Advanced statistical and mathematical modelling of a constructed wetland, treating abattoir waste water: A new modelling method for resilience analysis

By

John L. Byrne BSc (Hons)

A thesis presented for the Degree of Doctor of Philosophy

Department of Science and Health

Institute of Technology Carlow

September 2019

Supervisors: Dr. David N. Dowling and Dr. Dina Brazil

Declaration

I, the undersigned, cerify that the information I have provided in this thesis is correct and that I have read and am aware of my responsibilities in the Institute of Technology Carlow's Policy and Procedures for Postgraduate research, Students (Admissions, Supervision and Examination). I further confirm that I am unaware of any potential conflicts of interest that would compromise the Institute and/or the applicant in pusuit of the level 10 award sought.

I hereby declare that I am the author of the thesis and that no part has been submitted to any other instiurion for examination. I certify that, to my knowledge, this thesis does not violate any plagiarism standards or infringe on copyright of proprietary rights. Any ideas, knowledge, techniques, statements and any other included materials are rightlfully acknowledged in accordance with standard referencing practices. This document is a tru copy of my thesis to date and includes all final revision by my supervisors.

This is orginal work and all conclusions made are drawn from original work and data gathered by me.

Signature: _____

Name (Print): _____

Date of Submission: _____

Acknowledgements

I am greatly indebted to Dr David Dowling and Dr Dina Brazil for their patience, perseverance and persistence in this ardous endeavour. Their generous time and support throughout this research. Without their mentoring, this PhD would not have been completed.

To the technical staff within the Department of Science and Health for their support in training, equipment setups and calibration techniques, my sincere thanks.

To my colleagues within the EnviroCore group at the Institute of Technology Carlow, for listening to my endless discussions on various statistical, mathematical and complexity modelling regimes. Thanks for listening to me!!.

To my family (Mam,Dad and brothers) – the thesis is completed. Thank you all for standing by me so patiently.

To my poor long suffering wife Ann and my son Matthew, I'm finished.
{MFMLME}

Abstract:

The meat processing industry produces high strength toxic waste from the slaughtering processes. Conventional methods were to ship the blood waste products to other companies for treatment, leading to high costs in storage, transport and licensing. A constructed wetland (CW) is an artificial treatment system consisting of soil, plants and water used to passively treat contaminated wastewater. The use of constructed wetlands may provide a cost effective passive treatment for the high strength toxic waste by-products, with the potential to implement effective waste treatment and lowering the overall costs of storage and transport. The output treated wastewaters are expected to show an overall reduction in toxicity. Typical wastewater generated by the abattoir industry has a high organic content (both soluble and particulate) with high nutrient and microbial loads. There are limited studies on the application of constructed wetlands treating high strength wastewaters especially in the meat industry and within the scope of the Irish landscape even fewer.

The main objective of this research was the application of statstical and mathemathical modelling methodologies to provide an in-depth review of the constructed wetland treating abattoir wastewater. The research is based on a free water constructed wetland system, with twelve interconntected ponds treating abattoir wastewater. Samples were retrived from six locations. The dissolved air flotation (DAF) plant, pond 1, pond 6, pond 9, pond 12 and the local stream.

This work reviews the input (DAF) and output (stream) variables of the wetland, but sampling of specific internal locations within the wetland system; the aim was to understand what was happening inside the wetland system. This work delves into the internal processes i.e. pond 1, pond 6, pond 9 and pond 12. The methods employed used advanced statistical methods such as Canonical variate analysis (CVA), Discriminant Functional Analysis (DFA) and Principal Component Analysis (PCA). Soft-computing methods such as self-organising maps (SOMs), artificial neural networks (ANNs), fuzzy logic and Bayesian belief networks were utilised.

Canonical variate analysis (CVA) potentially revealed a second source of *E.coil* entering the wetland. Principal component analysis (PCA) highlights the dominance of indicator bacteria within the wetland system. The use of indicator

bacteria as a viable tracer to deterimine the hydraulic retention time (HRT), with modelled values between 55 days to 128 days depending on the season.
The use of Bayesian networks and sensitivity analysis on the wetland soils and sludge reveals that the wetlands soil porosity may be an issue in particular from pond 9 to pond 12. Self-organised criticality (SOC) analysis was employed, indicating the potential that indicator bacteria and nutrients are not being treated correctly within the CW and water depth was highlighted as an issue. The use of self-organising maps (SOMs) and artificial neural networks (ANNs) were employed to understand complex interactions within the CW. The SOM models can provide what-if scenarios. The use of coefficient of reliability (COR) and Cronbachs' Alpha which measure reliability of the CW. Indicating a questionable reliability of the CW and potentially show that indicator bacteria provide a good indicator as to the overall performance of the CW. Combining soft-computing methods such fuzzy logic and Bayesian network analysis into hybrid models to understand the CWs resilience. The models indicated issues in pond 9 to pond 12, and that the overall resilience of the CW was an undesirable system to treat the abattoir wastewater.

