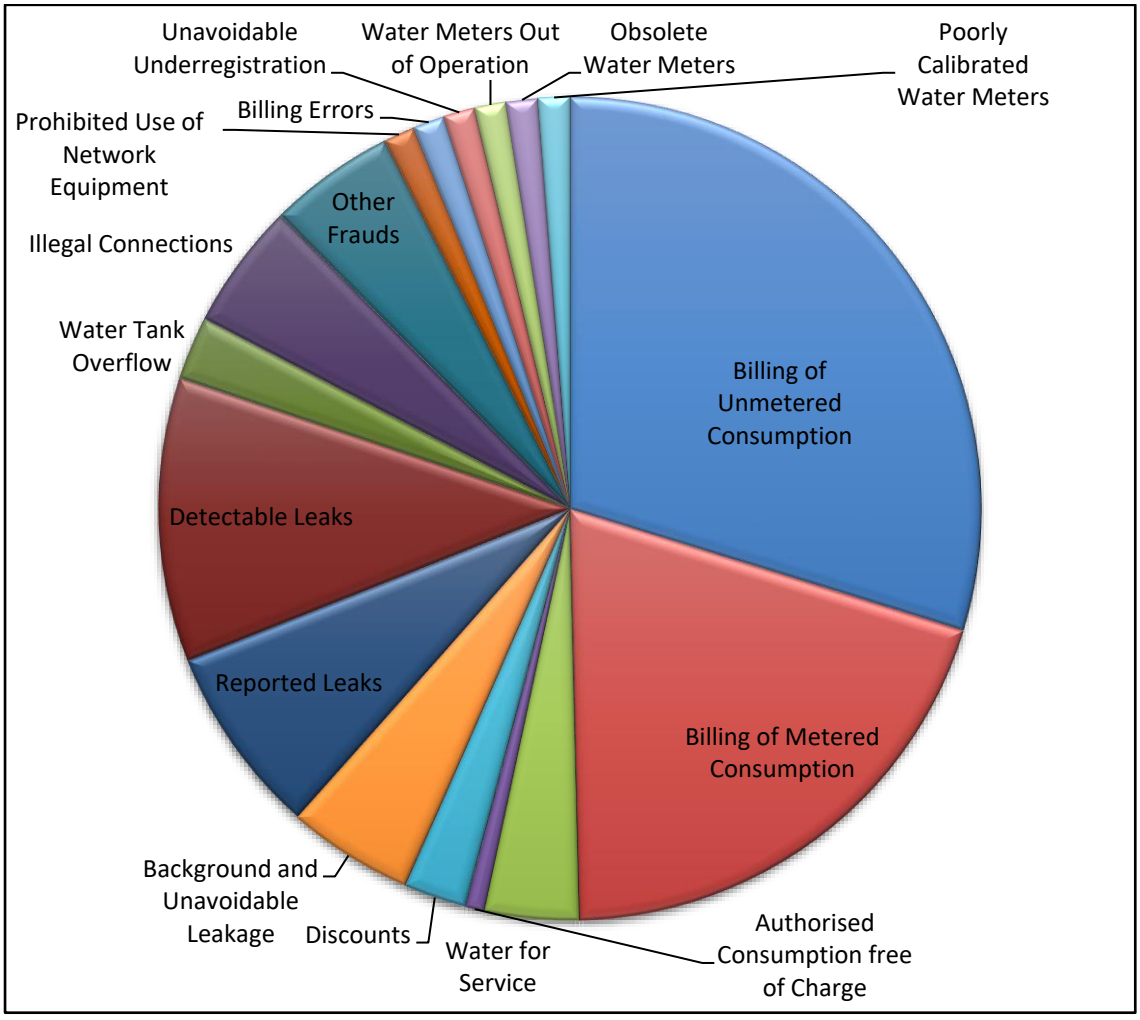


Figure 2.8: Circular Water Balance. Source: McKenzie and Wegelin (2009)

IWA Water Balance

Table 2.4 shows the International Water Association (IWA) standard water balance. The water balance uses terminology developed by the IWA and is the most robust and comprehensive approach (Liemberger *et al.* 2007).

Table 2.4: Standard IWA Water Balance. Source: McKenzie and Wegelin (2009)

<table>
<tr><td rowspan="8">System Input Volume</td><td rowspan="3">Authorised Consumption</td><td rowspan="2">Billed Authorised Consumption</td><td>Billed Metered Consumption</td><td rowspan="2">Revenue Water</td></tr>
<tr><td>Billed Unmetered Consumption</td></tr>
<tr><td rowspan="2">Unbilled Authorised Consumption</td><td>Unbilled Metered Consumption</td><td rowspan="6">Non Revenue Water</td></tr>
<tr><td>Unbilled Unmetered Consumption</td></tr>
<tr><td rowspan="5">Water Losses</td><td rowspan="2">Apparent Losses</td><td>Unauthorised Consumption</td></tr>
<tr><td>Customer Meter Inaccuracies</td></tr>
<tr><td rowspan="3">Real Losses</td><td>Leakage on Transmission and Distribution Mains</td></tr>
<tr><td>Leakage and Overflows at Storage Tanks</td></tr>
<tr><td>Leakage on Service Connections up to point of Customer Meter</td></tr>
</table>

South African Water Balance

Table 2.5 shows the South African Water Balance that is a modification of the IWA standard water balance. From the table it can be noted that non-recovered revenue and free basic water is included. Non-recovered revenue is billed water that ends up not being paid for. This is a universal phenomenon but its high levels in South Africa call for its inclusion in the water balance (McKenzie, Siqalaba and Wegelin 2012). The free basic water policy in South Africa ensures a free supply of 6kℓ per property per month.

Table 2.5: South African Water Balance Source: McKenzie and Wegelin (2009)

<table>
<tr><td rowspan="8">System Input Volume</td><td rowspan="4">Authorised Consumption</td><td rowspan="2">Billed Authorised Consumption</td><td>Billed Metered Consumption</td><td rowspan="2">Potential Revenue Water</td><td>Free basic</td></tr>
<tr><td>Billed Unmetered Consumption</td><td>Recovered revenue</td></tr>
<tr><td rowspan="2">Unbilled Authorised Consumption</td><td>Unbilled Metered Consumption</td><td>Non-Recovered revenue</td></tr>
<tr><td>Unbilled Unmetered Consumption</td><td colspan="2" rowspan="4">Non Revenue Water</td></tr>
<tr><td rowspan="4">Water Losses</td><td rowspan="2">Apparent Losses</td><td>Unauthorised Consumption</td></tr>
<tr><td>Customer Meter Inaccuracies</td></tr>
<tr><td rowspan="3">Real Losses</td><td>Leakage on Transmission and Distribution Mains</td></tr>
<tr><td>Leakage on Overflows at Storage Tanks</td></tr>
<tr><td>Leakage on Service Connections up to point of Customer Meter</td></tr>
</table>

2.7.4 Strategies to reduce Non-Revenue Water

The performance of a water distribution system can be measured by its level of NRW. Worldwide, methods and strategies have been developed and implemented to improve the management of NRW. Lambert and Fantozzi (2005) described the strategies prompted by the IWA Water Losses Task Forces (WLTF).

a. *Strategies to reduce Real Losses*

In most distribution systems Real Losses are the main focus in managing NRW. A variety of strategies are being implemented to help reduce Real Loss levels. Lambert and Fantozzi (2005) proposed a Real Loss management program made up of the four components shown in Figure 2.9.

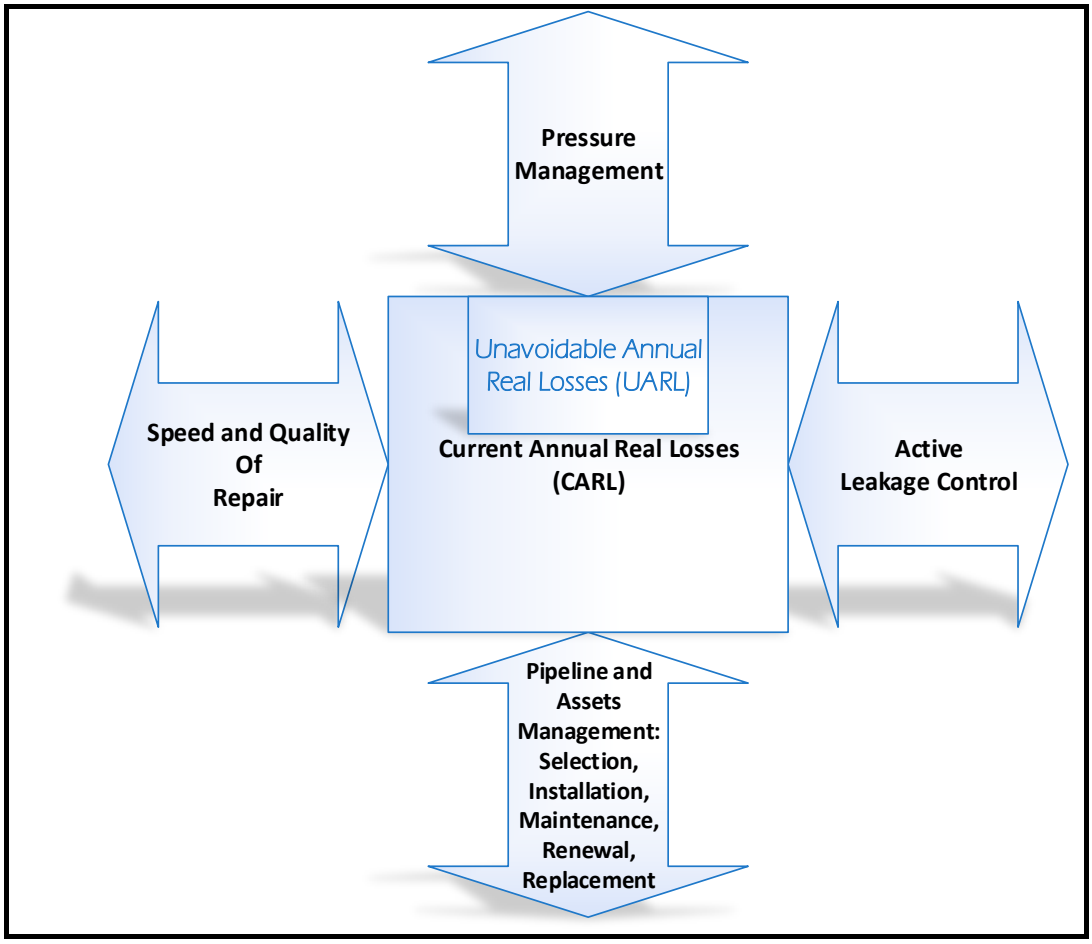


Figure 2.9: Components of a Real Loss Management Program. Source: Lambert and Fantozzi (2005)

As the losses are reduced it is crucial to consider the financial investment made in doing so. The rectangle labelled Unavoidable Annual Real Losses in Figure 2.9 examines what is known as the Economic Level of Leakages (ELL). In terms of this notion, all losses can be reduced to a minimal level which is not economical to reach. The Economic Level of Leakage is therefore the level of intervention for which benefits are equal to the costs associated with them.

b. Strategies to reduce Apparent Losses

Rizzo *et al.* (2007) presented the IWA's recommended apparent water loss control methodology shown in Figure 2.10. In their study they raised the issue of calculating the leakage component for Apparent Losses. This cannot be calculated using a single formula as water losses strongly depend on customer behaviour, the meter type and complexity of the network.

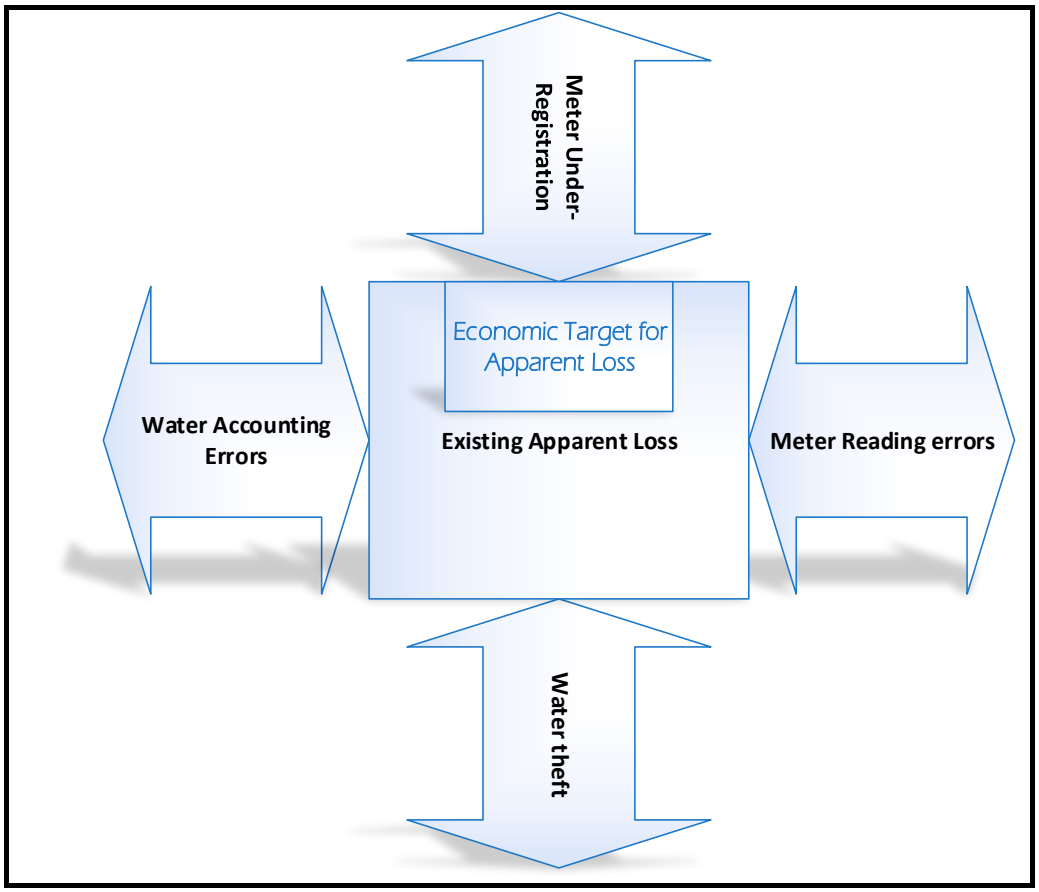


Figure 2.10: Components of an Apparent Loss Management Program. Source: Rizzo *et al.* (2007)

The human factor is clearly significant with regard to apparent losses, including customer behaviour as well as accounting errors and data handling (Rizzo *et al.* 2007).

Strategies that can be developed to minimise physical losses include formulation of a meter replacement policy that is strongly related to meter audits. This should be done in conjunction with investigations on illegal use and internal audits done to evaluate the accuracy of the accounting data (Farley and Liemberger 2005b).

2.7.5 Methods of assessing Non-Revenue Water

The most common methods used to assess NRW are the top-down approach, the bottom-up approach and component-based analysis (Li and Chen 2014). In the top-down approach, the first step is the estimation of the Apparent Losses and unbilled authorised consumption in the system. In estimating the Apparent Losses, unauthorised consumption is assumed to

be between 0.25% and 1%. Mutikanga *et al.* (2010) concluded that unauthorised consumption in developing countries might reach as high as 10% of water sales. Thus, this method is difficult to implement in developing countries.

The bottom-up approach relies on Minimum Night Flow (MNF) analysis. This calls for research to be done in the early hours of the morning when most consumers are not using water. The MNF leakage is then calculated from the difference between the estimated legitimate night uses from the MNF. This can be a challenge when implemented in an area where customers are able to fill water storage tanks during the time of the experiment (Kouziakis *et al.* 2013).

Component-based analysis is a data-driven approach that requires data availability and accuracy in a distribution system. Karadirek *et al.* (2012) recommended that component analysis should not be used on its own when calculating Real Losses due to uncertainties that increase and accumulate with the lack of data availability and accuracy. This can lead to unrealistic results (Mutikanga 2012).

2.8 Optimisation of Water Usage Analysis

Based on mathematical simulation, Almandoz *et al.* (2005) concluded that there is a direct relationship between pressure and Real Losses (leakage) in the water network system, particularly in the night- time flows. In this research the methodology involved measuring the two components of water loss. The components were (a) Physical losses or Real Losses that emanated from the pipe leakages and (b) losses that could not be accounted for from the bills and losses (Almandoz *et al.* 2005).

Giustolisi, Savic and Kapelan (2008) research was similar to that conducted by Almandoz *et al.* (2005) in that in the problem formulation, both studies considered pressure as the main component of physical losses. Mutikanga (2012) listed software application packages among the methods and tools that can be used to optimise water usage analysis.

2.9 Summary

This chapter reviewed the literature relevant to the synthesis of a model for optimising a potable water treatment plant and water usage analysis for the UWTP. It began by examining the commitment and level of success of certain countries, including South Africa, with regard to water treatment and the quality of the water produced. This was followed by a discussion on the theoretical aspects relating to the optimisation of water treatment plants. The methods that have been successfully used in case studies were also discussed. The application of genetic algorithms in the water industry was investigated, as well as the management of NRW and the methods employed to assess losses in the water distribution system. The following chapter focusses on the study area.

CHAPTER 3 – STUDY AREA

3.1 Overview of Umzinto Water Treatment Plant

In this study, the Umzinto Water Treatment Plant (UWTP) provided the data that were used. This treatment plant is owned by the Ugu District Municipality and managed by Umgeni Water which is amongst the leading water utilities in South Africa. Ramjatan *et al.* (2007)'s case study benchmarked the water utility against other utilities in Africa. The benchmarking initiatives used were Deloitte and Touche's Best Company to Work for Survey, the South African Association of Water Utilities (SAAWU) benchmarking initiative, and the Water Utility Partnership (WUP) survey, amongst others. The results showed that the utility was among the leading African utilities. Figure 3.1 shows the location of the UWTP in the KwaZulu-Natal Province.

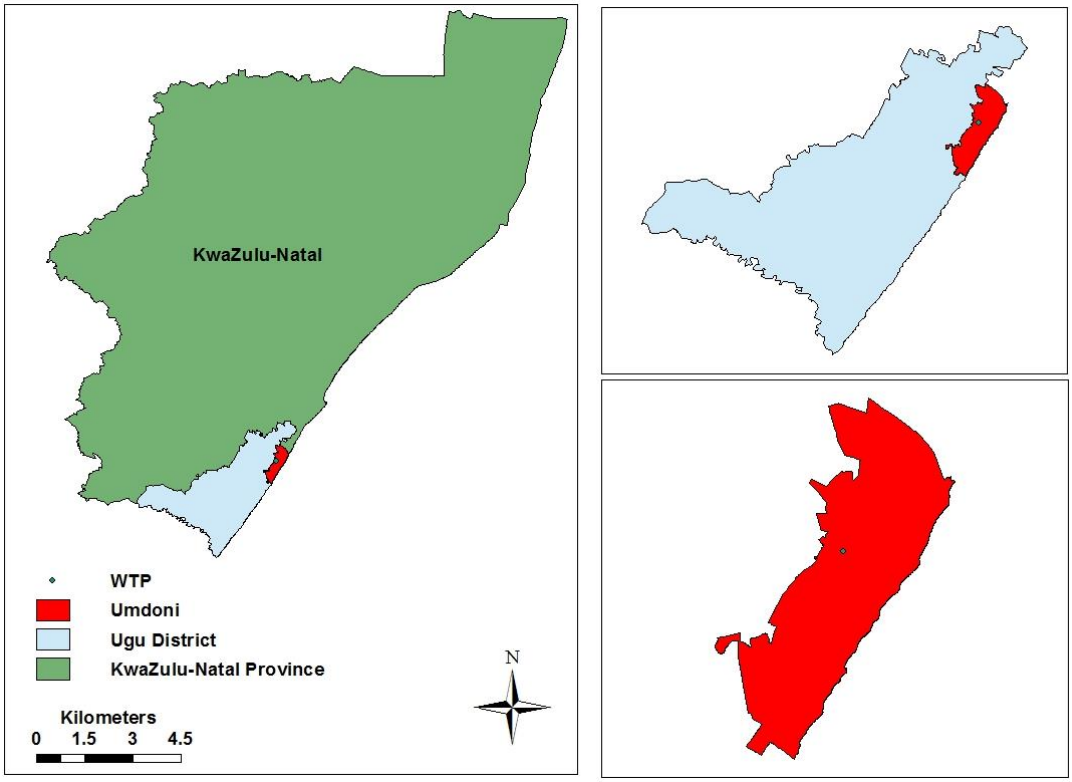


Figure 3.1: Location of the Umzinto Water Treatment Plant in Umdoni Local Municipality. Source of shape files: Internet web pages and Ugu District Municipality. Software platform: ESRI ArcGIS 10.1

The UWTP is located in the Umdoni area which has a population density of more than 260 persons per km^2. The population of the region has been increasing at a rate of 2.8%

per annum since 1996 and in 2007 it surpassed 70 000 (Brinkoff 2012). Figure 3.2 shows the general increase in the population as shown in the data from a website, (City Population2012). This has been a major driver of increased demand for portable water.

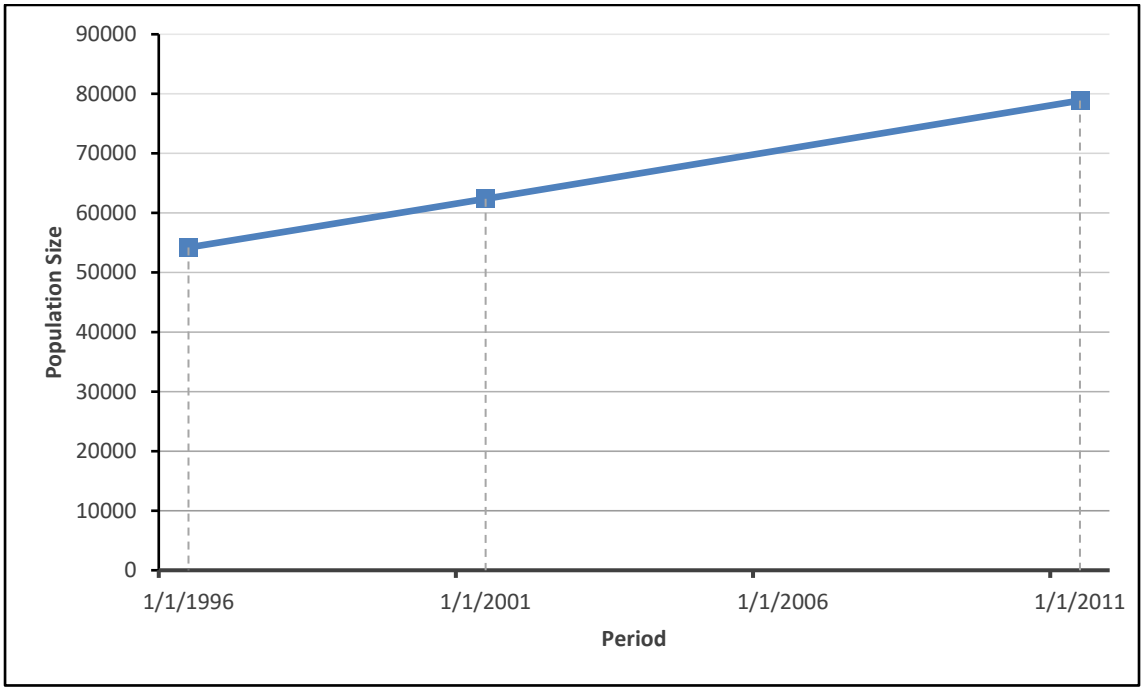


Figure 3.2: Population increase in Umdoni Local Municipality. Source of Data: City Brinkoff (2012)

A map showing the location of the UWTP and its raw water sources is presented in Figure 3.3.

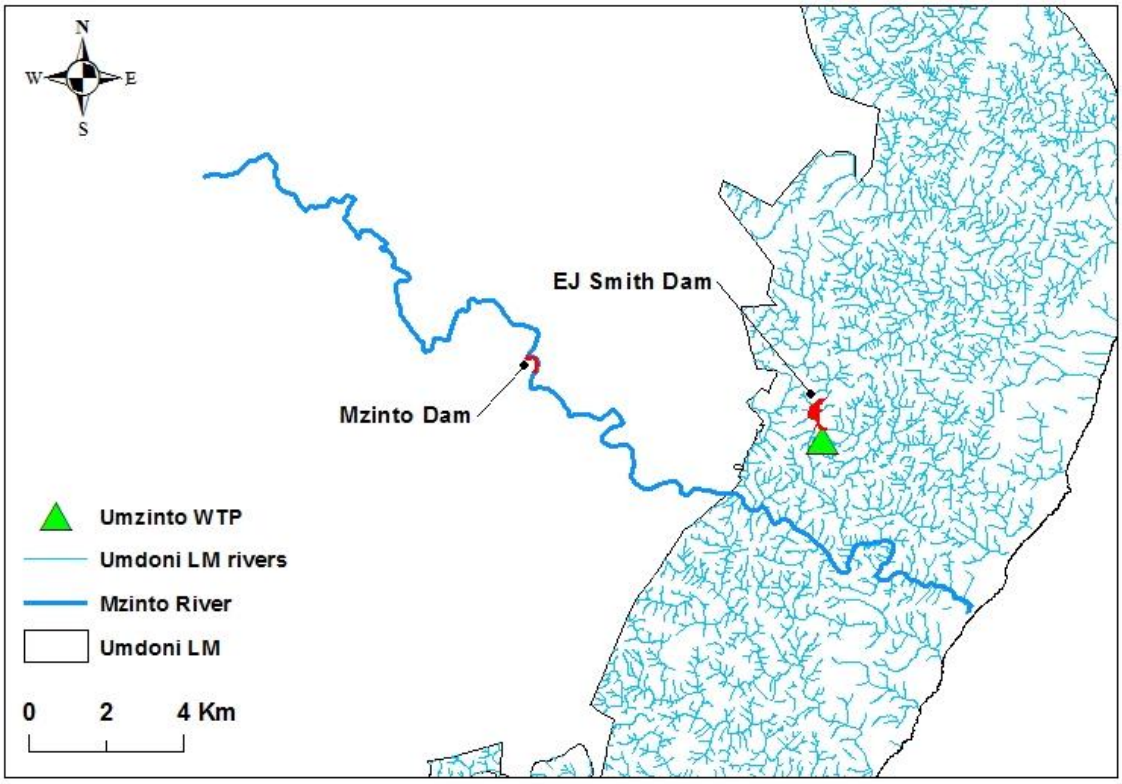


Figure 3.3: Location of the Umzinto Water Treatment Plant and its raw water sources. Source of shape files: Department of Water and Sanitation and Ugu District Municipality. Software platform: ESRI ArcGIS 10.1

The treatment plant abstracts its raw water from two sources, which are the EJ Smith Dam in the Mzimayi River and the Esperanza weir in the Umzinto River. The EJ Smith Dam has a raw water abstraction capacity of 5.0 ML/day (1.8 million m^3/a) that is usually of poor quality with the main contaminants being manganese, iron and *Escherichia coli*. The capacity of the Esperanza weir is 10.0 ML/day (3.6 million m^3/a) and its quality is better than that of the EJ Smith Dam (Ugu District Ugu District Municipality 2011).

3.2 Activities Requiring Water Supply

The water treatment plant supplies a number of sectors and its development is centred on their activities. Economic activities in the area include manufacturing, agriculture, tourism and retail and the Government.

The manufacturing sector comprises of companies and their employees. Companies require water for their operations. They also employ the majority of people in the region that need to use water and will pay their bills. The Government and the public service

also employ a significant percentage of the population and own a number of buildings that require a water supply. Most of the region's income comes from this sector. Sugarcane is the most important crop in the agricultural sector that also requires water and employs people within the region. Tourists visit the area for its coastline and beaches. Tourism is one of the cornerstones of the local economy and has great potential to expand. Office blocks, commercial developments, shops and restaurants also require water.

3.3 Plant Layout

The plant layout of the UWTP is shown in Figure 3.4.

Figure 3.4: Satellite image of the Umzinto Water Treatment Plant Source: Google Earth (2014)

The plant has the following four basic components:

1. *Flocculation channels*: The raw water comes from the EJ Smith Dam and the Esperanza weir. It flows through the flocculation channels as the chemicals are added. Polyelectrolyte dosing takes place, resulting in coagulation to form the flocs.

2. *Clarification (sedimentation) tanks*: These are responsible for settling the flocs that form during the flocculation process. This is done under gravity. The settled sludge is removed by frequent de-sludging of the tanks and sent to the sludge lagoons where the sludge is dried while the supernatant water is discharged back into the river. The plant was not designed for recycling of the wastewater.

3. *Rapid Gravity Sand Filtration*: The sand filters polish the clarified water under gravity. This process prepares the treated water for chlorination.

4. *Disinfection*: The filtered water then gravitates to the chorine contact tanks where chlorination takes place before the potable water is pumped to the command reservoirs in the Umzinto scheme area for distribution.

The hydraulic design capacity of the water treatment plant is only 13.6 ML/day or 5.0 million m^3/a (WSDP, 2008). The average annual flow rate of the treatment plant is estimated to be 9.1 Ml/day or 3.3 million m^3/a based on a peaking factor of 1.5.

3.4 Process Layout and Technologies Employed

The UWTP abstracts raw water from two different sources. The water then goes through the conventional treatment processes illustrated in the process layout in Figure 3.5.

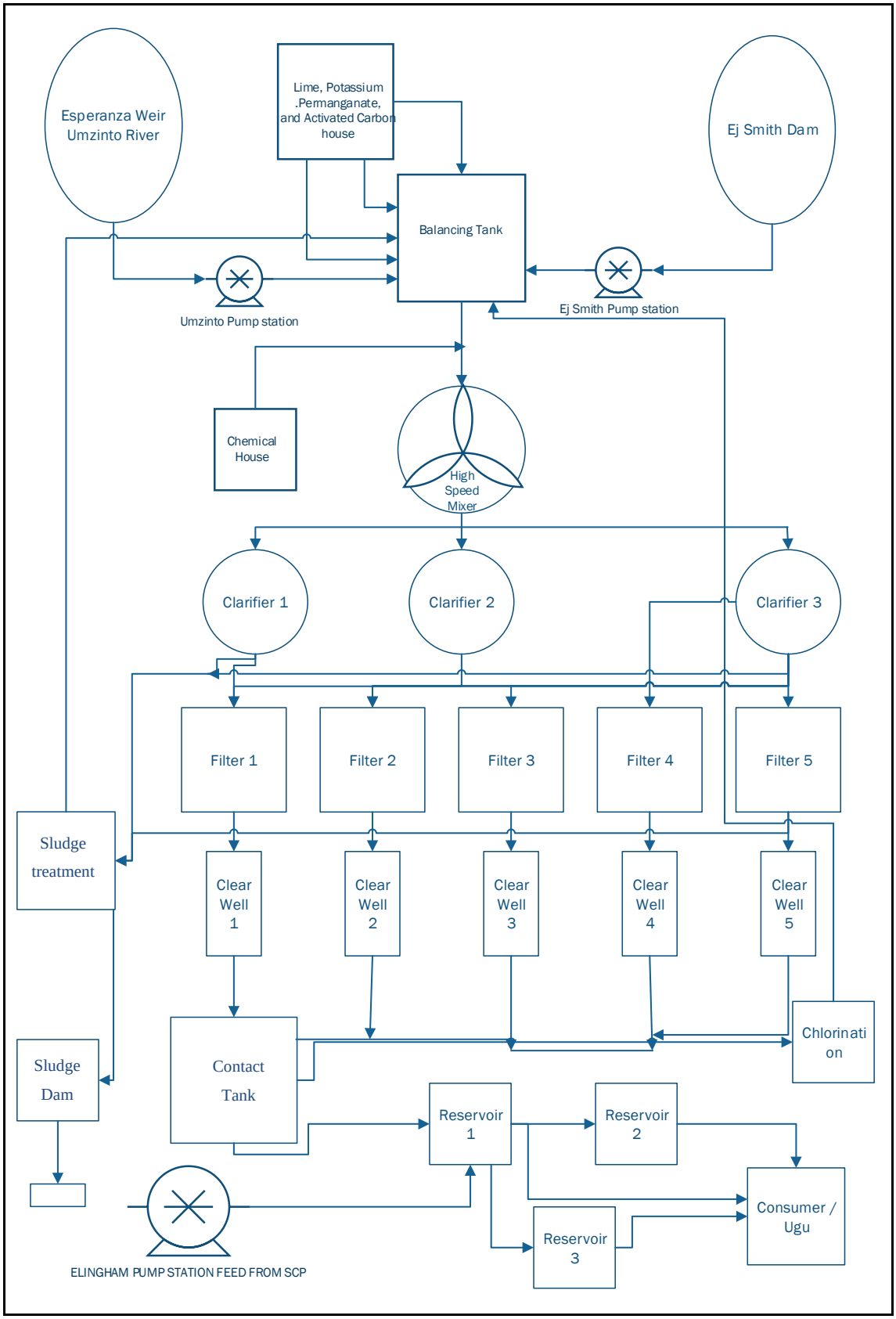


Figure 3.5: Umzinto Water Treatment Plant Process layout

In general a conventional water treatment plant comprises of the following five stages: Coagulation, Flocculation, Filtration, Disinfection and Storage. For the Umzinto plant, the first stage is a balancing tank.

3.4.1 Abstraction

As noted above, the treatment plant has two abstraction points, the EJ Smith Dam and the Esperenza weir. There are four draw-off levels at the EJ Smith Dam. Scientific Services monitors the water quality on a weekly basis and reports the results to the treatment plant. The results are used to determine the level to be used to obtain the best quality. The EJ Smith Dam abstraction point has five pumps which are serviced on a monthly basis. The number of pumps used is determined by the treatment's plant outflow or demand, the level of the reservoirs, and the number of pumps being used at the other abstraction point.

The Esperenza weir has 11 well points that are used to abstract water from the Umzinto River. This abstraction point has four pumps which are auto-controlled from the plant control room. The maximum number of pumps that can be used is three and the remaining one is reserved for backup. More water is abstracted from this source due to its good raw water quality.

3.4.2 Balancing tank

This is the first stage of water treatment at the UWTP. The balancing tank receives raw water from the two abstraction points and chemicals are added to condition the water for further treatment. The chemicals that are added are lime, potassium permanganate and activated carbon.

Lime is added to control the pH of the water. The treatment plant's procedures stipulate a formula for lime dosage which shows a relationship between the amounts of lime and the raw water inflows. This calculation is scheduled to be done every week and then the plant operators are instructed on the number of 25kg lime bags they should be adding to the raw water. Potassium permanganate is added to remove iron and manganese from the raw water. The potassium permanganate is stored in 50 kg bags at the treatment plant. To dose it, a solution should first be prepared. The solution is made by pouring 1

000 litres of water into a mixing tank and then adding 10 kg of potassium permanganate. The mixture will be ready for dosing after 30 minutes. The treatment plant's procedures require the drop test to be done to determine potassium permanganate dosage. The other chemical dosed at the balancing tank is Activated Carbon, which is added to control the taste and odour of the water. From the treatment plant's records activated carbon is rarely added.

3.4.3 Coagulation

Coagulation is usually the first main step in water treatment although in some treatment plants, depending on the quality of the raw water source, processes such as aeration, mixing and pre-oxidising may precede it (Hamilton and Schladow 1997). At the UWTP coagulation is preceded by the settling tank. The innovation is a well-demonstrated technology for the generous evacuation of shading and particulate matter including protozoa (e.g., *Cryptosporidium* oocysts and *Giardia* cysts), viruses, bacteria, and other micro-organisms. Iron, manganese, tastes and odours may also be removed from the water by this process (Franceschi *et al.* 2002).

Kulkarni and Chellam (2010) research showed that adding chlorine to water containing natural organic matter can trigger a reaction that reduces disinfection efficiency and forms chlorinated organic species, e.g., disinfection by-products (DBPs). The DBPs pose a great risk to health. Organic material left over in the water may contain micro-organisms which pose risks to public health (Kulkarni and Chellam 2010).

3.4.4 Flocculation

In this stage, water is received from the coagulation phase and is gently mixed. Gentle mixing allows the particles to agglomerate and form settleable flocs (*Water purification* 2013). Agglomeration during this stage gives greatest molecule contact to floc development, whilst minimizing turbulence and shear which may harm the flocs. The viability of flocculation relies on the delay time (or contact time) and blending conditions before any flocculants are included, the rate of treatment, water temperature and the blending conditions inside of the flocculation chamber (Franceschi *et al.* 2002). This is also supported by Beverly (2011).

3.4.5 Sedimentation

During this treatment process, the velocity of the water is reduced to a level lower than the suspension velocity. This allows the suspended particles to settle due to gravity. The process is also referred to as settling or clarification (Droste 1997).

Sedimentation follows coagulation and flocculation and precedes filtration in most water treatment designs (Crittenden *et al.* 2012; Mountain Empire Community College 2013), including the UWTP. With sedimentation at this stage chemicals are added in the coagulation to facilitate the removal of the resulting floc from the water. (O'Connor and O'Connor 2001) noted that 90% of the suspended particles from raw water, including bacteria, should be removed at this stage. The process will therefore decrease the concentration of suspended particles in the water, reducing the load on the filters (Mountain Empire Community College 2013).

It should be noted that, despite the fact that most water treatment plants use sedimentation, when the turbidity is very low there is no need for it. In such a scenario, coagulation and flocculation are used to produce very small flocs, which easily be filtered off (Mountain Empire Community College 2013).

3.4.6 Filtration

After most of the floc has been removed from the water during coagulation flocculation and sedimentation, it moves to the filtration stage. Filtration is described as the polishing stage of the treatment process (Rietveld *et al.* 2010). Between 1965 and 1996, 108 disease outbreaks were reported, 52 of which were caused by inadequate treatment of surface water, where chlorination was used, but not filtration (USEPA 1999). Such statistics prove the importance of filtration in the treatment process. At this treatment stage, about 99.5% of suspended particles comprise of floc, minerals and microorganisms (Gray 2010). The microorganism cysts and oocytes are highly resistant to chemical disinfectants such as chlorine; hence, the use of filtration can greatly reduce water-related diseases (Ashbolt 2004). Both filtration and disinfection are part of the multi-barrier approach recommended by the USEPA to prevent waterborne transmission of *Giardia* and *Cryptosporidium* (USEPA 1999, 2001).

During the treatment process, the bed will eventually clog with suspended particles as the bed porosity reduces. If this is not dealt with, the head loss through the bed and the shear on captured floc increases. Before these reach unacceptable levels, backwashing should be done to clear the bed. At the UWTP, this takes 20 minutes and is done at random intervals (in most cases once every 6hrs) depending on the quality of the water abstracted.

3.4.7 Disinfection

This chemical process is usually the last in the conventional treatment process (Kingham and Hoggart 1995). As noted in section 2.3.4, a large number of microorganisms are removed during the physical processes especially the filtration process. Realising that clear does not necessarily mean clean, dis-infection is done to kill or inactivate potentially harmful pathogens. Unlike sterilisation, an expensive process that can remove all the pathogens, disinfection will only eliminate selective pathogens (Fraise, Maillard and Sattar 2012). Disinfection is necessary to obtain safe drinking water, whilst sterilisation is an unnecessary expense. The chemical used at this stage is chlorine which is added relative to the residual chlorine measured in the clear well. The target residual chlorine range is 1.5-2.0 mg/L.

3.4.8 Storage

To ensure that disinfection is complete; the water is usually placed in a closed tank or reservoir. chlorine, chloramines, or chlorine dioxide are added to help protect both the end user and the distribution system (Mintz, Reiff and Tauxe 1995). The water then flows through pipes to homes, schools, farms, and businesses, etc.

3.5 Umzinto Water Distribution System

The supply area is covered by the water treatment plant and the South Coast Pipeline shown in Figure 3.6. The water treatment plant and the pipeline fill the command reservoirs and water is distributed to consumers from the reservoirs. The distribution system consisted of Pennington, Umzinto, Umzinto Flats, Ghandinagar, Shayamoya, Park Rynie, Scottburgh South, Scottburgh Central, Amandawe, Malangeni, Amahlongwa and Imfume. The first eight areas are urban supply zones and the other four are rural zones.

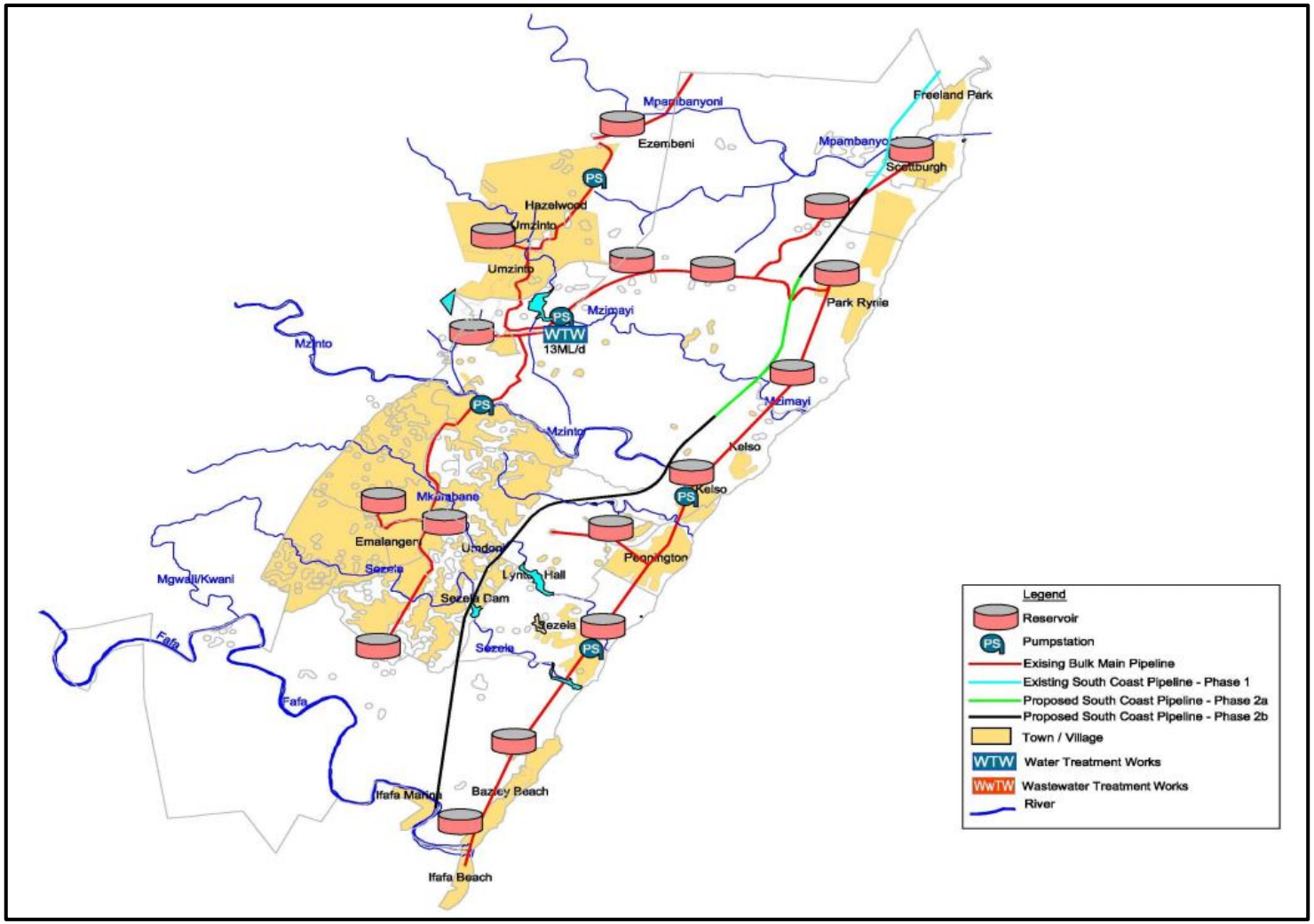


Figure 3.6: Schematic layout of the Umzinto Water Distribution System.Source : (Ugu District Municipality 2011)

3.6 Components of Non-Revenue Water

The components of NRW in the Umzinto Distribution System will generally include those discussed in chapter 2 under the South African Water Balance. The components are:

- Unpaid Water Bills (non-recovered revenue)
- Unbilled metered consumption (authorised consumption);
- Unbilled unmetered consumption (authorised consumption);
- Unauthorised consumption (Apparent Losses);
- Customer meter inaccuracies (Apparent Losses);
- Leakage on transmission and distribution mains (Real Losses);
- Leakage on overflows at storage tanks (Real Losses); and
- Leakage on service connections up to the point of the customer meter (Real Losses) (McKenzie, Siqalaba and Wegelin 2012).

For the purpose of this study, it is important to understand the components making up NRW in the Umzinto Water Distribution System.

3.6.1 Unpaid water bills

This is usually the result of bad debts in the water sector. The consumer is billed but fails to pay their bill. This water excludes the 6 kL/month free basic water that is given to every household. This data can only be obtained from Ugu District Municipality's financial records.

3.6.2 Unbilled metered consumption

This includes the water used in municipal buildings, parks and swimming pools. Another component is unreadable consumer meters. An example of meter that is unreadable is shown in Figure 3.7.

Figure 3.7: A customer meter which cannot be read

3.6.3 Unbilled unmetered consumption

This authorised component of unbilled water emanates from consumers that are billed at a flat rate. It adds to NRW when the actual consumption is more than the flat rate amount. In some cases consumers who are billed take advantage of this arrangement.

Stand pipes are installed in low income areas. Users do not pay for the water and hence stand pipes are a major contributor to unbilled metered consumption. Since most stand pipes are unmetered, the volume of water is often estimated and included in the water balance. Major losses occur due to vandalism and lack of maintenance. A common situation is shown in Figure 3.8; a spindle has been removed from a tap and the water simply runs.

Figure 3.8: Standpipe left running after the spindle was removed

Adding to unbilled unmetered consumption is the use of fire hydrants. Most fire hydrants are unmetered and are consequently subject to abuse by consumers through illegal connections. In some cases, taxi drivers use the hydrants to wash their vehicles. Beside these illegal activities, the volume of this portion of unmetered consumption includes the water lost when the hydrants are opened to reduce distribution pressure prior to maintenance.

3.6.4 Unauthorised connections

Like any other municipality, Ugu District Municipality has to deal with the problem of illegal connections. In South Africa, water connections are deemed illegal when the consumer connects directly to the water source without the municipality's consent. The most common forms involve bypassing meters, as shown in Figure 3.9, and connecting to air valves.

Figure 3.9: An illegal connection where a consumer meter has been bypassed

3.6.5 Customer meter inaccuracies

With time, customer meters lose accuracy. The accuracy of a meter is also dependent on the water quality. In the Umzinto Water Reticulation System age is the main determinant of meter accuracy. It is very difficult to estimate the volume of water that has been lost through an inaccurate meter.

3.6.6 Leakage on transmission and distribution mains

When there is excessive pressure, mains can burst and this can be a significant contributor to Real Losses in the Umzinto Water Distribution System. Figure 3.10 shows a leakage on a main distribution line. Another cause of losses from the mains is background leaks. To control these two sources of loss, the municipality has a pressure management programme in place and invests in the maintenance of the distribution system.

Figure 3.10: Water Leaking from the Mains Distribution Line

Community awareness programmes are run to help consumers understand their role in reducing water losses. This is essential in reducing or eliminating unreported leaks, such as that shown in Figure 3.11. If consumers are not educated about the importance of reporting water leaks, they will not do so unless their household taps run dry.

Figure 3.11: Water Leakage in a Community

3.6.7 Leakage on overflows at storage tanks

The municipality has a good reservoir balance system to avoid these leakages. The system includes meters, valves and control systems at the reservoirs. Although water losses cannot be reduced to zero, this type of water loss is very minor in the Umzinto Distribution System.

3.6.8 Leakage on service connections

In many frameworks, leakage from connections is by a wide margin the greatest source of physical leakage; frequently representing 80% or a greater amount of aggregate physical losses (Karadirek *et al.* 2012; McKenzie, Siqalaba and Wegelin 2012; Lambert, Fantozzi and Thornton 2013). In a water balance, water lost via service connection leakage is included in the calculation of the unavoidable annual Real Losses. Further, evaluation of this total leakage on service connections is therefore mainly dependant on the number of service connections within a network and its average operating pressure (McKenzie, Siqalaba and Wegelin 2012). Figure 3.12 is a typical example of leakage at a service connection.

Figure 3.12: A leak at a service connection

3.7 Summary

This chapter described the study area and presented an overview of the treatment plant's location relative to the raw water sources and the supply area. The technologies currently employed at the UWTP, Umzinto Water Distribution System were also described. NRW such as unpaid water bills, customer meter inaccuracies and leakage on transmission and distribution mains were also elucidated, paving the way for chapter 4 that describes the methodology used for its optimisation.

CHAPTER 4 – RESEARCH METHODS

4.1 Introduction

This study utilised exploratory, descriptive and explanatory research designs. Exploratory research was used to gather and analyse the literature on water treatment plants and their distribution systems. Meetings were held with experts in the fields of water treatment and water loss management. This enabled the researcher to understand what had already been done and the areas that required improvement. A descriptive design was employed to interpret the study's findings, while explanatory research was used to explain the reasons for the relationships that were observed during data analysis.

This study adopted a quantitative approach that allowed for the generation of statistical data on a large scale. Quantitative research offers significant advantages in terms of validity and the repeatability of results.

The research strategy involved the development of an optimisation model based on historical and observational data collected from the UWTP. This method reduced speculation, interpolation and the time taken to determine the water quality by environmental means. The study also sought to develop a tool for water usage analysis to help monitor and control NRW.

4.2 Data Collection

The instrument used to collect data involved observing the system, meetings and collection of historical data. Separate data collection was done for the water treatment plant and the distribution system.

4.2.1 Water treatment plant data

The water treatment data were collected from the UWTP's historical records. The chemical usage data available were on the dosage of chlorine, lime, activated carbon, and potassium permanganate. The dosing was done at the treatment plant and the data available for collection was from July 2006 to June 2013. For raw water quality, the available data were from August 1995 to December 2013, but due to a lack of dosage data for some of this

period, the study used raw water quality data from July 2006 to June 2013. The raw water quality data were obtained from the two abstraction points viz the EJ Smith Dam and Umzinto River.

4.2.2 Water distribution data

The same criteria used to collect data for the water treatment plant were used for water distribution. The researcher consulted with Ugu District Municipality which is responsible for distributing water from the UWTP.

4.3 Data Pre-processing

The raw water quality data were extracted from the Laboratory Information Management System (LIMS). As indicated in Magombo *et al.* (2015) extracted data were in the Excel format but its layout made it difficult to use for analysis. The first step in pre-processing the data required use of a relational database to make it usable in MATLAB as the tool for developing the optimisation model. In this study, Microsoft SQL Server Management Studio 2012 (MSQL2012) was used to create the relational database.

The process involved importing data from an Excel file into a MSQL2012 table. The graphic user interface for this process is shown in Figure 4.1 were the data source is being defined. In the MSQL2012 database the data had the same layout as that from the Excel file, but the management studio enabled the extraction to be done in the preferred format for analysis in MATLAB.

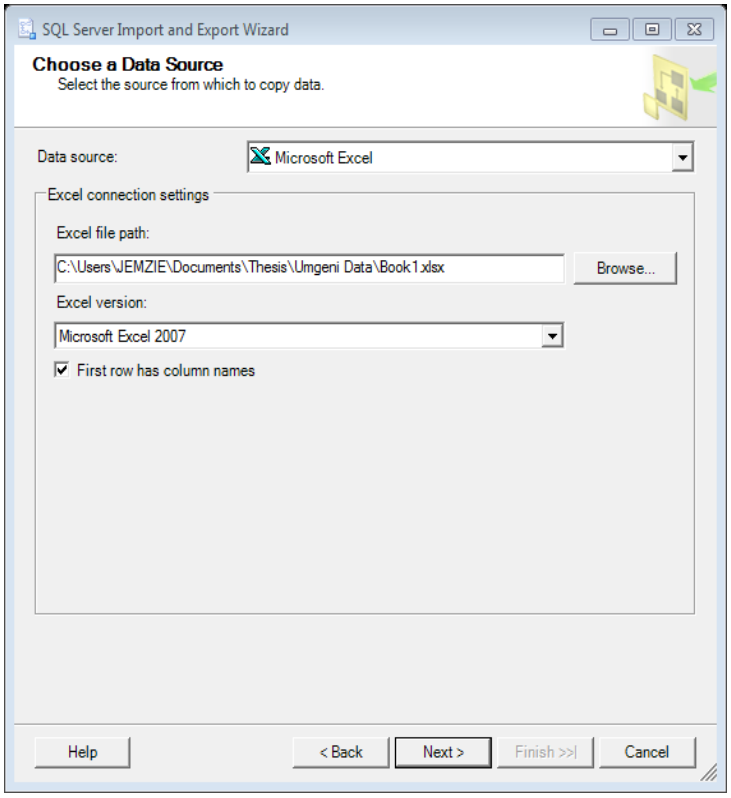


Figure 4.1: Definition of the data source to import into a SQL database. Software platform: Microsoft SQL Server 2012 Management Studio (Microsoft Corporation 2012a)

Figure 4.1 shows that the file path and version of Microsoft Excel also helped define the data source. After defining the data source the next step was to define the data destination which in this study was the Microsoft SQL Server 2012 database. This process is shown in Figure 4.2.

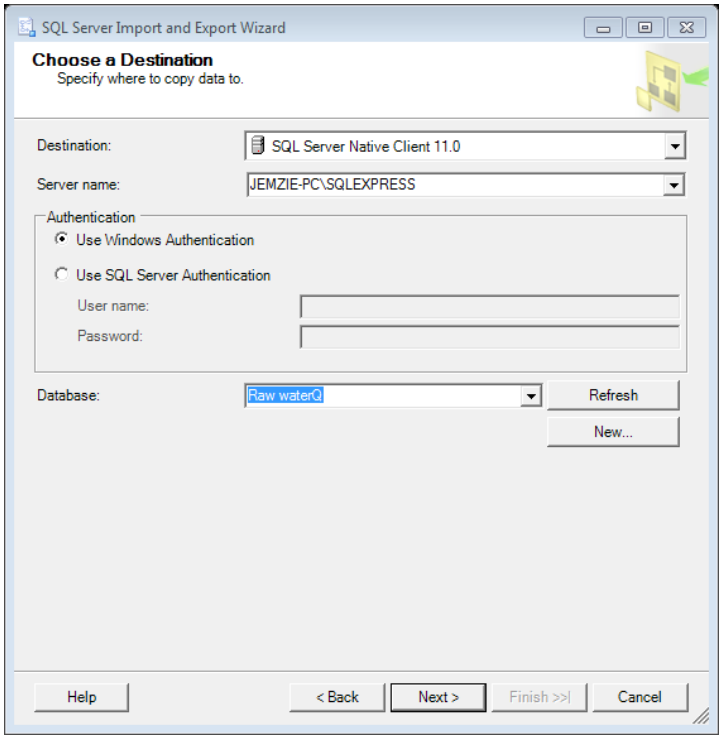


Figure 4.2: Data destination for the SQL server and authentication setting. Software platform: Microsoft SQL Server 2012 Management Studio (Microsoft Corporation 2012a)

Figure 4.2 shows that the server configuration needed to be defined in defining the destination and in this study, the SQL Server Native Client 11.0 was used. A new database storing the data was created and the authentication was set. The next step was to specify the amount of data that could be imported from the data source to the destination. This is shown in Figure 4.3.

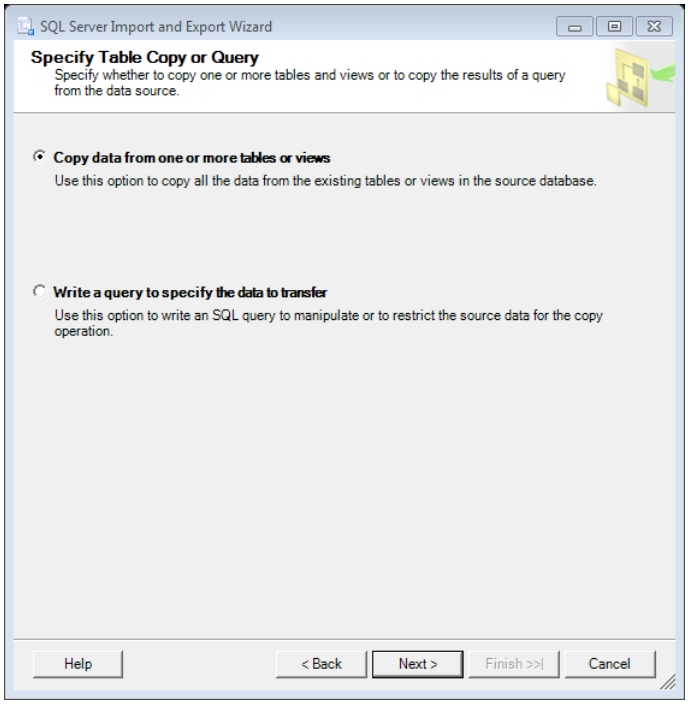


Figure 4.3: Specifying the amount of data to be copied from the data source. Software platform: Microsoft SQL Server 2012 Management Studio (Microsoft Corporation 2012a)

There are two methods for copying data from the source to the destination. These are shown in Figure 4.3. One option is to copy all the data from the source and the other is to write scripts to select the specific data to be copied. In this study, all data from the source were copied from the first sheet of the source. These were copied to a table named WaterQ in the destination database. This type of mapping is illustrated in Figure 4.4.

When the data has been successfully imported, the determinant and the sampling points are key in extracting the data. Appendix 1 shows the scripts that were used to extract data for each parameter for the two abstraction points. Further processing was required to deal with missing values, censored data and outlying values.

Firstly, the percentage of missing values from July 2006 to June 2013was analysed for each parameter. Parameters with more than 25% missing data were removed from the datasets. For the remaining parameters, pre-processing for the missing values was achieved using the cubic interpolation (Dzwairo *et al.* 2011). Advantages of using cubic interpolation includes the fact that it can provide a “smooth” interpolant and is usually more accurate compared to linear interpolation. Cubic interpolation also doesn’t get “wiggly” like higher-order polynomial interpolation can.

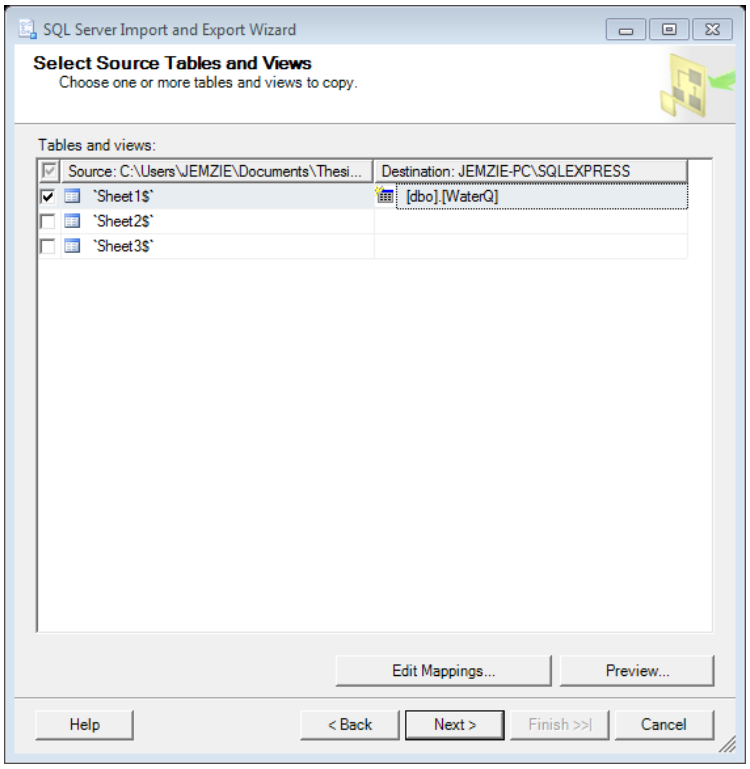


Figure 4.4: Mapping the source tables to the destination tables. Software platform: Microsoft SQL Server 2012 Management Studio (Microsoft Corporation 2012a)

In this study, Matlab software was used to perform the interpolation. This can be done by using the pchip and spline functions in Matlab. The two functions are defined as: a) pchip:- Piecewise Cubic Hermite Interpolating Polynomial and b) spline:- Cubic spline data interpolation (MathWorks 2014). The main differences between the two can be shown in Figure 4.5

The cubic spline data interpolation produces a smoother result and if the data supplied consists of values of a smooth function, it produces a more accurate curve. The cubic hermite has an advantage that it does not produce overshoots and has less oscillations if the data are not smooth. The two methods are equally expensive to evaluate but pchip give has an advantage of being easy to setup. An alternative method to use in filling out missing data was to choose one definite value (e.g. the average) to fill all the blank cells, but this is not a statistically recommended approach (Weiss and Indurkhya 2000; Kotsiantis, Kanellopoulos and Pintelas 2006; Magombo *et al.* 2015)

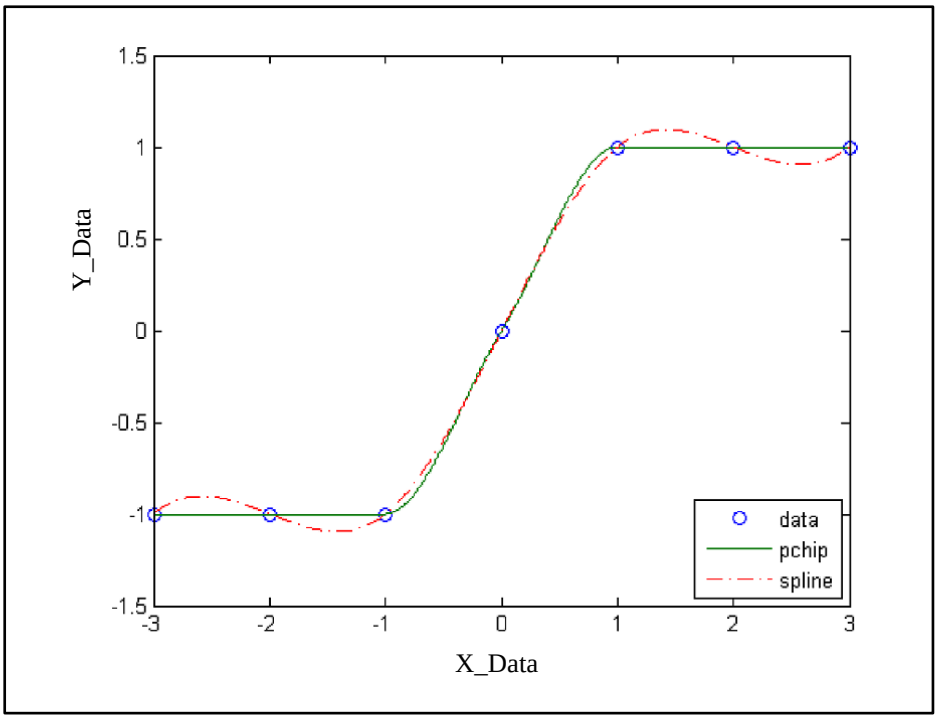


Figure 4.5: Difference between pchip and spline functions in Matlab.

In pre-processing the censored values, this study used the simple replacement technique. Censored values were replaced by a value equal to the limit of detection value divided by the square root of two (Magombo *et al.* 2015). For this process, an application designed and developed during this study was used and its interface is shown in Appendix 3.

In dealing with outliers, statistical methods were used for the initial identification of the outliers. After the statistical identification, comparison was made on the trends for related raw water quality parameter to justify the outliers. For example, if suspended solids had outliers, its trend line would be compared to the trend lines for colour or turbidity. A peak in suspended solids should result in a peak in the other parameters. Once the outliers were identified and justified, cubic hermite interpolation was used to determine values to replace them.

4.4 Methods for Objective 1

Objective 1: Identify and quantify key raw water quality parameters affecting treatment at the Umzinto Water Treatment Plant.

The first step in accomplishing this objective was to determine the parameters that were tested for by UWTP from July 2006 to June 2013. These data were extracted from the LIMS database for the two abstraction points. During this period, raw water quality parameters were sampled and these are shown in Appendix 2 together with the number of tests for each parameter. The key parameters were chosen from these measured parameters using the following three criteria:

1. Data availability and the effect of the detected level of the parameter measured.
2. The literature survey on water quality parameter selection.
3. Correlation of raw water quality parameters to the chemical dosages.

4.4.1 Data availability and the effect of the detected level of parameter measured

For the initial parameter selection, water quality data for 120 parameters were extracted for the two abstraction points, EJ Smith and Esperanza weir and processed as described under data pre-processing. The frequency of measurement for each parameter was assessed. Parameters that had no readings were automatically left out. The objective was to obtain a dataset where the frequency enabled further processing to establish a correlation.

4.4.2 Water quality parameter selection

In terms of the parameters selected, taking into account data availability and the literature review on water quality parameters in Chapter 2, 19 parameters, algal count, manganese (Mn), iron (Fe), Geosmin, *Escherichia coli*, 2-methylisoborneol, total coliforms, colour, conductivity, turbidity, suspended solids (SS), pH, temperature, total organic carbon (TOC), aluminium (Al), calcium (Ca), magnesium (Mg), hardness and taste were selected for further analysis to determine the key parameters affecting treatment at the plant.

In selecting the key parameters, consideration was also given to the WHO Water Quality Guidelines (WHO 2011) and the South African Water Quality Standards (SANS 241, 2011). The parameters were identified and plotted on line graphs and box plots.

Box and Whisker plots

The statistical parameters for box and whisker plots are illustrated in Figure 4.6.In Excel the parameters are defined by the minimum (min), first quartile (q1), median, third quartile (q3), and maximum (max), which are calculated as shown in Table 4.1.

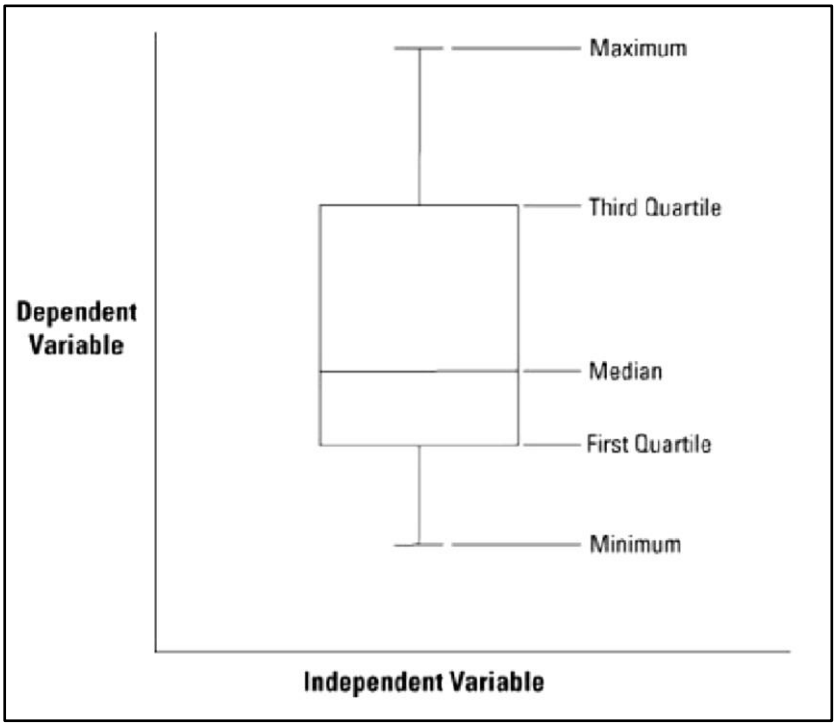


Figure 4.6: Statistical parameters on a Box and Whisker plot

Table 4.1: Excel functions used to determine parameters for the Box and Whisker plots

Statistical parameter	**Excel Function**
Minimum	Min(Range)
First Quartile	Quartile(Range,1)
Median	Median(Range)
Third Quartile	Quartile(Range,3)
Maximum	Max

In Table 4.1 the minimum represents the lowest observation in a range. The first quartile is also called the lower quartile or the 25th percentile (splits off the lowest 25% of data from the highest 75%). The median also called the second quartile or the 50th percentile (cuts data set in half). The third quartile also called the upper quartile or the 75th percentile (splits

off the highest 25% of data from the lowest 75%). The maximum represent the highest observation.

4.4.3 Correlation and regression of water quality parameters

To establish the relationship between the raw water quality parameters and chemical dosage, the raw water quality parameters measured at EJ Smith Dam and the Umzinto River were correlated to the chemical dosages at the UWTP. The chemical dosages considered were lime, polymer, chlorine and activated carbon dosage.

The Karl Pearson's correlation method was used to determine the presence and strength of the relationship between water quality parameters and the chemical dosages. According to Karl Pearson, the coefficient of correlation (r), between two parameters x and y is given by Equation 4.1.

$$r = \frac{\sum xy - n\left(\frac{\sum x}{n}\right)\left(\frac{\sum y}{n}\right)}{n\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}\sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}}$$ Equation 4.1

where $x = X - \overline{X}$, $y = Y - \overline{Y}$, $\overline{X} = \frac{\sum X}{n}$, $\overline{Y} = \frac{\sum Y}{n}$ and n is number of samples.

