Decision-making tools for establishment of improved monitoring of water purification processes

G O'Reilly

orcid.org 0000-0002-7576-1723

Thesis accepted in fulfilment of the requirements for the degree *Doctor of Philosophy in Environmental Sciences* at the North-West University

Promoter:	Prof CC Bezuidenhout
Assistant Promoter:	Dr JJ Bezuidenhout

Graduation May 2020

20728328

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude and appreciation to the following persons for their contributions and support towards the successful completion of this study:

The National Research Foundation (NRF) for their financial support.

Prof Carlos Bezuidenhout, words will never be enough to express my gratitude towards you Prof. Thank you for all the support, patience, constructive criticism and words of encouragement throughout the years. Thank you for always treating me with respect and never giving up on me, even during the tough times. You are truly the best at what you do and I was blessed to have you as my promotor.

Dr. Jaco Bezuidenhout, thank you for all your assistance with the statistics. Thank you for always having an open-door policy and for always being willing to help wherever you can. It has only been a pleasure to work with you.

Prof Nico Smit, thank you Prof Nico, not only for the financial support, but also for the words of encouragement, your kind heart and for always being understanding whenever I came to see you.

Prof. Sarina Claassens, thank you for all your support and words of encouragement throughout the past years. I have always looked up to you. Thank you for being a great role model and friend.

My parents, Stanley & Lynda O'Reilly. There are not enough words to say how thankful I am for all that you have done for me. Thank you for your unconditional love, your support in everything that I do and for being the best parents that anyone could ask for. I love you so much and I will be forever grateful for all your support throughout the years.

My fiancé, Marco van Staden. Thank you for all your love and support babs, especially during the last few months. Thank you for always believing in me and teaching me to never give up. I have learned so much from you and I thank God every day for giving me a life partner so understanding and loving to walk this journey with me. I couldn't have asked for anyone better. I love you with all my heart!!

To my family, Granny (ouma Sarie), Zynéldi, Stanley, Roos, Ryan, Chanté, Gavin, Blom, Owen, Chani-mae, Joe, Lychél, Liam, Luhan, Hanlie, OD, Tannie Anne-Marie and Anneme. Thank you for all your love and support throughout my studies. I am so blessed and grateful to have such an awesome family and support network. I love you all very much.

Roelof Coertze, Roels, thank you for your assistance with some of the statistical work. More importantly, thank you for always being there for me and for your friendship and support throughout the years. I really appreciate everything that you have done for me. Keeping it in the family… Tannie Linda and Oom Dirk, tannie Linda, firstly, thank you for your assistance with the proofreading. I really appreciate the help on such short notice. Secondly, thank you both for all your love and support during my studies. I really appreciate it.

To all my friends and work colleagues at the Department of Microbiology, thank you for all the words of encouragement, laughs and support throughout the years. A special thank you to Clarissa and Bianca. We started this journey together…and we have obviously not finished together.. hahaha… but nonetheless, thank you for your friendship, motivation, support and just for always being there when I needed you.

To Vasti, thank you for being my best friend for the past 17 years. I really appreciate your friendship and support throughout my studies. Thank you for always being there for me and for always phoning me. Hehe. Dankie maat.

To my work colleagues and friends, Adéle (Êdêlle) and Adriaan(nus). A HUGE thank you for all your support the past few months. Thank you for all the laughs and words of encouragement and understanding when things got a bit rough. I really appreciate all that you have done for me.

TABLE OF CONTENTS

Abstract

Source water is becoming a scarce resource in South Africa, especially with the extreme periods of drought that the country has faced the last few years. Continuous pollution of rivers and dams is drastically deteriorating the water quality. This puts more pressure on water purification plants to ensure the water is adequately treated and safe drinking water is produced. However, some water purification plants may not have the infrastructure or financial resources to produce drinking water of acceptable quality. Therefore, there is a need for inexpensive and rapid solutions to produce safe drinking water. In this study, three decision-making tools (Hazard Analysis Critical Control Point (HACCP) concept, Artificial Neural Networks, Evolutionary Algorithms and isolation of VBNC bacteria) were evaluated for the improved monitoring of treatment processes. The HACCP concept was evaluated at three water purification plants (Plant A, Plant B and Plant C). This study demonstrated that monitoring certain parameters after each step in the treatment process was useful to identify process failures. Results for Plant B indicated that due to a failure in the filtration process, unacceptable turbidity levels were present in the drinking water. The second decision-making tool investigated in this study was the application of ANNs and evolutionary algorithms (EAs) at Plant A, Plant B and Plant C. Results from this study indicated that the combination of ANNs and EAs resulted in the accurate prediction of electrical conductivity (EC) in the drinking water of Plant A and Plant C. Additionally, this study indicated the importance of consistent monitoring. A prediction model for Plant B could not be generated due to a lack of historical data. The application of EAs resulted in the formation of accurate predictive rule-sets for Plant A and Plant C. The last decision-making tool was the recovery of *Escherichia coli* (*E. coli*) in drinking water. Results of this study indicated that viable-but-nonculturable *E. coli* was recovered from drinking water. Two of the isolates were identified as *E. coli* O177 and O157. This is concerning as O177 and O157 are Shiga toxin-producing strains of *E. coli* and could pose a serious health risk. In addition, this study indicated the development of a new resuscitation method adapted from a current method. Based on all the results obtained, decision-making tools for the improved monitoring of water treatment processes was demonstrated.

Keywords: Hazard Analysis Critical Control Point; artificial neural networks; evolutionary algorithms; viable-but-nonculturable *Escherichia coli*

Preface

This thesis was written in article format. Chapter 3 has been published in *Water Science & Technology: Water Supply* under the title “Artificial neural networks: applications in the drinking water sector”. I (G. O’Reilly) was the main author of this article. The co-authors were C.C. Bezuidenhout and J.J. Bezuidenhout. Permission from the co-authors for submitting this article for degree purposes is given in Appendix B.

List of abbreviations

ANFIS	Adaptive network-based fuzzy inference system
ANNs	Artificial neural networks
BPNN	Back Propagation Neural Network
CART	Classification and regression tree
CCPs	Critical Control Points
CFN	Cascade Forward Networks
DBPs	Disinfection by-products
DWA	Department of Water Affairs
E. coli	*Escherichia coli*
EAs	Evolutionary algorithms
EC	Electrical conductivity
eWQMS	Electronic Water Quality Management System
GA	Genetic algorithm
GAC	Granulated Activated Carbon
GEP	Gene expression programming
GRNN	General Regression Neural Network
HACCP	Hazard Analysis Critical Control Point
MAPE	Mean Absolute Percent Error
MLP	Multi-layer Perceptrons
MLR	Multiple linear regression
MLSA	Membrane Lauryl Sulphate Agar
PAC	Poly aluminium chloride
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
RBF	Radial Basis Function

RDA	Redundancy analysis
RO	Reverse Osmosis
SANS	South African National Standard
SOM	Self-organizing map
SVM	Support vector machine
TDS	Total Dissolved Solids
VBNC	Viable-but-non-culturable
WHO	World Health Organization
WSPs	Water Safety Plans

List of Figures

List of Tables

CHAPTER 1

Introduction and problem statement

Source water is becoming a scarce resource in South Africa and with developments arising along our important rivers it can be expected to influence water quality negatively (Van der Laan *et al.*, 2012; Cullis *et al*., 2018). Rivers play an important role in socio-economic growth, but pollution due to human activity resulted in rivers not being in their natural state (Kinge & Mbewe, 2012). The quality of source water in South Africa is threatened by agricultural run-off, industrial and sewage effluent, untreated domestic wastewater and uncontrolled disposal of wastewater by informal settlements (Nyenje *et al.*, 2010; Sershen *et al.*, 2016).

Natural events, such as drought, may also have a negative impact on source water bodies (Deksissa *et al.*, 2003). Loss of water causes reduced flow in rivers which in turn rapidly increase total dissolved solids (TDS) concentrations and nutrient concentrations and deteriorate the water quality (Deksissa *et al.*, 2003; Cullis *et al*., 2018). During 2015 and 2016, South Africa experienced a combination of extreme drought and a strong El Niño event. The El Niño effect is a quasi-periodic intrusion of warm sea surface water into the central and eastern Pacific Ocean (Baudoin *et al.*, 2017). The drought had a negative impact on water availability and supply (Baudoin *et al.*, 2017). During 2017 and 2018, the City of Cape Town in the Western Cape Province experienced the consequences of major water shortages due to a severe drought. With human population and economic developments continuously increasing, the water quality of river systems will continue to decline, unless it is properly managed (Deksissa *et al.*, 2003; Cullis *et al*., 2018).

Deteriorating source water quality may have a negative impact on water purification plants. The type of treatment processes applied by water purification plants depend on the pollutants in the source water (Momba & Brouckaert, 2005). Advanced treatment processes may be necessary to remove contaminants from heavy polluted source water (Morrison *et al.*, 2012). The evaluation of water safety is often limited to occasional water quality tests and not much consideration is given to proactive water quality management (Godfrey & Howard, 2004; Douterelo *et al*., 2019). There are various proactive management approaches available which water treatment plants can use as decision-making tools to improve the monitoring of their treatment

processes, such the Hazard Analysis Critical Control Point (HACCP) concept, water safety plans (WSPs), artificial neural networks (ANNs), evolutionary algorithms (EAs) and unconventional microbial monitoring.

1.1 The HACCP concept

The HACCP concept is a preventative risk management approach that originated from the food industry to ensure food safety (Bartram, 2009). It focuses on identifying critical control points (CCPs) within a system. CCPs are points whose failure may lead to greater public health risk compared to other points (US Environmental Protection Agency, 2007). Identification of critical control points within the system will help to identify potential hazards closer to the source (US Environmental Protection Agency, 2007). This ensures that potential hazards can be prevented, removed or reduced to an acceptable level (Damikouka *et al.*, 2007; Dunn *et al.*, 2014). Implementation of the HACCP plan involves seven main principles (Martel *et al.*, 2006; Damikouka *et al.*, 2007):

- Principle 1: Perform a hazard analysis
- Principle 2: Determine the CCPs in the system
- Principle 3: Establish the critical limits
- Principle 4: Establish a monitoring system for the CCPs
- Principle 5: Establish corrective action procedures for possible CCP failures
- Principle 6: Validate the HACCP plan
- Principle 7: Establish a documentation system

The drinking water sector is becoming more aware of the limitations of end-point testing (Okeyo *et al*., 2011; Tsitsifli & Tsoukalas, 2019). The application of the HACCP concept at water purification plants promotes source-to-tap monitoring and may add the following benefits to the drinking water system: improve public health protection and regulatory compliance; improve the design and operation of the treatment process; better understanding of the risks and risk management; improve operator skills and improve work processes (US Environmental Protection Agency, 2007). Even though the HACCP concept has only been applied in the drinking water sector since the 1990's (US Environmental Protection Agency, 2007), its application in this sector has proven to be successful (Damikouka *et al.*, 2007; Okeyo *et al.*, 2011; Tavasolifar *et al.*, 2012; Tsitsifli & Tsoukalas, 2019).

1.2 Water Safety Plans

During 2004, the World Health Organization (WHO) introduced WSPs (World Health Organization, 2017a). WSPs are comprehensive risk assessment and management plans which identifies and prioritise potential threats to water quality from source to tap and implementing the best methods to alleviate threats to drinking water (Dunn *et al.*, 2014). WSPs consists of three key components: system assessment and design; operational monitoring; and management plans (Thompson & Majam, 2009; World Health Organization, 2017a).

The South African government implements water safety planning through the Blue Drop Certification Programme. The Blue Drop Certification Programme consists of six main sections with subsections containing specific requirements to which all water services institutions need to adhere. Each year the water services institutions are audited against these requirements and provided with a score to determine their improvements. Blue Drop Certification is awarded to water service institutions that achieve a score of ≥ 95% (Department of Water and Sanitation, 2014). There are five main objectives of the Blue Drop Certification Programme (Department of Water Affairs, 2014), namely:

- To establish an incentive-based monitoring plan for drinking water quality management;
- Encourage transparency and subsequent accountability;
- Provide the public with reliable information;
- Where applicable, improve the relationship between Water Services Authorities and Water Services Providers;
- Introduce a factor of excellence

Water safety planning carry the largest weight (35%) of the Blue Drop Certification Programme and consists of 5 main components (Department of Water Affairs, 2014), namely:

- Planning process
- Risk assessment
- Monitoring Programme (operational monitoring & compliance monitoring)
- Credibility of drinking water quality (DWQ) data
- Incident management

WSPs are based on the multiple-barrier and HACCP approaches (Dunn *et al.*, 2014; World Health Organization, 2017a) and can assist water purification plants in taking simple and cost-effective steps to improve their treatment process.

1.3 Artificial Neural Networks (ANNs) and Evolutionary Algorithms (EAs)

McCulloch and Pitts first introduced the concept of artificial neurons in 1943 describing how neurons in the brain might work (McCulloch & Pitts, 1943). They modelled a simple neural network by making use of electrical circuits (Clabaugh *et al.*, 2000). In 1949, neuropsychologist Donald Hebb highlighted the fact that every time neural pathways are used, it can reinforce each other (Hebb, 1949; Clabaugh *et al.*, 2000). This lasting effect is what is referred to as memory or learning (Hebb, 1949). During the 1950's there were some developments in developing these concepts. However, the 1960's and 1970's did not see much expansion of the research into artificial neural networks (Graupe, 2013). This was possibly due to flaws in the modelling techniques which resulted in reduced funding for research (Waskan, 2016).

However, interest in the field started again during the early 1980's. In 1982, John Hopfield's approach to neural networks was not only to model the brain, but through mathematical analysis he indicated how such networks could emulate phenomena such as memory and learning as found in neurobiological systems and how it could be applied to special classes of computational problems (Hopfield, 1982; Siddique & Adeli, 2013). This, however, was not the first time the mathematical analysis of neural networks was presented. Improvements to the mathematical frameworks underlying the training schemes of neural networks was developed by Paul Werbos in 1974, but was largely unnoticed at the time, as the work was only published as a Doctor of Philosophy thesis (Werbos, 1974; Siddique & Adeli, 2013). Hopfield further proposed that improvements to artificial neural networks would be possible with asynchronous parallel processing (Hopfield, 1982).

In 1982, Teuvo Kohonen developed a neural network, known as the Kohonen self-organising map (SOM) (Kohonen, 1982; Clabaugh *et al.*, 2000). SOMs are able to map multidimensional input data onto a neural network with a much lower dimension preserving the topological and metric relationships of the input data (Kohonen *et al.*, 1996; Bowden *et al.*, 2005; Farmaki *et al.*, 2013). In 1986, research into multiple layered neural networks began. Researchers faced the problem of extending the

Widrow-Hoff learning rule to multiple layers (Clabaugh *et al.*, 2000). This problem was solved by researcher David Rumelhart and colleagues during 1986 resulting in the back-propagation learning network being formulated (Clabaugh *et al.*, 2000; Graupe, 2013).

Since then computing potential for layered networks have increased considerably (Siddique & Adeli, 2013). Also, in 1986, two volumes on parallel distributed processing were published (Rumelhart & McClelland, 1986a; Rumelhart & McClelland, 1986b). This process proved the foundation to remove a significant training barrier for ANNs (Rumelhart & McClelland, 1986a; Rumelhart & McClelland, 1986b; Siddique & Adeli, 2013) and resulted in an increase in applications of ANNs.

When the quality of source water change, water purification plants usually rely on extra bench-scale studies or past experiences to solve subsequent problems (Veerapaneni *et al.*, 2010). However, this can be unreliable and time-consuming and may lead to costly downtime whilst operators are fixing the problems (Veerapaneni *et al.*, 2010). ANNs can be used as forecasting tools as they have the ability to learn certain patterns through a learning algorithm (Sarkar & Pandey, 2015). Unlike conventional statistical methods, ANNs can process large amounts of non-linear data quickly. As a result, underlying problems within a system can be identified and solutions implemented rapidly (Sarkar & Pandey, 2015).

However, ANNs have the disadvantage of being “black box” models, as they do not provide an equation or rule set to explain how the prediction models were developed (Yang *et al.*, 2014; Delpla *et al.*, 2019). This problem can be addressed by applying evolutionary algorithms. EAs recognise certain patterns in a dataset and search for appropriate representations of models through genetic variation and the principle of “survival of the fittest” (Welk *et al.*, 2008; Kim *et al.*, 2010). By applying genetic programming, the model architecture can then be defined by developing predictive rule sets (Welk *et al.*, 2008; Kim *et al.*, 2010).

1.4 Unconventional microbial monitoring

Microbiological contamination of drinking water is one of the most serious public health risks (Forstinus *et al.*, 2016). Faecal pollution in water resources deteriorates the aquatic ecosystem and subsequently poses a health risk to communities consuming the water (Ontumbi *et al.*, 2015). The World Health Organization (WHO) suggests that

Escherichia coli (*E. coli*) is the organism of choice to indicate faecal pollution (World Health Organization, 2017a). Even though *E. coli* forms part of the normal intestinal flora of human and animals, exposure to other parts of the body can cause serious infections (World Health Organization, 2017a). Pathogenic strains of *E. coli* may cause urinary tract infections, meningitis and acute diarrhoea (World Health Organization, 2017a).

The acceptable physico-chemical and microbiological limits for drinking water in South Africa are defined in the South African National Standard (SANS) 241:2015. The specification for *Escherichia coli* in drinking water is 0 CFU/100 ml. Various studies have found that *E. coli* could enter into a viable-but-non-culturable (VBNC) state (Zhang *et al.*, 2015; Lin *et al.*, 2017; Chen *et al.*, 2018). The VBNC state is defined as organisms not able to grow on conventional media, but they retain metabolic activity and may return to a viable state (Zhao *et al.*, 2013; Lin *et al.*, 2017). The presence of VBNC *E. coli* in drinking water may give water purification plants the false impression that all the *E. coli* cells have been removed. Therefore, monitoring drinking water for the presence of pathogenic organisms other than required by the SANS 241, such as VBNC *E. coli*, could be indicative of underlying failures in the treatment process.

Problem statement:

To ensure the continuous production of safe drinking water, water purification plants need to implement a holistic management approach (World Health Organization, 2017a). This approach requires water purification plants to identify certain risks within a system and to ensure that the control measures put in place to manage these risks are working effectively (World Health Organization, 2017a). Water purification plants in the metropolitan cities of South Africa have the infrastructure, capacity and efficiency to perform comprehensive water quality control (Damikouka *et al.*, 2007; Okeyo *et al.*, 2011). However, this is not the case for smaller cities, which have high incidences of poor water quality production (Momba *et al.*, 2004; Okeyo *et al.*, 2011). This may be due to financial and technical limitations (World Health Organization, 2017a). Therefore, there is a need for cost-effective solutions to ensure the production of safe drinking water.

1.5 Description of study sites

The current study focused on three water purification plants in the North West Province, which will be referred to as Plant A, B and C.

1.5.1 Plant A

Plant A forms part of the Mooi River catchment. The Mooi River catchment, situated in the North West and western Gauteng Provinces, has three major tributaries, namely the Wonderfontein Spruit, Loop Spruit and the Mooi River (Bezuidenhout, 2013; van der Walt & Nell, 2002). The Mooi River flows southwards from the Boons area east of Koster in the North West Province into Boskop Dam and Potchefstroom Dam until it joins the Vaal River (Venter *et al.*, 2013). Water quality of the Mooi River is influenced by agricultural and small-scale diamond mining activities (van der Walt & Nell, 2002). Upstream from the Mooi River, the Wonderfontein Spruit is negatively influenced by acid mine drainage (Venter *et al.*, 2013). During heavy rainfall periods, polluted water from the Wonderfontein Spruit flows into Potchefstroom and Boskop Dams (Venter *et al.*, 2013). Plant A applies basic water treatment methods. Figure 1.1 illustrates the treatment process flow of Plant A.

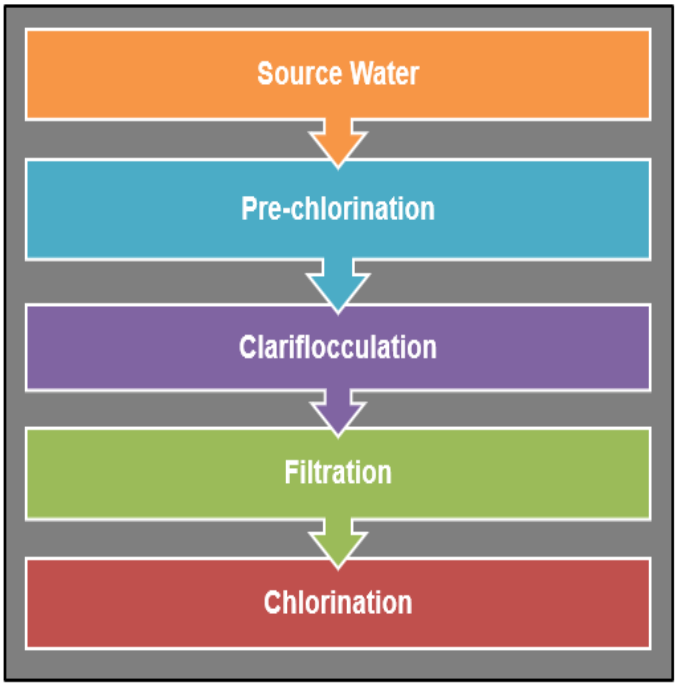


Figure 1.1: Water treatment process of Plant A.

1.5.2 Plant B

Plant B is situated in the eastern part of the North West Province and receive their source water from a groundwater recharge point called the Schoonspruit eye. The Schoonspruit stream flows from Ventersdorp through Klerksdorp and converges with the Vaal River west of Orkney (Lusilao-Makiese *et al.*, 2014). The Schoonspruit stream has been polluted by industrial activities, spillage from tailing facilities and artisanal mining activities from the local community Kanana near Orkney (Lusilao-Makiese *et al.*, 2014). Plant B also applies basic treatment methods. Figure 1.2 illustrates the treatment process flow of Plant B.

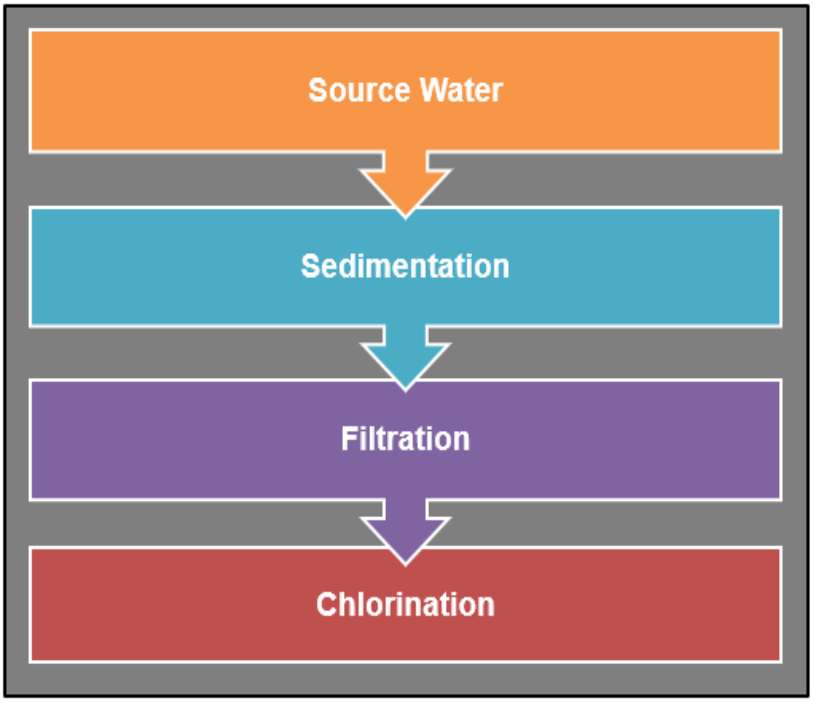


Figure 1.2: Water treatment process of Plant B.

1.5.3 Plant C

Plant C is situated on the river banks of the Vaal River in the eastern part of the North West Province. Source water is abstracted from the Vaal River. The Vaal River is one of the major rivers in South Africa. Continuous disposal of domestic and industrial waste into the Vaal River has deteriorated the water quality considerably (Jordaan & Bezuidenhout, 2013). It consists of three water management areas namely the Upper, Middle and Lower Vaal (Bezuidenhout, 2013; Department of Water Affairs and Forestry, 2009). The Middle Vaal catchment includes parts of the Free State and North West Provinces (Bezuidenhout, 2013; Department of Water Affairs and Forestry,

2009). Water quality of the Middle Vaal area is of concern and water treatment plants have had to upgrade their treatment systems due to high TDS levels and sporadic algal blooms (Bezuidenhout, 2013; Department of Water Affairs and Forestry, 2009). Plant C applies advanced treatment methods which are illustrated in Figure 1.3.

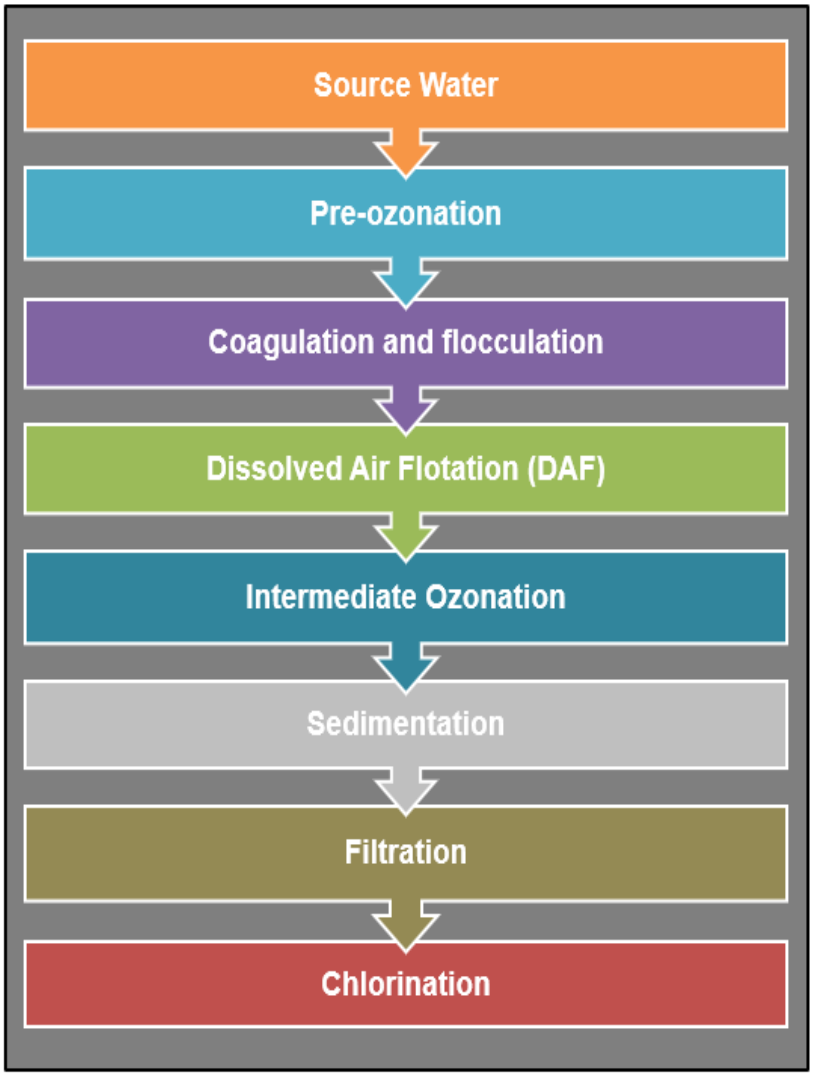


Figure 1.3: Water treatment process of Plant C.

Plant A, Plant B and Plant C were chosen as study sites for the current study due to their differences in size, management and treatment processes. Table 1.1 gives a description of the monitoring methods currently used by each water purification plant.

Table 1.1: Current monitoring methods of Plant A, Plant B and Plant C.

Water purification plant	Current monitoring methods
Plant A	• Raw water and final water monitoring only • Frequent monitoring of physico-chemical and microbiological parameters • Water samples tested at onsite laboratory
Plant B	• Raw water and final water monitoring only • Not frequent monitoring of physico-chemical and microbiological parameters • Water samples tested by an accredited laboratory
Plant C	• Applying HACCP concept, thus monitoring certain parameters after each point in the treatment process • Frequent monitoring of physico-chemical and microbiological parameters • Water samples tested at onsite laboratory

The aim of this study was to apply decision-making tools at three drinking water purification plants in the North West Province, for the establishment of improved monitoring of water purification processes. Specific objectives included:

i) To perform a comparative analysis of the HACCP concept at drinking water purification plants;
ii) To conduct a review of the applications of artificial neural networks in the drinking water sector;
iii) To apply artificial neural networks and evolutionary algorithms at drinking water purification plants;
iv) To perform the recovery of *Escherichia coli* in drinking water

1.6 Outline of thesis

Chapter 1 provides an overview of the quality of the source water in South Africa. It also provides information on the various decision-making tools available for the proactive management of water purification processes. Thereafter, the aim and

objectives are provided. The chapter concludes with descriptions of the study sites and ends with the outline of the thesis.

Chapter 2 reports on the comparative analysis of the HACCP concept at the drinking water purification plants. A general introduction is given, after which materials and methods, results, discussion and a conclusion follow. The effectiveness of the HACCP concept was illustrated by using historical and measured data. The use of HACCP as a decision-making tool to improve monitoring at water purification plants was illustrated.

Chapter 3 is a literature review of the applications of artificial neural networks in the drinking water sector. This chapter has been published in *Water Science & Technology: Water Supply* under the title “Artificial neural networks: applications in the drinking water sector” (the full reference is provided in the chapter).

Chapter 4 reports the application of artificial neural networks and evolutionary algorithms at drinking water purification plants. A general introduction is given, there after the materials and methods, results, discussion and a conclusion follow. The ANNs were applied successfully with high accuracy prediction models that were developed. Predictive rule sets for these models were also successfully developed using evolutionary algorithms. The combination of using ANNs and EAs as a decision-making tool to improve monitoring at water purification plants was indicated.

Chapter 5 reports on the recovery of *Escherichia coli* in drinking water. A general introduction is given, there after the materials and methods, results, discussion and a conclusion follow. VBNC *Escherichia coli* were recovered in the drinking water by means of a current method as described by a previous study, as well as a new method adapted from the first method.

Chapter 6 provides a summary of the conclusions of the stated objectives, after which recommendations for future research in this field are given.

References are provided after Chapter 6

CHAPTER 2

Comparative analysis of the HACCP concept at drinking water purification plants

2.1 Introduction

The HACCP concept was first introduced in the 1960's by the US Space Agency to ensure the safe production of food to be used during spaceflights (Hamilton *et al.*, 2006). It has been applied in the food industry since the 1980's (US Environmental Protection Agency, 2007). HACCP is still used today in the food industry as a preventative approach for food safety (Tsoukalas & Tsitsifli, 2018). It is a system that identifies biological, physical and chemical hazards at specific control points, which may potentially be harmful to human health (Tsoukalas & Tsitsifli, 2018). Evaluation of the entire water purification system from source to tap could lead to the understanding of existing problems and also identify potential problems (Dunn *et al.*, 2014). During 2004, the World Health Organization (WHO) published guidelines for drinking water quality which focussed on risk-based approaches, including HACCP (World Health Organization, 2004; Dunn *et al.*, 2014). This emphasized the inadequacy of traditional risk management systems that focus on end-point testing only (Tsoukalas & Tsitsifli, 2018).

Even though the HACCP concept has been applied in the water industry since the mid-1990's (US Environmental Protection Agency, 2007; Tsitsifli & Tsoukalas, 2019), it remains a concept that originated from the food industry. Hence, there was a need for a preventative management framework specifically designed for water purification plants. During various expert review meetings by the WHO in 2000 and 2001, the possible application of Water Safety Plans (WSPs) in the water sector was evaluated. WSPs are based on the HACCP principles and the value of applying WSPs in the water sector was highlighted during the revision of the WHO Guidelines for drinking-water quality leading to the 3rd edition (World Health Organization, 2004; Davison *et al.*, 2005).

In South Africa, the Department of Water Affairs (DWA) was mandated by Cabinet in 2003 to serve as Regulator of Water Services (Department of Water Affairs, 2014). As stipulated in Section 62 of the Water Services Act (no. 108 of 1997) of South Africa, the DWA has the duty to monitor water service providers to ensure safe drinking water

is produced (South Africa, 1997). To improve regulation of drinking water and improve drinking water quality, the Department then implemented the Electronic Water Quality Management System (eWQMS), which focussed on end-point testing of drinking water alone. This proved to be a major limitation of the eWQMS system. By the time health-related contaminants or process failures were identified in the drinking water, communities may have already been exposed (Department of Water Affairs, 2014). This emphasized the importance of preventative monitoring.

For this reason, a framework for the production of safe drinking water was established by the Department of Water Affairs in 2008, called the “Blue Drop Certification Programme for Drinking Water Quality Management regulation” (Department of Water Affairs, 2014). The Blue Drop Certification Programme is an incentive-based regulation programme consisting of six main sections (Department of Water and Sanitation, 2014): Water Safety Planning (35%); Drinking Water Quality Process Management and Control (8%); Drinking Water Quality Verifications (30%); Management, Accountability & Local Regulation (10%); Asset Management (14%); Water Use Efficiency & Water Loss Management (3%). Each section has specific requirements to which municipalities (as Water Services Authorities) must adhere to, as stated in section 62 of the Water Services Act no. 108 of 1997 (South Africa, 1997).

Water Safety Planning accounts for 35% of the Blue Drop Certification Programme and consists of five subsections, including a monitoring programme (Department of Water and Sanitation, 2014). The monitoring programme is divided to address operational monitoring and compliance monitoring separately. Under operational monitoring the required sites to monitor at least every 8 hours are: raw water, after filtration and final water.

Traditionally, water treatment process consists of screening, coagulation, flocculation, sedimentation, clarification and disinfection (Okeyo *et al.*, 2011; Hu *et al.*, 2018). All of these points in the treatment process have the potential to fail, which may lead to the production of poor quality drinking water. Therefore, not monitoring each step in the treatment process may lead to unnecessary increases in costs for water purification plants, as identifying the source of the failure may be time-consuming and laborious. Even though the Blue Drop Certification Programme added an extra monitoring point

(after filtration), should there be a failure in the treatment process, this may not be the point where the underlying problem is occurring.

The prevention of failures through the purification process, rather than correcting them, has become more important in recent decades (World Health Organization, 2011; Tsoukalas & Tsitsifli, 2018). Therefore, the application of the HACCP concept at water purification plants can be an effective decision-making tool for the evaluation of hazards in the treatment process.

Iceland has applied the HACCP concept at various water treatment plants since 1997 (Gunnarsdóttir & Gissurarson, 2008). The implementation of HACCP proved to be very successful as results indicated that the compliance with drinking water quality standards improved considerably. About 68% of the Icelandic population enjoy drinking water supplied by water purification plants that implemented the HACCP concept (Gunnarsdóttir & Gissurarson, 2008). The capital of Iceland, Reykjavík, had an improved mean compliance value for bacterial count from 94% for the years 1991 to 1997 to 99% for the years 1998 to 2006. Another town, Akureyri, had an improved mean compliance value for bacterial count from 88% for the years 1992 to 1999 to 99% for the years 2000 to 2004 (Gunnarsdóttir & Gissurarson, 2008).

Various water purification plants have evaluated the HACCP principles to develop a framework for the actual implementation of the HACCP concept. Neuilly-sur-Marne, a suburb in Paris, France, applied the HACCP principles to their drinking water purification plant (de Traversay, 2006). Treatment processes of the plant included coagulation/flocculation, sedimentation, rapid sand filtration, inter-ozonation, Granulated Activated Carbon (GAC) stage and chlorination. The critical control points identified were raw water, after sand filtration, after ozonation, after chlorination and water in the distribution. Historical data of turbidity and microbiological parameters were used to evaluate the HACCP plan. This study indicated that the evaluation of the HACCP plan helped to define the system to ensure continuous control and monitoring of critical control points (CCP's) could be implemented. It was also a microbiological quality control mechanism to ensure that the consumers are provided with safe drinking water.

In Iran, the HACCP principles were applied at a water treatment plant in Germi City (Khaniki *et al.*, 2009), as well as in Isfahan (Tavasolifar *et al.*, 2012). The water

treatment plant in Germi City makes use of traditional treatment processes which include screening, pre-chlorination, coagulation, flocculation, sedimentation, filtration and post-chlorination. A CCP decision tree was used to identify the critical control points at the plant which included raw water source, pre-chlorination, coagulation, flocculation, sedimentation, filtration and post-chlorination. Various problems and solutions were identified through the evaluation of the HACCP concept at the Germi plant. Their investigation concluded that changing the dose and type of coagulant as well as the conditions of mixing and sedimentation should improve the water quality of Germi water treatment plant.

The water treatment plant in Isfahan used historical data to develop their HACCP model. The eight CCP's identified were the intake point, coagulation and flocculation, sedimentation and middle disinfection, filtration, final disinfection and the distribution system. Results indicated that extensive control of five of the CCP's are needed to ensure water safety: coagulation, middle disinfection, filtration, final disinfection and the distribution system. The study concluded and recommended that the water treatment systems in the cities of Iran need to implement prerequisite programs to prepare them for the implementation of HACCP as the implementation of HACCP is essential.

Damikouka *et al.* (2007) evaluated the HACCP principles at the Aspropyrgos Water Treatment Plant in Greece. The aim of the study was to provide a framework for the steps that were taken in preparing a comprehensive HACCP plan for the water purification plant. The critical control points identified in the study included flocculation, filtration and chemical disinfection. The study concluded that even though the HACCP plan may not be entirely implemented at the water purification plant, the planning process itself highlighted certain areas where improvements could be implemented.

The aim of the current study was to evaluate the effectiveness of the HACCP concept at three water purification plants (Plant A, Plant B and Plant C) in the North West Province, South Africa, using historical and measured data. Plant C has already implemented the HACCP concept at their water purification plant, whereas Plant A and B mainly monitor source and drinking water.

2.2 Materials and Methods

2.2.1 Evaluation of historical data

Historical data was obtained from Plant A, Plant B and Plant C as electronic and/or hard copies. The turbidity data from 2009 to 2015 were analysed to determine the efficiency of the water treatment processes at Plant A, Plant B and Plant C over time. High turbidity levels can be an indication of the presence of chemical and biological particles, which may have a negative impact on the safety and aesthetic quality of the water produced by water purification systems (World Health Organization, 2017a). The average percentage turbidity removal and the lowest percentage turbidity removal for each year was determined.

2.2.2 Evaluation of measured data

Plant C already implements the HACCP concept. Therefore, we did not perform sampling at Plant C after each process, as this data could be provided by the operators of Plant C. Physico-chemical and microbiological parameters were measured at Plant A and B after each step in the treatment process during 2015 and 2016. The water was sampled in sterile glass bottles (Schott Duran, Germany) and transported on ice in cooler boxes to the North-West University Microbiology laboratory. Sampling was done according to the sampling guide by the Department of Water Affairs and Forestry, Department of Health and the Water Research Commission (Department of Water Affairs and Forestry, 2000). Table 2.1 indicates the sampling points of each plant. The water samples were analysed immediately upon arrival at the laboratory.

Electrical conductivity (EC), pH, salinity, total dissolved solids (TDS) and temperature were measured onsite using a PCSTestr 35 multi-meter (Eutech Instruments Pte Ltd, Singapore). Turbidity was measured using a Hach portable turbidimeter (model 2100P). The manufacturers' protocol was followed. Heterotrophic plate count (HPC) bacteria were isolated using R_2A agar (Merck, Germany) by means of the spread plate method. The plates were incubated at 25°C for 5 days. All the colonies present on the plate were counted.

Table 2.1: Different sampling points of Plant A and B.

	SAMPLING POINTS
PLANT A	Source water, pre-chlorination, clarifloccuation, disinfection, point of consumption (POC)
PLANT B	Source water, filtration (tank 1, 2 & 3), disinfection, POC

2.3 Results

2.3.1 Evaluation of historical turbidity data

Table 2.2 indicates the average percentage turbidity removal, as well as the lowest percentage turbidity removal for Plant A and C for each year from 2009 to 2015. As depicted in Table 2.2, the average percentage turbidity removal for Plant A from 2009 to 2015 were generally above 80%. However, each year had instances of very low percentages of turbidity removal. According to the South African National Standards (SANS 241:2015), the operational limit for turbidity in drinking water should be < 1 NTU. The lowest percentage turbidity removal for Plant A was during 2013, which was 8%. On this specific day, the turbidity in the source water was recorded as 3.95 NTU and 3.63 NTU in the drinking water (Table 2.2), which exceeded the SANS 241:2015 limit. During 2009, the lowest percentage turbidity removal was 60%. This was 25% less than the average turbidity removal of 85.94% for the year (Table 2.2). On this specific day, the turbidity in the drinking was 2.44 NTU, which also exceeded the SANS 241:2015 limit. The high turbidity levels of 3.63 NTU and 2.44 NTU in the drinking water of Plant A was an indication that the treatment process failed to reduce the turbidity to a safe level.

The average percentage turbidity removal for Plant C remained above 95% from 2009 to 2015 (Table 2.2). However, during 2011 a turbidity removal of 88% was obtained, which was 10% less than the average percentage turbidity removal of 98.72% for that year. This low percentage turbidity removal was recorded on 20 July 2011. Due to Plant C implementing the HACCP concept, turbidity levels after each step could be inspected to identify the possible reason for the low percentage turbidity on that specific day. Table 2.3 indicates the turbidity levels after each step in the treatment process during June, July and August 2011. As expected, the turbidity levels gradually decreased after each step in the treatment process. Similar results were obtained during the other months of 2011. However, on 20 July the turbidity increased from 3.7

NTU after ozonation to 4.7 NTU after settling West (Table 2.3), which was unusual. This had a negative effect on the remaining processes and the turbidity in the final water resulted in 1.4 NTU, which exceeded the SANS 241:2015 limit.

During 2014, the lowest percentage turbidity removal for Plant C was 90% with a turbidity level of 1.1 NTU recorded in the drinking water (Table 2.2). Similar to the incident during 2011, the turbidity level after each step in the treatment process could be inspected to determine the possible cause of failure. In this case, there was a drastic increase in the turbidity level after ozonation which could have contributed to the increase in the high turbidity level in the drinking water.

Limited historical data was available for Plant B. From the data obtained from 2011 to 2014, it was observed that the turbidity levels mostly increased in the drinking water and exceeding the SANS 241:2015 limit (Figure 2.1). However, during October 2013, the turbidity level was reduced from 0.5 NTU in the source water to 0.2 NTU in the drinking water. During May 2014 the turbidity was reduced by 80% from 1 NTU in the source water to 0.2 in the drinking water. The highest turbidity level in the drinking water was recorded at 6.6 NTU during May 2013. The continuous high turbidity levels in the drinking water from 2011 to 2014 was an indication of failure in the treatment process, which was concerning.

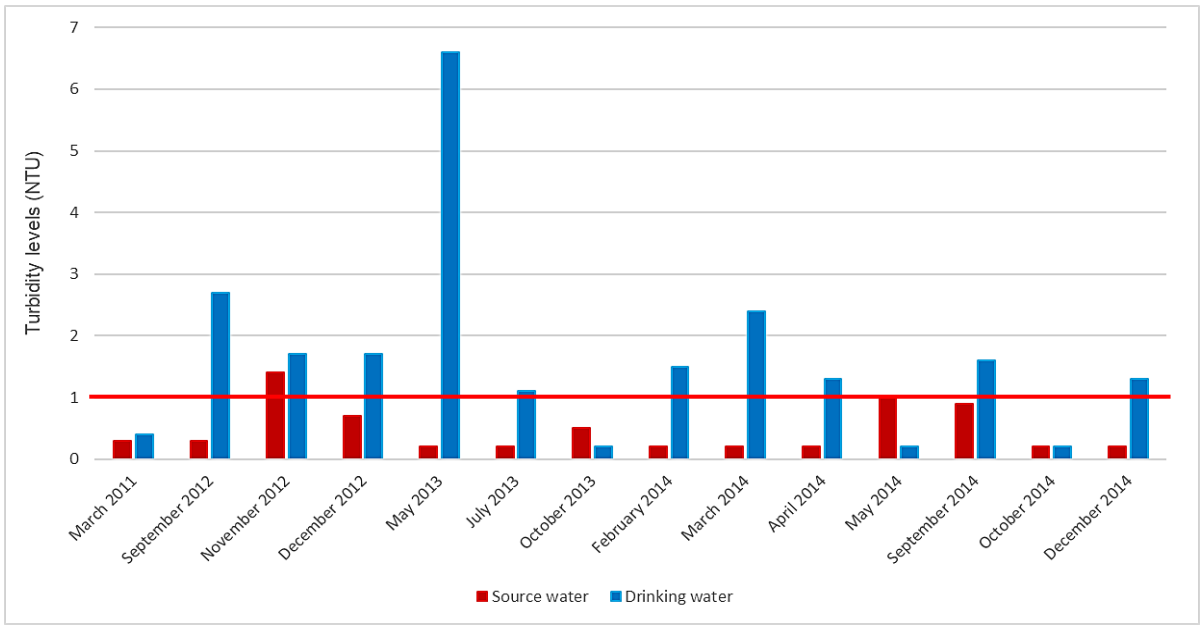


Figure 2.1: Turbidity levels in the source and drinking water of Plant B during 2011 to 2014. The red line indicates the SANS 241:2015 limit for turbidity in drinking water (< 1 NTU).