Table of Contents

Chapter 1. Introduction

Chapter 2. Material and Methods

Chapter 3. Establishment and multi-variate analysis of a constructed wetland dataset

Chapter 4. Constructed wetland: Traditional models

Chapter 5. Bio-aerosol analysis of the constructed wetland

Chapter 8. Section 1: The application of artificial neural networks (ANNs) to understand wetland behaviours

Chapter 8. Section 2: The application of self-organising maps (SOMs) to understand wetland behaviours

Chapter 9. Coeffiecient of Reiliability and Cronbachs' Alpha

Chapter 10. Hybrid fuzzy logic and de-centralised Bayesian belief networks

Section 1: Iterative and reductive Principal Component Analysis

Chapter 10

Section 2: F-IND: Fuzzy indices of the environmental variables of the constructed wetland

Chapter 10

Section 3: Creation of two fuzzy logic decentralised Bayesian belief networks

Chapter 11. Conclusions and recommendations

List of Abbreviations

A-D Test – Anderson-Darling Test

ANOVA – Analysis of Variance

ANN – Artificial Neural Networks

ASM – Activated Sludge Model

BBN – Bayesian Belief Networks

BAT - Best Available Technology

BATNEEC - Best Available Technology Not Exceeding Costs

BOD – Biological Oxygen Demand

BN- Bayesian Network

BTS – Bartlett's Test of Sphericity

CCI – Correctly Classified Instances

CDA – Canonical Discriminant Analysis

CFU – Colony Forming Units

COD – Chemical Oxygen Demand

COR – Coefficient of Reliability

CSTR – Continuous Stirred Reactor

CV – Coefficient of Variation

CVA- Canonical Variate Analysis

CR – Correlation ratio

CW – Constructed Wetland

DA – Discriminant Analysis

DAF – Dissolved Air Floatation

DAG – Direct Acyclic Graph

DDS – Dissipative Dynamic System

DF – Dispersed Flow

DFA – Discriminant Functional Analysis

DFM – Dispersed Flow Model

DO – Dissolved Oxygen

ELV – Emission Limit Values

EPA – Environmental Protection Agency

FC – Faecal Coliform

FCW – Farm Constructed Wetland

FI – Fuzzy Indexing

FIDOL – Frequency, Intensity, Duration, relative Offensiveness and Location

F-IND – Fuzzy Indices

FL – Fuzzy Logic

FOG – Fats Oil Grease

FWS – Free Water Surface

GA – Genetic Algorithm

GP – Genetic Programming

GSA – Global Sensitivity Analysis

HDPE – High Density Polyethylene

HFCW- Horizontal Flow Constructed Wetland

HRT – Hydraulic Retention Time

ICW – Integrated Constructed Wetland

IPPC – Integrated Pollution Prevention and Control

KMO - Kaiser-Meyer-Olkin

K-S test- Kolmogorov-Smirnov Test

LBE – Load Based Entropy

LINT- Linear Interpolation

MAC – MacConkey Agar

MANOVA- Multivariate Analysis of Variance

MBM – Meat and Bone Meal

MLGA – Membrane Lactose Glucuronide Agar

PCA – Principal Component Analysis

PCA – Plate Count Agar

PDF – Probability Distribution Function

PF – Plug Flow

WFD – Water Framework Directive

WWTP – Wastewater Treatment Plant

WQI – Water Quality Index

RBDs – River Basin Districts

RTD – Residence Time Distribution

SA – Sensitivity Analysis

SAS – Surface Air System

SCE – Soft Scaled Entropy

SOC – Self Organised Criticality

SF – Surface Flow

SOM – Self Organizing Maps

SSF – Sub Surface Flow

SVM – Support Vector Machines

TIS – Tank is Series

TSA – Trypticase Soy Agar

VFCW – Vertical Flow Constructed Wetland

WEI- Wetland Efficiency Index

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Constructed wetlands – a general introduction

Constructed wetlands (CWs) are anthropogenic engineered wastewater treatment systems that encompass multiple treatment principles including biological, chemical and physical processes, mimicking processes occurring in natural wetlands (Kadlec and Knight, 1996; Vymazal, 2005; Zhao *et al*., 2018). Because of their imitation of natural ecological degradation processes, the CWs can be more cost effective than chemical and mechanical systems which are typified by wastewater treatment plants (WWTPs). Constructed wetlands can be classified as passive systems due to the fact that no external energy inputs are undertaken to improve the bioremediation processes. Once constructed, the wetlands are planted with reeds and grasses, the treatment of wastewaters is achieved due to the environmental and climate processes. In contrast active systems such as (WWTPs), use chemical dosing agents such as iron sulphate and alum which are introduced to help remove large particulates. Air diffusers and pumps are also employed to reduce the toxicity of the wastewater to futher improve the WWTPs performance. Regarding the use of CWs for wastewater treatment in Ireland, the technology was considered to be in its infancy when compared to North America and other European countries (Healy and Cawley, 2002).

Constructed wetlands (CWs) are capable of treating municipal wastewaters (Cameron *et al*, 2003), (Babatunde *et al*., 2008, Babatunde *et al*., 2016). The successful application of CWs for the treatment of domestic wastewater, has led to a multitude of other uses for CWs for treating wastewater from different sources including acid mine, abattoir, agricultural, storm water, landfill leachate, motorway run off, phosphorus removal in steel slags and the potential reduction of antibiotics and antibiotic bacteria from livestock wastewater (Mitsch and Wise, 1998, Poggi-Varaldo *et al*., 2004; Harrington *et al*., 2005; Yazdi and Scholtz, 2010; Koumanova and Lavrova, 2010; Verhoeven *et al*., 2012; Seo *et al*., 2016; Almeida *et al*., 2019). In Europe constructed wetlands are becoming a common and viable alternative treatment system and in rural areas 95% of the wetlands are

subsurface flow wetlands. It is predicted that the number of these systems will reach 10,000 in Europe in early 2000's (Platzer, 2000).