If the value of $r=1$ for the two parameters X and Y it means they have a strong correlation. Table 4.2 shows some descriptions for the most common values of r.

Table 4.2: Correlation descriptive scale. Source: (Rowntree 2004)

Coefficient ranges	**Correlation description**
0.0	No correlation
+1 and -1	Perfect positive and perfect negative correlation, respectively
0.0 to 0.2	Very weak, negligible correlation
0.2 to 0.4	Weak, low correlation
0.4 to 0.7	Moderate correlation
0.7 to 0.9	Strong, high, marked correlation
0.9 to 1.0	Very strong, very high correlation.

4.5 Methods for Objective 2

Objective 2: Design a genetic algorithm for the potable water treatment process control.

For this objective Matlab 2014, a software developed by Mathworks, was used to develop algorithms for the optimisation of water treatment. Using the results obtained from the key raw water parameters selected in objective 1, the steps taken to develop the model comprised the following:

- Definition of the objective function
- Definition of the objective function variables
- Model constraints
- Genetic Algorithm parameter selection.

4.5.1 Definition of the objective function

This study employed the general cost equation from economics as a cost function. The water treatment cost was taken as the sum of operational cost, depreciation cost, and levies and tax. This is mathematically shown in Equation 4.2.

$$WC = OC + DC + TX \qquad \text{Equation 4.2}$$

where *WC* = Water Treatment Cost

OC = Operational Cost

DC = Depreciation Cost

TX = Tax and Levies.

Equation 4.2 is also used by Abdullahi and Abdulkarim (2010). In this water treatment cost model, this study focused on the component of operational cost since the other components are relatively fixed and easy to manage. The operational cost is a combination of the costs associated with labour, energy, maintenance, administration and the cost of chemicals. This is mathematically shown in Equation 4.3.

$$OC = LC + EC + MC + AC + CC \qquad \text{Equation 4.3}$$

Where *OC* = Operational Cost

LC = Labour Cost

EC = Energy Cost

MC = Maintenance Cost

AC = Administrative Costs

CC = Chemical Costs

Stevenson and Sum (2009) noted that operational cost is the sum of fixed and variable cost shown in Equation 4.4.

$$OC = VC + FC \qquad \text{Equation 4.4}$$

Where *OC* = Operational Cost

VC = Variable Cost

FC = Fixed Cost.

In equation 4.4 the fixed cost is given by the sum of labour, energy, maintenance and administration costs. This means that the variable costs will be the chemical costs. This was the scope of this study. The water treatment was optimised by minimising the costs associated with chemical dosage. At the UWTP the chemical cost is mainly driven by the dosage of lime, potassium permanganate, activated carbon, chlorine and polymer. In mathematical terms this is given by Equation 4.5:

$$CC = D_1P_1 + D_2P_2 + D_3P_3 + D_4P_4 + D_5P_5 \qquad \text{Equation 4.5}$$

Where CC = Chemical Cost

D_1 = chlorine dosage

P_1 = Price of chlorine

D_2 = Lime dosage

P_2 = Price of lime

D_3 = Potassium permanganate dosage

P_3 = Price of potassium permanganate

D_4 = Polymer dosage

P_4 = Price of polymer

D_5 = Activated carbon dosage

P_5 = Price of activated carbon

The objective function for the study is therefore expressed as shown in Equation 4.6.

$$Minimise\ CC = D_1P_1 + D_2P_2 + D_3P_3 + D_4P_4 + D_5P_5 \quad \text{Equation 4.6}$$

The prices of the chemicals were obtained from the water treatment plant's purchases report. The variables in the objective function were defined from the analysis done in the first objective and their definition is given in section 4.5.2.

4.5.2 Definition of the objective function variables

The variables in the objective function were derived from the correlation information from objective 1, which identifies and quantifies the key raw water quality parameters affecting treatment at the UWTP. These variables were used to develop four chemical dosage prediction models that were then combined to produce a cost prediction model governed by the objective function given in section 4.5.1. Table 4.3 shows the relationship between chemical dosage and raw water quality parameters. The shaded cells show the relationships that had a correlation.

Table 4.3: Relation between chemical dosage and raw water quality parameters

Quality Parameter	Symbol in Equations	Chlorine	Polymer	Lime	Activated Carbon	Potassium Permanganate
Algal count1	AC_1		■		■	
Alga count2	AC_2		■		■	
Alkalinity1	Alk_1		■		■	■
Alkalinity2	Alk_2		■			■
Coliforms1	Col_1		■			■
Coliforms2	Col_2					
Colour1	C_1		■			■
Colour2	C_2	■	■	■		■
Conductivity1	CO_1	■	■	■		
Conductivity2	CO_2		■		■	
E.coli1	Eco_1					
E.coli2	Eco_2		■			
Fe1	Fe_1		■			■
Fe2	Fe_2		■		■	■
Hardness1	H_1		■			
Hardness2	H_2		■		■	

Mn1	Mn_1	■		■		
Mn2	Mn_2		■			
pH1	pH_1		■			
pH2	pH_2		■			
SS1	SS_1		■			
SS2	SS_2		■			■
Temperature1	T_1	■	■			
Temperature2	T_2	■	■			
TOC1	TC_1		■	■		■
TOC2	TC_2	■	■		■	■
Turbidity 1	Tu_1		■			
Turbidity2	Tu_2		■			

In Table 4.3, the number in front of a parameter indicates the raw water abstraction point associated with it. The number 1 represents parameter observations from EJ Smith Dam and the number 2 represents parameter observations from Umzinto River. For example Algal count1 represent observations from the EJ Smith Dam abstraction point and Algal count2 represent observations from the Umzinto River abstraction point.

To develop a model, there was a need to consider the relationship between dosage of a particular chemical and each of the raw water quality parameters it is correlated to (Abdullahi and Abdulkarim 2010). The relationships were obtained from scatter plots and trendlines created using a curve fitting tool in Microsoft Excel 2013. This approach is recommended in the drinking water advice note by EPA (2014).

On the scatter plots, the y values would represent chemical dosage and the x values would represent the raw water quality parameters. The excel curve fitting tool allows the creation and display of equaions for the trendlines created, ranging from the basic linear to the more complicated polynomial curves. It also calculates and display the coefficient of determination (R^2) for a plotted curve. This value can be used to show how close modelled data is to actual data. R^2 values lie between 0 and 1, with 1 meaning the modelled data closely approximates the actual.

Chemical dosage is usually determined by more than one parameter hence in this section the R^2 values are expected to be low, but, can be used to show how much a particular parameter contributes to the determination of a chemical dosage. The curves displayed were used to determine the general equations that would be used to create the chemical

dosage models. Due to the scattering of the data curves created are linear, similar to the approach used by . In this study, the equations for the individual parameters related to a chemical dosage are not taken as the final models but are combined to develop an equation that would be further optimised in Matlab using the genetic algorithm toolbox (MathWorks 2012).

Relationship between chlorine dosage and raw water quality parameters

Table 4.3 shows that chlorine is related to colour2, conductivity1, Mn1, temperature1, temperature2, and TOC2. This can be mathematically shown as:

$$D_1 = f(C_2, CO_1, M_1, T_1, T_2, TC_2) \quad \text{Equation 4.7}$$

Figure 4.7 shows the relationships between chlorine dosage and colour of the raw water from the Umzinto abstraction point. One of the reasons of using chlorine as a good disinfectant is its ability to remove colour.

In the equation shown in Figure 4.7, the y values represent the chlorine dosage and the x values represent colour. This equation was used to formulate a general equation relating chlorine dosage to the raw water colour for the Umzinto River abstraction point. This is shown in Equation 4.8. The coefficient of determination for the relationship shows that colour is a major determinant of chlorine dosage but is not the only factor determining chlorine dosage.

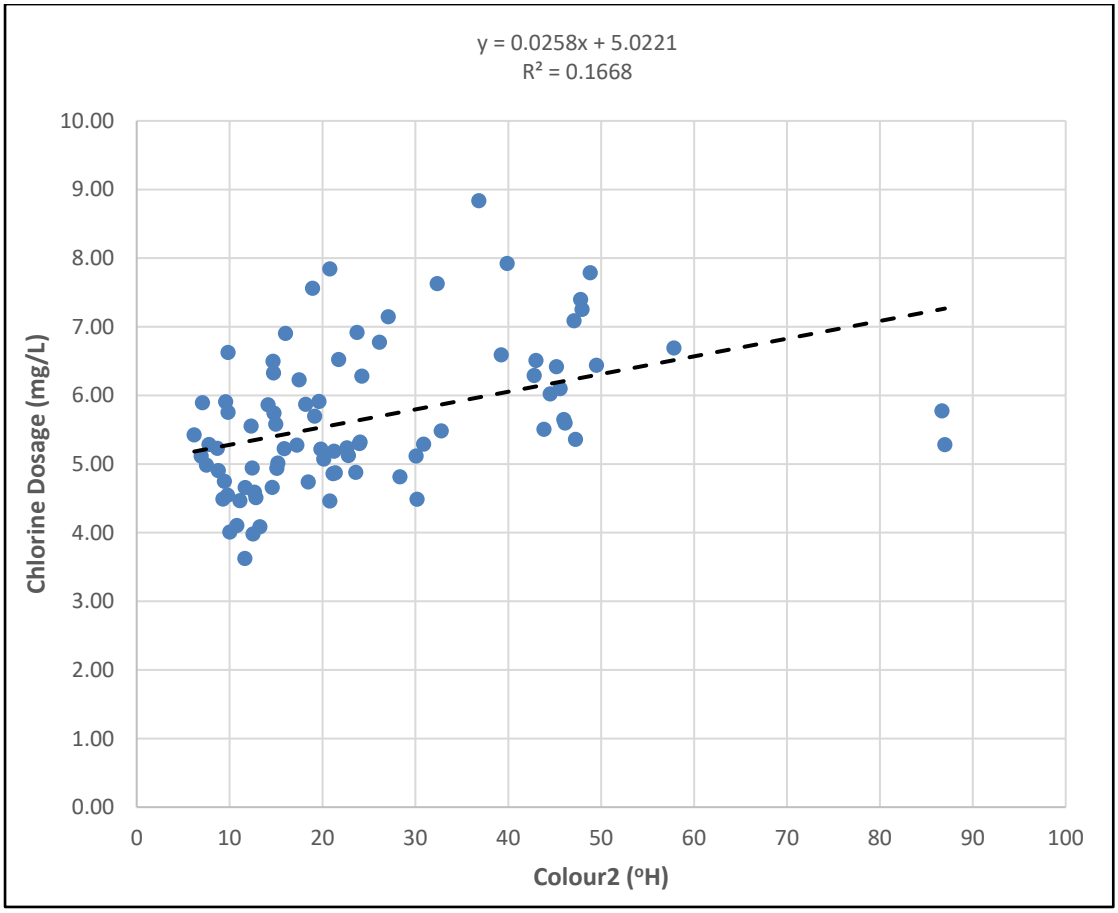


Figure 4.7: Relationship between chlorine dosage and colour2 for the period between July 2006 to June 2013

$$D_1 = K_1 C_2 + K_2 \qquad \text{Equation 4.8}$$

In Equation 4.8, *D1* represents chlorine dosage and *C* represents colour. The subscripts on *C* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Figure 4.8 show the relationship between the dosage of chlorine and conductivity in the Ej Smith's Dam. In the scatter plot chlorine dosage is shown to be related to conductivity using a linear with equation with a low coefficient of determination. This means that conductivity is not major determinant of chlorine dosage but the Pearson correlation carried out in the first objective of the study show that it is worth including.

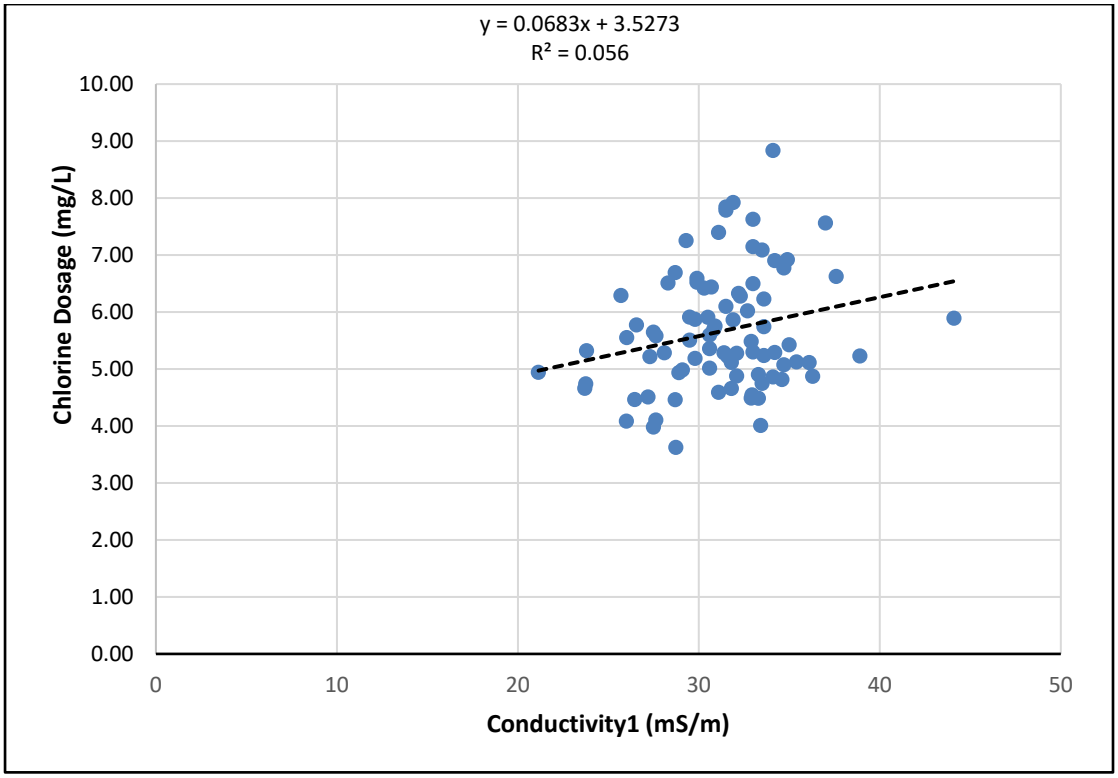


Figure 4.8: Relationship between chlorine dosage and conductivity1 for the period between July 2006 to June 2013

In the equation, the y values represent the chlorine dosage and the x values represent conductivity. The general equation for raw water conductivity at the Ej Smith's Dam abstraction point is shown in Equation 4.9.

$$D_1 = K_3 CO_1 + K_4 \qquad \text{Equation 4.9}$$

In Equation 4.9, D_1 represents chlorine dosage and CO represents conductivity. The subscripts on CO show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.7 shows that chlorine dosage depends on the manganese from the EJ Smith Dam. The scatter plot for the relationship is shown in Figure 4.9 and the linear equation for the trendline representing the relationship has low coefficient of determination.

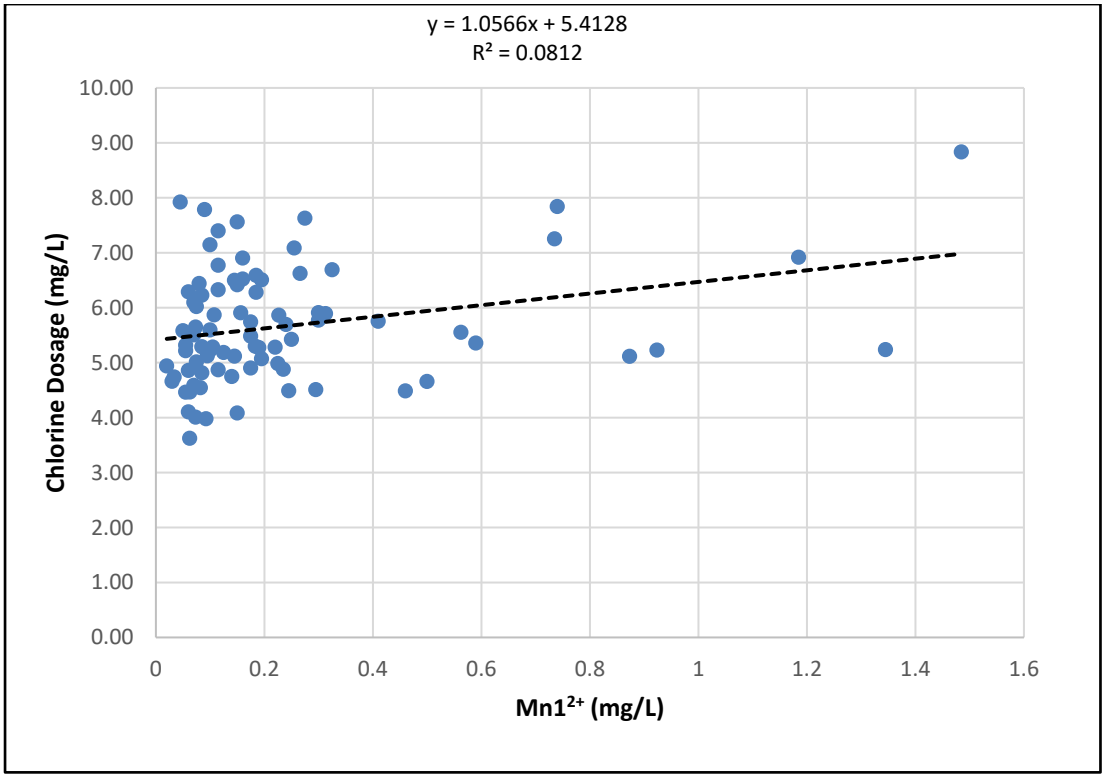


Figure 4.9: Relationship between chlorine dosage and manganese1 for the period between July 2006 to June 2013

In the equation shown in Figure 4.9, the y values represent the chlorine dosage and the x values represent manganese. The general equation for the relationship between chlorine dosage and manganese1 is shown in Equation 4.10.

$$D_1 = K_5 Mn_1 + K_6 \qquad \text{Equation 4.10}$$

In Equation 4.10, *D1* represents chlorine dosage and *Mn* represents manganese. The subscripts on *Mn* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.7 shows that chlorine dosage depends on both the raw water temperature from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.10 and 4.11 respectively. For both abstraction points, the coefficient of determination show that temperature is a major determinant of chlorine dosage. This could be related to the fact that chlorine vaporises at high temperatures and hence increase

chlorine demand in the summer period and low demand in winter (Veenstra and Schnoor 1980).

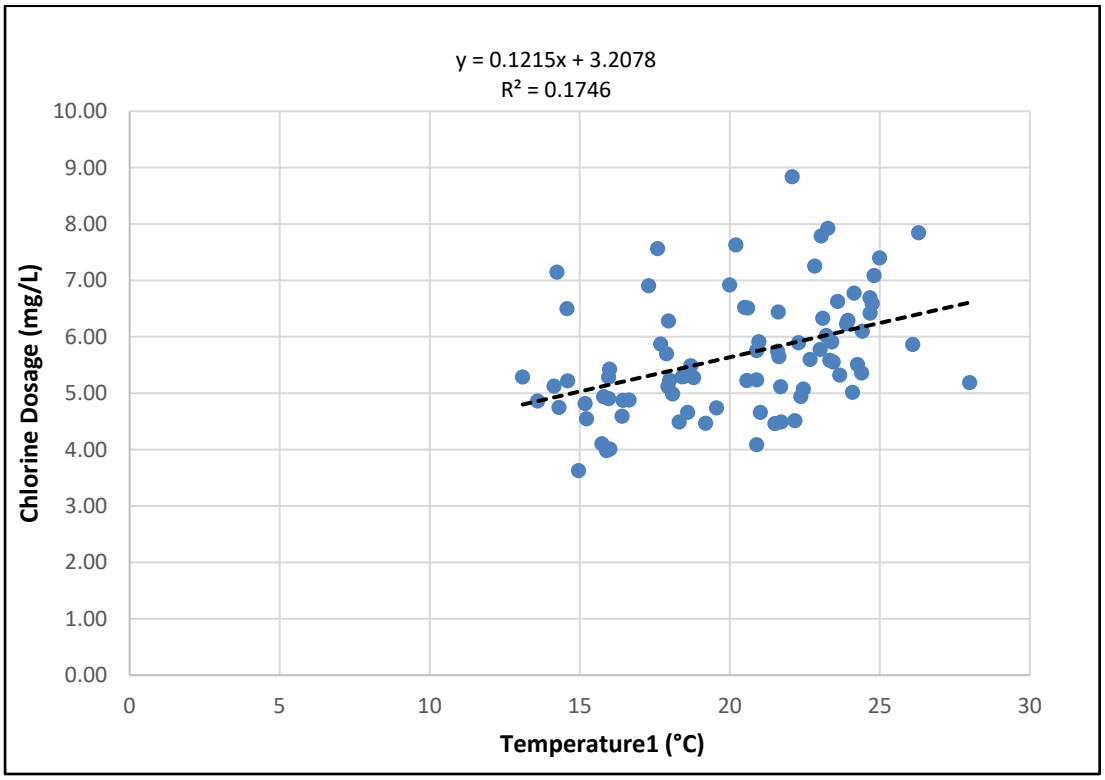


Figure 4.10: Relationship between chlorine dosage and temperature1 for the period between July 2006 to June 2013

In the equations shown in Figure 4.10 and 4.11, the y values represent the chlorine dosage and the x values represent temperature. These equations were then used to formulate general equations relating chlorine dosage to temperature. The general equation for the relationship between chlorine dosage and temperature1 is shown in Equation 4.11 and that for temperature2 is shown in Equation 4.12.

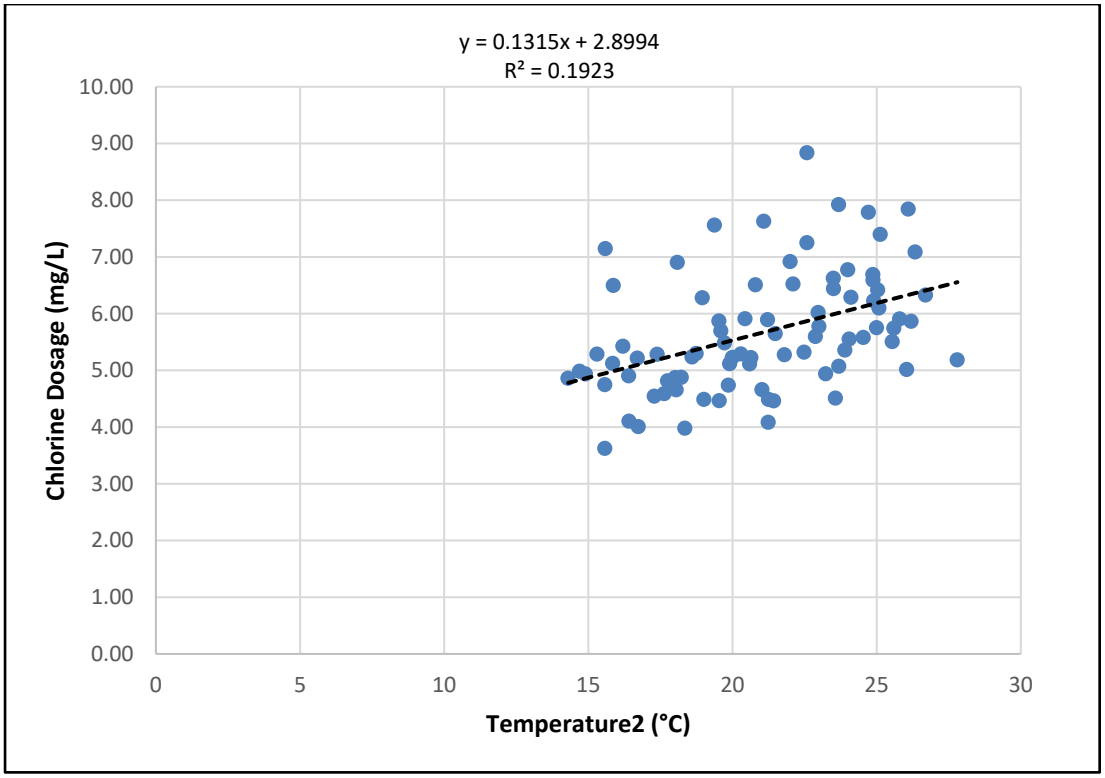


Figure 4.11: Relationship between chlorine dosage and temperature2 for the period between July 2006 to June 2013

.

$$D_1 = K_7 T_1 + K_8 \qquad \text{Equation 4.11}$$

$$D_1 = K_9 T_2 + K_{10} \qquad \text{Equation 4.12}$$

In Equation 4.11 and 4.12, *D1* represents chlorine dosage and *T* represents temperature. The subscripts on *T* show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.7 shows that chlorine dosage depends on the total organic carbonyl (TOC) from the Umzinto River. The scatter plot for the relationship is shown in Figure 4.12. The coefficient of determination for the relationship shows that chlorine is major determinant of chlorine dosage.

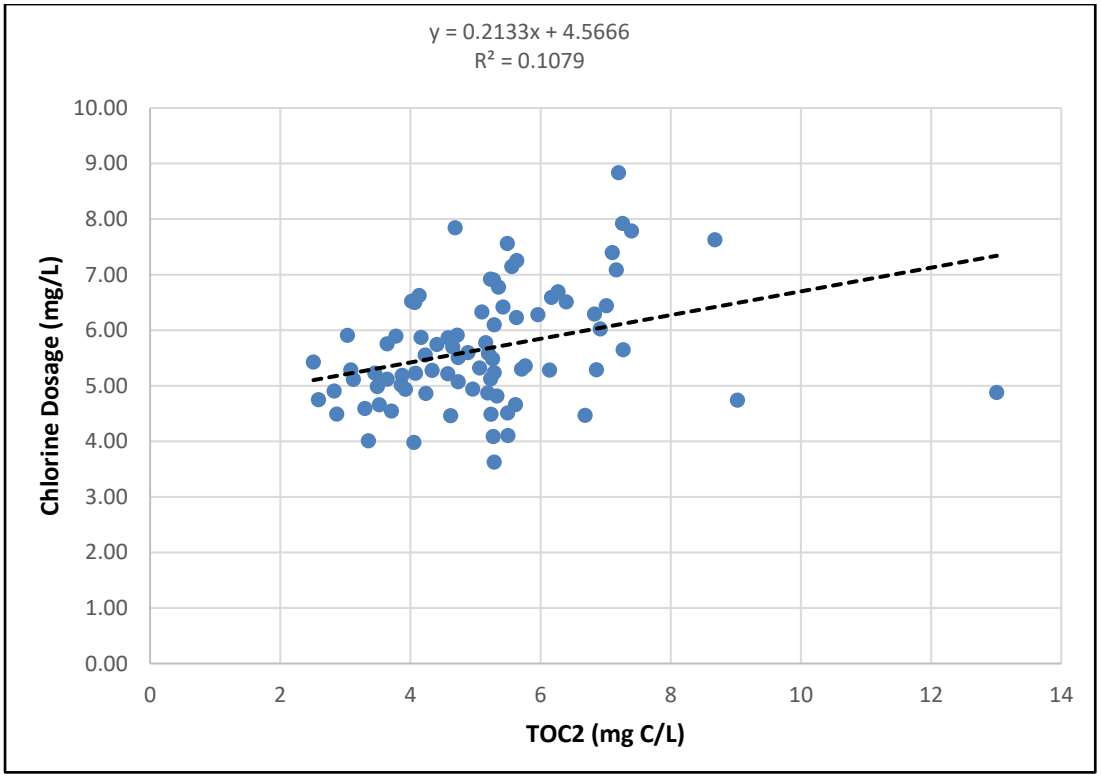


Figure 4.12: Relationship between chlorine dosage and TOC2 for the period between July 2006 to June 2013

In the equation shown in Figure 4.12, the y values represent the chlorine dosage and the x values represent TOC. The general equation for the relationship between chlorine dosage and TOC1 is shown in Equation 4.13.

$$D_1 = K_{11}TC_2 + K_{12} \qquad \text{Equation 4.13}$$

In Equation 4.13, *D1* represents chlorine dosage and *TC* represents TOC. The subscripts on *TC* show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

To obtain the final model for calculating chlorine dosage , the general equations for the individual parameters are combined to produce one general equation whose coefficients would established using genetic algorithms in (MathWorks 2014). This means adding together equations 4.8, 4.9, 4.10, 4.11, 4.12 and 4.13 to obtain Equation 4.14.

$$D_1 = K_1C_2 + K_3CO_2 + K_5M_1 + K_7T_1 + K_9T_2 + K_{11}TC_2 + K_{13}$$ Equation 4.14

Relationship between lime dosage and raw water quality parameters

The lime dosage model was derived in the same manner in which the chlorine dosage model was derived. A scatter plot was done for each parameter that lime dosage was dependent on. The lime dosage model was dependent on the levels of colour, conductivity, manganese and total organic carbonyl (TOC). This can be mathematically shown as:

$$D_2 = f(C_2, CO_1, M_1, TC_2)$$ Equation 4.15

Equation 4.15 shows that lime dosage depends on the raw water colour from the Umzinto River. The scatter plot for the relationship is shown in Figure 4.13. From the coefficient of determination it can be gathered that colour has no major influence on the lime dosage.

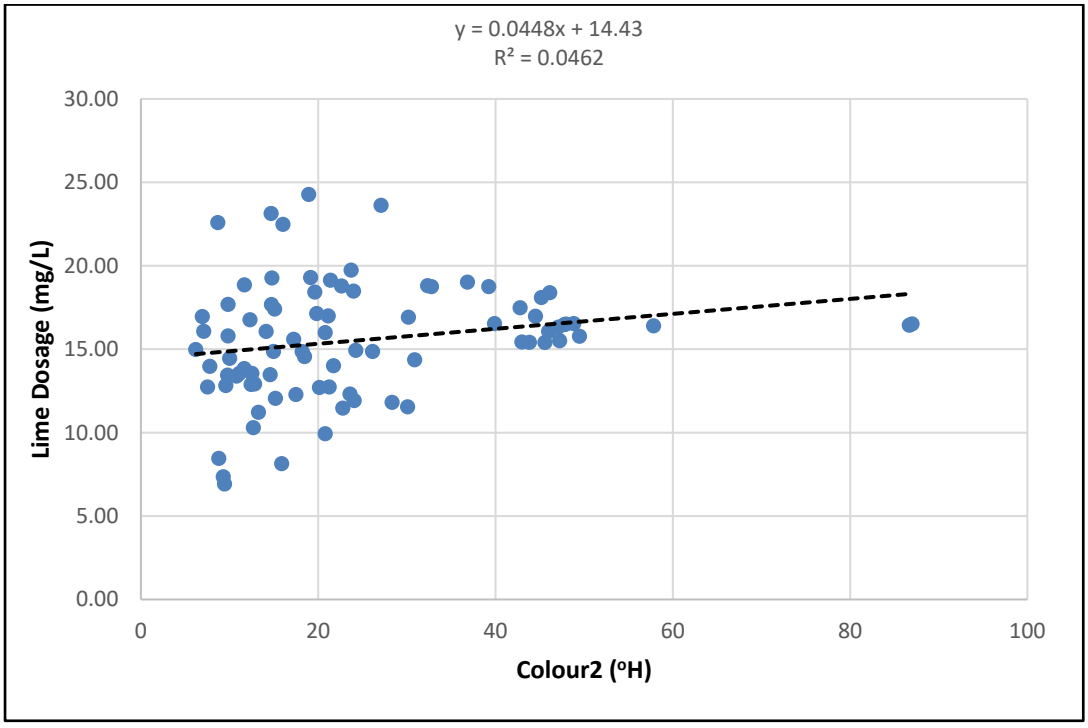


Figure 4.13: Relationship between lime dosage and colour2 for the period between July 2006 to June 2013

In the equation shown in Figure 4.13, the y values represent the lime dosage and the x values represent colour. The general equation for the relationship between lime dosage and colour2 is shown in Equation 4.16.

$$D_2 = K_1 C_2 + K_2 \qquad \text{Equation 4.16}$$

In Equation 4.16, D_2 represents lime dosage and C represents colour. The subscripts on C show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.15 shows that lime dosage depends on conductivity from the EJ Smith Dam. Figure 4.14 show the scatter plots for the relationship. From the scatter plot a linear trendline was fit and the relationship had a low coefficient of determination.

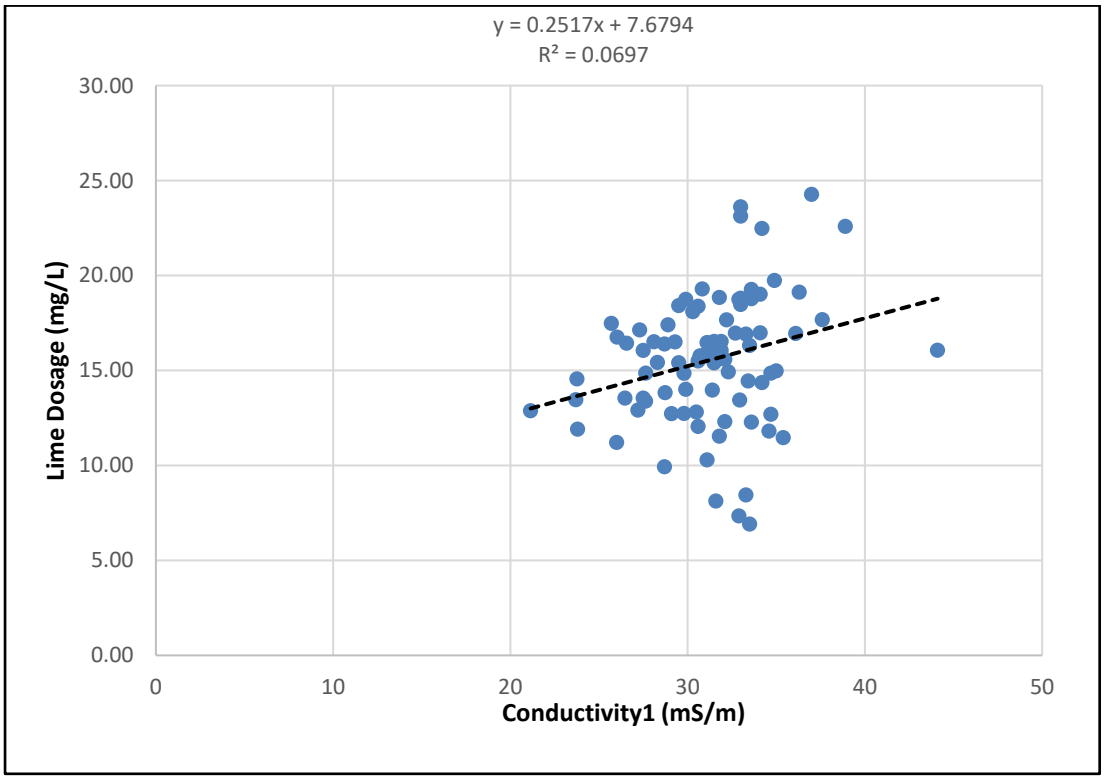


Figure 4.14: Relationship between lime dosage and conductivity1 for the period between July 2006 to June 2013

In the equation shown in Figure 4.14, the y values represent the lime dosage and the x values represent conductivity. The equation was used to formulate a general equation relating lime dosage to conductivity. The general equation for the relationship between lime dosage and conductivity1 is shown in Equation 4.17.

$$D_2 = K_3 CO_1 + K_4$$ Equation 4.17

In Equation 4.17, D_2 represents lime dosage and *CO* represents conductivity. The subscripts on *CO* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.15 shows that lime dosage depends on the manganese from the EJ Smith Dam. The scatter plot for the relationship are shown in Figure 4.15. The coefficient of determination shows that manganese is a major contributor in determining lime dosage.

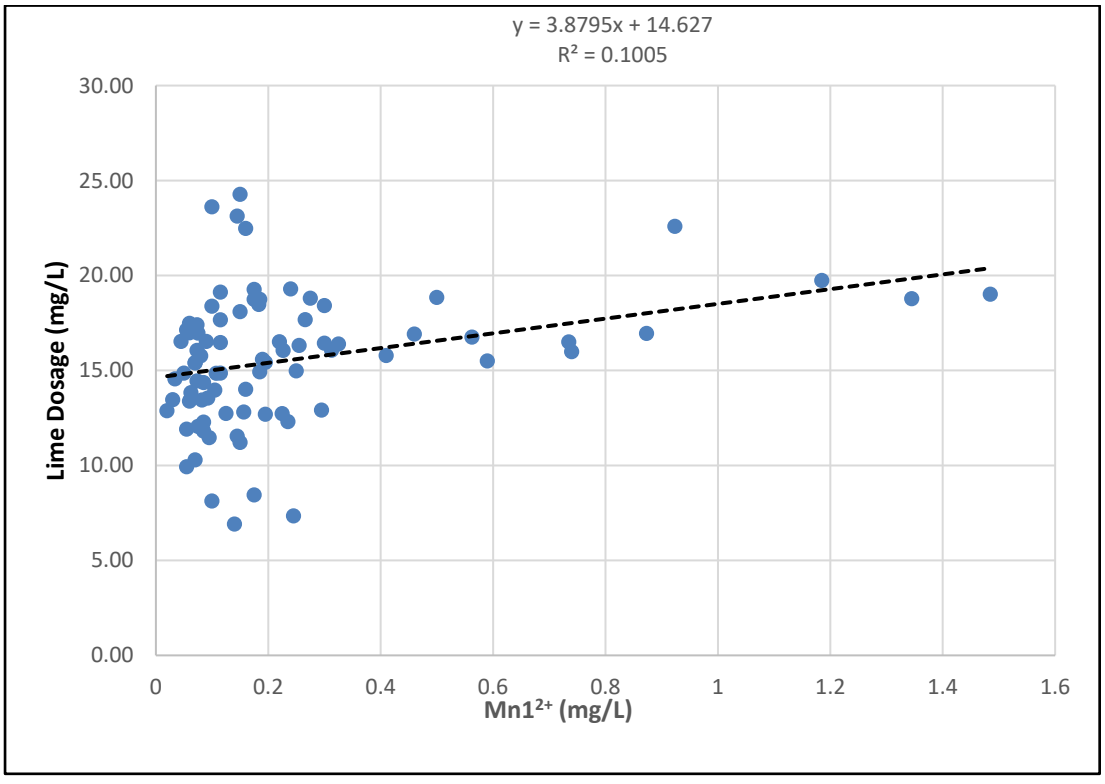


Figure 4.15: Relationship between lime dosage and manganese1 for the period between July 2006 to June 2013

In the equation shown in Figure 4.15, the y values represent the lime dosage and the x values represent manganese. The general equation for the relationship between lime dosage and manganese1 is shown in Equation 4.18.

$$D_2 = K_5 Mn_1 + K_6 \qquad \text{Equation 4.18}$$

In Equation 4.18, D_2 represents lime dosage and *Mn* represents manganese. The subscripts on *Mn* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.15 shows that lime dosage depends on total organic carbonyl (TOC) from the Umzinto River. The coefficient of determination shows that TOC is a major determinant of lime dosage but is not the only factor determining lime dosage.

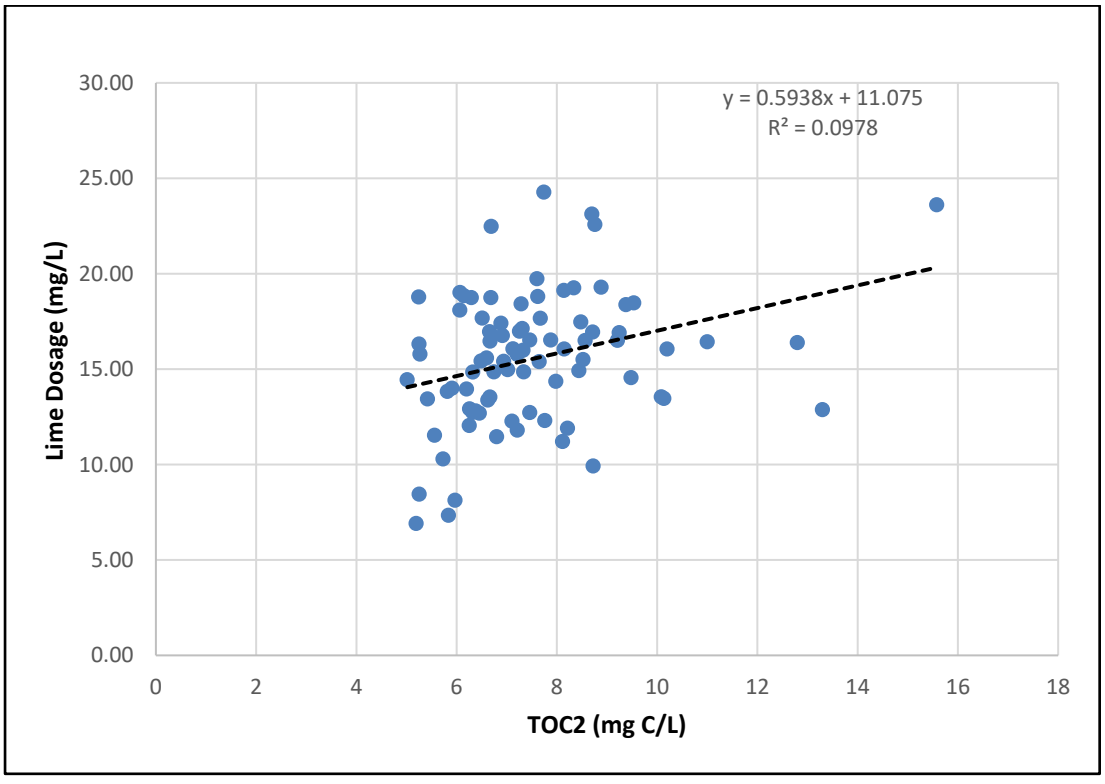


Figure 4.16: Relationship between lime dosage and TOC2 for the period between July 2006 to June 2013

In the equation shown in Figure 4.16, the y values represent the lime dosage and the x values represent TOC. This equation is then used to formulate a general equation relating lime dosage to TOC. The general equation for the relationship between lime dosage and TOC2 is shown in Equation 4.19.

$$D_2 = K_7 TC_2 + K_8 \quad \text{Equation 4.19}$$

In Equation 4.19, D_2 represents lime dosage and *TC* represents TOC. The subscript on *TC* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

To obtain the final model for calculating lime dosage , the general equations for the individual parameters are combined to produce one general equation whose coefficients would established using genetic algorithms in (MathWorks 2014). This means adding together equations 4.16, 4.17, 4.18 and 4.19 to obtain Equation 4.20.

$$D_2 = K_1 C_2 + K_3 CO_1 + K_5 M_1 + K_7 TC_2 + K_9 \quad \text{Equation 4.20}$$

Relationship between potassium *permanganate dosage and raw water quality parameters*

From the objective 1 results, potassium permanganate was seen to have a good correlation with alkalinity, coliforms, colour, iron, suspended solids and total organic carbonyl (TOC). This relationship is represented by equation 4.38 in the form

$$D_3 = f(Alk_1, Alk_2, C_1, C_2, Col_1, Fe_1, Fe_2, SS_2, TC_1, TC_2) \quad \text{Equation 4.21}$$

Equation 4.21 shows that potassium permanganate dosage depends on both the alkalinity from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.17 and 4.18 respectively. For both abstraction points potassium permanganate dosage is shown to reduce with increasing alkalinity. The coefficient of determination suggest that alkalinity has no major influence on the potassium permanganate.

In the equations shown in Figure 4.17 and 4.18, the y values represent the potassium permanganate dosage and the x values represent alkalinity. These equations were then used to formulate general equations relating potassium permanganate dosage to alkalinity. The

general equation for the relationship between potassium permanganate dosage and alkalinity1 is shown in Equation 4.22 and that for alkalinity2 is shown in Equation 4.23.

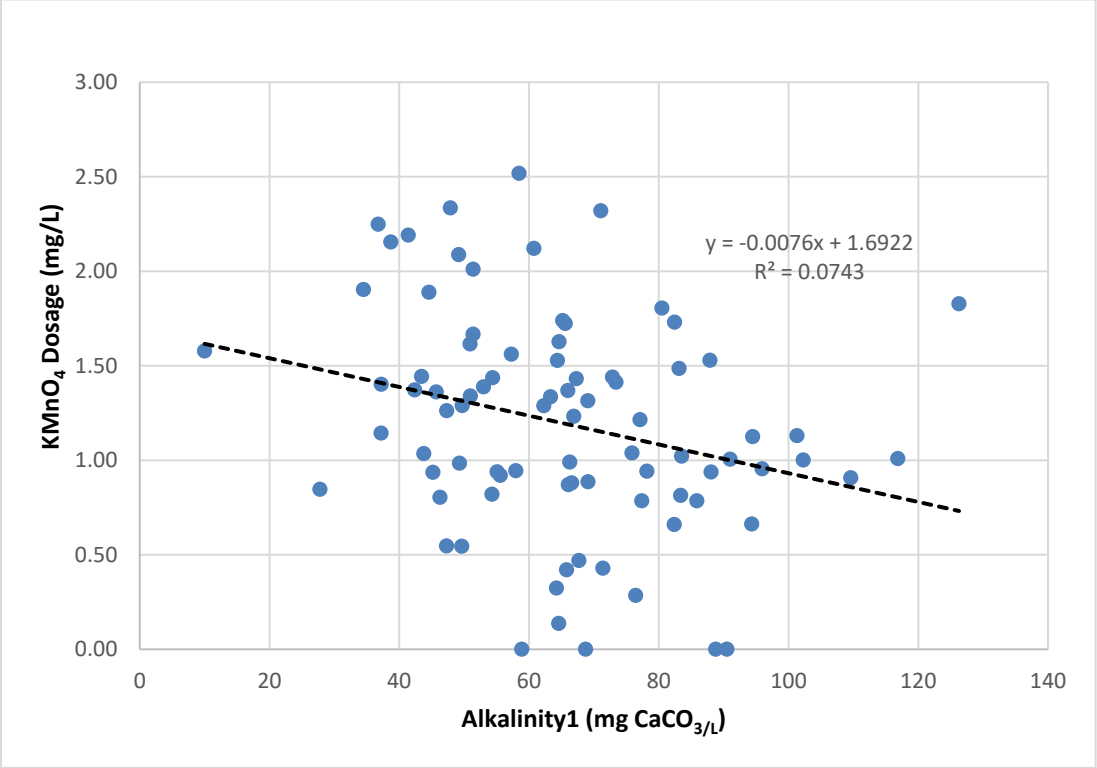


Figure 4.17: Relationship between potassium permanganate dosage and alkalinity1 for the period between July 2006 to June 2013

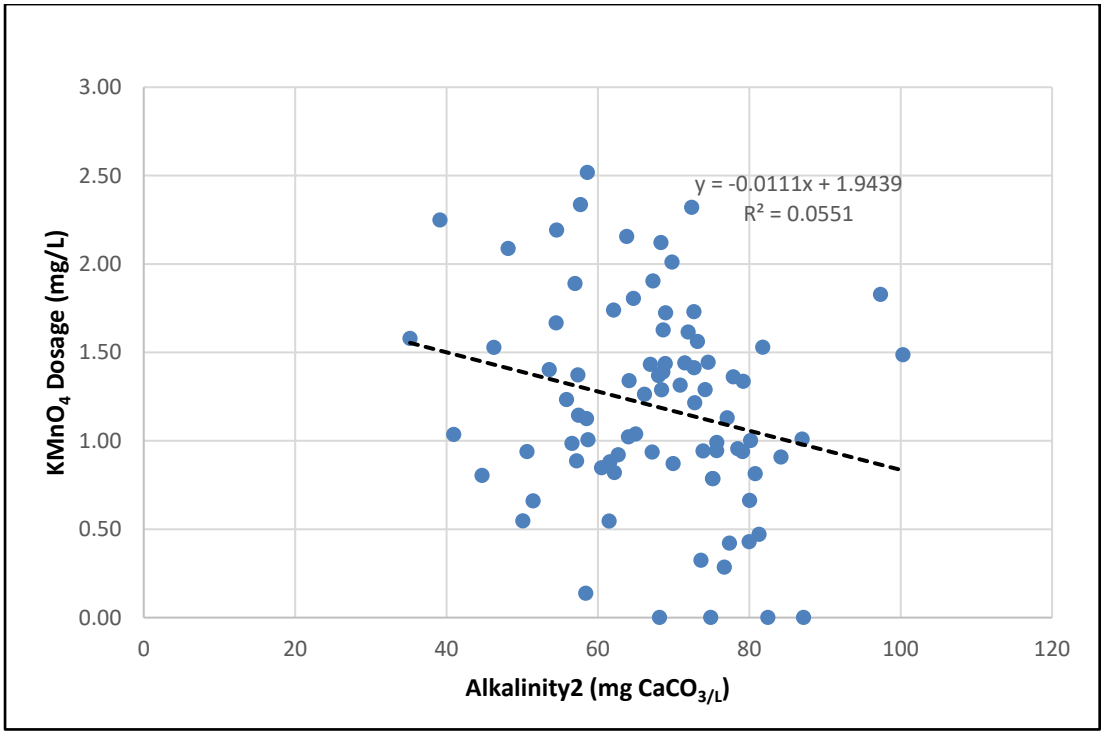


Figure 4.18: Relationship between potassium permanganate dosage and alkalinity2 for the

period between July 2006 to June 2013

$$D_3 = K_1 Alk_1 + K_2 \quad \text{Equation 4.22}$$

$$D_3 = K_3 Alk_2 + K_4 \quad \text{Equation 4.23}$$

In Equation 4.22 and 4.23, D_3 represents potassium permanganate dosage and *Alk* represents algal count. The subscripts on *Alk* show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.21 shows that potassium permanganate dosage depends on coliforms from the EJ Smith Dam. Figure 4.19 shows the scatter plot for the relationship with a coefficient of determination suggesting coliforms to be one of factors influencing potassium permanganate dosage.

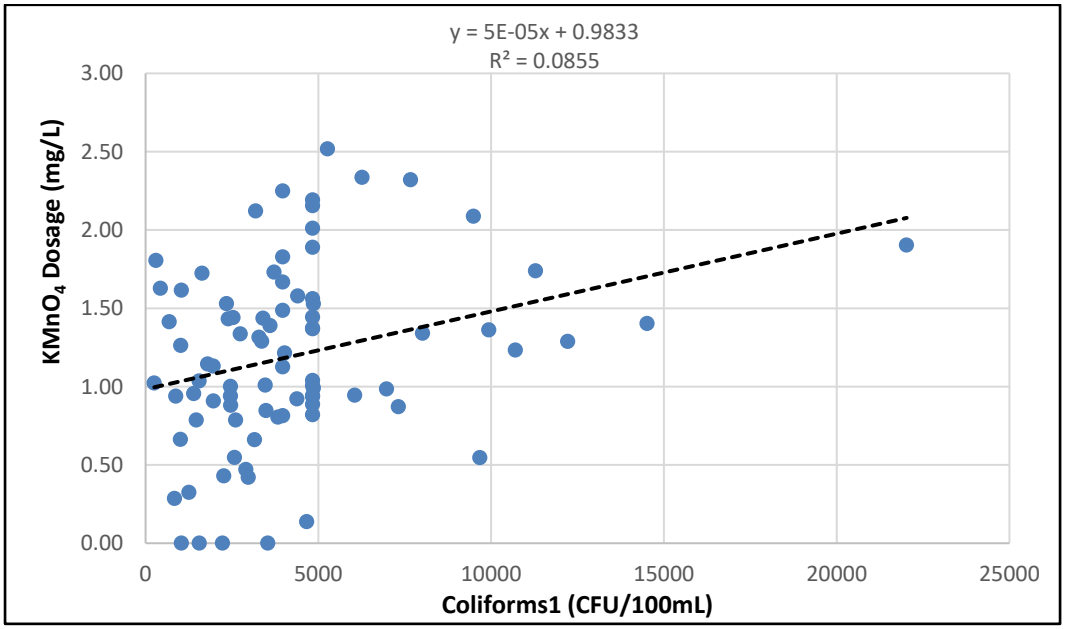


Figure 4.19: Relationship between potassium permanganate dosage and coliforms1 for the period between July 2006 to June 2013

In the equation shown in Figure 4.19, the y values represent the potassium permanganate dosage and the x values represent coliforms. This equation was then used to formulate a general equation relating potassium permanganate dosage to coliforms. The general

equation for the relationship between potassium permanganate dosage and coliforms1 is shown in Equation 4.24.

$$D_3 = K_5 Col_1 + K_6 \qquad \text{Equation 4.24}$$

In Equation 4.57, D_3 represents potassium permanganate dosage and *Col* represents coliforms. The subscripts on *Col* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.21 shows that potassium permanganate dosage depends on both the raw water colour from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.20 and 4.21, respectively.

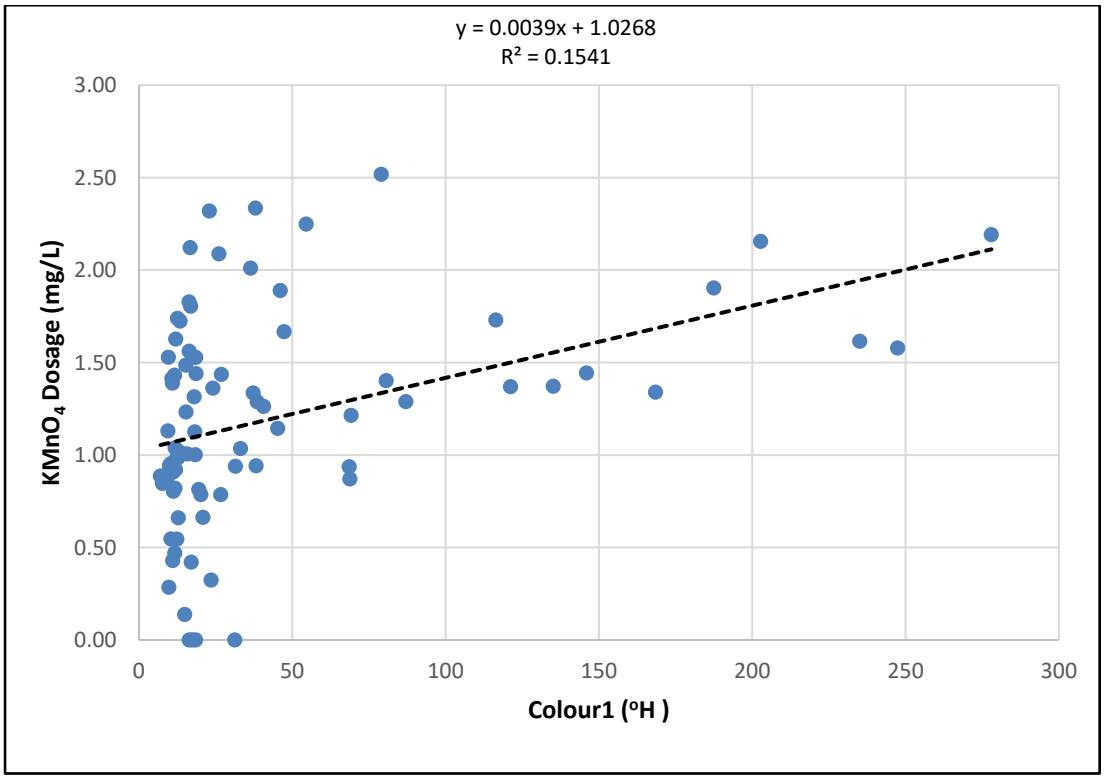


Figure 4.20: Relationship between potassium permanganate dosage and colour1 for the period between July 2006 to June 2013

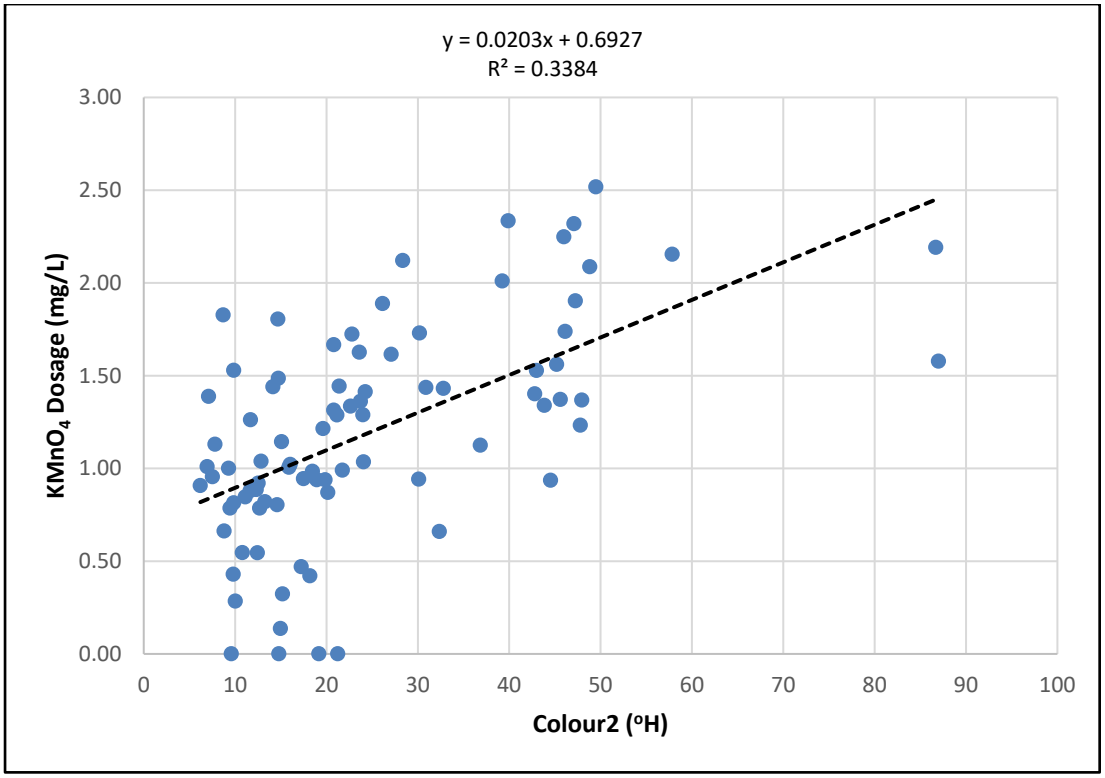


Figure 4.21: Relationship between potassium permanganate dosage and colour2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.20 and 4.21, the y values represent the potassium permanganate dosage and the x values represent colour. These equations were then used to formulate general equations relating potassium permanganate dosage to colour. The general equation for the relationship between potassium permanganate dosage and colour1 is shown in Equation 4.25 and that for colour2 is shown in Equation 4.26.

$$D_3 = K_7 C_1 + K_8 \qquad \text{Equation 4.25}$$

$$D_3 = K_9 C_2 + K_{10} \qquad \text{Equation 4.26}$$

In Equation 4.25 and 4.26, D_3 represents potassium permanganate dosage and C represents colour. The subscripts on C show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.21 shows that potassium permanganate dosage depends on both the iron from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.22 and 4.23, respectively.

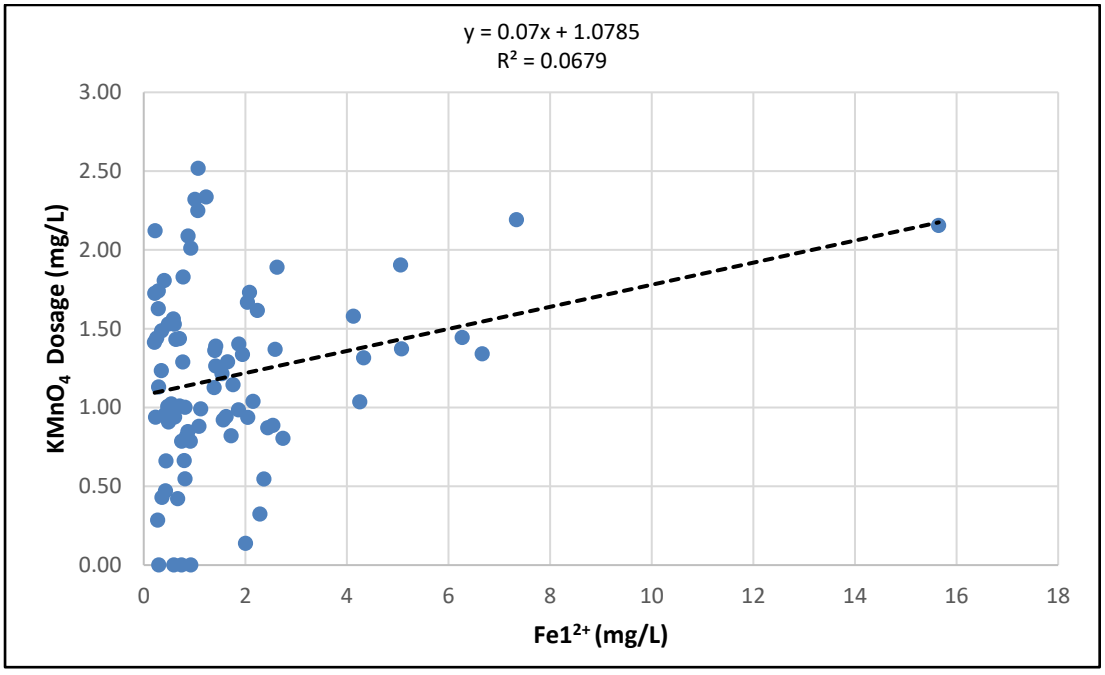


Figure 4.22: Relationship between potassium permanganate dosage and iron1 for the period between July 2006 to June 2013

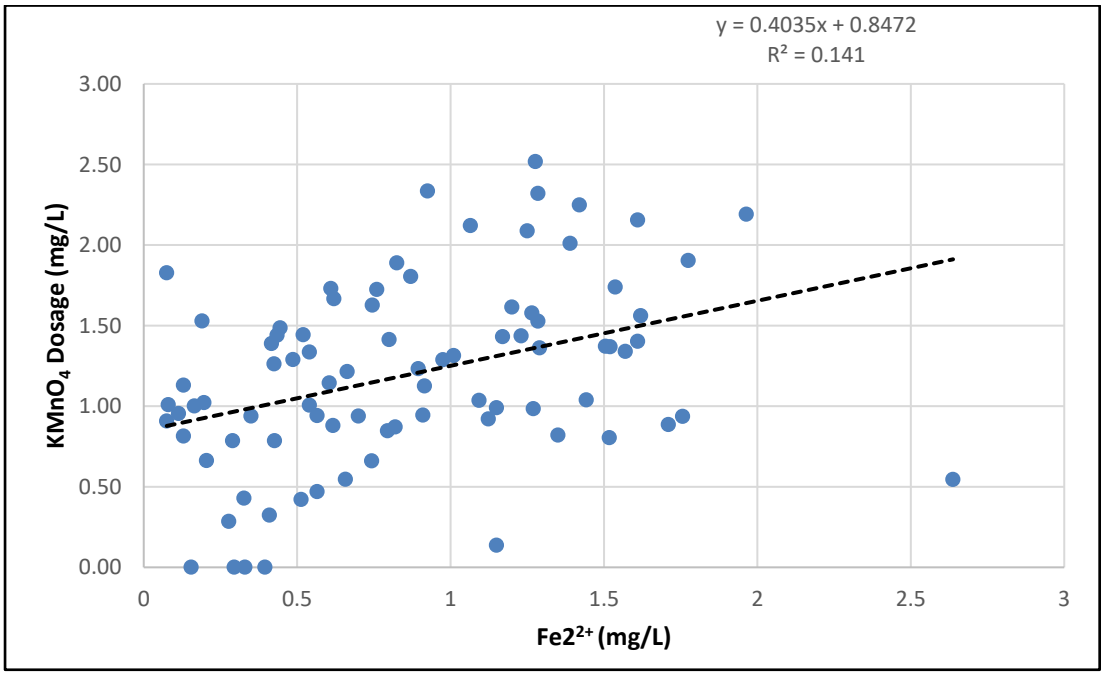


Figure 4.23: Relationship between potassium permanganate dosage and iron2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.22 and 4.23, the y values represent the potassium permanganate dosage and the x values represent iron. These equations were then used to formulate general equations relating potassium permanganate dosage to iron. The general equation for the relationship between potassium permanganate dosage and iron1 is shown in Equation 4.27 and that for iron2 is shown in Equation 4.28.

$$D_3 = K_{11}Fe_1 + K_{12} \quad \text{Equation 4.27}$$

$$D_3 = K_{13}Fe_2 + K_{14} \quad \text{Equation 4.28}$$

In Equation 4.27 and 4.28, D_3 represents potassium permanganate dosage and *Fe* represents iron. The subscripts on *Fe* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.21 shows that potassium permanganate dosage depends on both the raw water suspended solids from the Umzinto River. Figure 4.24 shows the coefficient of determination for this relationship.

In the equation shown in Figure 4.24, the y values represent the potassium permanganate dosage and the x values represent suspended solids. The general equation for the relationship between potassium permanganate dosage and suspended solids2 is shown in Equation 4.29.

$$D_3 = K_6SS_2 + K_7 \quad \text{Equation 4.29}$$

In Equation 4.29, D_3 represents potassium permanganate dosage and *SS* represents algal count. The subscripts on *SS* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

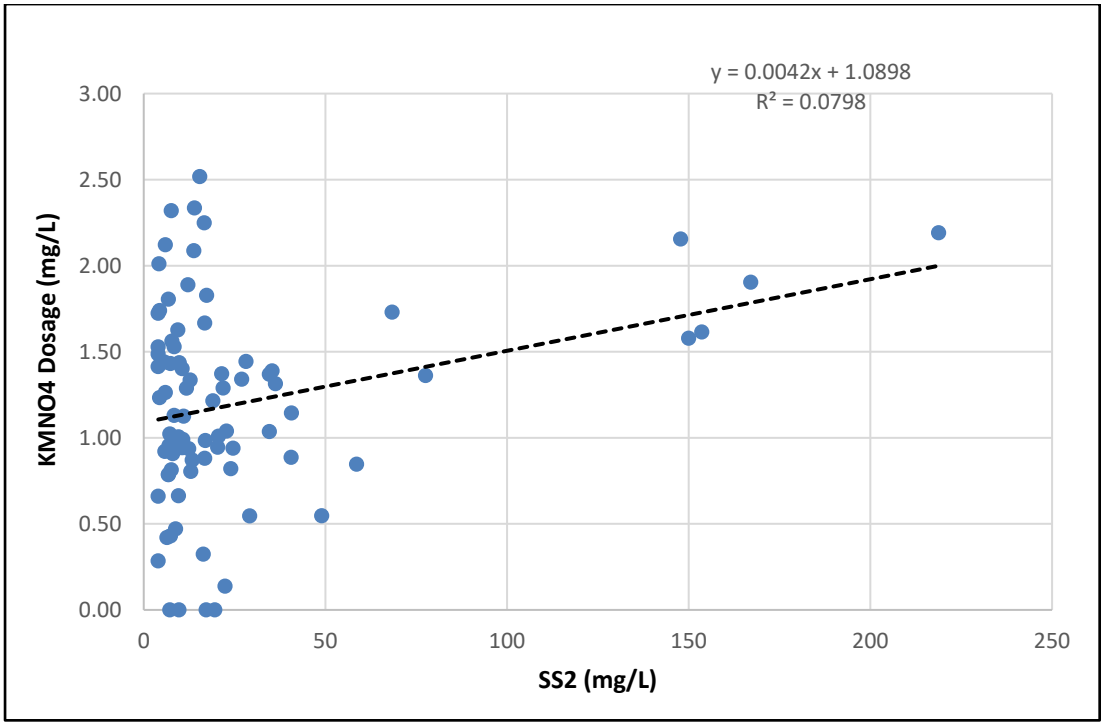


Figure 4.24: Relationship between potassium permanganate dosage and suspended solids2 for the period between July 2006 to June 2013

Equation 4.21 shows that potassium permanganate dosage depends on both the total organic carbonyl (TOC) from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure4.25 and 4.26, respectively.