Table 2.2: Table indicating the average % turbidity removal, as well as the lowest % turbidity removal for Plant A and C for each year from 2009 to 2015.

PLANT	YEAR	AVERAGE % TURBIDITY REMOVAL FOR THE YEAR	LOWEST TURBIDITY REMOVAL		
			Turbidity (NTU) in source water	Turbidity (NTU) in final water	% removal
PLANT A	2009	85.94%	6.17	2.44	60%
	2010	81.84%	1.77	0.91	49%
	2011	84.62%	1.69	0.98	42%
	2012	78.03%	1.44	0.81	44%
	2013	85.12%	3.95	3.63	8%
	2014	82.93%	1.48	0.81	45%
	2015	85.45%	1.20	0.59	51%
PLANT B	2011	0%	0.3	0.4	0%
	2012	0%	1.4	1.7	0%
	2013	0%	0.5	0.2	60%
	2014	0%	1	0.2	80%
PLANT C	2009	98.37%	13	0.6	95%
	2010	98.66%	11	0.6	94%
	2011	98.72%	12	1.4	88%
	2012	97.38%	15	1	93%
	2013	96.91%	16	0.8	95%
	2014	98.25%	11	1.1	90%
	2015	96.72%	9.2	0.56	94%

Table 2.3: Table indicating the SANS 241:2015 specifications for physico-chemical and microbiological parameters.

PARAMETER	SANS 241:2015 STANDARD
ELECTRICAL CONDUCTIVITY	< 170 mS/m
PH	> 5 & < 9.7 pH units
TOTAL DISSOLVED SOLIDS	< 1200 mg/L
TURBIDITY	Operational: < 1 NTU Aesthetic: < 5 NTU
HETEROTROPHIC PLATE COUNT BACTERIA	< 1000 CFU/ml

Table 2.4: Turbidity levels measured after each treatment process of Plant C during June, July and August 2011.

DATE	SOURCE	ACD	AFL	AO	ASW	ASE	AFIL	FW
01 JUNE	25	24	3,5	8,3	2,4	2,5	0,2	0,2
08 JUNE	18	18	20	7,8	2,5	2,6	0,4	0,4
15 JUNE	31	32	11	11	6,3	5,6	1,0	1,7
22 JUNE	20	21	10	9,8	2,3	3,4	0,4	0,4
06 JULY	17	16	5,0	7,7	3,7	3,8	0,8	0,6
13 JULY	12	11	4,7	4,8	2,7	2,1	0,4	0,5
20 JULY	12	10	5,4	3,7	4,7	3,9	1,0	1,4
27 JULY	16	13	7,5	7,3	3,6	3,0	0,8	0,8
03 AUGUST	16	18	3,5	2,7	3,4	3,4	0,7	0,6
10 AUGUST	14	16	5,1	4,7	3,4	2,7	0,8	0,8
17 AUGUST	10	10	2,3	2,4	1,8	1,9	0,8	0,5
24 AUGUST	13	21	4,7	3,6	1,9	2,1	0,6	0,5
31 AUGUST	16	26	3,0	2,8	2,0	2,3	0,6	0,5

ACD: after chemical dosing; AFL: after dissolved air flotation; AO: after ozonation; ASW: after settling west; ASE: after settling east; AFIL: after filtration; FW: final water.

2.3.2 Evaluation of measured turbidity data

Physico-chemical and microbiological parameters were measured at Plant A and B after each step in the treatment process during 2015 and 2016. Table 2.4 indicates the SANS 241:2015 standards for the physico-chemical and microbiological parameters measured in the current study. As indicated in Table 2.5 and 2.6, Plant A indicated no irregularities from the sampling data obtained for EC, pH and TDS as the levels were within the SANS 241:2015 limits. Turbidity was removed gradually after clariflocculation and chlorination, which was expected. However, during March 2016 (Table 2.5) the turbidity level after chlorination were not removed to the accepted SANS 241:2015 limit of < 1 NTU. The HPC levels also exceeded the SANS 241:2015 limits (< 1000 CFU/ml) after chlorination during March 2016 (2800 CFU/ml) and at the point of consumption during November 2016 (1750 CFU/ml).

The EC, pH and TDS levels of Plant B were within the SANS 241:2015 limits during all sampling periods (Table 2.7 to 2.10). A drastic increase in the turbidity level after

filtration, chlorination and in the final water during August 2015 was concerning (Table 2.7). Subsequently, sampling during November 2015, March 2016 and November 2016 was performed after each filtration tank to determine if the problem occurred in all three filtration tanks. As indicated in Table 2.8, 2.9 and 2.10, the turbidity levels had an increase after each filtration tank. These high levels persisted after chlorination and in the final drinking water. The turbidity levels in the drinking water exceeded the SANS 241:2015 limit during all sampling periods, which was concerning. The highest turbidity level in the drinking water of Plant B was during November 2016 (2.04 NTU). The HPC levels were within the specified SANS 241:2015 limits after chlorination during August 2015, November 2015 and November 2016. However, during November 2015 (Table 2.8) and March 2016 (Table 2.9), the HPC levels had a drastic increase at the point of consumption, exceeding the allowable limits.

Table 2.5: Physico-chemical and microbiological results after each step in the treatment process of Plant A for March 2016.

PARAMETER	SOURCE WATER	PRE-CHLORINATION	CLARIFLOCCULATION	CHLORINATION	POC
EC (mS/m)	73.70	74.30	73	75.80	69
pH (pH units)	7.90	7.80	7.60	7.40	8.05
Salinity (ppm)	302	306	304	306	300
TDS (mg/L)	495	494	494	499	490
Temperature (°c)	23.40	23.80	23.60	24.00	22.20
Turbidity (NTU)	3,87 ±0,59	4,38 ±0,42	2,32 ±0,13	**1,63** ±0,09	0.53 ±0,03
HPC bacteria (CFU/ml)	TNTC	800 ± 141.42	450 ± 70.71	**2800** ± 68.87	290 ± 1.53

TNTC: Too numerous to count; ± Standard deviation

Table 2.6: Physico-chemical and microbiological results after each step in the treatment process of Plant A for November 2016.

PARAMETER	SOURCE WATER	PRE-CHLORINATION	CLARIFLOCCULATION	CHLORINATION	POC
EC (mS/m)	65.30	65.70	65.70	66.50	66.90
pH (pH units)	8.43	8.11	8.20	7.99	8.04
Salinity (ppm)	285	286	285	287	291
TDS (mg/L)	464	467	467	468	470
Temperature (°c)	20.30	20.60	20.60	20.20	22.00
Turbidity (NTU)	4.84 ± 0.20	6.27 ± 0.32	3.18 ± 0.11	0.58 ± 0.03	0.59 ± 0.06
HPC bacteria (CFU/ml)	10 425 ± 1449.57	1150 ± 212.13	800 ± 282.84	900 ± 0.0	**1750** ± 70.71

± Standard deviation

Table 2.7: Physico-chemical and microbiological results after each step in the treatment process of Plant B for August 2015.

PARAMETER	SOURCE WATER	AFTER FILTRATION	CHLORINATION	POC
EC (mS/m)	59.10	59	59.20	59.60
pH (pH units)	8.32	8.33	8.27	8.35
Salinity (ppm)	228	229	228	230
TDS (mg/L)	418	418	420	422
Temperature (°c)	18.40	18.20	17.80	17.10
Turbidity (NTU)	0.31 ± 0.04	**1.04** ± 0.16	**1.18** ± 0.12	**1.04** ± 0.21
HPC bacteria (CFU/ml)	5385 ± 1120	0	0	0

± Standard deviation

Table 2.8: Physico-chemical and microbiological results after each step in the treatment process of Plant B for November 2015.

PARAMETER	SOURCE WATER	AFTER FILTRATION			CHLORINATION	POC
		1	2	3		
EC (mS/m)	52.20	52	51.80	52.10	52.50	50.50
pH (pH units)	8.60	8.46	8.44	8.43	8.58	8.46
Salinity (ppm)	227	224	225	226	226	221
TDS (mg/L)	371	368	368	369	371	358
Temperature (°c)	21.10	21.40	20.70	20.80	21.20	27.20
Turbidity (NTU)	0.94 ± 0.25	**1.54** ± 0.21	**2.09** ± 0.30	**1.71** ± 0.23	**1.75** ± 0.39	**2.04** ± 0.06
HPC bacteria (CFU/ml)	5457 ± 1379	2458 ± 2394	555 ± 64	738 ± 159	720 ± 537	**3280** ± 3352

± Standard deviation

Table 2.9 Physico-chemical and microbiological results after each step in the treatment process of Plant B for March 2016.

PARAMETER	SOURCE WATER	AFTER FILTRATION			CHLORINATION	POC
		1	2	3		
EC (mS/m)	52.60	51.70	52.30	52.50	52.80	53.0
pH (pH units)	8.56	8.53	8.67	8.62	8.54	8.52
Salinity (ppm)	227	222	222	224	227	228
TDS (mg/L)	374	365	371	372	373	376
Temperature (°c)	18.40	17.60	17.60	18.10	17.50	21.8
Turbidity (NTU)	0.20 ± 0.05	**2.26** ± 0.10	**1.03** ± 0.14	**0.91** ± 0.26	**2.33** ± 0.02	**1.26** ± 0.06
HPC bacteria (CFU/ml)	14150 ± 1600	5200 ± 613	3883 ± 24	1475 ± 389	**2000** ± 500	**14425** ± 1175

± Standard deviation

Table 2.10: Physico-chemical and microbiological results after each step in the treatment process of Plant B for November 2016.

PARAMETER	SOURCE WATER	AFTER FILTRATION			CHLORINATION	POC
		1	2	3		
EC (mS/m)	48.20	47.80	48.10	48.20	48.60	47.80
pH (pH units)	8.57	8.24	8.5	8.54	8.43	8.51
Salinity (ppm)	208	206	207	208	208	207
TDS (mg/L)	341	340	342	342	344	339
Temperature (°c)	19.10	19	18.70	18.60	19.10	23.80
Turbidity (NTU)	0.37 ± 0.02	**3.36** ± 0.02	**1.41** ± 0.02	**1.21** ± 0.06	**2.25** ± 0.02	**1.46** ± 0.04
HPC bacteria (CFU/ml)	25833 ± 19045	100 ± 0.0	850 ± 778	0	200 ± 141	100 ± 0.0

± Standard deviation

2.4 Discussion

The main goal of water purification plants is to provide safe drinking water to their consumers (Ho *et al.*, 2010). To achieve this goal, the quality of the drinking water should be monitored. Traditionally, the focus of water purification plants is to focus on monitoring the final drinking water and correcting any failures after it has occurred. However, this technique has proved to be inadequate (Tsoukalas & Tsitsifli, 2018). In cases where water quality is not up to standard, identifying which treatment process failed can be difficult (Okeyo *et al.*, 2011). The WHO suggests that a more holistic approach should be followed. This involves the assessment of risks from the source, throughout the water distribution network, to the consumer (World Health Organization, 2017a). In the present study, the HACCP concept was evaluated at three water purification plants (Plant A, Plant B and Plant C), in the North West Province, South Africa.

Firstly, the efficiency of the treatment process at each plant was determined by evaluating historical data from 2009 to 2015, using turbidity levels as an indicator. A study by Ho *et al*. (2010) indicated that turbidity can be a useful parameter to indicate how effective a treatment process is performing. In their study, turbidity and colour were used to determine the quality of drinking water from four different water treatment processes in parallel. The four treatment process streams (full- and pilot scale) consisted of a combination of conventional treatment methods (coagulation, flocculation and sand filtration) and advanced treatment methods (high rate magnetic ion exchange, granular activated carbon, microfiltration and nanofiltration). The effectiveness of each treatment process was evaluated by determining the reduction in the turbidity levels.

Secondly, based on the results obtained from the historical data, the HACCP concept was implemented at Plant A and Plant B by measuring physico-chemical and microbiological parameters after each step in the treatment process. Plant C already implemented the HACCP concept.

Historical data of Plant A, Plant B and Plant C indicated that there were instances where the treatment process failed to adequately remove turbidity from the source water. Even though Plant A had a steady turbidity removal averaging above 80%, during 2009 and 2013, turbidity levels exceeded the SANS 241 limit in the drinking

water, which was concerning (Table 2.2). The HPC levels also exceeded the SANS 241 limit of < 1000 CFU/ml after chlorination and at the point of consumption (Table 2.2). As a result of Plant A only monitoring source and drinking water, determining which of the steps in the treatment process was not efficient in removing the turbidity or HPC levels to an acceptable level, could not be established. Therefore, the implementation of the HACCP concept at Plant A could be a valuable tool to identify the failure in the treatment system timeously.

Evaluation of the historical data for Plant B indicated some serious treatment issues. Most of the time the turbidity levels in the drinking water of Plant B were higher than in the source water. Various physico-chemical and microbiological parameters were measured after each step in the treatment process of Plant B during 2015 and 2016. Even though the EC, pH, salinity and TDS levels did not raise any concern, it was clear that the turbidity level drastically increased after filtration. This trend was seen during all sampling periods of 2015 and 2016. Operators of Plant B could use this information to adjust or change the treatment process to ensure the turbidity levels are within the SANS 241:2015 specification. Applying the HACCP concept at Plant B proved to be useful by indicating where the failures in the treatment processes were.

Plant C, however, monitored turbidity throughout the treatment system. During July 2011 Plant C had a process failure resulting in a turbidity level of 1.4 NTU in the drinking water, exceeding the SANS 241:2015 limit. Each step in the treatment process could be evaluated to determine where the failure in the treatment process was. An unusual increase in turbidity after settling west was noted, which could have contributed to the unsatisfactory turbidity removal for that day. Even though this was an isolated incidence, should it have occurred continuously, the operators of Plant C would have been able to identify and rectify the problem rapidly due to monitoring after each process. This highlighted the importance of continuous monitoring from source to tap.

Similar results were obtained by a study by Okeyo *et al.* (2011). In their study the HACCP concept was applied at three water treatment plants in Gauteng Province, South Africa. Critical control points for the study included raw water sources before coagulation/flocculation, after filtration, after chlorination and at the point of use. Physico-chemical parameters that were measured included turbidity, pH, temperature

and residual chlorine. Microbiological analysis included total and faecal coliforms. Results from the study indicated that the average percentage turbidity removal after sedimentation ranged from 41.1% to 80.5% and from 75.1 % to 97.8% after filtration. Microbiological results indicated the percentage removal of indicator bacteria ranged from 35.8% to 86.1% after sedimentation and after filtration from 74.2% to 97.3%. The results illustrated that pre-disinfection processes failed to remove pathogenic bacteria. The study concluded that applying the HACCP concept, various problems pertaining to water quality were identified within each water treatment plant and preventative measures could be put in place.

Another study by Jagals & Jagals (2004), applied the HACCP concept at two water purification plants in the Free State Province, South Africa. The aim of their study was to evaluate the effectiveness of the treatment process by measuring faecal coliforms and turbidity. The critical control points identified in the study were raw water, sedimentation, filtration and disinfection. Results from the study indicated that the sedimentation and filtration process underperformed with regards to the adequate removal of turbidity and faecal coliforms. The disinfection process, however, removed 100% of the faecal coliforms. The study concluded that obtaining good results in the treated water may mask the presence of process failures. Monitoring only raw and treated water may thus not be sufficient to ensure the production of safe drinking water. Therefore, the authors suggested that a comprehensive HACCP plan should be implemented at both water purification plants, as the preliminary implementation of the HACCP concept proved to be successful in identifying treatment failures.

2.5 Conclusion

The limitations of end product testing are becoming more evident in the water sector around the world (Okeyo *et al.*, 2011). HACCP is a preventative approach, which is more reliable and cost-effective than the reactive approach of focusing on treated water (Dunn *et al.*, 2014; Tsitsifli & Tsoukalas, 2019). Identifying and monitoring critical control points within a treatment process may be useful to identify water quality issues within the treatment plant which may lead to poor water quality production. From previous literature it was evident that the use of HACCP can improve the purification process, as well as improve bacterial compliances, which in turn ensures safe drinking water.

In the current study, evaluating the historical data of Plant A and B highlighted the limitations of only performing end-point testing, as the cause of previous process failures could not be established. The effectiveness of applying the HACCP concept was clearly indicated by the results obtained from Plant B during 2015 and 2016. By monitoring certain parameters after each step in the treatment process, the sources of treatment failures were identified. This information may be useful to the plant operators in finding a solution and preventing future occurrences. The application of the HACCP concept at water purification plants may thus be an effective decision-making tool to ensure the continuous production of safe drinking water.

2.6 Recommendations

- Increase sampling frequencies after each treatment process to obtain more data which may increase the probability to detect treatment process failures.
- Include more water purification plants with similar treatment processes to evaluate the effectiveness of the HACCP concept.
- Implement training programs at local municipalities to educate them on the advantages of applying the HACCP concept.

CHAPTER 3

Artificial Neural Networks: Applications in the drinking water sector

This chapter has been published in *Water Science & Technology: Water Supply:*
O'Reilly, G., Bezuidenhout, C.C. and Bezuidenhout, J.J., 2018. Artificial neural networks: applications in the drinking water sector. *Water Science and Technology: Water Supply*, *18*(6), pp.1869-1887.

3.1 Introduction

Supplying adequate and safe drinking water to communities is currently one of the most important challenges that water purification facilities in developing and developed countries face (Badejo *et al.*, 2015; Gray & Vawda, 2016). To ensure treatment processes are efficient, knowledge of certain contaminants in the water is very important to ensure they are correctly removed (Meng *et al.*, 2015). Most of the water treatment plants use purification technologies developed decades ago (Trussel, 2005). Traditional water purification methods include flocculation, sedimentation, sand filtration and chlorination (Rigobello *et al.*, 2013). Even though these purification methods may be effective, deterioration of source water quality may require advanced treatment methods to ensure that the water is effectively purified (van der Hoek *et al.*, 2014; Ang *et al.*, 2015; Meng *et al.*, 2015). Advanced treatment methods usually include membrane filtration, reverse osmosis, ozonation, activated carbon and advanced oxidation (van der Hoek *et al.*, 2014; Meng *et al.*, 2015). However, advanced treatment technologies are costly (Houtman, 2010), which means poor quality raw water is thus more expensive to treat (Adgar *et al.*, 2000; Ang *et al.*, 2015).

To ensure safe drinking water is produced, the implementation of an effective water quality management program is important, but should not only be compliance driven. For compliance to water quality standards, such as those prescribed by the World Health Organization (WHO), United States Environmental Protection Agency (EPA) and the South African National Standards (SANS 241:2015), the levels of specific parameters are determined. Analysing a large number of variables in aquatic systems can be complex which makes the monitoring of water quality challenging (Ranković *et al.*, 2010; Antanasijevic *et al.*, 2013; Chen & Liu, 2014), particularly for small water supply authorities. Typical monthly monitoring of water quality parameters may also lead to missing values in the data set (Tabari & Hosseinzadeh Talaee, 2015). In areas

where pollution episodes regularly occur, preventative methods, such as automatic monitoring, is an option (Iglesias *et al.*, 2014). However, automatic monitoring could also be costly and time consuming, particularly when the pollution events are sporadic (Iglesias *et al.*, 2014). Thus, qualitative and quantitative decisions based on real data are a challenge for environmental engineers monitoring water quality (Lermontov *et al.*, 2009; Tabari & Hosseinzadeh Talaee, 2015). Water quality modelling is thus a valuable tool to ensure optimum water quality management (Antanasijevic *et al.*, 2013; Vieira *et al.*, 2013).

ANNs are modelling approaches that could be used in predicting the impacts of deteriorating water quality on drinking water purification processes. This could then be used to identify critical parameters as well as steps in the purification processes to be monitored or to be addressed. Whenever there are drastic changes in the water quality, water purification facilities usually rely on past experience or extra bench-scale testing to resolve the problems (Veerapaneni *et al.*, 2010). However, ANNs can be a useful tool for managing certain aspects of the water treatment operation (Veerapaneni *et al.*, 2010). This is due to the fact that ANNs have the ability to do predictions in systems where information on particular interrelationships is inadequate (Veerapaneni *et al.*, 2010). Hidden relationships in historical data can be revealed by using ANNs, which assists in the forecasting of water quality (Najah *et al.*, 2013).

These approaches (ANNs) are not new to the water sector where it has been applied as modelling and forecasting tools (Wu *et al.*, 2014). It has found applications in water engineering, environmental sciences and ecological sciences since the 1990's (Palani *et al.*, 2008; Najah *et al.*, 2013; Antanasijević *et al.*, 2019). Advantages that ANNs bring to water quality modelling include: (i) Model building do not require a physics-based algorithm - this makes the modelling approach faster and more flexible; (ii) Non-linear relationships can be handled properly and without any effort (Tabari & Hosseinzadeh Talaee, 2015); and (iii) user experiences and knowledge can be incorporated in construction of a model (Zhang & Stanley, 1997). The aim of this review is to give an overview on the principles of ANNs, the application of ANNs in the water sector, the current scenario with regard to drinking water and future prospects.

3.2 Principles of ANNs

Artificial Neural Networks are computational techniques that mimic some operational features of the human brain (Haykin, 2009; Vicente *et al.*, 2012; Salari *et al.*, 2018). ANNs are not programmed like conventional computer programs, but they have mechanisms which can learn certain data or patterns (Sarkar & Pandey, 2015). Data in ANNs are connected to each other by weights parallel to synapses (Seth, 2015). Training of the ANN is done by adjusting these connections through a learning algorithm (Seth, 2015). ANN modelling usually consists of the following steps: data collection, data analysis and training of the neural network (Antanasijevic *et al.*, 2013). ANNs can identify intricate nonlinear relationships between input and output data sets (Antanasijevic *et al.*, 2013; Najah *et al.*, 2013). There are various types of Artificial Neural Networks available, but the most commonly used are: Multi-layer Perceptrons (MLPs), Radial Basis Function (RBF), General Regression Neural Network (GRNN), Cascade Forward Networks (CFN) and Kohonen's self-organisations maps (Farmaki *et al.*, 2013; Wu *et al.*, 2014).

MLPs are the most commonly used feed-forward neural networks (Piotrowski *et al.*, 2015; Tabari & Hosseinzadeh Talaee, 2015). The architecture of a typical feed-forward network (Figure 3.1) contains an input layer, a hidden layer and an output layer (Piotrowski *et al.*, 2015; Salami Shahid & Ehteshami, 2016). The neurons in one layer are connected to the next layer, but the neurons of the same layer are not connected to each other (Najah *et al.*, 2013). The back-propagation training algorithm is used most commonly with MLPs (Vicente *et al.*, 2012). During training of the model actual and target output values are compared. By using the back-propagation algorithm, errors resulting from the comparison are propagated backwards through the network, adjusting the weight values and the errors are minimised (Abdulkadir *et al.*, 2012). A feed-forward network trained with the back-propagation algorithm can be referred to as Back Propagation Neural Network (BPNN) (AL-Allaf, 2012). Since the introduction of the feed-forward ANN, research into the application of ANNs has thrived (Maier & Dandy, 2000).

Radial Basis Function (RBF) is similar to the MLP neural network consisting of three layers: an input layer, hidden layer (known as kernel) and an output layer (Hannan *et al.*, 2010). Just like MLPs, each layer is connected to the next layer, but with RBF, the neurons in the hidden and output layer are interconnected to each other by weights

(Sharma *et al.*, 2003; Farmaki *et al.*, 2010). The General Regression Neural Network is a variation of the Radial Basis Function network (May *et al.*, 2008; Hannan *et al.*, 2010). Unlike networks using the Back-propagation algorithm, GRNN does not need a repetitive training procedure (Hannan *et al.*, 2010). It estimates random functions between input and output neurons directly from the training data (Hannan *et al.*, 2010).

Cascade forward networks (Figure 3.2) are similar to feed-forward networks, but each layer is connected to the successive layers by means of a weight connection (Goyal & Goyal, 2011; AL-Allaf, 2012). In other words, not only is layer 1 connected to layer 2 and layer 2 connected to layer 3, but layer 1 is also connected to layer 3 by means of a weight connection (Goyal & Goyal, 2011; AL-Allaf, 2012). The back-propagation algorithm can also be used to update the weights of the layers (Chayjan, 2010; Goyal & Goyal, 2011). The Kohonen's self-organising map (Figure 3.3) consists of only an input layer and a Kohonen layer (Bowden *et al.*, 2005; Farmaki *et al.*, 2013). Each input element is connected to all the other neurons of the Kohonen layer (Farmaki *et al.*, 2013). These networks have the unsupervised ability to learn and organise data without being given associated output values for the input data, hence the term "self-organising" (Mukherjee, 1997; Farmaki *et al.*, 2013).

A Neuro-fuzzy network is a combination of artificial neural networks and fuzzy logic (Rani & Moreira, 2010). Fuzzy logic is a representation of knowledge (obtained from data analysis or expert knowledge) that is based on reasoning that is approximate rather than predicated logic (Christodoulou & Deligianni, 2010). For example, a set of objects or a scenario can be characterised by being true or false. If "true or false" are given values of "1 and 0", fuzzy logic allows grades of characterisation assigned to each object ranging between the values of 0 and 1, basically referred to as "degrees of truth" (Zadeh, 1965; Christodoulou & Deligianni, 2010). One of the most popular neuro-fuzzy methods is the adaptive network-based fuzzy inference system (ANFIS) (Rani & Moreira, 2010).

3.3 Application of ANNs in the drinking water sector

Since the late 1980's into the 1990's, ANNs have been applied in drinking water distribution system management to predict pipe pressure/leakage (Bargiela & Hainsworth, 1989; Vairavamoorthy & Lumbers, 1998), scheduling of booster disinfection (Boccelli *et al.*, 1998) and coagulation/flocculation (Zhang & Stanley,

1999). During the early 21st century ANNs became more popular and were applied to various applications such as membrane filtration (Cabassud *et al.*, 2002), predicting disinfection residuals (Gibbs *et al.*, 2003; Legube *et al.*, 2004), chemical dosing (Valentin & Denœux, 2001) and disinfection by-products (DPBs) (Milot *et al.* 2002).

However, from 2006 the application of ANNs in the drinking water sector has increased dramatically compared to the previous decade. Researchers applied ANNs in fields such as the prediction of water quality parameters, municipal water production/consumption, contamination events and operational costs. Some of the studies from 2006 - 2019 are summarized in Table 3.1.

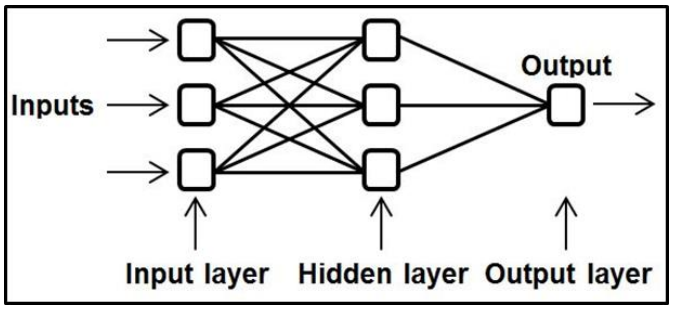


Figure 3.1: Architecture of a feed-forward network (adapted from Najah *et al.* 2013).

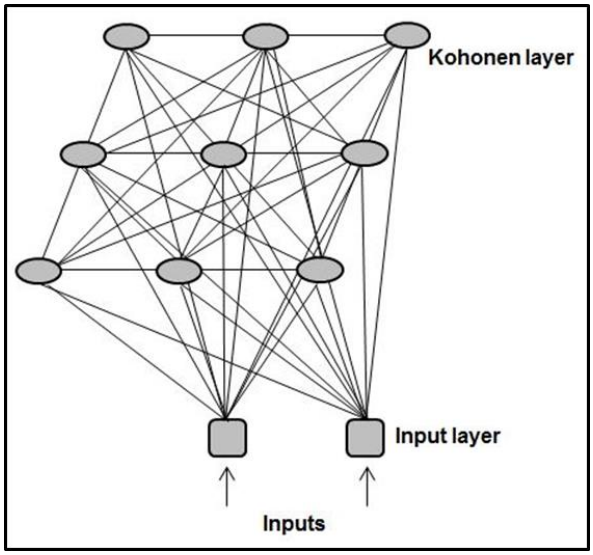


Figure 3 2: Architecture of Kohonen's self-organising map (adapted from Bowden *et al.* 2005).

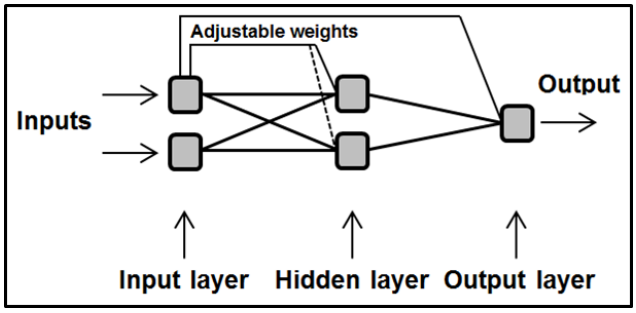


Figure 3.3: Architecture of a Cascade Forward network (adapted from AL-Allaf 2012).

3.3.1 Pipes/infrastructure

As seen in Table 3.1, the most popular area of interest was the prediction of pipes and infrastructure problems at water purification facilities. Several authors used ANNs to predict leakages of pipes (Table 3.1). Pressure in a water distribution network may play a role in aggravating leakage (Makaya & Hensel, 2015). This parameter (pressure) can be measured by sensors, but sensors cannot be placed at every node

which means leaks are not always detected (Ridolfi *et al.*, 2014). Measurements and data analysis of sensor monitoring as well as other leakage detection methods may be time and cost consuming (Wachla *et al.*, 2015). Therefore, ANNs could be used to predict pipe pressure, which may reduce water leakage. Mounce & Machell (2006); Nazif *et al.* (2010); Ridolfi *et al.* (2014) and Makaya & Hensel (2015) used MLP neural networks to detect leakages, whereas Ho *et al.* (2010) used RBF and Wachla *et al.* (2015) used ANFIS (Table 3.1). Results of these studies indicated that ANNs were successfully used to detect pipe leakages.

The study by Nazif *et al.* (2010) concluded that implementing the ANN model to optimize tank water levels can reduce leakage annually by 30%. Ridolfi and colleagues (2014) used ANNs to simulate water pressure at every node in the distribution system. The model is useful to determine which pressure monitoring sensors can be omitted from the distribution system without any loss of information. A study by Ho *et al.* (2010) included earthquake data as major input parameter to predict pipe breakage events and pipeline leakage problems (Table 3.1). LuoDong Township in Taiwan is frequently affected by earthquakes and was therefore chosen as the study site. It was noted that the inclusion of earthquake data yielded a higher prediction performance. The study concluded that the implementation of the model will not only address leakage problems, but also the labour requirement and costs involved in pipe replacement could be reduced.

Pipe failure may lead to financial loss due to repairs and maintenance being done (Jafar *et al.*, 2010). Traditional statistical methods have been used to determine pipe failures, but the main disadvantage of these methods is that they do not usually take all the parameters that may have an influence on pipe failure into account (Tabesh *et al.*, 2009). To overcome this challenge, Tabesh *et al.* (2009), Christodoulou & Deligianni (2010), Jafar *et al.* (2010) and Al-Barqawi & Zayed (2008) used ANNs to predict pipe failures (Table 3.1). Tabesh *et al.* (2009) used MLP and ANFIS neural networks, whereas Christodoulou & Deligianni (2010) used a neuro-fuzzy network and Jafar *et al.* (2010) and Al-Barqawi & Zayed (2008) used BPNNs.

Table 3.1: Summary of applications of ANNs in the drinking water sector (2006 - 2017).

	AUTHORS	ANN/MODEL TYPE	PREDICT/MODEL	INPUT VARIABLES	STUDY AREA
PIPES/ INFRASTRUCTURE	Mounce & Machell (2006)	MLP	Burst detection of pipes	Flow, pressure	United Kingdom
	Martínez *et al*. (2007)	Feed-forward	Operational control settings	Pump and valve settings, storage tank water level & demands of district metering areas	Spain
	Rao & Alvarruiz (2007)	Feed-forward	Operational control settings	Pump settings, valve settings, demands for the various district-metering areas, storage tank water levels	-
	Rao & Salomons (2007)	Feed-forward	Operational control settings	Pumping power, pressures, flows, costs, penalties	-
	Salomons *et al*. (2007)	Feed-forward	Operational control settings	Pumping status, valve settings, storage levels & demands of district metering areas	Israel
	Al-Barqawi & Zayed (2008)	BPNN	Performance of municipal water mains	Pipe length, size, age, type of material, depth, slope & type of sewer	Canada
	Tabesh *et al*. (2009)	MLP & ANFIS	Pipe failure and mechanical reliability	Pipe length, diameter, age, installation depth & hydraulic pressure	Iran
	Christodoulou and Deligianni (2010)	Neuro-fuzzy	Performance of pipes and failure analysis	Pipe parameters: previous breaks, length, materials & diameter Traffic parameters: traffic load, pipe's proximity to a highway, underground railway & roadway or block intersection	New York City (U.S.A) & Limassol (Cyprus)
	Ho *et al*. (2010)	RBF	Pipeline replacement and leakage	Pipe diameter, material & seismic factor (earthquakes)	China
	Jafar *et al.* (2010)	BPNN	Model the failure of the pipes	Historical failure, hydraulic pressure, characteristics, location of pipes & soil type	France
	Nazif *et al*. (2010)	MLP	Pipe pressure	Elevation of nodes, storage tank levels & demand at each node	Iran

Table 3.1 continued

	AUTHORS	ANN/MODEL TYPE	PREDICT/MODEL	INPUT VARIABLES	STUDY AREA
	Farokhzad *et al.* (2012)	MLP	Faults in centrifugal water pump	Mean, standard deviation, sample variance, kurtosis, skewness, root mean square, crest factor, slippage & fourth, fifth and sixth central moment	Iran
	Ridolfi *et al.* (2014)	Three layered feed-forward (MLP)	Pressure distribution	Water pressure	Italy
	Makaya & Hensel (2015)	MLP	Flow dynamics to detect leakage	Flow logging data	Zimbabwe
	Wachla *et al.* (2015)	ANFIS	Leakage detection	Water flow rates	Poland
	Kamiński, Kamiński, & Mizerski (2017)	MLP	Tool for renovation decisions in water supply	Pipe diameter, material, age, failure rate, forces affecting the pipeline	Poland
	Dawidowicz (2018)	MLP	Diameter of water distribution pipes	Length of the pipe, flow rate at the beginning, flow rate at the end, absolute roughness coefficient	Poland
	Jang *et al*. (2018)	MLP	Leakage ratio	Mean pipe diameter, pipe length, water supply quantity, deteriorated pipe ratio, demand energy ratio, no. of leaks	Korea
COAGULATION/ FLOCCULATION DOSAGE	Wu & Lo (2010)	MLP	PAC	Turbidity, temperature, colour, pH & coagulant dosage	Taiwan
	Gholikandi *et al.* (2011)	BPNN	PAC	Influent turbidity, Poly Aluminum Chloride (PAC) dosage & coagulant kinds	Iran
	Heddam *et al*. (2011)	RBF, GRNN	Aluminum sulphate	Turbidity, EC, pH, temperature, DO & ultraviolet absorption	Algeria

Table 3.1 continued

	AUTHORS	ANN/MODEL TYPE	PREDICT/MODEL	INPUT VARIABLES	STUDY AREA
	Dharman *et al*. (2012)	Feed-forward	PAC	Plant flow, raw water alkalinity, TOC, pH, total hardness, turbidity, iron, fluorides, hardness (calcium), temperature, polymer $FeCl_2$, flow at lock 10, pH adjustment, disinfectant (pre ammonia) & coagulant polymers	Kentucky
	Heddam *et al.* (2012)	ANFIS	Aluminum sulphate	Turbidity, EC, pH, DO, temperature & ultraviolet absorption	Algeria
	Naidoo & van der Walt (2013)	Feed-forward	Polymeric coagulant	Turbidity, pH, alkalinity & colour	South Africa
	León-Luque *et al.* (2016)	Not mentioned in article	Aluminum sulphate	Turbidity, pH, EC, temperature, alkalinity & colour	Colombia
FILTRATION EFFICACY	Chen & Kim (2006)	RBF, BPNN	Membrane filtration: predict permeate flux decline	Particle size, solution pH, transmembrane pressure, elapsed time & ionic strength	Hawaii
	Curcio *et al*. (2006)	Feed-forward	Membrane filtration: model permeate flux decay	Operating time, sampling time & inlet flow rate	Italy
	Griffiths & Andrews (2011)	MLP	Granular media filtration: predict post-filtration particle counts and settled water turbidity	Temperature, pH, filter flow rate, filter head loss, filter run time, settled water turbidity & pre-chlorination dosage	Canada
	Kabsch-korbutowics & Kutylowska (2011)	MLP	Membrane filtration: predict turbidity retention coefficient during ultrafiltration	Feed water turbidity, turbidity in the tank, pH, temperature in the tank, transmembrane pressure & permeate flux	Germany
	Tashaouie *et al*. (2012)	MLP	Performance of pressure filters	Turbidity, filtration rate & pressure	Iran

Table 3.1 continued

	AUTHORS	ANN/MODEL TYPE	PREDICT/MODEL	INPUT VARIABLES	STUDY AREA
	Madaeni *et al*. (2015)	MLP	Performance of reverse osmosis plant	Time, conductivity, transmembrane pressure & flow rate	Iran
	Corbatón-Báguena *et al.* (2016)	Feed-forward	Membrane filtration: permeate flux decline	Transmembrane pressure, cross-flow velocity, operating time, flux normalisation & fouling indicator	Spain
MUNICIPAL WATER DEMAND	Adamowski (2008)	MLP	Daily water demand	Water demand, temperature, rainfall data	Canada
	Firat *et al.* (2009)	GRNN, RBF, Feed-forward	Monthly water use	Average monthly water bill, population, number of households, gross national product, temperature, rainfall, humidity & inflation rate	Turkey
	Yurdesev & Firat (2009)	ANFIS	Monthly water use	Average monthly water bill, population, number of households, gross national product, temperature, rainfall, humidity & inflation rate	Turkey
	Adamowski & Karapataki (2010)	MLP	Weekly water demand	Water demand, temperature, rainfall data	Cyprus
	Firat *et al.* (2010)	GRNN, CCNN, Feed-forward	Monthly water consumption time series	Historical water consumption data	Turkey
	Ajbar & Ali (2015)	MLP	Monthly & annual water demand	City population, housing density, personal income, maximum monthly temperature & number of monthly visitors	Saudi Arabia
	Pacchin *et al.* (2019)	MLP	Short-term (hourly) water demand	Water demand	Italy

Table 3.1 continued

	AUTHORS	ANN/MODEL TYPE	PREDICT/MODEL	INPUT VARIABLES	STUDY AREA
DISINFECTION RESIDUALS	Bowden *et al.* (2006)	GRNN	Residual chlorine	Flow, turbidity, pH, temperature & chlorine	Australia
	Gibbs *et al.* (2006)	MLP, GRNN & SOM	Residual chlorine	Water temperature, flow, chlorine concentration, dissolved organic carbon (DOC), ultraviolet absorbance (UV_{254}) & time of measurement	Australia
	May *et al.* (2008)	GRNN	Residual chlorine	Distribution plant 1: Chlorine at different nodes, tank level, pump flow rate & pipe flow rate Distribution plant 2: Free chlorine, pH, turbidity, tank level, outlet flow & temperature	Australia
	Soyupak *et al.* (2011)	MLP	Residual chlorine	pH, EC, turbidity, water flow rates, temperature & free residual chlorine	Turkey
	Wu *et al.* (2011)	GRNN	Residual chlorine and free ammonia levels	Chlorine & free ammonia	Australia
	Cordoba *et al.* (2014)	MLP	Residual chlorine	Temperature, pH, turbidity, flow, initial chlorine & free chlorine	Czech Republic
	Zounemat-Kermani *et al*. (2018)	MLP, RBF	Residual chlorine	Water flow, daily chlorine consumption, daily residual chlorine	Iran
WATER QUALITY	Mustonen *et al*. (2008)	SOM	Water quality changes	Particle measurement & EC	Finland
	Vicente *et al.* (2012)	MLP	Nitrate, manganese, sodium and potassium	pH & conductivity	Portugal

Table 3.1 continued

	AUTHORS	ANN/MODEL TYPE	PREDICT/MODEL	INPUT VARIABLES	STUDY AREA
	Juntunen *et al.* (2013)	SOM	Various physico-chemical parameters	Wells water level & flow, lake flow, lake surface level, total inflow, solvent water, lime feed, $KMnO_4$ dose, Al dose, CO_2 feed, temperature & flow Raw water: turbidity, pH, alkalinity, EC, COD, iron, manganese & aluminium Treated water: pH, alkalinity, EC, free chlorine, COD, aluminium, iron, manganese & calcium	Finland
	Rak (2013)	MLP	Turbidity	Raw water turbidity, water flow, water retention level, daily rainfall & reservoir temperature	Poland
	Gaya *et al.* (2017)	Feed-forward	Turbidity	Influent parameters: Free CO_2, calcium, suspended solids, hardness, chloride, conductivity, TDS, pH, turbidity	Nigeria
DBPS	Kulkarni and Chellam (2010)	BPNN	THM, haloacetic acids (HAA), total organic halide (TOX)	UV_{254}, contact time, temperature, pH, TOC, bromium & chlorine dose	U.S.A
	Ye *et al.* (2011)	BPNN	THM, HAA	Residence time, water temperature, pH, UV_{254}, TOC, bromium concentration & residual free chlorine	China
	Singh & Gupta (2012)	Feed-forward, RBF	THM	pH, temperature, contact time, Br concentration & dissolved organic carbon normalized chlorine dose (Cl_2/DOC)	India
	Karadurmuş *et al*. (2018)	Feed-forward	Bromate removal	Particle size, amount of activated carbon, height & diameter of column, volumetric flowrate, initial concentration	Turkey

Table 3.1 continued

	AUTHORS	ANN/MODEL TYPE	PREDICT/MODEL	INPUT VARIABLES	STUDY AREA
ORGANIC MATTER REMOVAL	Bieroza *et al.* (2011)	SOM, BPNN	-	Organic matter fluorescence data	United Kingdom
	Bieroza *et al*. (2012)	SOM, BPNN	-	Organic matter fluorescence data	United Kingdom
CONTAMINATION EVENT	Perelman *et al*. (2012)	BPNN	-	EC, pH, temperature, turbidity, total chlorine & TOC	U.S.A (CANARY database)
	Arad *et al*. (2013)	BPNN	-	EC, pH, temperature, turbidity, total chlorine & TOC	Israel/U.S.A (CANARY database)
ORGANIC & INORGANIC POLLUTANTS	Cauchi *et al.* (2011)	Feed-forward	Anthracene, naphtalene, phenanthrene, cadmium, lead & copper	Anthracene, phenanthrene, naphtalene, cadmium, lead & copper	United Kingdom
RESIDUAL ALUMINUM	Tomperi *et al.* (2013)	MLP	-	Raw water temperature, colour, pH, potassium permanganate ($KMnO_4$) & Poly-Aluminum Chloride/Potassium permanganate ratio (PAC/$KMnO_4$)	Finland
COST OF TREATMENT PLANT	Marzouk & Elkadi (2016)	MLP	Construction cost	Soil type, clarifier type & land property	Egypt
PERFORMANCE EFFICIENCY OF TREATMENT PLANT	Saha *et al.* (2017)	Not mentioned in article	Most important parameter of a water treatment plant	Amount of intake water, time of treatment, discharge rate, amount of output water, efficiency of clariflocculator, filter bed, chlorination unit & channel efficiency	India

In all these studies, ANNs were successfully applied to predict pipe failures, which may reduce financial losses. Tabesh and colleagues (2009) also concluded that even though both the MLP and ANFIS models were able to predict pipe failures successfully, the MLP model slightly outperformed the ANFIS model. Christodoulou & Deligianni (2010) not only used pipe parameters as input variables, but also included traffic parameters as well (Table 3.1). This was due to the study areas being subjected to heavy traffic. The use of the neuro-fuzzy network made it possible for the authors to establish a repair-or-replace rule and to determine to which areas priority should be given. Inspection of existing water mains is costly and time-consuming. Therefore, Al-Barqawi & Zayed (2008) developed their model into a user friendly, web-based condition rating tool which will benefit municipal engineers, consultants and contractors.