Constructed wetlands (CWs) are considered an effective waste management system relating to livestock wastewater (Mendieta-Pino *et al*., 2019; Knight *et al*., 2000; Cronk, 1996). Furthermore, the closed loop design of the constructed wetlands increases the validity in the use of CW's in treating livestock wastewaters and do not negatively impact on the environment. Further use of CW's is the potential of utilising livestock waste as an energy product (Kinyua *et al.*, 2016), the recycling of nutrients (Petersen *et al*., 2007) and in drought stressed regions, the reuse of water (Scott *et al*., 2004). The rural location of the majority of livestock activity and the fact that wastewater is generated by these enterprises, CWs provide an appropriate substrate for further biological activity (Cantrell *et al*., 2008). In the mid 1990's in the Republic of Ireland, a wetland concept that integrated land and water management, respecting the local ecosystem and providing an aesthetic wetland system encompassing shallow-vegetated wetlands in the management of farmyard dirty water was realised (Harrington and Ryder, 2002). This approach was termed the 'Integrated Constructed Wetland' (ICW) concept (Harrington and Ryder, 2002). The fundamental objective of thw ICW is the sustainable and holistic management of a diversity of wastewater, including livestock wastewater with the associated land and water resources, adhering to the principles of adaptive management (Harrington *et al*., 2005).

The main factors that are key when considering the successful application of a constructed wetland treating livestock wastewater (Carty *et al*., 2008) are the Ammonium-N concentration of the influent waste entering the wetland, its effective removal through nitrification and de-nitrification and finally phosphorous capture and retention within the wetland. These inorganic ions are generally considered to be fundamental parameters in wetland functionality (Reddy *et al*., 1999; Guimaraes *et al*., 2016) and their interaction with local soils. The soil and inorganic ion interactions are key factors in the ICW concept and are critical with regard to livestock wastewater treatment and management.

1.2 Livestock wastewater and nutrients

Traditional methods of treating livestock wastewater involve waste retention using lagoons[1] and after a time, land spreading on fields. Due to high volume and increased production rates, these technologies are no longer sustainable (Cressie and Majure, 1997). Consideration for livestock wastewater management as in the context of climate change is of special relevance (Harrington and McInnes, 2009). Within the Irish context, rainfall patterns and variations, predicted by climate change are likely to increase volumes of run-off entering collection systems, such as integrated constructed wetlands (ICWs) (Department of the Environment, Heritage and Local Government, Ireland, 2010 a). A sustainable and effective constructed wetland system in treating livestock wastewaters must recognise these changing conditions (Harrington and McInnes, 2009) and other important parameters must be considered such as ammonium and phosphorous entering the constructed wetlands.

Ammonium-N can result in toxicity within the wetland system effecting the *in-situ* plants (Hill *et al*., 1997; Britto and Kronzuker, 2002; Zhao *et al*., 2016). Li *et al*., (2011) tested different plant species under varying levels of ammonia toxicity within a constructed wetland to ascertain which plant species could tolerate the highest ammonia toxicity level, with *Zizania latifolia* tolerating toxicity levels up to 300mg/L. Within the ICWs, the Ammonium-N entering the wetland needs to be controlled, to ensure the toxicity does not negatively impact the wetlands vegetation. With respect to wetland emergent plant species such as *Carex, Typha, Phragmites, Glyceria* the recommended Ammonium-N concentrations are <100mg/L for wastewaters emanating from the meat industry (Harrington, 2005). This is an optimum operating concentration for most ICW systems, input variations and pulses in the range of 280mg/L maybe tolerated for short periods (Harrington, 2005). Internal anaerobic wetland conditions may reduce the potential negative impact of these input variations and conditions. Despite the nitrogen content of the livestock wastewater being a toxic pollutant

[1] A lagoon is an active treatment pond/tank with the application of chemical, physical and biological processes.

in large concentrations, at lower concentrations it is a valuable added resource within ICW systems. Therefore, appropriate control and management of the influent wastewater into the wetland is of importance (Cronk, 1996; Harrington, 2005; van der Valk, 2006; Murray-Gulde *et al*., 2008; Kumwimba *et al*., 2018).
Phosphorous retention is maintained within the constructed wetlands through plant uptake, sorption onto soil particulates and organic litter deposition over time (Kadlec and Knight, 1996) and wetlands have developed a complex system of retaining phosphorous (Reddy *et al*., 1999). The emergent vegetation provides the organic litter through decomposition within ICWs, along with the influent. This organic matter holds the phosphorous, so long as decomposition is inhibited. The inhibition is secured through anaerobic conditions. The management of adequate water levels within individual ponds is critical in maintaining optimium water depths to enable vegetative growth and ensuring correct flows within the ponds. Correct design of an ICWs allows for adequate organic matter build up, typically 2-4cm/annum in the initial wetland ponds (Scholz *et al.*, 2007a). The Dunhill-Annestown project in County Waterford, Ireland continues to operate without intervention, accumulating a sustainable phosphorous, organic and mineral material that could be utilised in the local agricultural environment (Harrington and McInnes, 2009).