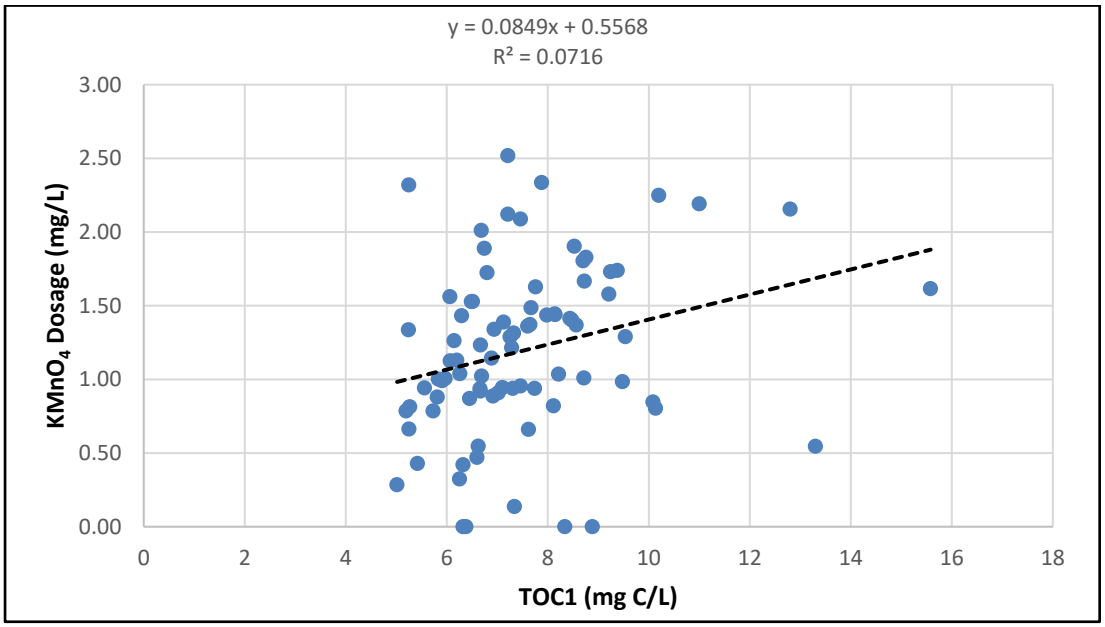


Figure 4.25: Relationship between potassium permanganate dosage and total organic carbon1 for the period between July 2006 to June 2013

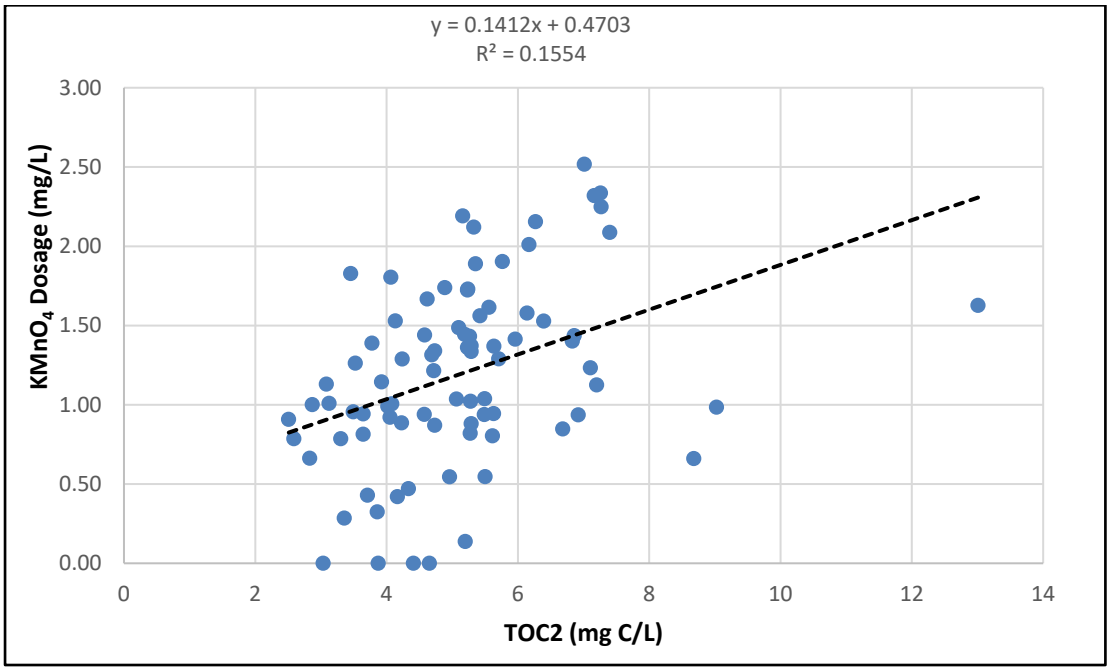


Figure 4.26: Relationship between potassium permanganate dosage and total organic carbon2 for the period between July 2006 to June 2013

In the equations shown in Figure4.25 and 4.26, the y values represent the potassium permanganate dosage and the x values represent TOC. These equations were then used to formulate general equations relating potassium permanganate dosage to TOC. The general equation for the relationship between potassium permanganate dosage and TOC1 is shown in Equation 4.30 and that for TOC2 is shown in Equation 4.31.

$$D_3 = K_6 TC_1 + K_7 \qquad \text{Equation 4.30}$$

$$D_3 = K_6 TC_1 + K_7 \qquad \text{Equation 4.31}$$

In Equation 4.30 and 4.31, D_3 represents potassium permanganate dosage and *TC* represents TOC. The subscripts on *TC* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

To obtain the final model for calculating potassium permanganate dosage, the general equations for the individual parameters are combined to produce one general equation

whose coefficients would be established using genetic algorithms in (MathWorks 2014). This means adding together equations 4.22, 4.23, 4.24, 4.25, 4.26, 4.27, 4.28, 4.30 and 4.31 to obtain Equation 4.32.

$$D_3 = K_1 Alk_1 + K_3 Alk_2 + K_5 C_1 + K_7 C_2 + K_9 Col_1 + K_{11} Fe_1 + K_{12} Fe_2 + K_{13} SS_2 + K_{14} TC_1 + K_{15} TC_2 + K_{16} \quad \text{Equation 4.32}$$

Relationship between polymer dosage and raw water quality parameters

Table 4.3 shows that the polymer dosage is related to the levels of algal count, alkalinity, coliforms, colour, conductivity, *Escherichia coli*, iron, hardness, manganese, pH, suspended solids, temperature, total organic carbon and turbidity. To develop a model for predicting the polymer dosage the method of scatter plots and basic curve fitting was used for each parameter.

An equation relating polymer dosage to the parameters is shown in Equation 4.33 below as:

$$D_4 = f(AC_1, AC_2, Alk_1, Alk_2, Col_1, C_1, C_2, CO_1, CO_2, Eco_2, Fe_1, Fe_2, H_1, H_2, M_2, pH_1, S_1, S_2, T_1, T_2, TC_1, TC_2, Tu_1, Tu_2) \quad \text{Equation 4.33}$$

Equation 4.33 shows that polymer dosage depends on both the algal count from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.27 and 4.28, respectively.

In the equations shown in Figure 4.27 and 4.28, the y values represent the polymer dosage and the x values represent algal count. These equations were then used to formulate general equations relating polymer dosage to algal count. The general equation for the relationship between polymer dosage and algal count1 is shown in Equation 4.34 and that for algal count2 is shown in Equation 4.35.

$$D_4 = K_1 AC_1 + K_2 \quad \text{Equation 4.34}$$

$$D_4 = K_3 AC_2 + K_4$$ Equation 4.35

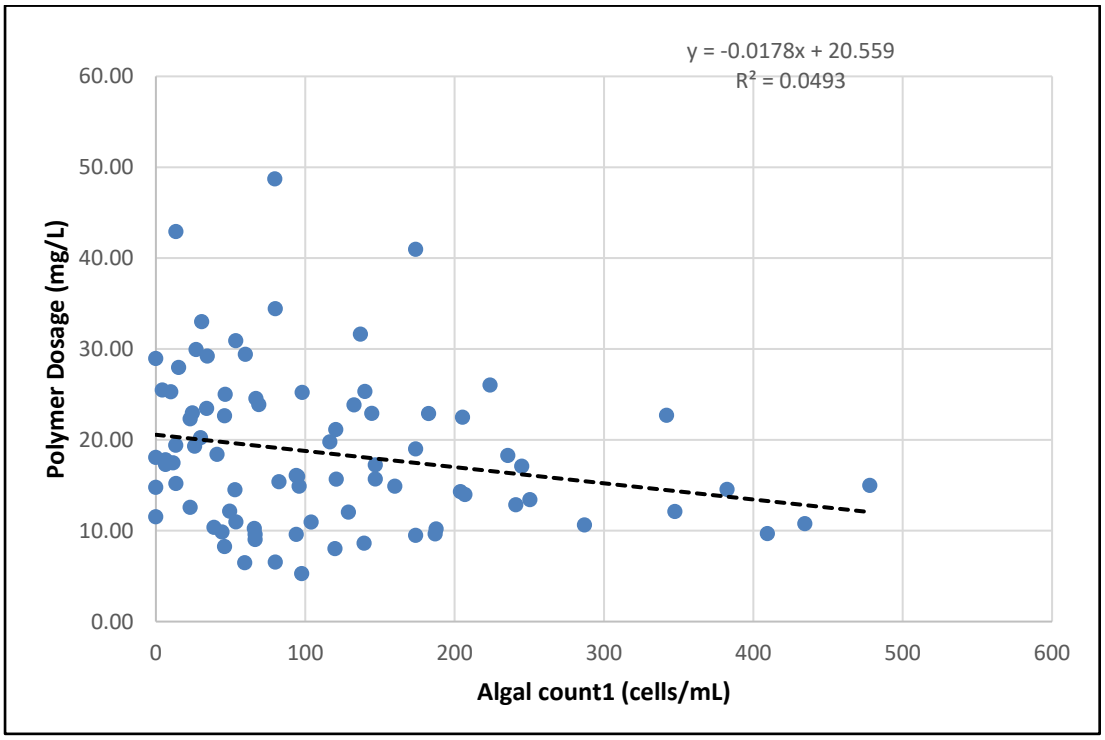


Figure 4.27: Relationship between polymer dosage and algal count1 for the period between July 2006 to June 2013

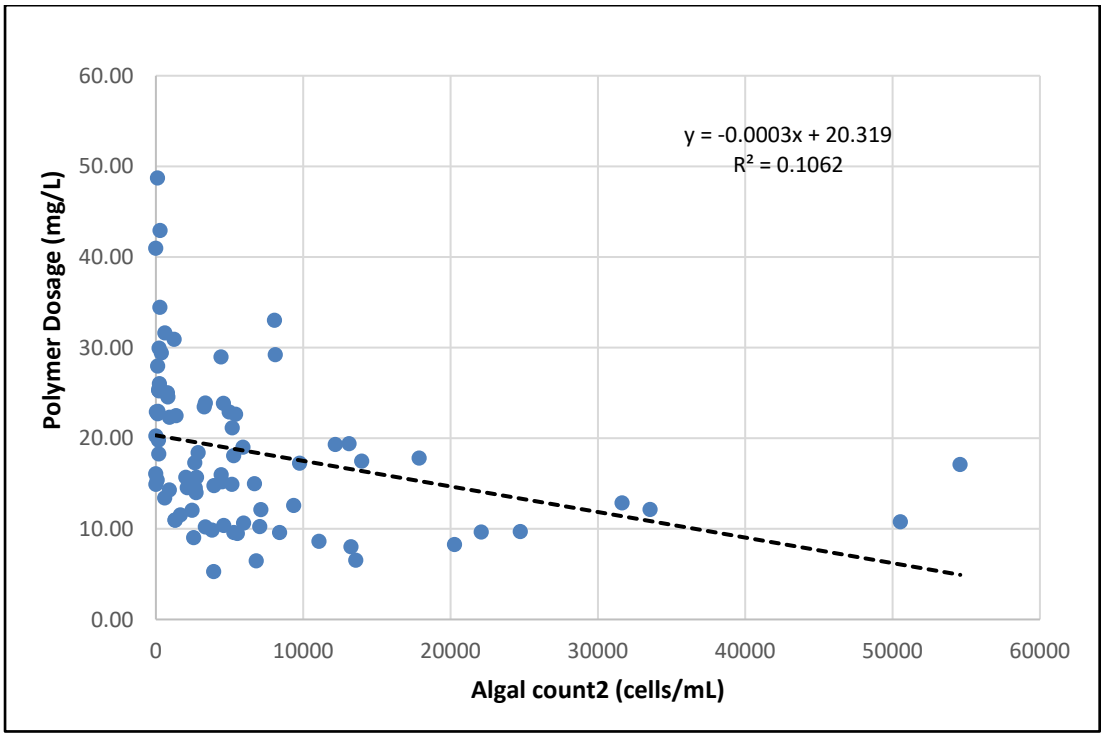


Figure 4.28: Relationship between polymer dosage and algal count2 for the period between July 2006 to June 2013

In Equation 4.34 and 4.35, D_4 represents polymer dosage and *AC* represents algal count. The subscripts on *AC* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on both the alkalinity from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.29 and 4.30, respectively. For both abstraction points, the relationship show a good coefficient of determination.

In the equations shown in Figure 4.29 and 4.30, the y values represent the polymer dosage and the x values represent alkalinity. These equations were then used to formulate general equations relating polymer dosage to alkalinity. The general equation for the relationship between polymer dosage and alkalinity1 is shown in Equation 4.36 and that for alkalinity2 is shown in Equation 4.37.

$$D_4 = K_5 Alk_1 + K_6 \qquad \text{Equation 4.36}$$

$$D_4 = K_7 Alk_2 + K_8 \qquad \text{Equation 4.37}$$

In Equation 4.36 and 4.37, D_4 represents polymer dosage and *Alk* represents alkalinity. The subscripts on *Alk* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

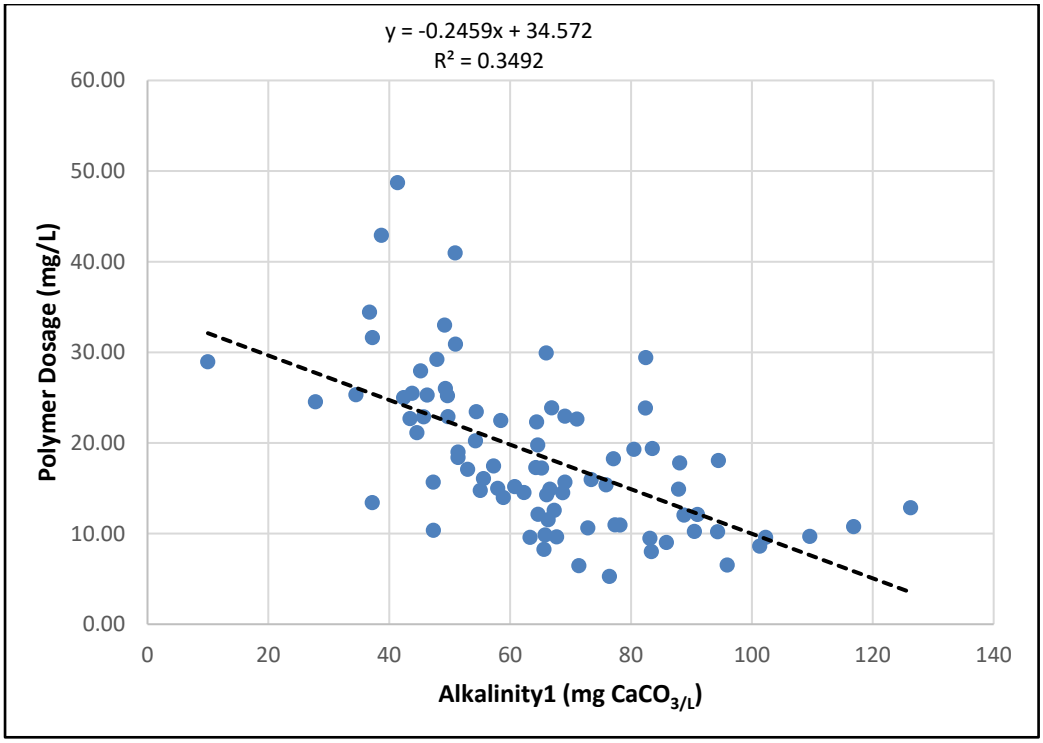


Figure 4.29: Relationship between polymer dosage and alkalinity1 for the period between July 2006 to June 2013

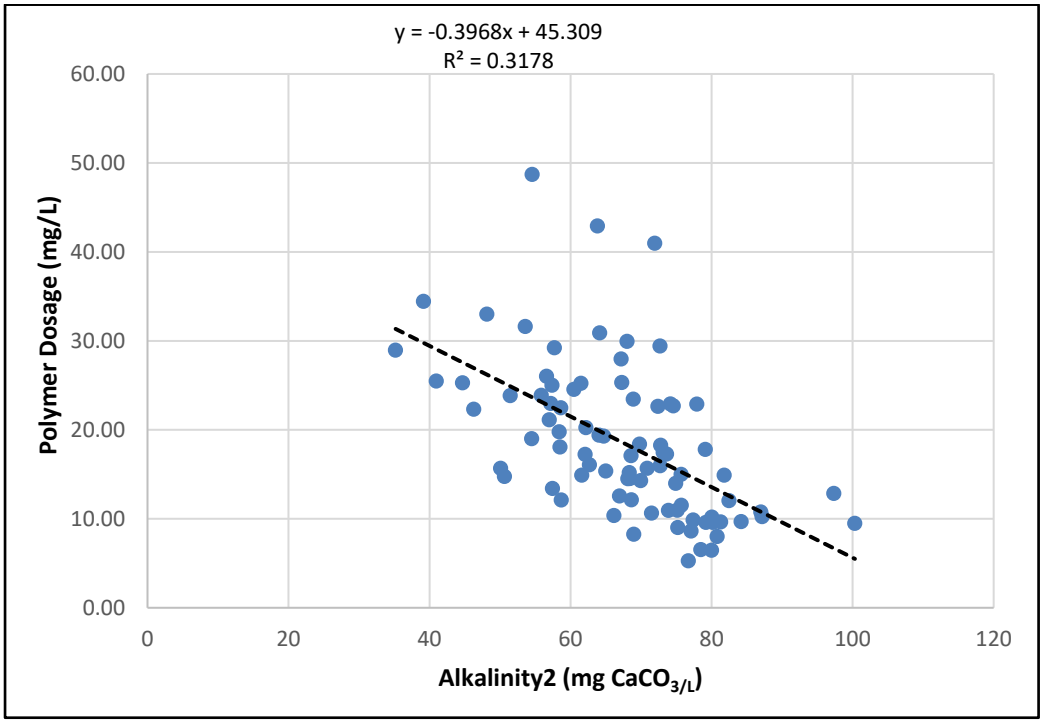


Figure 4.30: Relationship between polymer dosage and alkalinity2 for the period between July 2006 to June 2013

Equation 4.33 shows that polymer dosage depends on coliforms from the EJ Smith Dam. The scatter plot for the relationship is shown in Figure 4.31.

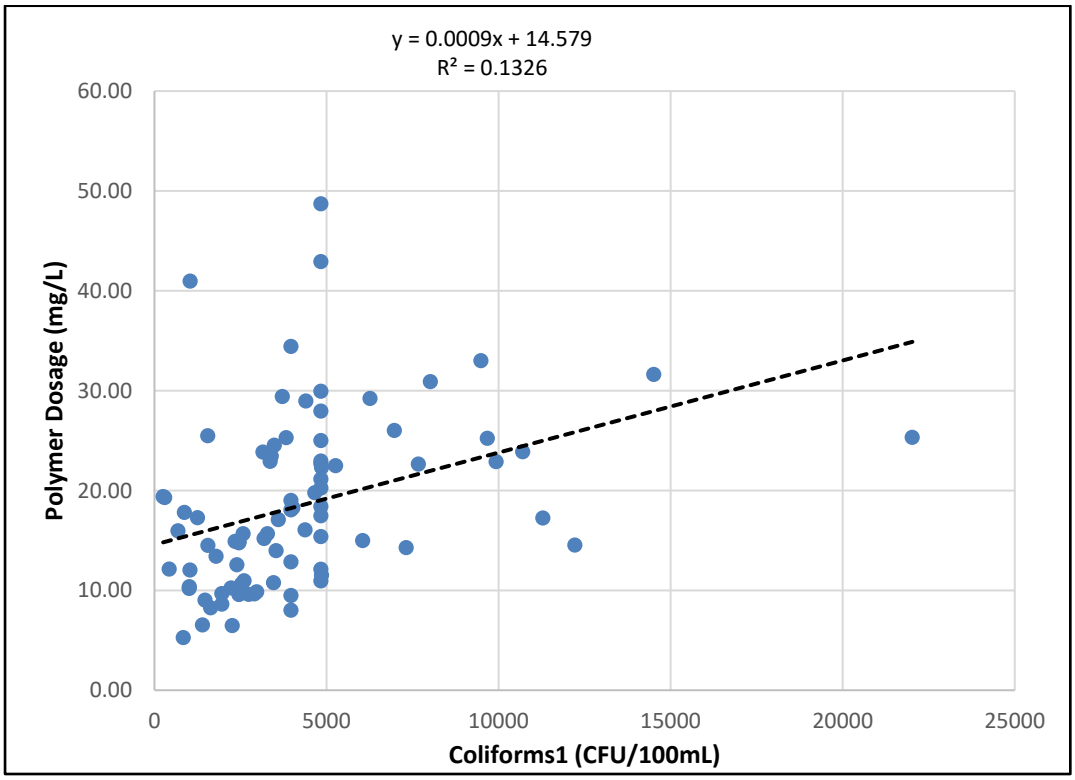


Figure 4.31: Relationship between polymer dosage and coliforms1 for the period between July 2006 to June 2013

In the equation shown in Figure 4.31, the y values represent the polymer dosage and the x values represent coliforms. The general equation for the relationship between polymer dosage and coliforms1 is shown in Equation 4.38.

$$D_4 = K_9 Col_1 + K_{10} \qquad \text{Equation 4.38}$$

In Equation 4.79 and 4.80, D_4 represents polymer dosage and *Col* represents coliforms. The subscripts on *Col* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on both the raw water colour from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown

in Figure 4.32 and 4.33, respectively.

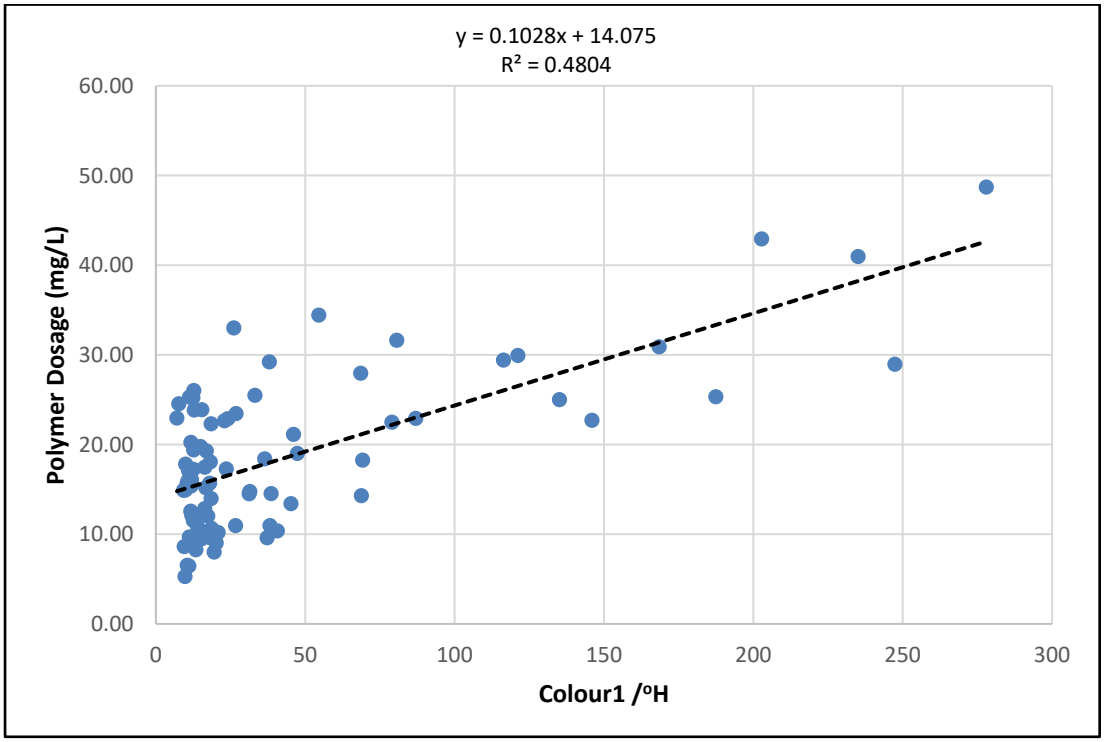


Figure 4.32: Relationship between polymer dosage and colour1 for the period between July 2006 to June 2013

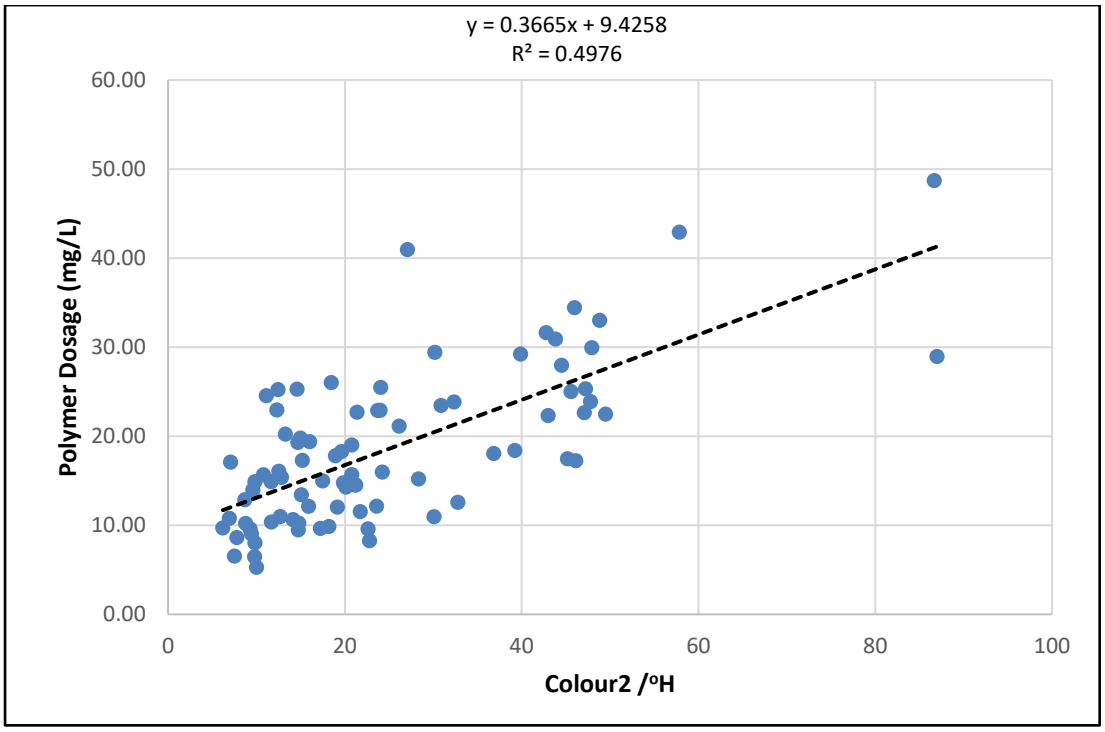


Figure 4.33: Relationship between polymer dosage and colour2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.32 and 4.33, the y values represent the polymer dosage and the x values represent colour. These equations were then used to formulate general equations relating polymer dosage to colour. The general equation for the relationship between polymer dosage and colour1 is shown in Equation 4.39 and that for colour2 is shown in Equation 4.40.

$$D_4 = K_{11}C_1 + K_{12} \quad \text{Equation 4.39}$$

$$D_4 = K_{13}C_2 + K_{14} \quad \text{Equation 4.40}$$

In Equation 4.39 and 4.40, D_4 represents polymer dosage and C represents colour. The subscripts on C show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on both the conductivity from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.34 and 4.35, respectively.

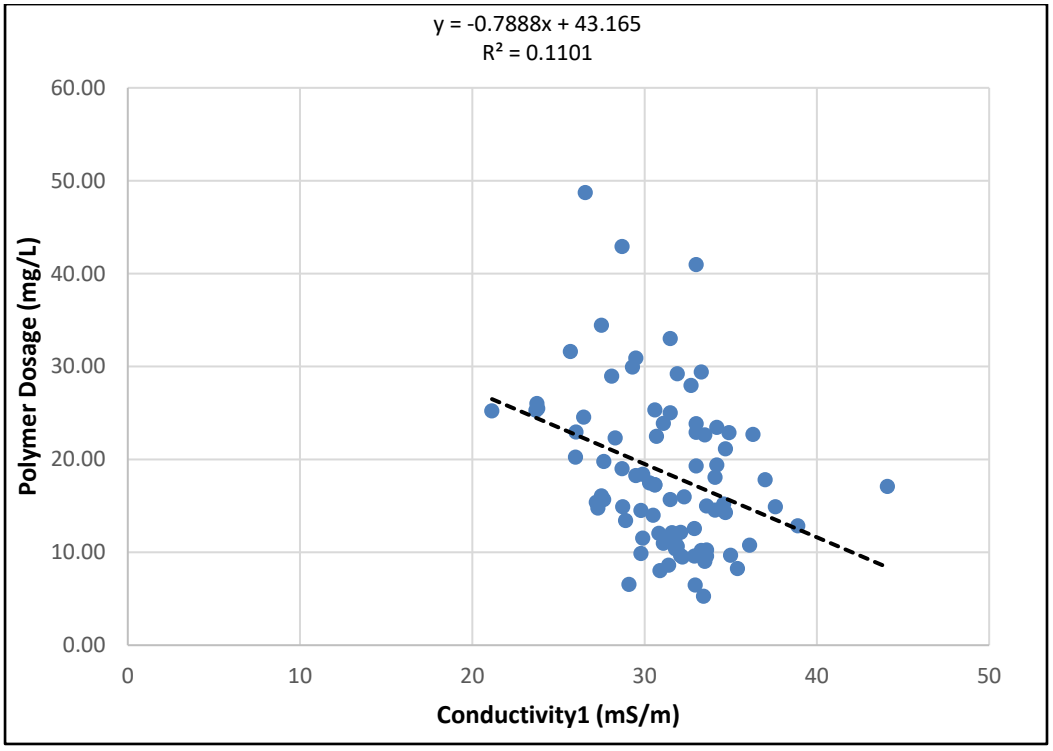


Figure 4.34: Relationship between polymer dosage and conductivity1 for the period between July 2006 to June 2013

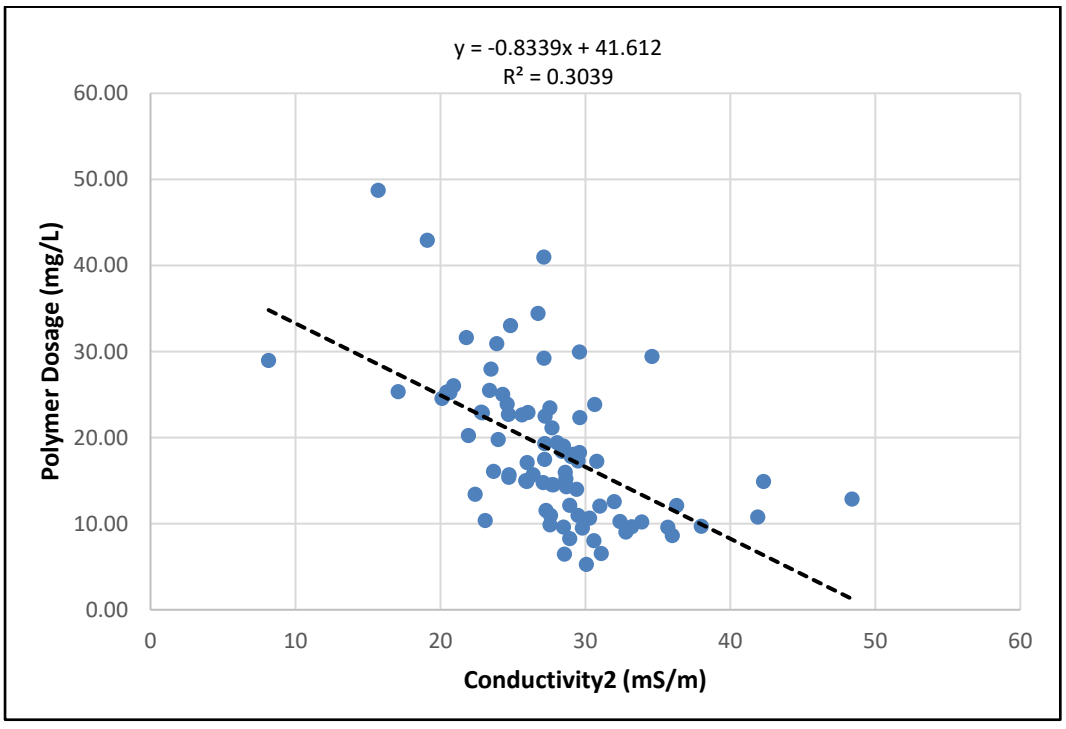


Figure 4.35: Relationship between polymer dosage and conductivity2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.34 and 4.35, the y values represent the polymer dosage and the x values represent conductivity. These equations were then used to formulate general equations relating polymer dosage to conductivity. The general equation for the relationship between polymer dosage and conductivity1 is shown in Equation 4.41 and that for conductivity2 is shown in Equation 4.42.

$$D_4 = K_{15}CO_1 + K_{16} \qquad \text{Equation 4.41}$$

$$D_4 = K_{17}CO_2 + K_{18} \qquad \text{Equation 4.42}$$

In Equation 4.41 and 4.42, D_4 represents polymer dosage and *CO* represents conductivity. The subscripts on *CO* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on *E.coli* from the Umzinto River. The scatter plot for the relationship is shown in Figure 4.36.

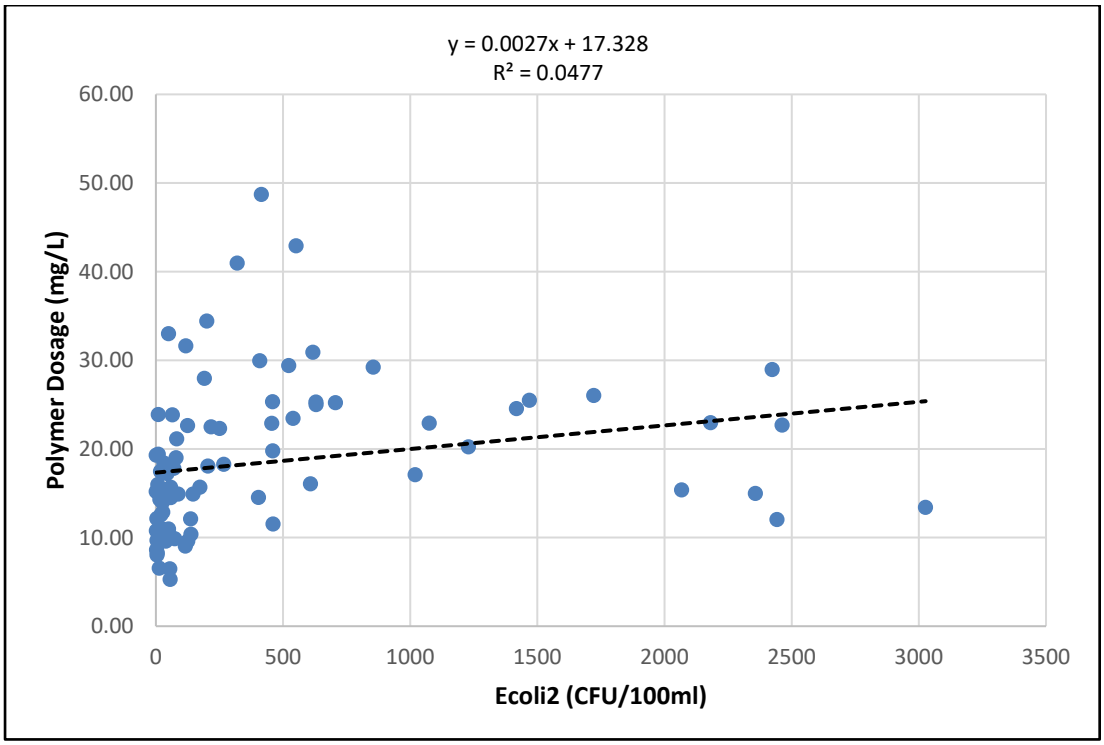


Figure 4.36: Relationship between polymer dosage and *Ecoli2* for the period between July 2006 to June 2013

In the equation shown in Figure 4.36, the y values represent the polymer dosage and the x values represent *E.coli*. The general equation for the relationship between polymer dosage and *E.coli*2 is shown in Equation 4.43.

$$D_4 = K_{19}Eco_2 + K_{20} \qquad \text{Equation 4.43}$$

In Equation 4.43, D_4 represents polymer dosage and *Eco* represents *E.coli*. The subscripts on *Eco* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on both the iron from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.37 and 4.38, respectively.

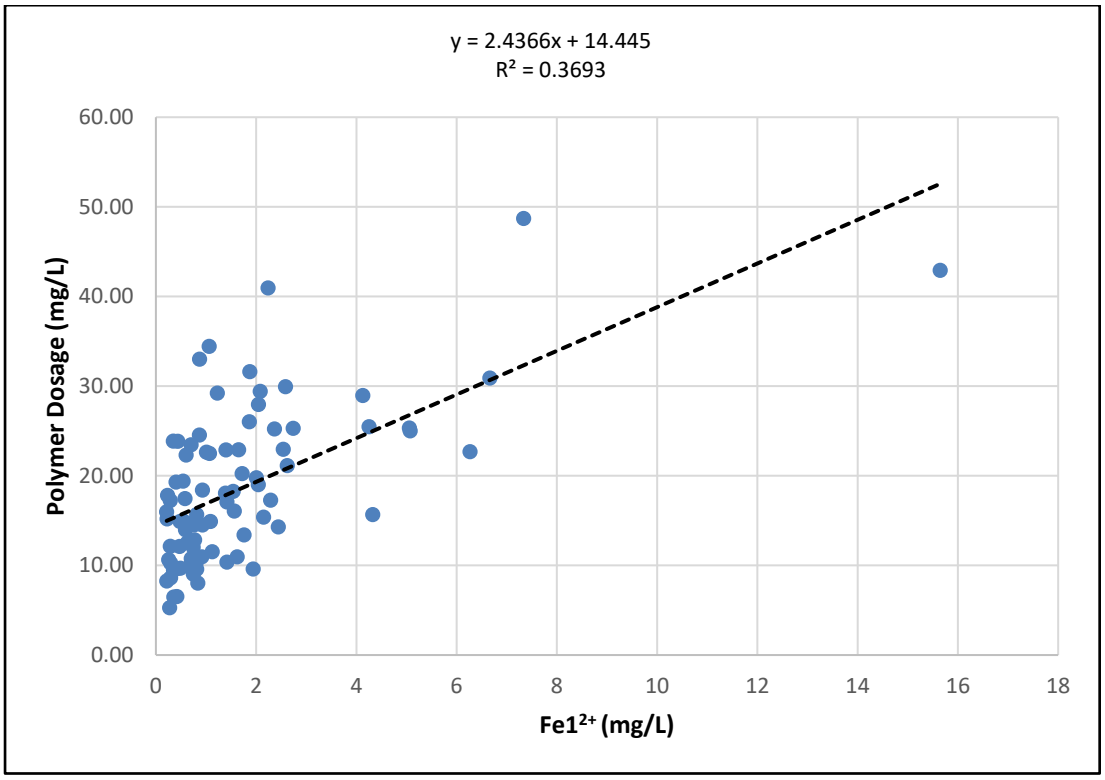


Figure 4.37: Relationship between polymer dosage and iron1 for the period between July 2006 to June 2013

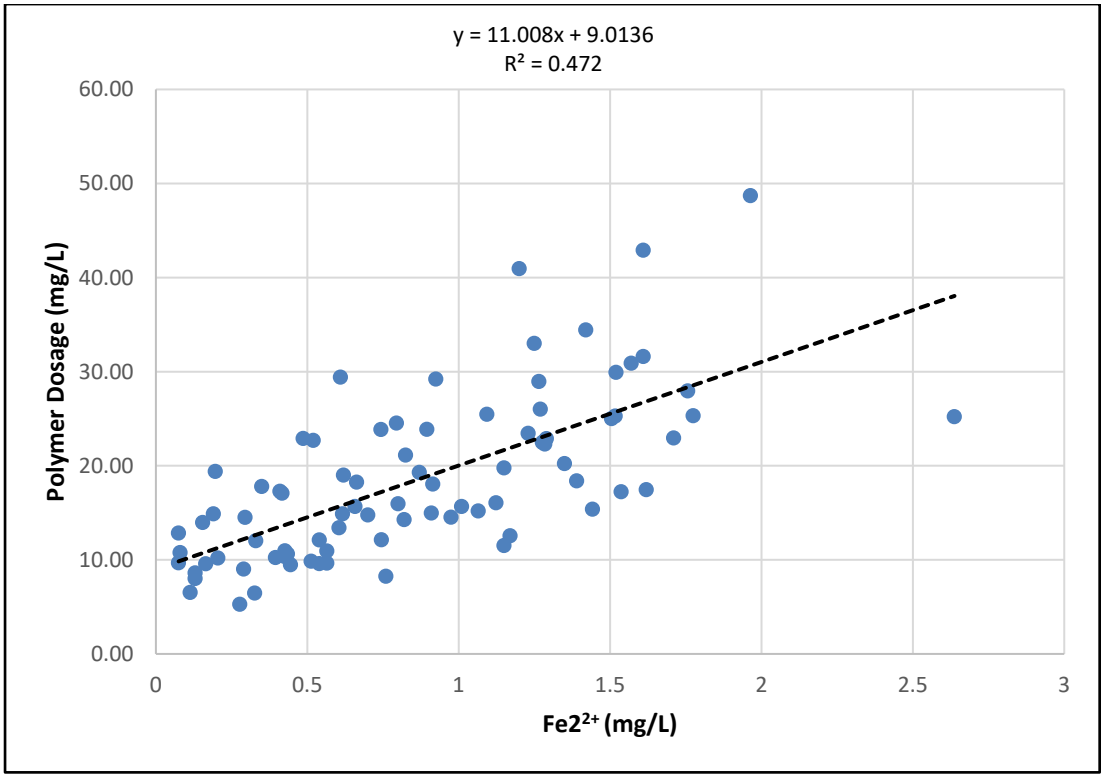


Figure 4.38: Relationship between polymer dosage and iron2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.37 and 4.38, the y values represent the polymer dosage and the x values represent iron. These equations were then used to formulate general equations relating polymer dosage to iron. The general equation for the relationship between polymer dosage and iron1 is shown in Equation 4.44 and that for iron2 is shown in Equation 4.45.

$$D_4 = K_{21}Fe_1 + K_{22} \qquad \text{Equation 4.44}$$

$$D_4 = K_{23}Fe_2 + K_{24} \qquad \text{Equation 4.45}$$

In Equation 4.44 and 4.45, D_4 represents polymer dosage and Fe represents iron. The subscripts on Fe show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on both the raw water hardness from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.39 and 4.40, respectively.

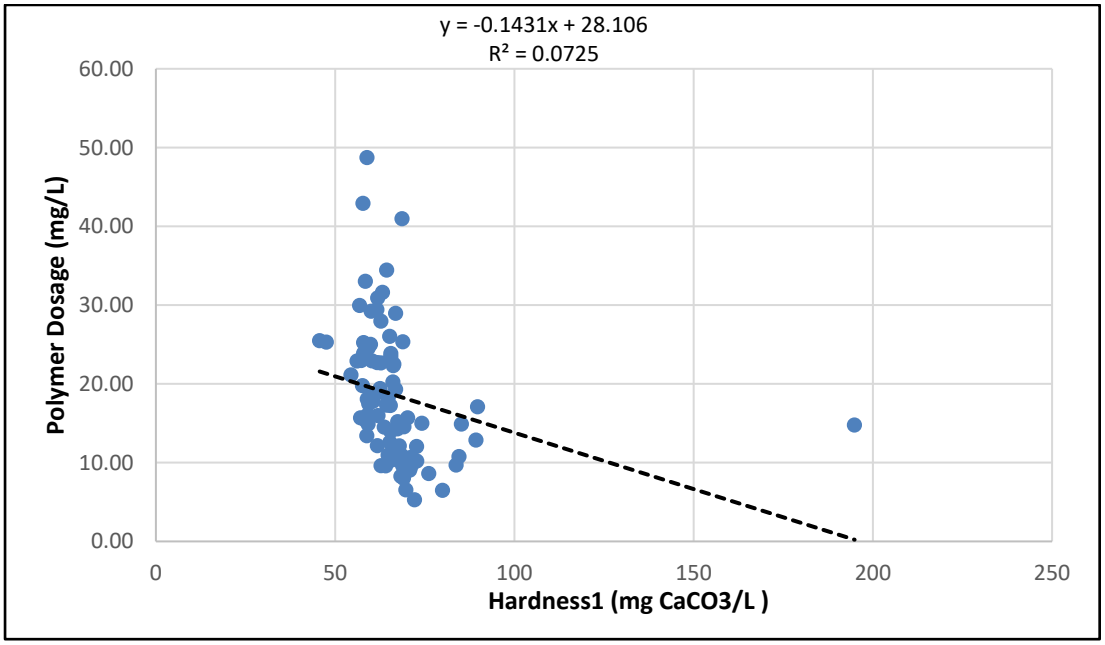


Figure 4.39: Relationship between polymer dosage and Hardness1 for the period between July 2006 to June 2013

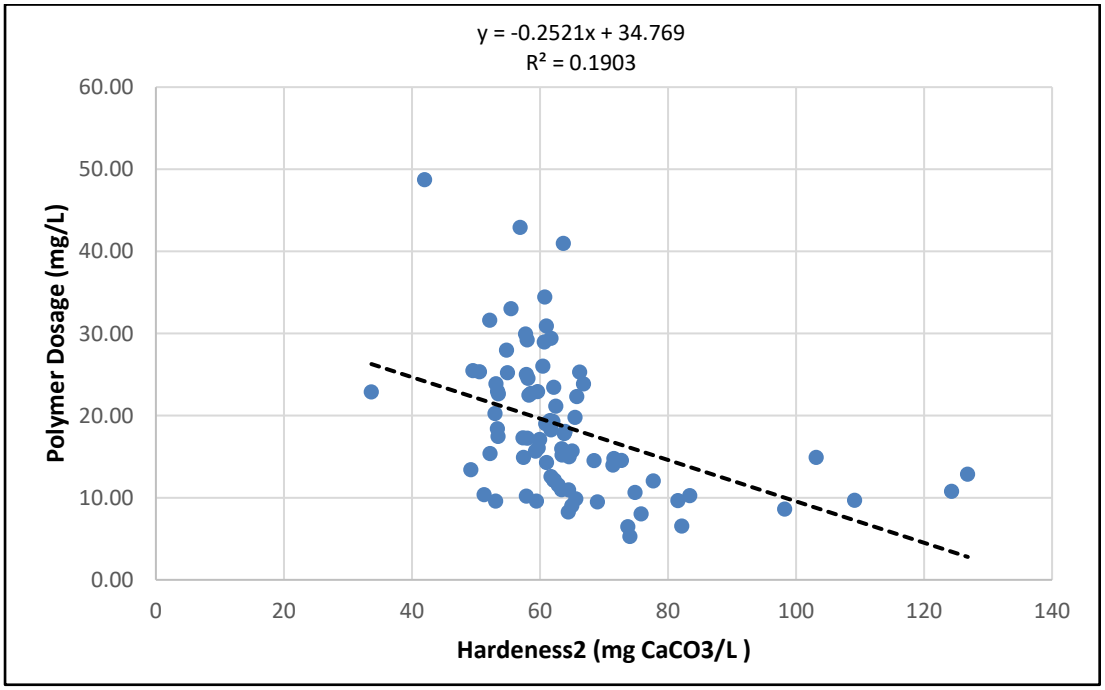


Figure 4.40: Relationship between polymer dosage and Hardness2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.39 and 4.40, the y values represent the polymer dosage and the x values represent hardness. These equations were then used to formulate general equations relating polymer dosage to hardness. The general equation for the relationship between polymer dosage and hardness1 is shown in Equation 4.46 and that for hardness2 is shown in Equation 4.47.

$$D_4 = K_{25}H_1 + K_{26} \quad \text{Equation 4.46}$$

$$D_4 = K_{27}H_2 + K_{28} \quad \text{Equation 4.47}$$

In Equation 4.46 and 4.47, *D4* represents polymer dosage and *H* represents hardness. The subscripts on *H* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on the manganese from the Umzinto River. A scatter plot for the relationship shown in Figure 4.41 shows that polymer dosage is related to the manganese.

In the equation shown in Figure 4.41, the y values represent the polymer dosage and the x values represent hardness. This equation was then used to formulate the general equation relating polymer dosage to manganese. The general equation for the relationship between polymer dosage and manganese2 is shown in Equation 4.48.

$$D_4 = K_{29}Mn_2 + K_{30} \quad \text{Equation 4.48}$$

In Equation 4.48, D_4 represents polymer dosage and *Mn* represents manganese. The subscripts on *Mn* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

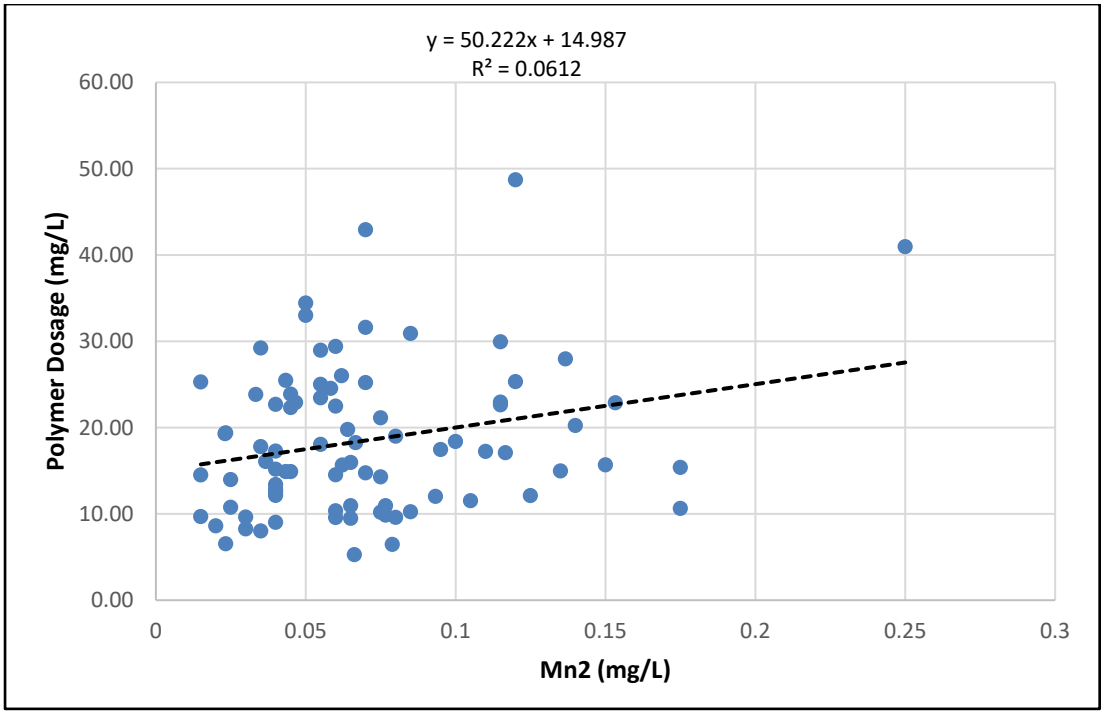


Figure 4.41: Relationship between polymer dosage and manganese2 for the period between July 2006 to June 2013

Equation 4.33 shows that polymer dosage depends on pH from the EJ Smith Dam. The relationships are shown in Figure 4.42.

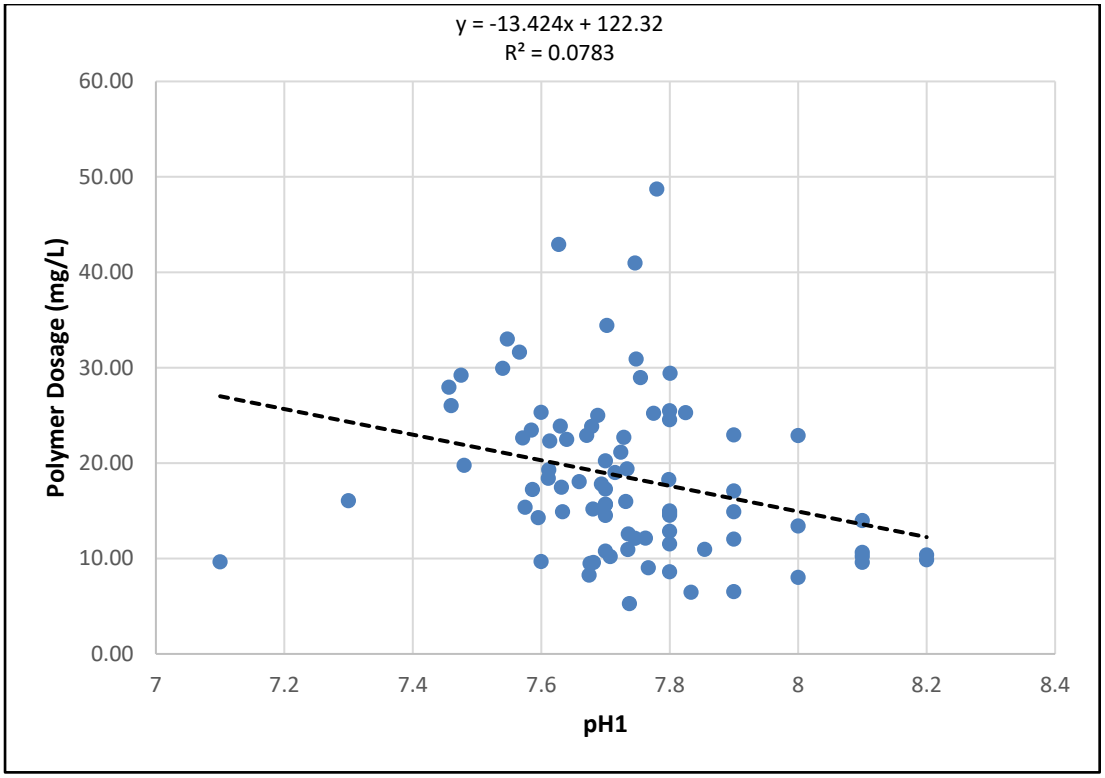


Figure 4.42: Relationship between polymer dosage and pH1 for the period between July 2006 to June 2013

In the equations shown in Figure 4.42, the y values represent the polymer dosage and the x values represent pH. The general equation for the relationship between polymer dosage and pH1 is shown in Equation 4.49.

$$D_4 = K_{31} pH_1 + K_{32} \qquad \text{Equation 4.49}$$

In Equation 4.49, D_4 represents polymer dosage. The subscripts on pH show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on both the suspended solids from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.43 and 4.44 respectively.

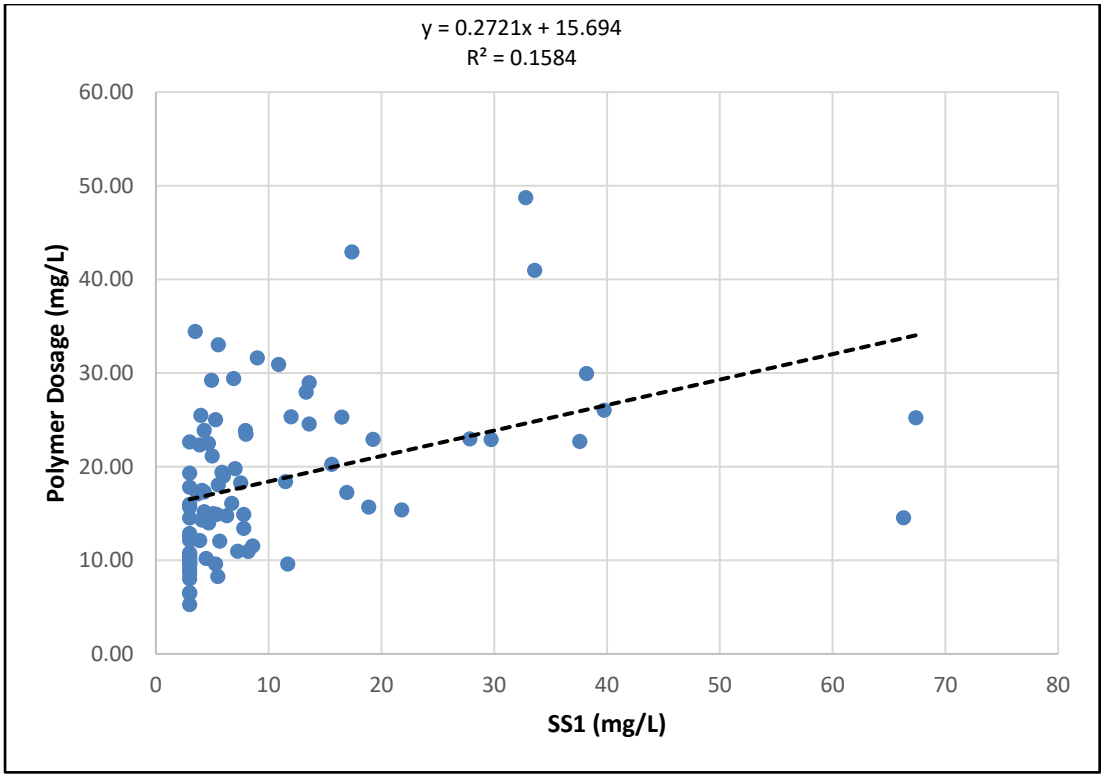


Figure 4.43: Relationship between polymer dosage and suspended solids1 for the period between July 2006 to June 2013

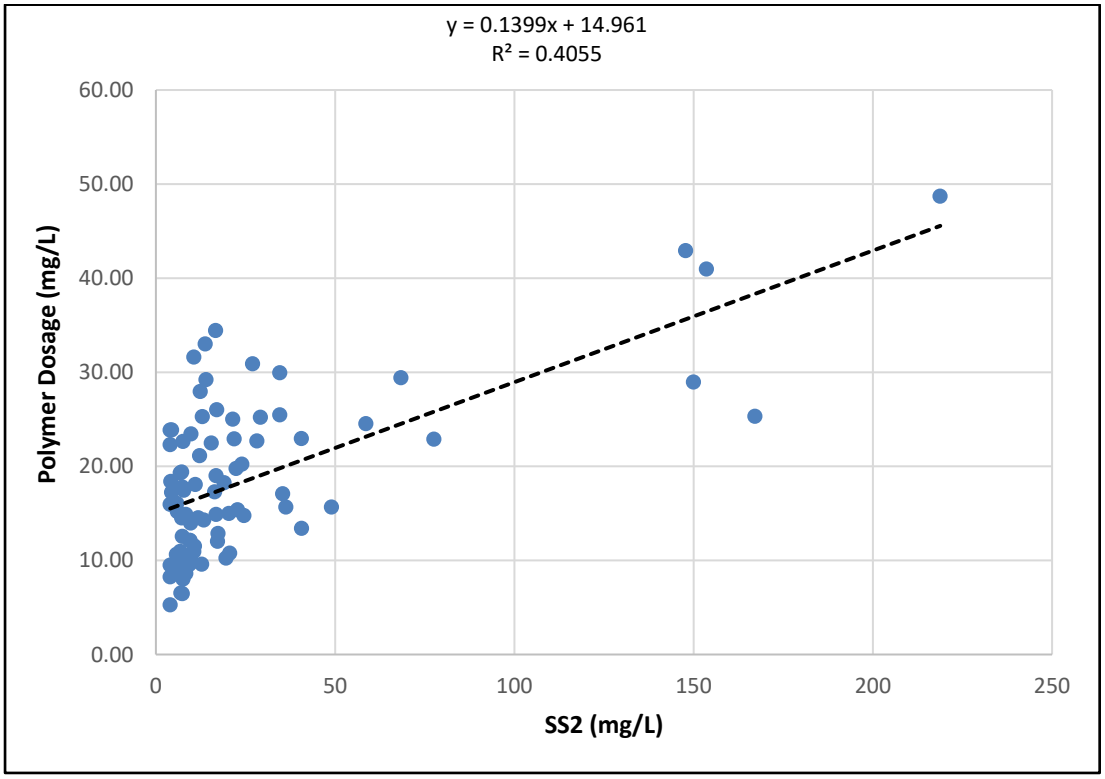


Figure 4.44: Relationship between polymer dosage and suspended solids2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.43 and 4.44, the y values represent the polymer dosage and the x values represent suspended solids. These equations were then used to formulate general equations relating polymer dosage to suspended solids. The general equation for the relationship between polymer dosage and suspended solids1 is shown in Equation 4.0 and that for suspended solids2 is shown in Equation 4.51.

$$D_4 = K_{33}SS_1 + K_{34} \quad \text{Equation 4.50}$$

$$D_4 = K_{35}SS_2 + K_{36} \quad \text{Equation 4.51}$$

In Equation 4.50 and 4.51, D_4 represents polymer dosage and *SS* represents suspended solids. The subscripts on *SS* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on raw water temperature from the EJ Smith Dam and that from Umzinto River. This relationship is shown by the scatter plots in Figure 4.45 and 4.45.

In the equations shown in Figure 4.45 and 4.46, the y values represent the polymer dosage and the x values represent raw water temperature. These equations were then used to formulate general equations relating polymer dosage to temperature. The general equation for the relationship between polymer dosage and temperatuure1 is shown in Equation 4.52 and that for temperature2 is shown in Equation 4.53.

$$D_4 = K_{37}T_1 + K_{38} \quad \text{Equation 4.52}$$

$$D_4 = K_{39}T_2 + K_{40} \quad \text{Equation 4.53}$$

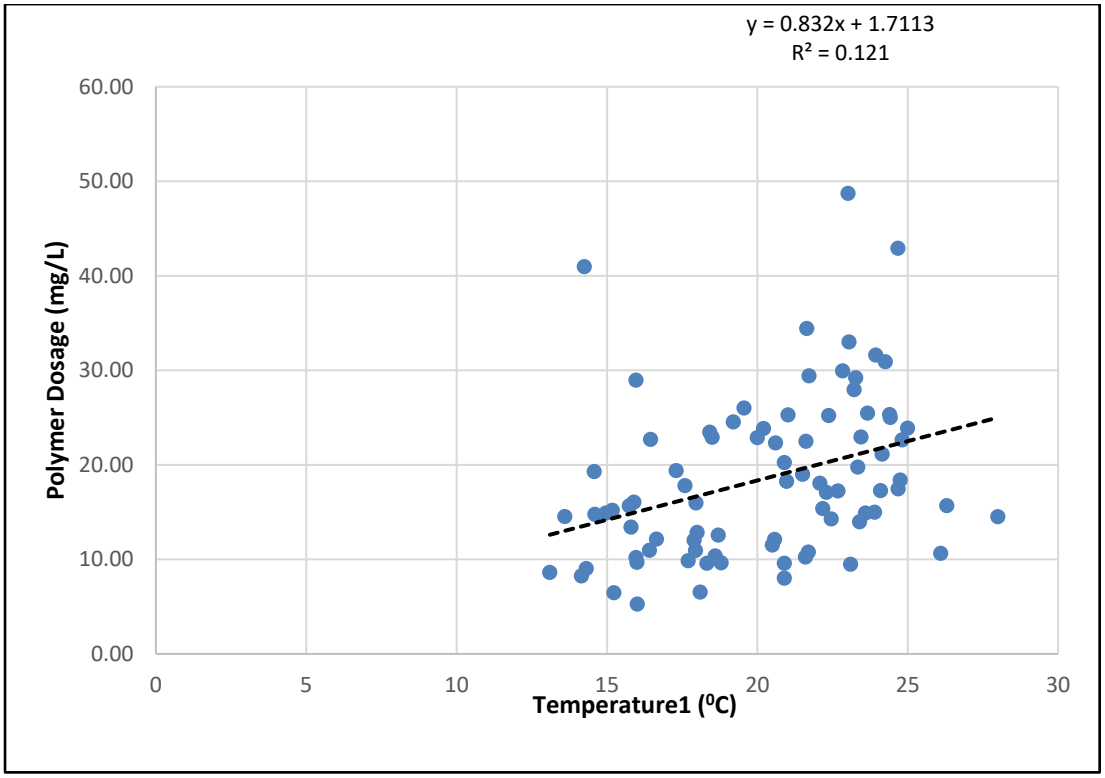


Figure 4.45: Relationship between polymer dosage and temperature1 for the period between July 2006 to June 2013

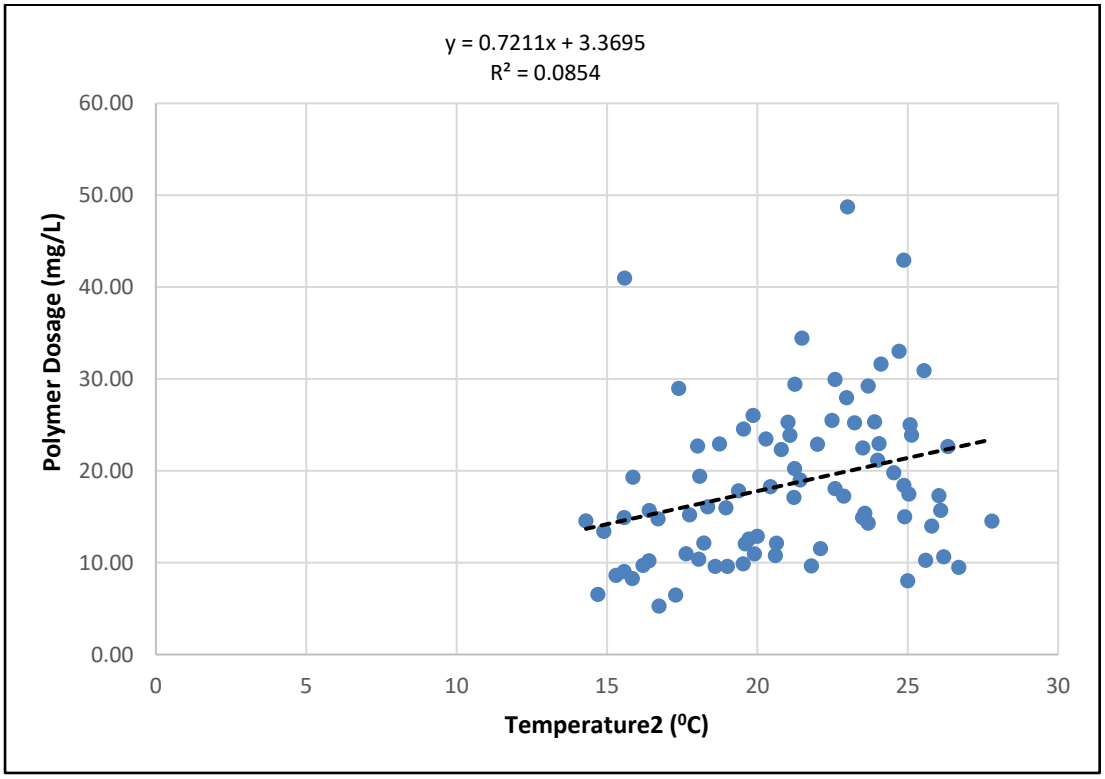


Figure 4.46: Relationship between polymer dosage and temperature2 for the period between July 2006 to June 2013

In Equation 4.52 and 4.53, D_4 represents polymer dosage and T represents raw water temperature. The subscripts on T show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

Equation 4.33 shows that polymer dosage depends on both the total organic carbonyl (TOC) from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.47 and 4.48 respectively.

In the equations shown in Figure 4.47 and 4.48, the y values represent the polymer dosage and the x values represent TOC. These equations were then used to formulate general equations relating polymer dosage to TOC. The general equation for the relationship between polymer dosage and TOC1 is shown in Equation 4.54 and that for TOC2 is shown in Equation 4.55.

$$D_4 = K_{41}TC_1 + K_{42}$$ Equation 4.54

$$D_4 = K_{43}TC_2 + K_{44}$$ Equation 4.55

In Equation 4.54 and 4.55, D_4 represents polymer dosage and TC represents TOC. The subscripts on TC show the point of abstraction as described at the beginning of section 4.5.2. The K values are constants that were determined during the optimisation process using a genetic algorithm.