Martínez *et al.* (2007) and Salomons *et al.* (2007) used ANNs to optimize operational control settings (Table 3.1). These two studies formed part of a Potable Water Distribution Management (POWADIMA) research project. Previous studies by Rao & Alvarruiz (2007) and Rao & Salomons (2007) indicated that it was possible to form a near-optimal control process for a small, hypothetical water distribution network by using ANNs (Table 3.1). The next step was to apply the methodology from that study to a real network. Hence, the first of the two case studies was performed by Salomons *et al.* (2007) and included data from the Haifa-A distribution network located on Mount Carmel in Israel (Table 3.1). The second case study was performed by Martínez *et al.* (2007) and included the Valencia water distribution network in Spain (Table 3.1). Haifa-A is smaller than the Valencia distribution network and due to its geographic convenience and relationship with the Municipal Department of Water, Sewage and Drainage, it was chosen to be the first of the two case studies. Results of both studies indicated that ANNs were useful tools to optimize operational control settings, which could reduce annual operating costs by around 25% for Haifa-A and 17% for Valencia.

Centrifugal pumps play a significant role in the production process and early detection of faults may help to prevent system shutdowns, human fatalities and material damage (Farokhzad *et al.* 2012). Vibration signals are often used in fault diagnosis systems of rotating machinery. However, human expertise to convert vibration data into maintenance information is sometimes unavailable (Farokhzad *et al.* 2012). Therefore, Farokhzad *et al.* (2012) applied a MLP network to predict faults in centrifugal water

pumps by using vibration condition monitoring. The study concluded that the ANN was able to predict faults, based on vibration differences, with 100% accuracy.

Kamiński *et al.* (2017) used a feed-forward MLP network as a decision-making tool for renovation needs of a water supply in Poland (Table 3.1). Failure of distribution pipes contribute to half of all failures in a water supply system (Kamiński *et al.* 2017). Therefore, to avoid breakdowns the pipes need to be kept in a good condition. The study indicated that, should expert human advice be absent, artificial neural networks could be successfully implemented to aid in the formation of renovation plans. This will ensure that the water purification plant is maintained and can operate efficiently.

Designing water treatment plants involves hydraulic calculations to determine the diameter of the pipes that will be distributing the water. Even though computer programs are available to select the diameters, the task of choosing the final diameters lies with the designer (Dawidowicz, 2018). A study by Dawidowicz (2018) applied a MLP neural network for the prediction of the diameter of water pipes after hydraulic calculations were made (Table 3.1). Data from 36 water distribution systems was used. Although the final diameter selection will be done by the person undertaking the hydraulic calculations, the study concluded that the MLP network was able to predict the pipe diameters accurately, which could assist the designer to select appropriate pipe parameters.

One of the factors influencing water treatment management is the leakage of tap water supply (Huang *et al.*, 2018; Jang *et al.*, 2018). Jang *et al*. (2018) applied principal component analysis (PCA) and ANN for the estimation of leakage ratio (Table 3.1). Results indicated that using the parameters calculated by PCA improved the accuracy of the ANN model. The study concluded that the combination of PCA-ANN may be useful in predicting leakage, which may improve the operation and management of water distribution systems.

3.3.2 Coagulation/flocculation dosage

The application of ANNs in the area of coagulation management during water purification has increased. The required coagulation dosage is usually determined by using traditional jar tests. However, jar tests can be time consuming and water samples have to be taken regularly, relying on manual intervention. If the quality of the raw water changes, operators have to perform a new jar test (Lamrini *et al.*, 2005). In

earlier studies, the prediction of aluminium sulphate dosages was the main coagulant that was predicted, but over the past decade, poly aluminium chloride (PAC) has also proven to be popular (Table 3.1). Feed-forward networks remained the ANN of choice in these studies, but other ANNs were also explored. Heddam *et al.* (2011) compared RBF and GRNN for predicting aluminum sulphate dosing at a drinking water treatment plant in Boudouaou, Algeria. Results indicated that the GRNN consistently outperformed the RBF network. The study concluded that GRNN is an effective tool for modelling coagulant dosage and can be a timesaving option when compared to the usual jar tests.

ANNs have the advantage of being efficient in adapting and learning, but have the negative aspect of being "black box" models due to their lack of transparency in making certain decisions (Delpla *et al.*, 2019). Fuzzy logic, on the other hand, is not efficient in learning, but have the advantage of approximate reasoning (Heddam *et al.* 2012). ANFIS combines the advantages of these two methods making it a very efficient modelling tool. For this reason, Heddam *et al.* (2012) performed a study where aluminum sulphate was predicted, but they used ANFIS as modelling tool. The same water treatment plant and input variables were used as in the 2011 study (Heddam *et al.*, 2011) (Table 3.1). It was found that ANFIS was able to predict the coagulant dosage successfully and the authors suggested that ANFIS might also be used instead of jar tests due to its quick responsive tools, low cost and applicability in a real-time process.

Wu & Lo (2010) and Dharman *et al.* (2012) used feed-forward networks to predict optimal PAC dosage, whereas Gholikandi *et al.* (2011) used a BPNN (Table 3.1). Results of the studies indicated that the various ANNs were able to predict PAC dosage levels accurately. Wu & Lo (2010) also concluded that the prediction model is useful when information on influent water quality is not provided. Dharman *et al.* (2012) noted that the ANN model outperformed the multiple linear regression (MLR) model and provides a quicker response to changing influent data. Therefore, time-consuming jar tests should be used only to crosscheck the validity of ANN predictions during periodic re-training of the model.

3.3.3 Filtration efficacy

During the early 21st century, ANNs were used to predict the efficacy of membrane filtration in water purification facilities. Over the past decade, studies in this area continued, but the performance of granular media filtration and pressure filters have also been included (Table 3.1). Feed-forward networks proved to be the most popular ANN to predict membrane fouling (Table 3.1). Membrane fouling may lead to increased energy, operational and maintenance costs (Gao *et al.*, 2011). Therefore, Curcio *et al.* (2006) and Corbatón-Báguena *et al.* (2016) used feed-forward networks to predict permeate flux decay which proved to be successful. Chen & Kim (2006), however, applied RBF and BPNN models. In their study, a comparison was made between these two ANNs and between the ANNs and a multiple regression method. Results indicated that the RBF neural network outperformed the BPNN and multiple regression models and was able to predict permeate flux with a limited number of training points.

In the studies by Griffiths & Andrews (2011) and Tashaouie *et al.* (2012), both used MLP to determine the performance of granular media filtration and pressure filtration, respectively (Table 3.1). Even though the type of filters used varied, results of both studies indicated that ANNs were able to successfully predict the efficacy of the filters. The ANN models established by Griffiths & Andrews (2011), were implemented into an online optimization application and installed at the Elgin Area water purification facility in Canada to monitor and optimize filtration conditions. In the study by Kabsch-Korbutowicz & Kutylowska (2011) MLP was used to determine the turbidity retention coefficient after integrated coagulation/ultrafiltration process (Table 3.1). Results indicated that the ANN was able to predict the turbidity retention coefficient successfully and that transmembrane pressure played a major role in the prediction model. The authors suggested that the created model can be used for forecasting quality parameters of permeate in hybrid processes, but the conditions of the membrane processes and input variables should be similar.

The operating conditions of reverse osmosis (RO) are very important to ensure efficient performance of other processes such as membrane filtration (Madaeni *et al.* 2015). Madaeni and colleagues (2015) used a MLP network to determine the performance of a RO plant by predicting process performance degradation (Table 3.1). The study concluded that the ANN was able to accurately predict long-term

performance degradation, which is useful for RO process control. Determining the efficacy of filtration is important, because membrane fouling or ineffective filtration may lead to deterioration in the produced water quality (Chen & Kim 2006; Griffiths & Andrews 2011).

3.3.4 Municipal water demand

One of the areas where the application of ANNs has increased is the prediction of municipal water demand. Globally source water has become stressed due to factors such as climate change, population growth and increased water consumption (Adamowski & Karapataki, 2010). For planning and management of water resources, it is important to know what the future needs for drinking water may be (Ajbar & Ali, 2015). Various authors have used ANNs to predict short- and long-term water demands (Table 3.1). Adamowski (2008) used a MLP network to predict daily water demand in the Ottawa West Center pressure zone in Canada. Summer water demand levels in this region indicated an increase from 67.8 ML/day in 1993 to 109.3 ML/day in 2002, which was an indication of the variability in the water demand. For this reason, and the fact that research into daily water prediction was limited, the authors were motivated to use an ANN to develop a prediction model. Results indicated that the ANN was able to predict daily water demand and outperformed the MLR model. The study also concluded that the daily water demand correlated better with rainfall occurrence rather than rainfall levels. The latter statement was later challenged by Adamowski & Karapataki (2010). Their challenge was based on a study by Bougadis *et al.* (2005) which arrived at a different conclusion. Adamowski & Karapataki (2010) compared different MLP networks with a MLR model to predict weekly water demand for two regions in Cyprus (Table 3.1). Results of the study concurred with those of Adamowski (2008).

Firat and colleagues applied various ANN models for the prediction of monthly water demand during 2009 and 2010 for the metropolitan area of Izmir, Turkey (Table 3.1). In the study by Firat *et al.* (2009), GRNN, RBF and feed-forward neural networks were compared. This study was followed by a study by Yurdusev & Firat (2009) where similar input variables were used, but the ANFIS network was applied (Table 3.1). The studies concluded that the GRNN and ANFIS model with three input variables (monthly water bill, population, monthly average temperature) gave the best results for forecasting monthly water consumption. From these studies, Firat *et al.* (2010)

identified the need to compare GRNN, cascade correlation neural network (CCNN) and feed-forward neural networks for modelling monthly water consumption time series (Table 3.1). Various combinations of historical monthly water consumption values were used as input data. Results indicated that the CCNN outperformed the other models and was able to successfully forecast monthly water consumption time series.

Ajbar & Ali (2015) predicted monthly and annual water demand for Mecca city, Saudi Arabia (Table 3.1). Saudi Arabia is an arid country, which depends on costly desalination plants to satisfy water demands. With a large number of tourists visiting Mecca city every year and a lack of effective water management policies, the authors saw the importance to predict the future water demand. The MLP model was able to predict monthly and annual water demands successfully. This may be a useful tool for optimal operation of urban water systems. However, the authors stated that municipal data might be influenced by unforeseen leaks, changing policies and social habits.

More recently, Pacchin *et al.* (2019) compared data-driven and pattern-based techniques for the forecasting of hourly water demand (Table 3.1). For the ANN model, a MLP network was used. Results of the study indicated that both data-driven and pattern-based models were able to forecast hourly water demand. Even though preference was made to using the pattern-based models, the ANN model was able to forecast accurate results with minimal differences compared to the other models used.

3.3.5 Disinfection residuals

Applications of ANNs to determine residual chlorine levels have also increased during the past decade, especially in Australia (Table 3.1). In many Australian studies, the GRNN was the preferred ANN. Bowden *et al.* (2006) used a GRNN to forecast chlorine residuals in the Myponga distribution system in South Australia. Results indicated that the GRNN model was able to forecast chlorine levels very accurately for up to 72 hours in advance. Their study also concluded that the GRNN outperforms the MLR model. Based on these results, May *et al.* (2008) and Wu *et al.* (2011) used GRNNs in their studies as well (Table 3.1). Even though the main focus of the study by May *et al.* (2008) was the improvement of the methodology in developing ANN models, the authors also found the GRNN to be successful in predicting residual chlorine levels.

In the study of Gibbs *et al.* (2006), a comparison between MLP, GRNN and SOM was made for the prediction of residual chlorine levels in the Hope Valley distribution system, South Australia. Results of this study, however, found the MLP model to consistently outperform the other models. Soyupak *et al.* (2011) and Cordoba *et al.* (2014) also used MLPs in their studies (Table 3.1) and found that it was able to predict residual chlorine levels successfully, but Cordoba *et al.* (2014) concluded that the model from their study can only be used to predict chlorine decay for that specific study area.

Some distribution systems use chloramines as disinfectant, which may cause free ammonia levels in the water. Nitrifying bacteria can use the free ammonia as nutrient source which may cause nitrate levels in the water to increase and have various health effects in humans (Wu *et al.* 2011). Therefore, Wu *et al.* (2011) used a GRNN not only to predict residual chlorine levels, but free ammonia levels as well (Table 3.1). Results indicated that the GRNN was able to predict chlorine levels, but due to noisy and inaccurate ammonia data, the model performed poorly for the prediction of free ammonia. The authors suggested accurate free ammonia analysers are required to obtain accurate data for the development of a successful ANN model.

Zounemat-Kermani and colleagues (2018) investigated the use of two ANN models (MLP & RBF), a support vector machine (SVM) and a classification and regression tree (CART) for the prediction of chlorine levels in four distribution systems (Table 3.1). Their study indicated that the CART model outperformed the SVM and RBF models. However, the MLP model was superior to the other models. The ANN model provided the accuracy needed by disinfection systems to predict residual chlorine levels.

3.3.6 Water quality

Over the past decade, interest into the prediction of water quality parameters has increased. Online sensors are able to measure various water quality parameters continuously. However, this means large amounts of data with different time measurements are accumulated which makes pinpointing abrupt changes in water quality challenging (Mustonen *et al.*, 2008). Therefore, Mustonen and colleagues (2008) used a SOM to evaluate water quality changes of online data due to biofilm detaching in a pilot drinking water distribution system (Table 3.1). Results indicated the SOM was able to separate sudden changes in the data from normal data. The

authors suggested their research could be used to develop alert systems or prediction models for controlling water quality.

Data obtained during a water treatment process may be complex due to the non-linear relationships of all the variables (Juntunen *et al.*, 2013). Hence, Juntunen and colleagues (2013) also used a SOM to model water quality in a treatment process (Table 3.1). The study concluded that the SOM was able to comprehensively indicate important characteristics of large data sets. This can be useful to determine the most essential states of water treatment systems, to predict the performance of the process and to use it as a graphical monitoring tool (Juntunen *et al.* 2013). In the study by Vicente *et al.* (2012), the authors used a MLP network to predict nitrate, manganese, sodium and potassium (measured less frequently) using only pH and conductivity (measured more frequently) as input variables (Table 3.1). Results indicated that the MLP model successfully predicted the four parameters with conductivity being the most important input variable.

Turbidity is one of the basic parameters for assessing water quality. During rainfall seasons or spring thawing, water levels may raise and increase turbidity levels. To predict turbidity allows operators to optimize treatment methods in advance. Rak (2013) and Gaya *et al.* (2017) used neural networks to predict turbidity in a treatment plant (Table 3.1). Rak (2013) used a MLP to predict turbidity during the treatment process. Results of the study indicated that the ANN was able to predict turbidity levels successfully. The study also concluded that the model could be useful to predict other parameters, such as pH and colour. Gaya *et al.* (2017) used a Hammerstein-Weiner model and a neural network to predict turbidity in a water treatment plant. The study concluded that the feed-forward neural network outperformed the Hammerstein-Weiner model. The neural network was able to predict turbidity accurately and had a Mean Absolute Percent Error (MAPE) of 12.82%, whereas the Hammerstein-Weiner model had a MAPE of -45.17%. Even though both these studies predicted turbidity, different input parameters were used (Table 3.1).

3.3.7 Disinfection by-products

DBPs may form during the disinfection process and may pose a health risk to consumers. In the studies by Kulkarni & Chellam (2010) and Ye *et al.* (2011), BPNNs were used to predict various DBPs (Table 3.1). These studies had similar input

variables and results for both studies indicated that ANNs were able to predict DBPs successfully.

In the study by Singh & Gupta (2012), however, two different ANNs were compared with support vector machine (SVM) and gene expression programming (GEP) models to predict THM (Table 3.1). Even though all the models were able to predict THM, the study concluded that the SVM slightly outperformed the ANN and GEP models. It was also found that pH followed by contact time had the highest effect on THM formation. Nevertheless, ANNs were useful tools to predict DBP levels and may assist drinking water facilities during design and operation decisions to meet the required DBP standards (Kulkarni & Chellam 2010).

During ozonation bromide may be transformed into carcinogenic bromate (World Health Organization, 2017a; Karadurmuş *et al.*, 2018). The removal of this disinfection by-product is thus important. Karadurmuş *et al.* (2018) estimated the percentage of bromate removal by using ANN (Table 3.1). Their study concluded that the feed-forward ANN was able to predict bromate removal from drinking water. Therefore, the use of neural networks could be an effective tool to improve the accuracy of bromate removal.

3.3.8 Organic matter removal

Even though research into the removal of DBPs has been done, another contributing factor to the formation of DBPs is organic matter. This is due to chlorine reacting with organic matter present in the water which could lead to the formation of THMs (Bieroza *et al.* 2012). Usually organic matter is removed during treatment processes such as coagulation, flocculation, clarification, filtration and granular activated carbon processes. However, these processes may sometimes only reduce the level of organic matter. Methods for quantification of organic matter are laborious (Bieroza *et al.* 2011). Therefore, Bieroza and colleagues (2011; 2012) used ANNs to predict the levels of organic matter removal by using fluorescence data (Table 3.1).

The study during 2011 provided the first insight for using different data mining techniques, where advanced multiway analysis (parallel factor analysis [PARAFAC], principal component analysis [PCA] & partial least squares [PLS]) and ANN approaches (BPNN & SOM) were compared (Bieroza *et al.* 2011). Results indicated that little difference between advanced and conventional peak-picking methods. In a

follow up study during 2012, the authors used the same data, but added the stepwise regression (SR) calibration algorithm (Bieroza *et al.* 2012). Results were similar than the previous study, indicating that PLS and BPNN models are both useful to predict organic matter removal. However, the study also indicated that, unlike the peak-picking methods, the SOM model enables advanced interpretation of fluorescence data.

3.3.9 Contamination events

ANNs have also been applied for the prediction of contamination events. Perelman *et al.* (2012) applied a BPNN network to predict possible contaminants in a water distribution system, based on online data (Table 3.1). An event detection algorithm using Bayesian analysis was established to detect abnormal behaviour of water quality parameters when exceeding a fixed threshold value. The algorithm was able to numerically and graphically indicate the possibility of a quality fault based on single and multiple measured water quality time series. The authors, however, stated that the model's performance needed improvement and a dynamic threshold method should be analysed. Arad *et al.* (2013) aimed to improve the study by Perelman and colleagues. Even though the same type of ANN and input variables were used (Table 3.1), Arad and colleagues (2013) included online and offline data and implemented the dynamic thresholds method by utilizing genetic algorithm (GA), where after Bayesian analysis were used to detect contamination event probability. The study concluded that with appropriate preparation, the method may be implemented at any water distribution system and may also provide statistical and visual indications of contaminant events. It was also noted that the dynamic threshold method was superior to the fixed threshold method.

3.3.10 Organic & inorganic pollutants

In the study by Cauchi *et al.* (2011), three polynuclear aromatic hydrocarbons and three heavy metals were quantified and predicted using a feed-forward neural network (Table 3.1). These parameters were selected due to their use in industrial processes and correlation with industrial sites. Their presence in water is of great concern as they have various health effects (Cauchi *et al.* 2011). When a water sample is measured with an analytic instrument, it is possible for pollutants with similar properties to have overlapping peaks, which makes it difficult to distinguish between them. To overcome this problem, Cauchi and colleagues applied a feed-forward ANN. Results indicated

that the ANN was able to accurately quantify and predict these pollutants simultaneously.

3.3.11 Residual aluminium, cost & performance efficiency

MLP networks have also been applied to predict residual aluminium levels, to determine the construction cost as well as the performance efficiency of water treatment plants (Table 3.1). Water treatment plants can use aluminium salts as a coagulation chemical. High levels of residual aluminium may have several health effects (World Health Organization, 2003). Tomperi *et al.* (2013) compared MLR and MLP models for the prediction of residual aluminium (Table 3.1). Even though both models were able to predict residual aluminium levels fairly accurate, the MLR model outperformed the MLP model. It was also concluded that raw water temperature, $KMnO_4$ and PAC/$KMnO_4$-ratio had the highest correlation with residual aluminium. The authors suggested the models could be used to create an early-warning system to give additional information to process operators.

With construction of a new water treatment plant, preliminary information on the costs is not always available. In Egypt, stakeholders often need to estimate construction costs which lead to high estimation variability (Marzouk & Elkadi, 2016). Therefore, Marzouk & Elkadi (2016) used a MLP network to model construction costs (Table 3.1). Various models were developed and the model with the lowest MAPE was chosen. In this case, the best model had a MAPE value of 21.18%, which is considered reasonable for cost estimation. The study concluded that the ANN was able to successfully predict the cost estimation, which may reduce the resources and time spent on the estimation process. The authors also suggested that detailed estimates could be compared by using this model as a benchmark.

Poor water quality and water shortages are two major challenges that India are continually facing. The optimization of water treatment processes and the prediction of water quality plays an important role to ensure good quality water is supplied to consumers (Saha *et al.*, 2017). Therefore, Saha *et al.* (2017) used a Non-structural Fuzzy Decision Support System (NSFDSS) as well as a neural network to determine the performance efficiency of a water treatment plant (Table 3.1). NSFDSS is a multi-criteria decision-making method (MCDM) which determines the comparative weight between parameters. In this study, the aim was not to compare the NSFDSS with the

neural network, but rather to use the neural network to determine the index weights by training the model after which the model output was predicted. Results indicated that the ANN was able to successfully predict the model output. The study concluded that the efficiency of the clariflocculator was the most significant parameter.

3.4 Application of ANNs in the water sector: Scenario in South Africa

In South Africa, the application of ANNs in the water sector is very limited, especially in the drinking water sector. Studies pertaining to environmental water include the prediction of: **streamflow** (Ilunga & Stephenson, 2005; Katambara & Ndiritu, 2009; Kagoda *et al.*, 2010; Van Vliet *et al.*, 2012; Onyari & Ilunga, 2013; Oyebode *et al.*, 2015); **reservoir capacity** (Adeloye & De Munari, 2006; Adeloye, 2009); Adeloye 2009); **rainfall data** (Hughes *et al.*, 2006; Nkuna & Odiyo, 2011); **river-runoff** (Steynor *et al.*, 2009); **water demand** (Msiza *et al.*, 2007) and **water temperature** (Van Vliet *et al.*, 2012). Studies relevant to drinking water only include the prediction of chemical dosing which was performed by Naidoo & van der Walt (2013). In their study, a feed-forward network was used to determine the chemical dosing in order to improve budgeting, accuracy and reliability of the distribution plant (Table 3.1). This was a case study that was done by the water company, Rand Water, in South Africa. Results indicated that the ANN was able to correctly predict the chemical dosing levels for lime, polymer and chlorine, even during periods where raw water quality spikes in turbidity, pH, alkalinity and colour levels were experienced. Limited ANN studies highlight the research gap regarding the application of ANNs in South African water purification facilities.

3.5 Conclusion

The limitations of end product testing are becoming more evident in the water sector around the world (Okeyo *et al.* 2011). Where deterioration in available raw water quality takes place, it is often difficult to identify which step in the water purification step is not working up to standard (Okeyo *et al.* 2011). The cost of advanced treatment may be unaffordable to some water purification facilities (Brookes *et al.*, 2014). Modelling and future projections, on the other hand, may not only help to improve water quality, but may also help to determine which other treatment options will be worth the investment (Brookes *et al.* 2014). Modelling techniques are increasingly playing important roles when it comes to water management decisions (Scholten *et al.*, 2007; Salami Shahid & Ehteshami, 2016).

This review indicated that ANNs are efficient forecasting tools in the water sector. From the literature, it was evident that the most popular neural network was MLP. However, it was also observed that ANNs were mainly used as prediction tools or studies were performed in order to compare or improve modelling techniques. None of the studies developed these models to be a decision-making tool, except Al-Barqawi & Zayed (2008) with the development of a web-based condition rating tool and Griffiths & Andrews (2011) with the development of a software package to monitor filtration conditions. It was also evident that the applications of ANNs in the water sector of South Africa are limited. With the current drought as well as pollution, the quality of environmental water in South Africa is deteriorating. It may thus be to the advantage of drinking water production facilities to use statistical approaches to ensure that safe drinking water of good potable quality is produced. In addition, based on international examples, there are opportunities for employing ANNs as a tool in decision-making.

CHAPTER 4

Application of Artificial Neural Networks and Evolutionary Algorithms at drinking water purification plants

4.1 Introduction

Although water purification companies strive to supply safe drinking water to consumers, they are often faced with various financial and operational challenges (Meng *et al.*, 2015). Even though plant operators are trained to deal with basic operational problems, some operators may lack in knowledge of basic treatment processes (Swartz, 2000; Momba & Brouckaert, 2005). Moreover, challenges are exacerbated when experienced operators become too complacent, especially when the plant has not experienced any failures for long periods (Hrudey & Hrudey, 2014). This may lead to the operator missing warning signals of possible system failures. Should there be a failure in the treatment process, there could be serious economic consequences for water purification plants (Hrudey & Hrudey, 2014).

To overcome basic treatment challenges and improve water quality, water purification companies may implement advanced water treatment methods or employ automated monitoring systems (Iglesias *et al.*, 2014). However, these advanced technologies are expensive, and some water purification companies may not have the financial resources to implement such costly solutions (Meng *et al*., 2015). Even though there are various methods available that can be used as decision-making tools to monitor water quality, there is a need for an inexpensive and rapid options. The combination of artificial neural networks (ANNs) and evolutionary algorithms (EAs) may be the solution for this problem.

ANNs are computational techniques that have mechanisms that are able to learn and adjust certain patterns in a dataset (Sarkar & Pandey, 2015). Predictions can be made using the current dataset by training the ANN through a learning algorithm (Seth, 2015). However, even though ANNs are very effective in creating forecasting models, they are limited in providing an equation or rule-set to provide an understanding of how the results were obtained (Bezuidenhout *et al.*, 2013). Evolutionary algorithms (EAs) consist of algorithms that rely on irregular operators to simulate recombination and mutation to create new solutions out of a population of alternative solutions (Nicklow *et al.*, 2009). These solutions then compete for the survival of selection

process (Nicklow *et al.*, 2009). EAs find a solution by imitating the processes of evolution, natural selection and genetic variation, which are based on the Darwinian evolutionary theory of “survival of the fittest” (Welk *et al.*, 2008; Swanepoel *et al.*, 2016). EAs have been used to determine predictive rule sets by applying genetic programming to optimise the “if-then” rules and genetic algorithms to optimise the parameters of the rule sets (Recknagel *et al.*, 2008; Swanepoel *et al.*, 2016).

In Chapter 3 an extensive review of ANNs was presented. However, the literature revealed that the application of ANNs in the drinking water sector of South Africa has been limited. Therefore, the aim of the study was to apply ANNs and EAs at three water purification plants in the North West Province, ultimately generating a predictive rule set which water purification plants may use as a decision-making tool for water quality monitoring. Various parameters can be measured during operational monitoring. However, the World Health Organization (2017a) suggests that parameters for operational monitoring should be easily measurable to provide a timely performance of the system, allowing quick and appropriate response to be taken by operators should it be necessary. Some of the suggested source water parameters include turbidity, pH, electrical conductivity and algal growth, whereas treatment parameters include turbidity, pH, colour and ultraviolet intensity (World Health Organization, 2017a).

In the current study, turbidity and electrical conductivity in the final water (after chlorination) were chosen as the parameters to be predicted. Turbidity is a measurement of organic and inorganic matter in water (World Health Organization, 2017a). Monitoring turbidity levels is important, because high turbidity levels may influence the efficiency of the disinfection process (World Health Organization, 2017a). Turbidity can easily be measured by plant operators using low-cost, portable, automated instruments (World Health Organization, 2017b). Electrical conductivity (EC) measures the capability of water to pass an electrical current caused by the inorganic chemicals and dissolved salts present in the water (US Environmental Protection Agency, 2016). Water bodies usually have a relatively constant conductivity. Once the conductivity range has been established, a sudden increase in EC could be an indication of pollution or an unwanted discharge into the water body (US Environmental Protection Agency, 2016). Electrical conductivity can also be easily measured using portable automated instruments.

4.2 Materials and methods

Historical data from 2009 to 2015 of three water purification plants (Plant A, Plant B and Plant C) in the North West Province were obtained. For the ANN analysis, Forecaster XL (Alyuda Research LLC) was used. Input parameters for the neural network were selected by first identifying source water parameters likely to influence the turbidity and electrical conductivity levels in the drinking water. These parameters included various heavy metals, chemical determinands (ammonia, chloride, fluoride, nitrate, nitrite, sulphate, TOC, chlorophyll, sodium, phosphate, magnesium), microbiological determinands (*E. coli*, total coliforms and HPC), turbidity, total dissolved solids, EC, hardness and pH. These identified parameters were then reviewed and parameters containing large gaps in the dataset were excluded. To determine which of the final selected source water parameters had the most influence on turbidity and electrical conductivity in the drinking water, multivariate analysis, specifically redundancy analysis (RDA), was performed using Canoco for Windows 4.5 (Ter Braak & Smilauer, 1998). These parameters were then used as the initial input parameters during the ANN modelling process (Table 4.1). The datasets for Plant A and Plant C used during the ANN modelling process are given in Appendix A.

The dataset for Plant B was extremely small containing a total of only 11 data points from the historical data obtained. Forecaster XL recommends that the training set should have at least 10 times as many records as input columns. This was not the case for Plant B as Forecaster XL warned that the number of data in the dataset was critically low, which could have caused inaccurate forecasting models. Hence, forecasting for Plant B could not be performed. Various combinations of input parameters were used during the ANN modelling processes to improve the accuracy of the model. Rule sets for Plant A and Plant C were generated using Multimodelling software, which is a hybrid evolutionary algorithm for ecoinformatics (Bezuidenhout *et al.*, 2013). Multiple rule sets were generated and evaluated (n>10). Rule sets were discarded when the THEN or ELSE statement were always valid, i.e. the IF statement didn't allow for alternate decision branching. Other rule sets were also discarded when the IF statement was judged to be non-sensical, i.e. selected for a value of the parameters that would never be found in the actual data. The R-squared value as well as how closely the fit matched events in the data served as additional selection for the suitability of the generated rulesets.

4.3 Results

4.3.1 Redundancy Analysis

Figure 4.1 and Figure 4.2 indicate the RDA ordination diagrams for Plant A and Plant C respectively. Angles between the vectors are an indication of the degree of correlation between the variables. If the arrows point in a similar direction, they are predicted to be positively correlated with each other (Lepš & Šmilauer, 2003). Hence, the narrower the angle between two variables, the larger the positive correlation between them.

As depicted in Figure 4.1, total coliforms, *E. coli*, turbidity, chloride, sulphate, pH and EC had a close correlation with EC in the drinking water of Plant A. Parameters closely correlated with turbidity in the drinking water were total coliforms, *E. coli*, iron, copper and turbidity. For Plant C (Figure 4.2), sulphate, sodium, chloride, pH, EC, total organic carbon, total chlorophyll, phosphate and total coliforms were closely correlated with EC in the drinking water. Total organic carbon, pH, chloride, sodium, EC and sulphate were closely correlated with turbidity in the drinking water of Plant C.

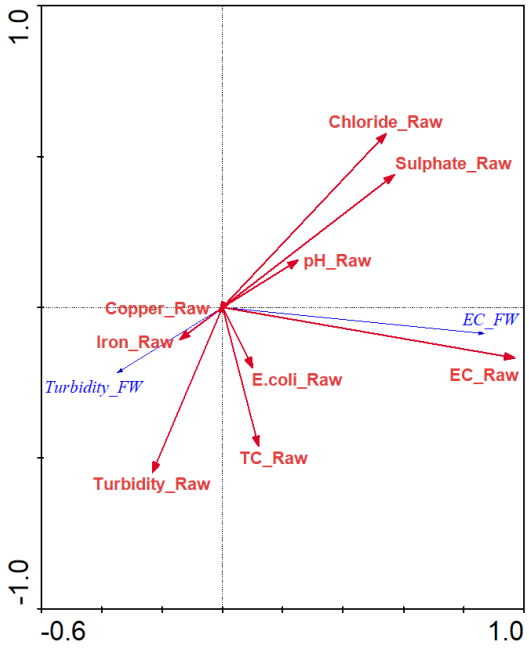


Figure 4.1: RDA ordination diagram illustrating the correlation between the source water parameters (red vectors) and drinking water parameters (blue vectors) for Plant A. TC: total coliforms; FW: final water.

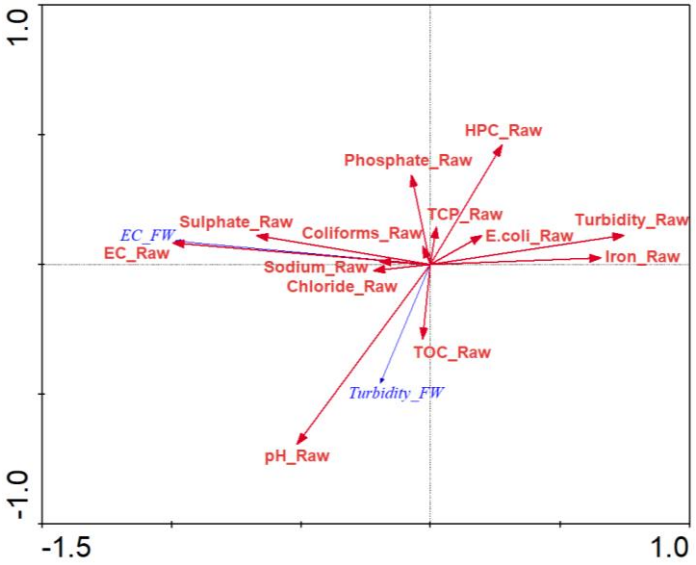


Figure 4.2: RDA ordination diagram illustrating the correlation between the source water parameters (red vectors) and drinking water parameters (blue vectors) for Plant C. TOC: Total organic carbon; TCP: Total chlorophyll; HPC: Heterotrophic plate count bacteria; FW: final water.

4.3.2 Artificial Neural Networks

Table 4.1 indicates the final selected input parameters for each of the output parameters for Plant A and Plant C. These input parameters produced ANN models with the highest percentage of good forecasts. For Plant A, Figure 4.3 (A) and Figure 4.4 (A) indicate the comparison between the actual and forecasted data of the ANN models for EC and turbidity respectively. The network structure (input: hidden: output layers) generated for the EC model for Plant A was 7: 13: 1 with 100% good forecasts for both the training and test sets. The most influential input parameter was EC (Figure 4.3 - B). The network structure generated for the turbidity model of Plant A was 5: 27: 1 with 32% and 21% good forecasts for the training and test sets respectively. Iron was the most influential input parameter (Figure 4.4 (B)).

For Plant C, Figure 4.5 (A) and Figure 4.6 (A) indicate the comparison between the actual and forecasted data of the ANN models for EC and turbidity respectively. The network structure generated for the EC model was 6: 7: 1 with 98% and 100% good forecasts for the training and test set respectively. EC was the most influential input parameter (Figure 4.5 (B)). The turbidity model for Plant C generated a network structure of 3: 16: 1 with 56% and 44% good forecasts for the training and test set respectively. The most influential parameter for this model was pH (Figure 4.6 (B)).

Table 4.1: Input parameters used in the ANN modelling of EC and turbidity in the drinking water of Plant A and C.

	INPUT PARAMETERS (SOURCE WATER)	**OUTPUT PARAMETERS**
PLANT A	Total coliforms, *E. coli*, pH, EC, chloride, sulphate, turbidity	EC in drinking water
	Total coliforms, *E. coli*, iron, copper, turbidity	Turbidity in drinking water
PLANT C	Total organic carbon, pH, EC, sodium, chloride, sulphate	EC in drinking water
	pH, EC, sodium	Turbidity in drinking water

E. coli: *Escherichia coli*

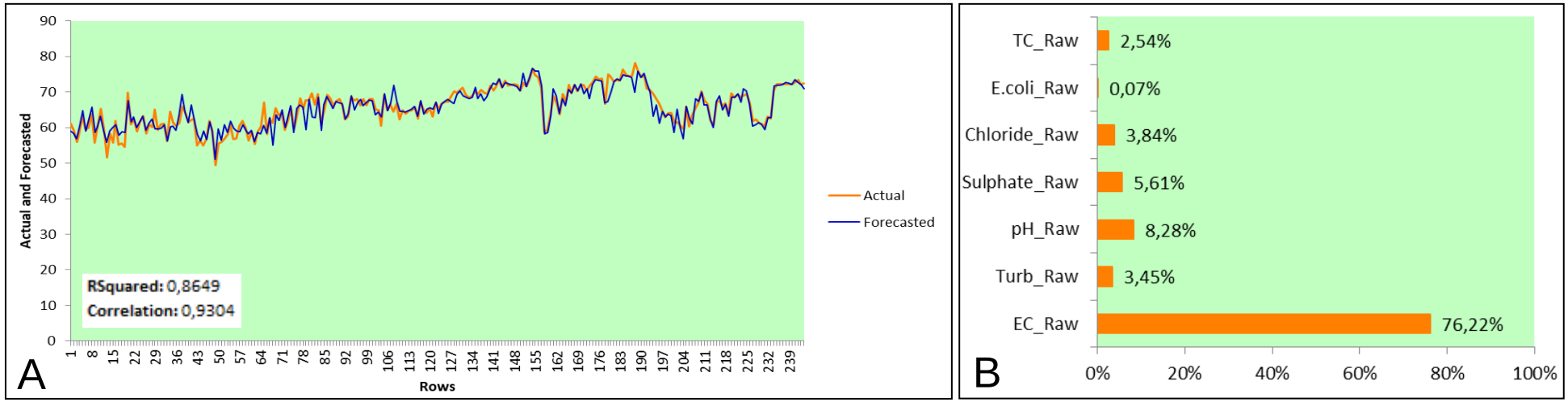


Figure 4.3: (A): Graph indicating the actual and forecasted data for EC in the drinking water of Plant A. (B): Percentage contribution of the various input parameters.

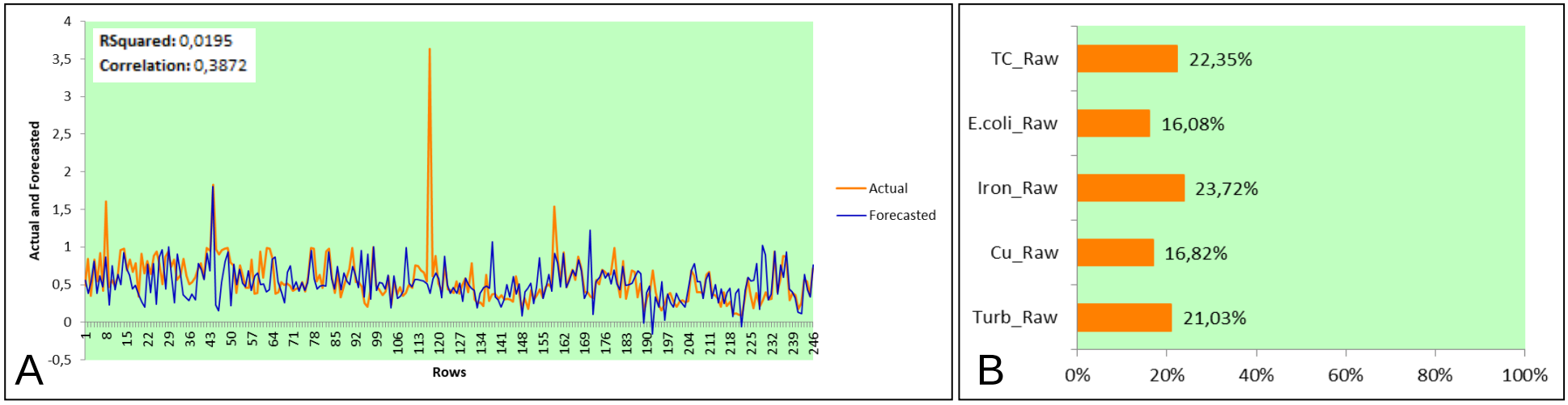


Figure 4.4: (A): Graph indicating the actual and forecasted data for turbidity in the drinking water of Plant A. (B): Percentage contribution of the various input parameters.

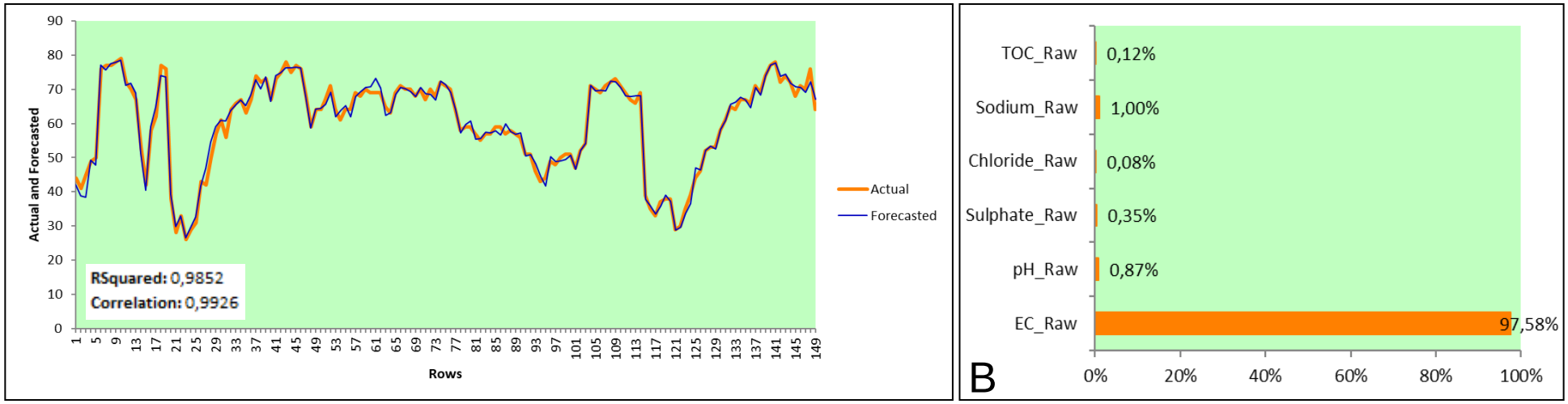


Figure 4.5: (A): Graph indicating the actual and forecasted data for EC in the drinking water of Plant C. (B): Percentage contribution of the various input parameters.

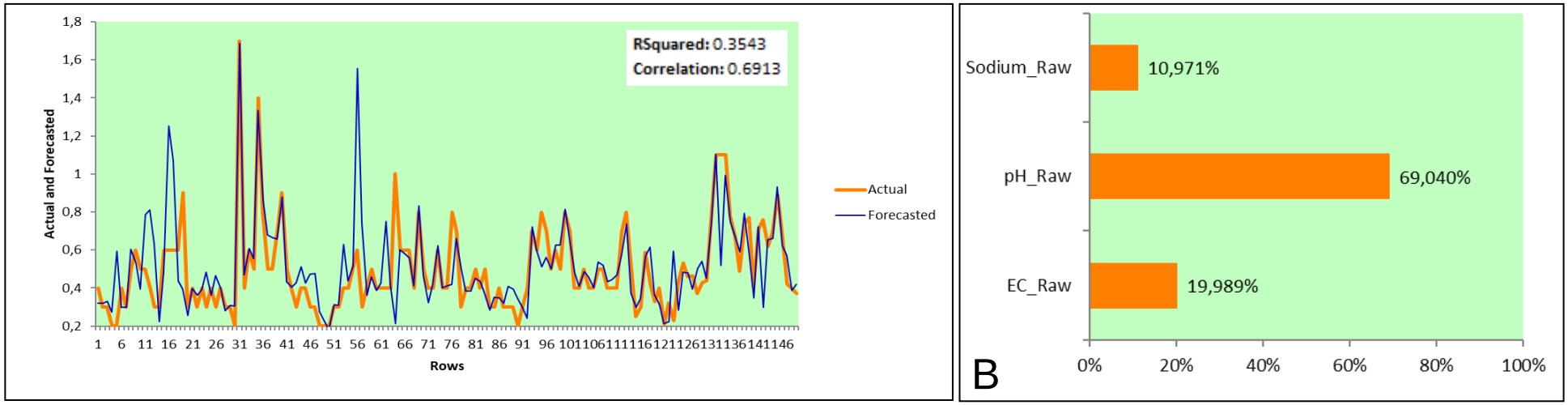


Figure 4.6: (A): Graph indicating the actual and forecasted data for turbidity in the drinking water of Plant C. (B): Percentage contribution of the various input parameters.