1.3 General design of constructed wetlands: an overview

The two main types of CW are: (1) Surface Flow (Free Water Surface (FWS) flow or Integrated Constructed Wetland) and (2) Subsurface Flow (SSF) CWs (Kadlec and Knight, 1996, Kadlec *et al*., 2000; Haberl *et al*., 2003). The surface flow (SF) are densely vegetated systems that also include open water surfaces that closely mimic natural wetlands. They require more land than subsurface flow wetlands and are generally easier and cheaper to design and build in comparison to subsurface flow CWs (Van Deun and Van Dyck, 1994), see Fig 1.1. The term Integrated Constructed Wetland (ICW) was defined by (Harrington and Ryder, 2002), as an unlined free surface flow constructed wetland, cleansing and managing water flow from farmyards and other wastewater sources.
In a FWS system the wastewater flows into the wetland via an inlet source passing through several wetland lagoons (cells) to an outlet point and in some cases the

wastewater may be lost to either to evapotranspiration or infiltration within the wetland system (Knight *et al*., 1999; Scholz *et al*., 2007a). FWS wetlands are usually not preferred in cold climates as they tend to freeze over in winter, which results in lower contaminant removal levels. In general, an FWS CW's performance is lower during the winter months because the rates of chemical and biological reactions slow, as the water temperature lowers (ITRC,2003 a). For example: in a municipal wastewater polishing system using a free water surface wetland, it was reported that a decrease in recognisable denitrification[2] in wastewater, though a dramatic reduction in biological activity is observed during winter months (ITRC, 2003 b).

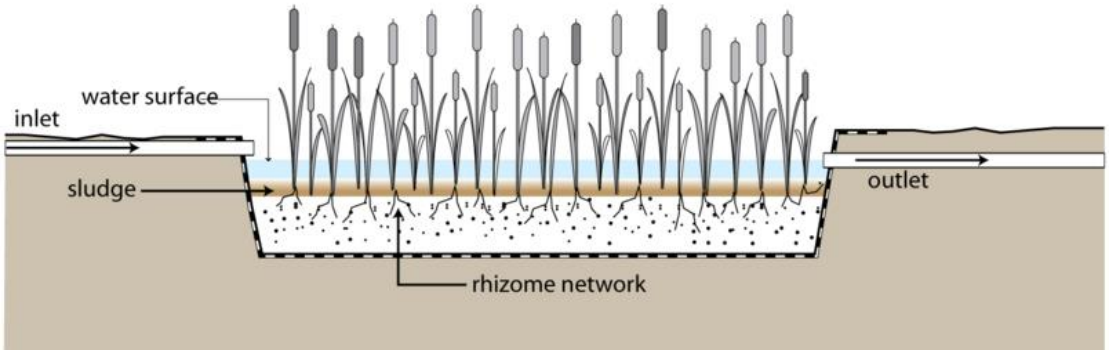


Figure 1.1: Typical configuration of a free water surface wetland (after Sandec and Eawag, 2009 a)

In the subsurface flow (SSF) constructed wetland, no free water level is visible. SSF CWs are sub-divided into two types (1) Horizontal subsurface flow (HSSF), see Figure 1.2 and (2) Vertical subsurface flow constructed wetland (VSSF), see Figure 1.3, which depends on the direction of the water flow through the porous media (sand or gravel). One drawback to SSF CWs is that clogging can occur in the porous media and generally the use of SSF's are confined to mechanically pre-treated wastewater (i.e. low content of particulate matter). In comparison to surface flow wetland, the contact area of water with bacteria and substrate is much

[2] Denitrification is a microbial process that reduces oxidised forms of nitrogen such as of nitrate, (NO_3^-) or nitrite (NO_2^-) to nitrogen gas (N_2).

larger, which therefore decreases the land area requirement of SSF CWs (Vymazal *et al*., 1998; Haberl *et al*., 2003) while improving the removal efficiency of the CW system. Sub-surface flow (SSF) CWs are often called anaerobic wetlands since treatment occurs in the lower layers of the wetland where anoxic (no oxygen) conditions are present (ITRC, 2003a). The SSF wetlands are often constructed using a composite of organic and non-organic materials that may include wood chips, compost, manure, hay, limestone and/or a bacterial inoculum, which are typically indigenous bacteria that can survive in these anoxic conditions.

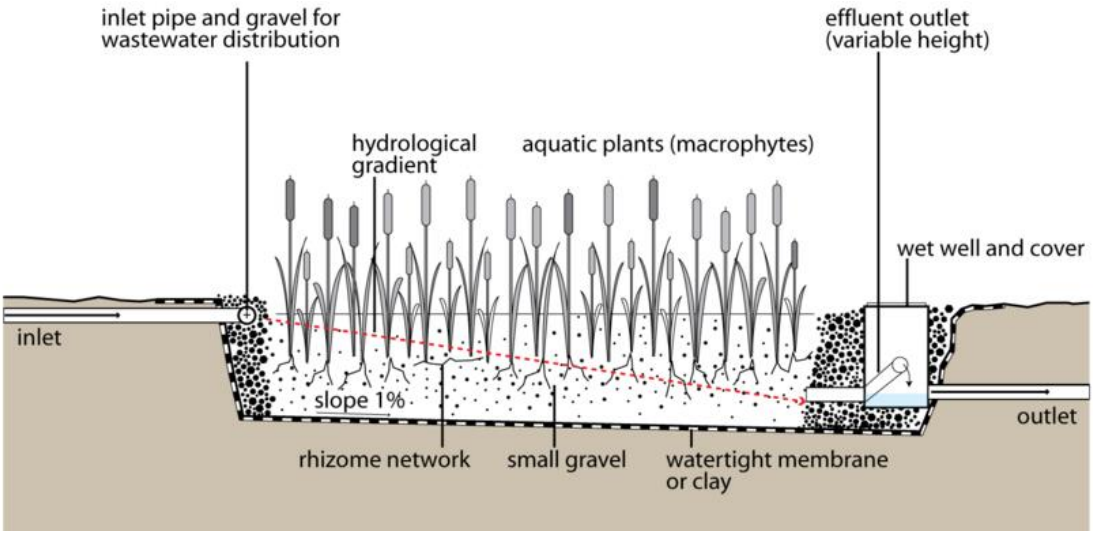


Figure 1.2**:** Typical configuration of a horizontal subsurface flow constructed wetland (after Sandec and Eawag, 2009b)