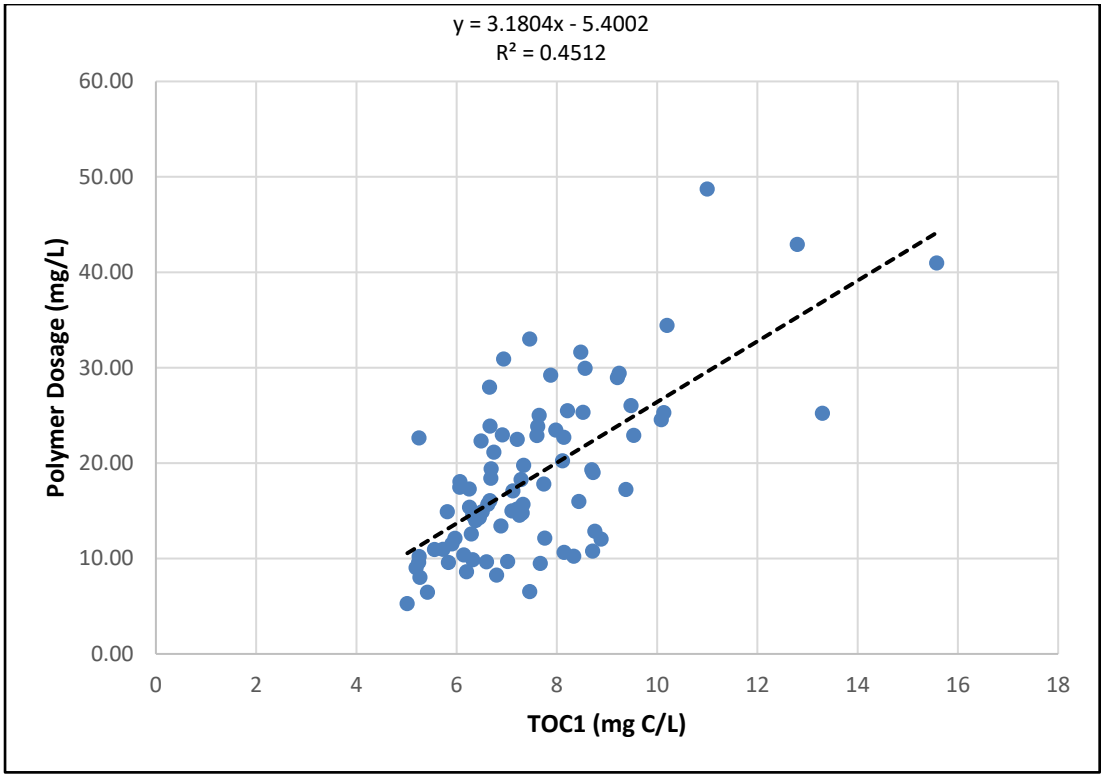


Figure 4.47: Relationship between polymer dosage and total organic carbon1 for the period between July 2006 to June 2013

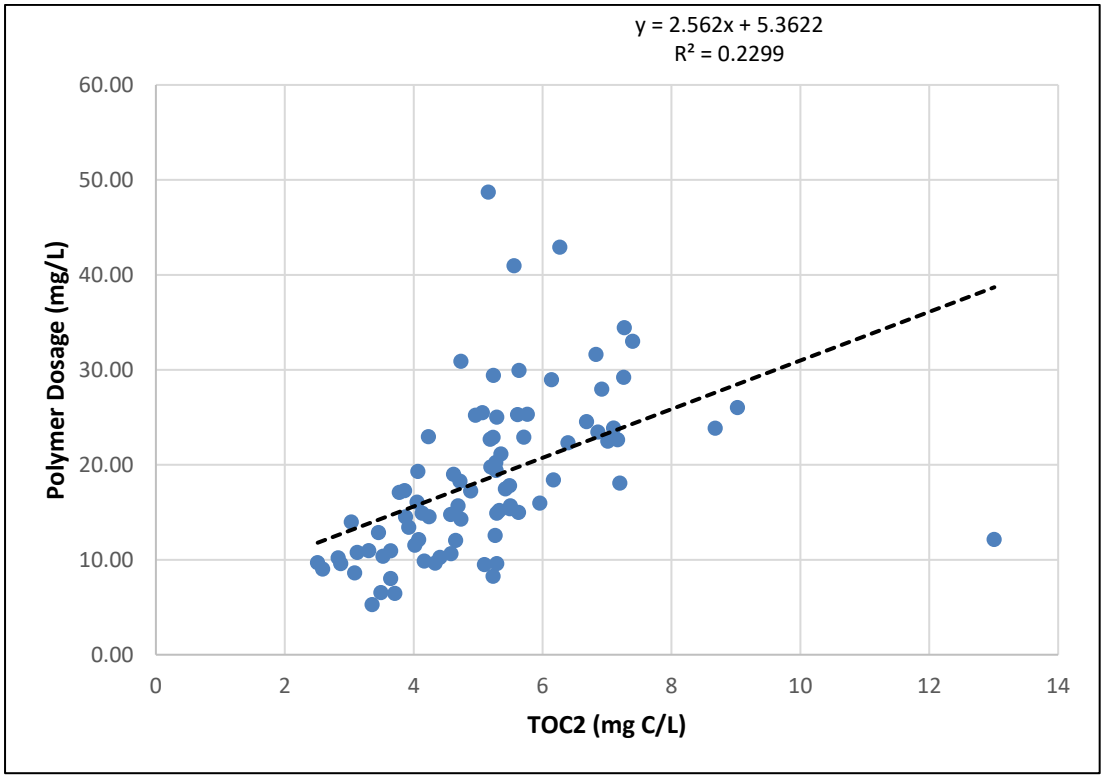


Figure 4.48: Relationship between polymer dosage and total organic carbon2 for the period between July 2006 to June 2013

Equation 4.33 shows that polymer dosage depends on both the turbidity from the EJ Smith Dam and that from Umzinto River. Scatter plots for the relationships are shown in Figure 4.49 and 4.50 respectively.

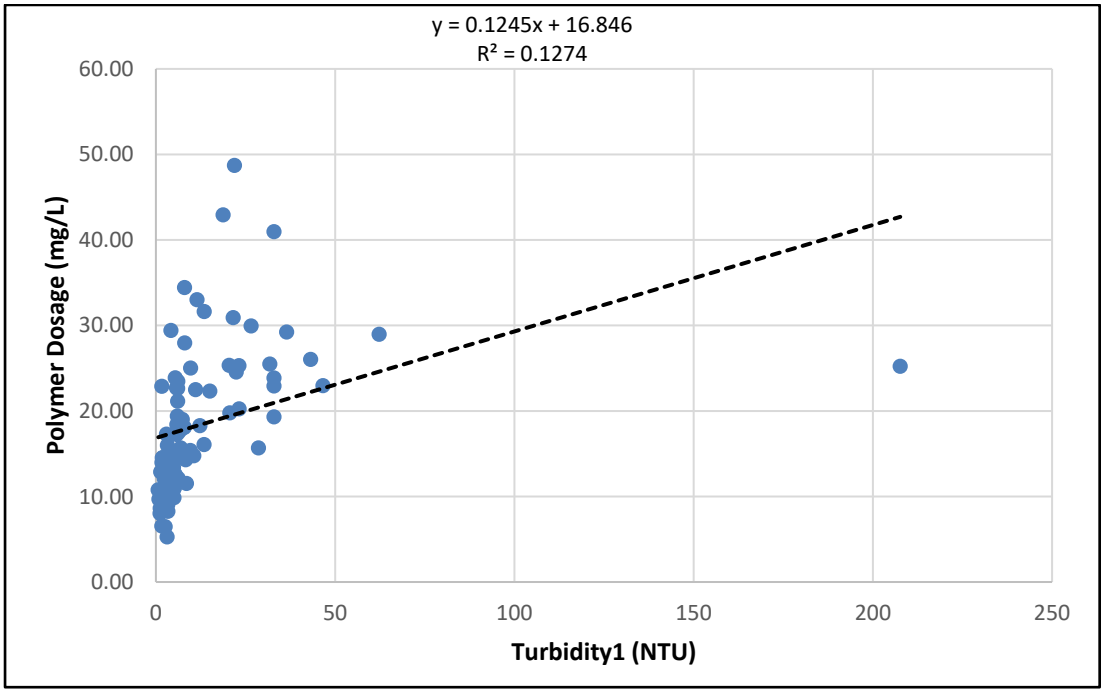


Figure 4.49: Relationship between polymer dosage and turbidity1 for the period between July 2006 to June 2013

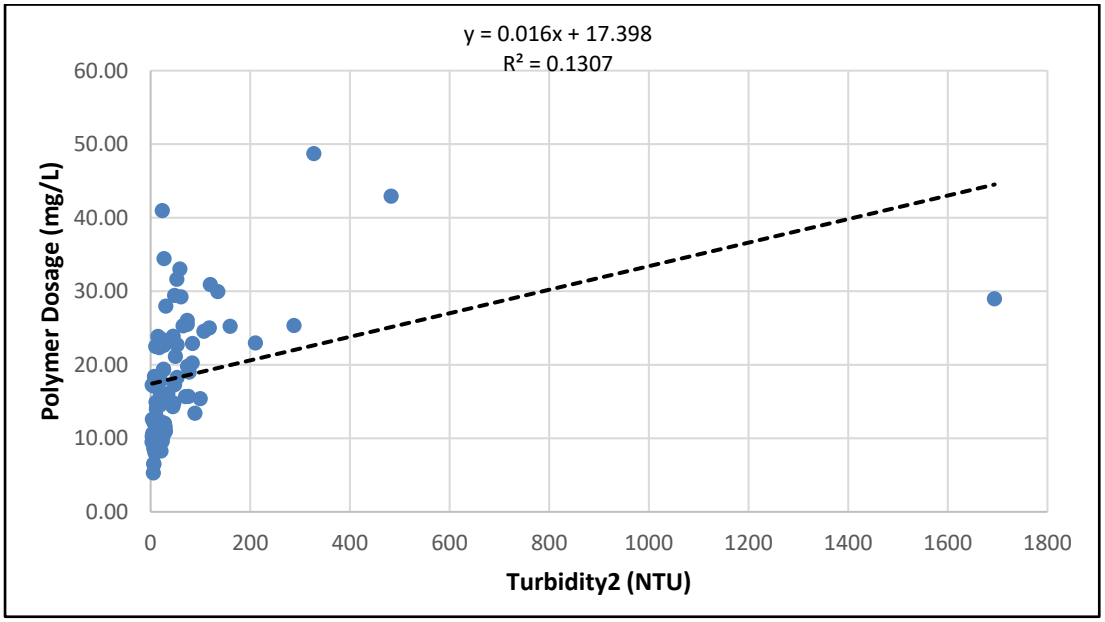


Figure 4.50: Relationship between polymer dosage and turbidity2 for the period between July 2006 to June 2013

In the equations shown in Figure 4.49 and 4.50, the y values represent the polymer dosage and the x values represent turbidity. These equations were then used to formulate general equations relating polymer dosage to turbidity. The general equation for the relationship between polymer dosage and turbidity1 is shown in Equation 4.56 and that for turbidity2 is shown in Equation 4.57.

$$D_4 = K_{45}Tu_1 + K_{46} \quad \text{Equation 4.56}$$

$$D_4 = K_{47}Tu_2 + K_{48} \quad \text{Equation 4.57}$$

In Equation 4.56 and 4.57, D_4 represents polymer dosage and *Tu* represents turbidity. The subscripts on *Tu* show the point of abstraction as described at the beginning of section 4.5.2. The *K* values are constants that were determined during the optimisation process using a genetic algorithm.

To obtain the overall relationship for the polymer dosage with the parameters it correlated to, equations 4.34 to 4.57 are added together. The overall relationship is shown as equation 4.58 below:

$$D_4 = K_1AC_1 + K_3AC_2 + + K_5Alk_1 + K_7Alk_2 + K_9Col_1 + K_{11}C_1 + K_{13}C_2 + K_{15}CO_1 + K_{17}CO_2 + K_{19}Eco_2 + K_{21}Fe_1 + K_{23}Fe_2 + K_{25}H_1 + K_{27}H_2 + K_{29}M_2 + K_{31}pH_1 + K_{33}S_1 + K_{35}S_2 + K_{37}T_1 + K_{39}T_2 + K_{41}TC_1 + K_{43}TC_2 + K_{45}Tu_1 + K_{47}Tu_2 + K_{49} \quad \text{Equation 4.58}$$

4.5.3 Model constraints

The constraints for the model were defined in relation to the parameter levels observed at the two abstraction points. Table 4.4 shows the constraints that were defined for the model.

4.5.4 Genetic algorithm parameter selection

The genetic algorithm toolbox in Matlab prompts the user to define options for running the model. The options would determine the performance of a developed model. In this study the developed models used the options shown in Table 4.5 below.

Table 4.4: Constraints for the parameters in the chemical dosage models

Parameter	min	max
Algal count2 (cells/mL)	0	54594
Algal count1 (cells/mL)	0	478
Alkalinity1 (mg $CaCO_3$/L)	10	126
Alkalinity2 (mg $CaCO_3$/L)	35	100
Coliforms1 (CFU/100mL)	253	22025
Coliforms2 (CFU/100mL)	128	49231
Colour1 (°H)	7	278
Colour2 (°H)	6	87
Conductivity1 (mS/m)	21	44
Conductivity2 (mS/m)	8	48
Escherichia coli1 (CFU/100mL)	5	1071
Escherichia coli2 (CFU/100mL)	2	3027
Fe1 (mg/L)	0	16
Fe2 (mg/L)	0	3
Hardness1 (mg$CaCO_3$/L)	46	195
Hardness2 (mg$CaCO_3$/L)	34	127
Mn1 (mg/L)	0	1
Mn2 (mg/L)	0	0
pH1	7	8
pH2	7	8
SS1 (mg/L)	3	67
SS2 (mg/L)	4	219
Temperature1 (°C)	13	28
Temperature2 (°C)	14	28
TOC1 (mg C/L)	5	16
TOC2 (mg C/L)	3	13
Turbidity 1 (NTU)	1	208
Turbidity2 (NTU)	3	1694
Chlorine (mg/L)	4	9
Poly (mg/L)	5	49
Lime (mg/L)	7	24
Activated Carbon (mg/L)	0	9
Potassium Permanganate (mg/L)	0	3

Table 4.5: Options for running genetic algorithm models

Matlab GA Option	**Value**
Selection Function	Stochastic
Population Type	Double
Elite count	0.05*min(max(10*nvars,40),100)
Crossover Fraction	0.5
Population Size	100
Stall Generation Limit	125
Generations	500
Plot Functions	@gaplotbestf
Output Functions	custOutput

4.6 Methods for Objective 3

Objective 3: Evaluate the Distribution System's Non-Revenue Water

To assesses the NRW for the period July 2013 to June 2014, the approach used was a combination of the top-down and component-based approaches. While Chapter 2 reviewed other methods, these are inappropriate for the Umzinto Distribution System unless they are modified or combined. The top-down approach makes an assumption on the input volume that leads to underestimation of Apparent Losses. The approach assumes that the unauthorised consumption will be at least 0.25% to 1% of the input volume. Farley and Liemberger (2005a) showed that the bottom-up approach can be applied to distribution systems that meet particular hydraulic conditions. This made it difficult to apply it to the Umzinto Distribution System.

The limitation for the third method (component-based analysis) is its well-known uncertainty (Farley and Liemberger 2005a). Component-based analysis is data driven; hence, its accuracy will depend on the accuracy and completeness of the data. For the Umzinto Distribution System, some values were estimated and this affected the accuracy. To mitigate against this, it is usually advocated that this approach not be used on its own.

4.6.1 Evaluation Steps

The first step was to determine the NRW by subtracting the volume of billed water for the Umzinto Distribution System from the volume of the water produced by the UWTP. The next step involved the calculation of the NRW components. The evaluation steps are illustrated in Figure 4.51.

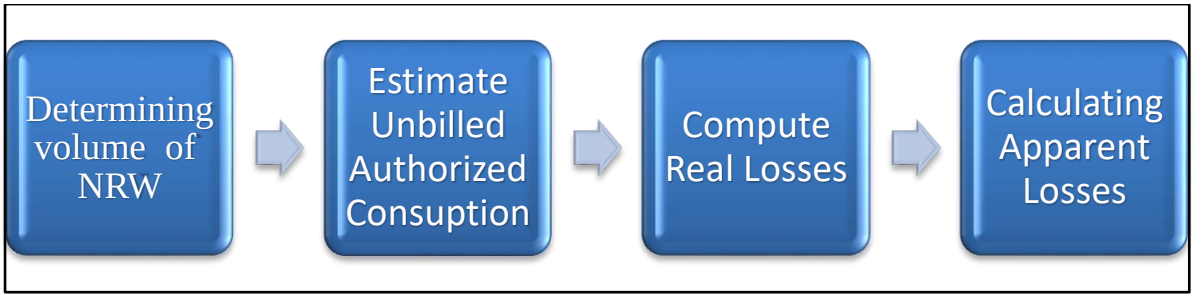


Figure 4.51: Steps involved in the calculation of Non-Revenue Water

4.6.2 Determining the volume of Non-Revenue Water

It should be noted that the billing month does not coincide with the production month. The end date for the billing month was in the middle of a normal calendar month and that for the production was at the end. Meters also had inaccuracies. This meant that, before the NRW was determined, adjustments were to be made to accommodate these factors. The produced water volume was obtained from the SIV data readings and adjustments were made to account for the inaccuracies. The volume of the billed consumption was obtained for both metered and unmetered connections. After correcting the time lag, the adjusted NRW volume was obtained as follows:

$$Adjusted\ NRW = Adjusted\ Produced\ Water - Adjusted\ Billed\ Water.$$

Equation 4.59

Figure 4.52 shows the general framework for determining NRW. The AWWA (2009) recommended a period of at least one year for assessing NRW levels. This would make it possible to take seasonal variations into consideration. In this study the data used were for the year ending December 2013.

The time frame of this study was not adequate for accurate analysis of input meters. To cater for this, input meter inaccuracy was assumed to be at 7% under-registration. This approach was also used by Arregui *et al.* (2006). The estimation was justified by the absence of a preventive maintenance program for the Umzinto Water Distribution System. The meters are examined or maintained on a corrective basis, that is, when they have stopped running or when the readings registered are lower than expected. Furthermore, some of the meters were not installed according to their design specifications.

Billed metered and unmetered consumption

Data for the billed consumption was obtained from the billing records of Ugu District Municipality. The records consisted of metered consumption and the estimated unmetered consumption data. Customers with estimated unmetered consumption were charged a flat rate.

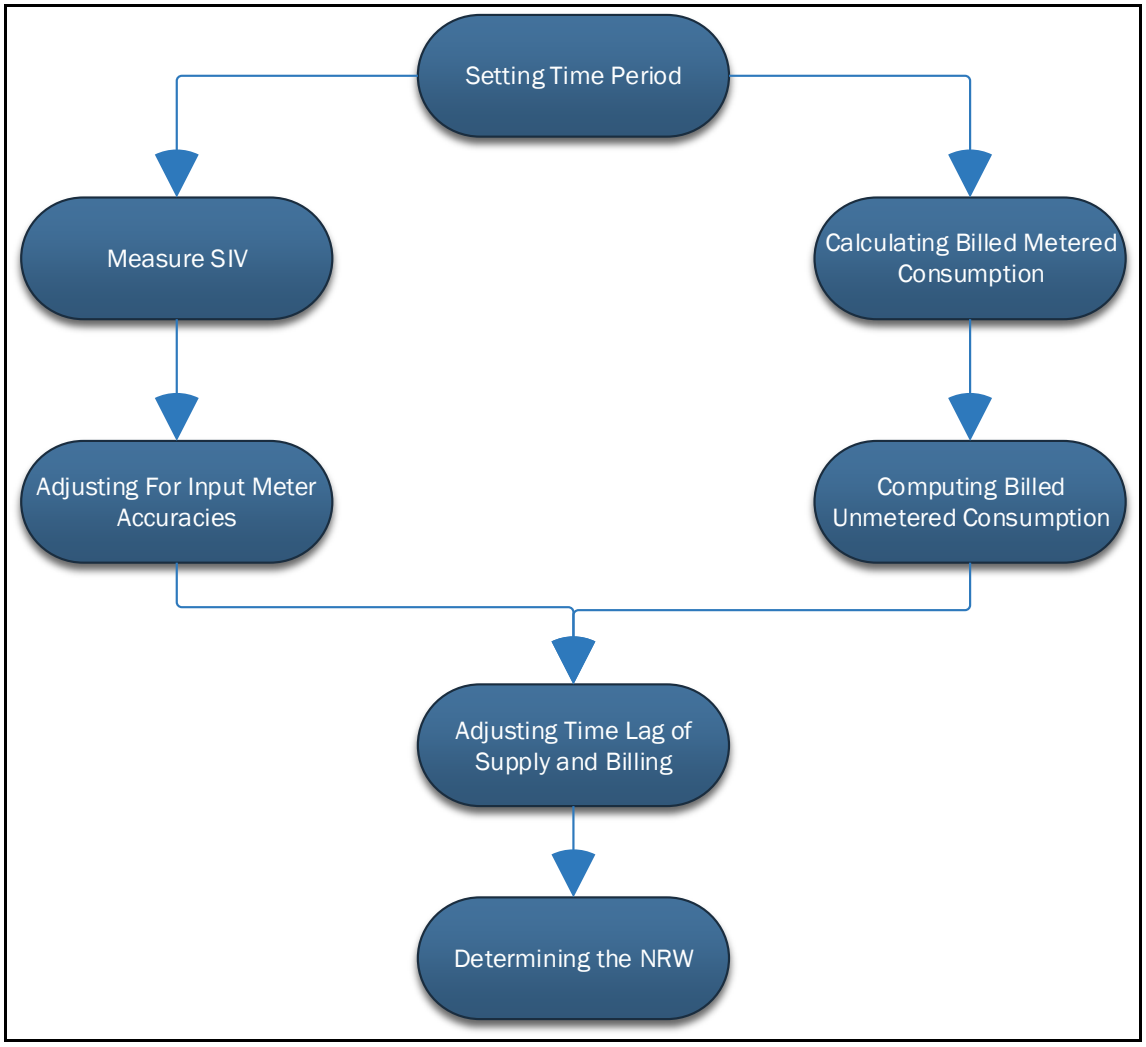


Figure 4.52: Determining Non-Revenue Water

Lag time adjustment

The water utility's billing dates do not coincide with the date on which the meters are read. Indeed, the meters are read and billing only takes place a month later. To resolve this problem, the production data used were offset by a month so that it coincided with billing. For example, when the production data used were for July 2013 to June 2014, the billing data would be for August 2013 to July 2014.

NRW volume was then determined by calculating the difference between the adjusted volume of produced water and the adjusted volume of billed water.

4.6.3 Estimating unbilled authorised consumption

This authorised component consisted of metered and unmetered water. Unbilled metered consumption was mainly due to use within the utility's buildings. The unmetered component included consumers that were billed at a flat rate, free basic water standpipes and other measures aimed at assisting communities. Actual consumption for the consumers charged at a flat rate was far higher than the flat rate amount. In some cases, consumers that were billed took advantage of this arrangement.

Stand pipes are installed in low income areas. Users do not pay for this water and hence it is a major contributor to unbilled metered consumption. For the study, while most stand pipes were unmetered, the volume of water was often estimated and included in the water balance. Major losses occurred through stand pipes due to vandalism and lack of maintenance (Ugu District Municipality 2011).

Unbilled unmetered consumption also included the water that was used to clean pipes during maintenance, to fight fires and to supply some special institutions. In some cases, water is supplied to the community using tankers, and this consumption was estimated by the number of tankers per year for each administrative zone. The final figure for unbilled authorised consumption was estimated using data from the utility's internal reports for the period June 2013 to July 2014.

4.6.4 Calculating Real and Apparent Losses

Once the Apparent Losses were defined and the unbilled authorised consumption was estimated, Real Losses were calculated using the following formula:

$$Real\ Losses = NRW - Apparent\ Losses - Unbilled\ authorised\ consumption. \quad \text{Equation 4.60}$$

4.7 Methods for Objective 4

Objective 4: Develop a model for water usage analysis

The key in achieving this objective was the establishment or development of software that improves the input and reporting of data for the Umzinto Distribution System. In achieving this, the study focused on the following three important aspects of software development:

- Software interface development
- Software platforms and languages
- Software functionality.

4.7.1 Software interface development

During interface design, the guiding principle was that the software should be easy to learn. This was achieved by eliminating the need for the user to know the formulae involved in the calculation of NRW components. In order to minimise software learning the application was designed to include components for data input and components for reporting. These were built to function as they would in commonly used programs such as spread sheets and word processors. In this way the user does not require any special skills when using the water usage analysis model. Figure 4.53 shows the flow diagram of the water usage analysis model and the components that the interface supports.

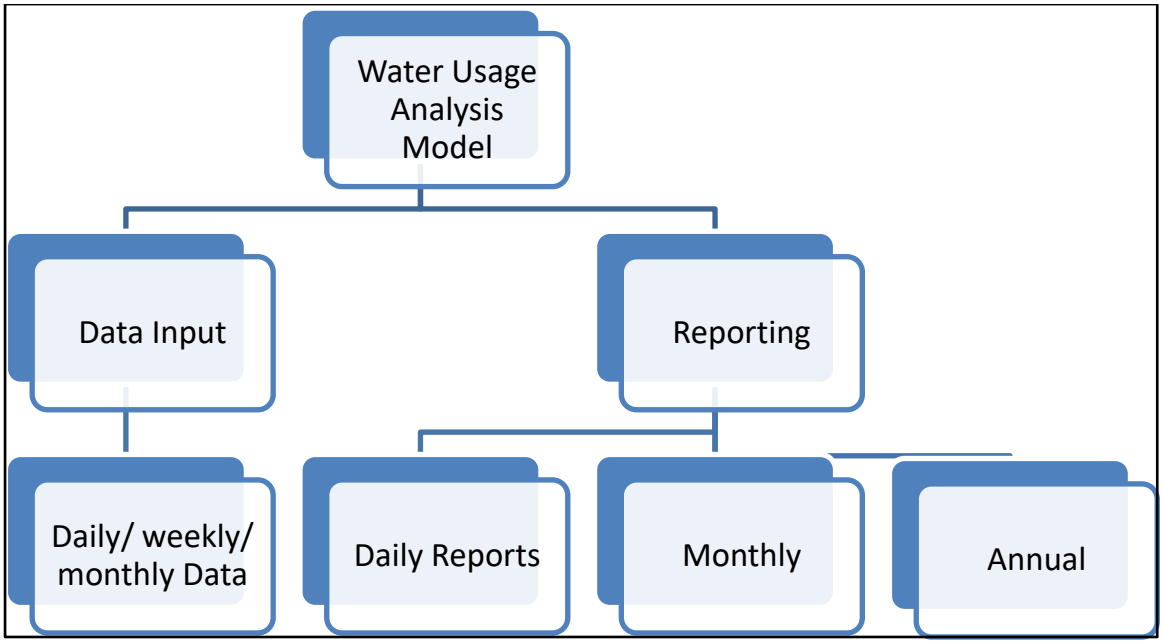


Figure 4.53: Flow diagram for the Water Usage Analysis Model's capabilities

Data Input

The data input was designed to enable the user to enter daily data for the different components of NRW. The software also allowed for the import of spread sheets containing weekly and monthly data. The spread sheets would be created in Microsoft Excel, which is a very common program. The user would be able to generate templates to assist in the creation of weekly and monthly data spread sheets.

Reporting

The reporting function was designed to enable the user to generate reports by simply clicking buttons. It offered graphical representation of the activities in the distribution system. The application produced graphs for the trends in the SIV data, reservoir over flows, leakages, billed consumption and unbilled consumption. Of great importance was the software's ability to compute the water balance at any given time using the rolling year method.

4.7.2 Software platforms and languages

The Microsoft Visual Studio development environment was used to develop the water usage analysis model. This is because it is easy to design interfaces in this environment and

the development program offers an interactive algorithm for debugging. In addition to the interface design, the software was linked to a relational database developed in Microsoft Server 2012. For speed the program was developed to be deployed onto Windows operating systems that include Windows 7 and 8.

4.8 Model Calibration

Model calibration involves estimation of the various constants and parameters making up a model's general structure. Model coefficients and constants are estimated by simultaneously solving equations for formulated from inputting the observed values of both the dependent and independent variables into the model's general structure (Bates *et al.* 2014). In this research, model calibration was important in achieving objective 2, which was to design a genetic algorithm for the potable water treatment process control. As noted previously, the chemical dosage data and the water quality parameter data for the period 2011 to 2015 were collected.

In the model development, the first step was to determine the general relationship between chemical dosage and each parameter. This was done using the curve fitting technique illustrated in section 4.5.2. This process helped in coming up with the general structure of the model. The model structure was then solved using the genetic algorithm toolbox in Matlab. In determining the solution, the collected data were used to develop column vectors in the model equation. The objective function of the algorithm was then set in a way that ensured that its solution would have the greatest probability or maximum likelihood of being accurate. The objective function was defined in a way that allowed constants for the model equation to be obtained whilst ensuring that the difference between the predicted value and actual value was minimal.

4.9 Validity and Reliability

In simple terms, the reliability of a research instrument can be described as its repeatability. According to Peat (2001), reliability usually describes the consistency of a research method, while validity describes the suitability of the instrument used. The research instrument should measure what it is designed to measure and should perform as it is

designed to perform. Cook (2015) defined this as the best available approximation to the truthfulness of a given proposition. It is the answer to the simple question, 'were we right?'

4.9.1 Validity

This research study considered two forms of validity, content validity and criterion-related validity.

- Content validity measures the extent to which the research instrument used is appropriate to answer the research questions. Content validity mainly depends on judgement. To ensure content validity for this research study, the research topic was fully defined and the objectives of the study were laid out. In addition to consulting experts in the fields of operation and management of water treatment plants the researcher conducted an extensive background study on the area. The combination of the consultations and the literature review on the use of genetic algorithms in modelling helped to determine the usefulness of the research and its instrument.

- Criterion-related validity usually describes the relevance of a certain measure. Pennington (2003) described it as a measure of how well a variable or set of variables can predict an outcome. Trochim and Donnelly (2005) observed that, in using criterion validity, predictions are made based on the research construct or theory. To assess the quality of a criterion measure, relevance, freedom from bias, reliability and availability should be considered. This can be achieved by using statistical analysis such as correlation. In this study, the model results were statistically analysed and compared to the actual observed data as presented in Chapter 5.

4.9.2 Reliability

To ascertain reliability, the researcher had to be aware of the sampling and testing methods at the UWTP. In conjunction with validity, reliability can help to assess the integrity of the historical data collected. In the event of errors in the datasets it is easier for a researcher to make decisions on missing values, censored values and the presence of outliers. These

errors are common in water quality datasets but the method for treatment greatly depends on the reliability of the method employed to do the sampling and tests.

4.10 Elimination of Bias

Bias cannot be completely eliminated in research. In this study, the data used were recorded by different people, drastically minimizing the effect of one person doing the recording. Data for both the day and night shifts were used.

4.11 Summary

This research study sought to synthesize a model for optimising a water treatment plant and water usage analysis. This chapter discussed the quantitative approach adopted and the manner in which data were collected for both the water treatment plant and the distribution system. The methods employed to achieve each of the study's four objectives were clearly explained. Since the study required the development of a model, historical data were used. The chapter also discussed the steps which were taken to overcome issues relating to reliability and validity, as well as the limitations of the study and how these were overcome to successfully complete the study.

CHAPTER 5 – RESULTS AND DISCUSSION

5.1 Introduction

This chapter presents and discusses the results of the study in accordance with the four specific objectives outlined in Chapter 1. The first section presents the results obtained during the data pre-processing stage. It also describes a Microsoft Excel Add-in that was developed to ease the pre-processing of water treatment data.

5.2 Data Pre-processing

The purpose of data pre-processing was to produce data sets that could be used to develop a genetic algorithm model for water treatment dosage prediction. The major tasks carried out in data pre-processing were converting the data into analysable format and dealing with missing values, censored data and outliers.

MSSQL was used to convert the data into an analysable format. The procedure for this process was described in Chapter 4 and the results are shown in Figure 5.1.

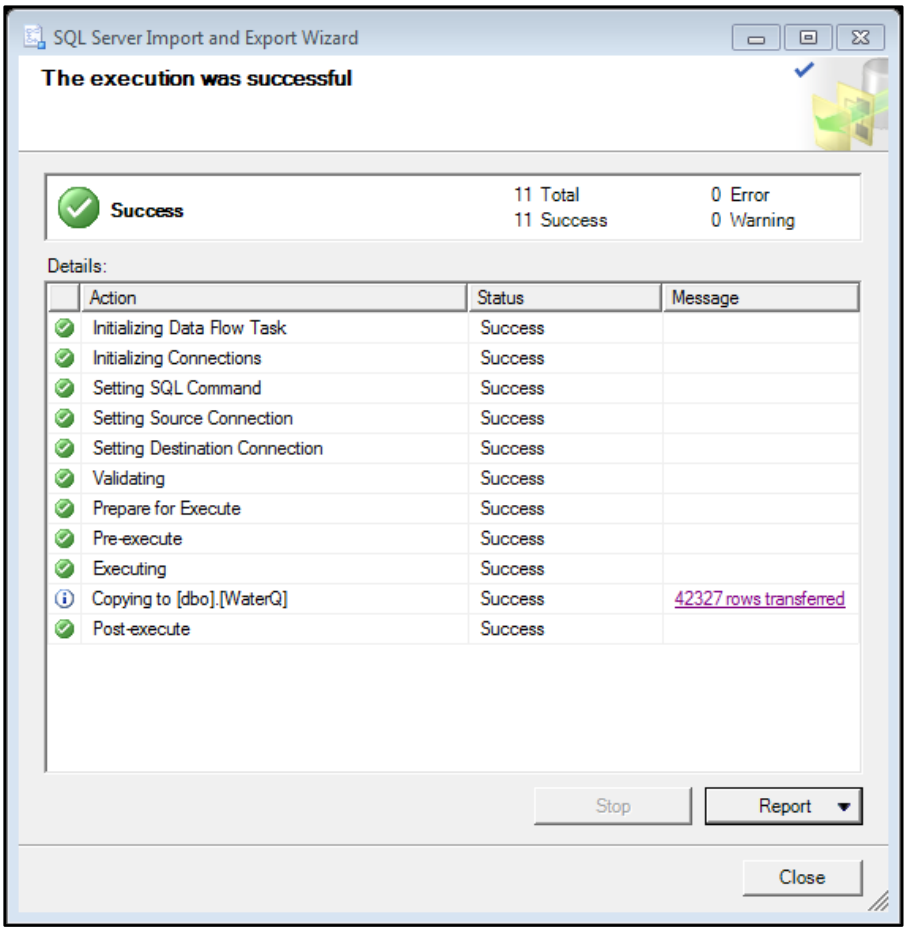
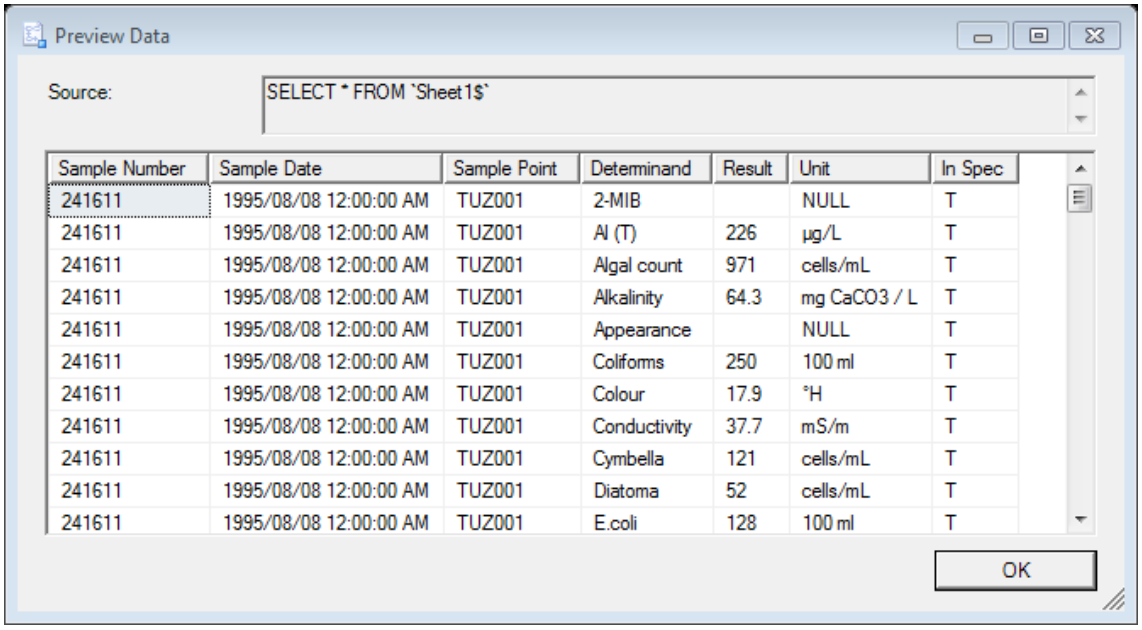


Figure 5.1: Importing data from Excel into Microsoft SQL 2012

As shown in Figure 5.1 the data import was executed in 11 stages. The execution was successful but during the process it was essential to make sure that the data source had all the information required in the destination (the created database table). In the import process 42327 rows were transferred from Excel to a database table in the format shown in Figure 5.2.



Sample Number	Sample Date	Sample Point	Determinand	Result	Unit	In Spec
241611	1995/08/08 12:00:00 AM	TUZ001	2-MIB		NULL	T
241611	1995/08/08 12:00:00 AM	TUZ001	Al (T)	226	µg/L	T
241611	1995/08/08 12:00:00 AM	TUZ001	Algal count	971	cells/mL	T
241611	1995/08/08 12:00:00 AM	TUZ001	Alkalinity	64.3	mg CaCO3 / L	T
241611	1995/08/08 12:00:00 AM	TUZ001	Appearance		NULL	T
241611	1995/08/08 12:00:00 AM	TUZ001	Coliforms	250	100 ml	T
241611	1995/08/08 12:00:00 AM	TUZ001	Colour	17.9	°H	T
241611	1995/08/08 12:00:00 AM	TUZ001	Conductivity	37.7	mS/m	T
241611	1995/08/08 12:00:00 AM	TUZ001	Cymbella	121	cells/mL	T
241611	1995/08/08 12:00:00 AM	TUZ001	Diatoma	52	cells/mL	T
241611	1995/08/08 12:00:00 AM	TUZ001	E.coli	128	100 ml	T

Figure 5.2: Water quality table created in Microsoft SQL 2012

The imported data contained censored and missing values. To handle these, a VBA program was developed to fill in the missing values and automatically correct the censored data into an analysable form. Data cleaning algorithms were applied which can fill in missing values, identify outliers and smooth out noisy data, and correct inconsistent data. Careful integration of the data was also done to reduce/avoid redundancies and inconsistencies and improve data quality. The resultant VBA program is shown in Appendix 3.

5.3 Results for Objective 1

The first objective was to:

Identify and quantify key raw water quality parameters affecting treatment at the Umzinto Water Treatment Plant.

For the period between July 2006 to June 2013, 116 raw water quality parameters were obtained. These were reduced to 14 operational water quality parameters by applying the data pre-processing techniques articulated in Chapter 4. Outliers were removed by plotting and examining the statistics using box and whisker plots described in Section 4.4.2. An example of a box and whisker plot is shown in Figure 5.3 for the algal count of the EJ Smith

Dam's data (EJS) and Umzinto River's data (UMZ). The plots illustrate the extent of outlier density in each variable as indicated by the points extending beyond the whiskers. In addition, they show the range of each variable and, consequently, the efficiency of the plant treatment.

The line graphs or trend lines provided information on the temporal variation of the parameter values. For objective 1, the results are divided into three sections: the temporal analysis of water quality parameters; temporal analysis of the chemical dosages; and correlation of water quality parameters to chemical dosages.

5.3.1 Temporal analysis of water quality parameters

Algal count

Figure 5.3 shows the box and whisker plot for the raw water algal count at the two abstraction points of the UWTP while Figure 5.4 shows the temporal analysis and the descriptive statistics are shown in Table 5.1.

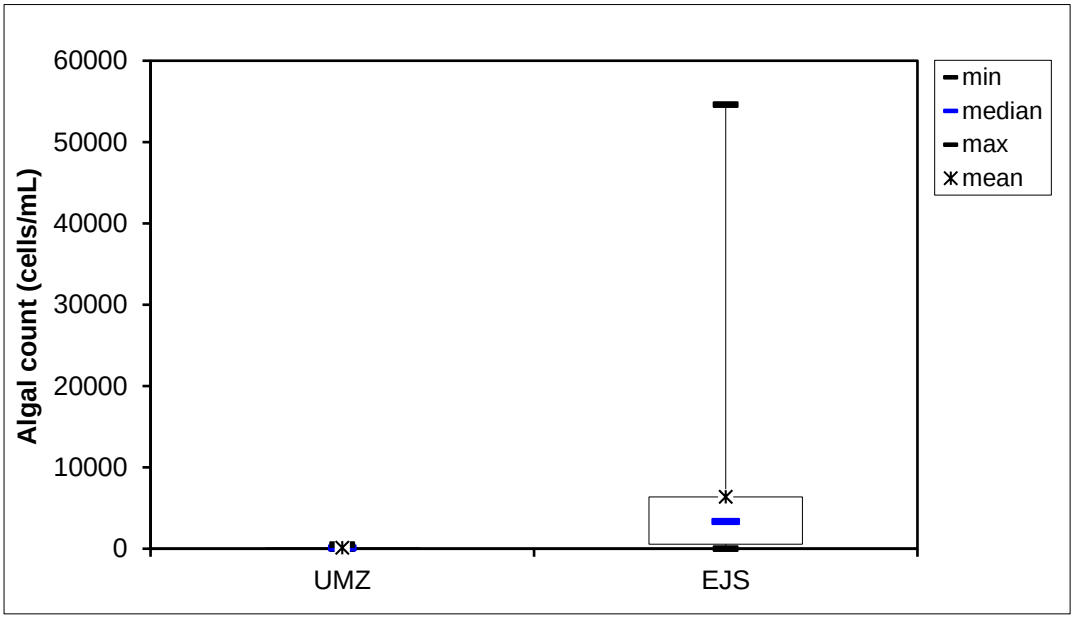


Figure 5.3: Raw water algal count

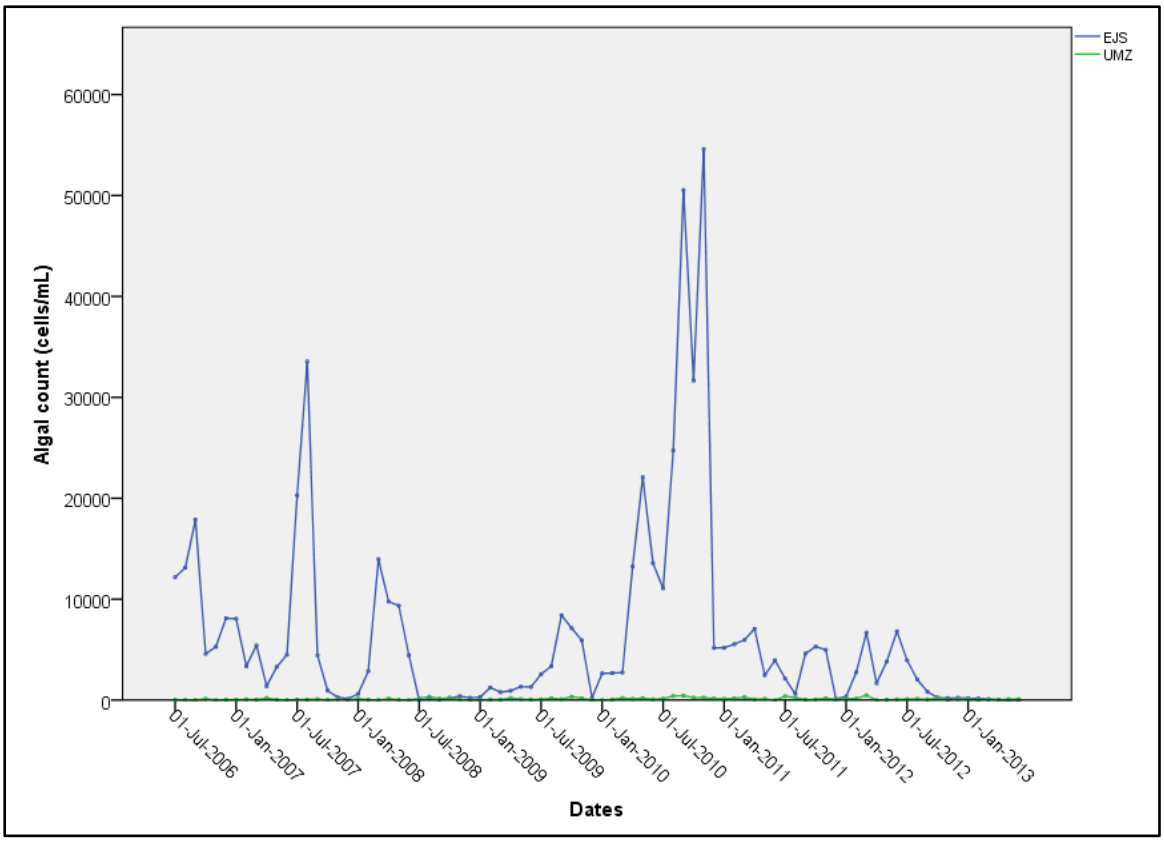


Figure 5.4: Temporal analysis of raw water algal count

Table 5.1: Descriptive statistics for algal count

Abstraction points	Mean (cells/mL)	Median (cells/mL)	Std. Deviation (cells/mL)	Range (cells/mL)	Minimum (cells/mL)	Maximum (cells/mL)
UMZ	115	82	108	478	0	478.0
EJS	6369	3325	9928	54594	0	54594

The box and whisker plot shows that the algal count levels from the EJ Smith Dam's raw water are very high, ranging from 0 to 54594 cells/mL. In contrast , the Umzinto River raw water algal count ranges from 0 to 478 cells/mL. The difference between the algal count levels is an indication of activities close to the abstraction points (Rangeti *et al.* 2015).

The first concern with high levels of algae is its effect on the aesthetic properties of the water. It usually results in unpleasant colour and odour, that is, the water will be unacceptable aesthetically. Dealing with this implies an increase in treatment costs.

Table 5.1 shows that the mean for the algal count for the water from the EJ Smith Dam between 2006 and 2013 is 6369 cells/mL. This results in the clogging of the filters, which

in turn increases the operational costs associated with backwashing and filtration time. The literature notes that high levels of algae often result in increased use of powdered activated carbon, a chemical used to treat the toxins produced by some algae species (Dixon *et al.* 2011; Sevilla *et al.* 2012). This directly affects treatment costs.

The high monthly mean value of algal count observed in the raw water from the EJ Smith Dam suggests a need for more coagulants and disinfectants. When chlorine is used, it reacts with the algae to produce dis-infection by-products, which poses a risk to cancer to the population that drinks the water (Chowdhury, Rodriguez and Sadiq 2011; Richardson and Postigo 2012; Henry 2013). Another concern is that this reaction tends to increase the amount of chlorine used.

Despite the difference in levels, both trends show the same seasonal pattern, with an increase in algae during the rainy season and low values in winter when there is little or no rainfall.

Alkalinity

Figure 5.5 shows the box and whisker plot for raw water alkalinity at the two abstraction points of the UWTP. Figure 5.6 shows the temporal analysis while the descriptive statistics are shown in Table 5.2.

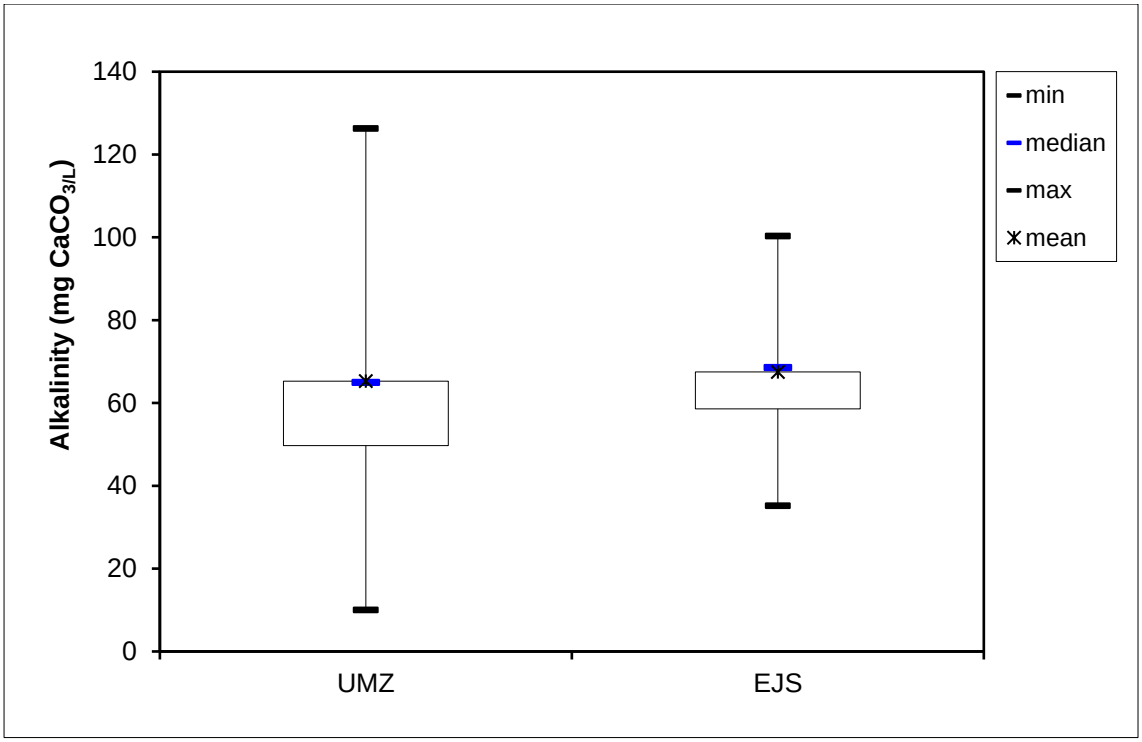


Figure 5.5: Raw water alkalinity

The box and whisker plot in Figure 5.5 shows that the mean and median alkalinity levels for the two abstraction points are almost similar. It also shows that the Umzinto River's alkalinity had a wider range than that of the EJ Smith Dam. Table 5.2 shows that the Umzinto River's range is 116.3 mg$CaCO_3$/L while that for the EJ Smith Dam is 65.1 mg$CaCO_3$/L. The differences in the range can be attributed to the difference in flow between a river and a dam (Wesanekar *et al.* 2014; Wing 2014).

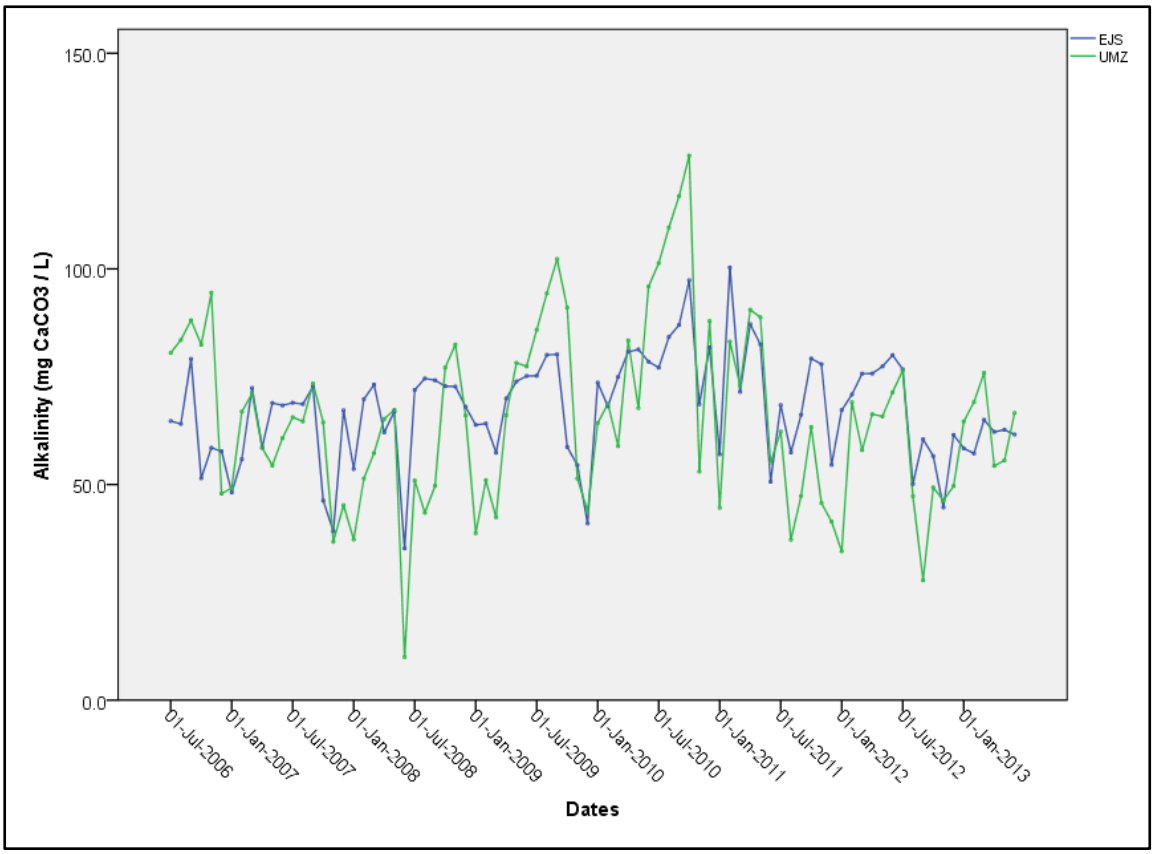


Figure 5.6: Temporal analysis of raw water alkalinity

Table 5.2: Descriptive statistics of raw water alkalinity for the period between July 2006 to June 2013

Abstraction Point	Mean (mg $CaCO_3$/L)	Median (mg $CaCO_3$/L)	Std. Deviation (mg $CaCO_3$/L)	Range (mg $CaCO_3$/L)	Minimum (mg $CaCO_3$/L)	Maximum (mg $CaCO_3$/L)
UMZ	65.3	64.9	20.6	116.3	10.0	126.3
EJS	67.5	68.5	12.2	65.1	35.2	100.3

The mean and median values of alkalinity shown in Table 5.2 demonstrate that the two sources of abstraction were highly alkaline. This is similar to other South African rivers that have been studied (Gose *et al.* 2013; Dabrowski *et al.* 2014).This means that the raw water had significant capacity to buffer the pH and would result in low lime dosages. The line graphs in Figure 5.6 show that alkalinity had an irregular variation at the two abstraction points. For better control of the treatment processes the pH needs to be kept within a specific range depending on the treatment stage. This suggests the need to consider the control of this parameter during treatment.

Colour

Figure 5.7 shows the box and whisker plot for the colour of the raw water at the two abstraction points of the UWTP. Figure 5.8 shows the temporal analysis and the descriptive statistics are shown in Table 5.3.

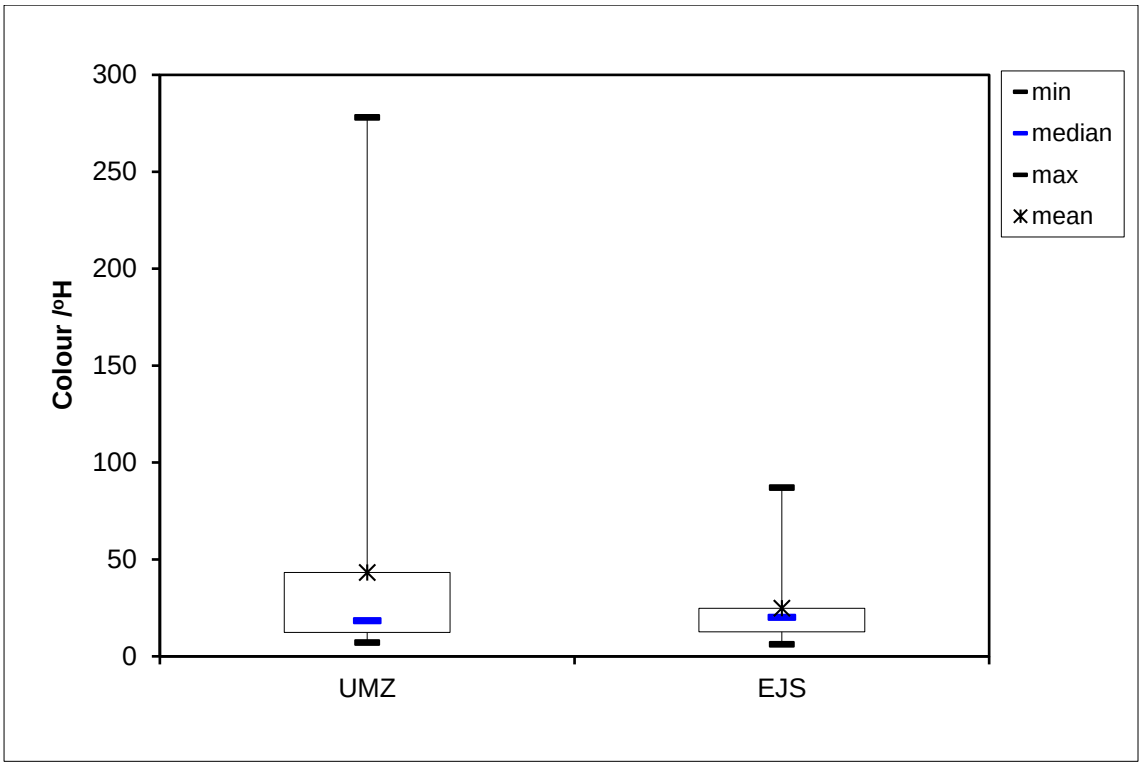


Figure 5.7: Raw water colour

As expected from literature, the raw water from the Umzinto River shows a wider range of colour than that from the EJ Smith Dam (Ugu District Ugu District Municipality 2011). The majority of the observations for both abstraction points are below 50 ^{0}H. This means that the main concern in treating the raw water's colour are the sudden peaks observed in the Umzinto River raw water. These are shown in Figure 5.8.

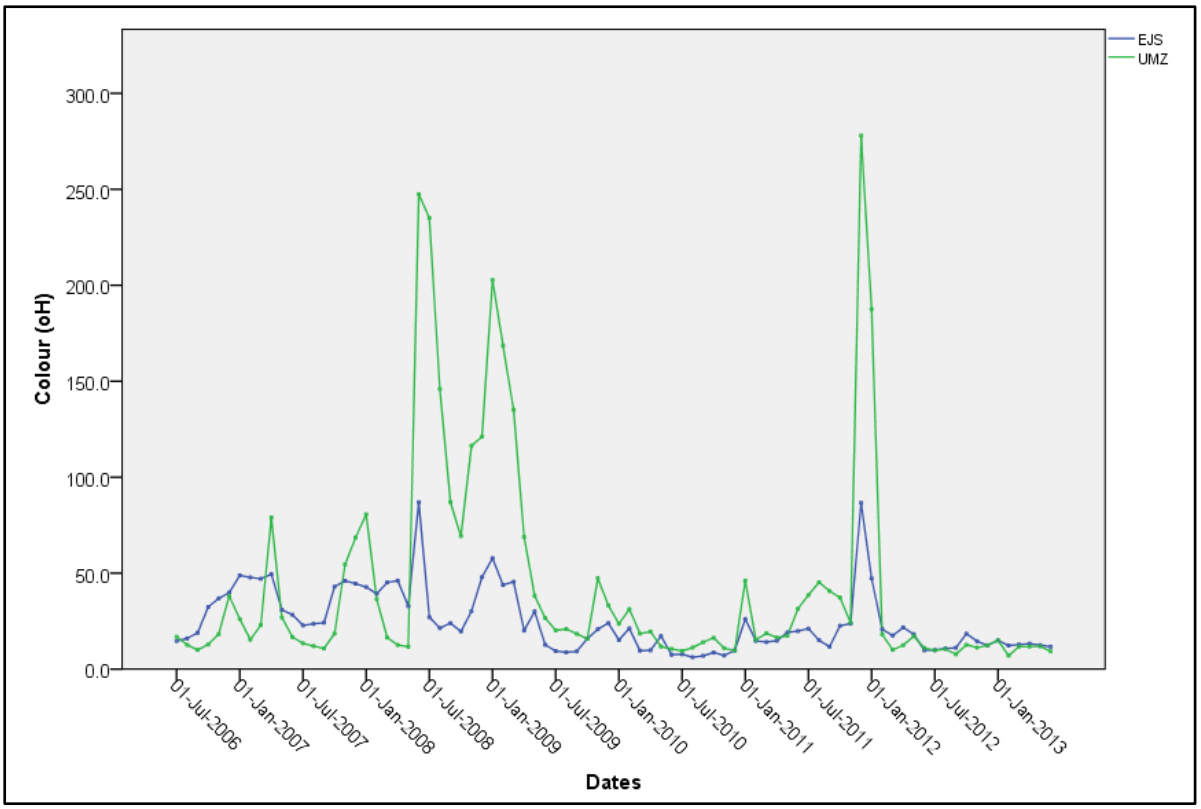


Figure 5.8: Temporal analysis for the raw water colour

Table 5.3: Descriptive statistics for the raw water colour for the period between July 2006 to June 2013

Abstraction Point	Mean (°H)	Median (°H)	Std. Deviation (°H)	Range (°H)	Minimum (°H)	Maximum (°H)
UMZ	43.3	18.3	57.9	271.0	7.1	278.0
EJS	24.8	20.0	16.5	80.8	6.2	87.0

Table 5.3 shows that the EJ Smith Dam's raw water has a lower mean value for colour and its small range suggests that the colour is much more stable in the dam than the river. The high mean and range in the Umzinto River suggest that it is the major contributor to the cost of colour correction chemicals. It should, however, be noted that while colour in itself has no adverse effect on human health (DWAF 1996), due to stipulations by the South African Bureau of Standards (SABS, 2011), the colour of portable water should be less than 15^0H in order to be classified as Class I.

Conductivity

Figure 5.9 shows the box and whisker plot for the raw water conductivity at the two abstraction points of the UWTP. Figure 5.10 shows the temporal analysis and the descriptive statistics are shown in Table 5.4.

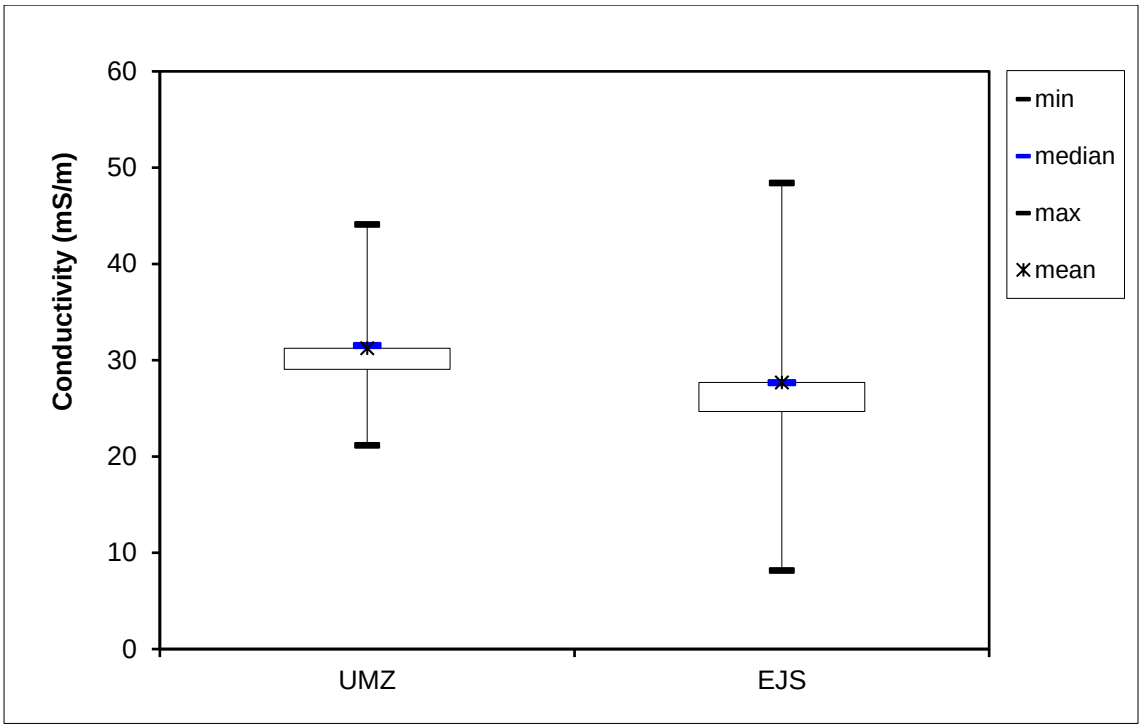


Figure 5.9: Raw water conductivity

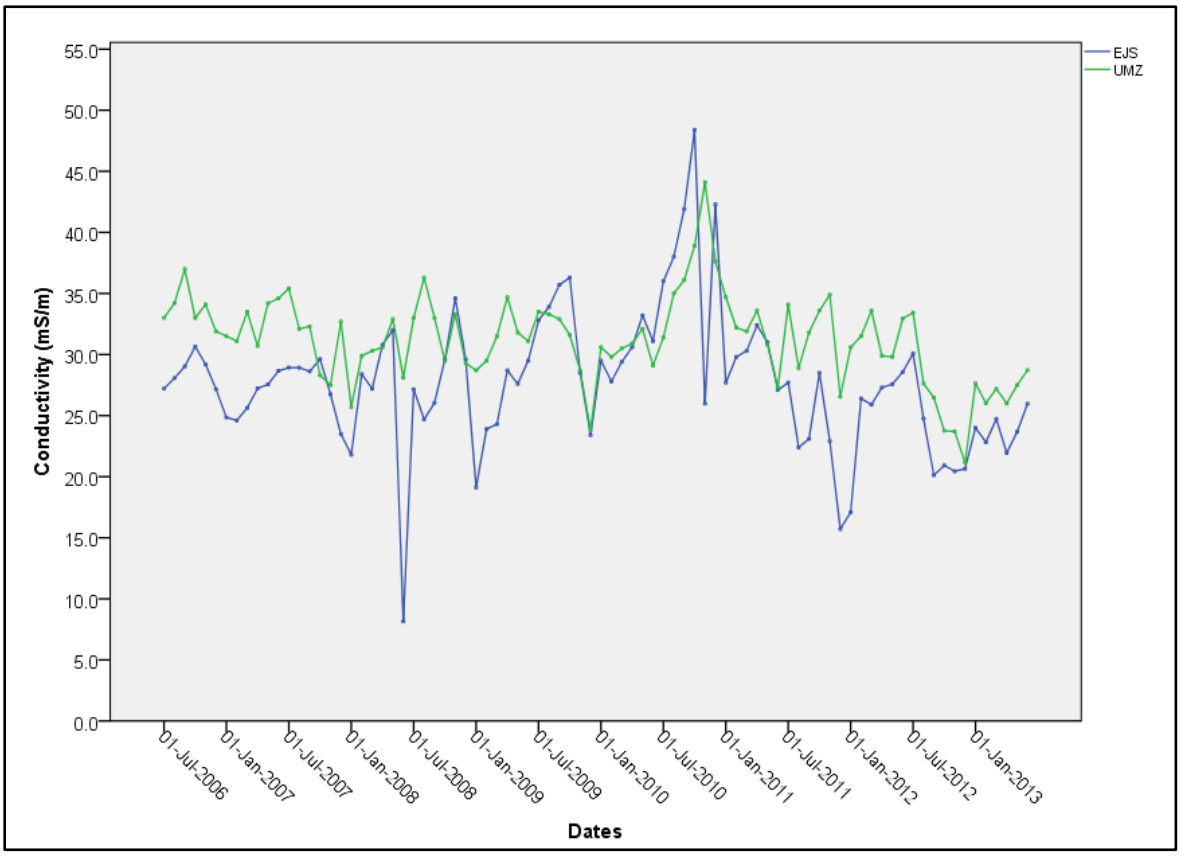


Figure 5.10: Temporal analysis of raw water conductivity

Table 5.4: Descriptive statistics of raw water conductivity for the period between July 2006 to June 2013

Abstraction Points	Mean (mS/m)	Median (mS/m)	Std. Deviation (mS/m)	Range (mS/m)	Minimum (mS/m)	Maximum (mS/m)
UMZ	31.2	31.5	3.6	23.0	21.1	44.1
EJS	27.7	27.7	5.7	40.3	8.2	48.4

The box and whisker plot in Figure 5.9 and Table 5.4 show that the EJ Smith conductivity had a much wider range than the Umzinto River. This is due to the very low conductivity value in June 2008 and a peak in November 2008. Overall, the electrical conductivity for the EJ Smith Dam's raw water was lower than that of Umzinto River from 2006 to 2012.

From the temporal analysis in Figure 5.10, the conductivity showed irregular variation. Despite this, the maximum for the values recorded from the two abstraction points (48.1 mS/m for the EJ Smith Dam and 44.1mS/m for the Umzinto River) are way below the South African guidelines for raw water (i.e., <170 mS/m) (DWAF 1996). This suggests that conductivity should not greatly influence the treatment costs at the UWTP.

Escherichia coli

Figure 5.11 shows the box and whisker plot for *Escherichia coli* at the two abstraction points of the UWTP. Figure 5.12 shows the temporal analysis and the descriptive statistics are shown in Table 5.5.