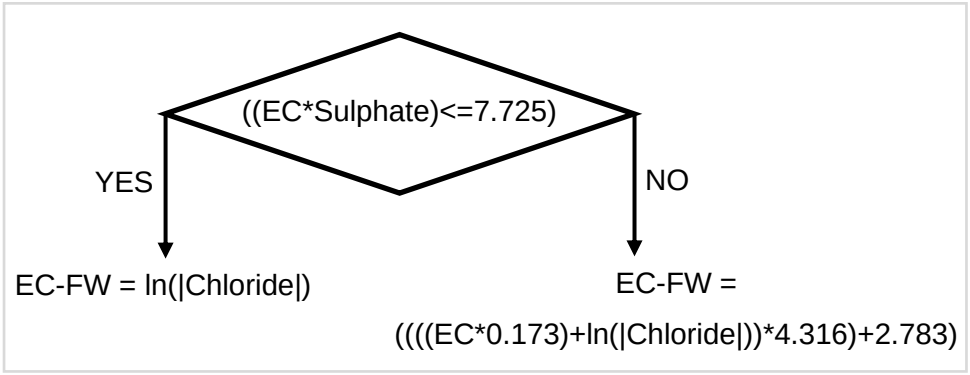
4.3.3 Evolutionary Algorithms

Due to the high accuracy of the ANN models for EC, IF-THEN-ELSE rule sets for Plant A and C for this parameter were generated successfully. However, the accuracy of the ANN models for turbidity for both plants were too low for rule sets to be developed. Two rule sets each for Plant A and C for EC were generated. Parameters used in the rule sets are source water parameters used to predict drinking water EC (EC-FW). A summary of the results obtained from these rule sets are illustrated in Figures 4.7 to 4.10.

Plant A: Rule set 1:

Rule set 1 for Plant A identified the IF criterion of this prediction model is determined by the combination of source water parameters EC and Sulphate (Figure 4.7 - A). Should the IF rule be true, the THEN rule applies, which only includes chloride for the EC-FW prediction model. Should the IF rule be false, the ELSE rule applies, which include EC and Chloride for the EC-FW prediction model. Rule set 1 showed an R^2 value of 0.751 which corresponds to a 75,10% accuracy (Figure 4.7 (B).

(A)



(B)

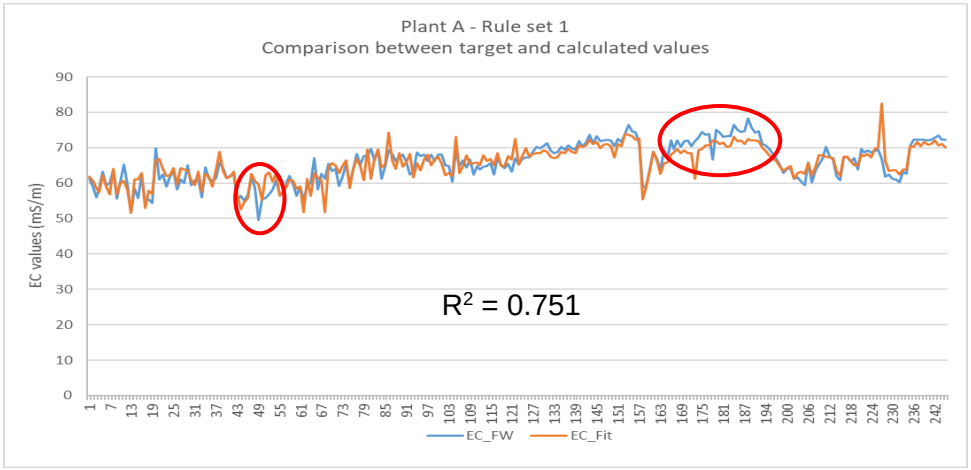


Figure 4.7 - (A): IF-THEN-ELSE model. (B): Line graph indicating the comparison between the target and calculated values. EC_FW represents the target values for the evolutionary algorithm. EC_Fit represents the calculated values from the rule set. The red circles indicate the over- and underestimations by the model.

Plant A: Rule set 2:

Rule set 2 for Plant A identified the IF criterion of this prediction model is determined by the source water parameter pH (Figure 4.8 - A). Should the IF rule be true, the THEN rule applies, which indicates EC-FW = 3.119. Should the IF rule be false, the ELSE rule applies which include EC for the EC-FW prediction model. Testing the rule set showed an R^2 value of 0.743 which corresponds to a 74,30% accuracy (Figure 4.8 (B)).

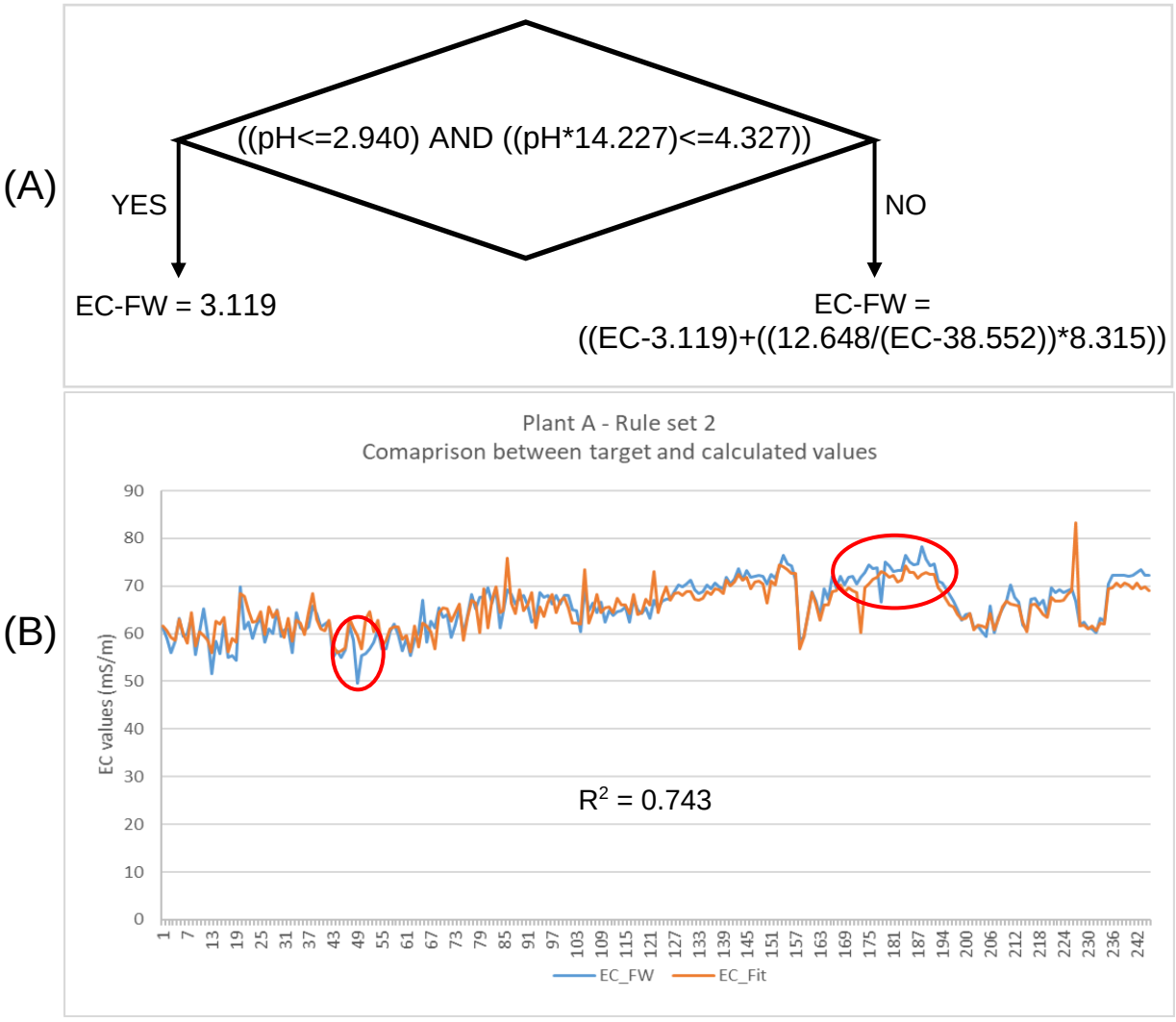


Figure 4.8 - (A): IF-THEN-ELSE model. (B): Line graph indicating the comparison between the target and calculated values. EC_FW represents the target values for the evolutionary algorithm. EC_Fit represents the calculated values from the rule set. The red circles indicate the over- and underestimations by the model.

Figure 4.7 (B) and Figure 4.8 (B) illustrate the comparison between the target values (EC_FW) and the calculated values (EC_Fit) from rule set 1 and rule set 2 respectively. As illustrated in Figure 4.7 (B) and Figure 4.8 (B), the prediction models for both rule sets were very similar with less than 1% difference between the R^2 values. Over- and underestimations were obtained at the similar data points. For this reason, results for rule set 1 will be explained as very similar results were obtained for rule set 2.

As illustrated in Figure 4.7 (B), at data point 49 there was an overestimation by the prediction model. The target value of EC at data point 49 was 49 mS/m, whereas the prediction value was calculated at 59 mS/m. At data points 166 to 191 there was an underestimation by the prediction model. The target value of EC at data point 166 was 72 mS/m, whereas the prediction value was calculated at 68 mS/m. At data point 191 the target value of EC was 74 mS/m, whereas the prediction value was calculated at 72 mS/m. Similar predictions were made for points in between data points 166 and 191. Even though occasional over- and underestimations were illustrated, for the most part, the predicted values of the rule set followed the same trend as the target values with high accuracy (Figure 4.7 (B) & Figure 4.8 (B)).

Plant C: Rule set 1:

Rule set 1 for Plant C identified the IF criterion of this prediction model is determined by the source water parameter Chloride (Figure 4.9 - A). Should the IF rule be true, the THEN rule applies, which only includes EC for the EC-FW prediction model. Should the IF rule be false, the ELSE rule applies, which includes TOC for the EC-FW prediction model. Testing the rule set showed an R^2 value of 0.983 which corresponds to an accuracy of 98.3% (Figure 4.9 (B)).

Plant C: Rule set 2:

Rule set 2 for Plant C identified the IF criterion of this prediction model is determined by the source water parameter TOC (Figure 4.10 - A). Should the IF rule be true, the THEN rule applies, which only includes EC for the EC-FW prediction model. Should the IF rule be false, the ELSE rule applies, which also includes EC for the EC-FW prediction model. Testing the rule set showed an R^2 value of 0.980, corresponding to an accuracy of 98.0% (Figure 4.10 (B)).

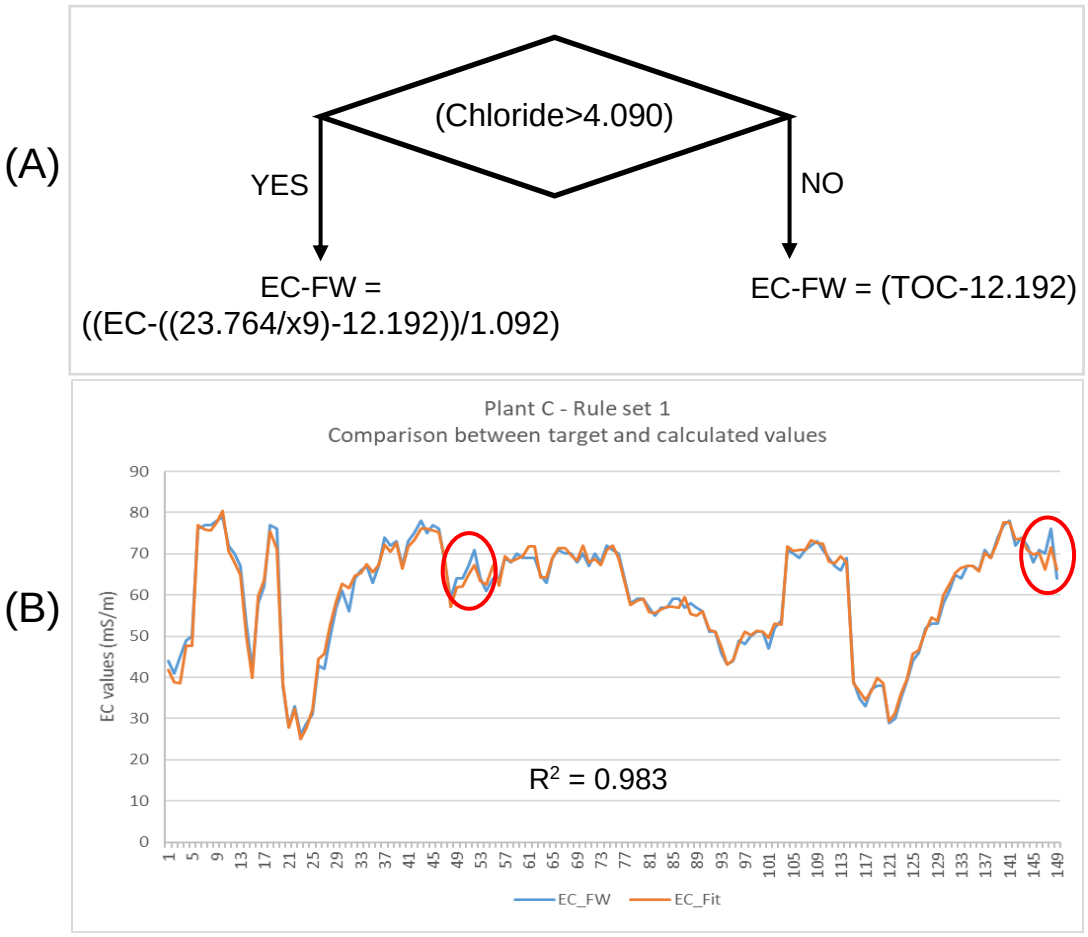


Figure 4.9 - (A): IF-THEN-ELSE model. (B): Line graph indicating the comparison between the target and calculated values. EC_FW represents the target values for the evolutionary algorithm. EC_Fit represents the calculated values from the rule set. The red circles indicate the underestimations by the model.

Figure 4.9 (B) and Figure 4.10 (B) illustrate the comparison between the target values (EC_FW) and the calculated values (EC_Fit) from rule set 1 and rule set 2 respectively. As illustrated in Figure 4.9 (B) and Figure 4.10 (B), the prediction models for both rule sets were very similar with only a 0.3% difference between the R^2 values. Underestimations were obtained at similar data points. For this reason, results for rule set 1 will be explained as very similar results were obtained for rule set 2.

As illustrated in Figure 4.9 (B), at data points 52 and 148 there were underestimations by the prediction model. The target value of EC at data point 52 was 71 mS/m, whereas the prediction value was calculated at 67 mS/m. The target value of EC at data point 148 was 76 mS/m, whereas the prediction value was calculated at 72 mS/m. Even though these slight underestimations were illustrated, for the most part, the

predicted values of the rule set followed the same trend as the target values with high accuracy (Figure 4.7 (B) & Figure 4.8 (B)).

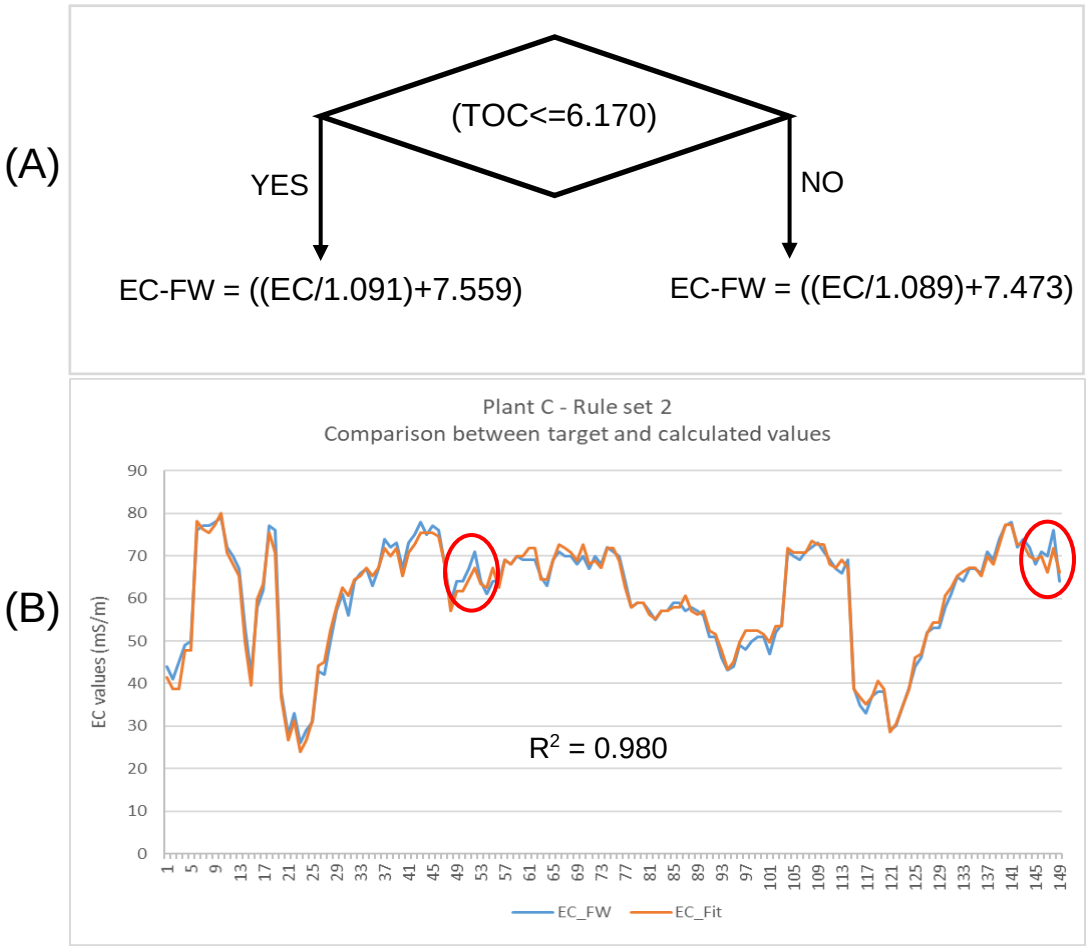


Figure 4.10 - (A): IF-THEN-ELSE model. (B): Line graph indicating the comparison between the target and calculated values. EC_FW represents the target values for the evolutionary algorithm. EC_Fit represents the calculated values from the rule set. The red circles indicate the underestimations by the model.

4.4 Discussion

4.4.1 Redundancy analysis

The use of redundancy analysis (RDA) as a screening step for the selection of input parameters for the ANN model, gave a quick and accurate insight as to which of the source water variables correlated with EC and turbidity in the drinking water. Correlations with EC (Figures 4.1 & 4.2) proved to be valid as the use of these variables as input parameters for the ANN models resulted in highly accurate forecasting of EC in the drinking water (Figures 4.3 - A & 4.5 - A). A study by Li *et al.* (2018) also combined RDA and ANN to predict nutrient removal in tidal flow constructed wetlands. In their study results of the RDA indicated that nutrient removal

was influenced by different factors. Information obtained from the RDA assisted the authors in selecting the main factors for the ANN model simulations. Jang *et al.* (2018) used Principal Component Analysis (PCA) and ANN to estimate the leakage ratio in water distribution systems. The results of their study indicated that the combination of PCA and ANN gave more accurate results than the single ANN simulation. Therefore, the use of linear ordination techniques (RDA or PCA) as a pre-step during the selection of input parameters for ANN modelling, may save time and effort and improve the accuracy of the ANN.

4.4.2 Artificial Neural Networks

The aim of the current study was to apply ANNs and EAs to three water purification plants in the North West Province, South Africa. However, due to insufficient data from Plant B, modelling of EC or turbidity could not be performed for this plant. The World Health Organization (2017a) stated that proper monitoring is an important requirement to ensure the production of safe drinking water. This is also an essential component of the South African Blue Drop requirements comprising 30% of the 35% in the water safety section (Department of Water and Sanitation, 2014). Inadequate monitoring of water quality at Plant B consequently led to the inability to use ANNs and EAs as a decision-making tool, which could have improved their water quality management. This finding emphasises the importance of continuous water quality monitoring as it provides valuable information necessary to make thorough and informative water management decisions with regards to current and emerging problems (Myers, 2019).

However, ANN models were developed for Plant A and Plant C for the prediction of EC and turbidity in the drinking water. The EC model for Plant A yielded 100% good forecasts for the training and the test sets. The R^2 value for this model was 0.8649 (Figure 4.3 (A)), corresponding to an 86.49% accuracy. The main influential parameters for predicting EC in the drinking water was EC (76.22%), pH (8.28%), sulphate (5.61%) and chloride (3.84%) (Figure 4.3 (B)). The EC model for Plant C yielded 98% and 100% good forecasts for the training and test sets respectively. The R^2 value for this model was 0.9852 (Figure 4.5 (A)), corresponding to 98.52% accuracy. EC was the most influential parameter with a contribution of 97.58%, followed by sodium (1.00%) and pH (0.87%). Correlation coefficients of 0.9304 for Plant A and 0.9926 for Plant C indicated a very strong relation between the actual and the forecasted data sets.

The ANN models for the prediction of turbidity for both Plant A and C were not as successful as the EC models. The turbidity model for Plant A yielded 32% and 21% good forecasts for the training and test sets respectively, with a correlation coefficient of 0.3872 (Figure 4.4 (A)). The R^2 value for this model was 0.0195, thus corresponding to an extremely low accuracy of 1.95%. The most influential parameters for this model were iron (23.72%), total coliforms (22.35%) and turbidity (21.03%). The turbidity model for Plant C performed slightly better yielding 56% and 44% good forecasts for the training test and test sets respectively. There were three main parameters contributing to this model of which pH had the highest contribution (69.04%), followed by EC (19.98%) ad sodium (10.97%). The correlation coefficient for this model was 0.6913 and the R^2 value was 0.3543 (Figure 4.6 (A)), which corresponded to a low accuracy of 35.43%.

Similar results were obtained in a study by Tabari & Hosseinzadeh Talaee (2015), who used ANNs for the prediction of missing river water quality data using existing data. Radial basis function (RBF) and multilayer perceptron (MLP) networks were used to predict missing data of 13 water quality parameters, including turbidity. Results of their study indicated that even though the ANN networks were able to predict the water quality parameters successfully, the worst performance by both networks was for turbidity. The authors argued that the cause of the low accuracy of turbidity prediction could have been due to the influence of additional factors, such as runoff, inflow and algae. The authors recommended that MLP and RBF should not be used for the prediction of turbidity data.

Kazemi *et al.* (2018) compared the use of nonlinear autoregressive exogenous (NARX) and feed-forward neural networks for the prediction of turbidity in water distribution trunk mains. NARX is an ANN-based model and is suitable for time-series multi-step forecasting. Results of the study indicated that the feed-forward neural network was not as effective in predicting the turbidity as the NARX model, which outperformed the feed-forward network with a R^2 value of 0.97 versus a R^2 value of 0.66 for the feed-forward network.

The influential parameters contributing to the EC models for Plant A and Plant C included EC, pH, sulphate, chloride and sodium. Parameters which contributed to the prediction of the turbidity models for Plant A and Plant C included iron, total coliforms,

pH, turbidity, EC and sodium. The importance of these parameters in water are as follows:

- **Electrical conductivity**

  Inorganic chemicals and dissolved salts have the capability to cause an electrical current in the water (US Environmental Protection Agency, 2016). Electrical conductivity is a measurement of this electrical current (US Environmental Protection Agency, 2016). An increase in EC levels may indicate an occurrence of pollution or system failure (US Environmental Protection Agency, 2016).

- **pH**

  pH is a measurement of the acidity or alkalinity of a water source (Oram, 2014). The normal range for surface water is usually between 6.5 and 8.5 and for drinking water 7.0 (Oram, 2014; World Health Organization, 2017a). Water with a pH < 6.5 may contain metals such as iron, copper, lead and zinc and high pH levels (> 8.5) may be an indication that the water is hard which may cause aesthetic problems (Oram, 2014). pH is one of the most important operational parameters for water purification plants, as various treatment processes are influenced by the pH level (World Health Organization, 2017a).

- **Sulphate**

  Sulphates naturally occur in various minerals and their presence in water may be attributed to industrial wastewater discharges (World Health Organization, 2017a). The presence of sulphate in drinking water may cause taste and odour problems (World Health Organization, 2017a).

- **Chloride & Sodium**

  Even though there is no health concern for sodium and chloride in drinking water, high concentrations of these parameters in drinking water may give a salty taste to the water (World Health Organization, 2017a).

- **Iron**

  Iron in drinking water may be caused by the presence of ferrous iron in anaerobic groundwater (World Health Organization, 2017a). Iron encourages the growth of “iron bacteria”. Although there is no health-based guideline for iron, levels above

0.3 mg/L may stain laundry and develop a turbid colour in water (World Health Organization, 2017a).

- **Turbidity**

  Turbidity describes the cloudiness of water due the presence of suspended solids, chemical and organic particles, and microorganisms (World Health Organization, 2017a). Poor source water quality, ineffective treatment of drinking water and biofilms may contribute to high levels of turbidity (World Health Organization, 2017a). Therefore, turbidity can be used to indicate the quality of a water source (World Health Organization, 2017a).

Results from the current study indicated that the forecasting models for EC for Plant A and Plant C were very accurate. However, the computation steps behind these models are not known. This is due to ANNs not having the ability to document and explain the conclusions they reach (Christodoulou & Deligianni, 2010). This problem was solved by applying EAs. The disadvantage of EAs is that they sometimes need numerous evaluations to reach ideal solutions (Magnier & Haghighat, 2010), which may be time consuming. Therefore, the combination of ANNs and EAs was applied. ANNs were used as a preliminary step in the selection process due their ability to rapidly evaluate various dataset combinations. Successful combinations derived from the ANNs were further refined by the EA step, until the most applicable rule sets were obtained.

4.4.3 Evolutionary algorithms - predictive rule sets generated

Water purification plants are often faced with unexpected loss of equipment or financial difficulties that restrict them from performing optimum monitoring of all the necessary parameters. Therefore, two rule sets each for Plant A and Plant C were selected which will give the water purification plant more than one option for predicting EC.

4.4.3.1 Rule sets developed for Plant A

The main role players for rule set 1 of Plant A were Sulphate, Chloride and EC. Rule set 2 was more basic and contained only two role players, namely EC and pH. Comparing the target and calculated values resulted in R^2 values of 0.751 for rule set 1 and 0.743 for rule set 2 (Figure 4.7 (B) and Figure 4.8 (B)). This corresponds to a respective accuracy of 75.10% and 74.30% for the rule sets. Even though rule set 2

produced slightly less accurate predictions, it offered the water purification plant the option to use easy measurable and low-cost parameters to predict EC in the drinking water.

4.4.3.2 Rule sets developed for Plant C

Rule set 1 for Plant C consisted of three main role players, namely Chloride, EC and TOC, whereas EC and TOC were the two main role players in rule set 2. Comparing the target and calculated values, both rule sets produced highly accurate EC predictions (Figures 4.9 (B) and Figure 4.10 (B)). The R^2 values were 0.983 for rule set 1 and 0.980 for rule set 2. This corresponds to 98% accuracy for both the rule sets.

The results from the rule sets generated for Plant A and Plant C, indicated that EAs had been successfully applied to predict EC in drinking water. Predictions for Plant C had a higher accuracy than Plant A. Evolutionary algorithms have been applied in various water quality sectors, such as river water quality predictions (Liu *et al.*, 2013; Burchard-Levine *et al.*, 2014; Azad *et al.*, 2018), and water treatment plant design (Magnier & Haghighat, 2010; Abkenar *et al.*, 2015; Bi *et al.*, 2015; McClymont *et al.*, 2015). However, the application of EAs for the prediction of drinking water quality parameters has been limited to the prediction of chlorine concentration (Kurek & Brdys, 2007; Kim *et al.*, 2014), prediction of disinfection by-products (Singh & Gupta, 2012) and the optimisation of reverse osmosis (Murthy & Vengal, 2006).

A review of recent literature has indicated that the application of ANNs and EAs for the prediction of drinking water quality parameters (such as EC and turbidity) using source water parameters as input variables has not been documented (O'Reilly *et al.*, 2018). The only study relating to this type of application was a study performed recently by Zhang *et al.* (2019). In their study a hybrid ANN and genetic algorithm were used to predict the average water production of a drinking water treatment plant in China. Eleven raw water parameters were selected as input variables. Results indicated that combining ANN with genetic algorithm increased the R^2 value from 0.71, obtained from the individual ANN model, to 0.93. Their study concluded that combining ANN and EAs could be used as a decision-making tool, thereby increasing the resilience of drinking water treatment plants.

4.5 Conclusion

The combined application of ANNs and EAs for the prediction of drinking water quality parameters was successful in this study. Plant A and Plant C differed from each other in many respects. For example, they have different sources of raw water, different treatment processes are applied, different monitoring parameters are used, and the overall management procedures also differ. Therefore, the rule sets developed for Plant A and Plant C are different and purification plant specific. The advantage of using EAs is that the rule sets generated can be applied by Plant A and Plant C without requiring complicated computer software. Programs such as Microsoft Excel, which is easily accessible, can be used to insert the algorithms after which the relevant data can be captured and analysed by plant operators.

The turbidity predictions for Plant A and C were not as successful as the EC predictions. However, EC is an important water quality parameter, not only because it indicates changes in water quality, but also because it plays an important role in water resources management and health studies (Al-Mukhtar & Al-Yaseen, 2019). Combining ANNs and EAs resulted in 98% accurate EC predictions for both Plant A and Plant C. Comparing predicted EC values with current measured values in the final drinking water may serve as a warning system for possible process failures. The rule sets may also be implemented to indicate the maximum allowable levels for certain parameters in the source water, which may enable plant operators to plan ahead and be prepared to make any adjustments to the treatment process, should these levels become a reality. The use of hybrid modelling will thus not only assist in understanding the current system dynamic, but will also give future scenarios which will aid in management decisions (Parrott, 2011).

Results from this study indicated that ANNs and EAs were successfully applied to predict certain water quality parameters in drinking water of South Africa. Hence, the use of ANNs and EAs as decision-making tools by water purification plant operators, may improve the quality and production of drinking water in this water-scarce country.

4.6 Recommendations

- More data and work in this field is required as studies in the drinking water sector are limited.

- Investigate the application of other ANN models, such as RBF, GRNN and CFN to generate prediction models for drinking water parameters.

CHAPTER 5

Recovery of E*scherichia coli* in drinking water

5.1 Introduction

Drinking water distribution systems are not sterile and systems contain unique microbiomes (Pinto *et al.*, 2012). However, these systems should be free from pathogens originating from human and animal faecal origin (World Health Organization, 2017a). *Escherichia coli* is used as a surrogate for monitoring the potential of these systems containing any pathogens and a world-wide standard is 0 CFU/100 ml for any potable water (World Health Organization, 2017a).

Microorganisms in raw water, used for drinking water production as well as during the drinking water purification process may be exposed to various harmful environmental stresses which may cause some cells to be lethally injured, whilst other organisms may survive (Pienaar *et al.*, 2016). Continued exposure to these environmental stresses may ultimately lead to death of the injured and uninjured cells. The injured cells may not be able to grow on selective culture media, but may still pose a health risk (Pienaar *et al.*, 2016). If the injured cells are removed from the stressed environment and placed into an environment with favourable conditions, they may recover and be culturable again (Pienaar *et al.*, 2016; Chen *et al.*, 2018). There have been various suggestions to describe the state microorganisms enter once they are exposed to environmental stresses. Pinto *et al.* (2015) described four non-growing states: sporulation, persistence, dormancy and viable-but-non-culturable (VBNC). This VBNC state may thus affect the culturability of the bacterial population in drinking water.

The concept of VBNC microorganisms was first introduced by Xu and colleagues (Xu *et al.*, 1982; Oliver, 2016). In their study, they illustrated that *Escherichia coli* (*E. coli*) and *Vibrio cholerae* were not able to be recovered with standard culture methods, but they were still viable (Xu *et al.*, 1982). This means that under adverse environmental conditions, some microorganisms can stay in a dormant state and may return to a metabolic active state (Lin *et al.*, 2017). The particular environmental stresses that could lead to a VBNC state include high osmotic concentrations, temperature fluctuations, presence of heavy metals and starvation among others (Pinto *et al.*, 2015; Oliver, 2016). A study by Vora *et al.* (2005) reported that human pathogenic *Vibrio*

spp. were able to express toxin and virulence genes while in the VBNC state (Pinto *et al*., 2015).

Roszak and colleagues introduced the concept of resuscitation of VBNC cells (Roszak *et al.*, 1984). The non-culturable state is also defined by cells not able to grow on conventional media, but they have intact cell membranes, decreased respiration rates and they retain low metabolic activity (Oliver, 2016; Pienaar *et al.*, 2016). According to Pinto *et al*. (2015) there are differences between the concepts of resuscitation, revival and recovery of VBNC cells. Resuscitation is the process where cells become culturable again through the reversal of metabolic and physiological processes. Revival refers to injured cells returning to an active state by being released from their stressed environments. Recovery, on the other hand, means the injured cells are transferred back to an active state by the addition of media supplements containing hydrogen peroxide-degrading compounds.

Various studies have illustrated that some disinfectants used during the treatment of water may induce the VBNC state in bacteria (Chen *et al.*, 2018). Some of the disinfection methods include ultraviolet treatment (Zhang *et al.*, 2015), chloramination (Liu *et al.*, 2009), monochloramine (Alleron *et al.*, 2008; Turetgen, 2008; Chen *et al.*, 2018), as well as chlorination (Moreno *et al.*, 2007; Lin *et al.*, 2017; Chen *et al.*, 2018). Chlorination is one of the most common disinfection processes used by water purification plants (Lin *et al*., 2017). The efficiency of chlorination is generally measured by counting the number of colonies enumerated on specific microbiological media (Lin *et al* 2017). This may be a health challenge, because the VBNC cells may not be enumerated on standard microbiological media, but may still pose a health risk.

Besides being an indicator species *E. coli* also consists of both pathogenic and non-pathogenic strains (Pienaar *et al*., 2016). There are various reports demonstrating that *E. coli* could enter into the VBNC state (Zhang *et al.*, 2015; Lin *et al.*, 2017; Chen *et al.*, 2018). A study by Chen *et al.* (2018) indicated that *E. coli* cells treated with chlorine and chloramine were induced into the VBNC state. More importantly, the study also concluded that the VBNC cells were able to express the virulence gene *gapA*. This is concerning, because the *gapA* gene has been identified as a housekeeping gene of the pathogenic *E. coli* O157:H7 strain (Jandu *et al.*, 2009; Rajkhowa *et al.*, 2010). Lin *et al.* (2017) illustrated that levels of 0.5 mg/L chlorine were able to induce *E. coli* into

the VBNC state. Their study also indicated that the VBNC *E. coli* cells showed broad spectrum persistence to nine different antibiotics. A study by Zhang *et al.* (2015) indicated that *E. coli* cells exposed to UV radiation were able to enter into a VBNC state. Their study concluded that the virulence gene *gadA* was expressed by the VBNC cells, which meant that the VBNC cells could still pose a health risk.

The presence of pathogenic *E. coli* in drinking water is thus a major concern for water treatment plants, especially if these cells are in a VBNC state and could be resuscitated again. Various resuscitation methods for *E. coli* in drinking water have been identified. Some of these include the addition of sodium pyruvate and amino acids to media, the presence of supernatants, temperature changes and heat stable autoinducers (Reissbrodt *et al.*, 2002; Bjergbaek & Roslev, 2005; Asakura *et al.*, 2008; Pinto *et al.*, 2011; Pinto *et al.*, 2015; Ding *et al.*, 2017). However, some resuscitation methods can be very complicated involving complex conditions or factors that must be adhered to (Arana *et al.*, 2007). Other rapid and sensitive VBNC detection methods such as flow cytometry, fluorescent microscopy, Polymerase Chain Reaction (PCR) and autoradiography have the disadvantage of needing expensive equipment (Pienaar *et al.*, 2016).

Even though the discovery of VBNC bacteria has highlighted some shortcomings of using culture-based methods, the simplicity and cost-effectiveness of these methods cannot be underestimated (Sardessai, 2005). Resuscitation of VBNC bacteria using culture-based methods has previously been documented (Dufour *et al.*, 1981; LeChevallier *et al.*, 1983; Özkanca *et al.*, 2009; Zhao *et al.*, 2013). However, recent studies in this field have been limited.

The aim of this study was to screen drinking water for the presence of stressed or injured *E. coli* by recovering these cells with culture-based methods. The resuscitated cells were verified for the presence of *E. coli* by using PCR and targeting the 16S rRNA gene and *E. coli* housekeeping genes *mdh*, *lacZ* and *uidA* (Ram & Shanker, 2005; Banu *et al.*, 2010; Igwaran *et al.*, 2018). The 16S rRNA gene is commonly present in bacteria (Janda & Abbott, 2007) and has also been used to demonstrate that the DNA is amplifiable. The *mdh* gene encodes the enzyme malate dehydrogenase (Park *et al*., 1995), *lacZ* encodes the enzyme β-D-galactosidase (Ram & Shanker, 2005) and *uidA* encodes the enzyme β-D-glucuronidase (Janezic *et al.*, 2013).

5.2 Materials and Methods

5.2.1 Sampling

Sampling was conducted at three water purification plants and their distribution systems (Plant A, B and C) during 2015, 2016 and 2017. Water samples were collected at the water purification plants directly after the chlorination step as well as in the distribution at a point of consumption. The water was sampled in sterile glass bottles (Schott Duran, Germany) and transported on ice in cooler boxes to the North-West University Microbiology laboratory. Sampling was done according to the sampling guide by the Department of Water Affairs and Forestry, Department of Health and the Water Research Commission (Department of Water Affairs and Forestry, 2000). The water samples were analysed immediately upon arrival at the laboratory.

5.2.2 Physico-chemical analyses

The pH of the water samples was taken onsite using a PCSTestr 35 multi-meter (Eutech Instruments Pty Ltd, Singapore). The concentration of free chlorine (method 8021) and total chlorine (method 8167) was measured using the Hach Lange DR 2800 spectrophotometer (Güler & Alpaslan, 2009) and turbidity was measured using a Hach portable turbidimeter (model 2100P). The manufacturers' protocols for the various methods were followed.

5.2.3 Recovery of *E. coli* cells from drinking water

For the culture-based method to resuscitate *E. coli* from drinking water, two methods were used. Method A was an existing method as described by Özkanca *et al.* (2009). Method B was a method adapted from Method A.

5.2.3.1 Method A: Method as described by Özkanca *et al*. (2009):

Hundred millilitres of water samples were filtered through 0.45 µm sterile membrane filters (PALL Corporation, SA). This was done in triplicate for each water sample. The membranes were placed on nutrient agar and incubated at 37°C for 8 hours. After incubation on nutrient agar, the membranes were transferred to selective Membrane Lauryl Sulphate Agar (MLSA; Oxoid, UK) and incubated at 44°C for 12 hours. Yellow presumptive *E. coli* colonies on the membranes were counted.

5.2.3.2 Method B: Adapted from Özkanca *et al*. (2009):

Hundred millilitres of water samples were filtered through 0.45 µm sterile membrane filters (PALL Corporation, SA). This was done in triplicate for each water sample. One

set membranes were placed on nutrient agar and one set on nutrient agar-plus respectively. The nutrient agar-plus contained twice the amount of yeast extract and peptone, compared to ordinary nutrient agar. These were incubated at 37°C for 8 hours. After the pre-incubation step, the membranes were transferred to MLSA and m-FC agar (Merck, Germany) respectively and incubated at 44°C for 12 hours. Control plates included membranes incubated directly on m-FC and MLSA media after filtration, without resuscitation. Incubation was at 44°C for 24 hours. Yellow colonies on MLSA medium and blue colonies on m-FC were counted.

The methods were initially tested during 2015 on drinking water from Plant B only. This was due to Plant B having various treatment challenges.

5.2.4 Preparation of cells for DNA extraction

Colonies that were enumerated on the membranes were streaked out onto selective media (MLSA and m-FC) to isolate pure single colonies. However, the colonies did not grow after streaking onto the selective media. Membranes with colonies were placed in 20 ml nutrient broth during the subsequent analysis periods. A vortex was used to dislocate the colonies from the membranes. The membranes were then removed and the nutrient broths were incubated at 37°C for 24 hours.

5.2.5 DNA isolation and identification

A chemagic DNA extraction kit for bacteria was used to extract the DNA from the nutrient broths according to the manufacturer's protocol (Perkin Elmer, USA). The quantity and quality of the DNA was determined by using a NanoDrop One Micro-UV-Vis Spectrophotometer (Thermofisher Scientific, USA). The DNA was amplified using a PCRmax Alpha Cycler 1 thermal cycler (Bibby Scientific, UK). Four genes were targeted in the present study (Table 5.1): 16S rRNA, *mdh*, *lacZ* and *uidA*. All PCR amplicons were confirmed using agarose electrophoresis.

5.2.5.1 PCR conditions

Table 5.1 indicates the primers used in the study. Each reaction tube consisted of 1 x PCR master mix (Thermo Scientific, USA), forward and reverse primers (0.2 µM of each) (Thermo Scientific, USA) and 0.5 mM $MgCl_2$ (Thermo Scientific, USA). Nuclease-free water was added to attain a total volume of 25 µl. Forty to eighty nanogram DNA was used as template in each reaction.

(a) Monoplex PCR - 16S rRNA gene

The PCR cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, 35 cycles of denaturation at 95°C for 30 seconds, annealing at 54°C for 30 seconds and extension at 72°C for 1 minute. Final extension was at 72ºC for 5 minutes.

(b) Monoplex PCR - *uidA* gene

In this case, forward and reverse primers were at 1 µM of each (Inqaba Biotec). The PCR cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, 40 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 1 minute and extension at 72°C for 2 minutes. Final extension was at 72ºC for 10 minutes.

(c) Multiplex & monoplex PCR – *mdh & lacZ* genes

The PCR cycling conditions were as follows: initial denaturation at 95°C for 7 minutes, 30 cycles of denaturation at 95°C for 30 seconds, annealing at 57°C for 30 seconds and extension at 72°C for 1 minute. Final extension was at 72ºC for 7 minutes. For the multiplex PCR, DreamTaq master mix was used instead of 1 x PCR master mix.

5.2.5.2 Sequencing

PCR products obtained by amplification of the 16S rRNA, *uidA* and *mdh* genes were sequenced by Inqaba Biotec. For the identity of the sequences, BLASTN searches were performed on the National Center of Biotechnology Information (NCBI) nucleotide database. 16S rRNA sequences were submitted to the GenBank database under accession numbers MN658429 - MN658454.

5.2.6 Statistical analyses

Statistical analyses were performed with Statistica 13.0 (TIBCO Software Inc. 2017). One-way analysis of variance (ANOVA), followed by Tukey's test were used to determine statistically significant differences ($p < 0.05$) between the sampling periods, source of water and media used.

5.2.7 Phylogenetic analysis

16S rRNA sequences obtained in this study as well as nucleotide sequences previously identified (reference sequences) were aligned with the Clustal W algorithm using MEGA X (Kumar *et al.*, 2018). Phylogenetic analyses were performed by using the Neighbor-Joining method of Tamura and Nei (Tamura & Nei, 1993) using MEGA X (Kumar *et al.,* 2018). Sequences < 850 bp were not included in the analysis.

Table 5.1: Primers used

GENE	PRIMERS	SEQUENCE	FRAGMENT LENGTH (BP)	REFERENCE
16S rRNA	27F	AGAGTTTGATCMTGGCTCAG	1 465	Lane, 1991
	1492R	TACGGYTACCTTGTTACGACTT		
uidA	uidF	CCGATCACCTGTGTCAATGT	380	Bower *et al.*, 2005
	uidR	GTTACCGCCAACGCGCAATA		
mdh	mdhF	GGTATGGATCGTTCCGACCT	301	Tarr *et al.*, 2002; Omar *et al.*, 2010
	mdhR	GGCAGAATGGTAACACCAGAGT		
lacZ	lacZF	CTGGCGTAATAGCGAAGAGG	228	Ram & Shanker, 2005
	lacZR	GGATTGACCGTAATGGGATATG		

5.3 Results

5.3.1 Resuscitation of VBNC *E. coli*

In the present study, using the method as described by Özkanca *et al.* (2009), only 1 CFU/100 ml was enumerated during November 2015 and 0 CFU/100 ml during December 2015 for Plant B (Table 5.3). Using the combination of nutrient agar and nutrient agar-plus with MLSA also gave poor results with only 1 CFU/100 ml enumerated during December 2015 (Table 5.3). This suggests that the method was not as efficient to resuscitate VBNC *E. coli* cells directly from drinking water. However, modifying the method and transferring the membranes from nutrient agar to m-FC agar yielded a higher number of colonies during November and December 2015 for Plant B (Table 5.3). For this reason, VBNC cells were recovered during 2016 and 2017 using only the modified method.