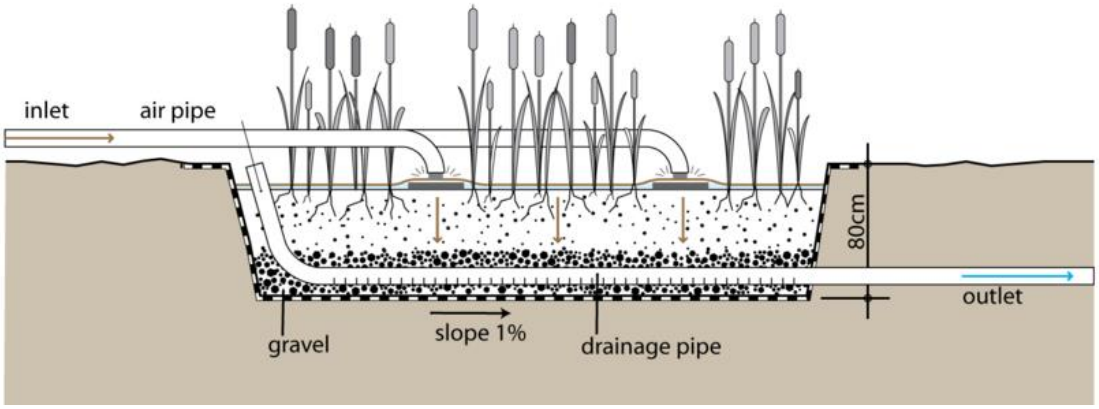


Figure 1.3: Typical configuration of a vertical subsurface flow constructed wetland (after Sandec and Eawag, 2009c)

Below are general wetland design strategies, with respect to vegetated surface flow wetlands.

(1) The quality of the influent wastewaters is not acutely toxic so as to be lethal, to maintain macrophytic vegetation.
(2) The hydraulic retention time and flow-through velocity enable capture and retention of nutrients.
(3) The wetland infrastructure can maintain adequate water depth to maintain hydric soil conditions., i.e. the wetland soil is permanently or seasonally saturated with water, to maintain anaerobic conditions in the wetland system.

Considering the above third point, it should be noted that biological clogging of soil pores will tend to decrease hydraulic conductivity (Ragusa *et al*., 1994; Magesan *et al*., 2000). Hydraulic conductivity in relation to wetlands is a property of plants, soil and rocks, that describes the ease with which water can move through plant material and soils. Correct management of the wetland must be applied to ensure the hydraulic conductivity will not be greatly impacted. To reduce the potential of clogging and to maintain optimum hydraulic conductivity, the use of sediment traps can be applied (Ellis *et al*., 2003) to reduce the amount of particulates entering the wetland.

A constructed wetland must be sized and configured to intercept and adequately treat the influent wastewater, having adequate hydraulic retention time (HRT) and appropriate velocity flow, to enable phosphorous capture and retention with suitable anaerobic-aerobic conditions for the transformation of nitrogen (Kadlec and Knight, 1996). Given the nature of abattoir wastewater, particularly its high suspended solid content and nutrient loadings, it would seem that surface flow wetlands systems have an advantage in managing the suspened solids and nurients loadings as theses sytems have less operational demand and greater capacities for more comprehensive treatment, in particular the retention and removal of phosphorous (Kadlec and Knight, 1996).

By segmenting the constructed wetland into a sequenced number of cells (ponds), water flows are easily managed and controlled. With careful management during the design phase of the wetland each pond can develop its own plant and animal

communities adding considerable biodiversity and aesthetic values without impacting wastewater treatment.

The design and size of a wetland for treating livestock wastewater is largely determined by volumetric flow through the individual ponds (Harrington and McInnes, 2009).

1.4 Ecosystem services

Ecosystem services are the multi-faceted and varied benefits that humans freely gain from indirect and direct contributions from the natural environment, leading to an improved quality of life. In Europe, wetlands are of high importance, interacting and controlling floods, water flow, water quality and cost-effectiveness in water treatment https://biodiversity.europa.eu/topics/ecosystem-services [accessed 03 2019]. Wetlands are among the most valuable ecosystems on the planet (Mitsch and Gosselink, 2015). Constanza *at al*., (1997) used ecosystem estimators that highlighted the importance of freshwater wetlands and these wetlands have a disproportionately high value in comparison with lakes, rivers, forests and grasslands. The main ecosystem high value estimators that highlight the wetlands importance are (i) waste treatment, (ii) water supply and (iii) disturbance. Points (i) and (ii) are connected, in that wetlands utilise large volumes of water and treat the waters as they pass through them and are therefore susceptible to climatic changes. The third point, disturbance or ecological disturbance emphasises that wetlands are susceptible to ecological events such as flooding, drought, loss of habitat and vegetation (Constanza *at al*., 1997).