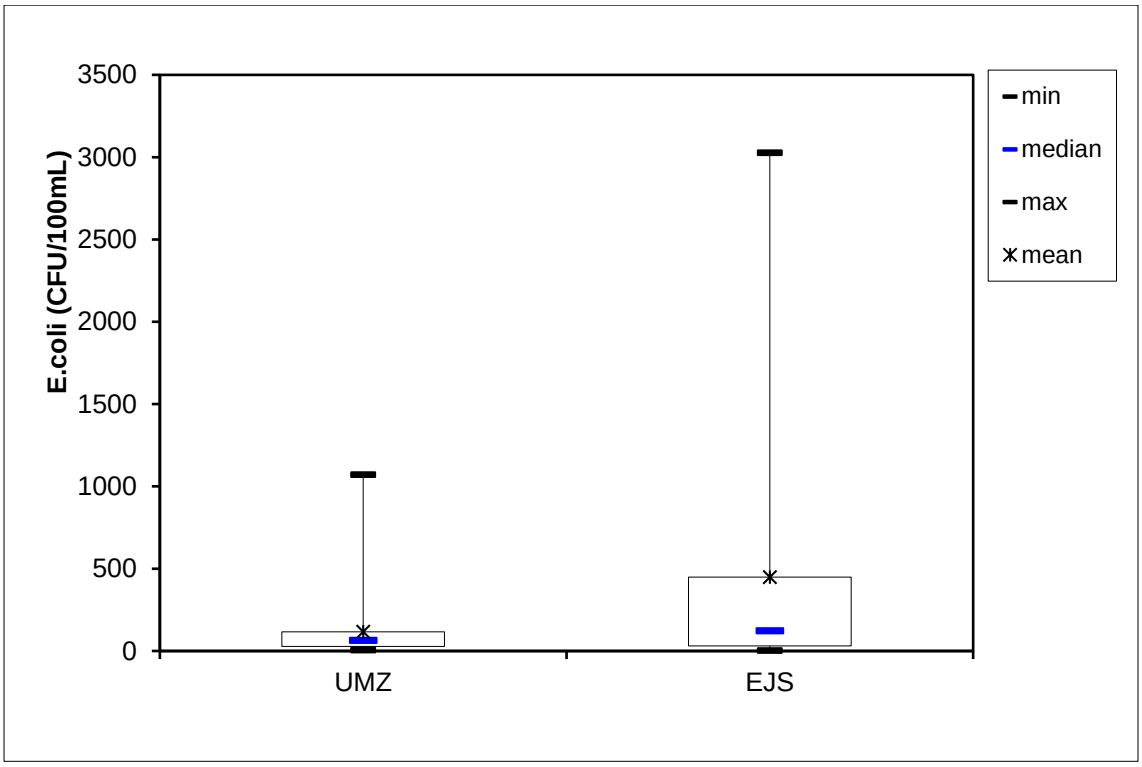


Figure 5.11: Raw water *Escherichia coli*

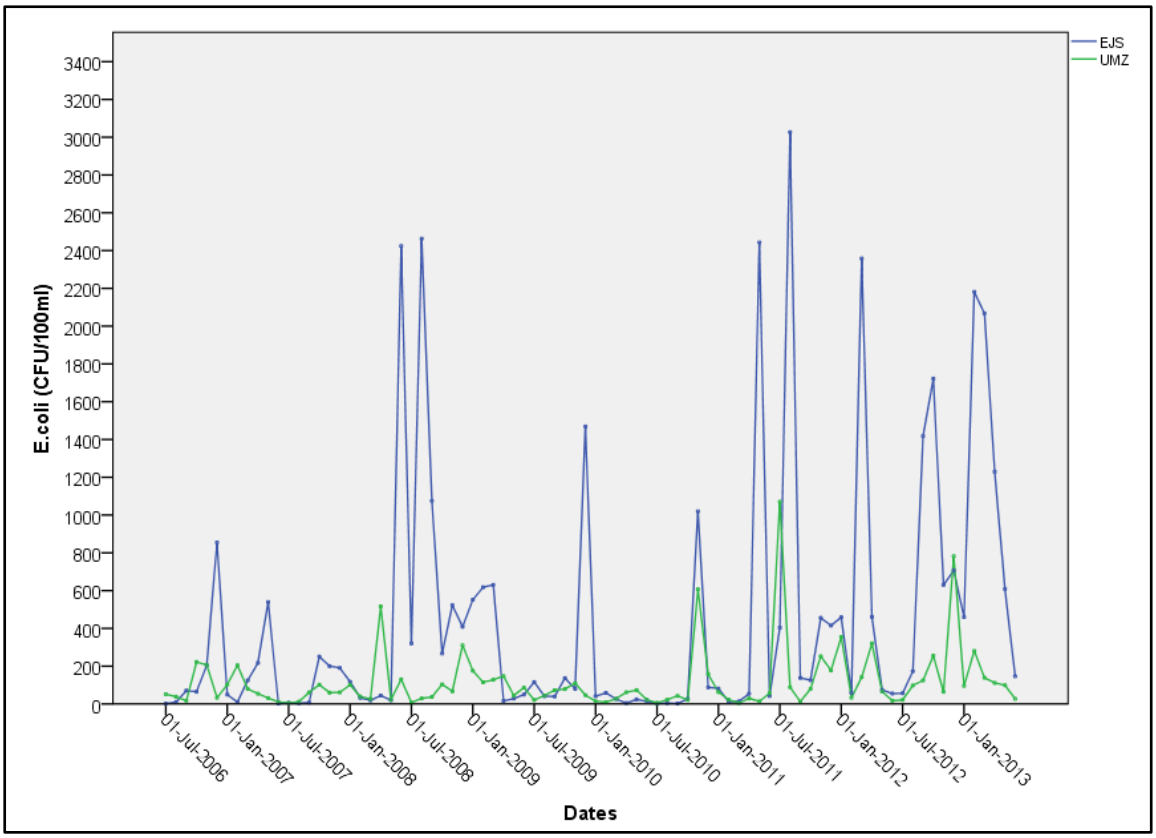


Figure 5.12: Temporal analysis of raw water *Escherichia coli*

Table 5.5: Descriptive statistics of raw water *Escherichia coli* for the period between July 2006 to June 2013

Abstraction point	Mean (CFU/100 mL)	Median (CFU/100 mL)	Std. Deviation (CFU/100 mL)	Range (CFU/100 mL)	Minimum (CFU/100 mL)	Maximum (CFU/100 mL)
Escherichia coli1	117	64	168	1066	5	1071
Ecoli2	449	122	705	3025	2	3027

The time-series plots in Figure 5.12 indicate seasonal fluctuation for the Umzinto River's *E. coli* levels. The EJ Smith Dam's raw water does not depict the same type of variation due to its irregular peaks throughout the year. The EJ Smith Dam is supplied by the Mzimayi River which passes through the town of Umzinto; hence, the irregular peaks could be attributed to discharge from broken wastewater distribution pipes flowing into the river.

South African guidelines stipulate that there should be no detection of *Escherichia coli* per 100 ml of drinking water (DWAF 1996). This means the *Escherichia coli* count range for both abstraction points makes the water unsafe for drinking without proper disinfection.

The high levels of *Escherichia coli* suggest that treatment should be much more effective prior to disinfection otherwise this will result in a high demand for chlorine which in turn increases the treatment costs.

Iron

Figure 5.13 shows the box and whisker plot for the raw water's iron content at the two abstraction points of the UWTP. Figure 5.14 shows the temporal analysis and the descriptive statistics are shown in Table 5.6.

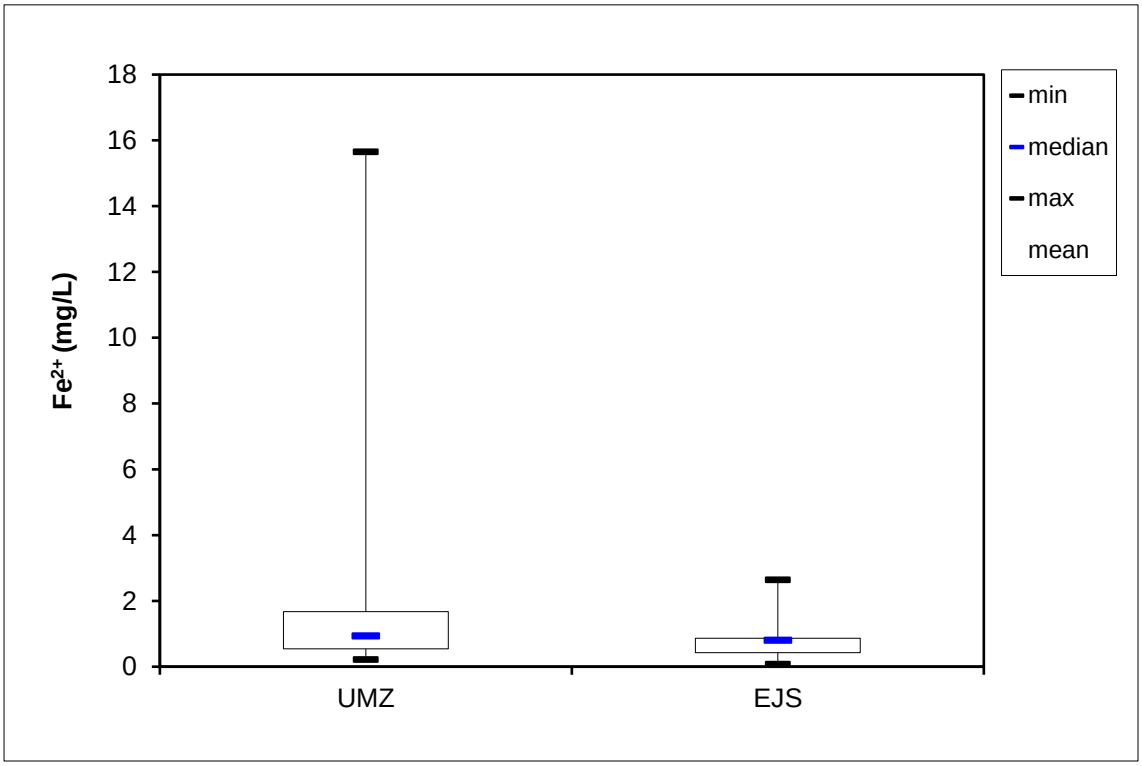


Figure 5.13: Raw water iron concentration

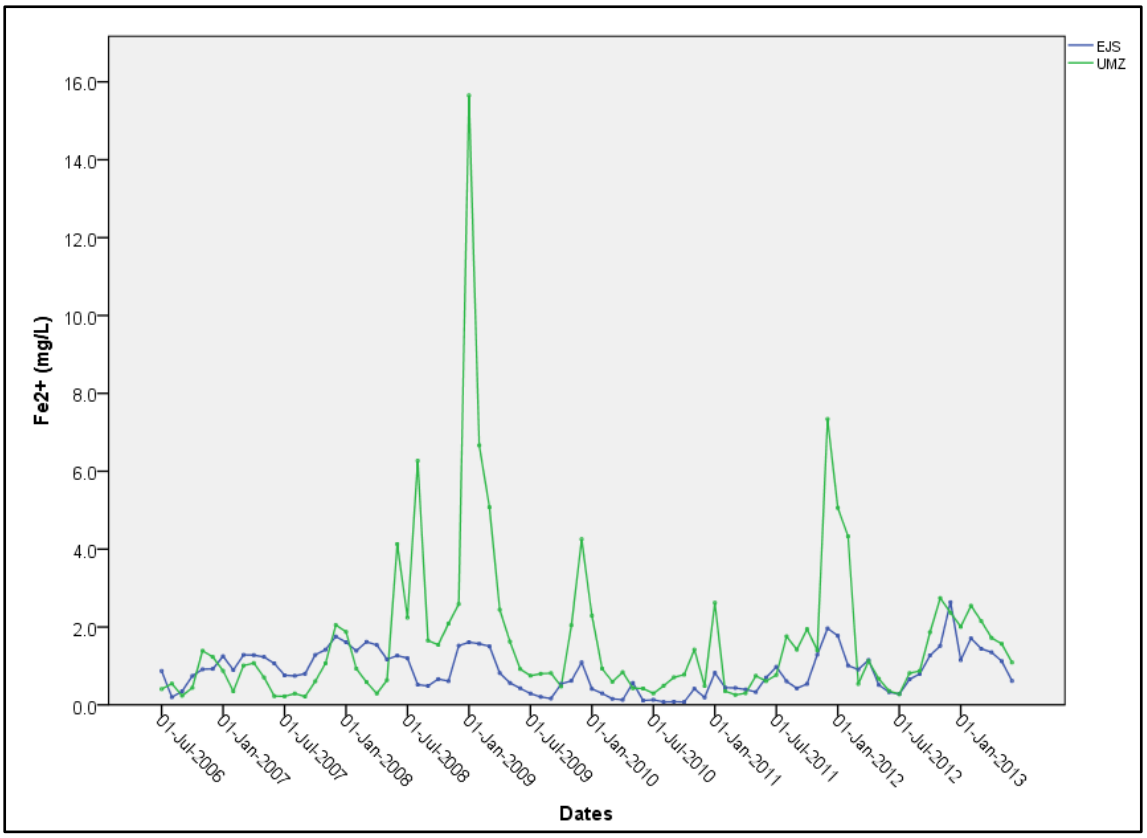


Figure 5.14: Temporal analysis of raw water iron concentration

Table 5.6: Descriptive statistics of raw water iron concentration for the period between July 2006 to June 2013

Abstraction Point	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
UMZ	1.7	0.9	2.1	15.4	0.2	15.7
EJS	0.9	0.8	0.5	2.6	0.1	2.6

In the South African guidelines for raw water, the Department of Water and Sanitation (DWAF 1996) stipulates a limit of less than < 2 mg/L for iron concentration. The mean values shown in Table 5.6 suggest that, most of the time, the iron concentration levels were in acceptable ranges. However, it is important to consider the peaks that were observed for Umzinto River. Figure 5.14 shows that the Umzinto River's raw water experiences some irregular peaks of iron content. This suggests the need to control the dosage of oxidising agents such as potassium permanganate. Iron pose risk of a condition called haemochromatosis. This condition is commonly caused by eating acid foodstuffs cooked in kitchenware made of iron for too long. Naturally, water does not contain excessive high concentrations of iron, hence iron poisoning is very rare (DWAF 1996).

Total Hardness

Figure 5.15 shows the box and whisker plot for raw water total hardness at the two abstraction points of the UWTP. Figure 5.16 shows the temporal analysis and the descriptive statistics are shown in Table 5.7.

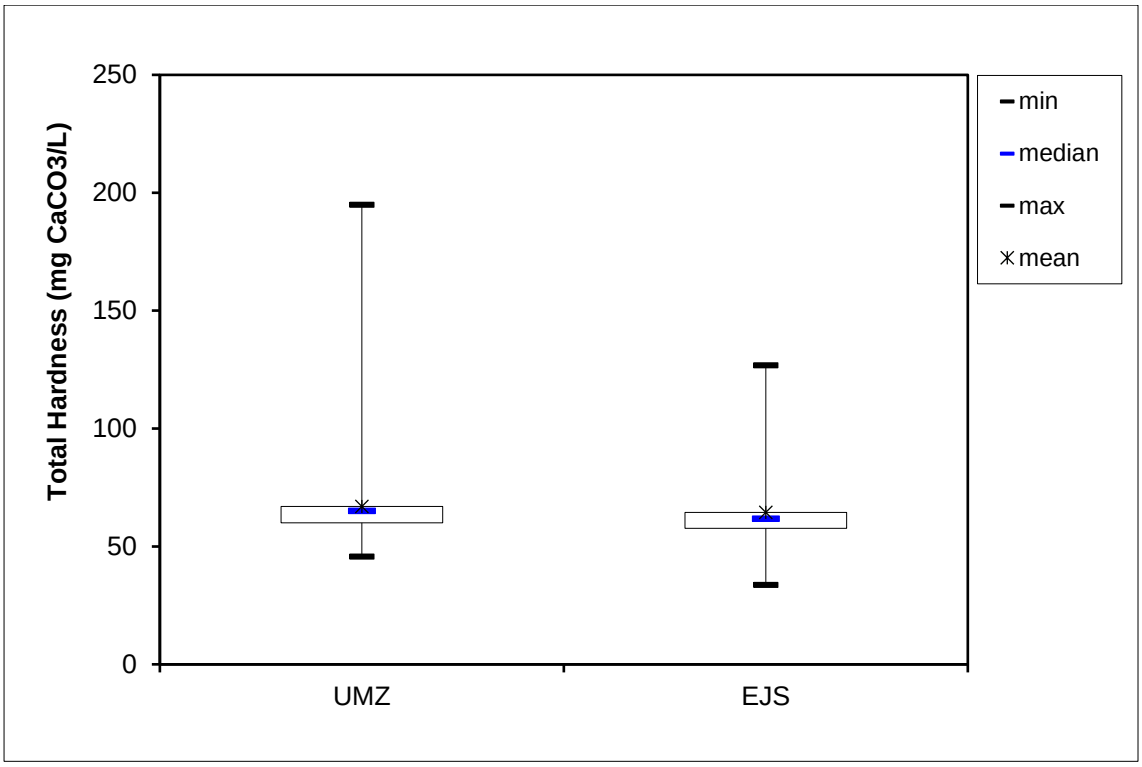


Figure 5. 15: Raw water total hardness

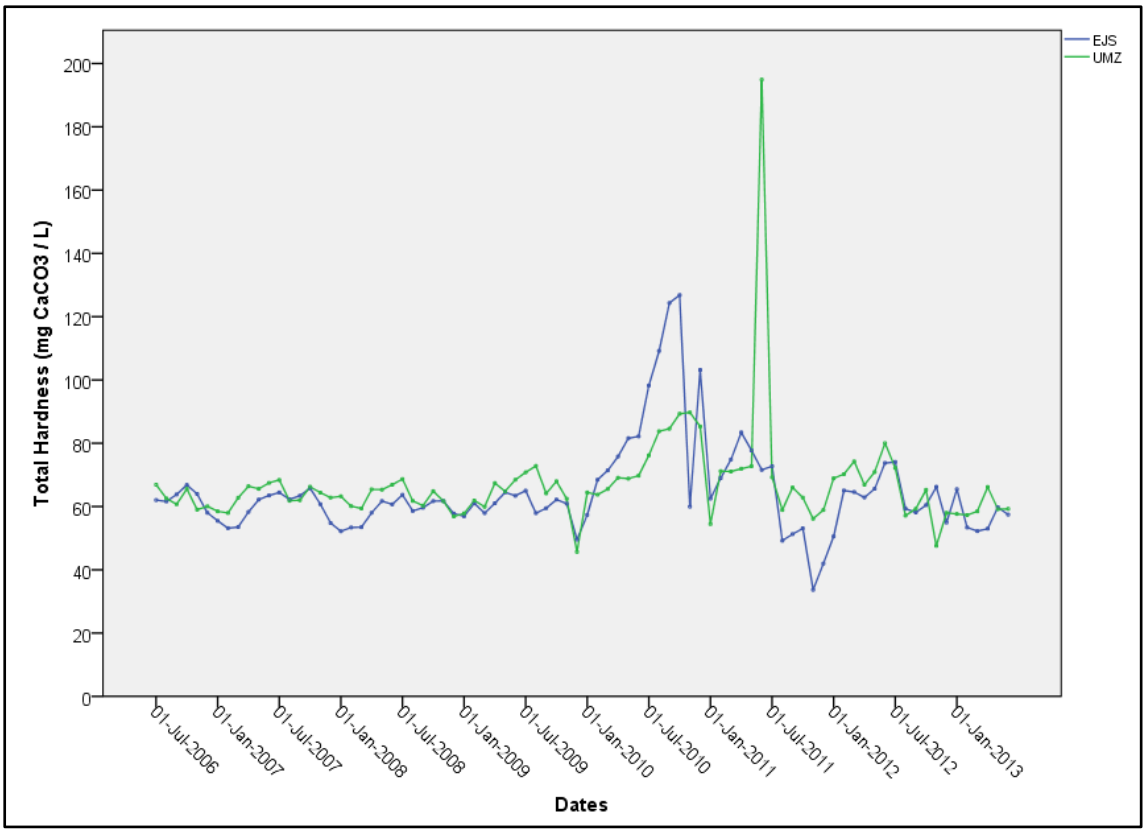


Figure 5.16: Temporal analysis of raw water total hardness

Table 5.7: Descriptive statistics of raw water total hardness for the period between July 2006 to June 2013

Abstraction Point	Mean ($mgCaCO_3$ /L)	Median ($mgCaCO_3$ /L)	Std. Deviation ($mgCaCO_3$ /L)	Range ($mgCaCO_3$ /L)	Minimum ($mgCaCO_3$ /L)	Maximum ($mgCaCO_3$ /L)
UMZ	67.0	65.1	16.2	149.2	45.7	194.9
EJS	64.4	61.7	14.9	93.2	33.7	126.8

The box and whisker plot shows that the mean and median values for the two abstraction points are almost similar meaning the two sources of abstraction have the same levels of total hardness. A very high value is noted in June 2011; this can be attributed to the high range value noted in Table 5.7. While South African standards do not set a limit for total hardness, it should be noted that this parameter is significantly linked to the taste of the water and consumer perceptions. Total hardness can also influence domestic uses of water such as laundry and bathing because it facilitates the formation of insoluble salts of long-chain fatty acids, the chief component of soaps. This makes total hardness important to control. This is also noted in the drinking water guidelines and a target range of 50 to 100 $mgCaCO_3/L$ is stipulated (DWAF 1996).

Manganese

Figure 5.17 shows the box and whisker plot for the raw water's manganese content at the two abstraction points of the UWTP. Figure 5.18 shows the temporal analysis and the descriptive statistics are shown in Table 5.8.

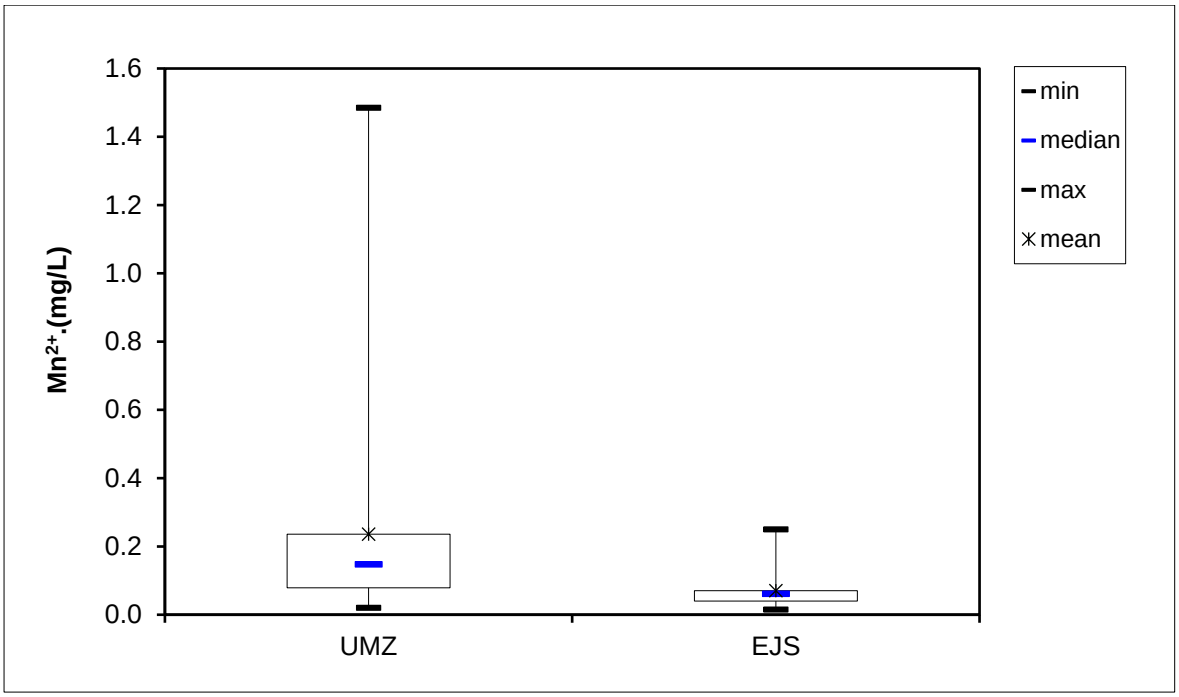


Figure 5.17: Raw water manganese content

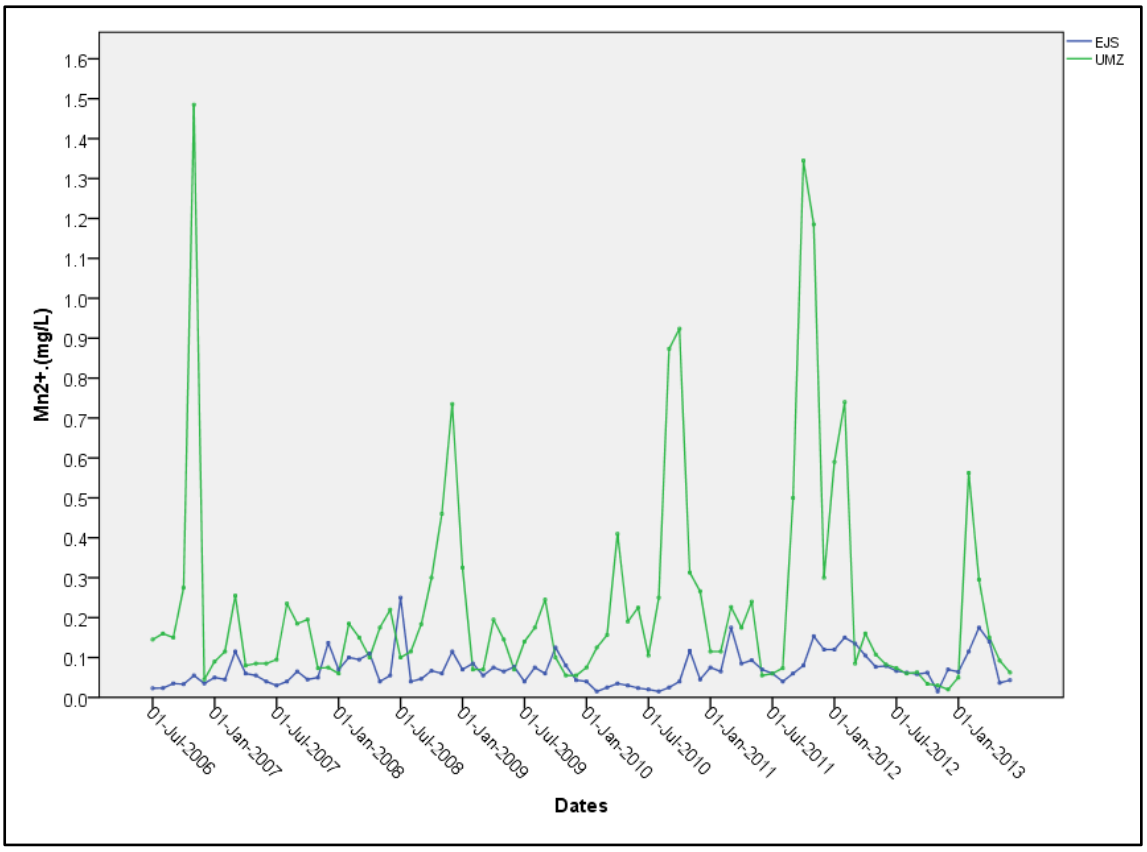


Figure 5. 18: Temporal analysis of raw water manganese content

Table 5.8: Descriptive Statistics for the raw water manganese content for the period between July 2006 to June 2013

Abstraction Points	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
UMZ	0.24	0.15	0.28	1.47	0.02	1.49
EJS	0.07	0.06	0.04	0.24	0.02	0.25

The South African standard for drinking water has two limits for the manganese content. One is for the aesthetic properties of the water and the other is for health purposes. The permissible aesthetic limit for manganese is <0.1 mg/L and the one for health is <0.5 mg/L (SABS, 2011). The temporal analysis in Figure 5.18 shows that the water from EJ Smith Dam is within the health permissible limit. However, of concern is the fact that the water from the Umzinto River has a seasonal pattern and can have peaks as high as 1.49 mg/L. Such peaks suggest the need to control the levels of raw water manganese content. The drinking water guidelines stipulate a target range between 0 and 0.05 to mitigate against both the anaesthetic and health effects of manganese (DWAF 1996).

pH

Figure 5.19 shows the box and whisker plot for the raw water pH at the two abstraction points of the UWTP. Figure 5.20 shows the temporal analysis and the descriptive statistics are shown in Table 5.9.

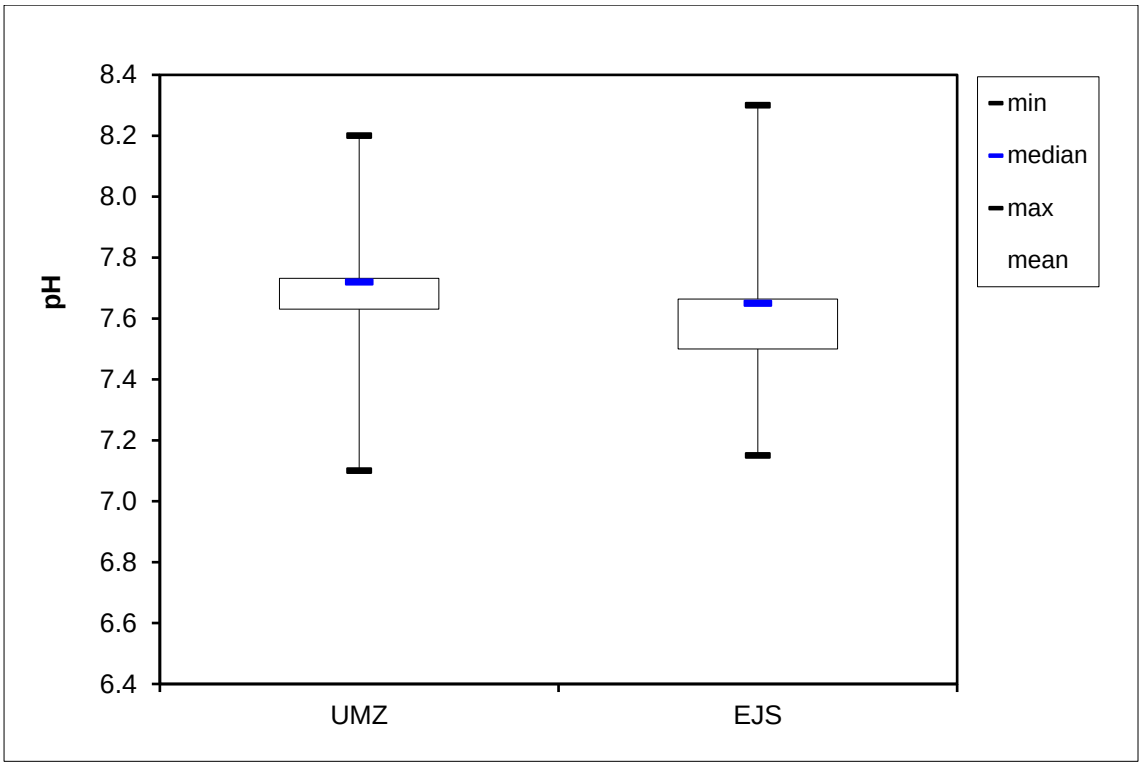


Figure 5.19: Raw water pH

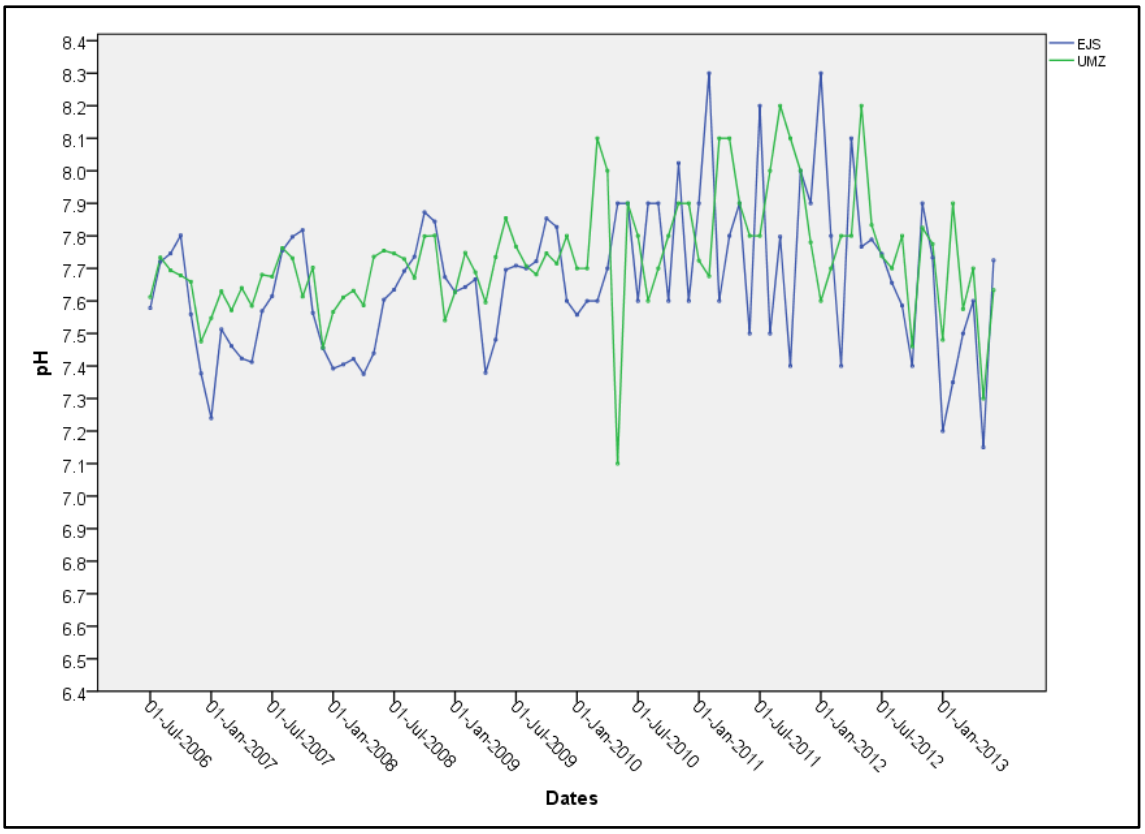


Figure 5.20: Temporal analysis of raw water pH

Table 5.9: Descriptive statistics for the raw water pH for the period between July 2006 to June 2013

Abstraction point	Mean	Median	Std. Deviation	Range	Minimum	Maximum
UMZ	7.7	7.7	0.2	1.1	7.1	8.2
EJS	7.7	7.6	0.2	1.2	7.2	8.3

The mean and median values for the two abstraction points are similar. The pH values range from 7.1 to 8.3. The pH values above 7.0 show that the raw water from the two abstraction points is alkaline. The South African standards stipulate a range between 5 and 9.7, which means that the two raw water sources are within that range (SABS, 2011). This should be a major benefit in the control of lime dosage and disinfection by chlorine. The WHO (2006) stipulates that disinfection using chlorine is more effective at pH values below 8. The levels of pH experienced are not very critical to be controlled. In the raw water guidelines the stipulated target is between 6.0 and 9.0. At this pH there is No significant effects on health due to toxicity of dissolved metal ions and protonated species, or on taste (DWAF 1996).

Suspended solids

Figure 5.21 shows the box and whisker plot for the raw water's suspended solids content at the two abstraction points of the UWTP. Figure 5.22 shows the temporal analysis and the descriptive statistics are shown in Table 5.10.

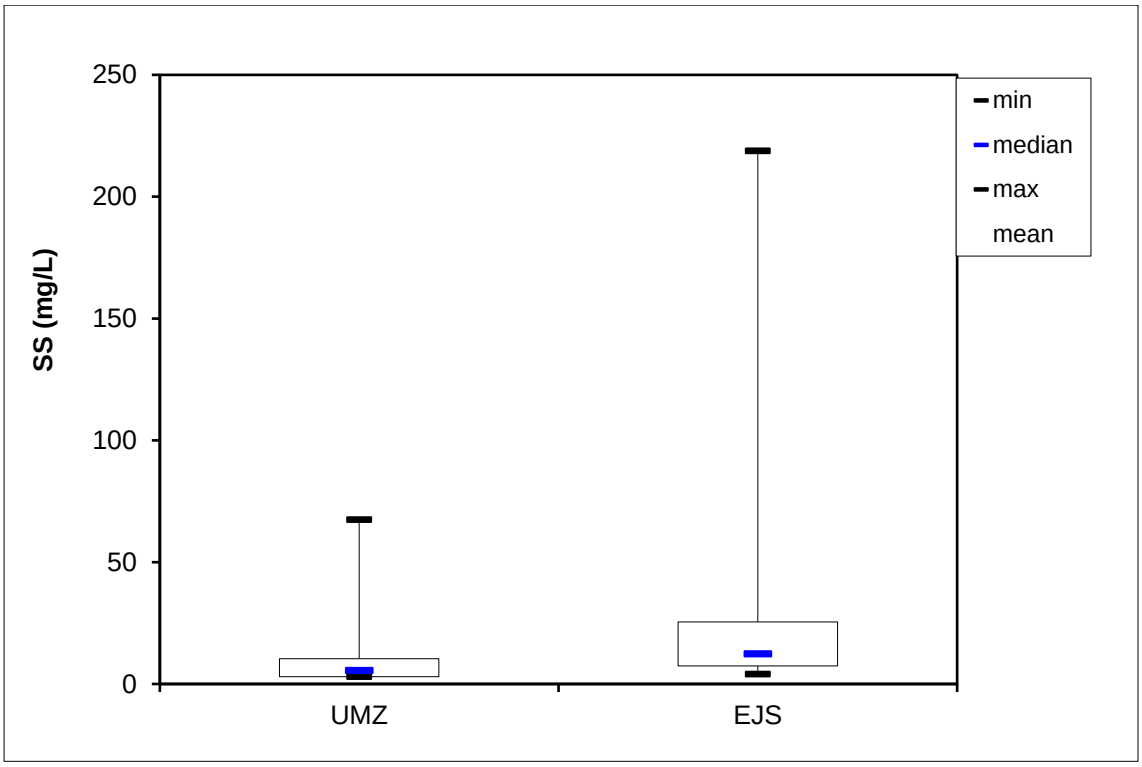


Figure 5.21: Raw water suspended solids

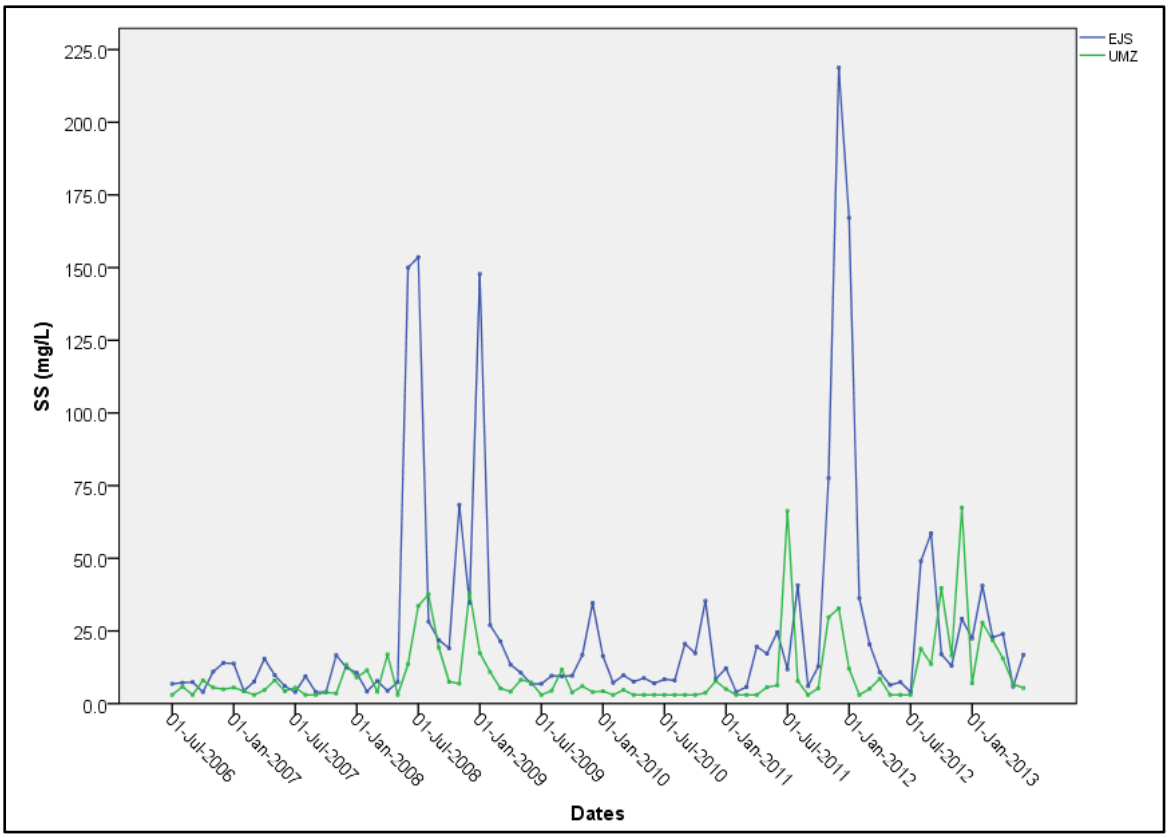


Figure 5.22: Temporal analysis of raw water suspended solids

Table 5.10: Descriptive statistics of raw water suspended solids for the period between July 2006 to June 2013

Abstraction Point	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
UMZ	10.4	5.5	12.6	64.4	3.0	67.4
EJS	25.5	12.3	39.1	214.8	4.0	218.8

The box and whisker plot shows that the range for suspended solids for the raw water from EJ Smith Dam is wider than that from Umzinto River. The temporal analysis shows peaks in 2008, 2009 and 2012 that correlate to the temporal analysis of colour and *Escherichia coli*. These peaks could be a result of direct discharge from broken wastewater pipes into the Mzimayi River. The South African standards and guidelines do not set limits for suspended solids for raw and drinking water (DWAF, 1996; SABS,2011). Therefore, to control suspended solids, it is important to consider its correlation with other drinking water quality parameters (Shroff *et al.* 2015).

Total coliforms

Figure 5.23 shows the box and whisker plot for total coliforms at the two abstraction points of the UWTP. Figure 5.24 shows the temporal analysis and the descriptive statistics are shown in Table 5.11.

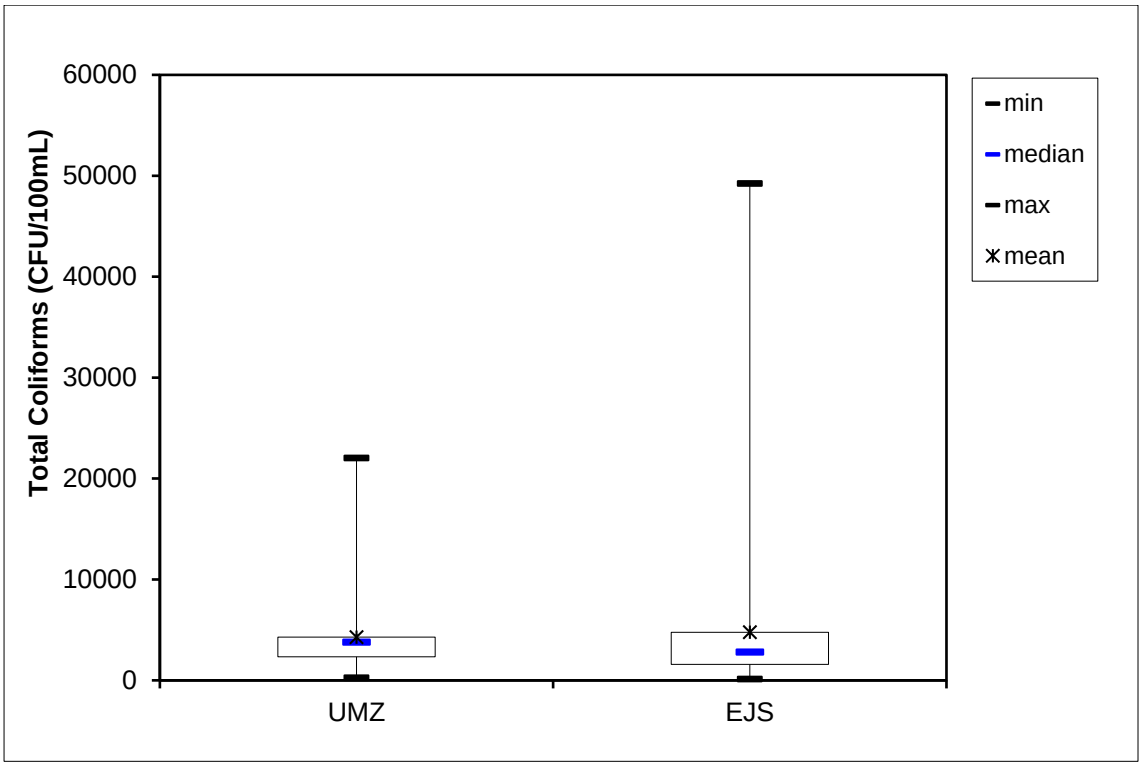


Figure 5.23: Raw water total coliforms

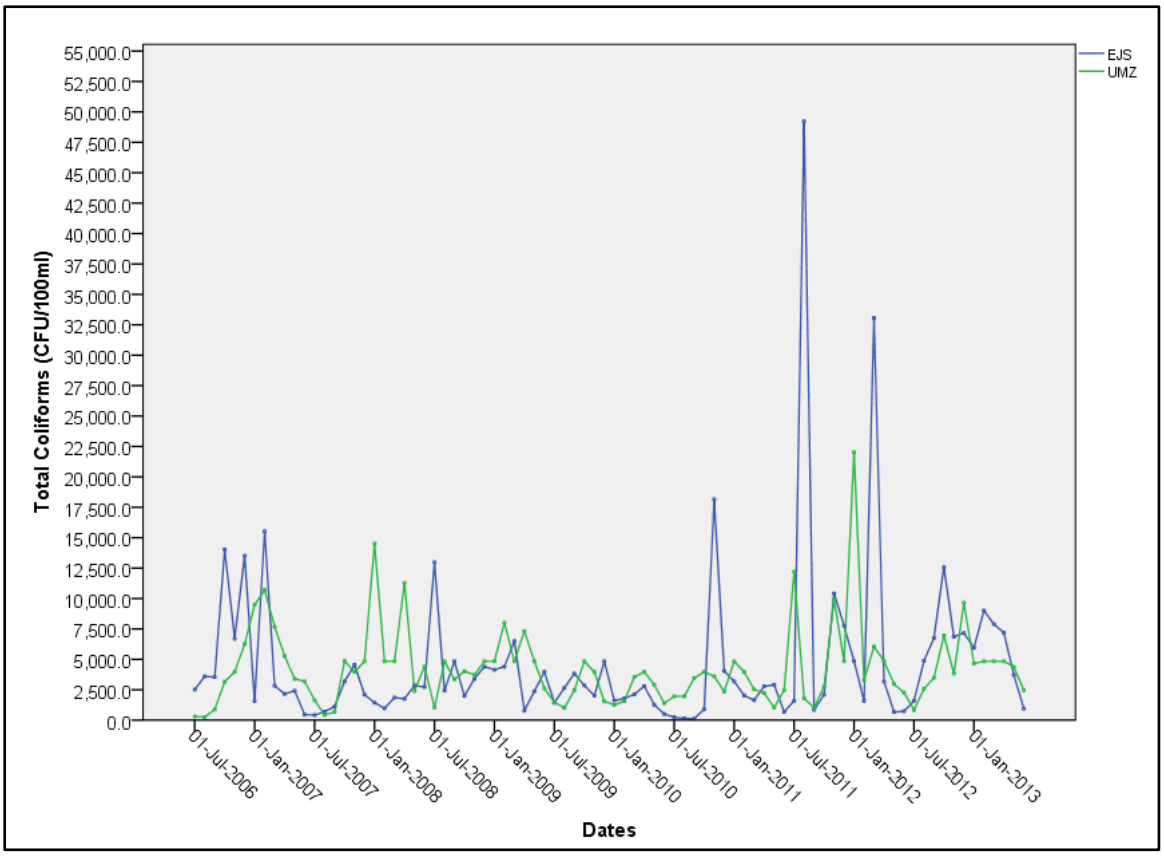


Figure 5.24: Temporal analysis of raw water total coliforms

Table 5.11: Descriptive statistics of raw water total coliforms for the period between July 2006 to June 2013

Abstraction Point	Mean (CFU/100 mL)	Median (CFU/100 mL)	Std. Deviation (CFU/100 mL)	Range (CFU/100 mL)	Minimum (CFU/100 mL)	Maximum (CFU/100 mL)
UMZ	4278	3775	3392	21773	253	22025
EJS	4762	2793	6921	49103	129	49231

According to the South African Bureau of Standards (SABS, 2011), the permissible limit for coliforms is 10 CFU/100ml. The box and whisker plot therefore shows that the water from the two sources is unsafe for drinking before it is treated. The temporal analysis shows seasonal variations; levels of total coliforms increase during the rainy season. In the raw water quality guidelines total coliforms refer to all bacteria which produce colonies with a typical metallic sheen within 20 - 24 hours of incubation at 35 °C on m-Endo agar (DWAF 1996). The presence/absence of total coliforms in water shows how safe the water is, as they include faecal related bacteria. However, the same guideline document suggests that most of the bacteria in this group may originate from growth in the marine environment.

Bacteria generally originating from faecal sources, such as *Salmonella spp., Vibrio cholerae, Campylobacter jejuni, Yersinia enterocolitica* and pathogenic *Escherichia coli*, pose a risk to diseases such as typhoid fever, cholera and gastroenteritis (DWAF 1996).

Temperature

Figure 5.25 shows the box and whisker plot for the water temperature at the two abstraction points of the UWTP. Figure 5.26 shows the temporal analysis and the descriptive statistics are shown in Table 5.12.

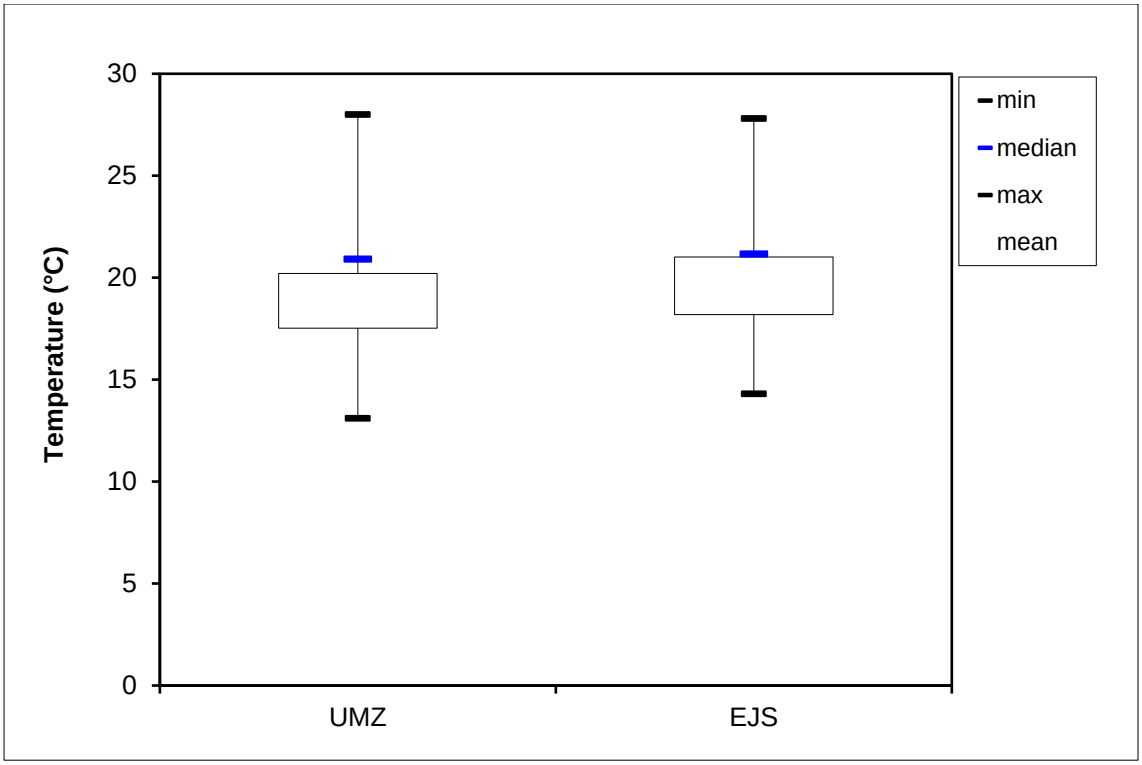


Figure 5.25: Raw water temperature

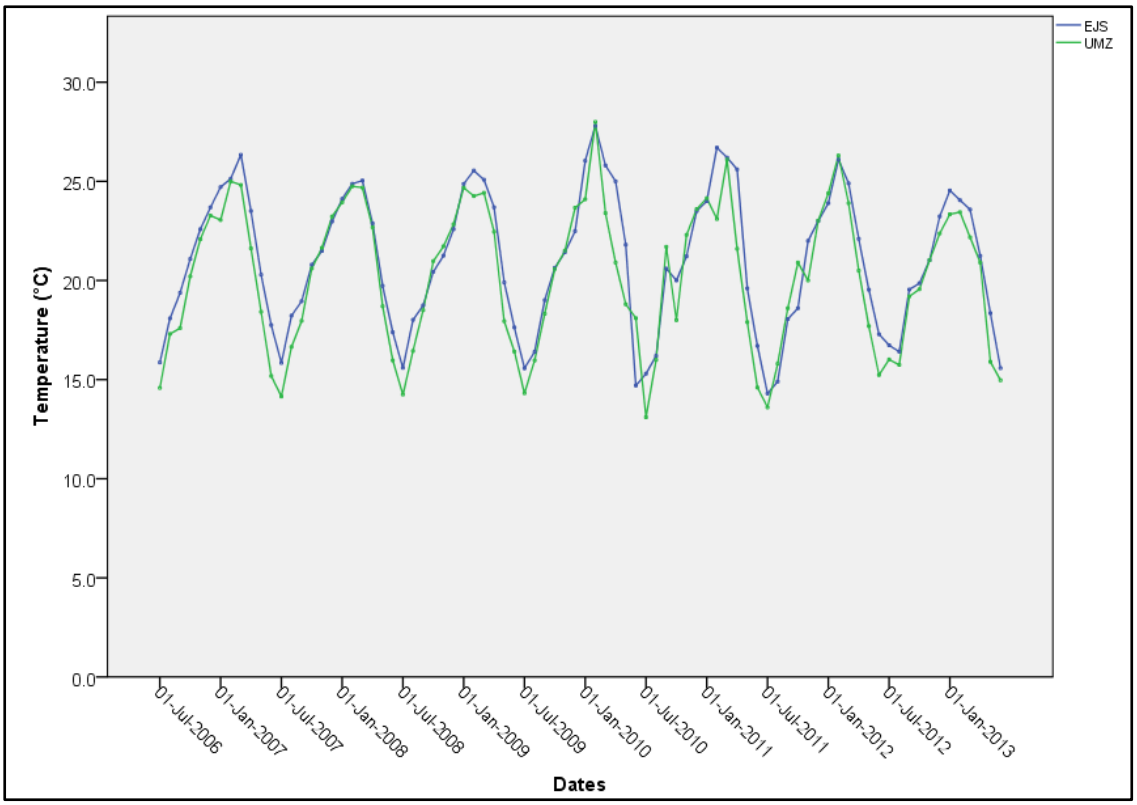


Figure 5.26: Temporal analysis of raw water temperature

Table 5.12: Descriptive statistics of raw water temperature for the period between July 2006 to June 2013

Abstraction point	Mean (°C)	Median (°C)	Std. Deviation (°C)	Range (°C)	Minimum (°C)	Maximum (°C)
UMZ	20.2	20.9	3.6	14.9	13.1	28.0
EJS	21.0	21.2	3.5	13.5	14.3	27.8

As expected, raw water temperature tends to show seasonal variations, with highs in summer and lows in winter. There is not much that can be done to control raw water temperature but the effects that it can have on coagulation and disinfection should be taken into account. Studies have shown that low raw water temperatures such as those in Figure 5.26, slow down the process of coagulation and flocculation (Hanson and Cleasby 1990; Tae-Hwan *et al.* 1997; Upadhyay and Agarwal 2008). This has a major influence on the efficiency of reducing turbidity. Furthermore, studies by Hanson and Cleasby (1990) and Rossini, Garrido and Galluzzo (1999) showed that low temperatures result in the formation of flocs that can easily break up when subjected to shear force.

Veenstra and Schnoor (1980) noted that chlorine can be lost during summer due to radiation. Larger dosages of chlorine are thus required to compensate for radiation loss. High temperatures can also promote algae blooms.

Total Organic Carbon

Figure 5.27 shows the box and whisker plot for total organic carbon at the two abstraction points of the UWTP. Figure 5.28 shows the temporal analysis and the descriptive statistics are shown in Table 5.13.

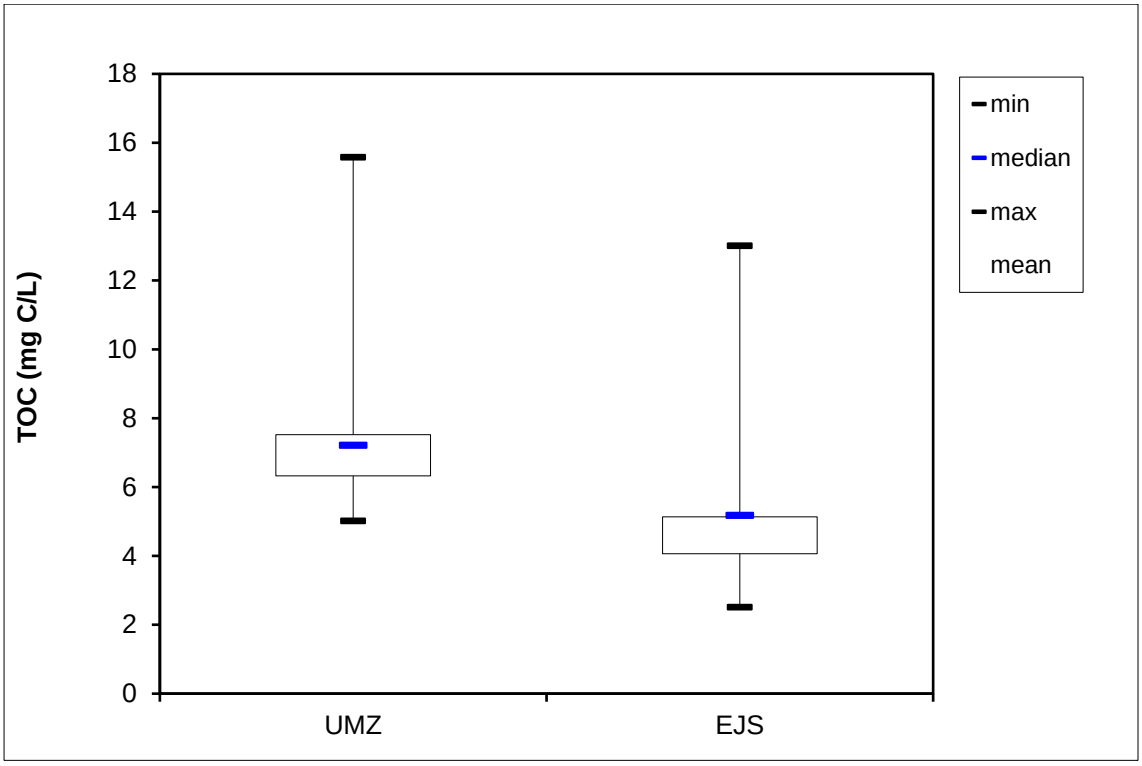


Figure 5.27: Raw water total organic carbon

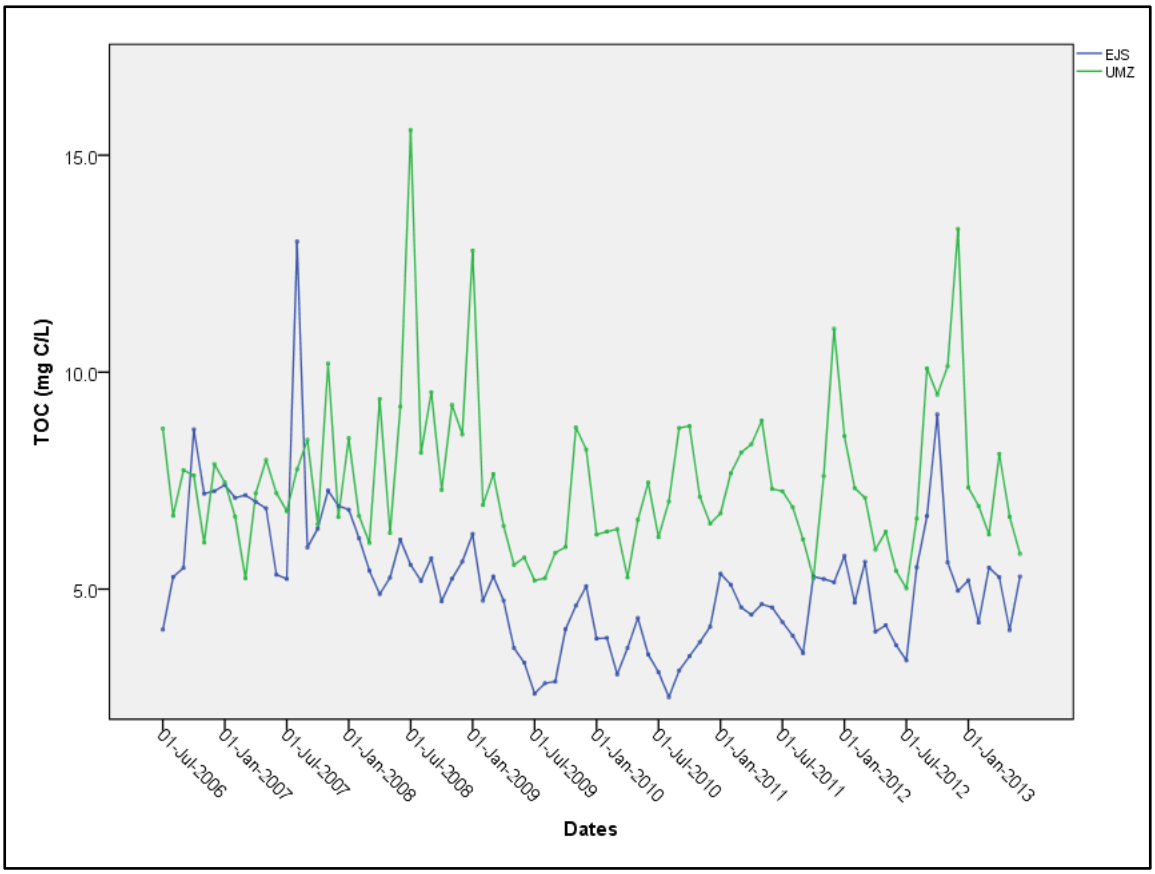


Figure 5.28: Temporal analysis of raw water total organic carbon

Table 5.13: Descriptive statistics of raw water total organic carbon for the period between July 2006 to June 2013

Abstraction Point	Mean (mg C/L)	Median (mg C/L)	Std. Deviation (mg C/L)	Range (mg C/L)	Minimum (mg C/L)	Maximum (mg C/L)
UMZ	7.5	7.2	1.8	10.6	5.0	15.6
EJS	5.1	5.2	1.6	10.5	2.5	13.0

The South African standard stipulates a limit of 10 mg/L (SABS, 2011). The box and whisker plot, and the temporal analysis show that most of the observations were within this limit. Those that were above would fall within the maximum allowable limit for a given period. Mamba *et al.* (2009) noted that this standard limit is between 10 and 70 mg/L for treated water, which would make this water safe according to this parameter. The raw water guideline states that total organic carbon is the sum of both the dissolved and suspended organic carbon, however, for drinking water purposes the dissolved organic carbon component is of more concern as it has a direct relationship with organic material in water (DWAF 1996). This type of material can pass through coagulation, flocculation,

sedimentation and filtration, to the disinfection stage. If the disinfectant is chlorine, disinfection by-products may be formed and pose a great risk to health (Hrudey 2009; Richardson and Postigo 2012).

Turbidity

Figure 5.29 shows the box and whisker plot for turbidity at the two abstraction points of the UWTP. Figure 5.30 shows the temporal analysis and the descriptive statistics are shown in Table 5.14.

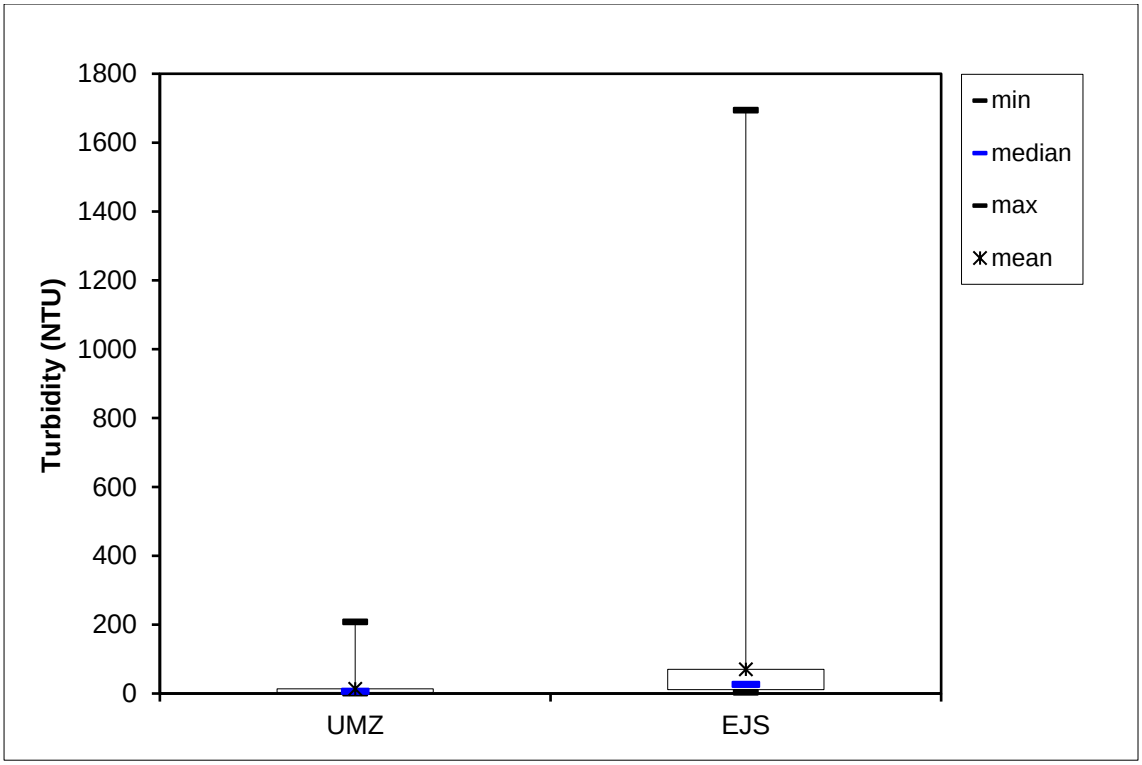


Figure 5.29: Raw water turbidity

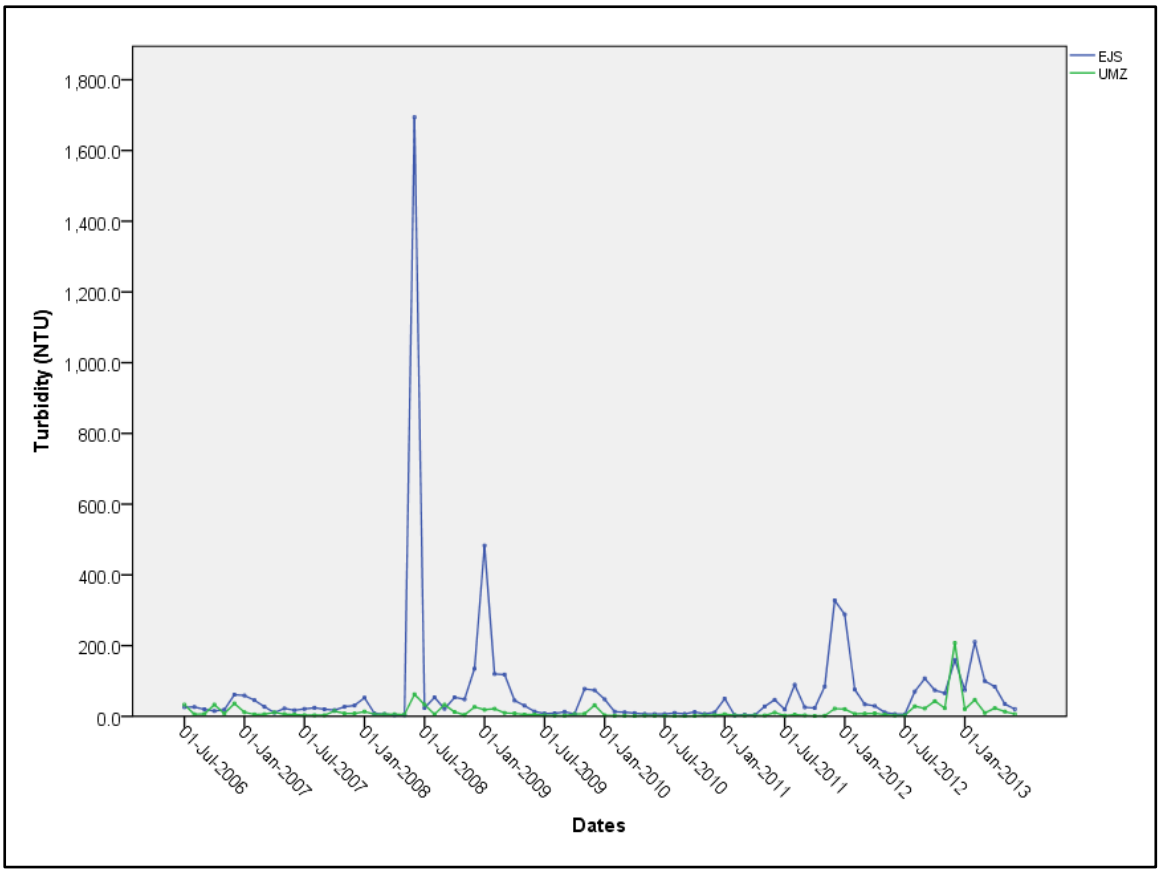


Figure 5.30: Temporal analysis of raw water turbidity

Table 5.14: Descriptive statistics of raw water turbidity for the period between July 2006 to June 2013

Abstraction Point	Mean (NTU)	Median (NTU)	Std. Deviation (NTU)	Range (NTU)	Minimum (NTU)	Maximum (NTU)
UMZ	13.5	6.1	24.6	207.0	0.7	207.7
EJS	70.3	25.9	194.0	1690.9	3.2	1694.0

The South African standard limit for drinking water turbidity is 0 to 1 NTU (SABS, 2011). The calculated means and medians therefore show that the water from the two abstraction points is unsafe before treatment. In the water quality guidelines consumption of highly turbid water is shown to noted have no direct health effects (DWAF 1996). However, a major concern is that, high levels of turbidity will reduce the effectiveness of disinfection. They also increase the chances of formation of disinfection by-products, such as trihalomethanes, halogenic acetic acids, chlorine hydrates, chloropicrin, halofuranones and bromohydrins, when using chlorine. These pose a risk of cancer to the consumer. Another concern is the fact that high turbidity inhibits the action of disinfectant on pathogenic

bacteria (DWAF 1996). Figure 5.30 shows that the EJ Smith Dam's turbidity should be of major concern because of its irregular peaks.