VBNC cells were recovered from drinking water directly after chlorination and at the point of consumption during all the sampling periods for Plant A and B (Table 5.2 & 5.3). No colonies were isolated from Plant C. This may be due to Plant C applying more efficient advanced treatment processes (Figure 1.3) and that the disinfection process is more successful. No colonies were isolated on the control plates, except for very low levels during April 2016 (2nd sampling) at Plant A (Table 5.2) and March 2016 at Plant B (Table 5.3).

As indicated in Table 5.2, Plant A had the highest recovery of VBNC cells in April 2016 during the second sampling period (Too numerous to count - TNTC). These high levels were isolated from drinking water after chlorination. During the same sampling period, 182 CFU/100 ml were enumerated at the point of consumption. Plant B had the highest recovery of VBNC cells during November 2016 (Table 5.3). A total of 177 CFU/100 ml were enumerated at the point of consumption, which is concerning.

Figures 5.1 & 5.2 are interval plots indicating the statistical differences between the sources of sampling, media used and sampling periods for Plant A and B respectively. Statistical analyses indicated that there was no statistical difference between the sources of sampling for Plant A ($p = 0.26435$) and for Plant B ($p = 0.18187$). There was also no statistical difference between using nutrient agar or nutrient agar-plus for Plant A ($p = 0,88912$) and for Plant B ($p = 0,94071$). However, there were statistical differences between the sampling periods for Plant A and Plant B ($p < 0.05$). For Plant A, results from the samples taken during April 2016 (2nd sampling) were statistically different from March 2016, April 2016 (1st sampling) and November 2016. For Plant B, results from the samples taken during November 2016 was statistically different from December 2015, March 2016 and September 2017.

5.3.2 Physico-chemical results

According to SANS 241:2015, pH levels for drinking water should be between 5 & 9.7 pH units, turbidity levels < 1 NTU and free chlorine levels < 5 mg/L. These mentioned physico-chemical parameters were all within the SANS 241:2015 limits (Table 5.2). For Plant B, the pH and free chlorine levels were within the SANS 241:2015 limit. However, the turbidity levels for Plant B were greater than the 1 NTU cut-off for SANS 241:2015 (Table 5.3). This is concerning as turbidity may be caused by the presence of chemical and biological particles, including pathogenic organisms (World Health Organization, 2017a). The free chlorine levels of Plant A and Plant B were below 0.1 mg/L after chlorination and at the point of consumption (Table 5.2 & 5.3). The physico-chemical parameters measured at Plant C were all within the SANS 241:2015 limits.

5.3.3 Detection and sequencing results of the 16S, *uidA, mdh* and *lacZ* genes

Figures 5.3, 5.4 and 5.5 are negative images of agarose gels indicating the PCR results for the 16S rRNA, *uidA* and *mdh* genes for Plant A and Plant B. Only 16S rRNA PCR results for November 2016 were obtained for Plant A (Figure 5.3). The isolates

from Plant B, however, amplified the 16S rRNA, *uidA* and *mdh* genes respectively. Figures 5.4 & 5.5 indicate the agarose gels for the *uidA* and *mdh* genes respectively. Similar results as indicated in Figure 5.3 were obtained for Plant B.

PCR and sequence results of the DNA isolated from the membranes are indicated in Table 5.4, 5.5 and 5.6. Sequence results for Plant A are indicated in Table 5.4. The various colonies on the membranes were identified as Gram-positive *Lysinibacillus fusiformis, Lysinibacillus parviboronicapiens, Paenibacillus* sp. and *Bacillus* sp. No direct identification could be provided for the specific colonies as mixed DNA was isolated. Despite various optimizations, the 16S rRNA PCR's for Plant A for March and April 2016, as well as the *uidA*, *mdh* and *lacZ* PCR's for November 2016 were not successful.

Table 5.5 indicates the PCR and sequencing results of the DNA from Plant B that were enumerated on the control plates (isolated on m-FC agar without resuscitation) during March 2016. The 16S rRNA results identified *Aeromonas veronii*. However, PCR's targeting the *E. coli* housekeeping gene, *uidA,* were successful and sequence results identified *E. coli* as a member of the colonies on the membranes.

Table 5.6 contains the PCR and sequencing results for the resuscitated isolates from Plant B. The 16S rRNA sequencing results for March 2016 identified *Enterobacter* sp. and *Aeromonas veronii* among the colonies. However, the PCR's for the housekeeping genes of *E. coli*, *uidA* and *mdh,* were successful. Sequencing results for these genes confirmed that *E. coli* was a member of the colonies on the membranes. One of the sequences was identified as *E. coli* O177:H21. This was a disconcerting finding as *E. coli* serogroup O177 is a Shiga toxin producing group of *E. coli* (Beutin *et al.*, 2005; Montso *et al.*, 2019) and thus poses a risk to human health.

Analyses of the samples collected in November and September 2016 indicated that the 16S rRNA for Plant B had various species of *Paenibacillus*, *Lysinibacillus* and *Aeromonas.* Amongst the resuscitated species were also *Bacillus cereus* and *Mixta intestinalis* (Table 5.6). Only one of the eight isolates (V31) amplified the *uidA* gene during November 2016, confirming the presence of *E. coli* (VBNC) at the point of consumption. During September 2017, two of the 16S rRNA sequences at the point of consumption (V37 & V38) were identified as *E. coli* (Table 5.6). One of these were identified as *E. coli* O157. This is concerning as *E. coli* O157 is a toxin producing

pathogen and has been associated with various gastrointestinal illness outbreaks (Rangel *et al.*, 2005; Manning *et al.*, 2008; Saxena *et al.*, 2015).

5.3.4 Phylogenetic analysis

Figure 5.6 illustrates the phylogenetic relationships among the VBNC sequences obtained during this study. The Neighbor-Joining tree produced two distinct clusters with Cluster A consisting of Gram-positive groups and Cluster B Gram-negative groups. Cluster A contains three Gram-positive reference strains Lysinibacillus sp. (MG757657), *Bacillus cereus* (FJ982661.1) and Paenibacillus sp. (AF273740.1). Cluster B contains three Gram-negative reference strains *Aeromonas* sp. (MF185210.1), *Mixta intestinalis* (MN061005.1) and *Escherichia coli* (J01859.1). The reference strains in Cluster A and B were grouped separately from each other in each cluster with bootstrap values between 93%-100% for each reference strain. The bootstrap values for the various isolates from Plant A and Plant B were all above 50%. Two of the isolates from Plant B (V37 and V38) were grouped with the reference strain *Escherichia coli* (J01859.1) with 100% bootstrap support for this relationship.

Table 5.2: Average VBNC counts and physico-chemical results for Plant A.

Sampling Period	Source	VBNC COUNTS			PHYSICO-CHEMICAL PARAMETERS		
		m-FC Agar	Nutrient Agar & m-FC Agar	Nutrient Agar Plus & m-FC Agar	pH	Turbidity	Free Chlorine
		Average CFU/100 ml & STDEV			pH units	NTU	mg/L
March 2016[a]	Raw Water	-	-	-	**8.11***	**2.77***	**-**
	After chlorination[a]	**0** ± 0.00	**16**[a] ± 4.93	**23**[a] ± 5.03	**7.40**	**0.53** ± 0.03	**1.23**
	Point of consumption[a]	**0** ± 0.00	**2**[a] ± 1.00	**3**[a] ± 3.21	**8.05**	**0.37** ± 0.19	**0.07**
April 2016 (1)[a]	Raw Water	**-**	**-**	**-**	**8.30***	**1.49***	**-**
	After chlorination[a]	**0** ± 0.00	**12**[a] ± 2.65	**15**[a] ± 2.65	**7.86***	**0.51***	**0.03***
	Point of consumption[a]	**0** ± 0.00	**1**[a] ± 0.58	**2**[a] ± 1.0	**-**	**-**	**-**
April 2016 (2)[b]	Raw Water	**-**	**-**	**-**	**8.48***	**3.29***	**-**
	After chlorination[a]	**2** ± 1.15	**TNTC**[a]	**TNTC**[a]	**7.12***	**0.40***	**0.70***
	Point of consumption[a]	**1** ± 0.58	**182**[a] ± 21.46	**202**[a] ± 13.08	**-**	**-**	**-**
November 2016[a]	Raw Water	**-**	**-**	**-**	**8.43***	**4.84** ± 0.20	**-**
	After chlorination[a]	**0** ± 0.00	**13**[a] ± 3.61	**15**[a] ± 2.65	**7.99**	**0.58** ± 0.03	**2.16** ± 0.06
	Point of consumption[a]	**0** ± 0.00	**12**[a] ± 2.00	**12**[a] ± 1.53	**8.04**	**0.59** ± 0.06	**0.69** ± 0.01

Mean values ± standard deviations represent the results obtained from triplicate samples. Statistical differences are indicated by alphabetical letters. The same letters indicate no statistical differences ($p > 0.05$). **TNTC**: Too numerous to count; ***** Measured by municipality

Table 5.3: Average VBNC counts and physico-chemical results for Plant B.

Sampling Period	Source	VBNC COUNTS						PHYSICO-CHEMICAL PARAMETERS		
		m-FC Agar	Nutrient Agar & m-FC Agar	Nutrient Agar Plus & m-FC Agar	MLSA	Nutrient Agar & MLSA	Nutrient Agar Plus & MLSA	pH	Turbidity	Free Chlorine
		Average CFU/100 ml & STDEV						pH units	NTU	mg/L
November 2015[a]	Raw Water	**33** ± 15.50	**TNTC**	-	**37** ± 3.21	**TNTC**	-	**8.60**	**0.94** ± 0.25	**-**
	After chlorination[a]	**0**	**38** ± 6.11	-	**0**	**1** ± 0.0	-	**8.58**	**1.75** ± 0.39	**0.33**
December 2015[a]	After chlorination[a]	**0**	**20[a]** ± 2.89	**13[a]** ± 4.36	**0**	**0**	**1** ± 0.0	**-**	**-**	**-**
March 2016[a]	Raw Water	**-**	**-**	**-**	-			**8.56**	**0.20** ± 0.05	**-**
	After chlorination[a]	**1** ± 0.0	**3[a]** ± 1.73	**3[a]** ± 0.58	-	-	-	**8.54**	**2.33** ± 0.02	**1.28****
	Point of consumption[a]	**1** ± 0.0	**1[a]** ± 0.0	**2[a]** ± 1.41	-	-	-	**8.52**	**1.26** ± 0.06	**1.06****
November 2016[b]	Raw Water	**-**	**-**	**-**	-	-	-	**8.57**	**0.37** ± 0.02	**-**
	After chlorination[a]	**0** ±0.0	**50+[a]**	**50[a]** ± 18.38	-	-	-	**8.43**	**2.25** ± 0.02	**1.94**
	Point of consumption[a]	**0** ± 0.0	**156[a]** ± 0.0	**177[a]** ± 3.61	-	-	-	**8.51**	**1.46** ± 0.04	**0.01**
September 2017[a]	After chlorination[a]	**0**	**17[a]** ± 5.03	**16[a]** ± 2.08	-	-	-	**-**	**-**	**-**
	Point of consumption[a]	**0**	**3[a]** ± 1.73	**14[a]** ± 4.36	-	-	-	-	-	-

Mean values ± standard deviations represent the results obtained from triplicate samples. Statistical differences are indicated by alphabetical letters. The same letters indicate no statistical differences ($p > 0.05$). **TNTC**: Too numerous to count; ***** Measured by municipality; ** Total chlorine

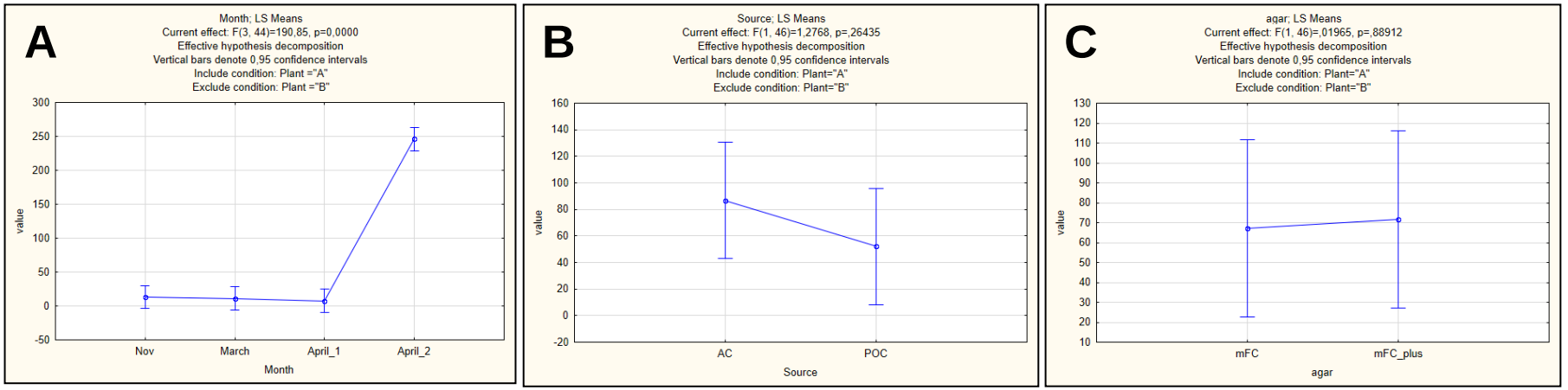


Figure 5.1: Interval plots for Plant A indicating the statistically significant differences between the months, source and media of VBNC counts. Plot A: April (2) had a statistically significant difference ($p < 0.05$) from the other sampling months. Plot B and C: There was no statistically significant difference ($p > 0.05$) between the source of sampling (AC: after chlorination and POC: point of consumption) as well as the media used.

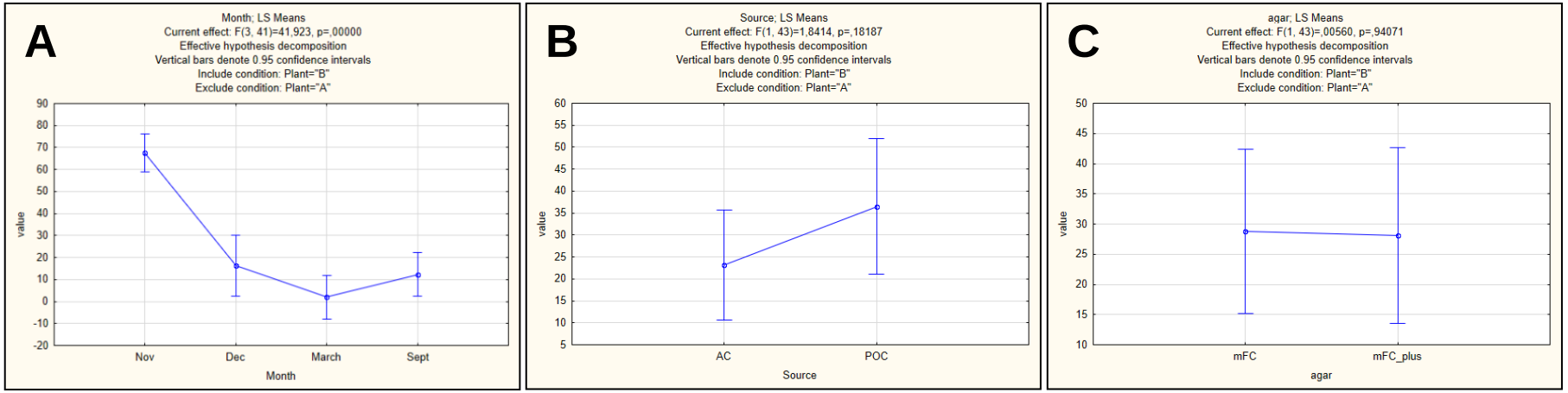


Figure 5.2: Interval plots for Plant B indicating the statistically significant differences between the months, source and media of VBNC counts. Plot A: November 2016 had a statistically significant difference ($p < 0.05$) from the other sampling months. Plot B and C: There was no statistically significant difference ($p > 0.05$) between the source of sampling (AC: after chlorination and POC: point of consumption) as well as the media used.

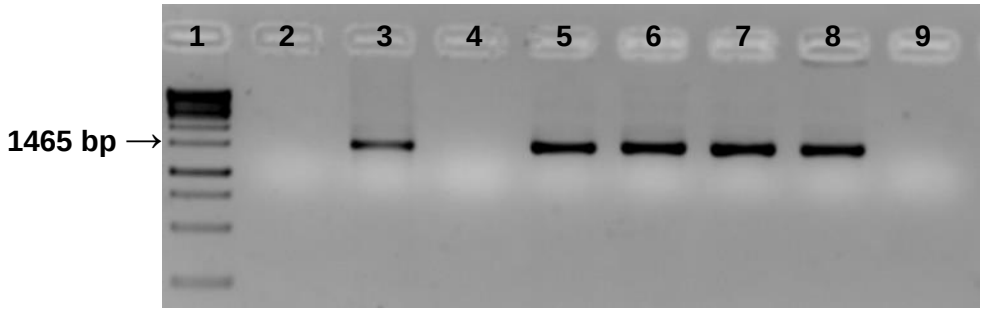


Figure 5.3: A negative image of a 1,5% (w/v) agarose gel indicating the 16S rRNA PCR results for November 2016 for Plant A. The first lane contains the 1 Kb molecular marker. The second lane contains the no template control and lane three the positive control (*E. coli* ATCC 10536). Lanes four to nine indicate the PCR products of the various isolates. The 16S rRNA gene has a fragment size of 1465 bp. Similar results were obtained for Plant B.

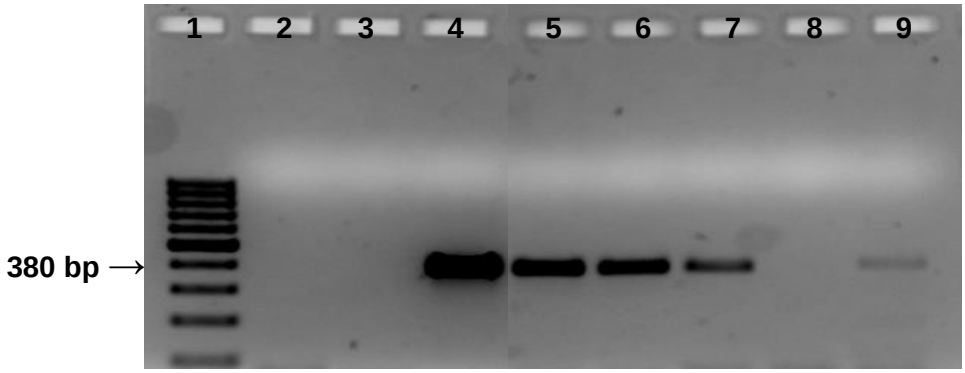


Figure 5.4: A negative image of a 1,5% (w/v) agarose gel (composite image) indicating the *uidA* PCR results for March 2016 for Plant B. The first lane contains the 100 bp molecular marker. The second and third lanes contain the no template controls. Lane four contains the positive control (*E. coli* ATCC 10536). Lanes five to nine indicate the PCR products of the various isolates. The *uidA* gene has a fragment size of 380 bp.

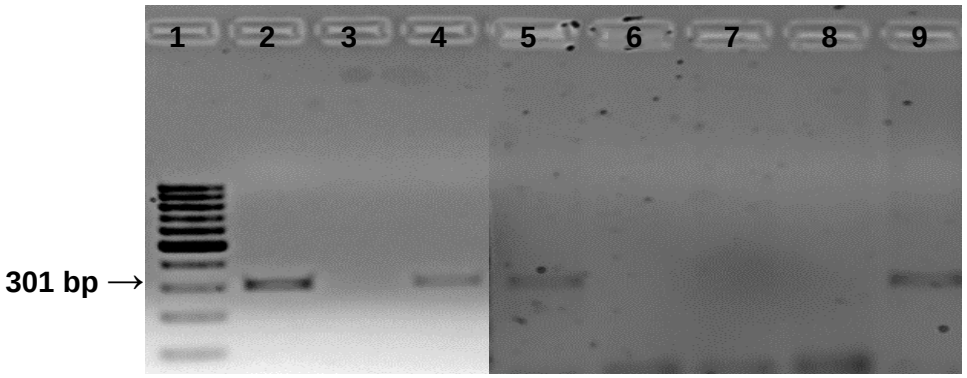


Figure 5.5: A negative image of a 1,5% (w/v) agarose gel (composite image) indicating the *mdh* PCR results for March 2016 for Plant B. The first lane contains the 100 bp molecular marker. The second lane contains the positive control (*E. coli* ATCC 10536). The third lane contains the no template control. Lanes four to nine indicate the PCR products of the various isolates. The *mdh* gene has a fragment size of 301 bp.

Table 5.4: PCR and sequencing results for Plant A.

SAMPLING PERIOD	ISOLATE	SOURCE	GENES TARGETED			ID
			16S	*uidA*	*mdh*	
NOVEMBER 2016	P1	After chlorination	✓	✓	✓	16S: *Lysinibacillus fusiformis* *uidA:* PCR unsuccessful *mdh*: PCR unsuccessful
	P2	After chlorination	✓	✓	✓	16S: Paenibacillus sp. *uidA:* PCR unsuccessful *mdh*: PCR unsuccessful
	P3	Point of consumption	✓	✓	✓	16S: *Lysinibacillus parviboronicapiens* *uidA:* PCR unsuccessful *mdh*: PCR unsuccessful
	P4	Point of consumption	✓	✓	✓	16S: Bacillus sp. *uidA:* PCR unsuccessful *mdh*: PCR unsuccessful

Table 5.5: PCR and sequencing results for Plant B – Isolates enumerated without resuscitation.

SAMPLING PERIOD	ISOLATE	SOURCE	GENES TARGETED			ID
			16S	*uidA*	*mdh*	
MARCH 2016	V10	Point of consumption	✓	✓	✓	16S: *Aeromonas veronii* *uidA*: ***Escherichia coli*** *mdh*: PCR unsuccessful
	V11	After chlorination	✓	✓	✓	16S: *Aeromonas veronii* *uidA*: ***Escherichia coli*** *mdh*: PCR unsuccessful
	V12	Point of consumption	✓	✓	✓	16S: *Aeromonas veronii* *uidA*: ***Escherichia coli*** *mdh*: PCR unsuccessful
	V14	After chlorination	✓	✓	✓	16S: *Aeromonas veronii* *uidA*: ***Escherichia coli*** *mdh*: PCR unsuccessful

Table 5.6: PCR and sequencing results for Plant B - Isolates enumerated with resuscitation.

SAMPLING PERIOD	ISOLATE	SOURCE	GENES TARGETED			ID
			16S	*uidA*	*mdh*	
MARCH 2016	V1	After chlorination	✓	✓	✓	16S: Enterobacter sp. *uidA:* ***Escherichia coli*** *mdh*: ***Escherichia coli***
	V2	After chlorination	✓	✓	✓	16S: *Aeromonas veronii* *uidA*: ***Escherichia coli* O177:H21** *mdh:* PCR unsuccessful
	V4	After chlorination	✓	✓	✓	16S: Problem with sequence *uidA:* ***Escherichia coli*** *mdh:* ***Escherichia coli***
	V6	Point of consumption	✓	✓	✓	16S: *Aeromonas veronii* *uidA*: ***Escherichia coli*** *mdh:* PCR unsuccessful
NOVEMBER 2016	V26	After chlorination	✓	✓		16S: *Paenibacillus taichungensis* *uidA:* PCR unsuccessful
	V27	After chlorination	✓	✓		16S: *Lysinibacillus macroides* *uidA:* PCR unsuccessful
	V28	After chlorination	✓	✓		16S: *Paenibacillus polymyxa* *uidA:* PCR unsuccessful
	V29	Point of consumption	✓	✓		16S: *Aeromonas hydrophila* *uidA:* PCR unsuccessful
	V30	Point of consumption	✓	✓		16S: *Lysinibacillus fusiformis* *uidA:* PCR unsuccessful
	V31	Point of consumption	✓	✓		16S: *Aeromonas veronii* *uidA*: ***Escherichia coli***
	V32	Point of consumption	✓	✓		16S: *Lysinibacillus fusiformis* *uidA:* PCR unsuccessful
	V33	Point of consumption	✓	✓		16S: *Paenibacillus typhae* *uidA:* PCR unsuccessful

Table 5.6 continue

SAMPLING PERIOD	ISOLATE	SOURCE	GENES TARGETED			ID
			16S	*uidA*	*mdh*	
SEPTEMBER 2017	V34	After chlorination	✓			*Bacillus cereus*
	V35	After chlorination	✓			*Lysinibacillus parviboronicapiens*
	V36	Point of consumption	✓			*Bacillus cereus*
	V37	Point of consumption	✓			***Escherichia coli***
	V38	Point of consumption	✓			***Escherichia coli* O157**
	V39	Point of consumption	✓			*Bacillus cereus*
	V40	Point of consumption	✓			*Mixta intestinalis*

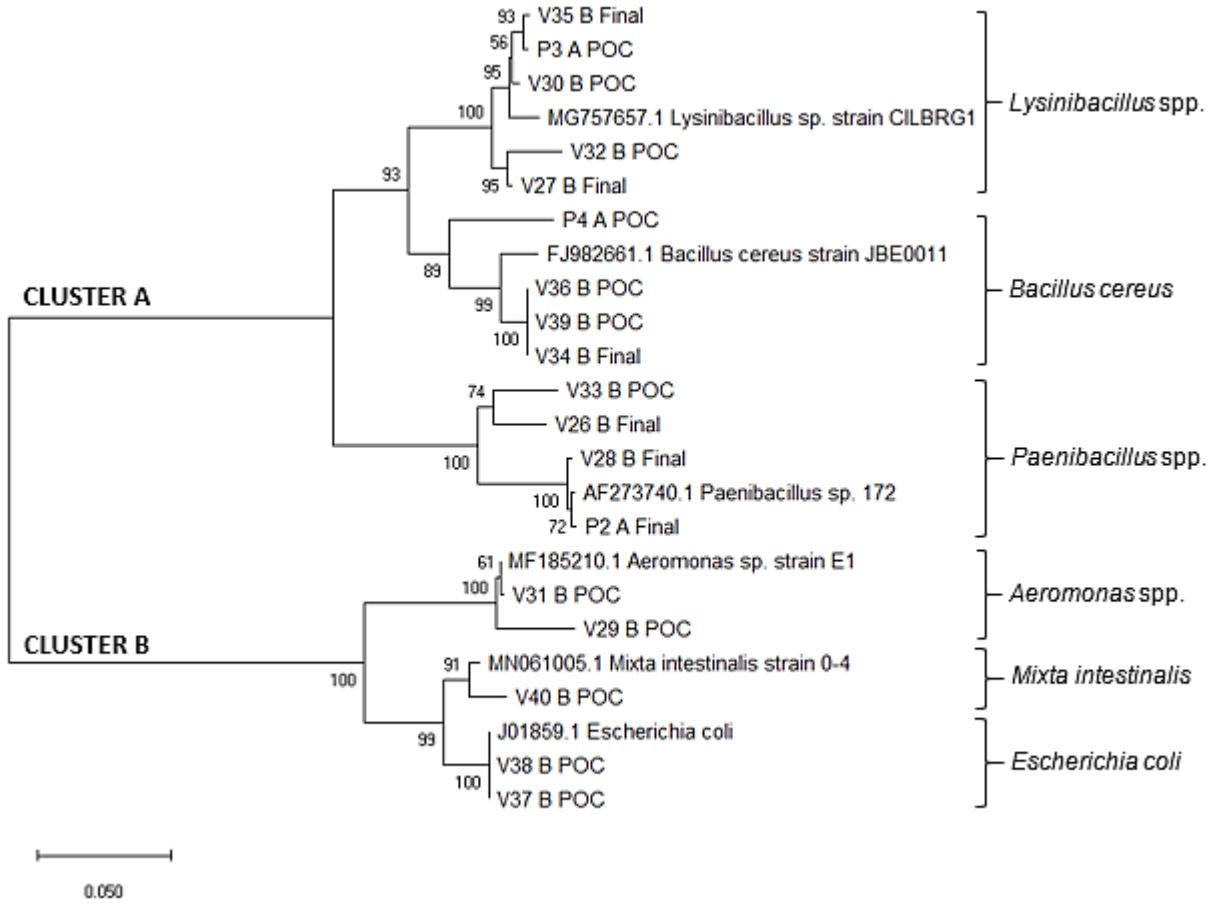


Figure 5.6: Phylogenetic relationships among the VBNC isolates identified in this study as inferred using the Neighbor-Joining method (MEGA X software). To calculate the evolutionary distances the Tamura-Nei method was used. Bootstrap values (in percentage) are indicated on each branch and were calculated from 10 000 replicates. The sum of the branch length was 0.77594586. Strain designations: Strains were designated as InPS, with In standing for Isolate name (e.g. V35/P3), P standing for the plant from which the isolates were enumerated (A or B) and S standing for the source of isolation (Final: directly after chlorination; POC: point of consumption). Reference cultures were indicated starting with their GenBank accession numbers followed by the Genus and specie names.

5.4 Discussion

5.4.1 Resuscitation of VBNC *E. coli*

Two methods were used for the resuscitation of VBNC *E. coli* cells from drinking water. The first method was a culture-based method as described by Özkanca *et al.* (2009). In their study, Özkanca *et al*. (2009) successfully resuscitated heat stressed VBNC *E. coli* cells by using a multi-well MPN method and membrane filter method. An average of 67 CFU/10ml were resuscitated on the MLSA medium. However, in the current study the method was not as successful in resuscitating stressed *E. coli* from drinking

water. Only one CFU/100 ml were enumerated during November and December 2015 respectively.

A possible reason for this may have been that Özkanca *et al*. (2009) induced the *E. coli* cells into a VBNC state and subsequently resuscitated these cells in a controlled environment without any other environmental factors playing a role. However, in the current study, the *E. coli* cells were enumerated from drinking water which were subject to other environmental factors, as well as the presence of background bacteria, which may have influenced the resuscitation on the MLSA medium. Özkanca *et al*. (2009) suggested the reason for the recovery of *E. coli* on some selective media and not on others may be due to the selective agents in the media. The stressed cells may be more penetrable affecting their recovery.

Method adapted from Özkanca *et al*. (2009):

Against this background, the aim of the current study was to use unaltered commercially available m-FC medium to resuscitate VBNC *E. coli* from drinking water. Our method was based on the method as described by Özkanca *et al*. (2009), but with additional nutrients being provided in the first step of resuscitation and unaltered m-FC medium instead of MLSA was used as selective media after initial resuscitation. The use of m-FC medium with membrane filtration technique was introduced by Geldreich and colleagues in 1965 (Geldreich *et al.*, 1965) and has been used ever since as an easy cost-effective way of isolating faecal coliforms from water samples (Mahesh & Prasanth, 2015; Bahgat *et al.*, 2018; Shamimuzzaman *et al.*, 2019). Therefore, the decision to use m-FC in the current study to resuscitate *E. coli* cells was plausible.

The use of m-FC medium for the resuscitation of stressed coliforms has previously been demonstrated: Grabow *et al.* (1981) evaluated the use of standard m-FC, modified m-FC, MacConkey and Teepol media for the isolation of faecal coliforms from water by means of membrane filtration. The m-FC medium was modified by removing the rosolic acid or adding resuscitation top layer which consisted of 13 g lactose broth, 15g agar and 1L distilled water. An amount of 2.6 ml of the lactose agar was poured on top of 6.5ml of standard m-FC medium. The plates were first incubated at 35°C for 2 hours and then at 44.5°C for 22 to 24 hours. Results of the study indicated that the modified m-FC agar yielded higher average counts than the standard m-FC medium.

This included the m-FC medium without rosolic acid as well as the m-FC medium containing the resuscitation layer.

Pagel *et al.* (1982) tested different medias for the recovery of faecal coliforms. In their study MacConkey, mTEC and m-FC media were used. The m-FC medium was used as is, but the method was also modified by leaving out the rosolic acid and adding a resuscitation step (m-FC2). The resuscitation step included incubating the m-FC2 plates for 2 hours at 35°C. The study indicated that recoveries of faecal coliforms on the m-FC2 media were significantly higher than the other media. However, the growth of non-targeted background bacteria was also higher on the m-FC2 medium.

In the study by Calabrese & Bissonnette (1990) amendments were made to m-FC, mT7, M-Endo and tryptone-glucose-yeast (TGY) extract media for the detection of chlorine-stressed coliform bacteria. Five different alterations were made to the media including supplementation with catalase, heat-inactivated catalase, sodium pyruvate, catalase-sodium pyruvate combination and acetic acid. Their study concluded that the supplementation of m-FC medium with dissolved crystalized bovine liver catalase, sodium pyruvate, or both significantly increased the recovery of coliform bacteria.

Even though the latter studies were able to resuscitate stressed coliforms using m-FC medium, in all three studies the m-FC medium had to be amended for the resuscitation to be successful. In the current study, however, m-FC medium was used as is for the resuscitation of *E. coli*. The combination of using nutrient agar and unaltered m-FC agar proved to be very successful as *E. coli* cells were enumerated after chlorination and at the point of consumption for Plant A and Plant B (Table 5.2 and 5.3). No statistical differences were shown between the use of nutrient agar or nutrient agar with added nutrients ($p > 0.05$) for both Plant A and Plant B. Therefore, it should be more cost effective to use standard nutrient agar for the initial resuscitation step.

There was also no statistical difference between the sampling sources (after chlorination and point of consumption). This could be an indication that the VBNC cells were distributed through the drinking water network and not removed by the free chlorine in the water. The WHO recommends that a minimum of 0.2 mg/L needs to be maintained in the drinking water distribution network up to the end to minimise the risk of contamination by harmful organisms in the network (World Health Organization, 2011; World Health Organization, 2017a). However, a study by Owoseni *et al.* (2017)

indicated that free chlorine levels of 0.50 mg/L were not as effective in removing *E. coli* cells from wastewater treatment plants. Their study concluded that *E. coli* cells were optimally removed at free chlorine levels of 1.5 mg/L.

Even though the free chlorine levels for Plant A and Plant B were within the acceptable SANS 241:2015 levels, VBNC *E. coli* were isolated in the presence of high levels of free chlorine (2.16 mg/L for Plant A and 1.94 mg/L for Plant B), as well as at low levels of free chlorine (0.03 mg/L for Plant A and 0.01 mg/L for Plant B). Similar results were obtained from a study by Chen *et al.* (2018). Their study indicated the ability of *E. coli* cells to enter a VBNC state in the presence of low and high levels of free chlorine (0.5, 1, 2, 3 and 4 mg/L). Denisova *et al.* (2014) also assessed the effect of chlorination on the viability of *E. coli* cells in drinking water. Their study indicated that *E. coli* cells were able to survive and retain metabolic activity at low and high chlorine concentrations (0.05 mg/L, 0.1 mg/L, 0.2 mg/L, 0.5 mg/L and 5 mg/L). They suggest that for the metabolic activity of *E. coli* cells to be rapidly reduced, 5 mg/L of free chlorine was necessary. For drinking water to be safe for consumers, free chlorine levels should not exceed 5 mg/L (World Health Organization, 2017a). Therefore, care should be taken when implementing high levels of free chlorine as this may pose a health risk to consumers (World Health Organization, 2017a).

In the current study, *E. coli* cells were also enumerated from Plant A and Plant B with the standard method of incubating the membranes on m-FC medium only, without any prior resuscitation. These findings suggest that the chlorination step was not efficient in removing the *E. coli* cells from the drinking water. According to the WHO, the effectiveness of chlorine may be influenced by a few factors including the chlorine concentration and contact time, the temperature of the water, the pH level and turbidity (World Health Organization, 2017a). Low levels of residual chlorine (<0.4 mg/L) usually requires longer contact time, whereas high levels of residual chlorine (>3 mg/L) requires less contact time. Disinfection at cooler water temperatures (<5°C) requires more contact time or a higher chlorine concentration. Higher water temperatures (>20°C) requires less contact time or a lower concentration of chlorine. Chlorine disinfection is also more effective at lower pH levels. Drinking water with pH levels higher than pH 8 requires longer contact time or a higher concentration of chlorine. Turbidity levels also influence the effectiveness of chlorine disinfection. High turbidity levels require higher chlorine levels for longer contact times.

Historical and measured data indicated that the pH levels of the raw water of Plant A and Plant B were above 8 pH units (Table 5.2 & 5.3). Turbidity levels of the raw water of Plant A were below 3 NTU during March 2016 and April (1) 2016, but during April (2) 2016 and November 2016, turbidity levels of the raw water exceeded 3 NTU (Table 5.2). For Plant B, the turbidity levels of the raw water were generally very low (< 1 NTU). However, after chlorination there was a steep increase in turbidity levels, exceeding the SANS 241:2015 standard of < 1 NTU (Table 5.3). This is an indication that there may have been contamination of organic and inorganic material during the treatment process of Plant B. The latter statement proved to be true as results from the study in Chapter 2 of this thesis indicated that turbidity levels of Plant B increased after filtration (see Tables 2.7 - 2.10). The high turbidity and pH levels may have contributed to the ineffectiveness of the chlorination which may have resulted in the isolation of *E. coli* in the drinking water of Plant A and B.

Statistical differences could be seen between the sampling periods for Plant A and Plant B. For Plant A, sampling during April (2) 2016 were statistically different ($p < 0.05$) from March 2016, April (1) 2016 and November 2016 (Figure 5.1A). This may be attributed to the highest number of VBNC cells recovered after chlorination and at the point of consumption during this sampling period (Table 5.2). For Plant B, sampling during November 2016 was statistically different ($p < 0.05$) from December 2015, March 2016 and September 2017 (Figure 5.2A). November 2016 was the sampling period that yielded the highest number of VBNC cells isolated after chlorination and at the point of consumption (Table 5.3).

5.4.2 Identification of VBNC cells

(a) Plant A

Although high levels of *E. coli* cells were enumerated on the membrane filters during resuscitation (Table 5.2), molecular identification of these isolates was unfortunately for the most part unsuccessful. Only four isolates from the November 2016 sampling period were able to amplify the 16S rRNA gene. Two of the isolates were identified as *Lysinibacillus fusiformis* and *Lysinibacillus parviboronicapiens* and the other two isolates as *Paenibacillus* sp. and *Bacillus* sp. (Table 5.4). Having to submerge the entire membrane filter into nutrient broth during the resuscitation step meant that all the bacterial cells present in the water could be trapped on the membrane and would, therefore, be able to multiply in the nutrient broth. This may have been the reason for

detecting these Gram-positive bacteria during 16S rRNA sequencing. Submerging the entire membrane in nutrient broth may also have caused unwanted contamination of the samples resulting in low quality DNA amplified and unsuccessful *uidA* and *mdh* PCR's. Phylogenetic analysis (Figure 5.6) grouped the isolates correctly on Cluster A with their corresponding reference cultures. Bootstrap values for these relationships were above 90%. Even though these isolates were not identified as *E. coli*, the presence of *Lysinibacillus* and *Paenibacillus* in drinking water is still concerning as these bacteria has been known to cause human illness (Castagnola *et al.*, 2001; Wenzler *et al.*, 2015).

(b) Plant B

Molecular identification for Plant B proved to be more successful as the 16S rRNA, *uidA* and *mdh* genes were amplified. As indicated in Table 5.6, the bacteria identified with 16S rRNA sequencing varied from *Aeromonas, Paenibacillus*, *Lysinibacillus*, *Bacillus cereus* and *Mixta intestinalis*. *Aeromonas* are generally found in aquatic environments and have been isolated from drinking water (Parker & Shaw, 2011; Miyagi *et al.*, 2017). *Aeromonas* are not usually enumerated using m-FC agar, however, Figueira *et al.* (2011) also enumerated *A. veronii* from a drinking water treatment plant in Northern Portugal with m-FC medium. *Aeromonas* species have been known to cause gastrointestinal and extraintestinal infections in humans (Gowda *et al.*, 2015; Shen *et al.*, 2018). The presence of *Aeromonas* in drinking water is thus concerning.

More concerning was the isolation of *E. coli* after chlorination as well as at the point of consumption. Previous studies have indicated that *E. coli* are able to survive in water treated with chlorine (Dungeni *et al.*, 2010; Denisova *et al.*, 2014; Owoseni *et al.*, 2017). In the present study, a total of seven isolates were identified as *E. coli* (Table 5.6). These isolates were identified through amplification of the 16S rRNA, *uidA* and *mdh* genes. Similar results were obtained by Zhang *et al.* (2015) which amplified the 16S rRNA gene from VBNC *E. coli.* Lin *et al.* (2017) also amplified the 16S rRNA gene from *E. coli* cells induced into a VBNC state by chlorination. Igwaran *et al.* (2018) identified *E. coli* in wastewater systems by targeting the *uidA* gene and Omar & Barnard (2010) confirmed the presence of *E. coli* in treated water from waste water treatment plants in South Africa by amplifying the *mdh* gene.

Two of the *E. coli* strains in the present study were identified as *E. coli* O177:H21 and O157 (Table 5.6). This is concerning as both *E. coli* O177 and O157 has been identified as Shiga toxin producing pathogens (Beutin *et al.*, 2005; Saxena *et al.*, 2015; Montso *et al.*, 2019). Montso *et al.* (2019) isolated *E. coli* O177 from South African cattle during 2017. Various outbreaks of diseases caused by the presence of O157 in drinking water contaminated by cattle faeces have been documented as these outbreaks caused various human illnesses and fatalities (Hunter, 2003). The presence of *E. coli* O177 in the drinking water of Plant B in the present study may have been due to faecal contamination. Cattle are also the primary source of *E. coli* O157 (Cobbold & Desmarchelier, 2000). However, Ateba & Mbewe (2011) isolated *E. coli* O157 from tap water in the North West Province, South Africa, which was an extremely disconcerting finding, especially as a study by Liu *et al.* (2009) had indicated that *E. coli* O157 cells exposed to chloramine in tap water were able to enter into a VBNC state within 15 minutes.

In the current study *E. coli* housekeeping genes *uidA* and *mdh* proved to be the most successful in positively identifying the resuscitated VBNC cells as *E. coli*. The housekeeping gene *lacZ* could not be amplified successfully. This may have been due to the presence of background bacteria which resulted in the amplification of low quality DNA and unsuccessful PCR's. The unsuccessful amplification of certain housekeeping genes could lead to the false assumption that no *E. coli* is present. Therefore, it is recommended that not only one housekeeping gene be targeted for the identification of VBNC *E. coli* cells.

The resuscitated cells can also be identified by using less expensive methods such as biochemical tests. The EnteroPluri-Test is a system containing 12 sectors of different media which can be used to identify *Enterobacteriaceae* and other Gram-negative, oxidase negative bacteria (Liofilchem, 2020). The API20E is a strip test kit that contain up to 20 miniature biochemical tests. The API20E kit can be used to identify Gram-negative and Gram-positive bacteria and yeast (Biomerieux, 2019).

The eradication of *E. coli* from drinking water is thus very important, because *E. coli* can pose serious health risks to humans when the contaminated water is consumed (Stauber *et al.*, 2016; Ercumen *et al.*, 2017; Navab-Daneshmand *et al.*, 2018). The isolation of *E. coli,* as well as other pathogens in the present study from water after

chlorination and at the point of consumption is an indication of failure in the treatment process for both Plant A and Plant B. This could have serious health implications to individuals consuming the drinking water.

5.5 Conclusion

VBNC *E. coli* was detected in the drinking water of two water treatment plants in the North West Province in the study period 2015 to 2017. Various studies have indicated that chlorination or chloramination of drinking water can induce *E. coli* into a VBNC state (Liu *et al.*, 2009; Liu *et al.*, 2010; Lin *et al.*, 2017; Chen *et al.*, 2018). Plant A and Plant B uses chlorine to disinfect their water. Therefore, the detection of VBNC *E. coli* in the drinking water of Plant A and Plant B may have been induced by the presence of chlorine. However, the possibility of other environmental stresses may have contributed to the VBNC state. The isolates from Plant A were able to amplify the 16S rRNA gene and isolates from Plant B were able to amplify the 16S rRNA, *uidA* and *mdh* genes after resuscitation. These findings proved the resuscitation of VBNC *E. coli* cells in drinking water. The presence of VBNC *E. coli*, especially O177 and O157, in drinking water of the North West Province, is very concerning. The false impression may exist that the drinking water in this region does not contain any *E. coli* and that it is safe for human consumption, but the detected presence of *E. coli* in the drinking water from two plants may have dire consequences for unsuspecting communities in this region.

The aim of this study was to use a relatively easy and cost-effective method to resuscitate VBNC *E. coli* from drinking water. The method as described by Özkanca *et al.* (2009) did not prove to be successful in resuscitating *E. coli* from drinking water. Therefore, the method was adapted and a new method developed. From previous studies it is evident that m-FC agar can be useful in resuscitating VBNC *E. coli* from water. However, in all the studies the m-FC agar was modified and not used as is. Results of the current study demonstrated that unaltered m-FC medium, together with a pre-incubation step on nutrient agar, can be used to isolate VBNC *E. coli* in drinking water. According to literature searched in several databases, no documentation could be found where this method was used.