Establishing a societal empathy and appreciation for wetland systems and when best to create and enhance them, is fundamental to their use and protection (Boyer and Polasky, 2004). A fundamental aspiration of the ICW system is to deliver sustainable water management without comprising on water quality, while focusing on biodiversity and landscape fit. Good landscape fit is an important feature towards achieving a valuable ecosystem, while cognisant to functionality, robustness and aesthetic. Critical components to achieve good landscape fit include topography and use of vegetation that are contiguous to the countryside. Within the design of the ICW system it is important to adopt curvilinear rather than rectilinear designs and sizing ponds with the correct design and taking into

consideration landscape features such as trees and other local natural aesthetics into the landscape fit. The use of irregular shapes or curvilinear forms improves the overall aesthetic and has been shown to enhance the biodiversity of constructed wetlands (Hansson *et al*., 2005).

Wetland systems and the biodiversity that they support are under threat (Millennium Ecosystem Assessment, 2005). The loss of wetlands has further decreased the ability of ecosystems to filter and decompose wastewaters (Millennium Ecosystem Assessment, 2005). Some of the main features of ICWs are to improve biodiversity and improve wastewater quality. This is achieved through correct design, utilising local species that are associated with fresh water wetlands thereby improving the local environment and specifically receiving water bodies, such as local streams (Harrington and McInnes, 2009).

Wetlands can play a significant role in reducing the impacts of CO_2 (Pant *et al*., 2003), (Moomaw *et al*., 2018). Wetlands sequester carbon either through plant (phragmites) dominated wetland systems and that in the long term, they can become effective sinks for greenhouse gases (Brix *et al*., 2001). The CO_2 can also be sequestered in their soils and decomposed matter (Moomaw *et al*., 2018).

Wetlands have a significant role in reducing the risk of flooding (Sather and Smith, 1984; Bullock and Acreman, 2003; Narayan *et al*., 2017; Leon *et al*., 2018). Using small wetlands in series adjacent to a river, is better in attenuating flooding in comparison to a single large wetland system downstream (Knight, 1992). Bradley *et al*., 2006, monitored small floodplain wetlands along the Lambourn river in the UK to model the hydrodynamics of water flow into the river and floodplains wetlands systems. Indicating the potential flood plain wetland system to mitigate against flooding. As the climate changes, the role of wetlands in reducing the risk of flooding could be increasingly important (Stern, 2006). Furthermore, the creation and restoration of wetlands has the potential to deliver a wide range of benefits to society (Silva *et al*., 2007; McInnes and Alexander, 2013), such as improved biodiversity and water quality.

Within the Irish context (Flood, 2012) reviews the potential economic impacts with respect to climate change and how Irish wetlands are vulnerable to climate change and how the Irish state must adapt to future climatic events. An essential concept to the ICW system is the need to maximise a range of benefits while

maintaining optimal treatment of livestock wastewater (Harrington and McInnes, 2009).

1.5 The concept of an integrated constructed wetland

The initial Integrated Constructed Wetland (ICW) concept was developed 1990's by Dr. Rory Harrington in collaboration with Dr. Miklas Scholz and Paul Carroll, and the concept published in 2007 (Scholz, et al., 2007a).

The definition of an ICW is as follows: a free surface flow constructed wetland which is unlined and can be used, for example, by farmyards and other wastewater sources to remove pollutants whilst controlling water flow. The design is such that the wetland infrastructure is integrated into the landscape while increasing biodiversity (Scholz et al., 2007a, b). The ICW concept has a number of objectives (1) aesthetically fits with the local environment (2) the wastewaters entering the ICW are treated by plants, that can adapt to local soils and (3) the biodiversity is improved. The ICW conforms to a sustainable system, not readily evident in other wetland treatment systems (Mustafa *et al*., 2009; Scholz *et al*., 2007a, b). The ICW design utilises larger land area with increased hydraulic residence time to provide optimum pollutant removal. As a result of the success of the ICW concept and design; the model has spread across Ireland, Northern Ireland and Scotland with the arrival of a subtype of ICW called the Farm Constructed Wetland (FCW) (Carty et al., 2008).

This ICW concept should not be confused with another "integrated constructed wetland system", which utilises the word, "integrated" as consisting of various types of wetland systems in series i.e. hybrid wetland systems (Vymazal, 2005; and Xiong *et al*., 2011).

1.6 Case Study – Dunhill-Annestown constructed wetland ICW project.

The Dunhill-Annestown stream catchment ICW project, located in County Waterford, Ireland has been the pivotal experiment in the realisation of the ICW concept to treat both farm yard wastewater and municipal wastewater. The initial ICW systems were treating farm yard wastewaters in the Annestown area, the ICW was subsequently expanded to treat munipical wastewaters from the village

of Dunhill in 2012. Sixteen independent ICW systems were established within the area, covering an approximate area of 2,5000 ha.

The systems were monitored for a range of parameters including biological, chemical and environmental data (Harrington *et al*., 2005; Scholz *et al*., 2007a; Mustafa, 2009). The key components monitored within the ICW system are related to nutrient loadings such as total and soluble phosphorous and nitrogen compounds, mainly Ammonium-N. The monitoring has shown that three of the twelve systems with large length to width ratios (8:1) gave poor performance in treatment (Harrington and McInnes, 2009). In contrast those systems with a ratio of 1:2 achieved better water quality and improved biodiversity within the wetland systems (Carty *et al*., 2008). The wetland area of an ICW is a fundamental factor in efficiently removing nutrients from the wastewater. Increasing the ICW area has been shown to positively benefit removal efficiencies of both ammonia and phosphorous and with the application of correct design, the ICW concept provides a robust mechanism for the treatment of farmyard wastewater while protecting local aquatic environments (Harrington and McInnes, 2009). The Dunhill-Annestown ICW project has provided regulators with a viable option in the treatment of livestock wastewaters, where environmental protection and the catchment of harmful wastewaters is improved, while providing improved landscape aesthetics, enhanced biodiversity. One of the ICW systems in the catchment area, was constructed for the village of Dunhill in County Waterford, with the construction of five treatment ponds in 2012 to support a growing population (Irish Water, 2019), see Figure 1.4.