5.3.2 Temporal analysis of chemical dosages

Chlorine dosage

Figure 5.31 shows the temporal analysis of chlorine dosage at the UWTP and its descriptive statistics are shown in Table 5.15.

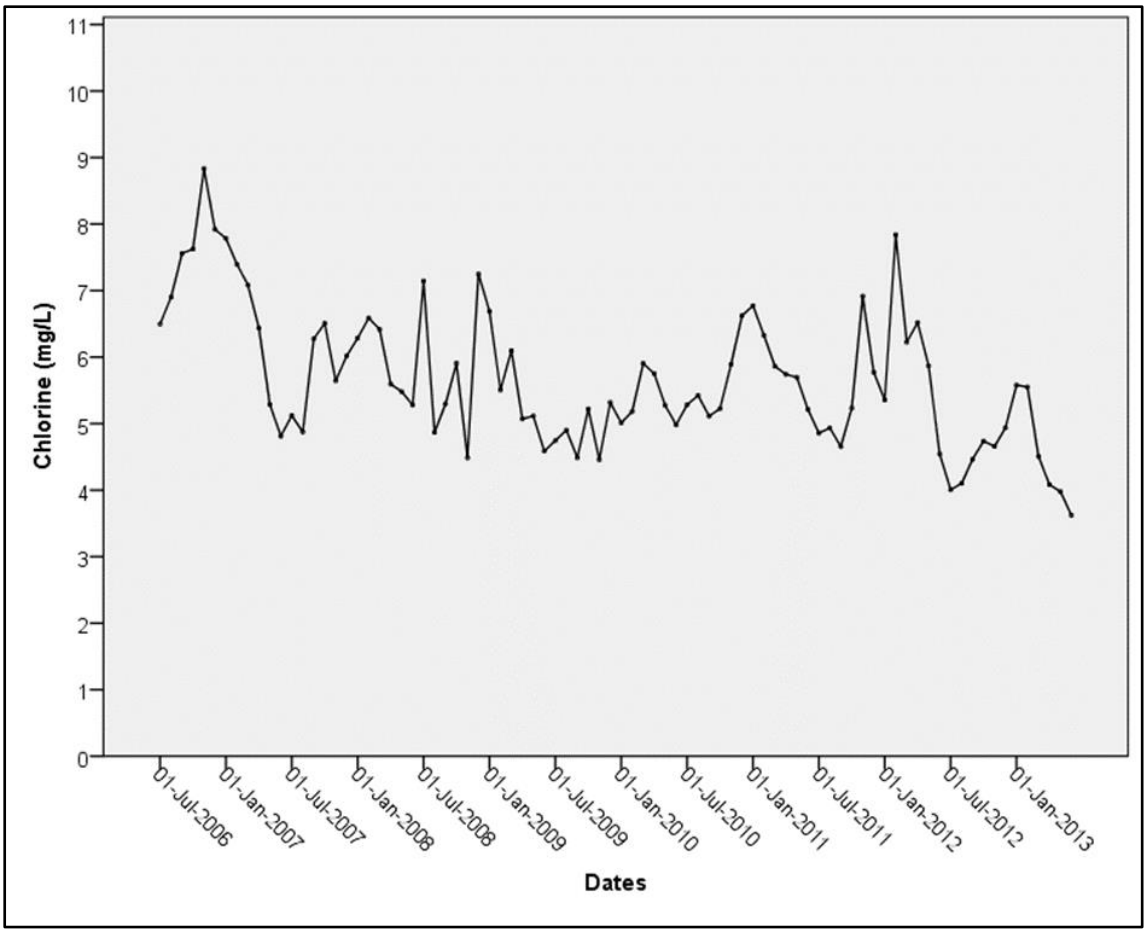


Figure 5.31: Temporal analysis of chlorine dosage

Table 5.15: Descriptive statistics for chlorine dosage for the period between July 2006 to June 2013

Chemical	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
Chlorine	5.7	5.5	1.04	5.21	3.62	8.83

At the UWTP, chlorine is used for disinfection. The temporal analysis shows that chlorine dosage follows seasonal variations. In winter when there is low or no rainfall, the dosages are low, but they increase during the rainy season. Veenstra and Schnoor (1980) noted that chlorine can be lost during summer due to radiation. Larger dosages of chlorine are thus required to compensate for radiation loss. High temperatures can also promote algae blooms, which has been supported in the research by Toroz and Uyak (2005). This has also been confirmed in recent talking about the increase in the formation of disinfection by products during the summer rains (Liu and Reckhow 2013; Uyak *et al.* 2014; Charisiadis *et al.* 2015). The overall trend shows that the dosage at the treatment plant has been decreasing since 2006, this could be due to the decreasing capacity at the UWTP.

Polymer dosage

Figure 5.32 shows the temporal analysis of polymer dosage at the UWTP and its descriptive statistics are shown in Table 5.16. For this study the polymer used between 2006 and 2013 was assumed to be a single type but in a practical scenario polymers change over the years. The changes are not presented for reasons of confidentiality.

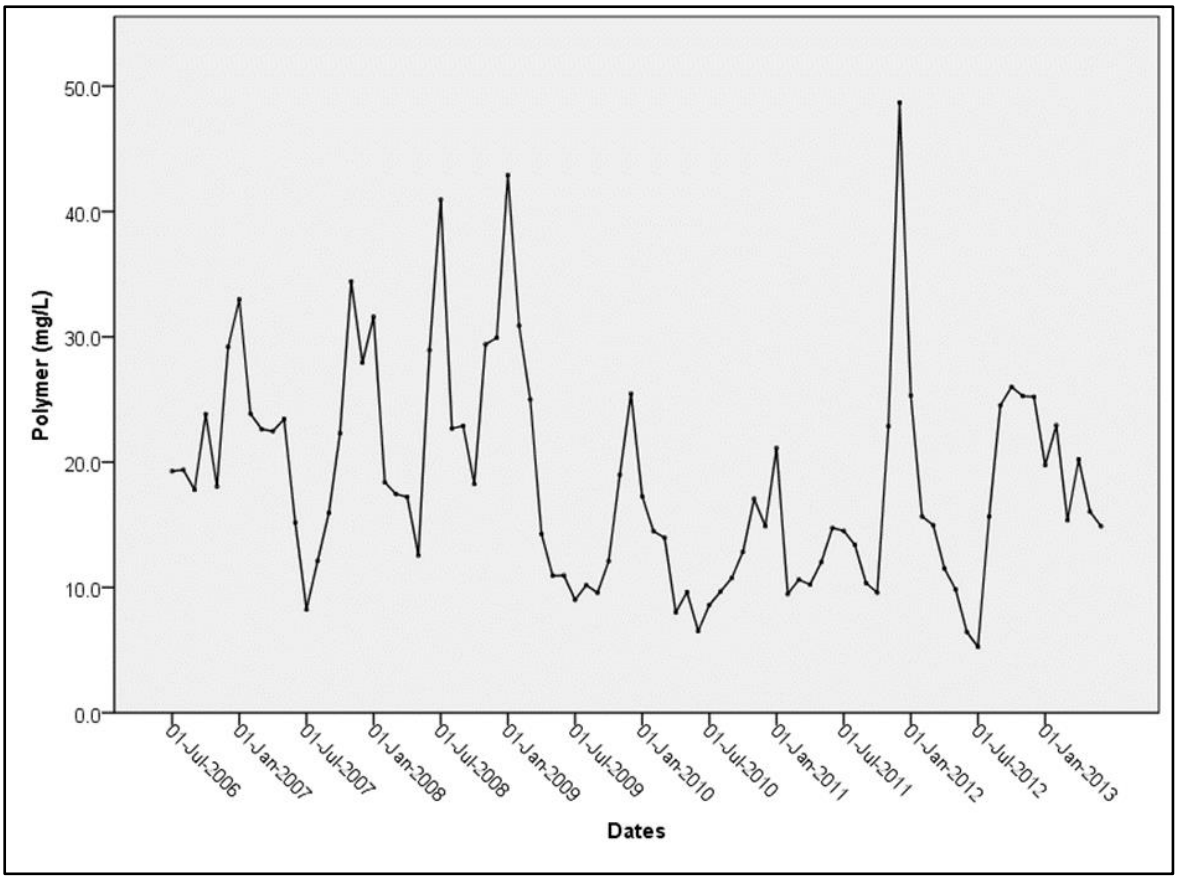


Figure 5.32: Temporal analysis of polymer dosage

Table 5.16: Descriptive statistics for polymer dosage for the period between July 2006 to June 2013

Chemical	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
Polymer	18.5	17.2	8.6	43.44	5.26	48.70

At the UWTP, polymer is used as a coagulant. The temporal analysis trend shows that the dosage experienced some peaks, and that these mainly occurred during the rainy season. In literature this has been attributed an increase in run off that consequentially increase natural organic matter flowing in to the rivers (Matilainen *et al.* 2002; Fearing *et al.* 2004; Kim and Yu 2005; Yan *et al.* 2006; Sillanpää and Matilainen 2014). Thus, there are seasonal variations in polymer dosage at the treatment plant.

Lime dosage

Figure 5.33 shows the temporal analysis of lime dosage at the UWTP and its descriptive statistics are shown in Table 5.17.

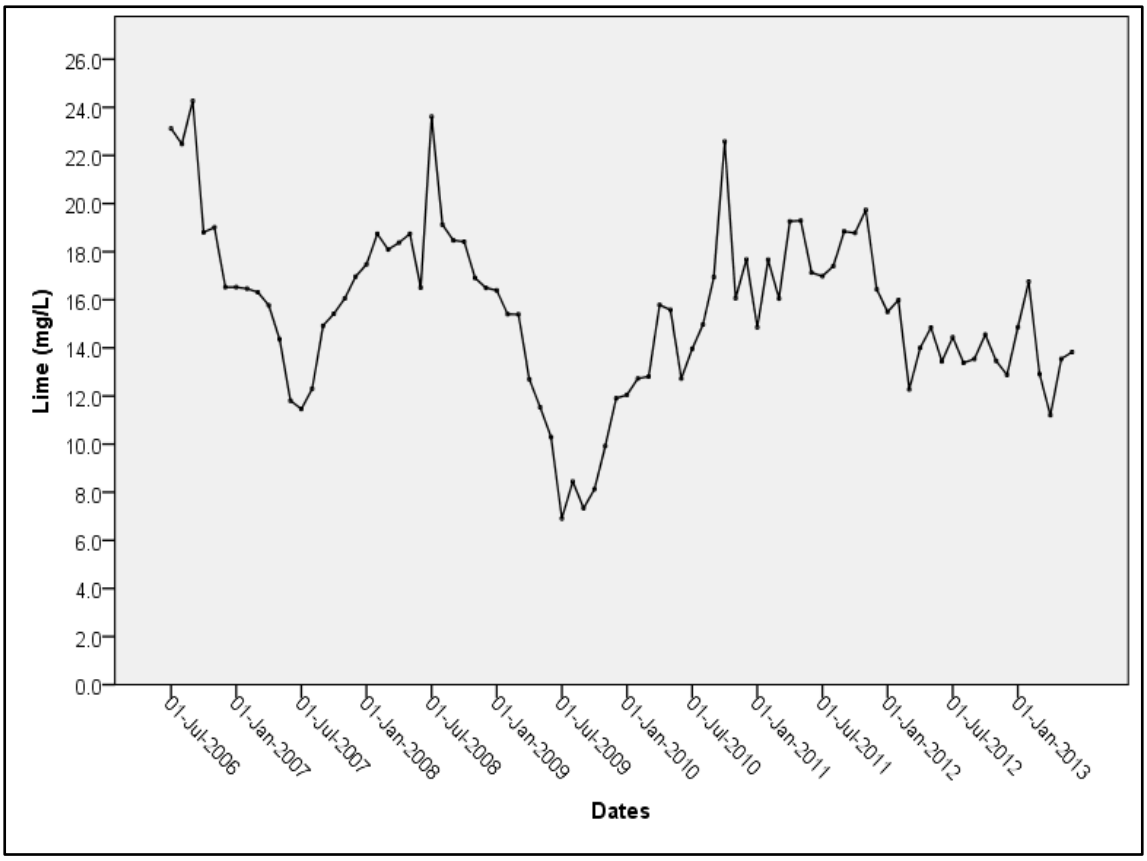


Figure 5.33: Temporal analysis of lime dosage

Table 5.17: Descriptive statistics for lime dosage for the period between July 2006 to June 2013

Chemical	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
Lime	15.5	15.8	3.4	17.4	6.9	24.3

Lime is used for the control of pH. The temporal analysis shows an irregular variation of lime at UWTP. This could have been caused by the irregular variation of the raw water pH.(Rashed *et al.* 2013). In addition to this relationship with pH, Farzadkia and Bazrafshan (2014) states that Lime dosage varies with the concentration of solids in the water. This is also a parameter that had an irregular variation and could be a major contributor for the variation experienced at UWTP.

Potassium permanganate

Figure 5.34 shows the temporal analysis of potassium permanganate dosage at the UWTP and its descriptive statistics are shown in Table 5.18.

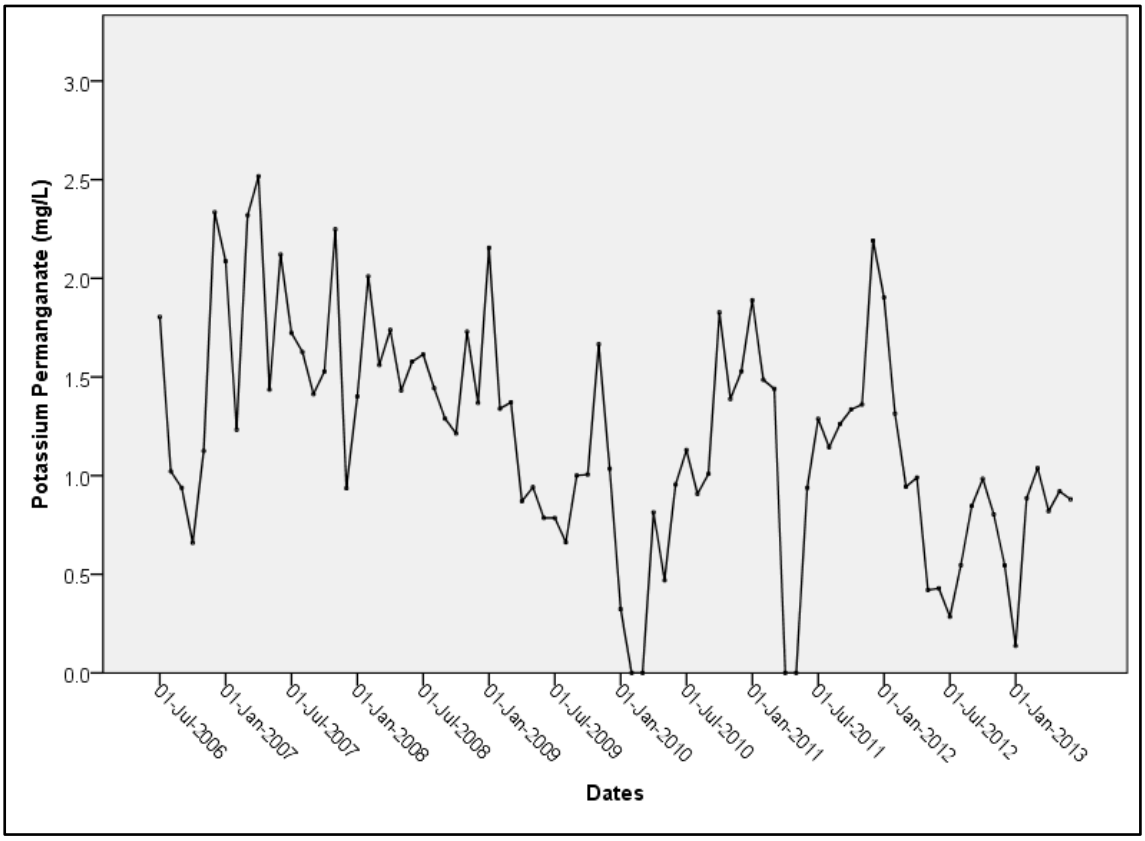


Figure 5.34: Temporal analysis of potassium permanganate dosage

Table 5.18: Descriptive statistics for potassium permanganate dosage for the period between July 2006 to June 2013

Chemical	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
Potassium Permanganate	1.2	1.2	0.6	2.5	0.0	2.5

Potassium permanganate is an oxidising agent. At the UWTP it is used to control the colour of the water and to remove iron and manganese. Bryant, Fulton and Budd (1992), suggested that it can also be used to control the formation of disinfection by-products such as trihalomethanes, halogenic acetic acids, chlorine hydrates, chloropicrin, halofuranones and bromohydrins. These can pose a great risk to cancer to the consumers (Ashbolt 2004). The temporal analysis shows irregular variation in potassium permanganate dosage and this suggests that it is important to control the parameters associated with it.

Activated Carbon

Figure 5.35 shows the temporal analysis of activated carbon dosage at the UWTP and its descriptive statistics are shown in Table 5.19.

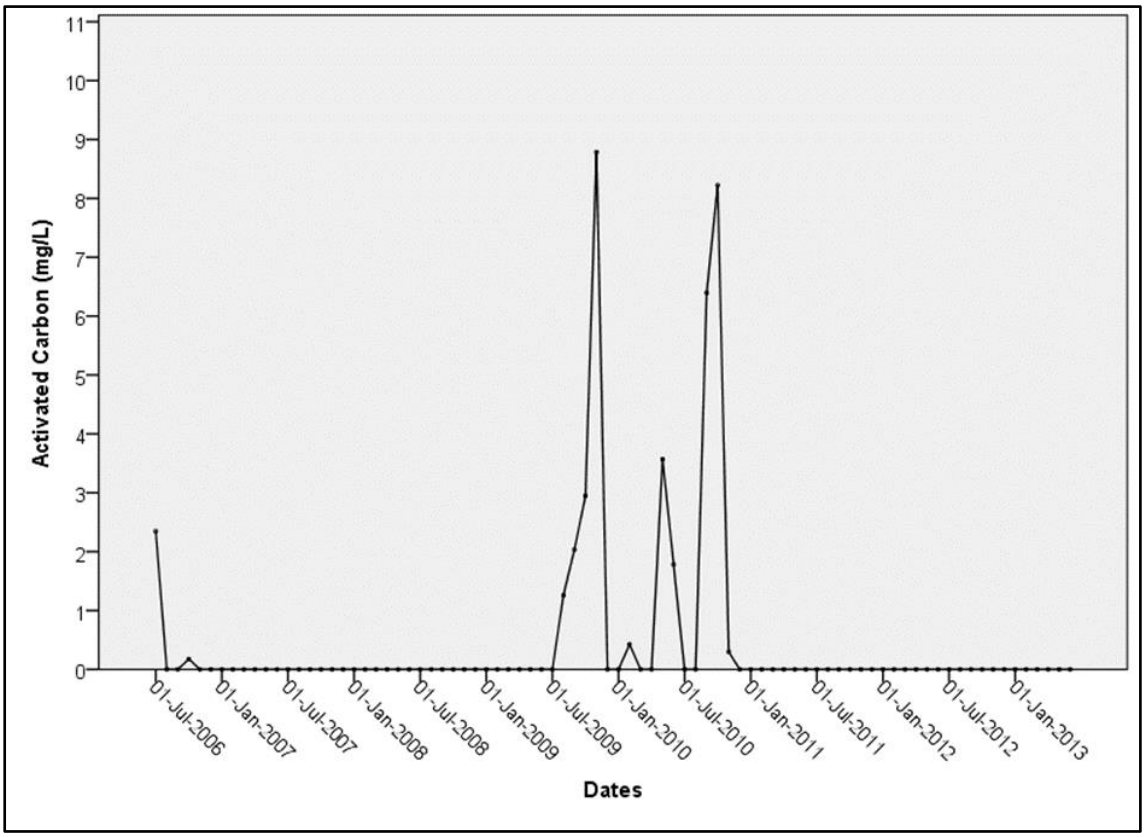


Figure 5.35: Temporal analysis of activated carbon dosage

Table 5.19: Descriptive statistics for activated carbon dosage for the period between July 2006 to June 2013

Chemical	Mean (mg/L)	Median (mg/L)	Std. Deviation (mg/L)	Range (mg/L)	Minimum (mg/L)	Maximum (mg/L)
Activated Carbon	0.46	0	1.57	8.79	0.00	8.79

Figure 5.35 shows that most of the time the observed dosage of activated carbon was zero. Activated carbon is used to control the taste and odour of the water. This is related to its total hardness. Comparing the total hardness trend to the activated carbon dosage suggest a possible error in the recordings of activated carbon.

5.3.3 Key raw water quality parameters affecting treatment

The final aspect that was used to select the key quality parameters affecting water treatment was correlation, shown in Table 5.20.

Table 5.20: Correlation of water quality parameters to the chemical dosage for the period between July 2006 to June 2013

		Chlorine	Polymer	Lime	Activated Carbon	Potassium Permanganate
Alga count	**UMZ**	-0.122	-0.222*	-0.029	0.312**	-0.060
	EJS	0.030	-0.326**	0.132	0.452**	0.063
Alkalinity	**UMZ**	0.023	-0.591**	-0.002	0.371**	-0.273*
	EJS	-0.059	-0.564**	0.099	0.209	-0.235*
Coliforms	**UMZ**	0.151	0.364**	0.038	-0.085	0.292**
	EJS	0.124	0.177	0.072	-0.129	-0.003
Colour	**UMZ**	0.096	0.693**	0.204	-0.104	0.393**
	EJS	0.408**	0.705**	0.215*	-0.208	0.582**
Conductivity	**UMZ**	0.237*	-0.332**	0.264*	0.182	0.149
	EJS	0.003	-0.551**	0.004	0.463**	-0.118
Escherichia coli	**UMZ**	0.028	0.168	0.031	-0.089	0.041
	EJS	-0.166	0.218*	0.028	-0.166	-0.073
Fe	**UMZ**	0.061	0.608**	0.035	-0.097	0.261*
	EJS	0.185	0.687**	0.074	-0.281**	0.376**
Hardness	**UMZ**	-0.085	-0.269*	0.063	0.132	-0.090
	EJS	-0.085	-0.436**	0.098	0.479**	-0.199
Mn	**UMZ**	0.285**	-0.053	0.317**	0.198	0.108
	EJS	0.165	0.247*	0.080	-0.121	0.128
pH	**UMZ**	-0.099	-0.280**	0.067	-0.090	-0.204
	EJS	-0.052	-0.161	0.003	0.147	-0.079
SS	**UMZ**	-0.092	0.398**	0.091	-0.142	0.029
	EJS	0.041	0.637**	0.162	-0.082	0.282**
Temperature	**UMZ**	0.418**	0.348**	0.036	-0.050	0.144
	EJS	0.439**	0.292**	0.049	-0.079	0.086
TOC	**UMZ**	0.100	0.672**	0.313**	0.069	0.268*
	EJS	0.328**	0.479**	0.181	-0.241*	0.394**
Turbidity	**EJS**	-0.019	0.357**	-0.007	-0.098	-0.041
	UMZ	-0.018	0.361**	0.029	-0.067	0.159
***. Correlation is significant at the 0.05 level (2-tailed).**						
****. Correlation is significant at the 0.01 level (2-tailed).**						

The temporal analysis shows that chlorine dosage had seasonal variations. Table 5.20 reveals that chlorine has a positive correlation with colour, electrical conductivity,

manganese, temperature and total organic carbon. These are the key parameters affecting the dosage of chlorine at the UWTP.

Like chlorine, the polymer dosage also showed seasonal variations. Table 5.20 shows that the dosage had correlation with almost all the parameters considered in this study. That means that, in terms of polymer dosage, the key raw water parameters affecting treatment are algal Count, alkalinity, coliforms, colour, conductivity, Escherichia coli, iron, Hardness, manganese, pH, suspended solids, temperature, total organic carbon and turbidity.

The literature notes that lime is mainly used for the control of pH. In this study, the temporal analysis showed irregular variations and, unexpectedly, it does not show a good correlation with the pH at the two sources of abstraction. However, Table 5.20 shows that at the UWTP, lime dosage has a positive correlation with colour, electrical conductivity, manganese and total organic carbon. According to the chemical dosage data from the treatment plant, these are the key raw water quality parameters affecting treatment.

Potassium permanganate shows positive correlation with alkalinity, total coliforms, colour, iron, suspended solids and total organic carbon. According to oxidising agent's dosage data, these are the key raw water quality parameters affecting water treatment at the plant.

Activated carbon can be used to effectively control the organic content of the water. It can also be used to improve the taste and odour. Table 5.20 shows that its dosage has correlation with algal Count, alkalinity, electrical conductivity and iron. These are the key parameters in determining the amount of activated carbon to be added during water treatment.

5.4 Results for Objective 2

The second objective was to:

Design a genetic algorithm for the potable water treatment process control.

For this objective water quality data for the period 2006 to 2013 were used. The data were used to develop models for the prediction of lime dosage, polymer dosage, chlorine dosage

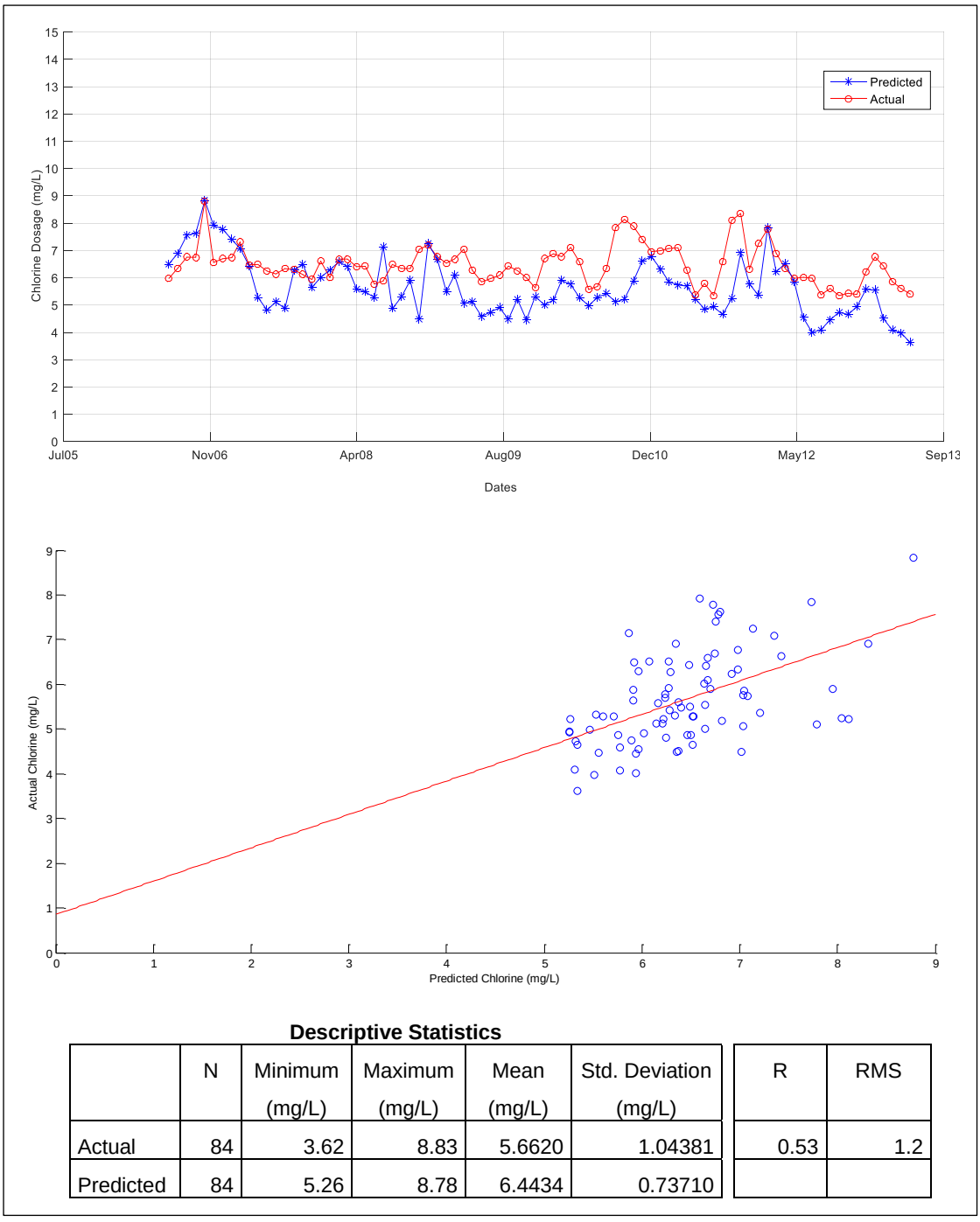
and potassium permanganate dosage. The models were developed using Matlab (MathWorks 2014) and the code used is shown in Appendix 4. These models were combined to develop a water treatment cost prediction model. The dosage prediction models were also used to develop a chemical dosage calculator that could replace the time-consuming jar tests. Jar tests are pilot-scale tests of the treatment chemicals used in a particular water plant. The process involved practically simulates the coagulation/flocculation process in a water treatment plant and helps operators calculate the right amount of chemicals to dose, and, thus improve the effectiveness of the water treatment (Arnold 2008). In the model results, R represents the linear correlation between predicted and actual chemical dosage data. The same statistical parameter was used by Mirsepassi (2004) and Chen and Sutcliffe (2012). According to Foster, Diamond and Jefferies (2014), when a predicted value of a model is close to the actual dosage, the R value will be close to 1. The RMS is the root mean square error between the predicted and actual chemical dosages data. For the RMS, Chai and Draxler (2014) supported its use in analysing models although other researchers such as Willmott, Matsuura and Robeson (2009) did not support its use, citing ambiguity as the cause. According to Kelley and Lai (2011) the RMS value is relative to the scale of data being analysed. In most cases it should be 5% of the maximum and 20% of the minimum. Exceptions can be found where the expected RMS value for a distribution is higher than the 5% and 20% limits. In this study, both R and RMS were used to describe the model fit.

5.4.1 Chlorine dosage prediction model

The prediction model was in the form shown in equation 4.14. The model produced a curve shown in Figure 5.36. The curve fit using actual data for the period July 2006 to June 2013 gave an R value of 0.53 and RMS value of 1.2. A good fit required that the RMS range falls within the range 0.44 to 0.8, which is 5% of the maximum and 20% of the minimum values within the data set. For this model, the RMS value was higher than expected meaning that it was not a good fit.

In theory chlorine is used for the purpose of disinfection, which means that the amount of chlorine dosed should have a correlation with bacteriological counts of the clarified water. Yet for this chlorine dosage prediction model *E.coli* and coliforms (represent bacteriological count), had a weak correlation with chlorine dosage (Table 5.20). The

parameters that showed significant correlations with chlorine dosage were colour, conductivity, manganese, temperature and total organic carbon. These do not directly relate to bacteriological counts, except for the total organic carbon. However, the presence of temperature in the model would confirm what was reported in Veenstra and Schnoor (1980). The researchers noted that chlorine can be lost during summer due to radiation, thus the need to increase chlorine dosage. This was also observed in later research (Marinovich *et al.* 2012; Wulff, Inniss and Clevenger 2013).



Descriptive Statistics

	N	Minimum (mg/L)	Maximum (mg/L)	Mean (mg/L)	Std. Deviation (mg/L)	R	RMS
Actual	84	3.62	8.83	5.6620	1.04381	0.53	1.2
Predicted	84	5.26	8.78	6.4434	0.73710		

Figure 5.36: Chlorine dosage prediction model

After solving for the constants, the model for determining chlorine dosage was as shown in Equation 5.1:

$$D_1 = 0.0001C_2 + 0.0665CO_2 + 0.00005M_1 + 0.0619T_1 + 0.0402T_2 + 0.2652TC_2 + 0.3270 \quad \text{Equation 5.1}$$

From Figure 5.36 it can be shown that the curve for the predicted data closely follows that of the actual, which indicates that the methodology used in prediction is relatively good according to R. However, the RMS value was higher than expected, which indicates that overall, the model had a bad fit. Thus focus should be given to the method that is currently being used to determine chlorine dosage at the UWTP. The current method involves adjusting the rotameter on the chlorinator until the final water has a residual chlorine level within the quality specifications of 1.5 mg/L to 2 mg/L. The main problem with this method is that the chlorine dose is set using the final water quality yet the raw water quality should be the major drive. It is therefore suggested that a proper method would enable the model to incorporate raw water quality parameters that are directly related to bacteriological count and hence produce a good fit. This is shown in a research by Abdullahi and Abdulkarim (2010), who developed a model that included bacteriological count as one of the parameters determining chlorine dosage.

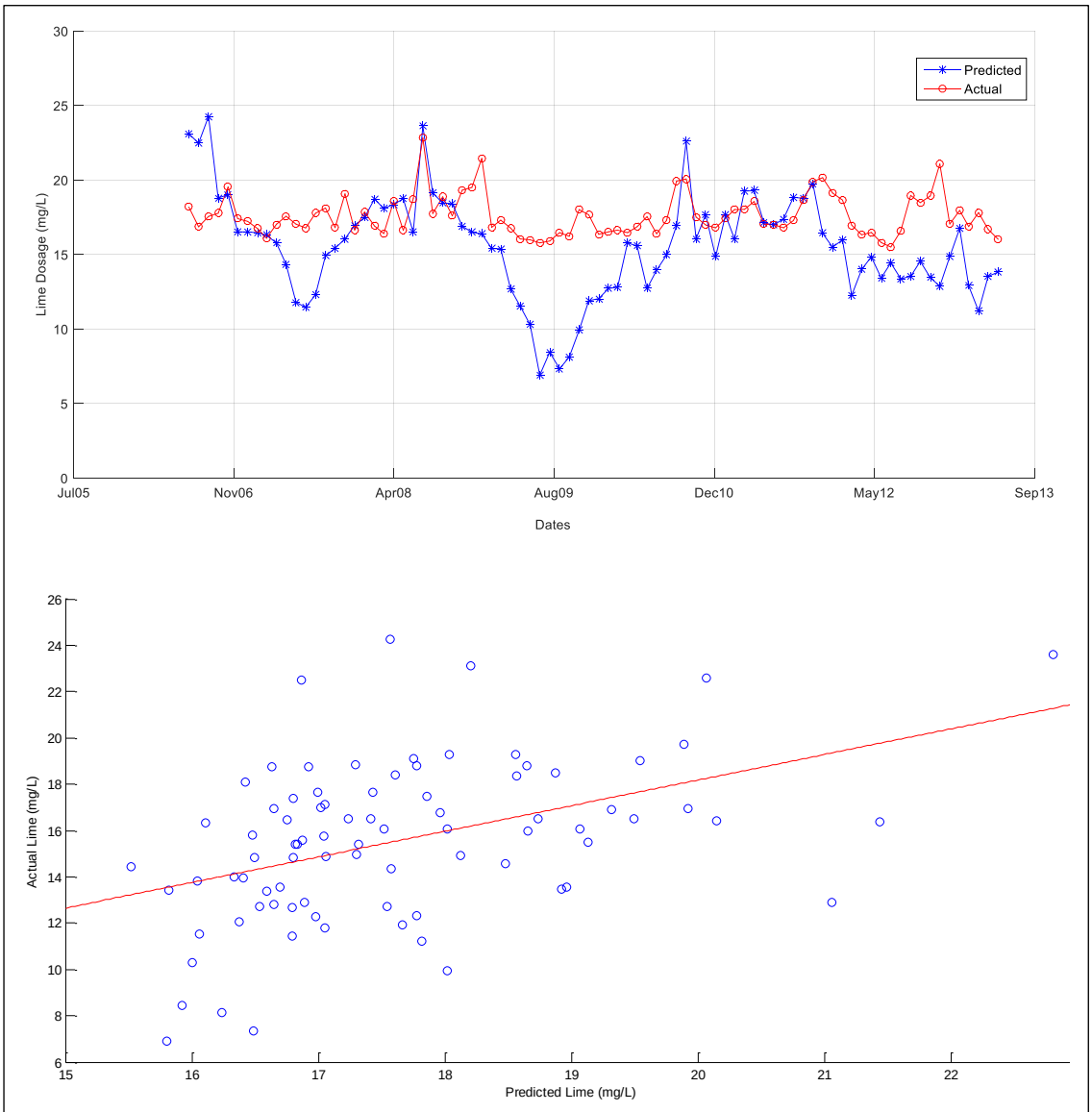
5.4.2 Lime dosage prediction model

The prediction model was in the form shown in equation 4.20. Figure 5.37 shows the results for the lime dosage prediction model. The curve fit, using actual data for the period July 2006 to June 2013 gave an R value of 0.44 and an RMS value of 3.72. A good fit required that the R value be close to 1 while the RMS range could be from 1.2 to 1.4, which is 5% of the maximum and 20% of the minimum values within the data set. For this model, the RMS value was high while the R value was low, meaning that the model had a bad fit.

In theory lime is added to control pH, hence pH and alkalinity are the parameters expected to show a good correlation with lime dosage (Abdullahi *et al.* 2012). In this study this was not the case because parameters that showed a good correlation with lime were colour, conductivity, manganese and total organic carbon (Table 5.20). According to Ersoz and Barrott (2012), an increase in pH should have resulted in a decrease in lime dosage.

A graphical comparison of actual dosage with predicted dosage for the same period showed that the model could predict fluctuations closely. However, it could not predict the peaks (Figure 5.37). This could have resulted in the observed low correlation coefficient. The major concern regarding this model was the unexpected result where pH and alkalinity showed no correlation with actual lime dosage. Similar to the chlorine dosage prediction

model, the lime dosage model could be improved by correcting the lime dosage determination method that is being used at the plant.



Descriptive Statistics

	N	Minimum (mg/L)	Maximum (mg/L)	Mean (mg/L)	Std. Deviation (mg/L)	R	RMS
Actual	84	6.91	24.27	15.5422	3.44570	0.44	3.72
Predicted	84	15.52	22.81	17.6168	1.36407		

Figure 5.37: Lime dosage prediction model

After solving for the constants the model for determining lime dosage was as shown in Equation 5.2:

$$D_2 = 0C_2 + 0.0001CO_1 + 3.5673M_1 + 0.5966TC_2 + 11.1186 \quad \text{Equation 5.2}$$

Currently, a wheelbarrow is used to transport the lime from the storage room to the dosing point where it is used to fill the hoper every morning. The hoper then feeds the lime into a channel that transports lime to the balancing tank. The amount of lime used is determined from the number of bags that has been utilised that day. The rate at which the lime is fed into the channel is determined by the pH meter, which is calibrated by an artisan every week. A major problem with this pH-control method is that when there is a sudden decrease or rise in pH during the day or week it takes time to adjust the dosage calibration accordingly. Use of a model would enable automatic calibration for the dosage.

5.4.3 Potassium permanganate dosage prediction model

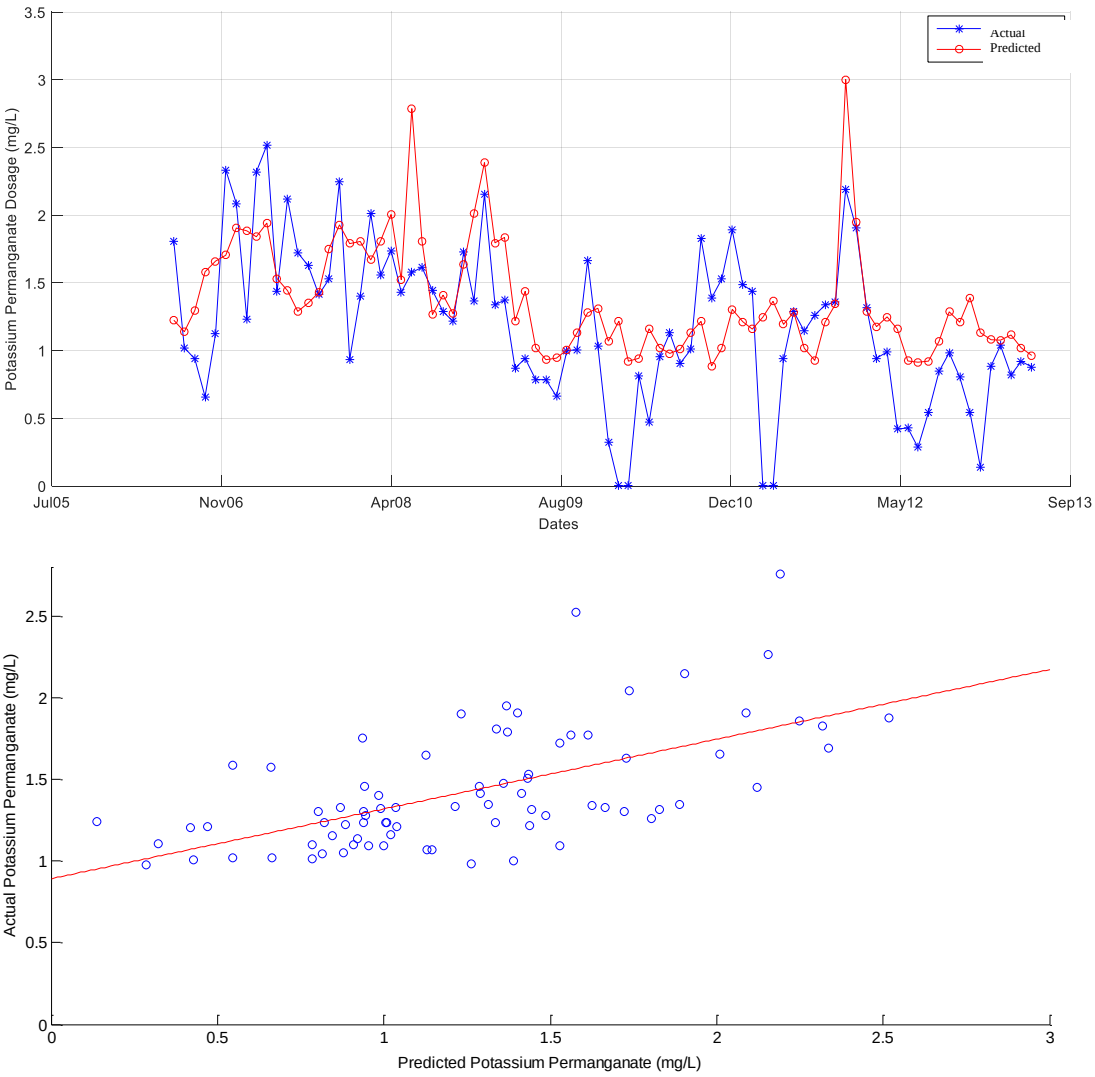
The prediction model was in the form shown in equation 4.20. Figure 5.38 shows the results for potassium permanganate dosage prediction model.

The curve fit using actual data for the period July 2006 to June 2013 gave an R value of 0.59 and RMS value of 0.51. A good fit required that the RMS range falls within the range 0 to 0.2, which is 5% of the maximum and 20% of the minimum values within the data set. For this model, the R value is relatively good. For the RMS value a comparison should be done with the actual data's standard deviation. The RMS value is close to the standard deviation of the actual data qualifying this model to be a good fit. When a model is good an RMS value should be close to the standard deviation of the actual data (Hyndman and Koehler 2006).

In theory potassium permanganate is used as an oxidant during water treatment (Willhite *et al.* 2012; Bing *et al.* 2013; Adcock *et al.* 2014). The parameters that were expected to be correlated to potassium permanganate dosage include coliforms, manganese and iron. These were included in the potassium permanganate dosage prediction model for this study.

The results for this model show that it can be used as a reliable tool for predicting potassium permanganate dosage. However, in Figure 5.38 it is shown that there are cases when the actual dosage value recorded was zero, yet during those periods parameters correlated to potassium permanganate were observed. This can be attributed to recording errors at the

water treatment plant. A suggested solution for this would be to automate the recording by including electronic instruments that can measure and record the raw water quality parameters.



Descriptive Statistics

	N	Minimum (mg/L)	Maximum (mg/L)	Mean (mg/L)	Std. Deviation (mg/L)	R	RMS
Actual	84	.00	2.52	1.1958	.57603	0.59	0.51
Predicted	84	.98	2.76	1.4200	.35851		

Figure 5.38: Prediction model for potassium permanganate dosage

After solving for the constants the potassium permanganate prediction model was as shown in Equation 5.3:

$$D_3 = 0.0022Alk_1 + 0.0018Alk_2 + 0.0002C_1 + 0.000004C_2 + 0.00005Col_1 + 0.0676Fe_1 + 0.0003Fe_2 + 0.0020SS_2 + 0.00002TC_1 + 0.0008TC_2 + 0.0043$$

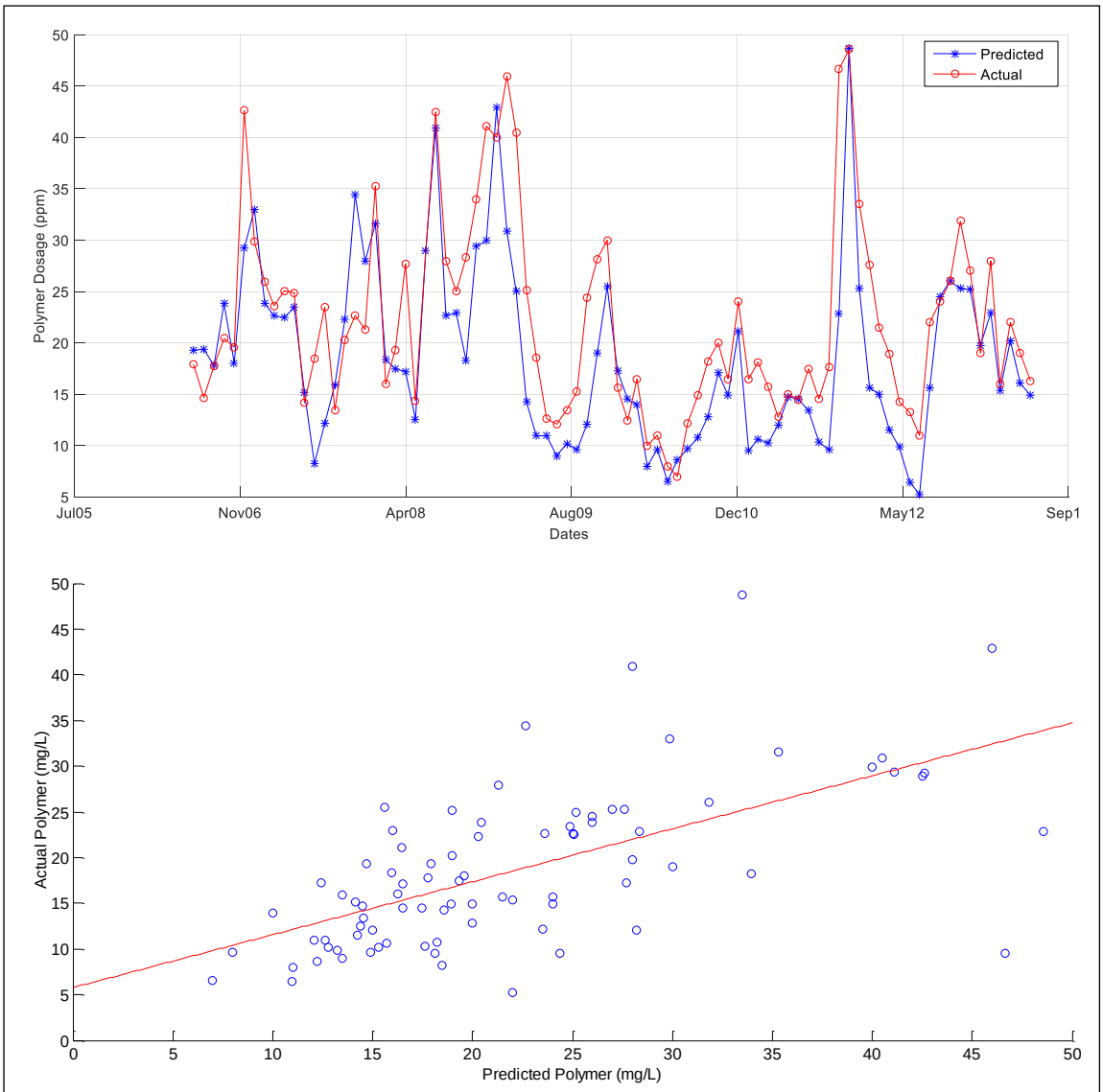
Equation 5.3

Currently, 10 kg of potassium permanganate is added to 100 L of water to prepare a solution that would be used to treat water in the balancing tank. The solution is added into the balancing tank via a pump. To calculate the dosage, the drop test is used. The tests require some mathematical calculations to be carried out. From the key raw water quality parameters affecting potassium permanganate dosage, the method being used at the plant effectively determines the dosage. However, using the model developed in this study will enable the treatment plant to save on the time and expertise required to carry out the drop test.

5.4.4 Polymer dosage prediction model

The prediction model was in the form shown in equation 4.20. Figure 5.39 shows the results of the polymer dosage prediction model. This model included all the parameters under investigation in this study. This is because the correlation showed that polymer dosage was dependent on all the parameters. Figure 5.39 also shows that the predicted dosage data follows the actual dosage very closely. The R value for this model was 0.64 and the RMS value of 8.5. A good fit required that the RMS range falls within the range 3.5 to 9, which is 5% of the maximum and 20% of the minimum values within the data set. This represented a good fit for the model and makes it a reliable tool to be used in predicting polymer dosage.

In theory polymer is used as a coagulant during water treatment (Devrimci, Yuksel and Sanin 2012; Sun *et al.* 2012; Bodlund 2013; Ramavandi 2014). The main parameters that were expected to be correlated to polymer dosage include turbidity, bacteriological count, algae, suspended solids and total organic carbon. These were all included in the polymer dosage prediction model for this study.



Descriptive Statistics

	N	Minimum (mg/L)	Maximum (mg/L)	Mean (mg/L)	Std. Deviation (mg/L)	R	RMS
Actual	84	5.26	48.70	18.5230	8.59099	0.64	8.5
Predicted	84	7.00	48.54	22.0101	9.31446		

Figure 5.39: Polymer dosage prediction model

This model had a good positive correlation value and the equation defining the model is as follows:

$$D_4 = 0.0001AC_1 + 0.0001AC_2 + 0.0001Alk_1 + 0.00003Alk_2 + 0.0000001Col_1 + 0.00004C_1 + 0.1884C_2 + 0.000004CO_1 + 0.0250CO_2 +$$

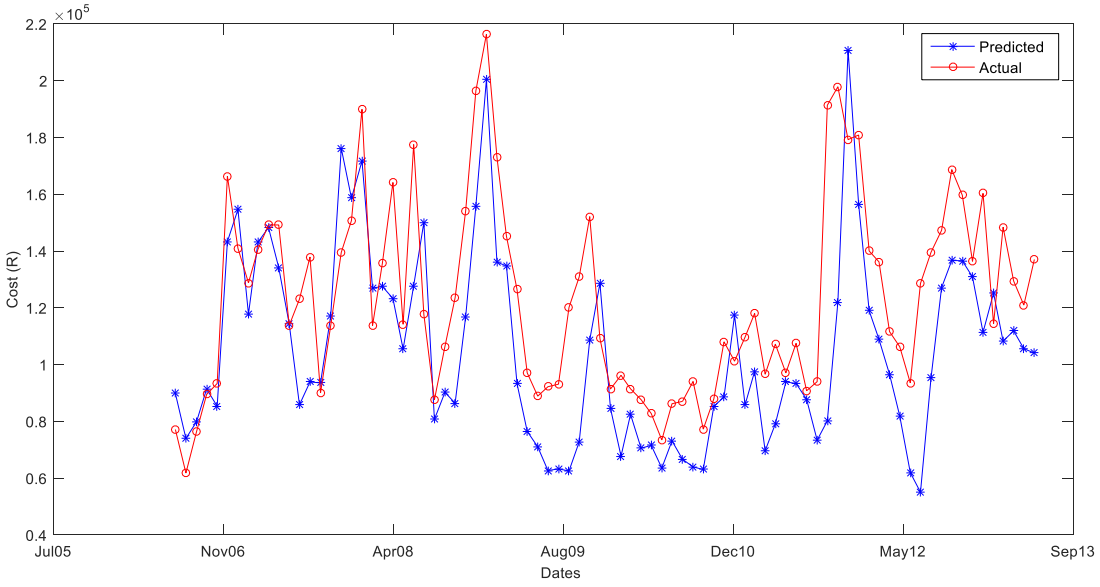
$$0.00001Eco_2 + 0.1144Fe_1 + 0.0204Fe_2 + 0.0105H_1 + 0.0175H_2 + 0.0424M_2 + 0.0001pH_1 + 0.2899S_1 + 0.0289S_2 + 0.0524T_1 + 0.0654T_2 + 0.00003TC_1 + 0.00004TC_2 + 0.0318Tu_1 + 0.1608Tu_2 + 0.00125$$

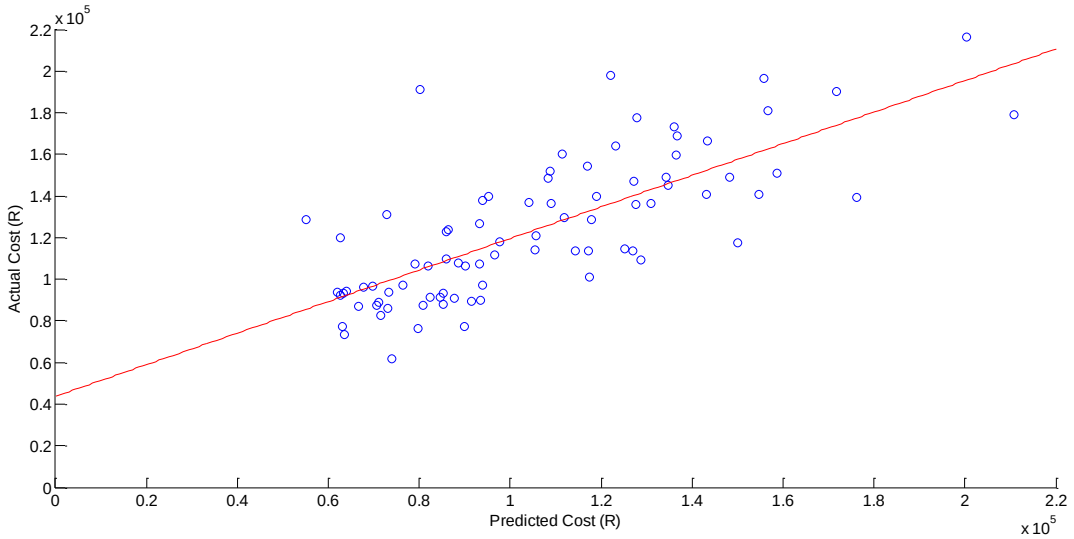
Equation 5.4

Currently the treatment plant uses the jar test to determine the optimum dosage of the polymer. The drop test is then used to calculate the amount of polymer dosed. This is a time and mathematically demanding process. Use of the polymer dosage prediction model would help reduce the time and eliminate the need for the calculations involved.

5.4.5 Cost prediction model

The water treatment cost prediction model was developed by substituting the results from the chemical dosage prediction models into equation 4.5. This was therefore a combination of the results obtained from the chemical dosage models to produce a cost function and hence would represent the overall optima for the models. The R value for this model was 0.75 and its RMS value was low relative to the actual data range. This shows a good fit as depicted on the graphs in Figure 5.40.





Descriptive Statistics

	N	Minimum (mg/L)	Maximum (mg/L)	Mean (mg/L)	Std. Deviation (mg/L)	R	RMS
Actual	84	61814.25	216371.19	123600.06	33877.42	0.75	30055.19
Predicted	84	55060.24	210718.26	105243.45	33351.15		

Figure 5.40: Chemical Cost prediction model

The combined model produced a good correlation with the actual cost data. The general equation of the model is in equation 4.5 and is also shown below.

$$CC = D_1P_1 + D_2P_2 + D_3P_3 + D_4P_4 + D_5P_5 \qquad \text{Equation 4.5}$$

The values for D_1, D_2 , D_3 and D_4 are from the solutions for the chemical dosage models. D_5 is removed from the cost function due to insufficient activated carbon dosage data.

5.5 Results for Objective 3

The third objective was to:

Evaluate the Distribution System's Non-Revenue Water.

5.5.1 Percentage of Non-Revenue Water in the System Input Volume

Non-Revenue Water in the Umzinto Water Distribution System makes up 28.1% of system input volume, amounting to 5017 kL/day. In Table 5.21, the total produced water is shown and adjusted for the period July 2013 to June 2014. The NRW and billed water are also shown. Figure 5.41 illustrates the percentage of NRW in the System Input Volume.

Table 5.21: Volume of Non-Revenue Water in the Umzinto Water Distribution System

Produced Water (kL/year)		**Total Billed Water (kL/year)**		**Non-Revenue Water**	
Measured	6 193 651	Metered	4 285 185	kL/year	1 831 060
Adjustment for Meter Accuracies	325 981	Unmetered	403 387	kL/day	5 017
Total	6 519 632	Total	4 688 572	% of system input volume	28.1

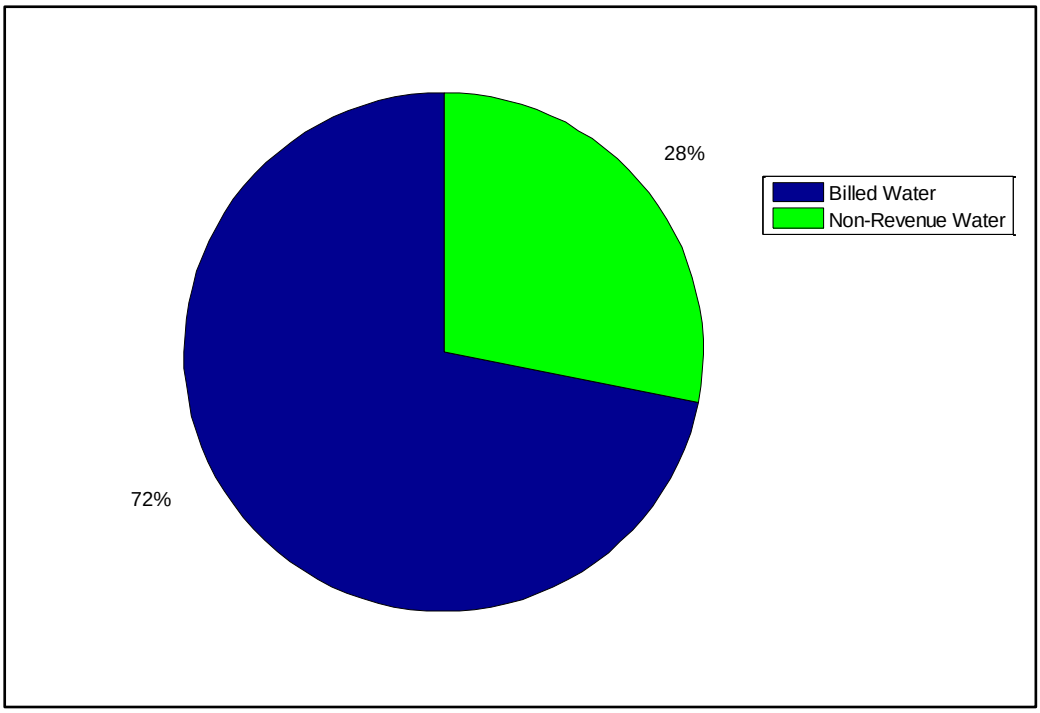


Figure 5.41: Percentage of Non-Revenue Water in the System Input Volume

5.5.2 Components of authorised consumption

Authorised consumption consisted of metered and unmetered, and billed and unbilled consumption. Table 5.22 shows that for 2014, the total volume of unbilled authorised consumption was 68364 kL/year which made up 1% of the system input volume. Figure 5.42 shows the components of consumption in the Umzinto Water Distribution System.

Table 5.22: Unbilled authorised consumption (metered and unmetered volumes)

Metered (kL/Year)		**Unmetered (kL/Year)**		**Total**	
Billing Database Shortfall	65 292	Municipal use, Fire water, and Flushing of Pipes and Reservoirs	3 072	kL/year	68 364
				kL/day	187
				% of System Input Volume	1

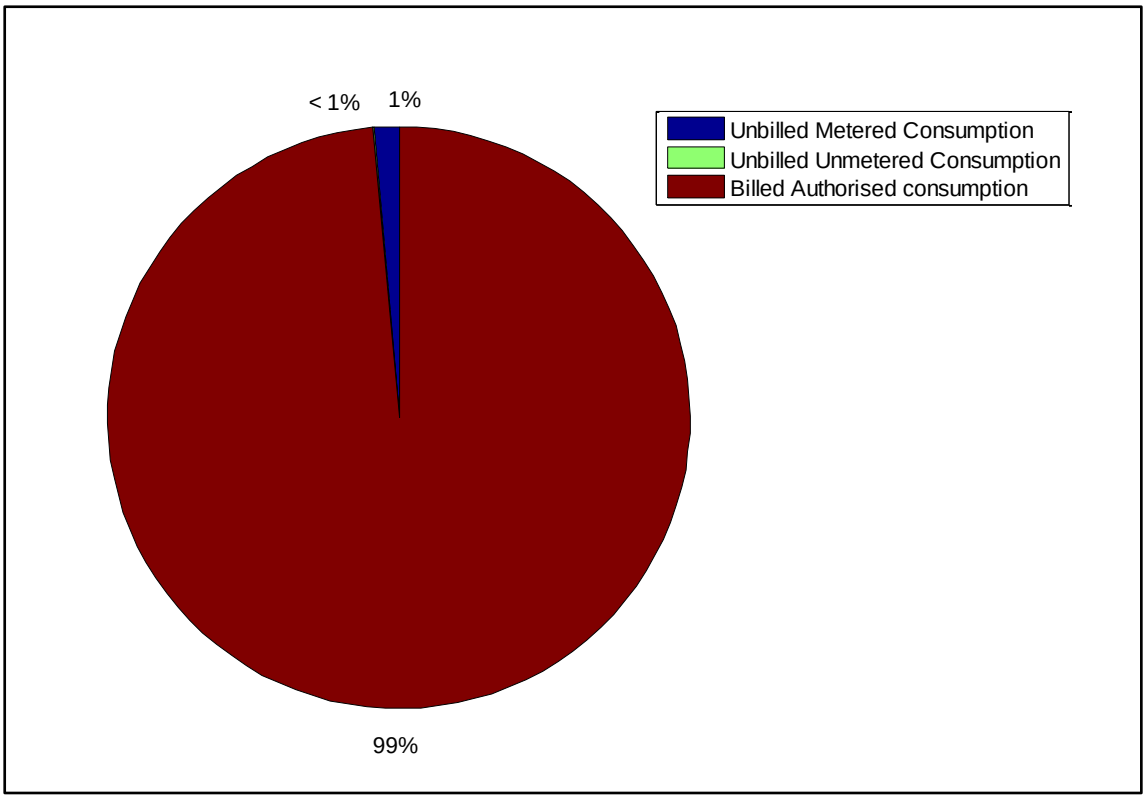


Figure 5.42: Components of Authorised Consumption

5.5.3 Apparent Losses

The apparent losses were determined using equation 4.60. Table 5.23 shows their percentage in the system input volume.

Table 5.23: Apparent Losses for the Umzinto Water Distribution System

kL/year	580 598
kL/day	1 591
% of System input Volume	9

5.5.4 Real Losses

Real losses were determined using recorded data for mains and distribution leaks, reservoir overflows and service connection leaks. Table 5.24 show the real losses in the distribution system and their percentage value in the system input volume.

Table 5.24: Real Losses for the Umzinto Water Distribution System

kL/year	1 182 098
kL/day	3 238
% of System input Volume	18

5.5.5 Components of Non-Revenue Water

Table 5.25 shows the System Input Volume components which are Billed Water and NRW. The NRW components are also show as a percentage of the System Input Volume. Figure 5.43 shows that Real Losses are the dominant component of NRW, constituting 65% of NRW and 18% of the System Input Volume. Apparent Losses make up 32% of NRW and are 9% of the System Input Volume. Unbilled authorised consumption is 1% of the System Input Volume.

Table 5.25: Components of Non-Revenue Water in the Umzinto Water Distribution System

System Input Volume (kL/day)	1 .07862
Billed Water (%)	72
Non-Revenue Water (%)	28
Apparent Loss (%)	9
Real Losses (%)	18
Unbilled Authorised Consumption (%)	1

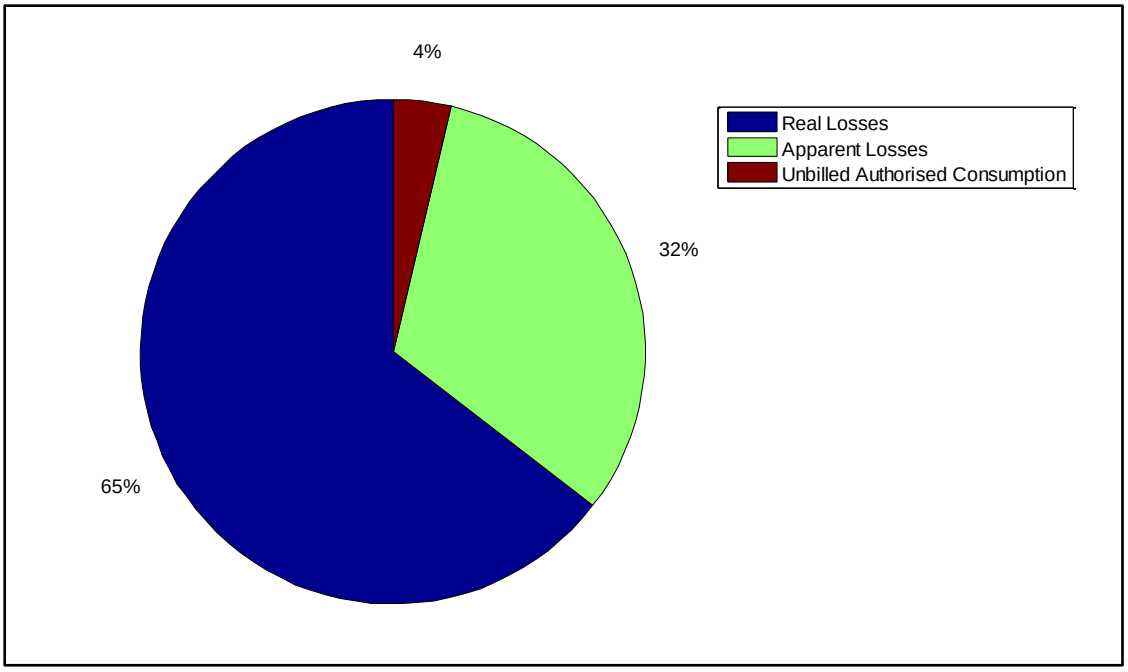


Figure 5.43: Components of Non-Revenue Water in the Umzinto Water Distribution System

5.5.6 Water balance for the Umzinto Distribution System

Table 5.26 illustrates the final results of the NRW assessment for the Umzinto Distribution System in kL/year. The Water Balance is different from the standard South African water balance in the sense that Free Basic water is taken as a component of billed authorised consumption.