5.6 Recommendations

- Verify the modified method by using ATCC *E. coli* cells that are manually induced into a VBNC state.
- Use purified colonies for the DNA extraction to eliminate interference from unwanted background bacteria.
- Due to financial limitations that some municipalities may face, it is recommended that further identification of the isolated VBNC cells could be performed using biochemical tests, such as the EnteroPluri-Test or API20E, instead of molecular identification.

CHAPTER 6

Conclusions and Recommendations

6.1 Introduction

South Africa is a water scarce country with average annual rainfall levels far below the world average of 860mm a year. Sixty five percent of the country receives less than 500mm of rainfall a year (Van Vuuren, 2015). Since 2015, the North West province, among others, has experienced what has been described as one of the worst droughts in 23 years (Van Vuuren, 2015). Droughts have a negative impact on river systems as reduced flow causes an increase in contaminants resulting in poor water quality (Deksissa *et al.*, 2003). However, it is not only natural events causing the deterioration of source water bodies in South Africa. Human-induced factors, such as discharges of municipal and industrial wastewater and agricultural and mining activities, also have a negative impact on source water (Van der Laan *et al.*, 2012). Should the challenge of water scarcity in South Africa continue, it may force water purification plants to use every available water resource, even those with very poor water quality (Morrison *et al.*, 2012).

The National Water Services Act (no. 108 of 1997) states that everyone has the right to basic water supply, i.e. *"the prescribed minimum standard of water supply services necessary for the reliable supply of a sufficient quantity and quality of water to households, including informal households, to support life and personal hygiene"* (South Africa, 1997). The Act also states that this basic water supply must not be harmful to health or well-being. Therefore, it is of utmost importance for water services providers, such as municipalities, to ensure the production of safe drinking water.

However, water purification plants are often faced with challenges, such as financial and operational issues, lack in training of operators and overall management issues (Momba & Brouckaert, 2005; Meng *et al*., 2015). These challenges put water purification plants under immense pressure to produce safe drinking water to communities (Badejo *et al.*, 2015; Gray & Vawda, 2016). Deterioration of source water and basic treatment challenges often means advanced treatment processes are necessary to achieve the goal of supplying safe and adequate drinking water (Iglesias *et al.,* 2014). However, advanced treatment may be costly and not all water purification companies have the financial resources and sustainability to implement advance

treatment (Iglesias *et al.,* 2014). There is thus a need for more cost effective and rapid solutions to monitoring drinking water and improve the overall treatment process.

The aim of this study was to establish improved monitoring of treatment processes at drinking water purification plants in the North-West Province. To achieve this aim, three decision-making tools, namely HACCP, predictive rule sets (ANNs an EAs) and unconventional microbial monitoring, were investigated.

6.2 Comparative analysis of the HACCP concept

The application of the HACCP concept was investigated at three drinking water purification plants (Plant A, Plant B, Plant C). The aim of the study was to evaluate the effectiveness of the HACCP concept using historical and measured data. Plant A and B only monitor source and drinking water, whereas Plant C already implemented HACCP concept.

Historical data from 2009 to 2015 was used to determine the average percentage turbidity removal and the lowest percentage turbidity removal for each year. Possible process failures could be identified by comparing the lowest and average percentage turbidity removals for each year. Plant A had an average percentage turbidity removal above 80% from 2009 to 2015. However, during 2013, an average percentage removal of only 8% was obtained. Historical data of Plant B indicated serious treatment issues as the turbidity levels mostly increased from the source water to the drinking water from 2011 to 2014. Due to Plant A and Plant B only performing end-point testing, the failures in the treatment process responsible for the insufficient removal of the turbidity could not be identified. Historical data of Plant C also indicated incidences where turbidity levels were reduced below the average level. The average percentage turbidity removal for Plant C from 2009 to 2015 remained above 95%. However, during 2011 a turbidity removal of 88% was obtained. The advantage of implementing the HACCP concept could be illustrated by the latter incident, as the historical data of Plant C was used to indicate where the treatment process failed on that day in removing the turbidity efficiently.

As a result of the historical data of Plant A and Plant B indicating possible treatment issues, physico-chemical and microbiological parameters were measured after each step in the treatment processes of Plant A and Plant B during 2015 and 2016. No irregularities in the treatment process of Plant A could be identified. However, for Plant

B a clear indication of where the treatment process failed could be illustrated. During all the sampling periods, it was noted that the turbidity levels increased after filtration and these high levels persisted after chlorination and at the point of consumption. This was concerning as the turbidity levels in the drinking water exceeded the SANS 241:2015 standard.

Historical and measured data from this study indicated that should a treatment process be underperforming, end-point testing may not be sufficient to identify which step in the treatment process is responsible for the failure. In addition, the study indicated that the implementation of the HACCP concept at drinking water purification plants may be an effective decision-making tool for plant operators and managers, as process failures can be identified, a solution be implemented and future occurrences prevented.

6.3 Artificial neural networks: Literature review of applications in the drinking water sector

The purpose of this review was to indicate the current status of the applications of ANNs in the drinking water sector, especially in South Africa. Firstly, the review gave an overview of the principles of ANNs, after which the application of ANNs in the water sector and specifically the drinking water sector was indicated. Lastly future prospects were given.

From the literature it was evident that ANNs have been applied in various fields of the water sector since the 1990's. Applications in the drinking water sector during that time were limited and most studies pertained to the prediction of pipe pressure/leakage, scheduling of booster disinfection and coagulation/flocculation dosages. During the last decade, the applications in the drinking water sector have increased dramatically.

This was not the case for South Africa, where applications in the water sector remained limited. Studies related to environmental water included the prediction of streamflow (six studies), reservoir capacity (two studies), rainfall data (two studies), river-runoff (one study), water demand (one study) and water temperature (one study). A study by Naidoo & van der Walt (2013) was performed in the drinking water sector in which ANNs were used to determine the chemical dosing levels.

This review indicated that ANNs are useful forecasting tools in the water sector. Also, the gap in research with regards to the application of ANNs in the drinking water sector

of South Africa was highlighted. International studies indicated that ANNs can be employed as decision-making tools.

6.4 Empirical application of artificial neural networks and evolutionary algorithms at drinking water purification plants

Previous studies have indicated that ANNs can be useful for the prediction of various parameters in the water sector. However, studies about the application of ANNs in the drinking water sector of South Africa are limited. Even though ANNs are efficient in adapting and learning, they have the disadvantage of being “black box” models as they do not provide an equation or rule-set to explain how the forecasting results were obtained (Bezuidenhout *et al*., 2013; Delpla *et* al., 2019). Evolutionary algorithms may be applied to overcome this obstacle as they can be used to determine predictive rule sets (Swanepoel *et al.,* 2016; Recknagel *et al.,* 2008).

Therefore, the aim of this study was to apply ANNs and EAs at three drinking water purification plants in the North West Province. Historical data from 2009 to 2015 of Plant A, Plant B and Plant C were obtained for the ANN modelling. However, due to limited historical data of Plant B, a prediction model could not be established. This highlighted the importance of regular monitoring of source and drinking water. For Plant A and Plant C electrical conductivity in the drinking water was successfully predicted. The prediction models for turbidity for Plant A and Plant C were not as successful. The percentage accuracy for the training and test sets were below 60%. The EC models for Plant A and Plant C, on the other hand, produced 100% good forecast for both the test and training sets for Plant A and 98% and 100%, respectively, for Plant B. EAs were then applied to determine predictive rule sets for EC in the drinking water of Plant A and Plant C. Results indicated that accurate predictive rule sets were generated for Plant A and Plant C. These rule sets may be applied by the water purification plants to serve as early warning systems should the predicted EC value differ significantly from the measured value. It may also be used to determine the maximum allowable levels for certain parameters in the source water, enabling plant operators to be prepared should these levels become a reality.

This study demonstrated that ANNs and EAs were successfully applied to predict water quality parameters in drinking water, subsequently generating predictive rule sets which may be used as decision-making tools.

6.5 Recovery of *Escherichia coli* in drinking water

Microbiological contamination of drinking water, especially faecal pollution, may pose serious health risks to consumers (Nwabor *et al*., 2016; Ontumbi *et al*., 2015). SANS 241:2015 requires water purification plants to monitor certain microbiological determinands in the drinking water, such as *E. coli*. The specification for *E. coli* according to SANS 241:2015 is 0 CFU/100 ml.

Recently, studies have indicated that *E. coli* is able to enter into a VBNC state (Zhang *et al.*, 2015; Lin *et al.*, 2017; Chen *et al.*, 2018). This means *E. coli* may not be able to grow on conventional media, but may still retain viability and pose a health risk to consumers. The presence of VBNC *E. coli* in drinking water may have serious consequences for water purification plants using conventional media to enumerate *E. coli*, as they may be under the false impression that the *E. coli* levels are zero. Therefore, the aim of this study was screen drinking water in the North West Province for the presence of stressed or injured *E. coli* by recovering these cells with culture-based methods.

Results from the study indicated that *E. coli* cells were recovered in the drinking water of water purification plants in the North West Province. The VBNC isolates were able to amplify the 16S rRNA, *uidA* and *mdh* genes after resuscitation. Two of the isolates were identified as *E. coli* O177 and O157. This is concerning as *E. coli* O177 and O157 are both Shiga toxin-producing pathogens which may cause serious illness (Farrokh *et al.*, 2012; Saxena *et al.*, 2015; Montso *et al.*, 2019). In addition to these results, a new culture-based method was developed for the resuscitation of VBNC *E. coli*. This study indicated that the use of unaltered m-FC media, together with a pre-incubation step on nutrient agar, can be used to isolate VBNC *E. coli* in drinking water. The Blue Drop Certification Programme and SANS 241:2015 provides a framework for water purification plants to ensure the production of safe drinking water. However, results from this study indicated the presence of unconventional pathogens in drinking water, which brings the current monitoring requirements set out in SANS 241:2015 into question. Screening for additional unconventional microbiological parameters may indicate possible process failures that would otherwise have been missed.

Data presented in this study indicated that there are various decision-making tools available which drinking water purification plants can use to improve monitoring of

water purification processes and improve the quality of drinking water delivered to consumers.

6.6 Recommendations

- In this study the application of the HACCP concept at Plant A indicated that no irregularities within the treatment process could be established. This might have been due to inadequate data obtained as a result of infrequent sampling. It is recommended to increase the sampling frequency to include data from a number of consecutive days. This may increase the possibility to detect irregularities within the treatment process. This does not have to include monitoring of all the parameters, but easy measurable and inexpensive parameters, such as EC, pH, turbidity and TDS.

- This study included three water purification plants of which one already implemented the HACCP concept (Plant C). It is recommended that more water purification plants are included to be examined. The application of the HACCP concept at water purification plants with similar treatment processes and raw water resources can be compared.

- Three decision-making tools were presented in the current study to improve the monitoring of the purification processes of water purification plants. It is recommended that the HACCP concept take priority for implementation at local municipalities as this decision-making tool may indicate underlying problems within the treatment system, which could improve the quality of the drinking water rapidly.

- Since the application of ANNs and EAs in the drinking water quality sector is very limited, more work and data in this important field is required. This will verify the deductions of the current limited study.

- Literature indicated that the most common used ANN is the feed-forward network. Therefore, this study used a feed-forward network to generate the prediction models. Even though high accuracy models for EC were developed with the feed-forward ANN, the turbidity models underperformed. It is recommended that the application of other ANN models, such as RBF, GRNN and CFN are investigated to compare their performance with the feed-forward network for the prediction of drinking water parameters, especially turbidity.

- For municipalities who have a database of historical data available, the application of ANN's and EA's will be a major advantage. On the other hand, for municipalities who do not monitor their water frequently, the effectiveness of ANN's and EA's indicated by the current study should be an encouragement for these local municipalities to start to build a database by increasing monitoring throughout the system.

- A modified method for the recovery of *E. coli* in drinking water was developed in this study. It is recommended that this modified method should be tested and verified using ATCC *E. coli* cells that is manually induced into a viable but nonculturable state by applying different types of stresses (i.e. different chlorine levels, high and low temperatures, high osmotic levels, starvation, etc.). This will determine if the method is able to resuscitate VBNC *E. coli* cells induced by various stresses or only specific stresses.

- The current study indicated that in an effort to dislocate the VBNC colonies from the membranes during the DNA preparation step (section 5.2.4), submerging the entire membrane into nutrient broth and using a vortex to dislocate the colonies, may have resulted in the growth of unwanted background bacteria and subsequent DNA amplification of these bacteria. It is recommended that the nutrient broths be streaked out onto nutrient agar or selective agar to try and isolate single colonies. These single colonies can then be identified through biochemical tests to screen for *E. coli* before further molecular tests are performed.

- The isolation of VBNC *E. coli* should be added by municipalities as part of their risk management programs to ensure that the water produced by the water purification plant does not contain any *E. coli* which may pose a health risk to the community consuming the water. Due to some the financial difficulties that some municipalities are facing, it is recommended that water purification plants test for VBNC *E. coli* at least seasonally or after a major challenge or pollution event took place.

REFERENCES

ABDULKADIR, T., SULE, B. & SALAMI, A. 2012. Application of artificial neural network model to the management of hydropower reservoirs along River Niger, Nigeria. *Annals of the Faculty of Engineering Hunedoara,* 10(3)**,** 419.

ABKENAR, S. M. S., STANLEY, S. D., MILLER, C. J., CHASE, D. V. & MCELMURRY, S. P. 2015. Evaluation of genetic algorithms using discrete and continuous methods for pump optimization of water distribution systems. *Sustainable Computing: Informatics and Systems,* 8, 18-23.

ADAMOWSKI, J. & KARAPATAKI, C. 2010. Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms. *Journal of Hydrologic Engineering,* 15(10)**,** 729-743.

ADAMOWSKI, J. F. 2008. Peak daily water demand forecast modeling using artificial neural networks. *Journal of Water Resources Planning and Management,* 134(2)**,** 119-128.

ADELOYE, A. 2009. The relative utility of regression and artificial neural networks models for rapidly predicting the capacity of water supply reservoirs. *Environmental Modelling & Software,* 24(10)**,** 1233-1240.

ADELOYE, A. & DE MUNARI, A. 2006. Artificial neural network based generalized storage–yield–reliability models using the Levenberg–Marquardt algorithm. *Journal of Hydrology,* 326(1)**,** 215-230.

ADGAR, A., COX, C. S. & BÖHME, T. J. 2000. Performance improvements at surface water treatment works using ANN-based automation schemes. *Chemical Engineering Research and Design,* 78(7)**,** 1026-1039.

AJBAR, A. & ALI, E. M. 2015. Prediction of municipal water production in touristic Mecca city in Saudi Arabia using neural networks. *Journal of King Saud University - Engineering Sciences,* 27(1)**,** 83-91.

AL-ALLAF, O. N. A. 2012. Cascade-forward vs. function fitting neural network for improving image quality and learning time in image compression system. Proceedings of the world congress on engineering, 2012 London, UK. 4-6.

AL-BARQAWI, H. & ZAYED, T. 2008. Infrastructure management: integrated AHP/ANN model to evaluate municipal water mains' performance. *Journal of Infrastructure Systems,* 14(4)**,** 305-318.

AL-MUKHTAR, M. & AL-YASEEN, F. 2019. Modeling water quality parameters using data-driven models, a case study Abu-Ziriq marsh in south of Iraq. *Hydrology,* 6(1)**,** 24.

ALLERON, L., MERLET, N., LACOMBE, C. & FRERE, J. 2008. Long-term survival of Legionella pneumophila in the viable but nonculturable state after monochloramine treatment. *Current Microbiology,* 57(5)**,** 497-502.

ALYUDA RESEARCH LLC. 2016. Alyuda Forecaster XL. [software]. http://www.alyuda.com/ [Accessed January 2016].

ANG, W. L., MOHAMMAD, A. W., HILAL, N. & LEO, C. P. 2015. A review on the applicability of integrated/hybrid membrane processes in water treatment and desalination plants. *Desalination,* 363, 2-18.

ANTANASIJEVIĆ, D., POCAJT, V., PERIĆ-GRUJIĆ, A. & RISTIĆ, M. 2019. Multilevel split of high-dimensional water quality data using artificial neural networks for the prediction of dissolved oxygen in the Danube River. *Neural Computing and Applications*.

ANTANASIJEVIC, D., POCAJT, V., POVRENOVIC, D., PERIC-GRUJIC, A. & RISTIC, M. 2013. Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study. *Environmental Science and Pollution Research,* 20(12)**,** 9006-9013.

ARAD, J., HOUSH, M., PERELMAN, L. & OSTFELD, A. 2013. A dynamic thresholds scheme for contaminant event detection in water distribution systems. *Water Research,* 47(5)**,** 1899-1908.

ARANA, I., ORRUNO, M., PEREZ-PASCUAL, D., SECO, C., MUELA, A. & BARCINA, I. 2007. Inability of *Escherichia coli* to resuscitate from the viable but nonculturable state. *FEMS Microbiology Ecology,* 62(1)**,** 1-11.

ASAKURA, H., KAWAMOTO, K., HAISHIMA, Y., IGIMI, S., YAMAMOTO, S. & MAKINO, S. 2008. Differential expression of the outer membrane protein W (OmpW) stress response in enterohemorrhagic *Escherichia coli* O157:H7

corresponds to the viable but non-culturable state. *Research in Microbioly,* 159(9-10)**,** 709-17.

ATEBA, C. N. & MBEWE, M. 2011. Detection of *Escherichia coli* O157:H7 virulence genes in isolates from beef, pork, water, human and animal species in the northwest province, South Africa: public health implications. *Research in Microbioly,* 162(3)**,** 240-8.

AZAD, A., KARAMI, H., FARZIN, S., SAEEDIAN, A., KASHI, H. & SAYYAHI, F. 2018. Prediction of water quality parameters using ANFIS optimized by intelligence algorithms (case study: Gorganrood River). *KSCE Journal of Civil Engineering,* 22(7)**,** 2206-2213.

BADEJO, A. A., NDAMBUKI, J. M., KUPOLATI, W. K., ADEKUNLE, A. A., TAIWO, S. A. & OMOLE, D. O. 2015. Appraisal of access to safe drinking water in southwest Nigeria. *African Journal of Science, Technology, Innovation and Development,* 7(6)**,** 441-445.

BAHGAT, M., SABER, W. & ZAKI, M. 2018. Bacteriological Quality of Water in Meet Khamis Drinking Water Plant, Egypt: Detection of Bacterial Pathogens and Contamination Sources. *Journal of Advances in Microbiology***,** 1-7.

BANU, O. K., GEORGE, B. T. & PAUL, J. 2010. Development of a competitive PCR assay for the quantification of total *Escherichia coli* DNA in water. *African Journal of Biotechnology,* 9(4).

BARGIELA, A. & HAINSWORTH, G. D. 1989. Pressure and flow uncertainty in water systems. *Journal of Water Resources Planning and Management,* 115(2)**,** 212-229.

BARTRAM, J. 2009. *Water safety plan manual: step-by-step risk management for drinking-water suppliers*, World Health Organization.

BAUDOIN, M.-A., VOGEL, C., NORTJE, K. & NAIK, M. 2017. Living with drought in South Africa: lessons learnt from the recent El Niño drought period. *International Journal of Disaster Risk Reduction,* 23, 128-137.

BEUTIN, L., KONG, Q., FENG, L., WANG, Q., KRAUSE, G., LEOMIL, L., JIN, Q. & WANG, L. 2005. Development of PCR assays targeting the genes involved in

synthesis and assembly of the new *Escherichia coli* O 174 and O 177 O antigens. *Journal of Clinical Microbiology,* 43(10)**,** 5143-9.

BEZUIDENHOUT, C. C. 2013. *A Large Scale Study on Microbial and Physico-Chemical Quality of Selected Groundwater and Surface Water in the North-West Province, South Africa.* WRC Report no. 1966/1/13.

BEZUIDENHOUT, J., LIEBENBERG, D., CLAASSENS, S. & VAN RENSBURG, L. 2013. Application of evolutionary algorithms to develop a rule set for assessing the rehabilitation status of asbestos mines in South Africa. *Environmental Earth Sciences,* 70(7)**,** 3267-3275.

BI, W., DANDY, G. C. & MAIER, H. R. 2015. Improved genetic algorithm optimization of water distribution system design by incorporating domain knowledge. *Environmental Modelling & Software,* 69, 370-381.

BIEROZA, M., BAKER, A. & BRIDGEMAN, J. 2011. Classification and calibration of organic matter fluorescence data with multiway analysis methods and artificial neural networks: an operational tool for improved drinking water treatment. *Environmetrics,* 22(3)**,** 256-270.

BIEROZA, M., BAKER, A. & BRIDGEMAN, J. 2012. New data mining and calibration approaches to the assessment of water treatment efficiency. *Advances in Engineering Software,* 44(1)**,** 126-135.

BIOMERIEUX. 2019. API Reference Guide [Online]. Available: https://www.biomerieux-usa.com/sites/subsidiary_us/files/18_api-ref-guide_v7.pdf [Accessed 17 February 2020].

BJERGBAEK, L. A. & ROSLEV, P. 2005. Formation of nonculturable *Escherichia coli* in drinking water. *Journal of Applied Microbiology,* 99(5)**,** 1090-8.

BOCCELLI, D. L., TRYBY, M. E., UBER, J. G., ROSSMAN, L. A., ZIEROLF, M. L. & POLYCARPOU, M. M. 1998. Optimal scheduling of booster disinfection in water distribution systems. *Journal of Water Resources Planning and Management,* 124(2)**,** 99-111.

BOUGADIS, J., ADAMOWSKI, K. & DIDUCH, R. 2005. Short-term municipal water demand forecasting. *Hydrological Processes,* 19(1)**,** 137-148.

BOWDEN, G. J., DANDY, G. C. & MAIER, H. R. 2005. Input determination for neural network models in water resources applications. Part 1 - background and methodology. *Journal of Hydrology,* 301(1-4)**,** 75-92.

BOWDEN, G. J., NIXON, J. B., DANDY, G. C., MAIER, H. R. & HOLMES, M. 2006. Forecasting chlorine residuals in a water distribution system using a general regression neural network. *Mathematical and Computer Modelling,* 44, 469-484.

BOWER, P. A., SCOPEL, C. O., JENSEN, E. T., DEPAS, M. M & MCLELLAN, S. L. 2005. Detection of Genetic Markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods. *Applied and Environmental Microbiology*, 71(12), 8305-8313.

BROOKES, J. D., CAREY, C. C., HAMILTON, D. P., HO, L., VAN DER LINDEN, L., RENNER, R. & RIGOSI, A. 2014. Emerging challenges for the drinking water industry. *Environmental Science & Technology,* 48(4)**,** 2099-2101.

BURCHARD-LEVINE, A., LIU, S., VINCE, F., LI, M. & OSTFELD, A. 2014. A hybrid evolutionary data driven model for river water quality early warning. *Journal of Environmental Management,* 143, 8-16.

CABASSUD, M., DELGRANGE-VINCENT, N., CABASSUD, C., DURAND-BOURLIER, L. & LAINE, J. 2002. Neural networks: a tool to improve UF plant productivity. *Desalination,* 145(1)**,** 223-231.

CALABRESE, J. P. & BISSONNETTE, G. K. 1990. Improved membrane filtration method incorporating catalase and sodium pyruvate for detection of chlorine-stressed coliform bacteria. *Applied and Environmental Microbiology.,* 56(11)**,** 3558-3564.

CASTAGNOLA, E., FIOREDDA, F., BARRETTA, M. A., PESCETTO, L., GARAVENTA, A., LANINO, E., MICALIZZI, C., GIACCHINO, R. & DINI, G. 2001. *Bacillus sphaericus* bacteraemia in children with cancer: case reports and literature review. *Journal of Hospital Infection,* 48(2)**,** 142-5.

CAUCHI, M., BIANCO, L. & BESSANT, C. 2011. The quantification of pollutants in drinking water by use of artificial neural networks. *Natural Computing,* 10(1)**,** 77-90.

CHAYJAN, R. A. 2010. Modeling of sesame seed dehydration energy requirements by a soft-computing approach. *Australian Journal of Crop Science,* 4(3)**,** 180-184.

CHEN, H. & KIM, A. S. 2006. Prediction of permeate flux decline in crossflow membrane filtration of colloidal suspension: a radial basis function neural network approach. *Desalination,* 192(1)**,** 415-428.

CHEN, S., LI, X., WANG, Y., ZENG, J., YE, C., LI, X., GUO, L., ZHANG, S. & YU, X. 2018. Induction of *Escherichia coli* into a VBNC state through chlorination/chloramination and differences in characteristics of the bacterium between states. *Water Research,* 142, 279-288.

CHEN, W. B. & LIU, W. C. 2014. Artificial neural network modeling of dissolved oxygen in reservoir. *Environmental Monitoring Assessment,* 186(2)**,** 1203-1217.

CHRISTODOULOU, S. & DELIGIANNI, A. 2010. A neurofuzzy decision framework for the management of water distribution networks. *Water Resources Management,* 24(1)**,** 139-156.

CLABAUGH, C., MYSZEWSKI, D. & PANG, J. 2000. *Neural Networks* [Online]. Available:https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html [Accessed 11 May 2016].

COBBOLD, R. & DESMARCHELIER, P. 2000. A longitudinal study of Shiga-toxigenic *Escherichia coli* (STEC) prevalence in three Australian dairy herds. *Veterinary microbiology,* 71(1-2)**,** 125-137.

CORBATÓN-BÁGUENA, M.-J., VINCENT-VELA, M.-C., GOZÁLVEZ-ZAFRILLA, J.-M., ÁLVAREZ-BLANCO, S., LORA-GARCÍA, J. & CATALÁN-MARTÍNEZ, D. 2016. Comparison between artificial neural networks and Hermia's models to assess ultrafiltration performance. *Separation and Purification Technology,* 170, 434-444.

CORDOBA, G. A. C., TUHOVČÁK, L. & TAUŠ, M. 2014. Using Artificial Neural Network Models to Assess Water Quality in Water Distribution Networks. *Procedia Engineering,* 70, 399-408.

CULLIS, J. D., ROSSOUW, N., DU TOIT, G., PETRIE, D., WOLFAARDT, G., DE CLERCQ, W. & HORN, A. 2018. Economic risks due to declining water quality in the Breede River catchment. *Water SA,* 44(3)**,** 464-473.

CURCIO, S., CALABRÒ, V. & IORIO, G. 2006. Reduction and control of flux decline in cross-flow membrane processes modeled by artificial neural networks. *Journal of membrane science,* 286(1)**,** 125-132.

DAMIKOUKA, I., KATSIRI, A. & TZIA, C. 2007. Application of HACCP principles in drinking water treatment. *Desalination,* 210(1-3)**,** 138-145.

DAVISON, A., HOWARD, G., STEVENS, M., CALLAN, P., FEWTRELL, L., DEERE, D. & BARTRAM, J. 2005. Water safety plans: managing drinking-water quality from catchment to consumer. Geneva: World Health Organization.

DAWIDOWICZ, J. A. 2018. Method for Estimating the Diameter of Water Pipes Using Artificial Neural Networks of the Multilayer Perceptron Type. 2nd International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2018), Atlantis Press.

DE TRAVERSAY, C. 2006. Challenging drinking water disinfection: How to face up to emerging waterborne pathogens? *Water Practice and Technology,* 1(2).

DEKSISSA, T., ASHTON, P. J. & VANROLLEGHEM, P. A. 2003. Control options for river water quality improvement: A case study of TDS and inorganic nitrogen in the Crocodile River (South Africa). *Water SA,* 29(2)**,** 209-218.

DELPLA, I., FLOREA, M. & RODRIGUEZ, M. J. 2019. Drinking Water Source Monitoring Using Early Warning Systems Based on Data Mining Techniques. *Water Resources Management,* 33(1)**,** 129-140.

DENISOVA, V., MEZULE, L. & JUHNA, T. 2014. The Effect of Chlorination on *Escherichia coli* Viability in Drinking Water. *Material Science and Applied Chemistry,* 30.

DEPARTMENT OF WATER AFFAIRS 2014. The blue drop and no drop handbook. South Africa.

DEPARTMENT OF WATER AFFAIRS AND FORESTRY. 2000. Quality of domestic water supplies. 1st ed. Volume 2: Sampling Guide. Pretoria, South Africa: Government Printer.

DEPARTMENT OF WATER AFFAIRS AND FORESTRY. 2009. INTEGRATED WATER QUALITY MANAGEMENT PLAN FOR THE VAAL RIVER SYSTEM: Task 8: Water Quality Management Strategy for the Vaal River System.

DEPARTMENT OF WATER AND SANITATION 2014. FACTSHEET: The Blue Drop: Highlights and Trends from 2009 to 2014. Gezina: Water Research Commission.

DHARMAN, S., CHANDRAMOULI, V. & LINGIREDDY, S. 2012. Predicting total organic carbon removal efficiency and coagulation dosage using artificial neural networks. *Environmental Engineering Science,* 29(8)**,** 743-750.

DING, T., SUO, Y., XIANG, Q., ZHAO, X., CHEN, S., YE, X. & LIU, D. 2017. Significance of Viable but Nonculturable *Escherichia coli*: Induction, Detection, and Control. *Journal of Microbiology and Biotechnology,* 27(3)**,** 417-428.

DOUTERELO, I., SHARPE, R. L., HUSBAND, S., FISH, K. E. & BOXALL, J. B. 2019. Understanding microbial ecology to improve management of drinking water distribution systems. *Wiley Interdisciplinary Reviews: Water,* 6(1)**,** e01325.

DUFOUR, A. P., STRICKLAND, E. R. & CABELLI, V. J. 1981. Membrane filter method for enumerating *Escherichia coli*. *Applied and Environmental Microbiology.,* 41(5)**,** 1152-1158.

DUNGENI, M., VAN DER MERWE, R. & MOMBA, M. 2010. Abundance of pathogenic bacteria and viral indicators in chlorinated effluents produced by four wastewater treatment plants in the Gauteng Province, South Africa. *Water SA,* 36(5).

DUNN, G., HARRIS, L., COOK, C. & PRYSTAJECKY, N. 2014. A comparative analysis of current microbial water quality risk assessment and management practices in British Columbia and Ontario, Canada. *Science of the Total Environment,* 468-469(544-52.

ERCUMEN, A., ARNOLD, B. F., NASER, A. M., UNICOMB, L., COLFORD, J. M., JR. & LUBY, S. P. 2017. Potential sources of bias in the use of *Escherichia coli* to

measure waterborne diarrhoea risk in low-income settings. *Tropical Medicine & International Health,* 22(1)**,** 2-11.

FARMAKI, E. G., THOMAIDIS, N. S. & EFSTATHIOU, C. E. 2010. Artificial Neural Networks in water analysis: Theory and applications. *International Journal of Environmental Analytical Chemistry,* 90(2)**,** 85-105.

FARMAKI, E. G., THOMAIDIS, N. S., SIMEONOV, V. & EFSTATHIOU, C. E. 2013. Comparative use of artificial neural networks for the quality assessment of the water reservoirs of Athens. *Journal of Water Supply: Research and Technology-AQUA,* 62(5)**,** 296-308.

FAROKHZAD, S., AHMADI, H., JAEFARI, A., ABAD, M. R. A. A. & KOHAN, M. R. 2012. Artificial neural network based classification of faults in centrifugal water pump. *Journal of Vibroengineering,* 14(4)**,** 1734-1744.

FARROKH, C., JORDAN, K., AUVRAY, F., GLASS, K., OPPEGAARD, H., RAYNAUD, S., THEVENOT, D., CONDRON, R., DE REU, K., GOVARIS, A., HEGGUM, K., HEYNDRICKX, M., HUMMERJOHANN, J., LINDSAY, D., MISZCZYCHA, S., MOUSSIEGT, S., VERSTRAETE, K. & CERF, O. 2012. Review of Shiga-toxin-producing *Escherichia coli* (STEC) and their significance in dairy production. *International Journal of Food Microbiology,* 162(2)**,** 190-212.

FIGUEIRA, V., VAZ-MOREIRA, I., SILVA, M. & MANAIA, C. M. 2011. Diversity and antibiotic resistance of Aeromonas spp. in drinking and waste water treatment plants. *Water Research,* 45(17)**,** 5599-5611.

FIRAT, M., TURAN, M. E. & YURDUSEV, M. A. 2010. Comparative analysis of neural network techniques for predicting water consumption time series. *Journal of Hydrology,* 384(1-2)**,** 46-51.

FIRAT, M., YURDUSEV, M. A. & TURAN, M. E. 2009. Evaluation of artificial neural network techniques for municipal water consumption modeling. *Water Resources Management,* 23(4)**,** 617-632.

FORSTINUS, N., IKECHUKWU, N., EMENIKE, M. & CHRISTIANA, A. 2016. Water and Waterborne Diseases: A Review. *International Journal of TROPICAL DISEASE & Health,* 12(4)**,** 1-14.

GAO, W., LIANG, H., MA, J., HAN, M., CHEN, Z.-L., HAN, Z.-S. & LI, G.-B. 2011. Membrane fouling control in ultrafiltration technology for drinking water production: a review. *Desalination,* 272(1)**,** 1-8.

GAYA, M., YUSUF, L., MUSTAPHA, M., MUHAMMAD, B., SANI, A., TIJJANI, A. T. A., WAHAB, N. & KHAIRI, M. 2017. Estimation of Turbidity in Water Treatment Plant using Hammerstein-Wiener and Neural Network Technique. *Indonesian Journal of Electrical Engineering and Computer Science,* 5(3)**,** 666-672.

GELDREICH, E. E., CLARK, H. F., HUH, C. B. & BEST, L. C. 1965. Fecal-coliform-organism medium for the membrane filter technique. *Journal-American Water Works Association,* 57(2)**,** 208-214.

GHOLIKANDI, G. B., DELNAVAZ, M. & RIAHI, R. 2011. Use of Artificial Neural Network for prediction of coagulation/flocculation process by PAC in water treatment plant. *Environmental Engineering & Management Journal (EEMJ),* 10(11).

GIBBS, M., MORGAN, N., MAIER, H., DANDY, G., HOLMES, M. & NIXON, J. 2003. Use of artificial neural networks for modelling chlorine residuals in water distribution systems. MODSIM 2003: Proceedings of the 2003 International Congress on Modelling and Simulation, Townsville, Australia, Citeseer, 789-794.

GIBBS, M. S., MORGAN, N., MAIER, H. R., DANDY, G. C., NIXON, J. B. & HOLMES, M. 2006. Investigation into the relationship between chlorine decay and water distribution parameters using data driven methods. *Mathematical and Computer Modelling,* 44(5)**,** 485-498.

GODFREY, S. & HOWARD, G. 2004. *Water Safety Plans (WSP) for Urban Piped Water Supplies in Developing Countries,* Leicestershire UK, Water, Engineering and Development Centre.

GOWDA, T. K., REDDY, V. R., DEVLEESSCHAUWER, B., ZADE, N. N., CHAUDHARI, S. P., KHAN, W. A., SHINDE, S. V. & PATIL, A. R. 2015. Isolation and Seroprevalence of Aeromonas spp. Among Common Food Animals Slaughtered in Nagpur, Central India. *Foodborne Pathogens and Diseases,* 12(7)**,** 626-30.

GOYAL, S. & GOYAL, G. K. 2011. Cascade and feedforward backpropagation artificial neural networks models for prediction of sensory quality of instant coffee flavoured sterilized drink. *Canadian Journal on Artificial Intelligence, Machine Learning and Pattern Recognition,* 2(6)**,** 78-82.

GRABOW, W., HILNER, C. & COUBROUGH, P. 1981. Evaluation of standard and modified M-FC, MacConkey, and Teepol media for membrane filtration counting of fecal coliforms in water. *Applied and Environmental Microbiology.,* 42(2)**,** 192-199.

GRAUPE, D. 2013. *Principles of Artificial Neural Networks*, World Scientific.

GRAY, A. & VAWDA, Y. 2016. Health Policy and Legislation. *In:* PADARATH A, K. J., MACKIE E, CASCIOLA J (ed.) *South African Health Review 2016.* Durban: Health Systems Trust.

GRIFFITHS, K. & ANDREWS, R. 2011. Application of artificial neural networks for filtration optimization. *Journal of Environmental Engineering,* 137(11)**,** 1040-1047.

GUNNARSDÓTTIR, M. J. & GISSURARSON, L. R. 2008. HACCP and water safety plans in Icelandic water supply: Preliminary evaluation of experience. *Journal of Water and Health,* 6(3)**,** 377.

HAMILTON, P. D., GALE, P. & POLLARD, S. J. 2006. A commentary on recent water safety initiatives in the context of water utility risk management. *Environment international,* 32(8)**,** 958-966.

HANNAN, S. A., MANZA, R. & RAMTEKE, R. 2010. Generalized regression neural network and radial basis function for heart disease diagnosis. *International Journal of Computer Applications,* 7(13)**,** 7-13.

HAYKIN, S. 2009. *Neural Networks and Learning Machines*, Pearson Prentice Hall, New Jersey, USA.

HEBB, D. O. 1949. *The Organization of Behavior: A Neuropsychological Theory,* New York, John Wiley & Sons, Inc.

HEDDAM, S., BERMAD, A. & DECHEMI, N. 2011. Applications of radial-basis function and generalized regression neural networks for modeling of coagulant dosage

in a drinking water-treatment plant: Comparative study. *Journal of Environmental Engineering,* 137(12)**,** 1209-1214.

HEDDAM, S., BERMAD, A. & DECHEMI, N. 2012. ANFIS-based modelling for coagulant dosage in drinking water treatment plant: a case study. *Environmental Monitoring and Assessment,* 184(4)**,** 1953-1971.

HO, C. I., LIN, M. D. & LO, S. L. 2010. Use of a GIS-based hybrid artificial neural network to prioritize the order of pipe replacement in a water distribution network. *Environmental Monitoring Assessment,* 166(1-4)**,** 177-89.

HOPFIELD, J. J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences,* 79(8)**,** 2554-2558.

HOUTMAN, C. J. 2010. Emerging contaminants in surface waters and their relevance for the production of drinking water in Europe. *Journal of Integrative Environmental Sciences,* 7(4)**,** 271-295.

HRUDEY, S. E. & HRUDEY, E. J. 2014. *Ensuring safe drinking water: learning from frontline experience with contamination,* Denver, Colorado, American Water Works Association.

HU, J., CHU, W., SUI, M., XU, B., GAO, N. & DING, S. 2018. Comparison of drinking water treatment processes combinations for the minimization of subsequent disinfection by-products formation during chlorination and chloramination. *Chemical Engineering Journal,* 335(352-361.

HUANG, P., ZHU, N., HOU, D., CHEN, J., XIAO, Y., YU, J., ZHANG, G. & ZHANG, H. 2018. Real-Time Burst Detection in District Metering Areas in Water Distribution System Based on Patterns of Water Demand with Supervised Learning. *Water,* 10(12).

HUGHES, D. A., DEMUTH, S., GUSTARD, A., PLANOS, E., SEATENA, F. & SERVAT, E. 2006. An evaluation of the potential use of satellite rainfall data for input to water resource estimation models in southern Africa. *International Association of Hydrological Sciences,* 308(75-80.

HUNTER, P. R. 2003. Drinking water and diarrhoeal disease due to *Escherichia coli*. *Journal of Water and Health,* 1(2)**,** 65-72.

IGLESIAS, C., MARTÍNEZ TORRES, J., GARCÍA NIETO, P. J., ALONSO FERNÁNDEZ, J. R., DÍAZ MUÑIZ, C., PIÑEIRO, J. I. & TABOADA, J. 2014. Turbidity prediction in a river basin by using artificial neural networks: a case study in northern Spain. *Water Resources Management,* 28(2)**,** 319-331.

IGWARAN, A., IWERIEBOR, B. & OKOH, A. 2018. Molecular Characterization and Antimicrobial Resistance Pattern of *Escherichia coli* Recovered from Wastewater Treatment Plants in Eastern Cape South Africa. *International Journal of Environmental Research and Public Health,* 15(6)**,** 1237.

ILUNGA, M. & STEPHENSON, D. 2005. Infilling streamflow data using feed-forward back-propagation (BP) artificial neural networks: application of standard BP and Pseudo Mac Laurin power series BP techniques. *Water SA,* 31(2)**,** 171-176.

JAFAR, R., SHAHROUR, I. & JURAN, I. 2010. Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. *Mathematical and Computer Modelling,* 51(9-10)**,** 1170-1180.

JAGALS, C. & JAGALS, P. 2004. Application of HACCP principles as a management tool for monitoring and controlling microbiological hazards in water treatment facilities. *Water Science and Technology,* 50(1)**,** 69-76.

JANDA, J. M. & ABBOTT, S. L. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *Journal of clinical microbiology,* 45(9)**,** 2761-2764.

JANDU, N., HO, N. K., DONATO, K. A., KARMALI, M. A., MASCARENHAS, M., DUFFY, S. P., TAILOR, C. & SHERMAN, P. M. 2009. Enterohemorrhagic *Escherichia coli* O157:H7 gene expression profiling in response to growth in the presence of host epithelia. *PLoS One,* 4(3)**,** e4889.

JANEZIC, K. J., FERRY, B., HENDRICKS, E. W., JANIGA, B. A., JOHNSON, T., MURPHY, S., ROBERTS, M. E., SCOTT, S. M., THEISEN, A. N. & HUNG, K. F. 2013. Phenotypic and genotypic characterization of *Escherichia coli* isolated from untreated surface waters. *The Open Microbiology Journal,* 7(9).

JANG, D., PARK, H. & CHOI, G. 2018. Estimation of leakage ratio using principal component analysis and artificial neural network in water distribution systems. *Sustainability,* 10(3)**,** 750.

JORDAAN, K. & BEZUIDENHOUT, C. C. 2013. The impact of physico-chemical water quality parameters on bacterial diversity in the Vaal River, South Africa. *Water SA,* 39(3).

JUNTUNEN, P., LIUKKONEN, M., LEHTOLA, M. & HILTUNEN, Y. 2013. Cluster analysis by self-organizing maps: an application to the modelling of water quality in a treatment process. *Applied Soft Computing,* 13(7)**,** 3191-3196.

KABSCH-KORBUTOWICZ, M. & KUTYLOWSKA, M. 2011. Use of artificial intelligence in predicting the turbidity retention coefficient during ultrafiltration of water. *Environment Protection Engineering,* 37(2)**,** 75-84.

KAGODA, P. A., NDIRITU, J., NTULI, C. & MWAKA, B. 2010. Application of radial basis function neural networks to short-term streamflow forecasting. *Physics and Chemistry of the Earth,* 35(13-14)**,** 571-581.

KAMIŃSKI, K., KAMIŃSKI, W. & MIZERSKI, T. 2017. Application of artificial neural networks to the technical condition assessment of water supply systems. *Ecological Chemistry and Engineering S,* 24(1)**,** 31-40.

KARADURMUŞ, E., TAŞKıN, N., GÖZ, E. & YÜCEER, M. 2018. Prediction of Bromate Removal in Drinking Water Using Artificial Neural Networks. *Ozone: Science & Engineering,* 41(2)**,** 118-127.

KATAMBARA, Z. & NDIRITU, J. 2009. A fuzzy inference system for modelling streamflow: case of Letaba River, South Africa. *Physics and Chemistry of the Earth,* 34(10)**,** 688-700.

KAZEMI, E., MOUNCE, S., HUSBAND, S. & BOXALL, J. 2018. Predicting turbidity in water distribution trunk mains using nonlinear autoregressive exogenous artificial neural networks. *Palermo, Sicily, 1st to 6th July*.

KHANIKI, G. R. J., MAHDAVI, M. & MOHEBBI, M. R. 2009. HACCP application for treatment of drinking water for Germi in Iran. *Journal of food, agriculture & environment,* 7(2)**,** 709-712.

KIM, D.-K., JEONG, K.-S., MCKAY, R. I. B., CHON, T.-S., KIM, H.-W. & JOO, G.-J. 2010. Model development in freshwater ecology with a case study using evolutionary computation. *Journal of Ecology and Environment,* 33(4)**,** 275-288.

KIM, H., KIM, S. & KOO, J. 2014. Prediction of chlorine concentration in various hydraulic conditions for a pilot scale water distribution system. *Procedia Engineering,* 70(934-942.

KINGE, C. W. & MBEWE, M. 2012. Bacterial contamination levels in river catchments of the North West Province, South Africa: Public health implications. *African Journal of Microbiology Research,* 6(7)**,** 1370-1375.

KOHONEN, T. 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics,* 43(1)**,** 59-69.

KOHONEN, T., OJA, E., SIMULA, O., VISA, A. & KANGAS, J. 1996. Engineering applications of the self-organizing map. *Proceedings of the IEEE,* 84(10)**,** 1358-1384.

KULKARNI, P. & 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)**,** 4202-10.