Figure 1.4: Dunhill ICW, Co Waterford, Ireland, Irish Water©.

1.7 Modelling wetland functions

Since the mid 1990's modelling strategies have been performed to simulate the internal processes occurring in constructed wetlands (CWs). CWs can be constructed to several designed based methodologies linked to water flow; which can be saturated/ unsaturated, vertical/horizontal and inclusive of all different combinations (Fonder and Headley, 2013). Where saturated/unsaturated refers to vertical/horizontal wetland systems whereby the lower part of the wetland is submerged underwater and the upper part of the wetland is dry or unsaturated.

The purpose of CWs are to remove and or reduce influent pollutants by a combination of physical, chemical and biological processes. The use of simple to complex modelling scenarios have been employed to understand these processes. In 2013, at the 5th International Symposium on Wetland Pollutant Dynamics and Control, (WETPOL 2013), a wide collection of simulation and modelling studies were presented. Within the context of these conference papers, dealing with modelling and simulations studies, the committee resolved these papers into three

main approaches. (A) Bio-kinetic models, (B) Process dedicated models and (C) Design support models (D) System Dynamic models (E) "soft computing"

(a) Bio-kinetic models such as Activated Sludge Models (ASMs), see Henze *et al*., (2000), are used for modelling vertical flow CWs. These systems are highly dynamic and complex due to the variability in flows and loadings within the CWs, Rousseau *et al*., (2004); Langergraber *et al*., (2009a). The HYDRUS® software package (Langergraber and Šimunek, 2005) which is based on the mathematical foundations of the ASMs (Henze *et al*., 2000), is a Windows based software program that can model wetland interactions in CW1-Dimesion (1D) CW2-D and CW3-D via water flow, heat and pollutant transport in soils.

(b) Process dedicated models are generally simpler than ASMs. They are based on simple kinetic interactions which model a single process related to the degradation/ transfer of a compound or related compounds (O_2, PO_4^{-x}, COD etc). The University of Salford, Civil Engineering Research group created a 1-D model for particulate transport in a vertical flow constructed wetland (VFCW) developed using COMSOL Multiphysics™, see Sani *et al*., (2013). This model was based on wastewater particulates with different sizes and composition. Using the model, the objective was to evaluate the particulates settling within a porous media. The model assumes a uniform water velocity throughout the saturated porous media.

(c) Design support models are based on the tank-in series (TIS), plug flow (PF) and dispersed flow modelling (DFM) regimes. These models are based on residence time distribution where (RTD) or hydraulic retention time (HRT), where a tracer is injected into the CW and subsequently detected at varying sample points and times. GPS-X® software is an example of the use modelling strategy, where the user can select a different hydraulic model such as TIS or PF regarding saturation conditions within biofilm development (Zeng *et al.,* 2013a). The initial phase determines the hydraulic variables such as number of tanks (ponds/lagoons) in series, average retention time, aspect ratio and dead volume. These variables are utilised in the GPX-S software to simulate the biodegradation processes (Zeng *et al.,* 2013b).

(d) System dynamic model is a design based on system thinking[3],i.e. mental models, that requires constructing "causal loop diagram" or "stock flow diagram" to form a system dynamics model for applications (Forrester, 1961, 1968). Constructing a system dynamincs model requires identifying the problem and developing a strategy to explain the cause of the problem within the construct of the model. The model is formed within a computer simulation program with simulation runs governed over time. (Chang *et al*., 2013). Common systems dynamic model programs used are PowerSim® and STELLA®. System dynamics has many applications, O'Regan and Moles, (2001a) used system dynamic modelling as a means of examining the effects of complex environmental, fiscal and corporate policies and developing investment decisions within the mining industry. System dynamics tools and techniques were used to understand the complex problem of the spread of Ebola virus (O'Regan and Moles, 2001b). System dynamics with multiple scenarios was used in tourism development study in Vietnam (Mai and Smith, 2018). Chang *et al*., (2013) used system dynamics with sensitivity analysis to model the efficacy of floating treatment wetlands in stormwater detention ponds.

(e) A fifth methodology using soft computing techniques should also be mentioned. This family of techniques are used to determine complex non-linear interactions occurring within the CWs. The majority of this thesis is dedicated to application of these methods, such as artificial neural networks (ANNs), fuzzy logic (FL), Bayesian belief networks (BBN) and self-organising maps (SOMs). Other techniques such as support vector machines (SVMs) and genetic programming (GP) which are not studied in this thesis, are also part of this family. Hart and Pollino (2009) developed a document for the Australian government with respect to utilising Bayesian modelling as a risk based environmental tool for water allocation, including rivers, floodplains and wetlands. Scholz et al., (2009) used SOMs as a predictive tool for an ICW treating agricultural runoff and Tsihrintzi *et al*., (2013) used fuzzy logic models to predict BOD and COD removal from a FWS constructed wetland. Mohammadpour *et al*., (2015) compared ANN and SVM in the prediction of water quality index in constructed

[3] Systems thinking is a holistic approach that focuses on the way that a system's parts interrelate and how systems work over time within the context of the whole system.

wetlands, with the outcome that SVMs were better equipped to predict an accurate water quality index (WQI). We will utilise Fuzzy logic, Bayesian belief networks, artificial neural networks and self-organising maps in this thesis.