Table 5.26: Water Balance for the Umzinto Water Distribution System in kL/year

<table>
<tr><td rowspan="8">System Input Volume
6 519 632 kl/year</td><td rowspan="4">Authorised Consumption
4 756 936 kl/year</td><td rowspan="3">Billed Authorised Consumption
4 688 572 kl/year</td><td>Exported Water
71 547 kl/year</td><td rowspan="3">Potential Revenue Water
4 688 572 kl/month</td></tr>
<tr><td>Billed Metered Consumption
4 285 185 kl/year</td></tr>
<tr><td>Free Basic Water (Standpipes)
331 840 kl/year</td></tr>
<tr><td rowspan="2">Unbilled Authorised Consumption
68 364 kl/year</td><td>Unbilled Unmetered Consumption
3 072 kl/year</td><td rowspan="5">Non-Revenue Water
1 831 060 kl/year</td></tr>
<tr><td>Unbilled Metered Consumption
65 292 kl/year</td></tr>
<tr><td rowspan="4">Water Losses
1 762 696 kl/year</td><td rowspan="2">Apparent Losses
580 598 kl/year</td><td>Unauthorised Consumption
430 616 kl/year</td></tr>
<tr><td>Metering Inaccuracies
149 981 kl/year</td></tr>
<tr><td rowspan="3">Real Losses
1 182 098 kl/year</td><td>Mains and Distribution Leaks
969 321 kl/year</td></tr>
<tr><td>Reservoir Overflows
5 910 kl/year</td></tr>
<tr><td>Service Connection Leaks
206 867 kl/year</td></tr>
</table>

5.6 Results for Objective 4

The fourth objective was to:

Develop a model for water usage analysis.

The model was in the form of application software that allowed the user to input data on water usage, losses and billing. This data were stored in a relational database and was accessed using the analysis part of the software. The analysis enabled a view of the following:

- Standard South African Water Balance components
- Circular Water Balance components

- Authorised Consumption components
- Water Loss components
- Non-Revenue Water Components

5.6.1 Standard Water Balance

Figure 5.44 shows the default analysis screen of the water usage analysis model developed. This shows a Standard South African Water Balance. On this screen a date can be selected and this enables the model to calculate the water balance components for the year ending as at that date. The software also allows the user to change the units displayed by the water balance components. In Figure 5.44, the units displayed are kL/year, and can be changed to kL/day or the water balance can be represented in the form of percentages. The refresh button allows the recalculation of the water balance components. This is essential when the date has been changed.

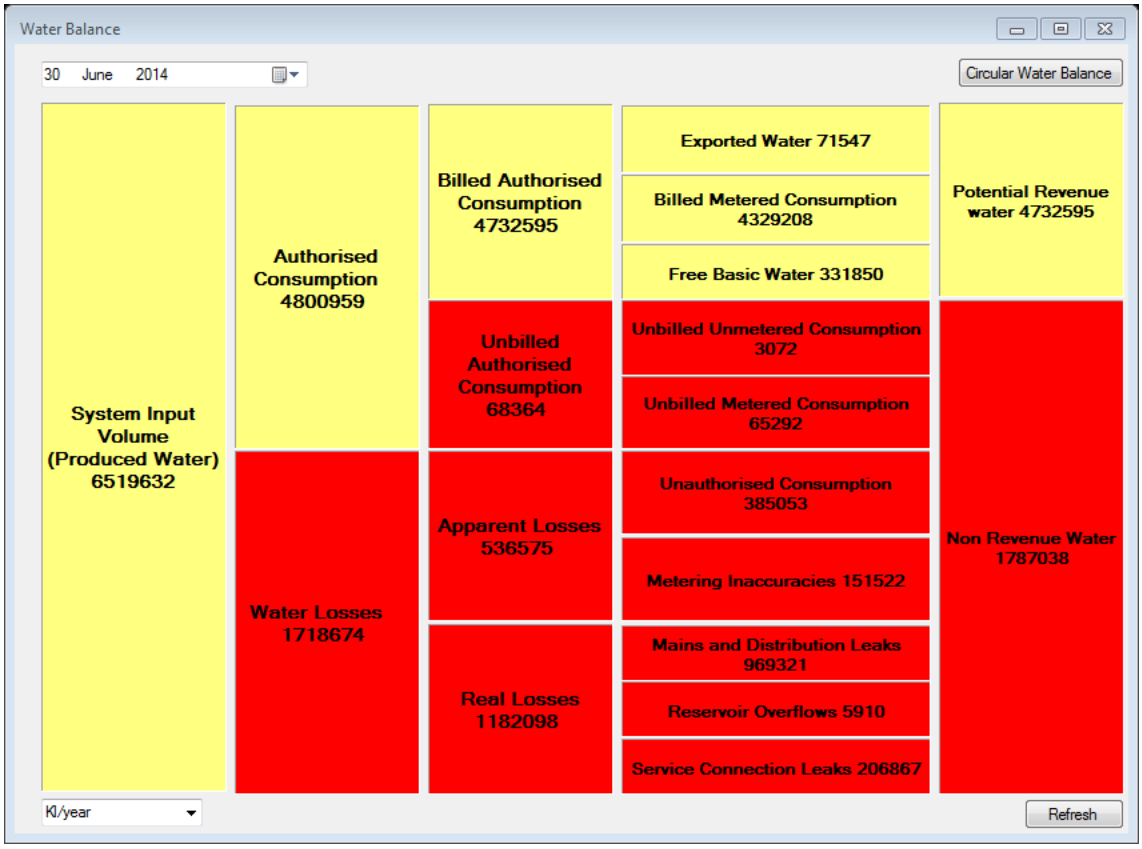


Figure 5.44: Graphic user interface for the Standard Water Balance

Further analysis is done by clicking on any of the components of the water balance. The Circular Water Balance is accessed by clicking on the button in the upper right hand corner of the screen.

5.6.2 Circular Water Balance

Figure 5.45 shows the graphic user interface of the Circular Water Balance in the water usage analysis model.

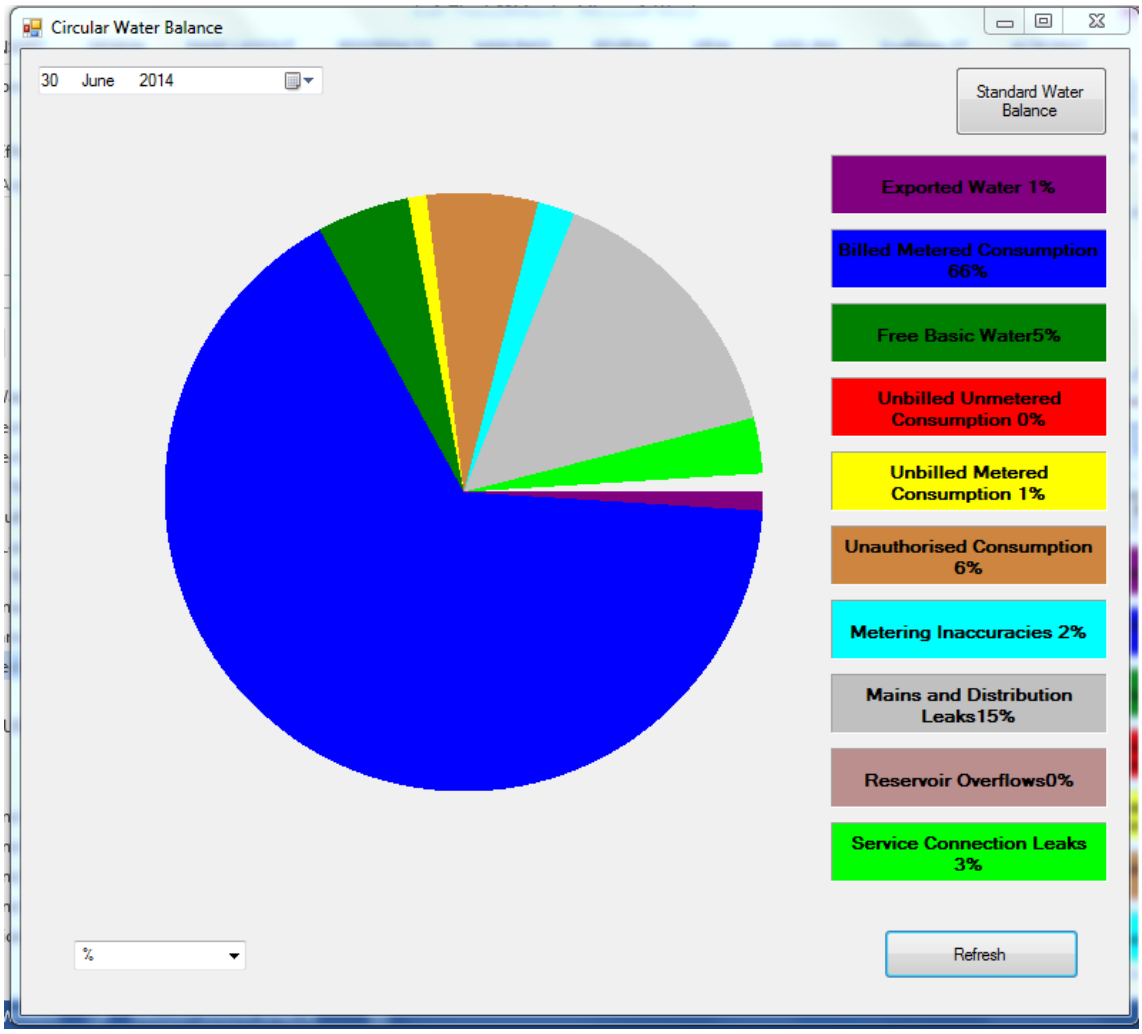


Figure 5.45: Graphic user interface for the Circular Water Balance

Like the Standard Water Balance in section 5.6.1, the Circular Water Balance offers the option of selecting a date and the water balance components are calculated for the year ending at that date. The refresh button enables the recalculation of the water balance components. The Standard Water Balance can easily be accessed by clicking on the button in the upper right-hand corner.

5.6.3 Authorised Consumption

Figure 5.46 shows a comparison of billed and unbilled authorised consumption. The screen is accessed by clicking on the Authorised Consumption label on the Standard Water Balance shown in Figure 5.44.

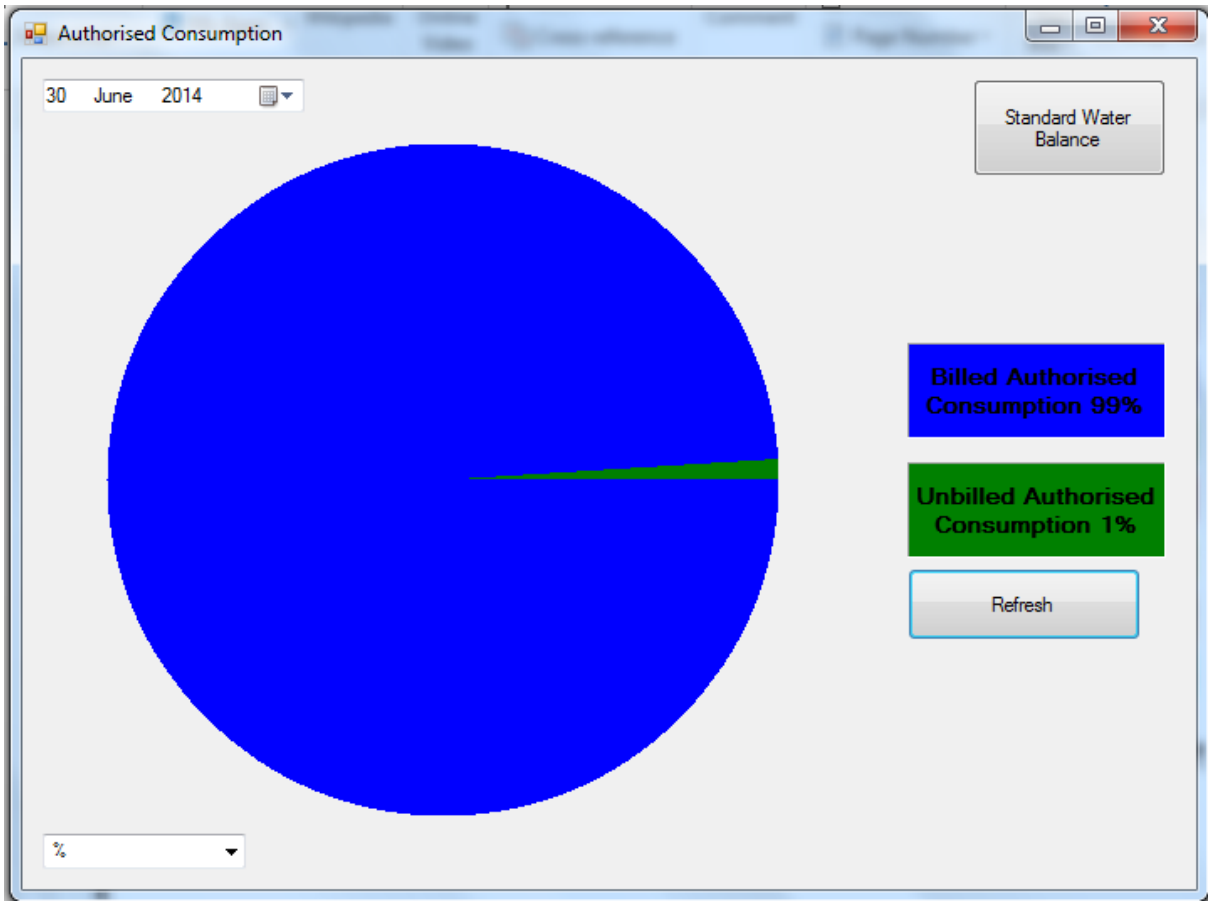


Figure 5.46: Graphic user interface for Authorised Consumption

In Figure 5.46 the Authorised Consumption components calculated were for the year ending 30 June 2014. The refresh button recalculates the components and access back to the Standard Water Balance screen is by clicking on the button in the upper right hand corner.

5.6.4 Water Losses

Figure 5.47 shows a comparison of Real Losses and Apparent Losses in a water distribution system. To access this screen, the user clicks on the Water Losses Label Shown in Figure 5.44.

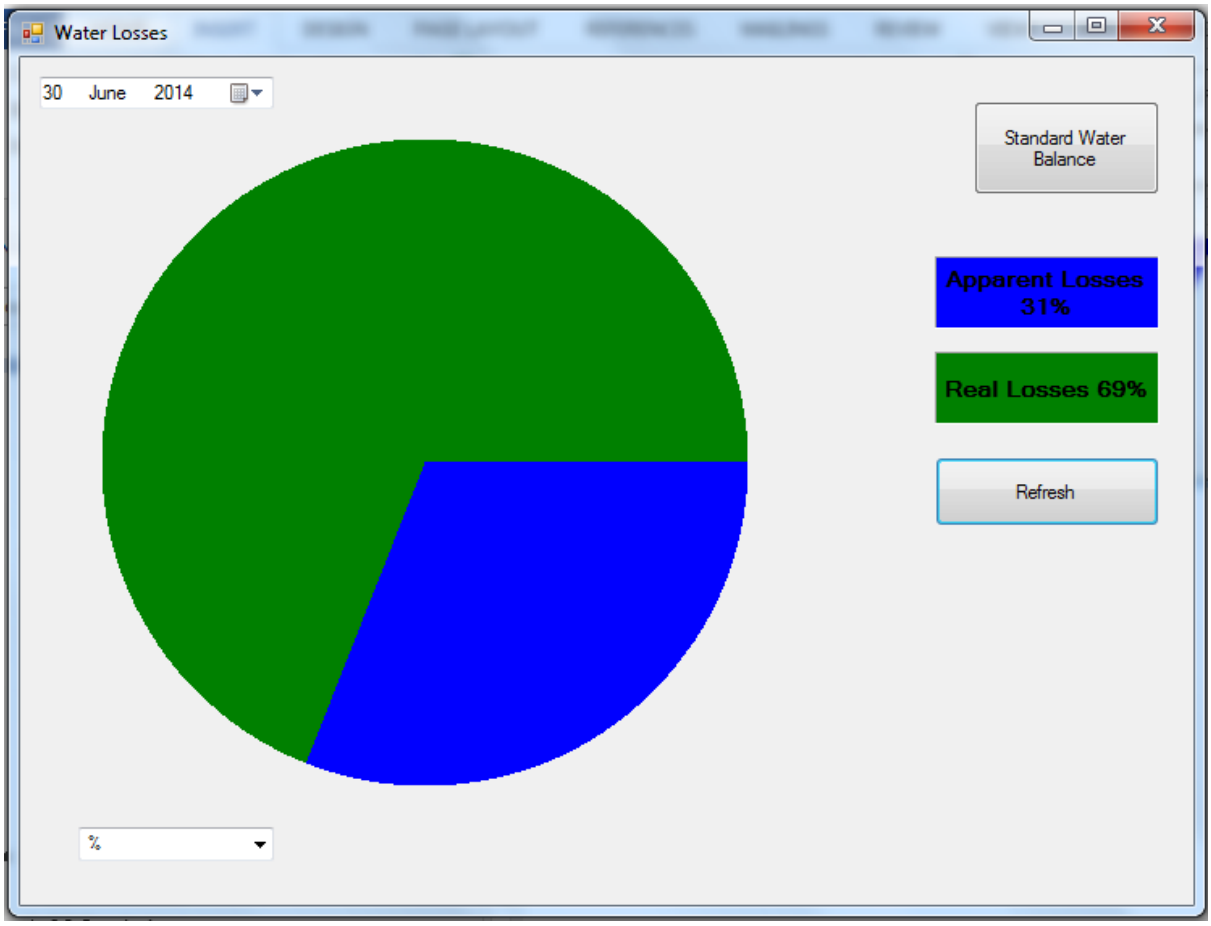


Figure 5.47: Graphic user interface for Water Losses

The remaining components of the Water Balance are shown in Appendix 5 and the code used to develop the complete software is shown in Appendix 6.

5.7 Summary

Statistical analysis of raw water quality parameters revealed the key raw water quality parameters affecting treatment at the UWTP. The trends showed that the EJ Smith Dam's water was of worse quality than that from the Umzinto River. This was clear for parameters such as *Escherichia coli*, colour, algal count, total coliforms and turbidity. The key parameters were then used to develop chemical dosage prediction models for polymer, lime, chlorine and potassium permanganate. The assessment of NRW in the UWTP revealed that Real Losses were the major component that had to be addressed. The level of NRW was calculated to be 28%.

CHAPTER 6 – CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

The first objective of this study was to:

Identify and quantify key raw water quality parameters affecting treatment at the Umzinto Water Treatment Plant.

The first step in determining the key parameters was to conduct a temporal analysis of raw water quality parameters and the chemical dosages at the UWTP. The study used descriptive statistics, trending and box plots. The box plots were constructed using a Microsoft Excel Adin developed during the data pre-processing stage and the trending was done using Matlab. This allowed for the comparison of raw water quality parameters from the two abstraction points of the UWTP. The analysis revealed that the EJ Smith Dam had poorer raw water quality compared to the Umzinto River. This was very clear when comparing the trends for algal count, *Escherichia coli*, suspended solids, total coliforms and turbidity.

The dosage trends for polymer and chlorine showed seasonal variations that correlated to health-related parameters. The lime, potassium permanganate and activated carbon dosage trends showed irregular variation. The selection of the key raw water quality parameters was greatly dependent on the chemical dosage being considered. The results showed that the key raw water quality parameters affecting chlorine dosage were colour, electrical conductivity, manganese, temperature and total organic carbon. These are the key parameters affecting the dosage of chlorine at the UWTP according to the historical data that was used. Concern is raised for the absence of parameters that are related to the bacteriological count, which is the main reason for chlorination.

According to historical data provided, the key raw water parameters affecting polymer dosage were algal count, alkalinity, coliforms, colour, conductivity, *Escherichia coli*, iron, hardness, manganese, pH, suspended solids, temperature, total organic carbon and turbidity. For lime dosage the key raw water quality parameters were colour, electrical

conductivity, manganese and total organic carbon. This is not what was expected, lime is added to control pH, so pH and alkalinity were expected to be among the key parameters. For the oxidising agent, potassium permanganate, the key raw water quality parameters were alkalinity, total coliforms, colour, iron, suspended solids and total organic carbon. The historical data for activated carbon suggested the key raw water quality parameters to be with algal count, alkalinity, electrical conductivity and iron. The frequency of observations however made it difficult to develop a reliable model for lime.

Considering the levels of algal count, *Escherichia coli*, suspended solids, total coliforms and turbidity, there was very high pollution emanating from communities and activities close to the raw water sources, especially the EJ Smith Dam.

The second objective was to:
Design a genetic algorithm for the potable water treatment process control.

This was important to ensure that drinking water is safe for consumption and to manage the efficiency of the water treatment plant in terms of chemical dosage. This study developed models for predicting chemical dosages for chlorine, lime, polymer and potassium permanganate. These models used the data obtained from the findings under objective 1, showing the relationship between chemical dosage and raw water quality. Using the four chemical dosage models, a treatment cost prediction model was also developed.

For the chlorine dosage prediction model the results showed an R value of 0.53 and RMS value of 1.2. A good fit required that the RMS range falls within the range 0.44 to 0.8, which is 5% of the maximum and 20% of the minimum values within the data set. For this model, the RMS value was higher than expected meaning that it was not a good fit. Concern in this model was on the absence of bacteriological count related parameters in the model. The UWTP should focus on implementing a method that allows chlorine dosage control relative to the raw water input quality. The current method involves adjusting the rotameter on the chlorinator until the final water has a residual chlorine level within the quality specifications of 1.5 mg/L to 2 mg/L. The main problem with this method is that the

chlorine dose is set using the final water quality which is not an optimal method of chlorine dosage.

The lime dosage prediction model had an R value of 0.44 and an RMS value of 3.72. A good fit required that the R value be close to 1 while the RMS range could be from 1.2 to 1.4, which is 5% of the maximum and 20% of the minimum values within the data set. For this model, the RMS value was high while the R value was low, meaning that the model had a bad fit. Again concern was on the parameter included in the model. Currently, the rate at which lime is fed into the balancing tank is determined by the pH meter, which is calibrated by an artisan every week. A major problem with this pH-control method is that when there is a sudden decrease or rise in pH during the day or week it takes time to adjust the dosage calibration accordingly. Use of a genetic algorithm model would enable automatic calibration for the dosage.

The potassium permanganate dosage prediction model had an R value of 0.59 and RMS value of 0.51. A good fit required that the RMS range falls within the range 0 to 0.2, which is 5% of the maximum and 20% of the minimum values within the data set. For this model, the R value is relatively good. For the RMS value a comparison should be done with the actual data's standard deviation. The RMS value is close to the standard deviation of the actual data qualifying this model to be a good fit. By using the method developed in this study, the treatment plant would benefit from the elimination of time and expertise required to carry out the drop test they are currently using.

The polymer dosage prediction model produced an R value of 0.64 and the RMS value was very low relative to the actual data range. This represented a good fit for the model and makes it a reliable tool to be used in predicting polymer dosage. Use of this dosage prediction model would help reduce the time and eliminate the need for the calculations involved when performing the jar tests. The model would also help the plant to include a variety of parameters related to coagulation and flocculation. Therefore, based on the afore mentioned arguments the models developed would help optimise water treatment by minimising the costs and time associated with chemical dosage at the treatment plant.

The cost prediction model produced an R value of 0.75 and its RMS value was very low relative to the actual data range. This represented a good fit for the model and makes it a reliable tool to be used in predicting dosage cost.

The third objective was to:

Assesse the Distribution System's Non-Revenue Water.

Evaluation of NRW is an integral phase in designing a management strategy to reduce and control water losses in a distribution system. This is important in ensuring that the greatest possible percentage of water treated by the treatment plant reaches the consumer and is converted into revenue. The process can be demanding and time consuming. To make it more efficient, there is need to prioritise components of NRW in a distribution system. This study assessed NRW components for the Umzinto Water Distribution System and suggested an approach which water utilities can use to break down NRW and draw up water balances efficiently.

The assessment of NRW for the Umzinto Water Distribution System showed that Real Losses are a dominant component of NRW. The components consisted of 65% Real Losses, 32% Apparent Losses and 3% Unbilled Authorised consumption.

The fourth objective was to:

Develop a model for water usage analysis.

This was achieved by designing a software application to evaluate NRW. This software determined all the important components of NRW that would take time to evaluate manually. It would also store historical data for the water distribution system and report on a rolling year basis. Implementation of this software would help minimise the errors associated with manual calculation of NRW and improve the availability of data for research and analysis.

6.2 Recommendations

The raw water quality parameters for the separate supplies are measured before they reach the balancing tank and this is used to calculate the dosage of chemicals in the balancing tank. This is prone to error as the supply from the two abstraction points can vary. There is a need to test the quality of the mixed water before the chemicals are dosed. This means that the treatment plant needs to add an additional buffer tank to monitor the combined raw water quality. This will also address the problems observed during modelling such as the lack of correlation between lime and pH or lime and alkalinity.

The analysis of NRW showed that Real Losses were a major challenge in the Umzinto Distribution System. There is need to develop a maintenance program to cater for leakage. Communities also need to be educated on the importance of reporting leakage in the network.

Further study is recommended on the development of chemical dosage models to automate drinking water treatment plants and water distribution systems. The methods used in this study relied on historical data from the water treatment plant and distribution system. A comparison of the models produced with laboratory experimental techniques would help to optimise models, which was beyond the current masters work.

REFERENCES

Abdullahi, M. E. and Abdulkarim, B. I. 2010. Development of a mathematical model for determining the quantity of chlorine required for water treatment. *Journal of Applied Sciences Research*, 6 (8): 1002-1007.

Abdullahi, M. E., Aloko, D., Baba, G. and Mohammed, J. 2012. Predictive model for lime dosage in water treatment plant. *International Journal of Scientific and Research Publications*, 2 (12): 1-5.

Abushams, I. 2001. Unaccounted for Water - Using the untapped resource to improve financial and technical performance in Jordan. *Water International*, 26 (4): 602-604.

Adcock, J. L., Barnett, N. W., Barrow, C. J. and Francis, P. S. 2014. Advances in the use of acidic potassium permanganate as a chemiluminescence reagent: a review. *Analytica Chimica Acta* 807: 9-28.

Adgar, A., Cox, C. and Böhme, T. 2000. Performance improvements at surface water treatment works using ANN-based automation schemes. *Chemical Engineering Research and Design*, 78 (7): 1026-1039.

Afshar, A., Shojaei, N. and Sagharjooghifarahani, M. 2013. Multiobjective calibration of reservoir water quality modeling using Multiobjective Particle Swarm Optimization (MOPSO). *Water Resources Management*, 27 (7): 1931-1947.

Alander, J. T. 1993. Indexed bibliography of most active researchers of genetic algorithms. *Artificial Intelligence*, 960: 961.

Almandoz, J., Cabrera, E., Arregui, F., Cabrera Jr, E. and Cobacho, R. 2005. Leakage assessment through water distribution network simulation. *Journal of Water Resources Planning and Management*, 131 (6): 458-466.

Altunkaynak, A., Özger, M. and Çakmakci, M. 2005. Water consumption prediction of Istanbul city by using fuzzy logic approach. *Water Resources Management*, 19 (5): 641-654.

Alvisi, S. and Franchini, M. 2006. Near-optimal rehabilitation scheduling of water distribution systems based on a multi-objective genetic algorithm. *Civil Engineering and Environmental Systems*, 23 (3): 143-160.

Alvisi, S. and Franchini, M. 2009. Multiobjective optimization of rehabilitation and leakage detection scheduling in water distribution systems. *Journal of Water Resources Planning and Management*, 135 (6): 426-439.

Andey, S. P. and Kelkar, P. S. 2009. Influence of intermittent and continuous modes of water supply on domestic water consumption. *Water Resources Management*, 23 (12): 2555-2566.

Arabi, M., Govindaraju, R. S. and Hantush, M. M. 2006. Cost-effective allocation of watershed management practices using a genetic algorithm. *Water Resources Research*, 42 (10): W10429.

Arnold, A. 2008. Evaluation and quantification of engineered flocs and drinking water treatability. Master of Applied Science, University of Waterloo.

Arregui, F., Cabrera, E., Cobacho, R. and Garcia-Serra, J. 2006. Reducing apparent losses caused by meters inaccuracies. *Water Practice & Technology*, 1 (04)

Ashbolt, N. J. 2004. Risk analysis of drinking water microbial contamination versus disinfection by-products (DBPs). *Toxicology*, 198 (1): 255-262.

Ashlock, D. 2006. *Evolutionary computation for modeling and optimization* New York: Springer Science & Business Media.

AWWA. 2009. *Water audits and loss control programs, (M36)*. Denver: American Water Works Association.

Bagajewicz, M. and Savelski, M. 2001. On the use of linear models for the design of water utilization systems in process plants with a single contaminant. *Chemical Engineering Research and Design*, 79 (5): 600-610.

Banga, J. R. and Casares, J. 1987. ICRS: Application to a wastewater treatment plant model. In: Proceedings of *IChemE Symposium Series*. Oxford, UK, Pergamon Press, 183-192.

Bartram, J. and Ballance, R. eds. 1996. *Water quality monitoring-A practical guide to the design and implementation of freshwater quality studies and monitoring programs*. London: Taylor & Francis.

Bates, D., Mächler, M., Bolker, B. and Walker, S. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*,

Bekele, E. and Knapp, H. V. 2010. Watershed modeling to assessing impacts of potential climate change on water supply availability. *Water Resources Management*, 24 (13): 3299-3320.

Béraud, B., Steyer, J. P., Lemoine, C., Latrille, E., Manic, G. and Printemps-Vacquier, C. 2007. Towards a global multi objective optimization of wastewater treatment plant based on modeling and genetic algorithms. *Water Science & Technology*, 56 (9): 109-116.

Beverly, R. P. 2011. *Filter troubleshooting and design handbook*. Denver: American Water Works Association.

Bhagwan, J., Wegelin, W., Mckenzie, R. and Wensley, A. 2014. Counting the lost drops: South Africa's study into Non-Revenue Water. *Water Practice & Technology*, 9 (4): 502-508.

Bing, Y., Yong, Y., Meiping, X., Liang, L. and Tao, L. 2013. Treatment of enhanced coagulation process of potassium permanganate pre-oxidation for raw water with High-Algae. *Water Purification Technology*, 2: 003.

Boccelli, D. L., Small, M. J. and Diwekar, U. M. 2007. Drinking water treatment plant design incorporating variability and uncertainty. *Journal of Environmental Engineering*, 133 (3): 303-312.

Bodlund, I. 2013. Coagulant protein from plant materials: Potential water treatment agent. Master of Science, Royal Institute of Technology.

Böhmelt, T., Bernauer, T., Buhaug, H., Gleditsch, N. P., Tribaldos, T. and Wischnath, G. 2013. Demand, supply, and restraint: Determinants of domestic water conflict and cooperation. *Global Environmental Change*, (0): 1-12.

Breese, S. 2006. Optimizing conventional water treatment plants. *AWWOA Annual Serminar* 15 March 2006.

Brinkoff, T. 2012. *City Population*. Available: http://www.citypopulation.de/php/southafrica-admin.php?adm2id=KZN212 (Accessed 18/11/2016).

Bryant, E. A., Fulton, G. P. and Budd, G. C. 1992. *Disinfection alternatives for safe drinking water*. New York: Van Nostrand Reinhold.

Buhaug, H. and Urdal, H. 2013. An urbanization bomb? Population growth and social disorder in cities. *Global Environmental Change*, 23 (1): 1-10.

Burn, D. H. and McBean, E. A. 1985. Optimization modeling of water quality in an uncertain environment. *Water Resources Research*, 21 (7): 934-940.

Carrasco, E., Rodrıguez, J., Punal, A., Roca, E. and Lema, J. 2004. Diagnosis of acidification states in an anaerobic wastewater treatment plant using a fuzzy-based expert system. *Control Engineering Practice*, 12 (1): 59-64.

Casey, T., Kearney, H. K., Kerr, H., Council, D. C. and Dublin, C. 2012. *The chlorine demand characteristics of Irish water supplies: Process design implications for disinfection and THM formation*: Ireland: CIWEM.

Castro, P., Teles, J. and Novais, A. 2009. Linear program-based algorithm for the optimal design of wastewater treatment systems. *Clean Technologies and Environmental Policy*, 11 (1): 83-93.

Chachuat, B., Roche, N. and Latifi, M. 2001. Dynamic optimisation of small size wastewater treatment plants including nitrification and denitrification processes. *Computers & Chemical Engineering*, 25 (4): 585-593.

Chai, T. and Draxler, R. R. 2014. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7 (3): 1247-1250.

Chang, S.-Y., Brill, E. D. and Hopkins, L. D. 1982. Use of mathematical models to generate alternative solutions to water resources planning problems. *Water Resources Research*, 18 (1): 58-64.

Charisiadis, P., Andra, S. S., Makris, K. C., Christophi, C. A., Skarlatos, D., Vamvakousis, V., Kargaki, S. and Stephanou, E. G. 2015. Spatial and seasonal variability of tap water disinfection by-products within distribution pipe networks. *Science of The Total Environment*, 506: 26-35.

Chen, F. and Sutcliffe, C. 2012. Pricing and hedging short sterling options using neural networks. *Intelligent Systems in Accounting, Finance and Management*, 19 (2): 128-149.

Chen, H. and Chang, N.-B. 1998. Water pollution control in the river basin by fuzzy genetic algorithm-based multiobjective programming modeling. *Water Science and Technology*, 37 (8): 55-63.

Chong, M. N., Jin, B., Chow, C. W. K. and Saint, C. 2010. Recent developments in photocatalytic water treatment technology: A review. *Water Research*, 44 (10): 2997-3027.

Chowdhury, S., Rodriguez, M. J. and Sadiq, R. 2011. Disinfection byproducts in Canadian provinces: associated cancer risks and medical expenses. *Journal of hazardous materials*, 187 (1): 574-584.

Cisty, M. 2010. Hybrid genetic algorithm and linear programming method for least-cost design of water distribution systems. *Water Resources Management*, 24 (1): 1-24.

Cook, T. D. 2015. Quasi-experimental Design. *Wiley Encyclopedia of Management*, 11: 1-2.

Cortes, U., Sanchez-Marre, M., Ceccaroni, L., R-Roda, I. and Poch, M. 2000. Artificial intelligence and environmental decision support systems. *Applied Intelligence*, 13 (1): 77-91.

Crittenden, J. C., Trussell, R. R., Hand, D. W., Howe, K. J. and Tchobanoglous, G. 2012. *MWH's Water Treatment: Principles and Design*. New Jersey: John Wiley & Sons.

Dabrowski, J., Dabrowski, J., Hill, L., MacMillan, P. and Oberholster, P. J. 2014. Fate, transport and effects of pollutants originating from acid mine drainage in the Olifants River, South Africa. *River Research and Applications*,

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

Darr, M. and Loflin, K. 2011. Water treatment plant optimization – What can it do for you? In: Proceedings of *NC AWWA-WEA Annual Conference*. North Carolina,

Dellana, S. A. and West, D. 2009. Predictive modeling for wastewater applications: Linear and nonlinear approaches. *Environmental Modelling & Software*, 24 (1): 96-106.

Department of Water Affairs (DWA). 2009. *Blue Drop Report 2009: Version 2 South African Drinking Water Quality Management Performance*. Pretoria: Government Printer.

Department of Water Affairs and Forestry (DWAF). 1996. *South African Water Quality Guidelines, 2nd ed., Vol. 1, Domestic water use*. Pretoria: Department of Water Affairs and Forestry.

Devrimci, H. A., Yuksel, A. M. and Sanin, F. D. 2012. Algal alginate: A potential coagulant for drinking water treatment. *Desalination*, 299: 16-21.

DeZuane, J. 1997. *Handbook of drinking water quality*. New York: Van Nostrand Reinhold.

Dixon, M. B., Richard, Y., Ho, L., Chow, C. W., O'Neill, B. K. and Newcombe, G. 2011. A coagulation–powdered activated carbon–ultrafiltration–Multiple barrier approach for removing toxins from two Australian cyanobacterial blooms. *Journal of hazardous materials*, 186 (2): 1553-1559.

Dowbiggin, W. 2012. Water treatment plant optimization — beyond 0.1 NTU. In: Proceedings of *NC AWWA-WEA Annual Conference*. North Carolina

Droste, R. L. 1997. *Theory and practice of water and wastewater treatment*. New York Wiley.

Duan, N., Mays, L. W. and Lansey, K. E. 1990. Optimal reliability-based design of pumping and distribution systems. *Journal of Hydraulic Engineering*, 116 (2): 249-268.

Dürrenmatt, D. J. and Gujer, W. 2012. Data-driven modeling approaches to support wastewater treatment plant operation. *Environmental Modelling & Software*, 30 (0): 47-56.

Dzwairo, B., Otieno, F. A. and Ochieng, G. M. 2010. Incorporating Surface Raw Water Quality into the Cost Chain for Water Services: Vaal Catchment, South Africa. *Research Journal of Chemistry and Environment*, 14 (1): 29-35.

Dzwairo, B., Otieno, F. A. and Ochieng, G. M. 2012. Rule-based agents for 7-days-ahead forecasting of chemical dosages. Paper presented at. Cape Town, 7 - 10 May. 1-4.

Dzwairo, B., Otieno, F. A., Ochieng, G. M., Letsoalo, M. E. and Leite, E. 2011. *Data reduction for water quality modelling, Vaal Basin*. INTECH Open Access Publisher.

EPA. 2014. *Drinking Water Advice Note No.15: Optimisation of chemical coagulant dosing at water treatment works*. Environmental Protection Agency.

Ersoz, M. and Barrott, L. 2012. *Best practice guide on metals removal from drinking water by treatment*. London: IWA Publishing.

ESRI. 2013. ArcGIS for Desktop: Student Edition Release 10.1 (computer software and manual).

Ezine. 2012. *What factors affect water quality*. Available: http://ezinearticles.com/?What-Factors-Affect-Water-Quality?-Here-Are-the-Top-5&id=2 (Accessed 14/11/2016).

Far, F., Yumei, D. and Zeng-chuan, D. 2001. Multi-objective optimization of water resources decision-making system is running genetic algorithms. *Journal of Hydraulic Engineering*, 9: 22-27.

Farley, M. 2001. *Leakage management and control*. Geneva: World Health Organisation.

Farley, M. and Liemberger, R. 2005a. Developing a Non-Revenue Water Reduction Strategy, Part 1: Investigating and assessing water losses. *Water Science & Technology: Water Supply*, 5 (1): 41-50.

Farley, M. and Liemberger, R. 2005b. Developing a non-revenue water reduction strategy: planning and implementing the strategy. *Water Science & Technology: Water Supply*, 5 (1): 41-50.

Farley, M., Wyeth, G., Ghazali, Z. M., Istandar, A., Singh, S., Dijk, N., Raksakulthai, V. and Kirkwood, E. 2008. *The Manager's Non-Revenue Water Handbook: A guide to understanding water losses*. Bangkok: Ranhill Utilities Berhad and the United States Agency for International Developement.

Farzadkia, M. and Bazrafshan, E. 2014. Lime Stabilization of Waste Activated Sludge. *Health Scope*, 16: 17.

Fearing, D. A., Banks, J., Guyetand, S., Eroles, C. M., Jefferson, B., Wilson, D., Hillis, P., Campbell, A. T. and Parsons, S. A. 2004. Combination of ferric and MIEX® for the treatment of a humic rich water. *Water Research*, 38 (10): 2551-2558.

Foster, L., Diamond, I. and Jefferies, J. 2014. *Beginning statistics: an introduction for social scientists*. Sage.

Fraise, A., Maillard, J.-Y. and Sattar, S. 2012. *Russell, Hugo and Ayliffe's principles and practice of disinfection, preservation and sterilization*. Malden: John Wiley & Sons.

Franceschi, M., Girou, A., Carro-Diaz, A. M., Maurette, M. T. and Puech-Costes, E. 2002. Optimisation of the coagulation–flocculation process of raw water by optimal design method. *Water Research*, 36 (14): 3561-3572.

Frauendorfer, R. and Liemberger, R. 2010. *The issues and challenges of reducing non-revenue water*. Mandaluyong: Asian Development Bank.

Freese, S., Nozaic, D., Smith, R. and Trollip, D. 2000. Powdered activated carbon: Can this be effectively assessed in the laboratory. In: Proceedings of *WISA 2000 Conference, Sun City, South Africa*.

Gen, M. and Cheng, R. 2000. *Genetic algorithms and engineering optimization*. New York: John Wiley & Sons.

George, B. A., Malano, H. M., Khan, A. R., Gaur, A. and Davidson, B. 2009. Urban water supply strategies for Hyderabad, India–Future scenarios. *Environmental Modeling & Assessment*, 14 (6): 691-704.

Giustolisi, O., Savic, D. and Kapelan, Z. 2008. Pressure-driven demand and leakage simulation for water distribution networks. *Journal of Hydraulic Engineering*, 134 (5): 626-635.

Goldberg, D. E. and Deb, K. 1991. A comparative analysis of selection schemes used in genetic algorithms. *Urbana*, 51: 61801-62996.

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

Gomes, R., Marques, A. A. and Sousa, J. 2013. District metered areas design under different decision makers' options: Cost analysis. *Water Resources Management*, 27 (13): 4527-4543.

Google Earth. 2014. Satellite image of the Umzinto Water Treatment Plant (computer software and manual).

Gose, W., Hanson, R., Harmer, R. and Seidel, E. 2013. Reconnaissance paleomagnetic studies of Mesoproterozoic alkaline igneous complexes in the Kaapvaal craton, South Africa. *Journal of African Earth Sciences*, 85: 22-30.

Gray, N. F. 2010. Chapter 10 - Water Treatment and Distribution. In: Gray, N. F. ed. *Water Technology (Third Edition)*. Oxford: Butterworth-Heinemann, 301-327.

Gumbo, B. 2004. The status of water demand management in selected cities of southern Africa. *Physics and Chemistry of the Earth, Parts A/B/C*, 29 (15): 1225-1231.

Gupta, A. K. and Shrivastava, R. K. 2008. Optimal design of water treatment plant under uncertainty using genetic algorithm. *Environmental Progress*, 27 (1): 91-97.

Gupta, A. K. and Shrivastava, R. K. 2010. Reliability-constrained optimization of water treatment plant design using genetic algorithm. *Journal of Environmental Engineering*, 136 (3): 326-334.

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

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

Hamilton, D. P. and Schladow, S. G. 1997. Prediction of water quality in lakes and reservoirs. Part I — Model description. *Ecological Modelling*, 96 (1–3): 91-110.

Hanson, A. T. and Cleasby, J. L. 1990. The effects of temperature on turbulent flocculation: fluid dynamics and chemistry. *Journal (American Water Works Association)*: 56-73.

Harrell, L. and Ranjithan, S. 2003. Detention pond design and land use planning for watershed management. *Journal of Water Resources Planning and Management*, 129 (2): 98-106.

Henry, H. 2013. EPA's Stage 2 Disinfection Byproducts Rules (DBPR) and Northern Kentucky Water: An Economic and Scientific Review. *Dose-Response*, 11 (4): dose-response. 12-056. Henry.

Honda, R. 2014. Removal of engineered nanomaterials through conventional water treatment processes. PhD, University of California Riverside.

Hongve, D., Riise, G. and Kristiansen, J. F. 2004. Increased colour and organic acid concentrations in Norwegian forest lakes and drinking water–a result of increased precipitation? *Aquatic Sciences*, 66 (2): 231-238.

Hrudey, S. E. 2009. Chlorination disinfection by-products, public health risk tradeoffs and me. *Water Research*, 43 (8): 2057-2092.

Huang, C. H., Chang, C. T., Ling, H. C. and Chang, C. C. 1999. A mathematical programming model for water usage and treatment network design. *Industrial Engineering Chemical Research*, 38: 2666-2679.

Hyeon, B., Dae-Won, C., Seng-Tai, L., Yejin, K. and Sungshin, K. 2004. Application of data mining and artificial modeling for coagulant dosage of water treatment plants corresponding to water quality. In: Proceedings of *IECON 2004 - 30th Annual Conference of IEEE*. Busan, South Korea, 2-6 Nov. 2004. Industrial Electronics Society, 2172-2177 Vol. 2173.

Hyndman, R. J. and Koehler, A. B. 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22 (4): 679-688.

Izquierdo, J., Pérez, R. and Iglesias, P. L. 2004. Mathematical models and methods in the water industry. *Mathematical and Computer Modelling*, 39 (11–12): 1353-1374.

Jagielska, I., Matthews, C. and Whitfort, T. 1999. An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems. *Neurocomputing*, 24 (1–3): 37-54.

Jang, J.-S. and Sun, C.-T. 1995. Neuro-fuzzy modeling and control. *Proceedings of the IEEE*, 83 (3): 378-406.

Jeffrey Yang, Y., Goodrich, J. A., Clark, R. M. and Li, S. Y. 2008. Modeling and testing of reactive contaminant transport in drinking water pipes: Chlorine response and implications for online contaminant detection. *Water Research*, 42 (6–7): 1397-1412.

Jeżowski, J., Bochenek, R. and Poplewski, G. 2007. On application of stochastic optimization techniques to designing heat exchanger-and water networks. *Chemical Engineering and Processing: Process Intensification*, 46 (11): 1160-1174.

Johnson, P. V. 1996. Unaccounted for Water puzzle: More than just leakage. *Florida Water Resources Journal*: 37-42.

Joon-Hong, S. and Ju-Jang, L. 2009. Prechlorination control using fuzzy logic and enhanced compact genetic algorithm. In: Proceedings of *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*. 5-8 July 2009. 228-231.

Kaini, P., Artita, K. and Nicklow, J. W. 2012. Optimizing structural best management practices using SWAT and genetic algorithm to improve water quality goals. *Water Resources Management*, 26 (7): 1827-1845.

Karadirek, I., Kara, S., Yilmaz, G., Muhammetoglu, A. and Muhammetoglu, H. 2012. Implementation of hydraulic modelling for water-loss reduction through pressure management. *Water Resources Management*, 26 (9): 2555-2568.

Karr, C. L. and Gentry, E. J. 1993. Fuzzy control of pH using genetic algorithms. *Fuzzy Systems, IEEE Transactions on*, 1 (1): 46.

Kelley, K. and Lai, K. 2011. Accuracy in parameter estimation for the root mean square error of approximation: Sample size planning for narrow confidence intervals. *Multivariate Behavioral Research*, 46 (1): 1-32.

Kim, H.-C. and Yu, M.-J. 2005. Characterization of natural organic matter in conventional water treatment processes for selection of treatment processes focused on DBPs control. *Water Research*, 39 (19): 4779-4789.

Kim, H., Kim, S. and Koo, J. 2014. Prediction of chlorine concentration in various hydraulic conditions for a pilot scale water distribution system. *Procedia Engineering*, 70: 934-942.

Kingdom, B., Liemberger, R. and Marin, P. 2006. The challenge of reducing Non-Revenue Water (NRW) in developing countries. *How the private sector can help: A look at performance-based service contracting, Watersupply and Sanitation Sector Board Discusion Paper Series*, 8: 1-52.

Kingham, T. J. and Hoggart, T. M. 1995. Chlorination control in a large water treatment works. In: Proceedings of *Application of Advanced PLC (Programmable Logic Controller) Systems with Specific Experiences from Water Treatment, IEE Colloquium on (Digest No.1995/112)*. 29 Jun 1995. 2/1-216.

Kotsiantis, S., Kanellopoulos, D. and Pintelas, P. 2006. Data preprocessing for supervised leaning. *International Journal of Computer Science*, 1 (2): 111-117.

Kouziakis, C., Kanakoudis, V., Tsitsifli, S., Doufas, A., Lappos, S., Zouboulis, A., Samaras, P., Kanakoudis, V. and Tsitsifli, S. 2013. Assessing the NRW level in the city of Kozani, Greece: The decisive role of the fixed charge in the water tarrifs. In: Proceedings of *1st EWaS-MED International Conference on "Improving Efficiency of Water Systems in a changing natural and financial environment"* Thessaloniki, 11-13.

Kramer, J. 2012. Alkalinity and acidity. *Water analysis*, 1: 85-135.

Kulkarni, P. and Chellam, S. 2010. Disinfection by-product formation following chlorination of drinking water: Artificial neural network models and changes in speciation with treatment. *Science of The Total Environment*, 408 (19): 4202-4210.

Kumar, S. 2010. Unaccounted for water and the performance of water utilities: an empirical analysis from India. *Water Policy*, 12 (5): 707-721.

Lagvankar, A. L. and Gemmell, R. S. 1968. A size-density relationship for flocs. *Journal (American Water Works Association)*: 1040-1046.

Lai, Y., Tu, Y., Yang, C., Surampalli, R. and Kao, C. 2013. Development of a water quality modeling system for river pollution index and suspended solid loading evaluation. *Journal of Hydrology*, 478: 89-101.

Lambert, A. and Fantozzi, M. 2005. Recent advances in calculating economic intervention frequency for active leakage control, and implications for calculation of economic leakage levels. *Water Science & Technology: Water Supply*, 5 (6): 263-271.

Lambert, A., Fantozzi, M. and Thornton, J. 2013. Practical approaches to modeling leakage and pressure management in distribution systems—Progress since 2005. In: Proceedings of *12th International Conference on Computing and Control for the Water Industry, CCWI2013*. 2-4.

Lansey, K. E. and Basnet, C. 1991. Parameter estimation for water distribution networks. *Journal of Water Resources Planning and Management*, 117 (1): 126-144.

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

Larson, T. and Henley, L. 1955. Determination of low alkalinity or acidity in water. *Analytical Chemistry*, 27 (5): 851-852.

Lawler, D., O'Melia, C. R. and Tobiason, J. E. 1980. Integral water treatment plant design. *Particulates in Water, edited by: Kavanaugh, M. and Leckie, J*: 353-388.

Lawrence, D. ed. 1991. *Handbook of genetic algorithms*. New York: Van Nostrand Reinhold.

Li, J. and Chen, G. 2014. Water footprint assessment for service sector: A case study of gaming industry in water scarce Macao. *Ecological Indicators*, 47: 164-170.

Liemberger, R., Brothers, K., Lambert, A., McKenzie, R., Rizzo, A. and Waldron, T. 2007. Water loss performance indicators. In: Proceedings of *International Conference on the WaterLoss, IWA, Bucharest*. 148-160.

Lim, S.-R. and Park, J. M. 2007. Environmental and economic analysis of a water network system using LCA and LCC. *AIChE Journal*, 53 (12): 3253-3262.

Liu, B. and Reckhow, D. A. 2013. DBP Formation in Hot and Cold Water Across a Simulated Distribution System: Effect of Incubation Time, Heating Time, pH, Chlorine Dose, and Incubation Temperature. *Environmental science & technology*, 47 (20): 11584-11591.

Liu, H., Rong, Q., Xu, L. and Zhang, Y. 2013. Application of Automatic Control System in Water Plant in China. *Journal of Convergence Information Technology*, 8 (5)

Liu, S., Butler, D., Brazier, R., Heathwaite, L. and Khu, S.-T. 2007. Using genetic algorithms to calibrate a water quality model. *Science of The Total Environment*, 374 (2–3): 260-272.

Loucks, D. P., Van Beek, E., Stedinger, J. R., Dijkman, J. P. and Villars, M. T. 2005. *Water resources systems planning and management: an introduction to methods, models and applications*. Paris: UNESCO.

Luo, Y., Nou, X., Yang, Y., Alegre, I., Turner, E., Feng, H., Abadias, M. and Conway, W. 2011. Determination of free chlorine concentrations needed to prevent Escherichia coli O157: H7 cross-contamination during fresh-cut produce wash. *Journal of Food Protection®*, 74 (3): 352-358.

Magombo, J., Dzwairo, B., Moyo, S. and Dewa, M. 2015. Data pre-processing for process optimization at a drinking water treatment plant in Ugu District Municipality, South Africa. *Environmental Economics*, 6 (1): 159-171.

Mamba, B., Krause, R., Matsebula, B. and Haarhoff, J. 2009. Monitoring natural organic matter and disinfection by-products at different stages in two South African water treatment plants. *Water SA*, 35 (1): 121-127.

Mamba, B., Rietveld, L. and Verberk, J. 2008. SA drinking water standards under the microscope: drinking water. *Water Wheel*, 7 (1): 24-27.

Marinovich, M., Campagner, P., Restani, P., Tragni, E. and Galli, C. 2012. Evaluation of the toxicological risk in animals following the ingestion of cheese obtained from milk with added formaldehyde. In: Proceedings of *Disease, Metabolism and Reproduction in the Toxic Response to Drugs and Other Chemicals: Proceedings of the European Society of Toxicology Meeting Held in Rome, March 28–30, 1983*. Springer Science & Business Media, 430.

Marunga, A., Hoko, Z. and Kaseke, E. 2006. Pressure management as a leakage reduction and water demand management tool: The case of the City of Mutare, Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, 31 (15): 763-770.

MathWorks. 2011. RMSE - Root mean square Error. *MATLAB Answers* (Blog). Available: https://www.mathworks.com/matlabcentral/answer/4774-rmse-root-mean-square-error (Accessed 12/11/2016).

MathWorks. 2012. *Genetic Algorithms Examples*. Available: http://www.mathworks.com/Genetic%20Algorithm%20Examples%20%20MATLAB%20&%20Simulink.htm (Accessed 12/11/2016).

MathWorks. 2014. Matlab 2014b (Genetic Algorithim Optimisation Toolbox) (computer software and manual).

Matilainen, A., Lindqvist, N., Korhonen, S. and Tuhkanen, T. 2002. Removal of NOM in the different stages of the water treatment process. *Environment International*, 28 (6): 457-465.

McKenzie, R. and Wegelin, W. 2009. Challenges facing the implementation of water demand management initiatives in Gauteng Province. *Water SA*, 35 (2): 168-174.

McKenzie, R. S., Siqalaba, Z. and Wegelin, W. 2012. *The state of Non-Revenue Water in South Africa* Pretoria: Water Research Commission.

Messai, A., Mellit, A., Guessoum, A. and Kalogirou, S. 2011. Maximum power point tracking using a GA optimized fuzzy logic controller and its FPGA implementation. *Solar energy*, 85 (2): 265-277.

Michalewicz, Z. and Schoenauer, M. 1996. Evolutionary algorithms for constrained parameter optimization problems. *Evol. Comput.*, 4 (1): 1-32.

Microsoft Corporation. 2012a. Microsoft SQL Management Studio (computer software and manual).

Microsoft Corporation. 2012b. Visual Studio (computer software and manual).

Mintz, E. D., Reiff, F. M. and Tauxe, R. V. 1995. Safe water treatment and storage in the home: a practical new strategy to prevent waterborne disease. *JAMA-Journal of the American Medical Association-US Edition*, 273 (12): 948-953.

Mirsepassi, A. 2004. Application of intelligent system for water treatment plant operation. *Iranian Journal of Environmental Health Science & Engineering*, 1 (2): 51-57.

Momba, M. N. B., Obi, C. L. and Thompson, P. 2009. Survey of disinfection efficiency of small drinking water treatment plants : challenges facing small water treatment plants in South Africa. *Water SA*, 35 (4): 485-493.

Momba, M. N. B., Tyafa, Z. and Makala, N. 2004. Rural water treatment plants fail to provide potable water to their consumers : the Alice water treatment plant in the Eastern Cape province of South Africa : research letter. *South African Journal of Science*, 100 (5 & 6): p.307-310.

Motiee, H., McBean, E. and Motiei, A. 2007. Estimating physical unaccounted for water (UFW) in distribution networks using simulation models and GIS. *Urban Water Journal*, 4 (1): 43-52.

Mountain Empire Community College. 2013. *Water/Wastewater Distance Learning*. Available: http://water.me.vccs.edu/ (Accessed 14/11/2016).

Muleta, M. K. and Nicklow, J. W. 2005. Decision support for watershed management using evolutionary algorithms. *Journal of Water Resources Planning and Management*, 131 (1): 35-44.

Mutikanga, H., Sharma, S., Vairavamoorthy, K. and Cabrera, E. 2010. Using performance indicators as a water loss management tool in developing countries. *Journal of Water Supply: Research and Technology—AQUA*, 59 (8): 471-481.

Mutikanga, H. E. 2012. Water loss management: tools and methods for developing countries. PhD, Delft University of Technology.

Ndwambi, R. 2009. Portable water treatment plants : water treatment. *Water & Sanitation Africa*, 4 (5): 71-73.

Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A. and Singh, A. 2009. State of the art for genetic algorithms and beyond in water resources planning and management. *Journal of Water Resources Planning and Management*, 136 (4): 412-432.

Nyemba, A., Manzungu, E., Masango, S. and Musasiwa, S. 2010. The impact of water scarcity on environmental health in selected residential areas in Bulawayo City, Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, 35 (13–14): 823-827.

O'Connor, J. T. and O'Connor, T. L. 2001. *Removal of microorganisms by Rapid Sand Filtration*. Columbia: H2O'C Engineering.

Obi, C., Igumbor, J., Momba, M. and Samie, A. 2008. Interplay of factors involving chlorine dose, turbidity flow capacity and pH on microbial quality of drinking water in small water treatment plants. *Water SA*, 34: 565-572.

Park, H. J. 2007. A study to develop strategies for proactive water-loss management. Doctor of Philosophy, Georgia State University and the Georgia Institute of Technology.

Park, S.-Y., Choi, J. H., Wang, S. and Park, S. S. 2006. Design of a water quality monitoring network in a large river system using the genetic algorithm. *Ecological Modelling*, 199 (3): 289-297.

Peat, J. 2001. *Health science research: A handbook of quantitative methods*. London: Sage.

Pennington, D. C. 2003. *Essential personality*. London: Anorld.

Perez-Pedini, C., Limbrunner, J. F. and Vogel, R. M. 2005. Optimal location of infiltration-based best management practices for storm water management. *Journal of Water Resources Planning and Management*, 131 (6): 441-448.

Piri, I., Homayoonnezhad, I. and Amirian, P. 2010. Investigation on optimization of conventional drinking water treatment plant. In: Proceedings of *Chemical, Biological and Environmental Engineering (ICBEE), 2010 2nd International Conference on*. 2-4 Nov. 2010. 304-310.

Platikanov, S., Martín, J. and Tauler, R. 2012. Linear and non-linear chemometric modeling of THM formation in Barcelona's water treatment plant. *Science of The Total Environment*, 432 (0): 365-374.

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

Pual, A., Rodrguez, J., Franco, A., Carrasco, E., Roca, E. and Lema, J. 2001. Advanced monitoring and control of anaerobic wastewater treatment plants: diagnosis and supervision by a fuzzy-based expert system. *Water Science & Technology*, 43 (7): 191-198.

Rajasekar, S., Philominathan, P. and Chinnathambi, V. 2006. Research methodology.

Ramaley, B. L., O'Melia, C. R., Wright, W. C. and Lawler, D. F. 1981. Integral analysis of water plant performance. *Journal of the Environmental Engineering Division*, 107 (3): 547-562.

Ramavandi, B. 2014. Treatment of water turbidity and bacteria by using a coagulant extracted from Plantago ovata. *Water Resources and Industry*, 6: 36-50.

Ramjatan, A., Dlamini, P., Tiba, F. and Pillay, M. 2007. *Water utilities and benchmarking: A case study of Umgeni Water*. Pietermaritzburg: Umgeni Water.

Rand, J. L., Gagnon, G. A. and Knowles, A. 2014. Establishing minimum free chlorine residual concentration for microbial control in a municipal drinking water distribution system. *Water Practice & Technology*, 9 (4): 491-501.

Rangeti, I., Dzwairo, B., Barratt, G. and Otieno, F. 2015. Ecosystem-specific water quality indices. *African Journal of Aquatic Science*: 1-8.

Rashed, I. G. A.-A., Afify, H. A., Ahmed, A. E.-M. and Ayoub, M. A. E.-S. 2013. Optimization of chemical precipitation to improve the primary treatment of wastewater. *Desalination and Water Treatment*, 51 (37-39): 7048-7056.

Reichold, L., Zechman, E. M., Brill, E. D. and Holmes, H. 2009. Simulation-optimization framework to support sustainable watershed development by mimicking the predevelopment flow regime. *Journal of Water Resources Planning and Management*, 136 (3): 366-375.

Richardson, S. D. and Postigo, C. 2012. Drinking water disinfection by-products. In: *Emerging organic contaminants and human health*. Springer, 93-137.

Rietveld, L. C., van der Helm, A. W. C., van Schagen, K. M. and van der Aa, L. T. J. 2010. Good modelling practice in drinking water treatment, applied to Weesperkarspel plant of Waternet. *Environmental Modelling & Software*, 25 (5): 661-669.

Rivett, U., Taylor, D., Chair, C., Forlee, B., Mrwebi, M., van Belle, J.-P. and Chigona, W. 2014. *Community engagement in drinking water supply management: a review* Cape Town: Water Research Commision.

Rizzo, A., Vermersch, M., John, S. G. S., Micallef, G. and Pace, R. 2007. Apparent water loss control: The way forward. *Water21, IWA Publishing*, 9 (4): 45-47.

Rodriguez, M. J. and Sérodes, J.-B. 1998. Assessing empirical linear and non-linear modelling of residual chlorine in urban drinking water systems. *Environmental Modelling & Software*, 14 (1): 93-102.

Rossini, M., Garrido, J. G. and Galluzzo, M. 1999. Optimization of the coagulation–flocculation treatment: influence of rapid mix parameters. *Water Research*, 33 (8): 1817-1826.

Rowntree, D. 2004. *Statistics without tears: A primer for non-mathematicians*. Scribner Book Company.

Sevilla, M., Falco, C., Titirici, M.-M. and Fuertes, A. B. 2012. High-performance CO 2 sorbents from algae. *RSC Advances*, 2 (33): 12792-12797.

Shariff, R., Cudrak, A. and Stanley, S. J. 2004. Lime softening clarifier modeling with artificial neural networks. *Journal of Environmental Engineering and Science*, 3 (S1): S69-S80.

Shariff, R., Cudrak, A., Zhang, Q., & Stanley, S. J. 2004. Advanced process control techniques for water treatment using artificial neural networks1. *Journal of Environmental Engineering and Science*, 3: S61-S67.

Shittu, O., Olaitan, J. and Amusa, T. 2008. Physico-chemical and bacteriological analyses of water used for drinking and swimming purposes in abeokuta, nigeria. *African Journal of Biomedical Research*, 11 (3)

Shroff, P., Vashi, R., Champaneri, V. and Patel, K. 2015. Correlation study among water quality parameters of groundwater of Valsad district of south Gujarat (India). *Journal of Fundamental and Applied Sciences*, 7 (3): 340-349.

Sillanpää, M. and Matilainen, A. 2014. NOM Removal by Coagulation. *Natural Organic Matter in Water: Characterization and Treatment Methods*: 55.

Simonovic, S. P. 2012. *Managing water resources: methods and tools for a systems approach*. UK: Earthscan.

Simpson, A. R., Dandy, G. C. and Murphy, L. J. 1994. Genetic algorithms compared to other techniques for pipe optimization. *Journal of Water Resources Planning and Management*, 120 (4): 423-443.

Sing, C. and Postlethwaite, B. 1997. pH control: handling nonlinearity and deadtime with fuzzy relational model-based control. In: Proceedings of *Control Theory and Applications, IEE Proceedings*. IET, 263-268.

Singh, K. P., Basant, A., Malik, A. and Jain, G. 2009. Artificial neural network modeling of the river water quality—A case study. *Ecological Modelling*, 220 (6): 888-895.

Skinner, B., Yuan, S., Huang, S., Liu, D., Cai, B., Dissanayake, G., Lau, H., Bott, A. and Pagac, D. 2013. Optimisation for job scheduling at automated container terminals using genetic algorithm. *Computers & Industrial Engineering*, 64 (1): 511-523.

Smeets, P., Medema, G. and Van Dijk, J. 2009. The Dutch secret: how to provide safe drinking water without chlorine in the Netherlands. *Drinking Water Engineering and Science*, 2 (1): 1-14.

South African Bureau of Standards. 2005. *South African National Standard (SANS 241) 2005: Drinking Water*. Pretoria: South African Bureau of Standards (SABS).

South African Bureau of Standards. 2011. *South African National Standard (SANS 241) 2011. Drinking Water Specification: Part 1: Microbiological, physical, aesthetic and chemical determinants*. Pretoria: South African Bureau of Standards (SABS).

Stevenson, W. J. and Sum, C. C. 2009. *Operations management*. Boston: McGraw-Hill / Irwin.

Sun, C., Yue, Q., Gao, B., Cao, B., Mu, R. and Zhang, Z. 2012. Synthesis and floc properties of polymeric ferric aluminum chloride–polydimethyl diallylammonium chloride coagulant in coagulating humic acid–kaolin synthetic water. *Chemical Engineering Journal*, 185: 29-34.

Tae-Hwan, H., Eui-Suck, N., Kwang-Bang, W., Kim, C. J. and Jeong-Woong, R. 1997. Optimization of coagulant dosing process in water purification system. In: Proceedings of *SICE '97. Proceedings of the 36th SICE Annual Conference. International Session Papers*. 29-31 Jul 1997. 1105-1109.

Tan, W., Lu, F., Loh, A. and Tan, K. 2005. Modeling and control of a pilot pH plant using genetic algorithm. *Engineering Applications of Artificial Intelligence*, 18 (4): 485-494.

Toroz, I. and Uyak, V. 2005. Seasonal variations of trihalomethanes (THMs) in water distribution networks of Istanbul City. *Desalination*, 176 (1): 127-141.

Trochim, W. M. and Donnelly, J. 2005. *Research methods: The concise knowledge base*. Cincinnati: Atomic Dog Publication.

Tsai, M.-J. and Chang, C.-T. 2001. Water usage and treatment network design using genetic algorithms. *Industrial & Engineering Chemistry Research* 40 (22): 4874-4888.

Tsompanakis, Y., Lagaros, N. D. and Papadrakakis, M. 2002. Innovative computational methods for shape optimization. In: Proceedings of *5th GRACM Congress on Computational Mechanics*.

Ugu District Municipality. 2007. *Intergrated Development Plan 2007-2012*. Port Shepstone: Ugu District Municipality.

Ugu District Municipality. Affairs, D. o. W. 2011. *First Stage Reconciliation Strategy For Umzinto Water Supply Scheme Area – Umdoni Local Municipality*. Port Shepstone: Ugu District Municipality.

Umgeni. 2015. *Umzinto Water Treatment Plant Operational Data 2010 - 2015*: Umgeni.

Upadhyay, A. and Agarwal, A. 2008. Controller design and analysis for automation of Chemical Water Treatment system. In: Proceedings of *Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on*. 17-20 Dec. 2008. 1625-1629.

USEPA. USEPA. 1999. *Giarda: drinking water health advisory*.(Notice EPA-822-R-99-008). Washington, United States: Environmental Protection Agency.

USEPA. USEPA. 2001. *Cryptosporodium: drinking water health advisory*.(Notice EPA-822-R-01-009). Washington, D.C: Environmental Protection Agency.

Uyak, V., Soylu, S., Topal, T., Karapinar, N., Ozdemir, K., Ozaydin, S. and Avsar, E. 2014. Spatial and Seasonal Variations of Disinfection Byproducts (DBPs) in Drinking Water Distribution Systems of Istanbul City, Turkey. *Environmental Forensics*, 15 (2): 190-205.

Van der Zaag, P. 2003. Urban Water Demand, IWRM 0.1 Lecture Notes. University of Zimbabwe.

Vasquez, J. A., Maier, H. R., Lence, B. J., Tolson, B. A. and Foschi, R. O. 2000. Achieving water quality system reliability using genetic algorithms. *Journal of Environmental Engineering*, 126 (10): 954-962.

Veenstra, J. N. and Schnoor, J. L. 1980. Seasonal variations in trihaiomethane levels in an Iowa River water supply. *Journal (American Water Works Association)*: 583-590.

Verrelli, D. I., Dixon, D. R. and Scales, P. J. 2009. Effect of coagulation conditions on the dewatering properties of sludges produced in drinking water treatment. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 348 (1–3): 14-23.

Wagner, J. M., Shamir, U. and Marks, D. H. 1988. Water distribution reliability: Analytical methods. *Journal of Water Resources Planning and Management*, 114 (3): 253-275.

Wang, Q. J. 1997. Using genetic algorithms to optimise model parameters. *Environmental Modelling & Software*, 12 (1): 27-34.

Wang, X., Lu, W., Cao, S. and Fang, D. 2007. Using time-delay neural network combined with genetic algorithms to predict runoff level of Linshan Watershed, Sichuan, China. *Journal of Hydrologic Engineering*, 12 (2): 231-236.

Wang, Y. P. and Smith, R. 1995. Wastewater minimization with flowrate constraints. *Transactions of the Institution of Chemical Engineers*, 73 (Part A): 889–904.

Warburton, A. 1983. Quasiconcave vector maximization: connectedness of the sets of Pareto-optimal and weak Pareto-optimal alternatives. *Journal of Optimization Theory and Applications*, 40 (4): 537-557.

Water purification. 2013. Available: http://www.waterpurifierhome.com (Accessed 16/11/2016).

Weiss, S. M. and Indurkhya, N. 2000. *Decision-rule solutions for data mining with missing values*. Springer.

Wesanekar, P., Shelake, N., Kulkarni, U. and Sakshi, K. 2014. Water analysis of Godawari river (Dhangar Takali to Vishnupuri Dam), Maharashtra, India. *Lakes: the mirrors of the earth*: 329.

Whitley, D. 2001. An overview of evolutionary algorithms: practical issues and common pitfalls. *Information and Software Technology*, 43 (14): 817-831.

WHO. 2006. *Guidelines for drinking-water quality: First addendum to volume 1, Recommendations*. Geneva: World Health Organization.

WHO. 2011. *Guidelines for Drinking-water Quality*. Geniva: World Health Organisation.

Willhite, C., Bhat, V., Ball, G. and McLellan, C. 2012. Emergency do not consume/do not use concentrations for potassium permanganate in drinking water. *Human & Experimental Toxicology*,

Willmott, C. J., Matsuura, K. and Robeson, S. M. 2009. Ambiguities inherent in sums-of-squares-based error statistics. *Atmospheric Environment*, 43 (3): 749-752.