KUMAR, S., STECHER, G., LI, M., KNYAZ, C. & TAMURA, K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution,* 35(6)**,** 1547-1549.

KUREK, W. & BRDYS, M. A. 2007. Genetic solver of optimization task of MPC for optimizing control of integrated quantity and quality in drinking water distribution systems. *IFAC Proceedings Volumes,* 40(9)**,** 292-297.

LAMRINI, B., BENHAMMOU, A., LE LANN, M.-V. & KARAMA, A. 2005. A neural software sensor for online prediction of coagulant dosage in a drinking water treatment plant. *Transactions of the Institute of Measurement and Control,* 27(3)**,** 195-213.

LANE, D. J. 1991. 16S/23S rRNA sequencing. In: Stackebrandt, E., Goodfellow, M. (Eds.), Nucleic Acid Techniques in Bacterial Systematics. John Wiley & Sons, New York, pp. 115-175.

LECHEVALLIER, M. W., CAMERON, S. C. & MCFETERS, G. A. 1983. New medium for improved recovery of coliform bacteria from drinking water. *Applied and Environmental Microbiology,* 45(2)**,** 484-492.

LEGUBE, B., PARINET, B., GELINET, K., BERNE, F. & CROUE, J. P. 2004. Modeling of bromate formation by ozonation of surface waters in drinking water treatment. *Water Research,* 38(8)**,** 2185-95.

LEÓN-LUQUE, A. J., BARAJAS, C. L. & PEÑA-GUZMÁN, C. A. 2016. Determination of the optimal dosage of aluminum sulfate in the coagulation-flocculation process using an artificial neural network. *International Journal of Environmental Science and Development,* 7(5)**,** 346-350.

LEPŠ, J. & ŠMILAUER, P. 2003. *Multivariate analysis of ecological data using CANOCO*, Cambridge university press.

LERMONTOV, A., YOKOYAMA, L., LERMONTOV, M. & MACHADO, M. A. S. 2009. River quality analysis using fuzzy water quality index: Ribeira do Iguape river watershed, Brazil. *Ecological Indicators,* 9(6)**,** 1188-1197.

LI, W., CUI, L., ZHANG, Y., CAI, Z., ZHANG, M., XU, W., ZHAO, X., LEI, Y., PAN, X., LI, J. & DOU, Z. 2018. Using a Backpropagation Artificial Neural Network to Predict Nutrient Removal in Tidal Flow Constructed Wetlands. *Water,* 10(1).

LIOFILCHEM. 2020. EnteroPluri-Test [Online]. Available: https://www.liofilchem.com/en/products/featured-products/enteropluri-test.html [Accessed 17 February 2020].

LIN, H., YE, C., CHEN, S., ZHANG, S. & YU, X. 2017. Viable but non-culturable *E. coli* induced by low level chlorination have higher persistence to antibiotics than their culturable counterparts. *Environmental Pollution,* 230(242-249.

LIU, S., TAI, H., DING, Q., LI, D., XU, L. & WEI, Y. 2013. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. *Mathematical and Computer Modelling,* 58(3-4)**,** 458-465.

LIU, Y., WANG, C., TYRRELL, G., HRUDEY, S. E. & LI, X. F. 2009. Induction of *Escherichia coli* O157:H7 into the viable but non-culturable state by chloraminated water and river water, and subsequent resuscitation. *Environmental Microbiology Reports,* 1(2)**,** 155-61.

LIU, Y., WANG, C., TYRRELL, G. & LI, X. F. 2010. Production of Shiga-like toxins in viable but nonculturable *Escherichia coli* O157:H7. *Water Research,* 44(3)**,** 711-8.

LUSILAO-MAKIESE, J., TESSIER, E., AMOUROUX, D., TUTU, H., CHIMUKA, L., WEIERSBYE, I. & CUKROWSKA, E. 2014. Seasonal distribution and speciation of mercury in a gold mining area, north-west province, South Africa. *Toxicological & Environmental Chemistry,* 96(3)**,** 387-402.

MADAENI, S., SHIRI, M. & KURDIAN, A. 2015. Modeling, optimization, and control of reverse osmosis water treatment in Kazeroon power plant using neural network. *Chemical Engineering Communications,* 202(1)**,** 6-14.

MAGNIER, L. & HAGHIGHAT, F. 2010. Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network. *Building and Environment,* 45(3)**,** 739-746.

MAHESH, V. & PRASANTH, S. 2015. Occurrence of pollution indicators in tropical perennial river of Periyar, Southern India. *International Journal of Pharmacological Research,* 5(10)**,** 264.

MAIER, H. R. & DANDY, G. C. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software,* 15(1)**,** 101-124.

MAKAYA, E. & HENSEL, O. 2015. Modelling flow dynamics in water distribution networks using artificial neural networks - a leakage detection technique. *International Journal of Engineering, Science and Technology,* 7(1)**,** 33-43.

MANNING, S. D., MOTIWALA, A. S., SPRINGMAN, A. C., QI, W., LACHER, D. W., OUELLETTE, L. M., MLADONICKY, J. M., SOMSEL, P., RUDRIK, J. T. & DIETRICH, S. E. 2008. Variation in virulence among clades of *Escherichia coli* O157: H7 associated with disease outbreaks. *Proceedings of the National Academy of Sciences,* 105(12)**,** 4868-4873.

MARTEL, K., KIRMEYER, G., HANSON, A., MULLENGER, J. & STEVENS, M. 2006. Application of HACCP for distribution system protection. American Water Works Association.

MARTÍNEZ, F., HERNÁNDEZ, V., ALONSO, J. M., RAO, Z. & ALVISI, S. 2007. Optimizing the operation of the Valencia water-distribution network. *Journal of Hydroinformatics,* 9(1)**,** 65-78.

MARZOUK, M. & ELKADI, M. 2016. Estimating water treatment plants costs using factor analysis and artificial neural networks. *Journal of Cleaner Production,* 112(Part 5)**,** 4540-4549.

MAY, R. J., DANDY, G. C., MAIER, H. R. & NIXON, J. B. 2008. Application of partial mutual information variable selection to ANN forecasting of water quality in water distribution systems. *Environmental Modelling & Software,* 23(10-11)**,** 1289-1299.

MCCLYMONT, K., KEEDWELL, E. & SAVIC, D. 2015. An analysis of the interface between evolutionary algorithm operators and problem features for water resources problems. A case study in water distribution network design. *Environmental Modelling & Software,* 69, 414-424.

MCCULLOCH, W. S. & PITTS, W. 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics,* 5(4)**,** 115-133.

MENG, S., GREENLEE, L. F., SHEN, Y. R. & WANG, E. 2015. Basic science of water: challenges and current status towards a molecular picture. *Nano Research,* 8(10)**,** 3085-3110.

MILOT, J., RODRIGUEZ, M. J. & SÉRODES, J. B. 2002. Contribution of neural networks for modeling trihalomethanes occurrence in drinking water. *Journal of Water Resources Planning and Management,* 128(5)**,** 370-376.

MIYAGI, K., SANO, K. & HIRAI, I. 2017. Sanitary evaluation of domestic water supply facilities with storage tanks and detection of Aeromonas, enteric and related bacteria in domestic water facilities in Okinawa Prefecture of Japan. *Water Research,* 119(171-177.

MOMBA, M. & BROUCKAERT, B. 2005. *Guidelines for ensuring sustainable effective disinfection in small water supply systems*, Water Research Commission Pretoria.

MOMBA, M., TYAFA, Z. & 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. *South African Journal of Science,* 100(5-6)**,** 307-310.

MONTSO, P. K., MLAMBO, V. & ATEBA, C. N. 2019. The First Isolation and Molecular Characterization of Shiga Toxin-Producing Virulent Multi-Drug Resistant Atypical Enteropathogenic *Escherichia coli* O177 Serogroup From South African Cattle. *Front Cellular Infection Microbiology,* 9, 333.

MORENO, Y., PIQUERES, P., ALONSO, J. L., JIMENEZ, A., GONZALEZ, A. & FERRUS, M. A. 2007. Survival and viability of *Helicobacter pylori* after inoculation into chlorinated drinking water. *Water Research,* 41(15)**,** 3490-6.

MORRISON, S., VENTER, A. & BARNARD, S. 2012. A case study to determine the efficacy of ozonation in purification processes. *Water SA,* 38(1)**,** 49-54.

MOUNCE, S. R. & MACHELL, J. 2006. Burst detection using hydraulic data from water distribution systems with artificial neural networks. *Urban Water Journal,* 3(1)**,** 21-31.

MSIZA, I. S., NELWAMONDO, F. V. & MARWALA, T. 2007. Artificial neural networks and support vector machines for water demand time series forecasting. 2007 IEEE International Conference on Systems, Man and Cybernetics, IEEE, 638-643.

MUKHERJEE, A. 1997. Self-organizing neural network for identification of natural modes. *Journal of Computing in Civil Engineering,* 11(1)**,** 74-77.

MURTHY, Z. & VENGAL, J. C. 2006. Optimization of a reverse osmosis system using genetic algorithm. *Separation Science and Technology,* 41(4)**,** 647-663.

MUSTONEN, S. M., TISSARI, S., HUIKKO, L., KOLEHMAINEN, M., LEHTOLA, M. J. & HIRVONEN, A. 2008. Evaluating online data of water quality changes in a pilot drinking water distribution system with multivariate data exploration methods. *Water Research,* 42(10)**,** 2421-2430.

MYERS, D. N. 2019. *Why monitor water quality?* [Online]. Available: https://water.usgs.gov/owq/WhyMonitorWaterQuality.pdf [Accessed 24 November 2019].

NAIDOO, P. & VAN DER WALT, J. J. 2013. Artificial neural networks as a chemical dosing budgeting tool. *Institute of Municipal Engineering of Southern Africa (IMIESA),* 38(10)**,** 152.

NAJAH, A., EL-SHAFIE, A., KARIM, O. A. & EL-SHAFIE, A. H. 2013. Application of artificial neural networks for water quality prediction. *Neural Computing and Applications,* 22(S1)**,** 187-201.

NAVAB-DANESHMAND, T., FRIEDRICH, M. N. D., GACHTER, M., MONTEALEGRE, M. C., MLAMBO, L. S., NHIWATIWA, T., MOSLER, H. J. & JULIAN, T. R. 2018. *Escherichia coli* Contamination across Multiple Environmental Compartments (Soil, Hands, Drinking Water, and Handwashing Water) in Urban Harare: Correlations and Risk Factors. *Am J Trop Med Hyg,* 98(3)**,** 803-813.

NAZIF, S., KARAMOUZ, M., TABESH, M. & MORIDI, A. 2010. Pressure management model for urban water distribution networks. *Water Resources Management,* 24(3)**,** 437-458.

NICKLOW, J., REED, P., SAVIC, D., DESSALEGNE, T., HARRELL, L., CHAN-HILTON, A., KARAMOUZ, M., MINSKER, B., OSTFELD, A. & 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.

NKUNA, T. & ODIYO, J. 2011. Filling of missing rainfall data in Luvuvhu River Catchment using artificial neural networks. *Physics and Chemistry of the Earth,* 36(14-15)**,** 830-835.

NWABOR, O. F., NNAMONU, E., MARTINS, P. & ANI, O. 2016. Water and waterborne diseases: a review. *International Journal of Tropical Diseases & Health,* 12(4)**,** 1-14.

NYENJE, P. M., FOPPEN, J. W., UHLENBROOK, S., KULABAKO, R. & MUWANGA, A. 2010. Eutrophication and nutrient release in urban areas of sub-Saharan Africa--a review. *Science of the Total Environment,* 408(3)**,** 447-55.

O'REILLY, G., BEZUIDENHOUT, C. & BEZUIDENHOUT, J. 2018. Artificial neural networks: applications in the drinking water sector. *Water Science and Technology: Water Supply,* 18(6)**,** 1869-1887.

OKEYO, A., MOMBA, M. & COETZEE, M. 2011. Application of the HACCP concept for the microbiological monitoring of drinking water quality: a case study of three

water treatment plants in the Gauteng Province, South Africa. *Trends in Applied Sciences Research,* 6(3)**,** 269-281.

OLIVER, J. D. 2016. The viable but nonculturable state for bacteria: status update. *Microbe,* 11(4)**,** 159-164.

OMAR, K. & BARNARD, T. 2010. The occurrence of pathogenic *Escherichia coli* in South African wastewater treatment plants as detected by multiplex PCR. *Water SA,* 36(2)**,** 172-176.

ONTUMBI, G., NYABERO, C. & SANG, C. C. 2015. An Assessment of Water Quality Variation On Human Health in the River Sosiani Catchment, Kenya. *International Journal of Innovation and Research in Educational Sciences,* 2(2)**,** 2349-5219.

ONYARI, E. & ILUNGA, F. 2013. Application of MLP neural network and M5P model tree in predicting streamflow: a case study of Luvuvhu catchment, South Africa. *International Journal of Innovation, Management and Technology,* 4(1)**,** 11-15.

ORAM, B. 2014. *The pH of Water* [Online]. Available: https://water-research.net/index.php/water-treatment/tools/the-ph-of-water [Accessed 10 December 2019].

OWOSENI, M., OLANIRAN, A. & OKOH, A. 2017. Chlorine Tolerance and Inactivation of *Escherichia coli* recovered from Wastewater Treatment Plants in the Eastern Cape, South Africa. *Applied Sciences,* 7(8).

OYEBODE, O., ADEYEMO, J. & OTIENO, F. 2015. Comparison of two data-driven modelling techniques for long-term streamflow prediction using limited datasets. *Journal of the South African Institution of Civil Engineering,* 57(3)**,** 9-19.

ÖZKANCA, R., SARIBIYIK, F., ISIK, K., SAHIN, N., KARIPTAS, E. & FLINT, K. P. 2009. Resuscitation and quantification of stressed *Escherichia coli* K12 NCTC8797 in water samples. *Microbiological Research,* 164(2)**,** 212-20.

PACCHIN, E., GAGLIARDI, F., ALVISI, S. & FRANCHINI, M. 2019. A Comparison of Short-Term Water Demand Forecasting Models. *Water Resources Management,* 33(4)**,** 1481-1497.

PAGEL, J. E., QURESHI, A., YOUNG, D. & VLASSOFF, L. 1982. Comparison of four membrane filter methods for fecal coliform enumeration. *Applied and Environmental Microbiology.,* 43(4)**,** 787-793.

PALANI, S., LIONG, S.-Y. & TKALICH, P. 2008. An ANN application for water quality forecasting. *Marine Pollution Bulletin,* 56(9)**,** 1586-1597.

PARK, S., COTTER, P.A. & GUNSALUS, R.P. 1995. Regulation of malate dehydrogenase (*mdh*) gene expression in *Escherichia coli* in response to oxygen, carbon, and heme availability. *Journal of Bacteriology*, 177(22), 6652-6656.

PARKER, J. L. & SHAW, J. G. 2011. Aeromonas spp. clinical microbiology and disease. *Journal of Infection,* 62(2)**,** 109-18.

PARROTT, L. 2011. Hybrid modelling of complex ecological systems for decision support: Recent successes and future perspectives. *Ecological Informatics,* 6(1)**,** 44-49.

PERELMAN, L., ARAD, J., HOUSH, M. & OSTFELD, A. 2012. Event detection in water distribution systems from multivariate water quality time series. *Environmental Science & Technology,* 46(15)**,** 8212-8219.

PIENAAR, J. A., SINGH, A. & BARNARD, T. G. 2016. The viable but non-culturable state in pathogenic *Escherichia coli*: A general review. *African Journal of Laboratory Medicine,* 5(1)**,** 1-9.

PINTO, A. J., XI, C. & RASKIN, L. 2012. Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environmental Science & Technology,* 46(16)**,** 8851-8859.

PINTO, D., ALMEIDA, V., ALMEIDA SANTOS, M. & CHAMBEL, L. 2011. Resuscitation of *Escherichia coli* VBNC cells depends on a variety of environmental or chemical stimuli. *Journal of Applied Microbiology,* 110(6)**,** 1601-11.

PINTO, D., SANTOS, M. A. & CHAMBEL, L. 2015. Thirty years of viable but nonculturable state research: unsolved molecular mechanisms. *Critical Review in Microbiology,* 41(1)**,** 61-76.

PIOTROWSKI, A. P., NAPIORKOWSKI, M. J., NAPIORKOWSKI, J. J. & OSUCH, M. 2015. Comparing various artificial neural network types for water temperature prediction in rivers. *Journal of Hydrology,* 529(Part 1)**,** 302-315.

RAJKHOWA, S., SCARIA, J., GARCIA, D. L., MUSSER, K. A., AKEY, B. L. & CHANG, Y.-F. 2010. Analysis of *Escherichia coli* O157 clinical isolates by multilocus sequence typing. *BMC Research Notes,* 3(1)**,** 343.

RAK, A. 2013. Water turbidity modelling during water treatment processes using artificial neural networks. *International Journal of Water Sciences,* 2(3)**,** 1-10.

RAM, S. & SHANKER, R. 2005. Computing TaqMan probes for multiplex PCR detection of *E. coli* O157 serotypes in water. *In silico biology,* 5(5, 6)**,** 499-504.

RANGEL, J. M., SPARLING, P. H., CROWE, C., GRIFFIN, P. M. & SWERDLOW, D. L. 2005. Epidemiology of *Escherichia coli* O157: H7 outbreaks, united states, 1982–2002. *Emerging infectious diseases,* 11(4)**,** 603.

RANI, D. & MOREIRA, M. M. 2010. Simulation–optimization modeling: a survey and potential application in reservoir systems operation. *Water Resources Management,* 24(6)**,** 1107-1138.

RANKOVIĆ, V., RADULOVIĆ, J., RADOJEVIĆ, I., OSTOJIĆ, A. & ČOMIĆ, L. 2010. Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia. *Ecological Modelling,* 221(8)**,** 1239-1244.

RAO, Z. & ALVARRUIZ, F. 2007. Use of an artificial neural network to capture the domain knowledge of a conventional hydraulic simulation model. *Journal of Hydroinformatics,* 9(1)**,** 15-24.

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

RECKNAGEL, F., VAN GINKEL, C., CAO, H., CETIN, L. & ZHANG, B. 2008. Generic limnological models on the touchstone: Testing the lake simulation library SALMO-OO and the rule-based Microcystis agent for warm-monomictic hypertrophic lakes in South Africa. *Ecological Modelling,* 215(1-3)**,** 144-158.

REISSBRODT, R., RIENAECKER, I., ROMANOVA, J. M., FREESTONE, P. P., HAIGH, R. D., LYTE, M., TSCHAPE, H. & WILLIAMS, P. H. 2002. Resuscitation

of Salmonella enterica serovar typhimurium and enterohemorrhagic *Escherichia coli* from the viable but nonculturable state by heat-stable enterobacterial autoinducer. *Applied and Environmental Microbiology,* 68(10)**,** 4788-94.

RIDOLFI, E., SERVILI, F., MAGINI, R., NAPOLITANO, F., RUSSO, F. & ALFONSO, L. 2014. Artificial neural networks and entropy-based methods to determine pressure distribution in water distribution systems. *Procedia Engineering,* 89, 648-655.

RIGOBELLO, E. S., DANTAS, A. D., DI BERNARDO, L. & VIEIRA, E. M. 2013. Removal of diclofenac by conventional drinking water treatment processes and granular activated carbon filtration. *Chemosphere,* 92(2)**,** 184-91.

ROSZAK, D., GRIMES, D. & COLWELL, R. 1984. Viable but nonrecoverable stage of *Salmonella enteritidis* in aquatic systems. *Canadian Journal of Microbiology,* 30(3)**,** 334-338.

RUMELHART, D. E. & MCCLELLAND, J. L. 1986a. *Parallel distributed processing: explorations in the microstructure of cognition. Volume 1: Foundations,* Cambridge, Massachusetts, The MIT Press.

RUMELHART, D. E. & MCCLELLAND, J. L. 1986b. *Parallel distributed processing: explorations in the microstructure of cognition. Volume 2: Psycological and biological models,* Cambridge, Massachusetts, The MIT Press.

SAHA, A. K., CHOUDHURY, S. & MAJUMDER, M. 2017. Performance efficiency analysis of water treatment plants by using MCDM and neural network model. *MATTER: International Journal of Science and Technology,* 3(1)**,** 27-35.

SALAMI SHAHID, E. & EHTESHAMI, M. 2016. Application of artificial neural networks to estimating DO and salinity in San Joaquin River basin. *Desalination and Water Treatment,* 57(11)**,** 4888-4897.

SALARI, M., SALAMI SHAHID, E., AFZALI, S. H., EHTESHAMI, M., CONTI, G. O., DERAKHSHAN, Z. & SHEIBANI, S. N. 2018. Quality assessment and artificial neural networks modeling for characterization of chemical and physical parameters of potable water. *Food and Chemical Toxicology,* 118(212-219.

SALOMONS, E., GORYASHKO, A., SHAMIR, U., RAO, Z. & ALVISI, S. 2007. Optimizing the operation of the Haifa-A water-distribution network. *Journal of Hydroinformatics,* 9(1)**,** 51-64.

SANS: SOUTH AFRICAN NATIONAL STANDARDS 241:2015. Drinking water. Part 1: Microbiological, physical, chemical, aesthetic and chemical determinands. SABS Standards Division. Pretoria, South Africa.

SARDESSAI, Y. N. 2005. Viable but non-culturable bacteria: their impact on public health. *Current Science Bangalore,* 89(10)**,** 1650.

SARKAR, A. & PANDEY, P. 2015. River water quality modelling using artificial neural network technique. *Aquatic Procedia,* 4(1070-1077.

SAXENA, T., KAUSHIK, P. & KRISHNA MOHAN, M. 2015. Prevalence of *E. coli* O157:H7 in water sources: an overview on associated diseases, outbreaks and detection methods. *Diagnostic Microbiology and Infectious Diseases,* 82(3)**,** 249-64.

SCHOLTEN, H., KASSAHUN, A., REFSGAARD, J. C., KARGAS, T., GAVARDINAS, C. & BEULENS, A. J. M. 2007. A methodology to support multidisciplinary model-based water management. *Environmental Modelling & Software,* 22(5)**,** 743-759.

SERSHEN, S., RODDA, N., STENSTRÖM, T.-A., SCHMIDT, S., DENT, M., BUX, F., HANKE, N., BUCKLEY, C. & FENNEMORE, C. 2016. Water security in South Africa: perceptions on public expectations and municipal obligations, governance and water re-use. *Water SA,* 42(3)**,** 456-465.

SETH, I. 2015. Use of artificial neural networks and genetic algorithms in urban water management: a brief overview. *Journal American Water Works Association,* 107(5)**,** 93-97.

SHAMIMUZZAMAN, M., NAYEEM, R. H., ARA, N., RAHMAN, M. M., JAHID, M. I. K. & HASAN, M. N. 2019. Physico-Chemical and Microbiological Quality Assessment of Supply Water around Dhaka City, Bangladesh. *Journal of Water Resource and Protection,* 11(03)**,** 280.

SHARMA, V., NEGI, S. C., RUDRA, R. P. & YANG, S. 2003. Neural networks for predicting nitrate-nitrogen in drainage water. *Agricultural Water Management,* 63(3)**,** 169-183.

SHEN, Y., XU, C., SUN, Q., SCHWARZ, S., OU, Y., YANG, L., HUANG, Z., EICHHORN, I., WALSH, T. R. & WANG, Y. 2018. Prevalence and genetic analysis of mcr-3-positive Aeromonas species from humans, retail meat, and environmental water samples. *Antimicrobial Agents and Chemotherapy,* 62(9)**,** e00404-18.

SIDDIQUE, N. & ADELI, H. 2013. *Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing,* United Kingdom, John Wiley & Sons.

SINGH, K. P. & GUPTA, S. 2012. Artificial intelligence based modeling for predicting the disinfection by-products in water. *Chemometrics and Intelligent Laboratory Systems,* 114(122-131.

SOUTH AFRICA. 1997. Water Services Act 108 of 1997.

SOYUPAK, S., KILIC, H., KARADIREK, I. E. & MUHAMMETOGLU, H. 2011. On the usage of artificial neural networks in chlorine control applications for water distribution networks with high quality water. *Journal of Water Supply: Research and Technology-AQUA,* 60(1)**,** 51.

STAUBER, C. E., WEDGWORTH, J. C., JOHNSON, P., OLSON, J. B., AYERS, T., ELLIOTT, M. & BROWN, J. 2016. Associations between Self-Reported Gastrointestinal Illness and Water System Characteristics in Community Water Supplies in Rural Alabama: A Cross-Sectional Study. *PLoS One,* 11(1)**,** e0148102.

STEYNOR, A., HEWITSON, B. & TADROSS, M. 2009. Projected future runoff of the Breede River under climate change. *Water SA,* 35(4)**,** 433-440.

SWANEPOEL, A., BARNARD, S., RECKNAGEL, F. & CAO, H. 2016. Evaluation of models generated via hybrid evolutionary algorithms for the prediction of Microcystis concentrations in the Vaal Dam, South Africa. *Water SA,* 42(2)**,** 243-252.

SWARTZ, C. 2000. Guidelines for the upgrading of existing small water treatment plants. *Water Research Commission*.

TABARI, H. & HOSSEINZADEH TALAEE, P. 2015. Reconstruction of river water quality missing data using artificial neural networks. *Water Quality Research Journal of Canada,* 50(4)**,** 326-335.

TABESH, M., SOLTANI, J., FARMANI, R. & SAVIC, D. 2009. Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling. *Journal of Hydroinformatics,* 11(1)**,** 1-17.

TAMURA, K. & NEI, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution,* 10(3)**,** 512-526.

TARR, C. L., LARGE, T. M., MOELLER, C. L., LACHER, D. W., TARR, P. I., ACHESON, D. W. & WHITTAM, T. S. 2002. Molecular characterization of a serotype O121:H19, a distinct Shiga toxin-producing clone of pathogenic *Escherichia coli. Infection and Immunity*, 70(12), 6853-6859.

TASHAOUIE, H. R., GHOLIKANDI, G. B. & HAZRATI, H. 2012. Artificial neural network modeling for predict performance of pressure filters in a water treatment plant. *Desalination and Water Treatment,* 39(1-3)**,** 192-198.

TAVASOLIFAR, A., BINA, B., AMIN, M. M., EBRAHIMI, A. & JALALI, M. 2012. Implementation of hazard analysis and critical control points in the drinking water supply system. *International Journal of Environmental Health Engineering,* 1(1)**,** 32.

TER BRAAK, C. J. & SMILAUER, P. 1998. CANOCO reference manual and user's guide to Canoco for Windows: software for canonical community ordination (version 4). Microcomputer Power, New York.

THOMPSON, P. & MAJAM, S. 2009. *The Development of a Generic Water Safety Plan for Small Community Water Supply: Report to the Water Research Commission*, Water Research Commission.

TOMPERI, J., PELO, M. & LEIVISKÄ, K. 2013. Predicting the residual aluminum level in water treatment process. *Drinking Water Engineering and Science,* 6(1)**,** 39-46.

TRUSSEL, R. R. 2005. *Water Treatment: Where have we been and where are we going?* [Online]. Available: http://www.trusselltech.com/uploads/media_items/historical-water-treatment-with-a-future-perspective.original.pdf [Accessed 16 June 2016].

TSITSIFLI, S. & TSOUKALAS, D. S. 2019. Water Safety Plans and HACCP implementation in water utilities around the world: benefits, drawbacks and critical success factors. *Environmental Science and Pollution Research***,** 1-13.

TSOUKALAS, D. S. & TSITSIFLI, S. 2018. A critical evaluation of Water Safety Plans (WSPs) and HACCP implementation in water utilities. Multidisciplinary Digital Publishing Institute Proceedings, 600.

TURETGEN, I. 2008. Induction of Viable but Nonculturable (VBNC) state and the effect of multiple subculturing on the survival ofLegionella pneumophila strains in the presence of monochloramine. *Annals of Microbiology,* 58(1)**,** 153-156.

US ENVIRONMENTAL PROTECTION AGENCY 2007. Hazard Analysis Critical Control Point (HACCP) Strategies for Distribution System Monitoring, Hazard Assessment and Control. Washington DC: HDR, Inc.

US ENVIRONMENTAL PROTECTION AGENCY. 2016. *Indicators: Conductivity* [Online]. Available: https://www.epa.gov/national-aquatic-resource-surveys/indicators-conductivity [Accessed 19 October 2019].

VAIRAVAMOORTHY, K. & LUMBERS, J. 1998. Leakage reduction in water distribution systems: optimal valve control. *Journal of Hydraulic Engineering,* 124(11)**,** 1146-1154.

VALENTIN, N. & DENŒUX, T. 2001. A neural network-based software sensor for coagulation control in a water treatment plant. *Intelligent Data Analysis,* 5(1)**,** 23-39.

VAN DER HOEK, J. P., BERTELKAMP, C., VERLIEFDE, A. R. D. & SINGHAL, N. 2014. Drinking water treatment technologies in Europe: state of the art – challenges – research needs. *Journal of Water Supply: Research and Technology - AQUA,* 63(2)**,** 124.

VAN DER LAAN, M., VAN ANTWERPEN, R. & BRISTOW, K. L. 2012. River water quality in the northern sugarcane-producing regions of South Africa and implications for irrigation: A scoping study. *Water SA,* 38(1).

VAN DER WALT, J. & NELL, B. 2002. Integrated catchment management: the Mooi River (Northwest Province, South Africa) as a case study. *Cuadernos de investigación geográfica,* 28**,** 109-126.

VAN VLIET, M., YEARSLEY, J., FRANSSEN, W., LUDWIG, F., HADDELAND, I., LETTENMAIER, D. & KABAT, P. 2012. Coupled daily streamflow and water temperature modelling in large river basins. *Hydrology and Earth System Sciences,* 16(11)**,** 4303-4321.

VAN VUUREN, L. 2015. Drought management – Strengthening our knowledge armoury. *The Water Wheel.* Water Research Commission.

VEERAPANENI, S., BUDD, G., BOND, R. & HORSLEY, M. 2010. Using neural networks to predict treatment process performance. *Journal American Water Works Association,* 102(4)**,** 38-40,42,44.

VENTER, A., BARNARD, S., DICKINSON, M. A., JANSE VAN VUUREN, S., LEVANETS, A. & TAYLOR, J. C. 2013. Planktonic algae and cyanoprokaryotes as indicators of ecosystem quality in the Mooi River system in the North-West Province, South Africa. *Water SA,* 39(5)**,** 707.

VICENTE, H., DIAS, S., FERNANDES, A., ABELHA, A., MACHADO, J. & NEVES, J. 2012. Prediction of the quality of public water supply using artificial neural networks. *Journal of Water Supply: Research and Technology - AQUA,* 61(7)**,** 446-459.

VIEIRA, J., FONSECA, A., VILAR, V. J., BOAVENTURA, R. A. & BOTELHO, C. M. 2013. Water quality modelling of Lis River, Portugal. *Environmental Science and Pollution Research,* 20(1)**,** 508-524.

VORA, G. J., MEADOR, C. E., BIRD, M. M., BOPP, C. A., ANDREADIS, J. D. & STENGER, D. A. 2005. Microarray-based detection of genetic heterogeneity, antimicrobial resistance, and the viable but nonculturable state in human pathogenic Vibrio spp. *Proceedings of the National Academy of Sciences,* 102(52)**,** 19109-19114.

WACHLA, D., PRZYSTALKA, P. & MOCZULSKI, W. 2015. A method of leakage location in water distribution networks using artificial neuro-fuzzy system. *IFAC-PapersOnLine,* 48(21)**,** 1216-1223.

WASKAN, J. 2016. *Connectionism* [Online]. The Internet Encyclopedia of Philosophy. Available: http://www.iep.utm.edu/connect/ [Accessed 15 June 2016].

WELK, A., RECKNAGEL, F., CAO, H., CHAN, W.-S. & TALIB, A. 2008. Rule-based agents for forecasting algal population dynamics in freshwater lakes discovered by hybrid evolutionary algorithms. *Ecological Informatics,* 3(1)**,** 46-54.

WENZLER, E., KAMBOJ, K. & BALADA-LLASAT, J. M. 2015. Severe Sepsis Secondary to Persistent Lysinibacillus sphaericus, Lysinibacillus fusiformis and Paenibacillus amylolyticus Bacteremia. *International Journal of Infectious Diseases,* 35, 93-5.

WERBOS, J. 1974. *Beyond regression: New tools for predicting and analysis in the behavioral sciences.* PhD thesis, Harvard University.

WORLD HEALTH ORGANIZATION 2003. Aluminium in drinking-water - Background document for development of WHO guidelines for drinking-water quality*,* Geneva, Switzerland.

WORLD HEALTH ORGANIZATION 2004. Guidelines for drinking-water quality. 3rd ed. Geneva: World Health Organization.

WORLD HEALTH ORGANIZATION 2011. Guidelines for drinking-water quality*,* Geneva, Switzerland, World Health Organization.

WORLD HEALTH ORGANIZATION 2017a. Guidelines for drinking water quality: fourth edition incorporating the first addendum. Geneva: World Health Organization.

WORLD HEALTH ORGANIZATION 2017b. Water quality and health-review of turbidity: information for regulators and water suppliers. World Health Organization.

WU, C. L. & CHAU, K. W. 2011. Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology,* 399(3-4)**,** 394-409.

WU, G.-D. & LO, S.-L. 2010. Effects of data normalization and inherent-factor on decision of optimal coagulant dosage in water treatment by artificial neural network. *Expert Systems with Applications,* 37(7)**,** 4974-4983.

WU, W., DANDY, G. C. & MAIER, H. R. 2011. Application of artificial neural networks to forecasting water quality in a chloraminated water distribution system. *19th International Congress on Modeling and Simulation.* Perth, Australia.

WU, W., DANDY, G. C. & MAIER, H. R. 2014. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. *Environmental Modelling & Software,* 54, 108-127.

XU, H.-S., ROBERTS, N., SINGLETON, F., ATTWELL, R., GRIMES, D. & COLWELL, R. 1982. Survival and viability of nonculturable *Escherichia coli* and *Vibrio cholerae* in the estuarine and marine environment. *Microbial ecology,* 8(4)**,** 313-323.

YANG, T. M., FAN, S. K., FAN, C. & HSU, N. S. 2014. Establishment of turbidity forecasting model and early-warning system for source water turbidity management using back-propagation artificial neural network algorithm and probability analysis. *Environmental Monitoring and Assessment,* 186(8)**,** 4925-4934.

YE, B., WANG, W., YANG, L., WEI, J. & XUELI, E. 2011. Formation and modeling of disinfection by-products in drinking water of six cities in China. *Journal of Environmental Monitoring,* 13(5)**,** 1271-1275.

YURDUSEV, M. A. & FIRAT, M. 2009. Adaptive neuro fuzzy inference system approach for municipal water consumption modeling: An application to Izmir, Turkey. *Journal of Hydrology,* 365(3)**,** 225-234.

ZADEH, L. A. 1965. Fuzzy sets. *Information and Control,* 8(3)**,** 338-353.

ZHANG, Q. & STANLEY, S. J. 1997. Forecasting raw-water quality parameters for the North Saskatchewan River by neural network modeling. *Water Research,* 31(9)**,** 2340-2350.

ZHANG, Q. & STANLEY, S. J. 1999. Real-time water treatment process control with artificial neural networks. *Journal of Environmental Engineering,* 125(2)**,** 153-160.

ZHANG, S., YE, C., LIN, H., LV, L. & YU, X. 2015. UV disinfection induces a VBNC state in *Escherichia coli* and *Pseudomonas aeruginosa*. *Environ Sci Technol,* 49(3)**,** 1721-8.

ZHANG, Y., GAO, X., SMITH, K., INIAL, G., LIU, S., CONIL, L. B. & PAN, B. 2019. Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network. *Water Research,* 164(114888).

ZHAO, F., BI, X., HAO, Y. & LIAO, X. 2013. Induction of viable but nonculturable *Escherichia coli* O157:H7 by high pressure CO2 and its characteristics. *PLoS One,* 8(4)**,** e62388.

ZOUNEMAT-KERMANI, M., RAMEZANI-CHARMAHINEH, A., ADAMOWSKI, J. & KISI, O. 2018. Investigating the management performance of disinfection analysis of water distribution networks using data mining approaches. *Environmental Monitoring Assessment,* 190(7)**,** 397.

APPENDIX A

Datasets of Plant A and Plant C used during ANN and EA analysis

Dataset for Plant A - EC forecasting

DATE	EC_Raw	Turbidity_ Raw	pH_Raw	Sulphate _Raw	Chloride _Raw	*E.coli*_Raw	TC_Raw	EC_FW
05/01/2010	59,7	4,67	8,37	124	29	192	1842	61,2
12/01/2010	58,6	3,09	8,22	124	26,4	18,3	613	59,1
19/01/2010	56,5	2,07	8,02	112	24	80	100	56
26/01/2010	55,4	1,69	8,27	108	22	185	7070	58,5
03/02/2010	61,6	3,15	8,23	108	23	235	8665	63,3
09/02/2010	57,5	1,89	8,22	100	22	146	2420	59,5
23/02/2010	54,6	3,01	8,29	100	22	180	12100	60
02/03/2010	63,3	5,31	8,14	112	26	155	12100	63,7
09/03/2010	53,4	3,51	8,28	108	27	66	2420	55,7
26/04/2010	58,1	8,41	8,15	104	25	40	99	60,5
04/05/2010	57,2	0,4	7,79	100	34	28	1986	65,3
11/05/2010	55,6	2,86	8,22	76	24	38	1120	59,7
18/05/2010	48,1	3,99	7,93	56	20	18	201	51,6
25/05/2010	61	6,09	8,03	44	19	69	816	58,4
01/06/2010	60,3	4,19	8,05	72	21	20	1553	55,8
08/06/2010	62,2	3,16	8,02	72	23	816	2420	62
15/06/2010	50,6	1,87	8,2	68	18	11	263	55,1
29/06/2010	56,1	2,22	8,06	76	21	6	1414	55,5
06/07/2010	55	1,96	7,93	88	21	8,5	1203	54,5
13/07/2010	68	1,77	8,14	88	20	24,6	1986	69,8
27/07/2010	67,3	2,09	7,95	92	25	88,2	2420	61
03/08/2010	63,8	3,75	7,97	68	23	13	1986	62,4
10/08/2010	60,9	3,48	8,08	72	24	5	197	59
17/08/2010	61,2	2,66	8,09	52	24	19	84	61,8
24/08/2010	63,5	2,26	8,13	60	26	1	206	63,4
31/08/2010	57,3	4,05	8,24	68	26	9	204	58,3
07/09/2010	64,8	6,64	8,11	88	21	15	411	61
14/09/2010	62	3,94	8	84	31	28	219	60,1
21/09/2010	63,9	3,83	8,12	96	21	19,9	980	65
28/09/2010	59	2,29	8,22	72	25	27	613	59,4
05/10/2010	56,5	3,51	8,13	84	29	21,3	517,2	60,9
12/10/2010	61,8	5,81	8,13	72	28	21	613	61,1
19/10/2010	55,2	5,77	8,22	72	24	50,4	488,4	56,1
26/10/2010	61,3	3,7	8,26	80	26	21	278	64,4
02/11/2010	60,6	4,65	8,15	84	24	12	1300	61,3
09/11/2010	57,4	3,6	8,2	72	22	43	1986	60,6
16/11/2010	63,2	5,34	8,08	112	25	34	1046	61,4
30/11/2010	67,9	6,1	8,15	112	35	138	1414	65,8
07/12/2010	61,5	5,75	8,15	108	30	110	691	64,2
14/12/2010	59	4,59	8,16	96	30	75	2599	61,5
21/12/2010	58,6	4,42	7,58	108	34	84	3466	62,1
28/12/2010	61,4	5,15	8,28	80	30	61	2420	62,6
04/01/2011	53,7	4,62	8,54	88	25	548	2420	55
11/01/2011	48,8	10	8,32	80	22	75	4813	56,4
18/01/2011	51,4	4,68	8,29	72	23	131	5654	55
25/01/2011	52,6	7,63	8,51	68	24	796	6212	56,7
01/02/2011	61,9	7,11	8,17	64	23	95	5199	61,9
08/02/2011	59,2	4,96	8,55	72	24	1728	6212	58,7
15/02/2011	57,2	4,35	8,73	68	27	222	9678	49,5
22/02/2011	52,2	3,63	8,49	80	23	180	9678	55,5
01/03/2011	61,5	6,64	8,45	72	23	1050	24196	55,9
08/03/2011	63,5	4,39	8,48	76	20	82	31062	56,9
15/03/2011	58,1	2,32	8,61	84	27	218	7945	58,2
22/03/2011	61,3	3,45	8,47	88	27	47	4820	60,9

Dataset for Plant A - EC forecasting (continue)

DATE	EC_Raw	Turbidity_ Raw	pH_Raw	Sulphate _Raw	Chloride _Raw	*E.coli*_Raw	TC_Raw	EC_FW
29/03/2011	53,6	4,82	8,34	80	23	54	6932	56,8
05/04/2011	55,7	2,34	8,28	76	22	24	2908	56,9
12/04/2011	59,1	3,19	8,14	88	17	31	2420	60,5
19/04/2011	59,5	5,52	8,1	88	25	68	3466	62
26/04/2011	59,6	1,26	8,28	76	22	29	1733	59,9
04/05/2011	55,8	2,62	8,23	92	25	112	1102	56,4
10/05/2011	57,1	7,5	8,17	84	23	152	549	59
17/05/2011	47,5	11,1	8,16	88	23	127	3080	55,4
24/05/2011	59,7	1,69	8,34	96	25	16	479	58,9
31/05/2011	53,1	4,69	8,34	84	26	13	435	60,3
14/02/2012	60,7	2,5	8,25	120	30	102	4840	67
21/02/2012	59,7	1,93	8,23	116	27	60	2420	58,2
28/02/2012	58,2	1,89	8,19	112	29	124	2420	62,6
06/03/2012	46,2	2,38	8,24	104	29	80	2420	61,3
13/03/2012	63,2	2,65	8,27	100	29	44	2420	65,5
20/03/2012	64,6	3,4	8,69	108	29	39	4839	63,5
28/03/2012	64,2	1,78	8,39	116	28	16	1553	64,1
03/04/2012	61	3,77	8,48	100	29	15	3106	59,3
10/04/2012	63,2	2,1	8,41	96	32	12	1842	61,9
17/04/2012	65,5	1,67	8,66	108	31	14	1096	65,4
24/04/2012	55,5	2,96	8,35	104	28	18	652	59,8
02/05/2012	62,2	2,19	8,62	112	31	13	1454	63,8
08/05/2012	66,5	2,45	8,25	48	28	36	3106	68,3
15/05/2012	65,8	1,36	8,57	120	20	6	409	65,2
22/05/2012	57,8	2,41	8,42	104	32	20	688	67,6
05/06/2012	69,2	1,9	8,82	116	32	2	606	67,6
12/06/2012	59,3	2,01	8,83	104	27	512	788	69,7
19/06/2012	66,5	1,88	8,53	104	28	167	4839	66,4
26/06/2012	69,6	2,01	8,87	84	30	2	551	69,5
03/07/2012	63,3	1,97	8,91	80	28	0	113	61,2
10/07/2012	63,9	1,82	8,95	112	31	6	72	64,6
17/07/2012	76,2	1,89	8,91	92	29	2	197	69,3
24/07/2012	65,8	1,96	8,96	112	29	35	112	68,2
31/07/2012	63,1	1,5	8,91	112	27	32	52	66,3
07/08/2012	68,8	1,48	8,92	64	27	8	37	67,5
14/08/2012	63,9	1,82	8,95	116	27	6	58	68,1
21/08/2012	65,1	2,03	8,98	112	28	10	68	66
29/08/2012	68,2	1,84	8,67	116	29	4	108	62,4
04/09/2012	59,4	2,16	8,87	112	29	2	236	63,2
11/09/2012	64,8	3,34	8,62	84	28	57	1034	68,7
18/09/2012	62,3	3,52	8,97	116	28	6	229	67,7
25/09/2012	66	2,62	8,97	108	30	8	498	68,1
02/10/2012	67,7	3,48	8,95	112	30	8	120	66,2
09/10/2012	63,3	5,83	9,01	84	32	30	1454	68
16/10/2012	66,1	3,06	8,95	128	29	42	449	66,2
23/10/2012	67,1	1,82	8,98	96	29	37	582	68
30/10/2012	64,6	1,22	9,11	112	28	236	977	68
06/11/2012	60,5	3,66	8,96	116	27	81	1034	65,1
13/11/2012	60,7	1,24	8,79	112	30	236	1732	64,8
20/11/2012	60,4	4,52	8,92	116	27	0	977	60,5
27/11/2012	73,6	1,7	8,87	112	35	160	4839	69,2
04/12/2012	60,6	4,16	8,74	124	31	113	2240	64,9
11/12/2012	63,5	1,78	9,16	120	30	231	4839	66,5
18/12/2012	67,8	3,94	8,72	124	28	117	4839	64,4
27/12/2012	63,3	3,61	8,67	124	35	238	2240	66,7
02/01/2013	64,5	1,45	8,42	124	30	240	4840	62,4
08/01/2013	64,8	1,84	8,62	124	30	131	3922	64,8
15/01/2013	63,3	1	8,74	116	33	276	5199	63,9
05/02/2013	66,8	1,92	8,54	108	33	176	5656	64,6
12/02/2013	65,2	1,73	8,89	104	30	162	7945	64,8
19/02/2013	65,2	3,27	8,8	104	35	132	2908	65,6