Wing, S. 2014. Reservoir sediment carbon along the Elwha River after dam removal. Master of Science, University of Washington.

Wu, W., Simpson, A. R., Maier, H. R. and Marchi, A. 2011. Incorporation of variable-speed pumping in multiobjective genetic algorithm optimization of the design of water transmission systems. *Journal of Water Resources Planning and Management*, 138 (5): 543-552.

Wu, Z. Y. and Sage, P. 2006. Water loss detection via genetic algorithm optimization-based model calibration. *ASCE 8th Annual International*, 5

Wulff, L., Inniss, E. and Clevenger, T. 2013. Importance of bromine-substituted DBP's in drinking water. *Journal of Water Resource and Protection*, 5 (8A): 28-34.

Yan, M., Wang, D., You, S., Qu, J. and Tang, H. 2006. Enhanced coagulation in a typical North-China water treatment plant. *Water Research*, 40 (19): 3621-3627.

Yandamuri, S. R., Srinivasan, K. and Murty, B. S. 2006. Multiobjective optimal waste load allocation models for rivers using nondominated sorting genetic algorithm-II. *Journal of Water Resources Planning and Management*, 132 (3): 133-143.

Yeh, C.-H. and Labadie, J. W. 1997. Multiobjective watershed-level planning of storm water detention systems. *Journal of Water Resources Planning and Management*, 123 (6): 336-343.

Zetland, D. J. 2011. Global water tariffs continue upward trend. *Global Water Intelligence*, 12 (9): 35-40.

Zinkus, G. A., Byers, W. D. and Doerr, W. W. 1998. Identify appropriate water reclamat technologies. *Chemical Engineering Progress*

APPENDICES

Appendix 1: The scripts that were used to extract each parameter for the two abstraction points

```
SELECT * FROM [dbo].[waterparam] where Determinand = pH;
OUTPUT TO 'c:\\data\\pH.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = conductivity;
OUTPUT TO 'c:\\data\\conductivity.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = nitrates;
OUTPUT TO 'c:\\data\\nitrates.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = Total phosphate;
OUTPUT TO 'c:\\data\\Total phosphate.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = temperature;
OUTPUT TO 'c:\\data\\temperature.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;

SELECT * FROM [dbo].[waterparam] where Determinand = turbidity;
OUTPUT TO 'c:\\data\\turbidity.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = algae;
OUTPUT TO 'c:\\data\\algae.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = Ecoli;
OUTPUT TO 'c:\\data\\Ecoli.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = DO;
OUTPUT TO 'c:\\data\\DO.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
SELECT * FROM [dbo].[waterparam] where Determinand = NH3;
OUTPUT TO 'c:\\data\\NH3.csv'
    FORMAT TEXT
    QUOTE '"'
    WITH COLUMN NAMES;
```

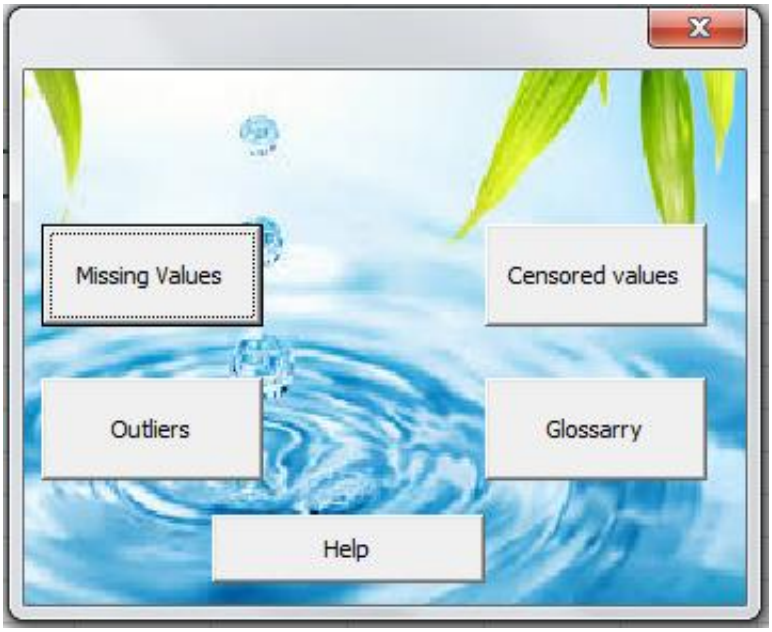
Appendix 2: Summary of water quality parameter observations at the Umzinto Water Treatment Plant

	Umzinto (TUZ0002)					Ej Smith (TUZ0001)			
	Expected Observations	Blanks	Observations	% Observations		Expected Observations	Blanks	Observations	% Observations
Algalcoun	850	98	752	88.47058824	Algalcoun	885	116	769	86.89265537
Mn	850	139	711	83.64705882	Mn	885	146	739	83.50282486
Fe	850	141	709	83.41176471	Fe	885	149	736	83.16384181
Geosmin	850	163	687	80.82352941	Colour	885	195	690	77.96610169
Ecoli	850	164	686	80.70588235	Ecoli	885	197	688	77.74011299
TwoMIB	850	165	685	80.58823529	Coliforms	885	198	687	77.62711864
Coliforms	850	165	685	80.58823529	Geosmin	885	262	623	70.39548023
Colour	850	172	678	79.76470588	TwoMIB	885	266	619	69.94350282
Conductiv	850	262	588	69.17647059	Cryptomo	885	284	601	67.90960452
Turbidity	850	330	520	61.17647059	Conductiv	885	335	550	62.14689266
SS	850	335	515	60.58823529	Turbidity	885	359	526	59.43502825
pHOS	850	339	511	60.11764706	Temperat	885	370	515	58.1920904
Temperat	850	341	509	59.88235294	TOC	885	372	513	57.96610169
TOC	850	355	495	58.23529412	Chlamydo	885	384	501	56.61016949
Alkalinity	850	368	482	56.70588235	pHOS	885	387	498	56.27118644
AIT	850	441	409	48.11764706	Chlorella	885	390	495	55.93220339
Odour	850	445	405	47.64705882	Alkalinity	885	396	489	55.25423729
Appearan	850	449	401	47.17647059	SS	885	396	489	55.25423729
Ca	850	515	335	39.41176471	Odour	885	457	428	48.36158192
Mg	850	515	335	39.41176471	Appearan	885	477	408	46.10169492
HardnessT	850	517	333	39.17647059	Melosira	885	477	408	46.10169492
F	850	518	332	39.05882353	AIT	885	495	390	44.06779661
Nitzschia	850	523	327	38.47058824	Crucigenia	885	527	358	40.4519774
Sulphate	850	524	326	38.35294118	F	885	539	346	39.0960452
Nitrate	850	542	308	36.23529412	Sulphate	885	543	342	38.6440678
TDS	850	545	305	35.88235294	Ca	885	546	339	38.30508475
Na	850	550	300	35.29411765	Mg	885	546	339	38.30508475
AMMONIA	850	566	284	33.41176471	HardnessT	885	549	336	37.96610169
Navicula	850	593	257	30.23529412	TDS	885	567	318	35.93220339
Chlorella	850	606	244	28.70588235	Nitrate	885	573	312	35.25423729
Cryptomo	850	619	231	27.17647059	Na	885	582	303	34.23728814
K	850	630	220	25.88235294	AMMONIA	885	583	302	34.12429379
Chlamydo	850	642	208	24.47058824	Trachelon	885	603	282	31.86440678
AMMONIA	850	647	203	23.88235294	Oocystis	885	623	262	29.60451977
Cocconeis	850	721	129	15.17647059	MnS	885	625	260	29.37853107
Cymbella	850	725	125	14.70588235	FeS	885	630	255	28.81355932
Corrosivit	850	739	111	13.05882353	Coelastru	885	653	232	26.21468927
LSI	850	745	105	12.35294118	K	885	656	229	25.87570621
RSI	850	745	105	12.35294118	Gonium	885	657	228	25.76271186
Cryptospo	850	748	102	12	PERCAMM	885	659	226	25.53672316
Nitrous	850	756	94	11.05882353	AMMONIA	885	660	225	25.42372881
Achnanth	850	756	94	11.05882353	Microcyst	885	676	209	23.61581921
Br	850	761	89	10.47058824	Scenedes	885	679	206	23.27683616
FeS	850	774	76	8.941176471	Nitzschia	885	693	192	21.69491525
MnS	850	774	76	8.941176471	Synedra	885	699	186	21.01694915
Synedra	850	781	69	8.117647059	Cyclotella	885	708	177	20
Cr	850	785	65	7.647058824	Euglena	885	719	166	18.75706215
Melosira	850	788	62	7.294117647	Other	885	740	145	16.38418079
Scenedes	850	790	60	7.058823529	Mallomor	885	746	139	15.70621469
Si	850	795	55	6.470588235	OdourSO	885	747	138	15.59322034
CNT	850	798	52	6.117647059	Oscillator	885	764	121	13.67231638
Cu	850	798	52	6.117647059	Cryptospo	885	764	121	13.67231638
Oocystis	850	805	45	5.294117647	Giardia	885	764	121	13.67231638
Trachelon	850	807	43	5.058823529	Corrosivit	885	770	115	12.99435028
Phenols	850	810	40	4.705882353	Anabaena	885	770	115	12.99435028
Zn	850	810	40	4.705882353	Cosmariu	885	776	109	12.31638418
CNF	850	812	38	4.470588235	RSI	885	777	108	12.20338983
Sb	850	813	37	4.352941176	LSI	885	778	107	12.09039548
Se	850	813	37	4.352941176	Nitrous	885	782	103	11.63841808
Cd	850	814	36	4.235294118	Fstrep	885	784	101	11.41242938
Ni	850	814	36	4.235294118	Br	885	799	86	9.717514124
Pb	850	814	36	4.235294118	Elakatoth	885	810	75	8.474576271
Oscillator	850	818	32	3.764705882	Closteriur	885	810	75	8.474576271
Gyrosigma	850	819	31	3.647058824	Ankistrod	885	821	64	7.231638418
Crucigenia	850	821	29	3.411764706	Navicula	885	822	63	7.118644068
As	850	821	29	3.411764706	Cr	885	829	56	6.327683616
Other	850	822	28	3.294117647	Peridiniur	885	831	54	6.101694915
Hardness0	850	822	28	3.294117647	Si	885	831	54	6.101694915
Ag	850	824	26	3.058823529	Dictyosph	885	833	52	5.875706215
Ba	850	824	26	3.058823529	Tetraedro	885	834	51	5.762711864
Cyclotella	850	825	25	2.941176471	Staurastru	885	835	50	5.649717514
B	850	825	25	2.941176471	CNT	885	839	46	5.197740113
V	850	826	24	2.823529412	Pediastru	885	842	43	4.858757062
Actinomyt	850	826	24	2.823529412	Quadrigul	885	842	43	4.858757062
Diatoma	850	828	22	2.588235294	Cu	885	844	41	4.632768362
Euglena	850	828	22	2.588235294	Phenols	885	848	37	4.18079096
Pesticides	850	828	22	2.588235294	Pandorina	885	849	36	4.06779661
Gomphon	850	828	22	2.588235294	Fragilaria	885	850	35	3.95480226
Microcysti	850	829	21	2.470588235	Stichococ	885	852	33	3.728813559
Hg	850	829	21	2.470588235	CNF	885	852	33	3.728813559

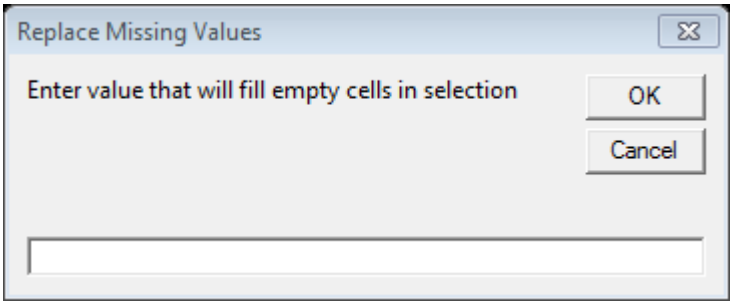
	Umzinto (TUZ0002)			
	Expected Observations	Blanks	Observations	% Observations
Tetraedro	850	830	20	2.352941176
Cosmariu	850	831	19	2.235294118
Quadrigul	850	831	19	2.235294118
Fragilaria	850	832	18	2.117647059
Ankistrod	850	836	14	1.647058824
PennateD	850	836	14	1.647058824
Stichococ	850	839	11	1.294117647
Co	850	839	11	1.294117647
Euastrum	850	839	11	1.294117647
Ceratium	850	840	10	1.176470588
Surirella	850	840	10	1.176470588
Anabaena	850	841	9	1.058823529
Twocp	850	841	9	1.058823529
Twonp	850	841	9	1.058823529
Fournp	850	841	9	1.058823529
pcp	850	841	9	1.058823529
Phenol	850	841	9	1.058823529
Coliphage	850	841	9	1.058823529
Coelastru	850	842	8	0.941176471
SOC	850	842	8	0.941176471
Spirogyra	850	842	8	0.941176471
Pteromon	850	843	7	0.823529412
Elakatothr	850	843	7	0.823529412
Fstrep	850	843	7	0.823529412
Closteriur	850	843	7	0.823529412
Mallomon	850	844	6	0.705882353
Gonium	850	845	5	0.588235294
Peridiniur	850	845	5	0.588235294
Pandorina	850	845	5	0.588235294
TKN	850	845	5	0.588235294
CCTwoone	850	846	4	0.470588235
CCThreeS	850	846	4	0.470588235
Chlorogon	850	846	4	0.470588235
Staurastru	850	846	4	0.470588235
Thoracom	850	846	4	0.470588235
Sphaerocy	850	847	3	0.352941176
Pediastru	850	847	3	0.352941176
Sideroceli	850	847	3	0.352941176
Schroeder	850	847	3	0.352941176
Kirchneric	850	847	3	0.352941176
Haematoc	850	848	2	0.235294118
Spermato	850	848	2	0.235294118
Phacus	850	848	2	0.235294118
Phacotus	850	848	2	0.235294118
Dictyosph	850	848	2	0.235294118
Merismop	850	848	2	0.235294118
Attheya	850	848	2	0.235294118
U	850	848	2	0.235294118
Chroococc	850	848	2	0.235294118
Cytopatho	850	848	2	0.235294118
Thalassios	850	848	2	0.235294118
Selenastru	850	848	2	0.235294118
Golenkini	850	848	2	0.235294118
Tabellaria	850	848	2	0.235294118
Centricdia	850	848	2	0.235294118
Stephano	850	849	1	0.117647059
Cl	850	849	1	0.117647059
Botryococ	850	849	1	0.117647059
DOOS	850	849	1	0.117647059
Volvox	850	849	1	0.117647059
Pleurococ	850	849	1	0.117647059
PERCAMM	850	849	1	0.117647059
Giardia	850	849	1	0.117647059
Dinobryor	850	849	1	0.117647059
Micrasteri	850	849	1	0.117647059
OdourSO	850	849	1	0.117647059
NitrateNit	850	849	1	0.117647059
TwomFou	850	849	1	0.117647059
FourcThre	850	849	1	0.117647059
Cymatopl	850	849	1	0.117647059
Lepocincli	850	849	1	0.117647059
Actinastru	850	849	1	0.117647059
Tetrastrur	850	849	1	0.117647059
Monoraph	850	849	1	0.117647059
Eudorina	850	849	1	0.117647059
Biddulphi	850	849	1	0.117647059
Micractini	850	849	1	0.117647059
Synechocy	850	849	1	0.117647059
Synura	850	849	1	0.117647059
Ankyra	850	849	1	0.117647059

	Ej Smith (TUZ0001)			
	Expected Observations	Blanks	Observations	% Observations
Cd	885	853	32	3.615819209
Ni	885	853	32	3.615819209
Sb	885	853	32	3.615819209
Zn	885	853	32	3.615819209
As	885	854	31	3.502824859
Se	885	854	31	3.502824859
Hardness(	885	856	29	3.276836158
Pb	885	858	27	3.050847458
Phacus	885	859	26	2.937853107
Ag	885	859	26	2.937853107
B	885	860	25	2.824858757
Ba	885	860	25	2.824858757
Schroeder	885	860	25	2.824858757
Achnanth	885	863	22	2.485875706
Pteromon	885	864	21	2.372881356
Botryococ	885	864	21	2.372881356
Pesticides	885	864	21	2.372881356
Hg	885	865	20	2.259887006
V	885	865	20	2.259887006
Actinastru	885	865	20	2.259887006
Cymbella	885	867	18	2.033898305
Thoracom	885	867	18	2.033898305
Sideroceli	885	868	17	1.920903955
Spermato	885	869	16	1.807909605
Monoraph	885	869	16	1.807909605
Sphaerocy	885	870	15	1.694915254
Dinobryor	885	870	15	1.694915254
Tetrastrur	885	870	15	1.694915254
Pleurococ	885	871	14	1.581920904
Lepocincli	885	871	14	1.581920904
Chroococc	885	872	13	1.468926554
Attheya	885	874	11	1.242937853
Coliphage	885	875	10	1.129943503
Twocp	885	876	9	1.016949153
Twonp	885	876	9	1.016949153
FourcThre	885	876	9	1.016949153
Fournp	885	876	9	1.016949153
pcp	885	876	9	1.016949153
Phenol	885	876	9	1.016949153
Ceratium	885	876	9	1.016949153
Cymatopl	885	877	8	0.903954802
SOC	885	878	7	0.790960452
Kirchneric	885	878	7	0.790960452
PennateD	885	878	7	0.790960452
Co	885	878	7	0.790960452
Euastrum	885	878	7	0.790960452
Diatoma	885	879	6	0.677966102
NitrateNit	885	879	6	0.677966102
Phacotus	885	880	5	0.564971751
Merismop	885	880	5	0.564971751
Micractini	885	880	5	0.564971751
Synura	885	880	5	0.564971751
TKN	885	881	4	0.451977401
Thalassios	885	881	4	0.451977401
Actinomy	885	881	4	0.451977401
Haematoc	885	882	3	0.338983051
Volvox	885	882	3	0.338983051
Gomphon	885	882	3	0.338983051
Micrasteri	885	882	3	0.338983051
Eudorina	885	882	3	0.338983051
Selenastru	885	882	3	0.338983051
Cocconeis	885	882	3	0.338983051
Biddulphi	885	882	3	0.338983051
Chlorogon	885	883	2	0.225988701
U	885	883	2	0.225988701
Cytopatho	885	883	2	0.225988701
Surirella	885	883	2	0.225988701
Synechocy	885	883	2	0.225988701
Ankyra	885	883	2	0.225988701
Stephano	885	884	1	0.11299435
CCTwoone	885	884	1	0.11299435
CCThreeS	885	884	1	0.11299435
Cl	885	884	1	0.11299435
DOOS	885	884	1	0.11299435
Gyrosigma	885	884	1	0.11299435
TwomFou	885	884	1	0.11299435
Golenkini	885	884	1	0.11299435
Spirogyra	885	884	1	0.11299435
Tabellaria	885	884	1	0.11299435
Centricdia	885	884	1	0.11299435

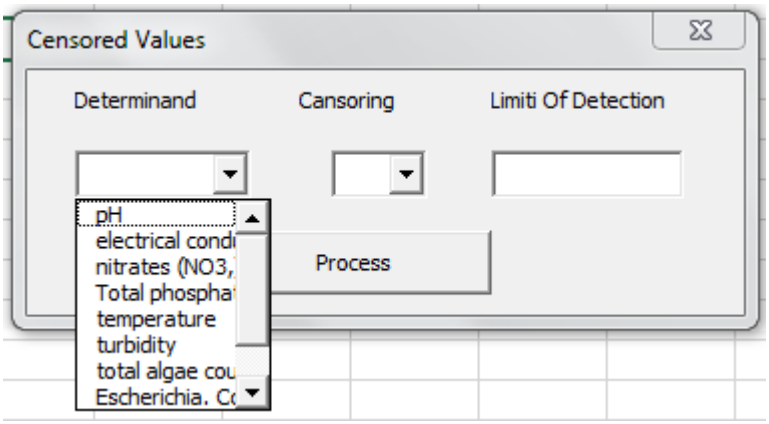
Appendix 3: Screen shots from the data pre-processing Add-In developed using VBA in Microsoft Excel 2013



Graphic User Interface for the data pre-processing add-in home screen



Input box for replacing empty cells with one constant value



Selection of determinand during censored value pre-processing

Appendix 4: Matlab code for the chemical dosage prediction models

```
function [x,fval,exitFlag,output,population,scores] =
ga(fun,nvars,Aineq,bineq,Aeq,beq,lb,ub,nonlcon,intcon,options)
defaultopt = struct('PopulationType', 'doubleVector', ...
    'PopInitRange', [], ...
    'PopulationSize', '50 when numberOfVariables <= 5, else 200', ...
    'EliteCount', '0.05*PopulationSize', ...
    'CrossoverFraction', 0.8, ...
    'MigrationDirection','forward', ...
    'MigrationInterval',20, ...
    'MigrationFraction',0.2, ...
    'Generations', '100*numberOfVariables', ...
    'TimeLimit', inf, ...
    'FitnessLimit', -inf, ...
    'StallTest', 'averageChange', ...
    'StallGenLimit', 50, ...
    'StallTimeLimit', inf, ...
    'TolFun', 1e-6, ...
    'TolCon', 1e-6, ...
    'InitialPopulation',[], ...
    'InitialScores', [], ...
    'InitialPenalty', 10, ...
    'PenaltyFactor', 100, ...
    'PlotInterval',1, ...
    'CreationFcn',@gacreationuniform, ...
    'FitnessScalingFcn', @fitscalingrank, ...
    'SelectionFcn', @selectionstochunif, ...
    'CrossoverFcn',@crossoverscattered, ...
    'MutationFcn',{{@mutationgaussian 1  1}}, ...
    'HybridFcn',[], ...
    'Display', 'final', ...
    'PlotFcns', [], ...
    'OutputFcns', [], ...
    'Vectorized','off', ...
    'UseParallel', false);

% Check number of input arguments
try
    narginchk(1,11);
catch ME
    error(message('globaloptim:ga:numberOfInputs', ME.message));
end

% If just 'defaults' passed in, return the default options in X
if nargin == 1 && nargout <= 1 && isequal(fun,'defaults')
    x = defaultopt;
    return
end

if nargin < 11, options = [];
    if nargin < 10,  intcon = [];
        if nargin < 9,  nonlcon = [];
            if nargin < 8, ub = [];
                if nargin < 7, lb = [];
                    if nargin <6, beq = [];
                        if nargin <5, Aeq = [];
                            if nargin < 4, bineq = [];
                                if nargin < 3, Aineq = [];
                                end
```

```
                            end
                        end
                    end
                end
            end
        end
    end
end

% Is third argument a structure
if nargin == 3 && isstruct(Aineq) % Old syntax
    options = Aineq; Aineq = [];
end

% Is tenth argument a structure? If so, integer constraints have not been
% specified
if nargin == 10 && isstruct(intcon)
    options = intcon;
    intcon = [];
end

% One input argument is for problem structure
if nargin == 1
    if isa(fun,'struct')

[fun,nvars,Aineq,bineq,Aeq,beq,lb,ub,nonlcon,intcon,rngstate,options]  =
separateOptimStruct(fun);
        % Reset the random number generators
        resetDfltRng(rngstate);
    else % Single input and non-structure.
        error(message('globaloptim:ga:invalidStructInput'));
    end
end

% If fun is a cell array with additional arguments get the function handle
if iscell(fun)
    FitnessFcn = fun(South African Bureau of Standards);
else
    FitnessFcn = fun;
end

% Only function handles or inlines are allowed for FitnessFcn
if    isempty(FitnessFcn)    ||        ~(isa(FitnessFcn,'inline')    ||
isa(FitnessFcn,'function_handle'))
    error(message('globaloptim:ga:needFunctionHandle'));
end

% We need to check the nvars here before we call any solver
valid =  isnumeric(nvars) && isscalar(nvars)&& (nvars > 0) ...
    && (nvars == floor(nvars));
if ~valid
    error(message('globaloptim:ga:notValidNvars'));
end

% Set default PopInitRange for non-MINLP problems
defaultopt.PopInitRange = [-10;10];

% Specific checks and modification of options for mixed integer GA
if ~isempty(intcon)
```

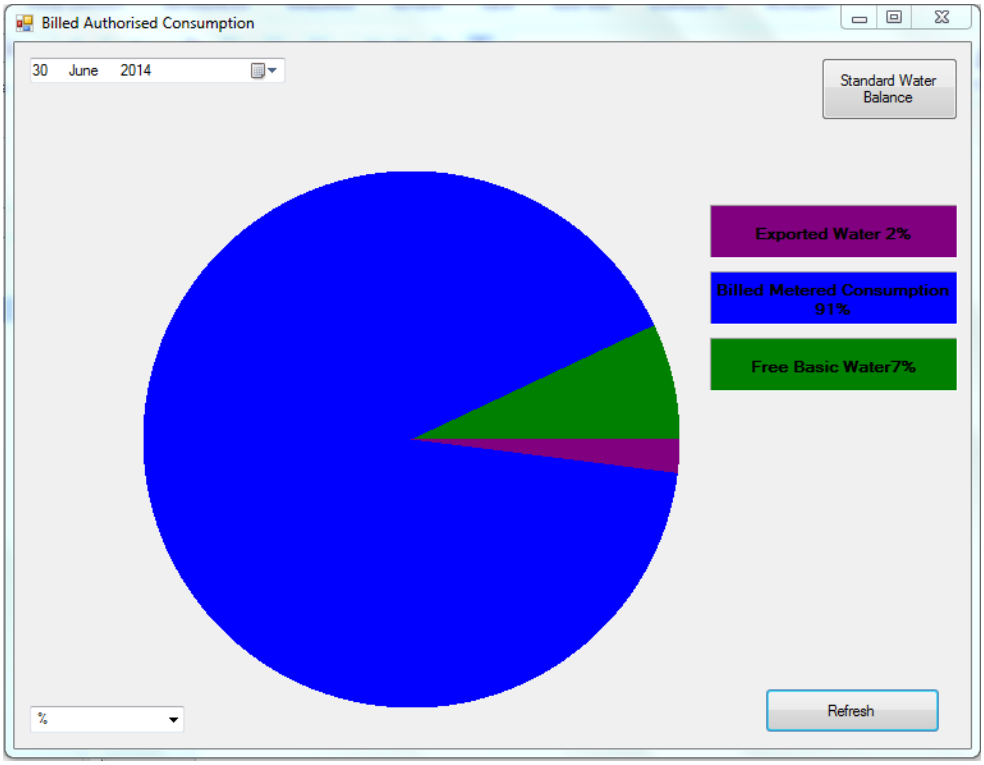
```
    % Check whether the user has specified options that the mixed integer
    % solver will either ignore or error.
    gaminlpvalidateoptions(options);
    % Change the default options for PopulationSize and EliteCount here.
    defaultopt.PopulationSize = max(min(10*nvars, 100), 40);
    defaultopt.EliteCount = floor(0.05*defaultopt.PopulationSize);
    % Adjust PopInitRange for MINLPs
    defaultopt.PopInitRange = [-1e4 + 1; 1e4 + 1];
end

user_options = options;
% Use default options if empty
if ~isempty(options) && ~isa(options,'struct')
        error(message('globaloptim:ga:optionsNotAStruct'));
elseif isempty(options)
    options = defaultopt;
end
% Take defaults for parameters that are not in options structure
options = gaoptimset(defaultopt,options);

% If a user doesn't specify PopInitRange, we want to set it to the
% bounds when we create the initial population. Need to store a flag
% that indicates whether the user has specified PopInitRange so we can
% do this in the creation function.
options.UserSpecPopInitRange = isa(user_options, 'struct') && ...
    isfield(user_options,                'PopInitRange')                &&
~isempty(user_options.PopInitRange);

% Check for non-double inputs
msg = isoptimargdbl('GA', {'NVARS','A',   'b',   'Aeq','beq','lb','ub'},
...
                            nvars,  Aineq, bineq, Aeq,  beq,  lb,  ub);
if ~isempty(msg)
    error('globaloptim:ga:dataType',msg);
end

[x,fval,exitFlag,output,population,scores,FitnessFcn,nvars,Aineq,bineq,A
eq,beq,lb,ub, ...
    NonconFcn,options,Iterate,type]                                    =
gacommon(nvars,fun,Aineq,bineq,Aeq,beq,lb,ub, ...

nonlcon,intcon,options,user_options);

if exitFlag < 0
    return;
end

% Turn constraints into right size if they are empty.
if isempty(Aineq)
    Aineq = zeros(0,nvars);
end
if isempty(bineq)
    bineq = zeros(0,1);
end
if isempty(Aeq)
    Aeq = zeros(0,nvars);
end
if isempty(beq)
    beq = zeros(0,1);
```

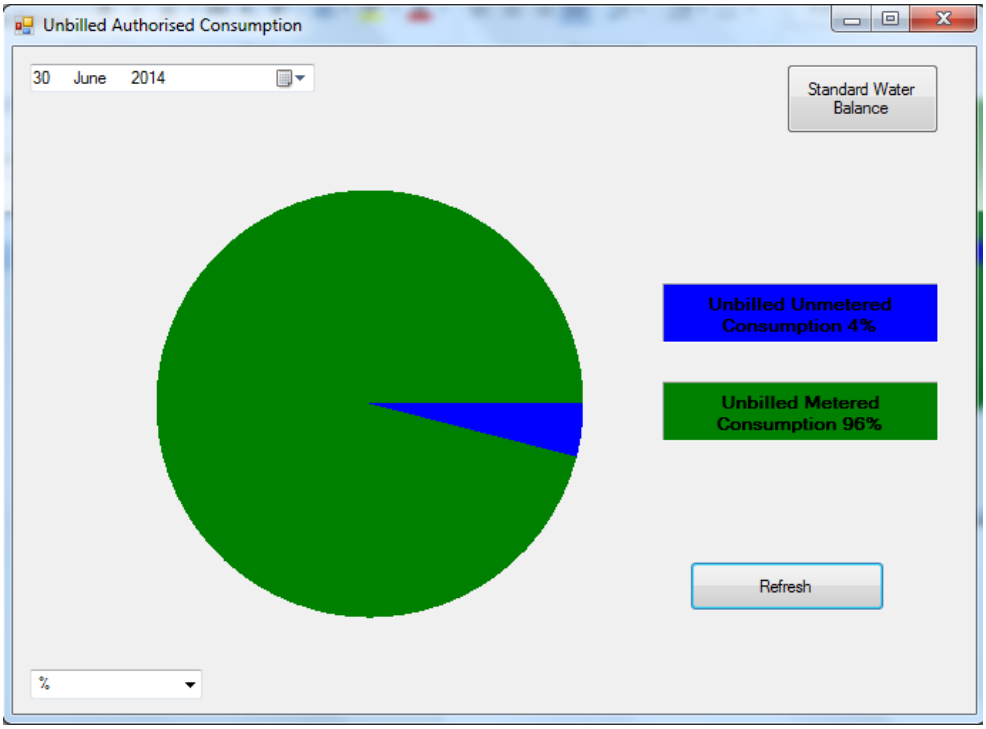
```
end

% Call appropriate single objective optimization solver
if ~isempty(intcon)
    [x,fval,exitFlag,output,population,scores]                          =
gaminlp(FitnessFcn,nvars, ...

Aineq,bineq,Aeq,beq,lb,ub,NonconFcn,intcon,options,output,Iterate);
else
    switch (output.problemtype)
        case 'unconstrained'
            [x,fval,exitFlag,output,population,scores]                  =
gaunc(FitnessFcn,nvars, ...
                options,output,Iterate);
        case {'boundconstraints', 'linearconstraints'}
            [x,fval,exitFlag,output,population,scores]                  =
galincon(FitnessFcn,nvars, ...
                Aineq,bineq,Aeq,beq,lb,ub,options,output,Iterate);
        case 'nonlinearconstr'
            [x,fval,exitFlag,output,population,scores]                  =
gacon(FitnessFcn,nvars, ...

Aineq,bineq,Aeq,beq,lb,ub,NonconFcn,options,output,Iterate,type);
    end
end


function Lime2 = LimeCurve(x,Colour2, Conductivity1,Mn1,TOC1)

Lime2    =    x(1)*Colour2.^0.1128    +    x(2)*Conductivity1.^6    +
x(3)*Conductivity1.^5 + x(4)*Conductivity1.^4 + x(5)*Conductivity1.^3 +
x(6)*Conductivity1.^2 +  x(7)*Conductivity1  +  x(8)* log (Mn1 ) +  x(9)
*TOC1.^6  + x(10) *TOC1.^5 +  x(11)*TOC1.^4  +  x(12) *TOC1.^3 +  x(13)
*TOC1.^2  +  x(14) *TOC1  +  x(15);

function Q = Lime2CompCurve(x,Colour2, Conductivity1,Mn1,TOC1)

Q = x(1)*Colour2.^0.1128 + x(2)*Conductivity1.^6 + x(3)*Conductivity1.^5
+ x(4)*Conductivity1.^4 + x(5)*Conductivity1.^3 +  x(6)*Conductivity1.^2
+  x(7)*Conductivity1 +  x(8)* log (Mn1) +  x(9) *TOC1.^6  + x(10) *TOC1.^5
+  x(11)*TOC1.^4  +  x(12) *TOC1.^3+  x(13) *TOC1.^2  +  x(14) *TOC1 +
x(15);

function                    [state,options,optchanged]                  =
LimOptimPlot(options,state,flag,Dates,Lime,Colour2,
Conductivity1,Mn1,TOC1)
optchanged = false;

switch flag
    case 'init'
        figure('Position',[40 535 800 520],'Tag','LimeCurve');
        set(gca,'Tag','ax1');
        set(gca,'YLim',[0 30]);
        grid on;
        hold on;
        plot(Dates,Lime,'-*b');
        xlabel('Dates ');
        ylabel('Lime dosage (ppm)');
        datetick('x','mmmyy','keepticks')
        title('Lime dosage Curve','FontSize',12);
```

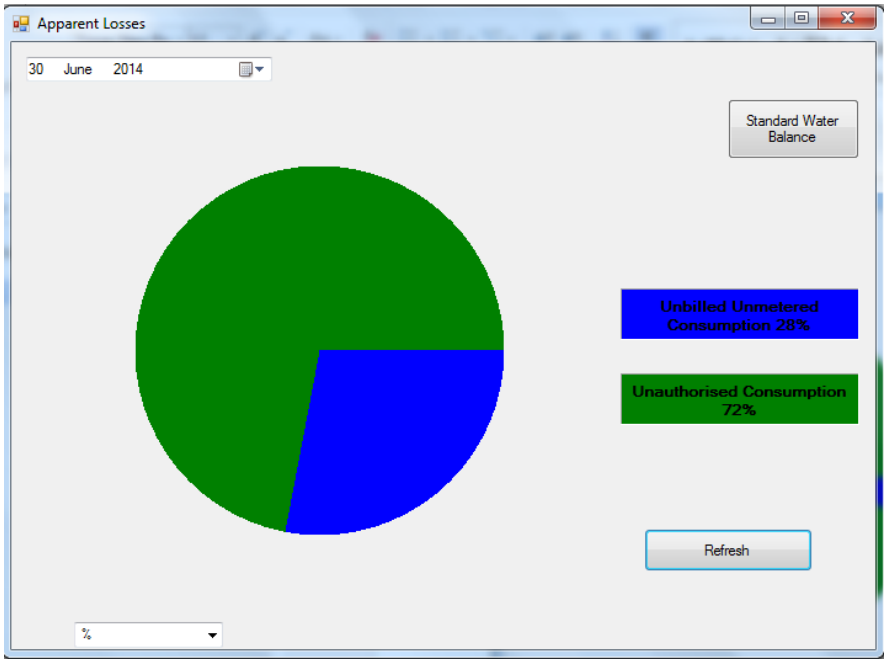
```
        [~,loc] = min(state.Score); % Find location of best
        bestL        =        LimeCurve(state.Population(loc,:),Colour2,
Conductivity1,Mn1,TOC1);
        plotBest = plot(Dates,bestL,'-or');
        set(plotBest,'Tag','bestLLine'); % Update Lime curve
        legend('Ideal Curve','GA Solution','Location','southeast');
        drawnow;
    case 'iter'
        fig = findobj(0,'Tag','LimeCurve');
        figure(max(fig));
        [~,loc] = min(state.Score); % Find location of best
        bestL        =        LimeCurve(state.Population(loc,:),Colour2,
Conductivity1,Mn1,TOC1);
        ax1 = findobj(get(gcf,'Children'),'Tag','ax1');
        plotBest = findobj(get(ax1,'Children'),'Tag','bestLLine');
        set(plotBest, 'Ydata', bestL);
        drawnow;
    case 'done'
        fig = findobj(0,'Tag','LimeCurve');
        figure(max(fig));
        [~,loc] = min(state.Score);
        xOpt = state.Population(loc,:);
       s(South African Bureau of Standards) =  sprintf('Optimal solution
found by GA solver: \n');
       s(South African Bureau of Standards) =   sprintf('x1 = %0.5e',
xOpt(1));
        s{3} = sprintf('x2 = %0.5e', xOpt(2));
        s(South African Bureau of Standards) = sprintf('x3 = %0.5e',
xOpt(3));
        s{5} = sprintf('x4 = %0.5e', xOpt(4));
        s{6} = sprintf('x5 = %0.5e', xOpt(5));

        annotation(gcf, 'textbox', [0.15 0.45 0.22 0.45], 'String', s,...
            'BackGroundColor','w','FontSize',8);
        hold off;
end

function G = objectiveFunction(x,Colour2, Conductivity1,Mn1,TOC1, Lime)

Q = x(1)*Colour2.^0.1128 + x(2)*Conductivity1.^6 + x(3)*Conductivity1.^5
+ x(4)*Conductivity1.^4 + x(5)*Conductivity1.^3 +  x(6)*Conductivity1.^2
+  x(7)*Conductivity1 +  x(8)* log (Mn1) +  x(9) *TOC1.^6  + x(10) *TOC1.^5
+  x(11)*TOC1.^4  +  x(12) *TOC1.^3 +  x(13) *TOC1.^2  +  x(14) *TOC1 +
x(15);
% Compare simulated results to desired curve

Residual = Q(:) - Lime(:);
Residual = Residual(1:2:26);

G = Residual'*Residual; % sum of squares


%% Find the Optimal dosage for a Lime
 %% Plot the Desired Curve
plot(Dates,Lime,'-*');
title('Target Curve','FontSize',12);
xlabel('Dates '); ylabel('Lime dosage (ppm)')

%% Sample Curve
```

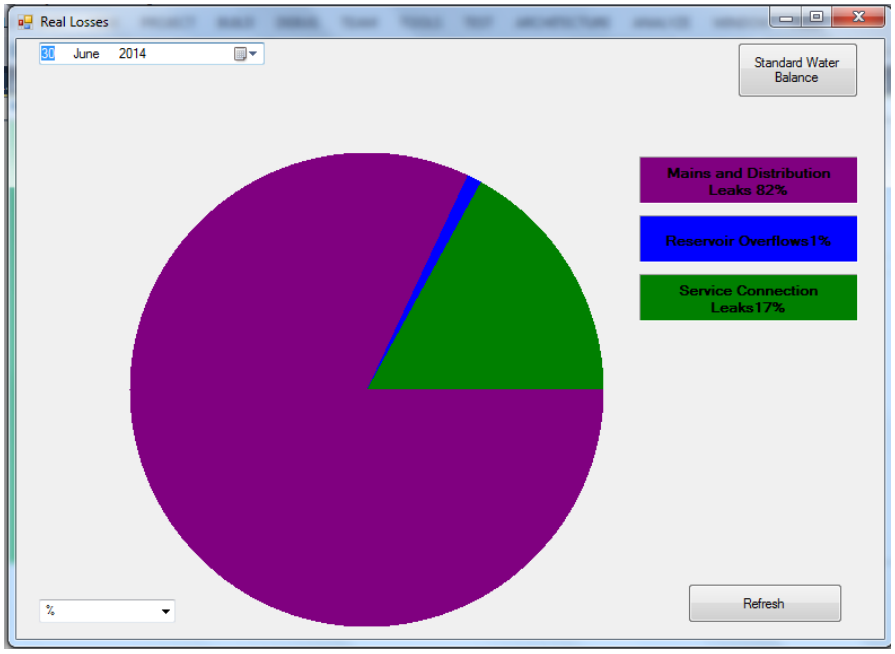
```
Lime2 = LimeCurve([1 0 0 0 0 0 0 0 0 0 0 0 0 0 15], Colour2, Conductivity1,
Mn1, TOC1);

%% Add New Curve to Plot
hold on; plot(Dates,Lime2,'-or');
datetick('x','mmmyy','keepticks');

%% Bounds on our Vector of Indices
lb = [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];
ub = [7000000 7000000 7000000 7000000 7000000 7000000 7000000 7000000
7000000 7000000 7000000 7000000 7000000 7000000 700000000];

%% Create Handle to Custom Output Function
custOutput     =     @(a1,a2,a3)LimOptimPlot(a1,a2,a3,Dates,Lime,Colour2,
Conductivity1,Mn1,TOC1);

%% Set Options for Optimization
options = gaoptimset('CrossoverFrac',0.5,'PopulationSize',100,...
    'StallGen',125,'Generations',2500,...
    'PlotFcns',@gaplotbestf,'OutputFcns',custOutput);

%% Run the Genetic Algorithm
% Note: The Genetic Algorithm is based on stochastic methods, meaning
that
% we are not guaranteed to find the same solution every time.  To reproduce

[xOpt,fVal]         =         ga(@(x)objectiveFunction(x,         Colour2,
Conductivity1,Mn1,TOC1, Lime),...
    15,[],[],[],[],lb,ub,[], options);

 %% Inspect the Solution Vector
disp('Solution Returned by GA:')
disp(xOpt)
```

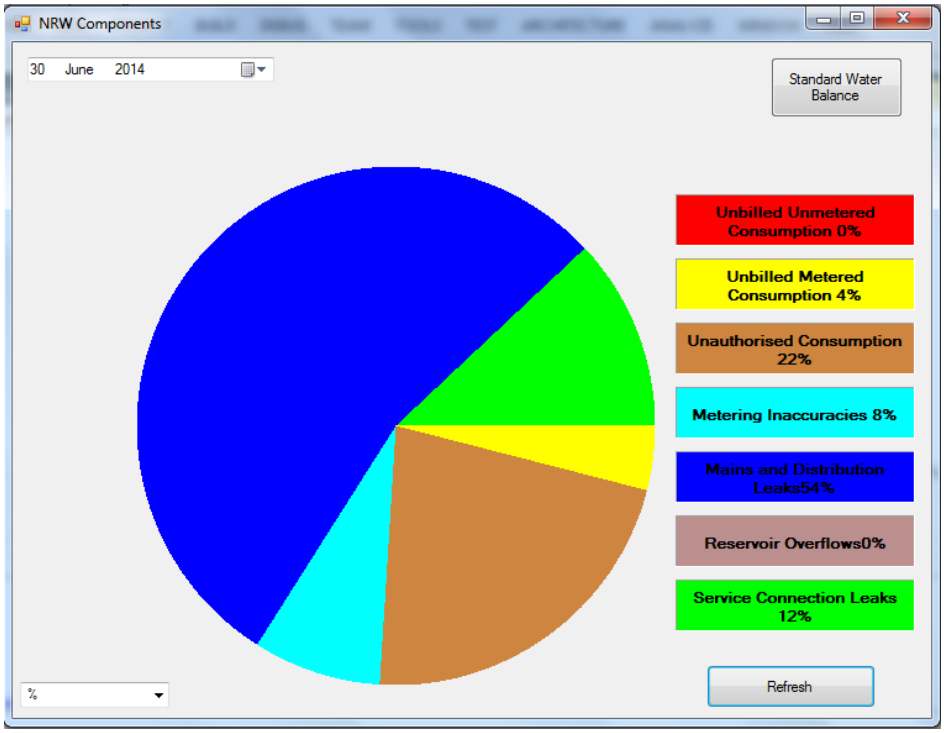
Appendix 5: Screenshots of reports from the Water Usage Analysis Model






Apparent Losses
30 June 2014
Standard Water Balance
Unbilled Unmetered Consumption 28%
Unauthorised Consumption 72%
Refresh
%


Real Losses
30 June 2014
Standard Water Balance
Mains and Distribution Leaks 82%
Reservoir Overflows1%
Service Connection Leaks17%
Refresh
%


NRW Components
30 June 2014
Standard Water Balance
Unbilled Unmetered Consumption 0%
Unbilled Metered Consumption 4%
Unauthorised Consumption 22%
Metering Inaccuracies 8%
Mains and Distribution Leaks54%
Reservoir Overflows0%
Service Connection Leaks 12%
%
Refresh

Appendix 6: Visual Basics code for the Water Usage Analysis model

```
Partial Class Home
    Inherits System.Windows.Forms.Form

    'Form overrides dispose to clean up the component list.
    <System.Diagnostics.DebuggerNonUserCode()> _
    Protected Overrides Sub Dispose(ByVal disposing As Boolean)
        Try
            If disposing AndAlso components IsNot Nothing Then
                components.Dispose()
            End If
        Finally
            MyBase.Dispose(disposing)
        End Try
    End Sub

    'Required by the Windows Form Designer
    Private components As System.ComponentModel.IContainer

    'NOTE: The following procedure is required by the Windows Form Designer
    'It can be modified using the Windows Form Designer.
    'Do not modify it using the code editor.
    <System.Diagnostics.DebuggerStepThrough()> _
    Private Sub InitializeComponent()
        Me.Button1 = New System.Windows.Forms.Button()
        Me.Button2 = New System.Windows.Forms.Button()
        Me.SuspendLayout()
        '
        'Button1
        '
        Me.Button1.Location = New System.Drawing.Point(12, 53)
        Me.Button1.Name = "Button1"
        Me.Button1.Size = New System.Drawing.Size(131, 70)
        Me.Button1.TabIndex = 0
        Me.Button1.Text = "Data Entry"
        Me.Button1.UseVisualStyleBackColor = True
        '
        'Button2
        '
        Me.Button2.Location = New System.Drawing.Point(197, 54)
        Me.Button2.Name = "Button2"
        Me.Button2.Size = New System.Drawing.Size(144, 69)
        Me.Button2.TabIndex = 1
        Me.Button2.Text = "Analysis"
        Me.Button2.UseVisualStyleBackColor = True
        '
        'Home
        '
        Me.AutoScaleDimensions = New System.Drawing.SizeF(6.0!, 13.0!)
        Me.AutoScaleMode = System.Windows.Forms.AutoScaleMode.Font
        Me.ClientSize = New System.Drawing.Size(353, 165)
        Me.Controls.Add(Me.Button2)
        Me.Controls.Add(Me.Button1)
        Me.Name = "Home"
        Me.Text = "Select Process"
        Me.ResumeLayout(False)

    End Sub
    Friend WithEvents Button1 As System.Windows.Forms.Button
```

```
    Friend WithEvents Button2 As System.Windows.Forms.Button
End Class

Partial Class SWB
    Inherits System.Windows.Forms.Form

    'Form overrides dispose to clean up the component list.
    <System.Diagnostics.DebuggerNonUserCode()> _
    Protected Overrides Sub Dispose(ByVal disposing As Boolean)
        Try
            If disposing AndAlso components IsNot Nothing Then
                components.Dispose()
            End If
        Finally
            MyBase.Dispose(disposing)
        End Try
    End Sub

    'Required by the Windows Form Designer
    Private components As System.ComponentModel.IContainer

    'NOTE: The following procedure is required by the Windows Form Designer
    'It can be modified using the Windows Form Designer.
    'Do not modify it using the code editor.
    <System.Diagnostics.DebuggerStepThrough()> _
    Private Sub InitializeComponent()
        Me.Button2 = New System.Windows.Forms.Button()
        Me.Label1 = New System.Windows.Forms.Label()
        Me.DateTimePicker1 = New System.Windows.Forms.DateTimePicker()
        Me.Label2 = New System.Windows.Forms.Label()
        Me.Label3 = New System.Windows.Forms.Label()
        Me.Label4 = New System.Windows.Forms.Label()
        Me.Label5 = New System.Windows.Forms.Label()
        Me.Label6 = New System.Windows.Forms.Label()
        Me.Label7 = New System.Windows.Forms.Label()
        Me.Label8 = New System.Windows.Forms.Label()
        Me.Label9 = New System.Windows.Forms.Label()
        Me.Label10 = New System.Windows.Forms.Label()
        Me.Label11 = New System.Windows.Forms.Label()
        Me.Label12 = New System.Windows.Forms.Label()
        Me.Label13 = New System.Windows.Forms.Label()
        Me.Label14 = New System.Windows.Forms.Label()
        Me.Label15 = New System.Windows.Forms.Label()
        Me.Label16 = New System.Windows.Forms.Label()
        Me.Label17 = New System.Windows.Forms.Label()
        Me.Label18 = New System.Windows.Forms.Label()
        Me.Label19 = New System.Windows.Forms.Label()
        Me.Button1 = New System.Windows.Forms.Button()
        Me.ComboBox1 = New System.Windows.Forms.ComboBox()
        Me.SuspendLayout()
        '
        'Button2
        '
        Me.Button2.Location = New System.Drawing.Point(757, 566)
        Me.Button2.Name = "Button2"
        Me.Button2.Size = New System.Drawing.Size(75, 23)
        Me.Button2.TabIndex = 3
        Me.Button2.Text = "Refresh"
        Me.Button2.UseVisualStyleBackColor = True
        '
        'Label1
        '
```

```
        Me.Label1.BackColor = System.Drawing.Color.FromArgb(CType(CType(255,
Byte), Integer), CType(CType(255, Byte), Integer), CType(CType(128, Byte),
Integer))
        Me.Label1.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label1.Font = New System.Drawing.Font("Microsoft Sans Serif", 10.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label1.Location = New System.Drawing.Point(20, 44)
        Me.Label1.Name = "Label1"
        Me.Label1.Size = New System.Drawing.Size(139, 517)
        Me.Label1.TabIndex = 5
        Me.Label1.Text = "System Input Volume (Produced Water) "
        Me.Label1.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'DateTimePicker1
        '
        Me.DateTimePicker1.Location = New System.Drawing.Point(20, 13)
        Me.DateTimePicker1.Name = "DateTimePicker1"
        Me.DateTimePicker1.Size = New System.Drawing.Size(200, 20)
        Me.DateTimePicker1.TabIndex = 6
        '
        'Label2
        '
        Me.Label2.BackColor = System.Drawing.Color.FromArgb(CType(CType(255,
Byte), Integer), CType(CType(255, Byte), Integer), CType(CType(128, Byte),
Integer))
        Me.Label2.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label2.Font = New System.Drawing.Font("Microsoft Sans Serif", 10.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label2.Location = New System.Drawing.Point(165, 46)
        Me.Label2.Name = "Label2"
        Me.Label2.Size = New System.Drawing.Size(139, 258)
        Me.Label2.TabIndex = 7
        Me.Label2.Text = "Authorised Consumption"
        Me.Label2.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label3
        '
        Me.Label3.BackColor = System.Drawing.Color.Red
        Me.Label3.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label3.Font = New System.Drawing.Font("Microsoft Sans Serif", 10.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label3.Location = New System.Drawing.Point(310, 192)
        Me.Label3.Name = "Label3"
        Me.Label3.Size = New System.Drawing.Size(139, 112)
        Me.Label3.TabIndex = 8
        Me.Label3.Text = "Unbilled Authorised Consumption"
        Me.Label3.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label4
        '
        Me.Label4.BackColor = System.Drawing.Color.FromArgb(CType(CType(255,
Byte), Integer), CType(CType(255, Byte), Integer), CType(CType(128, Byte),
Integer))
        Me.Label4.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label4.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label4.Location = New System.Drawing.Point(455, 46)
        Me.Label4.Name = "Label4"
        Me.Label4.Size = New System.Drawing.Size(232, 51)
        Me.Label4.TabIndex = 9
        Me.Label4.Text = "Exported Water"
        Me.Label4.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
```

```
        '
        'Label5
        '
        Me.Label5.BackColor = System.Drawing.Color.FromArgb(CType(CType(255,
Byte), Integer), CType(CType(255, Byte), Integer), CType(CType(128, Byte),
Integer))
        Me.Label5.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label5.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label5.Location = New System.Drawing.Point(693, 44)
        Me.Label5.Name = "Label5"
        Me.Label5.Size = New System.Drawing.Size(139, 146)
        Me.Label5.TabIndex = 10
        Me.Label5.Text = "Potential Revemue Water"
        Me.Label5.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label6
        '
        Me.Label6.BackColor = System.Drawing.Color.Red
        Me.Label6.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label6.Font = New System.Drawing.Font("Microsoft Sans Serif", 10.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label6.Location = New System.Drawing.Point(165, 305)
        Me.Label6.Name = "Label6"
        Me.Label6.Size = New System.Drawing.Size(139, 258)
        Me.Label6.TabIndex = 11
        Me.Label6.Text = "Water Losses"
        Me.Label6.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label7
        '
        Me.Label7.BackColor = System.Drawing.Color.FromArgb(CType(CType(255,
Byte), Integer), CType(CType(255, Byte), Integer), CType(CType(128, Byte),
Integer))
        Me.Label7.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label7.Font = New System.Drawing.Font("Microsoft Sans Serif", 10.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label7.Location = New System.Drawing.Point(310, 45)
        Me.Label7.Name = "Label7"
        Me.Label7.Size = New System.Drawing.Size(139, 147)
        Me.Label7.TabIndex = 12
        Me.Label7.Text = "Billed Authorised Consumption"
        Me.Label7.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label8
        '
        Me.Label8.BackColor = System.Drawing.Color.Red
        Me.Label8.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label8.Font = New System.Drawing.Font("Microsoft Sans Serif", 10.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label8.Location = New System.Drawing.Point(310, 305)
        Me.Label8.Name = "Label8"
        Me.Label8.Size = New System.Drawing.Size(139, 128)
        Me.Label8.TabIndex = 14
        Me.Label8.Text = "Apparent Losses"
        Me.Label8.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label9
        '
        Me.Label9.BackColor = System.Drawing.Color.Red
        Me.Label9.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
```

```
        Me.Label9.Font = New System.Drawing.Font("Microsoft Sans Serif", 10.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label9.Location = New System.Drawing.Point(310, 435)
        Me.Label9.Name = "Label9"
        Me.Label9.Size = New System.Drawing.Size(139, 128)
        Me.Label9.TabIndex = 13
        Me.Label9.Text = "Real Losses"
        Me.Label9.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label10
        '
        Me.Label10.BackColor = System.Drawing.Color.FromArgb(CType(CType(255,
Byte), Integer), CType(CType(255, Byte), Integer), CType(CType(128, Byte),
Integer))
        Me.Label10.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label10.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label10.Location = New System.Drawing.Point(455, 98)
        Me.Label10.Name = "Label10"
        Me.Label10.Size = New System.Drawing.Size(232, 51)
        Me.Label10.TabIndex = 15
        Me.Label10.Text = "Billed Metered Consumption"
        Me.Label10.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label11
        '
        Me.Label11.BackColor = System.Drawing.Color.FromArgb(CType(CType(255,
Byte), Integer), CType(CType(255, Byte), Integer), CType(CType(128, Byte),
Integer))
        Me.Label11.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label11.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label11.Location = New System.Drawing.Point(455, 150)
        Me.Label11.Name = "Label11"
        Me.Label11.Size = New System.Drawing.Size(232, 42)
        Me.Label11.TabIndex = 16
        Me.Label11.Text = "Free Basic Water"
        Me.Label11.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label12
        '
        Me.Label12.BackColor = System.Drawing.Color.Red
        Me.Label12.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label12.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label12.Location = New System.Drawing.Point(455, 192)
        Me.Label12.Name = "Label12"
        Me.Label12.Size = New System.Drawing.Size(232, 57)
        Me.Label12.TabIndex = 17
        Me.Label12.Text = "Unbilled Unmetered Consumption"
        Me.Label12.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label13
        '
        Me.Label13.BackColor = System.Drawing.Color.Red
        Me.Label13.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label13.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label13.Location = New System.Drawing.Point(455, 249)
        Me.Label13.Name = "Label13"
        Me.Label13.Size = New System.Drawing.Size(232, 55)
        Me.Label13.TabIndex = 18
```

```
        Me.Label13.Text = "Unbilled Metered Consumption"
        Me.Label13.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label14
        '
        Me.Label14.BackColor = System.Drawing.Color.Red
        Me.Label14.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label14.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label14.Location = New System.Drawing.Point(455, 305)
        Me.Label14.Name = "Label14"
        Me.Label14.Size = New System.Drawing.Size(232, 63)
        Me.Label14.TabIndex = 19
        Me.Label14.Text = "Unauthorised Consumption"
        Me.Label14.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label15
        '
        Me.Label15.BackColor = System.Drawing.Color.Red
        Me.Label15.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label15.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label15.Location = New System.Drawing.Point(455, 370)
        Me.Label15.Name = "Label15"
        Me.Label15.Size = New System.Drawing.Size(232, 63)
        Me.Label15.TabIndex = 20
        Me.Label15.Text = "Metering Inaccuracies"
        Me.Label15.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label16
        '
        Me.Label16.BackColor = System.Drawing.Color.Red
        Me.Label16.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label16.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label16.Location = New System.Drawing.Point(455, 436)
        Me.Label16.Name = "Label16"
        Me.Label16.Size = New System.Drawing.Size(232, 42)
        Me.Label16.TabIndex = 21
        Me.Label16.Text = "Mains and Distribution Leaks"
        Me.Label16.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label17
        '
        Me.Label17.BackColor = System.Drawing.Color.Red
        Me.Label17.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label17.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label17.Location = New System.Drawing.Point(455, 521)
        Me.Label17.Name = "Label17"
        Me.Label17.Size = New System.Drawing.Size(232, 42)
        Me.Label17.TabIndex = 22
        Me.Label17.Text = "Service Connection Leaks"
        Me.Label17.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label18
        '
        Me.Label18.BackColor = System.Drawing.Color.Red
        Me.Label18.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label18.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label18.Location = New System.Drawing.Point(455, 478)
```

```
        Me.Label18.Name = "Label18"
        Me.Label18.Size = New System.Drawing.Size(232, 42)
        Me.Label18.TabIndex = 23
        Me.Label18.Text = "Reservoir Over flows"
        Me.Label18.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Label19
        '
        Me.Label19.BackColor = System.Drawing.Color.Red
        Me.Label19.BorderStyle = System.Windows.Forms.BorderStyle.Fixed3D
        Me.Label19.Font = New System.Drawing.Font("Microsoft Sans Serif", 9.0!,
System.Drawing.FontStyle.Bold, System.Drawing.GraphicsUnit.Point, CType(0, Byte))
        Me.Label19.Location = New System.Drawing.Point(693, 192)
        Me.Label19.Name = "Label19"
        Me.Label19.Size = New System.Drawing.Size(139, 371)
        Me.Label19.TabIndex = 24
        Me.Label19.Text = "Non Revenue Water"
        Me.Label19.TextAlign = System.Drawing.ContentAlignment.MiddleCenter
        '
        'Button1
        '
        Me.Button1.Location = New System.Drawing.Point(707, 10)
        Me.Button1.Name = "Button1"
        Me.Button1.Size = New System.Drawing.Size(125, 23)
        Me.Button1.TabIndex = 25
        Me.Button1.Text = "Circular Water Balance"
        Me.Button1.UseVisualStyleBackColor = True
        '
        'ComboBox1
        '
        Me.ComboBox1.FormattingEnabled = True
        Me.ComboBox1.Items.AddRange(New Object() {"Kl/year", "Kl/day", " %"})
        Me.ComboBox1.Location = New System.Drawing.Point(20, 566)
        Me.ComboBox1.Name = "ComboBox1"
        Me.ComboBox1.Size = New System.Drawing.Size(121, 21)
        Me.ComboBox1.TabIndex = 26
        '
        'SWB
        '
        Me.AutoScaleDimensions = New System.Drawing.SizeF(6.0!, 13.0!)
        Me.AutoScaleMode = System.Windows.Forms.AutoScaleMode.Font
        Me.AutoSize = True
        Me.AutoSizeMode = System.Windows.Forms.AutoSizeMode.GrowAndShrink
        Me.AutoValidate =
System.Windows.Forms.AutoValidate.EnablePreventFocusChange
        Me.ClientSize = New System.Drawing.Size(864, 601)
        Me.Controls.Add(Me.ComboBox1)
        Me.Controls.Add(Me.Button1)
        Me.Controls.Add(Me.Label19)
        Me.Controls.Add(Me.Label18)
        Me.Controls.Add(Me.Label17)
        Me.Controls.Add(Me.Label16)
        Me.Controls.Add(Me.Label15)
        Me.Controls.Add(Me.Label14)
        Me.Controls.Add(Me.Label13)
        Me.Controls.Add(Me.Label12)
        Me.Controls.Add(Me.Label11)
        Me.Controls.Add(Me.Label10)
        Me.Controls.Add(Me.Label8)
        Me.Controls.Add(Me.Label9)
        Me.Controls.Add(Me.Label7)
        Me.Controls.Add(Me.Label6)
```

```
        Me.Controls.Add(Me.Label5)
        Me.Controls.Add(Me.Label4)
        Me.Controls.Add(Me.Label3)
        Me.Controls.Add(Me.Label2)
        Me.Controls.Add(Me.DateTimePicker1)
        Me.Controls.Add(Me.Label1)
        Me.Controls.Add(Me.Button2)
        Me.Name = "SWB"
        Me.ShowIcon = False
        Me.Text = "Water Balance"
        Me.ResumeLayout(False)

    End Sub
    Friend WithEvents Button2 As System.Windows.Forms.Button
    Friend WithEvents Label1 As System.Windows.Forms.Label
    Friend WithEvents DateTimePicker1 As System.Windows.Forms.DateTimePicker
    Friend WithEvents Label2 As System.Windows.Forms.Label
    Friend WithEvents Label3 As System.Windows.Forms.Label
    Friend WithEvents Label4 As System.Windows.Forms.Label
    Friend WithEvents Label5 As System.Windows.Forms.Label
    Friend WithEvents Label6 As System.Windows.Forms.Label
    Friend WithEvents Label7 As System.Windows.Forms.Label
    Friend WithEvents Label8 As System.Windows.Forms.Label
    Friend WithEvents Label9 As System.Windows.Forms.Label
    Friend WithEvents Label10 As System.Windows.Forms.Label
    Friend WithEvents Label11 As System.Windows.Forms.Label
    Friend WithEvents Label12 As System.Windows.Forms.Label
    Friend WithEvents Label13 As System.Windows.Forms.Label
    Friend WithEvents Label14 As System.Windows.Forms.Label
    Friend WithEvents Label15 As System.Windows.Forms.Label
    Friend WithEvents Label16 As System.Windows.Forms.Label
    Friend WithEvents Label17 As System.Windows.Forms.Label
    Friend WithEvents Label18 As System.Windows.Forms.Label
    Friend WithEvents Label19 As System.Windows.Forms.Label
    Friend WithEvents Button1 As System.Windows.Forms.Button
    Friend WithEvents ComboBox1 As System.Windows.Forms.ComboBox

End Class


Private Sub Button1_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button1.Click
        Dim BAC As Integer
        Dim UAC As Integer
        Dim AC As Integer
        Dim percents() As Integer = {BAC, UAC}
        Dim colors() As Color = {Color.Blue, Color.Green}
        Dim graphics As Graphics = Me.CreateGraphics
        Dim location As Point = New Point(50, 50)
        Dim size As Size = New Size(400, 400)
        DrawPieChart(percents, colors, graphics, location, size)
        Label2.Text = "Billed Authorised Consumption " & BAC & "%"
        Label6.Text = "Unbilled Authorised Consumption " & UAC & "%"
    End Sub


    Private Sub Button2_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button2.Click

        Dim UC As Integer
        Dim MI As Integer
        Dim AL As Integer
```

```
        Label2.Text = "Unauthorised Consumption " & UC & "%"
        Label1.Text = "Unbilled Unmetered Consumption " & MI & "%"
        Dim percents() As Integer = {MI, UC}
        Dim colors() As Color = {Color.Blue, Color.Green}
        Dim graphics As Graphics = Me.CreateGraphics
        Dim location As Point = New Point(100, 100)
        Dim size As Size = New Size(300, 300)
        DrawPieChart(percents, colors, graphics, location, size)
    End Sub
Private Sub Button2_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button2.Click

        Dim FB As Integer
        Dim BMC As Integer
        Dim EW As Integer
        Dim BAC As Integer
        Label11.Text = "Free Basic Water" & FB & "%"
        Label10.Text = "Billed Metered Consumption " & BMC & "%"
        Label4.Text = "Exported Water " & EW & "%"
        Dim percents() As Integer = {EW, BMC, FB}
        Dim colors() As Color = {Color.Purple, Color.Blue, Color.Green}
        Dim graphics As Graphics = Me.CreateGraphics
        Dim location As Point = New Point(100, 100)
        Dim size As Size = New Size(420, 420)
        DrawPieChart(percents, colors, graphics, location, size)
    End Sub


Private Sub Button2_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button2.Click
        Dim SCL As Integer
        Dim RO As Integer
        Dim MDL As Integer
        Dim MI As Integer
        Dim UAC As Integer
        Dim UMC As Integer
        Dim UUC As Integer
        Dim FB As Integer
        Dim BMC As Integer
        Dim EW As Integer
        Dim SI As Integer
        Label17.Text = "Service Connection Leaks " & SCL & "%"
        Label18.Text = "Reservoir Overflows" & RO & "%"
        Label16.Text = "Mains and Distribution Leaks" & MDL & "%"
        Label15.Text = "Metering Inaccuracies " & MI & "%"
        Label14.Text = "Unauthorised Consumption " & UAC & "%"
        Label13.Text = "Unbilled Metered Consumption " & UMC & "%"
        Label12.Text = "Unbilled Unmetered Consumption " & UUC & "%"
        Label11.Text = "Free Basic Water" & FB & "%"
        Label10.Text = "Billed Metered Consumption " & BMC & "%"
        Label4.Text = "Exported Water " & EW & "%"
        Dim percents() As Integer = {EW, BMC, FB, UUC, UMC, UAC, MI, MDL, RO, SCL}
        Dim colors() As Color = {Color.Purple, Color.Blue, Color.Green, Color.Red,
Color.Yellow, Color.Peru, Color.Cyan, Color.Silver, Color.RosyBrown, Color.Lime}
        Dim graphics As Graphics = Me.CreateGraphics
        Dim location As Point = New Point(100, 100)
        Dim size As Size = New Size(420, 420)
        DrawPieChart(percents, colors, graphics, location, size)
    End Sub
```

```
Private Sub Button2_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button2.Click

        Dim SCL As Integer
        Dim RO As Integer
        Dim MDL As Integer
        Dim RL As Integer
        Label11.Text = "Service Connection Leaks" & SCL & "%"
        Label10.Text = "Reservoir Overflows" & RO & "%"
        Label4.Text = "Mains and Distribution Leaks " & MDL & "%"
        Dim percents() As Integer = {MDL, RO, SCL}
        Dim colors() As Color = {Color.Purple, Color.Blue, Color.Green}
        Dim graphics As Graphics = Me.CreateGraphics
        Dim location As Point = New Point(100, 100)
        Dim size As Size = New Size(420, 420)
        DrawPieChart(percents, colors, graphics, location, size)
    End Sub
End Class


Private Sub Button1_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button1.Click
        Dim AC As Integer
        Dim WL As Integer
        Dim SI As Integer
        Dim percents() As Integer = {AC, WL}
        Dim colors() As Color = {Color.Blue, Color.Green}
        Dim graphics As Graphics = Me.CreateGraphics
        Dim location As Point = New Point(50, 50)
        Dim size As Size = New Size(400, 400)
        DrawPieChart(percents, colors, graphics, location, size)
        Label2.Text = "Authorised Consumption " & AC & "%"
        Label6.Text = "Water Losses " & WL & "%"
    End Sub


Private Sub Button2_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button2.Click

        Dim UMC As Integer
        Dim UUC As Integer
        Dim UAC As Integer
        Label2.Text = "Unbilled Metered Consumption " & UMC & "%"
        Label1.Text = "Unbilled Unmetered Consumption " & UUC & "%"
        Dim percents() As Integer = {UUC, UMC}
        Dim colors() As Color = {Color.Blue, Color.Green}
        Dim graphics As Graphics = Me.CreateGraphics
        Dim location As Point = New Point(100, 100)
        Dim size As Size = New Size(300, 300)
        DrawPieChart(percents, colors, graphics, location, size)
    End Sub


Private Sub Button1_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button1.Click
        Dim AL As Integer
        Dim RL As Integer
        Dim WL As Integer
        Dim percents() As Integer = {AL, RL}
        Dim colors() As Color = {Color.Blue, Color.Green}
        Dim graphics As Graphics = Me.CreateGraphics
```

```
        Dim location As Point = New Point(50, 50)
        Dim size As Size = New Size(400, 400)
        DrawPieChart(percents, colors, graphics, location, size)
        Label2.Text = "Apparent Losses " & AL & "%"
        Label6.Text = "Real Losses " & RL & "%"
    End Sub
```