Dataset for Plant A - EC forecasting (continue)

DATE	EC_Raw	Turbidity_ Raw	pH_Raw	Sulphate _Raw	Chloride _Raw	*E.coli*_Raw	TC_Raw	EC_FW
26/02/2013	63	3,95	8,65	112	34	114	3466	62,5
05/03/2013	67,7	4,61	8,77	108	33	119	4480	67,2
12/03/2013	62,8	2,86	8,85	108	34	81	1642	65
19/03/2013	63,2	3,85	8,66	104	32	364	9680	64,2
26/03/2013	66,6	2,28	8,96	112	33	117	9678	65,5
02/04/2013	65,1	2,87	8,63	104	34	336	7944	63,2
09/04/2013	73,2	3,63	8,79	104	32	9678	9678	67,1
16/04/2013	63,3	3,4	8,57	104	34	192	404	65,3
23/04/2013	67,1	2,34	8,74	44	30	63	1302	66,9
30/04/2013	69,5	1,9	7,8	108	33	50	802	67,3
07/05/2013	66,4	2,27	8,91	108	33	172	802	67,3
14/05/2013	67,9	2,03	8,83	84	30	0	4812	68,8
20/05/2013	68,2	3,39	8,79	112	30	44	566	70,3
28/05/2013	67,6	5,53	8,84	112	34	48	1102	69,92
04/06/2013	68,42	1,71	8,85	112	35	16	318	70,44
11/06/2013	68,53	3,3	8,84	108	32	4	1041	71,25
18/06/2013	66,7	3,64	8,19	108	32	4	1379	69,34
25/06/2013	66,4	3,31	8,74	104	30	29	507	68,46
02/07/2013	66,8	1,91	8,63	112	31	30	600	68,88
09/07/2013	68,4	1,46	8,45	104	32	25	198	70,3
16/07/2013	67,8	1,03	8,55	116	32	4	336	69,48
27/08/2013	69,2	13,9	8,61	112	34	8	344	70,6
04/09/2013	68,9	3,17	8,6	108	30	87	1230	69,91
10/09/2013	68,1	1,65	8,16	112	31	29	398	69,4
17/09/2013	71,1	2,06	8,6	104	31	134	2452	71,78
25/09/2013	69,9	1,63	8,71	112	35	60	1013	70,4
01/10/2013	70,8	1,45	8,65	120	33	160	1091	71,52
08/10/2013	72,5	1,41	8,62	120	35	254	1652	73,67
15/10/2013	71,2	1,07	8,57	100	33	192	1445	71,2
22/10/2013	71,8	6,24	8,75	108	35	344	2005	73,16
29/10/2013	69,2	6,57	8,97	112	35	178	2005	71,93
05/11/2013	70,7	4,37	8,5	120	35	222	12980	72,1
12/11/2013	71	2,91	8,63	112	34	130	6017	72,3
19/11/2013	70,3	2,94	8,73	120	33	216	9222	72,01
03/12/2013	65,7	21,8	8,71	108	35	48392	48392	70,5
10/12/2013	70,9	4,23	8,42	108	35	8212	48392	72,55
17/12/2013	70	1,81	8,4	116	35	946	20924	71,7
08/01/2014	74,6	2,84	8,45	116	36	1352	48392	73,9
14/01/2014	74,2	0,37	8,06	124	36	700	17328	76,5
21/01/2014	73,6	3,12	8,37	116	36	104	8704	74,75
28/01/2014	72,6	0,36	8,39	104	35	1112	15402	74,3
04/02/2014	72,8	2,29	8,45	116	37	4532	24196	71,17
11/02/2014	52,3	9,2	8,29	84	23	14540	48392	58,79
18/02/2014	56,46	1,48	8,3	96	26	2382	25994	59,31
25/02/2014	62,53	3,25	8,25	100	27	48392	48392	63,8
04/03/2014	68,26	2,64	8,34	108	33	406	18416	68,83
11/03/2014	65,47	1,9	8,55	112	30	7746	48392	66,9
18/03/2014	61,22	3,59	8,55	96	27	7746	48392	63,63
25/03/2014	65,25	6,86	8,36	84	27	48392	48392	69,48
01/04/2014	65,2	2,5	8,21	52	28	350	9768	67,26
08/04/2014	68,4	3,13	8,22	88	27	960	12262	72,03
15/04/2014	68,7	0,93	8,4	96	29	20924	48392	69,72
22/04/2014	70,02	2,5	8,25	100	29	350	9768	72,04
29/04/2014	68,17	2,6	8,26	100	30	2518	9222	70,33
06/05/2014	69,41	2,46	8,26	104	30	959	12033	71,80
13/05/2014	68,77	3,43	8,32	104	28	882	5794	72,00
20/05/2014	68,33	2,48	8,33	104	30	1553	19863	70,39
27/05/2014	57,8	1,86	8,3	88	35	1522	10462	71,80
17/06/2014	69,34	2,98	8,37	104	30	4010	4010	72,81
24/06/2014	70,21	3,84	8,43	108	28	2920	34660	74,37
01/07/2014	71,33	2,22	8,41	120	30	602	4962	73,64

Dataset for Plant A - EC forecasting (continue)

DATE	EC_Raw	Turbidity_ Raw	pH_Raw	Sulphate _Raw	Chloride _Raw	*E.coli*_Raw	TC_Raw	EC_FW
08/07/2014	71,72	0,43	8,56	108	28	51	780	73,89
15/07/2014	73,15	2,43	8,61	116	30	41	266	66,68
22/07/2014	72,64	3,4	8,57	108	31	1500	2247	75,03
29/07/2014	71,85	3,87	8,7	104	30	355	2613	74,33
05/08/2014	72,27	2,17	8,71	124	30	1782	4884	72,99
12/08/2014	70,78	2,25	8,23	116	30	20	1664	73,33
19/08/2014	71,25	1,52	8,34	104	29	309	2098	73,21
26/08/2014	74,41	1,33	8,29	108	30	52	1523	76,44
02/09/2014	72,9	1,57	8,33	120	30	20	882	75,07
09/09/2014	72,98	2,81	8,27	116	31	536	1191	74,43
16/09/2014	71,67	1,63	8,34	112	30	482	4286	74,63
23/09/2014	72,58	2,4	8,47	120	36	7701	24196	78,27
30/09/2014	72,99	3,61	8,3	120	31	2978	12031	75,71
07/10/2014	72,48	3,55	8,38	112	33	2014	24196	74,24
14/10/2014	72,5	2,31	8,44	104	32	5475	24196	74,70
04/11/2014	68,06	3,91	8,33	132	33	124	2005	
11/11/2014	69,46	3,52	8,39	120	35	111	2005	71,00
18/11/2014	68,18	3,01	8,43	112	35	906	4010	70,56
25/11/2014	66,85	2,35	8,45	108	34	374	6152	69,74
02/12/2014	65,3	2,29	8,42	120	34	518	6896	68,34
09/12/2014	64,84	4,7	8,60	112	33	690	8704	66,90
06/01/2015	62,91	3,18	9,01	112	32	73	9 208	64,95
13/01/2015	61,55	2,00	9,06	120	30	697	6 131	62,89
20/01/2015	61,86	2,18	9,00	80	34	85	9 208	64,12
27/01/2015	62,97	2,51	9,00	104	32	369	5 475	64,18
03/02/2015	58,65	2,22	8,34	104	30	355	17 329	61,27
10/02/2015	59,94	2,66	8,29	92	34	160	4 323	61,58
17/02/2015	59,70	2,73	8,31	116	39	315	2 142	60,45
24/02/2015	59,32	3,28	8,30	132	36	343	1 471	59,50
03/03/2015	63,04	2,62	8,45	104	38	486	978	65,79
10/03/2015	58,66	3,29	8,38	76	37	524	958	60,29
17/03/2015	61,65	3,33	8,45	124	37	620	780	63,44
24/03/2015	64,65	3,51	8,44	120	50	315	1 200	65,22
31/03/2015	66,18	2,97	8,15	104	37	536	826	66,94
07/04/2015	65,46	4,43	8,22	128	39	354	11 199	70,24
14/04/2015	65,12	2,31	8,19	120	39	473	599	67,61
21/04/2015	64,91	1,91	8,22	120	38	346	1 205	66,61
28/04/2015	60,78	2,19	8,24	80	34	70	762	61,92
05/05/2015	58,11	2,02	8,20	104	39	61	205	60,84
12/05/2015	65,31	2,69	8,30	128	40	605	755	67,24
19/05/2015	65,35	3,33	8,28	112	39	542	959	67,50
26/05/2015	64,16	0,59	8,26	124	37	430	1 301	66,12
02/06/2015	62,86	1,38	8,48	120	35	376	1 585	66,95
09/06/2015	62,18	0,80	8,27	116	39	538	914	63,79
17/06/2015	67,15	0,48	8,16	128	34	41	1 720	69,67
23/06/2015	66,24	1,95	8,04	120	36	288	1 993	68,55
30/06/2015	66,10	1,20	8,33	128	41	167	3 026	69,16
07/07/2015	66,36	2,39	8,32	112	32	119	313	68,63
14/07/2015	67,76	0,92	8,25	148	44	10	1 445	69,03
21/07/2015	68,63	0,30	8,30	128	41	20	406	69,55
28/07/2015	84,18	1,87	9,16	152	50	397	1 292	66,92
04/08/2015	59,67	2,23	9,19	168	79	10	189	61,86
11/08/2015	59,97	1,99	8,26	112	40	31	1 203	62,37
18/08/2015	58,96	2,66	8,18	104	50	31	189	61,26
25/08/2015	59,77	0,25	8,03	172	42	146	1 421	61,08
01/09/2015	58,54	3,19	8,32	116	41	72	15 531	60,27
08/09/2015	60,67	3,25	8,32	48	38	20	913	63,14
15/09/2015	60,26	5,04	8,27	56	41	171	657	62,54
22/09/2015	69,18	4,57	8,10	128	41	20	161	70,37
29/09/2015	69,42	3,90	7,84	124	37	63	1 106	72,28
06/10/2015	70,52	1,91	8,12	100	42	439	691	72,21

Dataset for Plant A - EC Data set for Plant A - EC forecasting (continue)

DATE	EC_Raw	Turbidity_ Raw	pH_Raw	Sulphate _Raw	Chloride _Raw	*E.coli*_Raw	TC_Raw	EC_FW
13/10/2015	69,47	2,89	8,24	120	39	122	845	72,28
20/10/2015	70,52	4,32	8,10	132	42	20	703	72,26
27/10/2015	70,05	5,15	8,10	116	38	20	789	72,14
03/11/2015	69,08	3,37	8,10	120	47	30	4 360	72,21
10/11/2015	70,54	5,00	7,99	128	46	31	638	72,87
17/11/2015	69,19	3,81	8,27	116	43	183	4 106	73,39
24/11/2015	69,49	4,16	8,30	124	45	122	3 076	72,29
01/12/2015	68,73	3,22	8,24	104	41	128	2 400	72,35

*TC: Total coliforms

Dataset for Plant A - Turbidity forecasting

Date	Turb_Raw	Cu_Raw	Iron_Raw	*E.coli*_Raw	TC_Raw	Turb_FW
05/01/2010	4,67	4	40	192	1842	0,49
12/01/2010	3,09	1	30	18,3	613	0,84
19/01/2010	2,07	0	0	80	100	0,35
26/01/2010	1,69	1	10	185	7070	0,83
03/02/2010	3,15	0	0	235	8665	0,54
09/02/2010	1,89	3	30	146	2420	0,92
23/02/2010	3,01	4	20	180	12100	0,42
02/03/2010	5,31	2	0	155	12100	1,61
09/03/2010	3,51	1	10	66	2420	0,47
26/04/2010	8,41	3	30	40	99	0,56
04/05/2010	0,4	1	10	28	1986	0,49
11/05/2010	2,86	1	10	38	1120	0,6
18/05/2010	3,99	7	60	18	201	0,96
25/05/2010	6,09	2	50	69	816	0,98
01/06/2010	4,19	2	20	20	1553	0,69
08/06/2010	3,16	3	0	816	2420	0,83
15/06/2010	1,87	3	10	11	263	0,67
29/06/2010	2,22	3	40	6	1414	0,79
06/07/2010	1,96	2	20	8,5	1203	0,34
13/07/2010	1,77	1	30	24,6	1986	0,91
27/07/2010	2,09	1	10	88,2	2420	0,65
03/08/2010	3,75	1	20	13	1986	0,81
10/08/2010	3,48	5	10	5	197	0,64
17/08/2010	2,66	3	10	19	84	0,88
24/08/2010	2,26	2	20	1	206	0,94
31/08/2010	4,05	1	10	9	204	0,76
07/09/2010	6,64	1	10	15	411	0,51
14/09/2010	3,94	1	0	28	219	0,87
21/09/2010	3,83	4	10	19,9	980	0,94
28/09/2010	2,29	0	20	27	613	0,74
05/10/2010	3,51	1	0	21,3	517,2	0,83
12/10/2010	5,81	2	10	21	613	0,56
19/10/2010	5,77	8	10	50,4	488,4	0,63
26/10/2010	3,7	0	30	21	278	0,84
02/11/2010	4,65	1	10	12	1300	0,62
09/11/2010	3,6	1	10	43	1986	0,51
23/11/2010	6,3	0	0	33	1733	0,53
30/11/2010	6,1	0	20	138	1414	0,59
07/12/2010	5,75	3	20	110	691	0,71
14/12/2010	4,59	2	10	75	2599	0,79
21/12/2010	4,42	9	60	84	3466	0,58
28/12/2010	5,15	2	10	61	2420	0,99
04/01/2011	4,62	2	10	548	2420	0,95
11/01/2011	10	2	50	75	4813	1,83
18/01/2011	4,68	1	0	131	5654	0,97
25/01/2011	7,63	1	0	796	6212	0,9
01/02/2011	7,11	2	80	95	5199	0,96

Dataset for Plant A - Turbidity forecasting (continue)

Date	Turb_Raw	Cu_Raw	Iron_Raw	*E.coli*_Raw	TC_Raw	Turb_FW
08/02/2011	4,96	45	40	1728	6212	0,98
15/02/2011	4,35	2	30	222	9678	0,99
22/02/2011	3,63	3	40	180	9678	0,8
01/03/2011	6,64	8	10	1050	24196	0,76
08/03/2011	4,39	5	0	82	31062	0,39
15/03/2011	2,32	3	20	218	7945	0,76
22/03/2011	3,45	3	40	47	4820	0,61
29/03/2011	4,82	7	40	54	6932	0,46
05/04/2011	2,34	7	40	24	2908	0,46
12/04/2011	3,19	0	30	31	2420	0,83
19/04/2011	5,52	3	10	68	3466	0,38
26/04/2011	1,26	0	0	29	1733	0,39
04/05/2011	2,62	2	40	112	1102	0,94
10/05/2011	7,5	2	50	152	549	0,59
17/05/2011	11,1	3	50	127	3080	0,99
24/05/2011	1,69	3	10	16	479	0,98
31/05/2011	4,69	2	80	13	435	0,84
08/02/2012	2,72	1	100	520	4839	0,38
14/02/2012	2,5	1	40	102	4840	0,4
21/02/2012	1,93	1	10	60	2420	0,53
28/02/2012	1,89	1	20	124	2420	0,5
06/03/2012	2,38	1	20	80	2420	0,52
13/03/2012	2,65	1	20	44	2420	0,49
20/03/2012	3,4	8	20	39	4839	0,42
28/03/2012	1,78	7	0	16	1553	0,45
03/04/2012	3,77	1	10	15	3106	0,45
10/04/2012	2,1	2	0	12	1842	0,53
17/04/2012	1,67	5	10	14	1096	0,41
24/04/2012	2,96	1	20	18	652	0,6
02/05/2012	2,19	2	10	13	1454	0,99
08/05/2012	2,45	2	10	36	3106	0,98
15/05/2012	1,36	2	10	6	409	0,54
22/05/2012	2,41	3	20	20	688	0,63
05/06/2012	1,9	1	70	2	606	0,46
12/06/2012	2,01	3	40	512	788	0,94
19/06/2012	1,88	4	10	167	4839	0,98
26/06/2012	2,01	3	30	2	551	0,59
03/07/2012	1,97	4	10	0	113	0,39
10/07/2012	1,82	1	10	6	72	0,67
17/07/2012	1,89	2	30	2	197	0,33
24/07/2012	1,96	1	10	35	112	0,47
31/07/2012	1,5	2	20	32	52	0,59
07/08/2012	1,48	1	30	8	37	0,71
14/08/2012	1,82	1	10	6	58	0,99
21/08/2012	2,03	5	10	10	68	0,53
29/08/2012	1,84	2	10	4	108	0,53
04/09/2012	2,16	1	10	2	236	0,48
11/09/2012	3,34	1	30	57	1034	0,25
18/09/2012	3,52	1	10	6	229	0,21
25/09/2012	2,62	2	20	8	498	0,42
02/10/2012	3,48	1	10	8	120	1
09/10/2012	5,83	1	40	30	1454	0,49
16/10/2012	3,06	0	40	42	449	0,44
23/10/2012	1,82	1	40	37	582	0,36
30/10/2012	1,22	1	10	236	977	0,41
06/11/2012	3,66	1	10	81	1034	0,61
13/11/2012	1,24	4	10	236	1732	0,24
20/11/2012	4,52	7	10	0	977	0,55
27/11/2012	1,7	3	20	160	4839	0,37
04/12/2012	4,16	1	10	113	2240	0,47
11/12/2012	1,78	1	10	231	4839	0,35
18/12/2012	3,94	1	50	117	4839	0,39

Dataset for Plant A - Turbidity forecasting (continue)

Date	Turb_Raw	Cu_Raw	Iron_Raw	*E.coli*_Raw	TC_Raw	Turb_FW
27/12/2012	3,61	2	20	238	2240	0,5
02/01/2013	1,45	1	50	240	4840	0,46
08/01/2013	1,84	1	20	131	3922	0,76
15/01/2013	1	1	10	276	5199	0,75
22/01/2013		2	10	142	4840	0,59
05/02/2013	1,92	2	0	176	5656	0,69
12/02/2013	1,73	2	0	162	7945	0,66
19/02/2013	3,27	4	0	132	2908	0,54
26/02/2013	3,95	4	0	114	3466	3,63
05/03/2013	4,61	2	10	119	4480	0,61
12/03/2013	2,86	2	10	81	1642	0,88
19/03/2013	3,85	2	0	364	9680	0,52
26/03/2013	2,28	5	10	117	9678	0,41
02/04/2013	2,87	0	20	336	7944	0,76
09/04/2013	3,63	2	10	9678	9678	0,44
16/04/2013	3,4	0	10	192	404	0,43
23/04/2013	2,34	4	10	63	1302	0,39
30/04/2013	1,9	2	10	50	802	0,54
07/05/2013	2,27	1	10	172	802	0,38
14/05/2013	2,03	2	10	0	4812	0,5
20/05/2013	3,39	12	10	44	566	0,58
28/05/2013	5,53	2	0	48	1102	0,4
04/06/2013	1,71	2	10	16	318	0,79
11/06/2013	3,3	1	40	4	1041	0,3
18/06/2013	3,64	2	0	4	1379	0,25
25/06/2013	3,31	3	10	29	507	0,26
02/07/2013	1,91	2	30	30	600	0,22
09/07/2013	1,46	1	20	25	198	0,63
16/07/2013	1,03	2	10	4	336	0,28
27/08/2013	13,9	0	10	8	344	0,37
04/09/2013	3,17	0	20	87	1230	0,38
10/09/2013	1,65	0	20	29	398	0,31
17/09/2013	2,06	1	10	134	2452	0,36
25/09/2013	1,63	0	20	60	1013	0,3
01/10/2013	1,45	2	40	160	1091	0,31
08/10/2013	1,41	11	10	254	1652	0,3
15/10/2013	1,07	1	30	192	1445	0,27
22/10/2013	6,24	3	20	344	2005	0,61
29/10/2013	6,57	2	30	178	2005	0,42
05/11/2013	4,37	12	10	222	12980	0,18
12/11/2013	2,91	7	10	130	6017	0,32
19/11/2013	2,94	2	0	216	9222	0,18
26/11/2013	2,85	6	10	148	2468	0,43
03/12/2013	21,8	2	30	48392	48392	0,3
10/12/2013	4,23	0	20	8212	48392	0,36
17/12/2013	1,81	4	10	946	20924	0,44
08/01/2014	2,84	1	10	1352	48392	0,33
21/01/2014	3,12	2	10	104	8704	0,47
28/01/2014	0,36	2	0	1112	15402	0,51
04/02/2014	2,29	2	30	4532	24196	0,46
11/02/2014	9,2	6	80	14540	48392	1,54
18/02/2014	1,48	3	20	2382	25994	0,81
25/02/2014	3,25	9	10	48392	48392	0,54
04/03/2014	2,64	1	10	406	18416	0,93
11/03/2014	1,9	0	20	7746	48392	0,47
18/03/2014	3,59	12	10	7746	48392	0,59
25/03/2014	6,86	1	10	48392	48392	0,7
01/04/2014	2,5	1	10	350	9768	0,56
08/04/2014	3,13	5	0	960	12262	0,87
15/04/2014	0,93	0	80	20924	48392	0,73
22/04/2014	2,5	8	10	350	9768	0,4
29/04/2014	2,6	5	20	2518	9222	0,4

Data set for Plant A - Turbidity forecasting (continue)

Date	Turb_Raw	Cu_Raw	Iron_Raw	*E.coli*_Raw	TC_Raw	Turb_FW
06/05/2014	2,46	1	30	959	12033	0,34
13/05/2014	3,43	1	20	882	5794	0,32
20/05/2014	2,48	3	30	1553	19863	0,56
27/05/2014	1,86	5	50	1522	10462	0,51
17/06/2014	2,98	3	10	4010	4010	0,7
24/06/2014	3,84	0	10	2920	34660	0,67
01/07/2014	2,22	1	10	602	4962	0,62
08/07/2014	0,43	0	40	51	780	0,64
15/07/2014	2,43	3	10	41	266	0,99
22/07/2014	3,4	4	10	1500	2247	0,54
29/07/2014	3,87	0	20	355	2613	0,33
05/08/2014	2,17	3	40	1782	4884	0,81
12/08/2014	2,25	4	20	20	1664	0,31
19/08/2014	1,52	6	20	309	2098	0,5
26/08/2014	1,33	1	40	52	1523	0,69
02/09/2014	1,57	5	10	20	882	0,67
09/09/2014	2,81	1	10	536	1191	0,33
16/09/2014	1,63	3	0	482	4286	0,52
23/09/2014	2,4	2	20	7701	24196	0,15
30/09/2014	3,61	0	10	2978	12031	0,32
07/10/2014	3,55	6	0	2014	24196	0,41
14/10/2014	2,31	4	60	5475	24196	0,69
04/11/2014	3,91	1	10	124	2005	0,4
11/11/2014	3,52	3	0	111	2005	0,24
18/11/2014	3,01	9	70	906	4010	0,16
25/11/2014	2,35	11	30	374	6152	0,24
02/12/2014	2,29	6	10	518	6896	0,35
09/12/2014	4,7	3	0	690	8704	0,39
06/01/2015	3,18	10	10	73	9 208	0,27
13/01/2015	2,00	2	0	697	6 131	0,21
20/01/2015	2,18	4	10	85	9 208	0,27
27/01/2015	2,51	4	10	369	5 475	0,30
03/02/2015	2,22	7	20	355	17 329	0,26
10/02/2015	2,66	3	20	160	4 323	0,28
17/02/2015	2,73	2	10	315	2 142	0,69
24/02/2015	3,28	0	20	343	1 471	0,62
03/03/2015	2,62	1	20	486	978	0,40
10/03/2015	3,29	3	20	524	958	0,40
17/03/2015	3,33	2	30	620	780	0,38
24/03/2015	3,51	1	10	315	1 200	0,64
31/03/2015	2,97	1	20	536	826	0,67
07/04/2015	4,43	4	10	354	11 199	0,37
14/04/2015	2,31	2	40	473	599	0,36
21/04/2015	1,91	1	30	346	1 205	0,31
28/04/2015	2,19	4	10	70	762	0,20
05/05/2015	2,02	1	30	61	205	0,40
12/05/2015	2,69	2	10	605	755	0,22
19/05/2015	3,33	2	0	542	959	0,27
26/05/2015	0,59	5	10	430	1 301	0,11
02/06/2015	1,38	1	10	376	1 585	0,12
09/06/2015	0,80	1	10	538	914	0,10
17/06/2015	0,48	1	20	41	1 720	0,09
23/06/2015	1,95	3	10	288	1 993	0,32
30/06/2015	1,20	8	10	167	3 026	0,59
07/07/2015	2,39	13	10	119	313	0,36
14/07/2015	0,92	12	0	10	1 445	0,19
21/07/2015	0,30	3	40	20	406	0,39
28/07/2015	1,87	27	20	397	1 292	0,21
04/08/2015	2,23	12	100	10	189	0,30
11/08/2015	1,99	10	0	31	1 203	0,40
18/08/2015	2,66	1	10	31	189	0,31
25/08/2015	0,25	1	10	146	1 421	0,32

Dataset for Plant A - Turbidity forecasting (continue)

Date	Turb_Raw	Cu_Raw	Iron_Raw	*E.coli*_Raw	TC_Raw	Turb_FW
01/09/2015	3,19	10	20	72	15 531	0,94
08/09/2015	3,25	6	10	20	913	0,40
15/09/2015	5,04	3	10	171	657	0,56
22/09/2015	4,57	22	0	20	161	0,88
29/09/2015	3,90	1	10	63	1 106	0,88
06/10/2015	1,91	2	10	439	691	0,29
13/10/2015	2,89	5	20	122	845	0,39
20/10/2015	4,32	0	0	20	703	0,36
27/10/2015	5,15	10	10	20	789	0,18
03/11/2015	3,37	0	20	30	4 360	0,25
10/11/2015	5,00	4	10	31	638	0,57
17/11/2015	3,81	2	10	183	4 106	0,55
24/11/2015	4,16	0	10	122	3 076	0,36
01/12/2015	3,22	23	10	128	2 400	0,72

*Turb: Turbidity; Cu: Copper; TC: Total coliforms

Dataset for Plant C - EC forecasting

Date	EC_Raw	pH_Raw	Sulphate_Raw	Chloride_Raw	Sodium_Raw	TOC_Raw	EC_FW
11/02/2009	37	7,61	85	32	23	7,0	44
25/02/2009	34	7,84	76	35	20	6,4	41
11/03/2009	34	7,95	83	55	22	5,9	45
01/04/2009	44	8,68	71	33	30	5,8	49
08/04/2009	44	8,96	74	35	33	5,8	50
29/07/2009	77	8,90	148	71	51	4,6	76
12/08/2009	75	8,81	137	63	56	5,6	77
02/09/2009	74	9,39	160	77	72	6,5	77
09/09/2009	76	9,46	152	72	70	6,8	78
30/09/2009	79	6,86	152	72	66	7,1	79
11/11/2009	69	8,64	112	67	60	6,1	72
09/12/2009	66	8,62	134	53	47	5,8	70
30/12/2009	63	8,47	123	53	51	5,5	67
20/01/2010	46	7,80	89	59	55	6,6	53
17/02/2010	35	8,10	49	58	55	6,5	42
07/07/2010	57	9,33	122	53	47	5,9	58
14/07/2010	61	9,54	121	49	46	6,5	62
11/08/2010	74	9,03	125	48	38	6,3	77
06/10/2010	69	8,60	129	31	30	6,9	76
22/12/2010	32	7,64	74	31	32	12,0	38
19/01/2011	21	7,98	23	60	64	8,4	28
26/01/2011	26	8,08	33	56	70	8,3	33
02/02/2011	18	8,00	17	58	55	8,1	26
09/02/2011	21	7,86	22	63	60	8,5	29
16/02/2011	26	8,15	26	61	60	7,9	31
16/03/2011	40	8,57	60	65	71	6,6	43
23/03/2011	41	8,06	53	65	60	7,1	42
30/03/2011	49	8,38	68	58	54	6,9	50
04/05/2011	55	8,72	97	53	46	7,2	57
11/05/2011	60	8,93	78	57	48	6,3	61
15/06/2011	58	9,20	85	51	55	8,1	56
22/06/2011	62	9,33	84	52	49	6,7	64
06/07/2011	63	9,30	89	49	44	6,0	66
13/07/2011	65	9,39	97	49	40	6,7	67
20/07/2011	63	9,25	105	45	46	6,4	63
27/07/2011	65	9,41	129	46	45	6,4	67
24/08/2011	70	9,25	118	44	47	7,1	74
31/08/2011	68	9,26	109	43	40	7,3	72
07/09/2011	70	9,18	109	80	39	8,2	73
28/09/2011	63	9,25	115	31	31	9,0	67
12/10/2011	69	8,69	120	25	27	8,7	73
19/10/2011	71	8,45	130	23	24	7,4	75

Dataset for Plant C - EC forecasting (continue)

Date	EC_Raw	pH_Raw	Sulphate_Raw	Chloride_Raw	Sodium_Raw	TOC_Raw	EC_FW
16/11/2011	74	9,02	105	10	16	7,9	78
30/11/2011	74	8,74	121	6	12	8	75
07/12/2011	74	9,02	124	14	12	6,6	77
14/12/2011	73	8,96	127	11	12	7,9	76
11/01/2012	66	8,74	115	49	46	7,8	68
01/02/2012	54	8,52	89	56	54	6,4	59
15/02/2012	59	8,34	96	58	54	6,5	64
29/02/2012	59	8,24	104	56	48	7,0	64
14/03/2012	62	8,77	111	62	63	6,9	67
28/03/2012	65	8,98	112	58	56	6,3	71
11/04/2012	61	9,04	109	57	55	6,1	64
18/04/2012	60	8,86	113	61	55	6,1	61
25/04/2012	65	9,01	108	57	55	6,2	64
02/05/2012	60	9,32	100	51	53	5,8	64
23/05/2012	67	9,07	101	59	70	7	69
30/05/2012	66	8,84	102	61	61	6,4	68
06/06/2012	68	8,60	102	61	52	4,6	70
13/06/2012	68	8,80	104	61	49	5,4	69
20/06/2012	70	8,23	104	56	42	6,2	69
27/06/2012	70	8,23	100	52	53	6,2	69
04/07/2012	62	8,89	97	54	51	6,0	65
11/07/2012	62	8,94	93	56	58	6,0	63
18/07/2012	67	9,23	104	54	60	6,0	69
25/07/2012	71	9,19	99	52	50	4,5	71
01/08/2012	70	9,02	104	51	53	5,5	70
08/08/2012	69	9,13	103	51	63	4,7	70
29/08/2012	67	8,72	97	51	47	5,3	68
05/09/2012	71	8,85	107	50	52	5,2	70
12/09/2012	66	8,64	91	51	52	6,0	67
19/09/2012	67	9,10	92	48	46	5,8	70
03/10/2012	65	9,06	87	53	54	6,3	68
24/10/2012	70	8,18	132	41	46	5,1	72
07/11/2012	70	8,67	124	40	39	6,6	71
21/11/2012	67	9,08	104	39	46	6,0	70
12/12/2012	60	8,92	97	52	52	8,2	64
09/01/2013	55	9,05	88	48	38	5,6	58
16/01/2013	56	9,11	93	44	43	5,7	59
23/01/2013	56	8,89	93	52	50	6,3	59
06/02/2013	53	8,61	86	63	52	5,6	57
13/02/2013	52	9,03	82	65	62	7,1	55
20/02/2013	54	8,73	84	64	58	5,4	57
27/02/2013	54	8,62	84	59	45	6,2	57
06/03/2013	55	8,53	84	61	60	4,9	59
13/03/2013	55	8,31	91	64	63	4,7	59
20/03/2013	58	8,33	107	65	64	4,5	57
03/04/2013	54	8,76	90	63	64	4,2	58
10/04/2013	53	8,73	94	63	60	4,6	57
17/04/2013	54	8,66	99	61	57	4,8	56
24/04/2013	49	8,19	83	60	59	4,8	51
08/05/2013	48	8,68	86	59	61	5,4	51
15/05/2013	44	9,35	88	57	52	5,3	46
22/05/2013	39	9,33	89	57	58	5,7	43
05/06/2013	41	9,00	105	55	56	4,9	44
19/06/2013	46	8,81	112	64	53	4,4	49
26/06/2013	49	8,79	111	56	59	4,5	48
03/07/2013	49	8,86	106	53	55	3,8	50
17/07/2013	49	8,89	116	55	54	4,6	51
24/07/2013	48	9,64	106	55	53	5,5	51
31/07/2013	46	9,55	105	49	47	5,8	47
07/08/2013	50	9,49	111	41	36	5,5	52
14/08/2013	50	8,98	107	39	36	5,4	54
04/09/2013	70	8,76	106	49	49	6,1	71

Dataset for Plant C - EC forecasting (continue)

Date	EC_Raw	pH_Raw	Sulphate_Raw	Chloride_Raw	Sodium_Raw	TOC_Raw	EC_FW
11/09/2013	69	8,78	106	53	53	5,9	70
25/09/2013	69	8,70	112	50	39	6,2	69
02/10/2013	69	8,86	105	46	42	6,3	71
09/10/2013	72	8,74	99	47	41	5,8	72
16/10/2013	71	8,84	105	47	46	6,3	73
23/10/2013	71	8,55	108	46	48	5,8	71
30/10/2013	66	9,06	89	45	43	6,2	69
06/11/2013	65	9,24	93	45	43	7,3	67
13/11/2013	67	9,01	98	46	44	7,1	66
20/11/2013	65	8,92	87	43	48	8	69
08/01/2014	34	7,57	49	25	17	6,0	39
22/01/2014	32	7,66	46	20	17	5,8	35
12/02/2014	30	7,44	49	36	18	5,3	33
19/02/2014	32	7,72	46	22	13	5,6	37
26/02/2014	36	7,92	53	25	23	5,2	38
05/03/2014	34	7,72	50	25	18	5,7	38
12/03/2014	23	7,19	29	31	13	7,5	29
19/03/2014	25	7,32	29	15	17	7,7	30
26/03/2014	30	7,53	40	21	16	8,5	35
02/04/2014	34	7,43	46	22	17	7,6	39
16/04/2014	42	8,34	80	30	28	5,6	44
23/04/2014	43	8,07	63	32	28	5,4	46
30/04/2014	48	9,06	75	39	32	5,3	52
14/05/2014	51	9,55	86	43	38	6,6	53
21/05/2014	51	9,63	94	46	31	5,3	53
04/06/2014	58	9,15	99	51	36	5	58
25/06/2014	60	8,66	105	53	48	6	61
02/07/2014	63	9,47	108	54	39	6	65
09/07/2014	64	9,16	107	54	51	7	64
16/07/2014	65	9,58	107	57	34	6	67
23/07/2014	65	9,19	106	59	51	6	67
13/08/2014	63	9,31	108	59	46	7	66
20/08/2014	68	8,82	115	56	56	7	71
27/08/2014	66	9,26	122	58	49	8	69
17/09/2014	71	8,90	147	57	57	7	74
08/10/2014	76	8,92	157	63	62	7	77
15/10/2014	76	8,57	154	63	41	7	78
22/10/2014	71	8,68	135	61	60	8	72
29/10/2014	71	9,55	136	62	64	9	74
05/11/2014	68	8,86	123	60	61	9	72
12/11/2014	67	8,60	121	58	49	8	68
19/11/2014	68	9,42	121	57	61	7	71
03/12/2014	64	8,29	115	51	49	6	70
10/12/2014	70	8,64	133	51	52	6	76
17/12/2014	64	8,90	127	45	36	6	64

*TOC: Total organic carbon

Dataset for Plant C - Turbidity forecasting

Date	EC_Raw	pH_Raw	Sodium_Raw	Turb_FW
11/02/2009	37	7,61	23	0,4
25/02/2009	34	7,84	20	0,3
11/03/2009	34	7,95	22	0,3
01/04/2009	44	8,68	30	0,2
08/04/2009	44	8,96	33	0,2
29/07/2009	77	8,90	51	0,4
12/08/2009	75	8,81	56	0,3
02/09/2009	74	9,39	72	0,5
09/09/2009	76	9,46	70	0,6
30/09/2009	79	6,86	66	0,5
11/11/2009	69	8,64	60	0,5

Dataset for Plant C - Turbidity forecasting (continue)

Date	EC_Raw	pH_Raw	Sodium_Raw	Turb_FW
09/12/2009	66	8,62	47	0,4
30/12/2009	63	8,47	51	0,3
20/01/2010	46	7,80	55	0,3
17/02/2010	35	8,10	55	0,6
07/07/2010	57	9,33	47	0,6
14/07/2010	61	9,54	46	0,6
11/08/2010	74	9,03	38	0,6
06/10/2010	69	8,60	30	0,9
22/12/2010	32	7,64	32	0,3
19/01/2011	21	7,98	64	0,4
26/01/2011	26	8,08	70	0,3
02/02/2011	18	8,00	55	0,4
09/02/2011	21	7,86	60	0,3
16/02/2011	26	8,15	60	0,4
16/03/2011	40	8,57	71	0,3
23/03/2011	41	8,06	60	0,4
30/03/2011	49	8,38	54	0,3
04/05/2011	55	8,72	46	0,3
11/05/2011	60	8,93	48	0,2
15/06/2011	58	9,20	55	1,7
22/06/2011	62	9,33	49	0,4
06/07/2011	63	9,30	44	0,6
13/07/2011	65	9,39	40	0,5
20/07/2011	63	9,25	46	1,4
27/07/2011	65	9,41	45	0,8
24/08/2011	70	9,25	47	0,5
31/08/2011	68	9,26	40	0,5
07/09/2011	70	9,18	39	0,7
28/09/2011	63	9,25	31	0,9
12/10/2011	69	8,69	27	0,5
19/10/2011	71	8,45	24	0,4
16/11/2011	74	9,02	16	0,3
30/11/2011	74	8,74	12	0,4
07/12/2011	74	9,02	12	0,4
14/12/2011	73	8,96	12	0,3
11/01/2012	66	8,74	46	0,3
01/02/2012	54	8,52	54	0,2
15/02/2012	59	8,34	54	0,2
29/02/2012	59	8,24	48	0,2
14/03/2012	62	8,77	63	0,3
28/03/2012	65	8,98	56	0,3
11/04/2012	61	9,04	55	0,4
18/04/2012	60	8,86	55	0,4
25/04/2012	65	9,01	55	0,5
02/05/2012	60	9,32	53	0,6
23/05/2012	67	9,07	70	0,3
30/05/2012	66	8,84	61	0,4
06/06/2012	68	8,60	52	0,5
13/06/2012	68	8,80	49	0,4
20/06/2012	70	8,23	42	0,4
27/06/2012	70	8,23	53	0,4
04/07/2012	62	8,89	51	0,4
11/07/2012	62	8,94	58	1,0
18/07/2012	67	9,23	60	0,6
25/07/2012	71	9,19	50	0,6
01/08/2012	70	9,02	53	0,6
08/08/2012	69	9,13	63	0,4
29/08/2012	67	8,72	47	0,8
05/09/2012	71	8,85	52	0,5
12/09/2012	66	8,64	52	0,4
19/09/2012	67	9,10	46	0,4
03/10/2012	65	9,06	54	0,6

Dataset for Plant C - Turbidity forecasting (continue)

Date	EC_Raw	pH_Raw	Sodium_Raw	Turb_FW
24/10/2012	70	8,18	46	0,4
07/11/2012	70	8,67	39	0,4
21/11/2012	67	9,08	46	0,8
12/12/2012	60	8,92	52	0,7
09/01/2013	55	9,05	38	0,3
16/01/2013	56	9,11	43	0,4
23/01/2013	56	8,89	50	0,4
06/02/2013	53	8,61	52	0,5
13/02/2013	52	9,03	62	0,4
20/02/2013	54	8,73	58	0,5
27/02/2013	54	8,62	45	0,3
06/03/2013	55	8,53	60	0,3
13/03/2013	55	8,31	63	0,4
20/03/2013	58	8,33	64	0,3
03/04/2013	54	8,76	64	0,3
10/04/2013	53	8,73	60	0,3
17/04/2013	54	8,66	57	0,2
24/04/2013	49	8,19	59	0,3
08/05/2013	48	8,68	61	0,4
15/05/2013	44	9,35	52	0,7
22/05/2013	39	9,33	58	0,6
05/06/2013	41	9,00	56	0,8
19/06/2013	46	8,81	53	0,7
26/06/2013	49	8,79	59	0,5
03/07/2013	49	8,86	55	0,6
17/07/2013	49	8,89	54	0,5
24/07/2013	48	9,64	53	0,8
31/07/2013	46	9,55	47	0,7
07/08/2013	50	9,49	36	0,4
14/08/2013	50	8,98	36	0,4
04/09/2013	70	8,76	49	0,5
11/09/2013	69	8,78	53	0,4
25/09/2013	69	8,70	39	0,4
02/10/2013	69	8,86	42	0,5
09/10/2013	72	8,74	41	0,5
16/10/2013	71	8,84	46	0,4
23/10/2013	71	8,55	48	0,4
30/10/2013	66	9,06	43	0,4
06/11/2013	65	9,24	43	0,7
13/11/2013	67	9,01	44	0,8
20/11/2013	65	8,92	48	0,5
08/01/2014	34	7,57	17	0,3
22/01/2014	32	7,66	17	0,3
12/02/2014	30	7,44	18	0,6
19/02/2014	32	7,72	13	0,4
26/02/2014	36	7,92	23	0,3
05/03/2014	34	7,72	18	0,4
12/03/2014	23	7,19	13	0,2
19/03/2014	25	7,32	17	0,3
26/03/2014	30	7,53	16	0,2
02/04/2014	34	7,43	17	0,4
16/04/2014	42	8,34	28	0,5
23/04/2014	43	8,07	28	0,5
30/04/2014	48	9,06	32	0,5
14/05/2014	51	9,55	38	0,4
21/05/2014	51	9,63	31	0,4
04/06/2014	58	9,15	36	0,4
25/06/2014	60	8,66	48	0,8
02/07/2014	63	9,47	39	1,1
09/07/2014	64	9,16	51	1,1
16/07/2014	65	9,58	34	1,1
23/07/2014	65	9,19	51	0,8

Dataset for Plant C - Turbidity forecasting (continue)

Date	EC_Raw	pH_Raw	Sodium_Raw	Turb_FW
13/08/2014	63	9,31	46	0,7
20/08/2014	68	8,82	56	0,5
27/08/2014	66	9,26	49	0,7
17/09/2014	71	8,90	57	0,8
08/10/2014	76	8,92	62	0,4
15/10/2014	76	8,57	41	0,7
22/10/2014	71	8,68	60	0,8
29/10/2014	71	9,55	64	0,6
05/11/2014	68	8,86	61	0,7
12/11/2014	67	8,60	49	0,9
19/11/2014	68	9,42	61	0,7
03/12/2014	64	8,29	49	0,4
10/12/2014	70	8,64	52	0,4
17/12/2014	64	8,90	36	0,4

*Turb: Turbidity

APPENDIX B

Permission letter

NWU®
NORTH-WEST UNIVERSITY
NOORDWES-UNIVERSITEIT
YUNIBESITI YA BOKONE-BOPHIRIMA

Private Bag X6001, Potchefstroom
South Africa 2520
Tel: 018 299-1111/2222
Web: http://www.nwu.ac.za
Microbiology
Email: 10926542@nwu.ac.za

24 February 2020

TO: Postgraduate administration

Permission to use "Artificial neural networks: applications in the drinking water sector" by G. O'Reilly, C. C. Bezuidenhout and J. J. Bezuidenhout (Published in Water Science and Technology, 2018; doi: 10.2166/ws.2018.016) as a chapter in thesis

We as the additional authors to the abovementioned article hereby give permission to use this in the thesis of Ms G O'Reilly. The candidate performed all the searches prepared the original draft and worked the comments of the co-authors into the final draft that was submitted. Ms O'Reilly was, under the guidance of the co-authors, also responsible for working the comments of the reviewers into the final version of the manuscript and was responsible to prepare the rebuttal. An updated version of the manuscript was in included in the thesis.

Yours sincerely

Dr JJ Bezuidenhout
Co-promotor

Prof CC Bezuidenhout
Promotor