1.8 Constructed wetland – black box

There are knowledge gaps, concerning the efficacy of constructed wetlands in treating wastewaters. Jahangir *et al*, (2016) investigated the knowledge gap in carbon and nitrogen dynamics in CWs treating wastewater. Kuschk *et al*., (2012) recognised that knowledge gaps in anaerobic microbiological processes in CWs and the emission of greenhouse gases within the rhizosphere[4] of wetlands is not well understood. With respect to bacterial populations interacting with soils, plants and BOD loading is still not well understood within CWs (Vacca *et al*., 2005; Faulwetter *et al*., 2009). Within the Irish context, there has been a significant increase in the use of CWs for wastewater treatment overt the last decade but there still remains knowledge gaps as to how efficiently they perform and how reliable and viable as a treatment option to treat domestic wastewater (O'Luanaigh and Gill, 2012). A further extension of the knowledge gaps within CWs, they are considered black boxes[5] and researchers agree that understanding the black-box nature of CW technology is difficult:

(i) Kadlec (in Cole, 1998): stated that wetlands are huge functioning mess and we do not have enough information about what happens inside them and we need to understand the internal processes.

(ii) Gearheart (in Cole 1998): stated that we know they work however we cannot model dynamic changes in the wetland system.

(iii) Langergraber *et al*., (2009b): stated that CWs are complex systems and often considered as "black boxes" and only little effort has been made to understand the main processes.

[4] Rhizosphere is the soil zone around the roots in which microbes present in the soil interact with the plant roots.

[5] Black box is a system which can be viewed in terms of inputs and outputs without any knowledge of internal workings.

The overall objective of this thesis was to develop multi-level mathematical/statistical methodologies to bridge the knowledge gap that exists in understanding the internal signals of wetland system i.e. to understand the black-box. In this, a more detailed picture can emerge concerning the reliability and resilience and thus the failure/s within CWs to potentially develop a strategy in the effective treatment of livestock wastewaters.

A global, socio-economic and environmental strategy is required in the management of livestock wastewaters (Chadwick *et al*., 2008) and a joined up strategy to address the issues of both agriculture point and diffuse pollutions sources is required (USEPA, 2008). Within the Irish context, little information exists regarding the use of CW's to treat high strength industrial waste such as abattoir effluent. In comparison, countries such as Canada, Australia, New Zealand and the USA have standard limits for slaughterhouse and meat processing plants wastewater discharges (Bustillo-Lecompte and Mehrvar, 2017). A disconnect exists between current Best Available Technologies (BAT) documents, which make little mention of the uses of CWs in the treatment of abattoir waste and the variables that are monitored within the BAT, no record of bacteria loadings are required. According to the Water Framework Directive (WFD, 2015), water bodies are classified on ecological and biological status. The European Environmental Agency (Nixon *et al*., 2003) classified Irelands water quality as poor with respect to biological performance. The document highlights the significant areas of pollution of water by nitrates, phosphates, pesticides and pathogens; habitat degradation and over-abstraction of water for irrigation.

1.9 Regulatory framework

Abattoir effluent is considered one of the most extreme of industrial waste streams for treatment. Animal by-products are classified in three different categories according to European regulation (CE) 1774/2002 which entered into effect in 2003. Class 1 concerns products suspected of carrying infectious diseases, as well as contaminated by illicit or dangerous substances. Cadavers from wild, domesticated or zoo animals also enter this category. Class 2 pertains to farm animals found dead or containing medicinal residues. Class 3 is attributed to animal remnants originating from abattoir processing of human food grade

certified live animals. This class can be valorized[6] (i.e. enhanced by artificial means), usually by separating fats from lean components, and producing combustible lipids and meat and bone meal (MBM) for feed or fertilizer. Class 2 products can also be valorized, but not for animal feed, whereas class 3 must be destroyed or co-incinerated (Sharrock *et al.*, 2009). The World Bank conducted a global study into livestock markets, slaughterhouses and related waste management systems (World Bank, 2009) and reported that wetlands were a treatment option chosen by meat processing plants in high-income countries, but as a secondary biological treatment. A final tertiary treatment option such as chlorination and/or ultraviolet irradiation is recommended. There is a need by industries in developed nations to adopt cost-intensive treatment methods including phytoremediation, land treatment and constructed wetlands (Javaid *et al*., 2015a). The authors highlight that lagoons can be classified as tertiary treatment, and that wastewater discharges from slaughterhouse without proper disinfection, leads to the occurrence of high populations of meat based bacteria such as *E.coli* and *Salmonella* (Javaid *et al*., 2015b). Therefore, it is important that key wetland parameters are monitored to ensure that local streams and rivers are not polluted with effluent from animal by-products, as Ireland has a history of poor compliance with EU regulation with respect to water quality (Nixon *et al*., 2003). The water framework directive (WFD), managed by the Irish EPA (EPA, 2009), provides a legal framework for the regulation of the biological status of Irish waters as of 2008.

A report published by the Department of the Environment, Heritage and Local Government, (2010 a) which is the first national guidance document to deal with integrated constructed wetlands (ICWs) in Ireland, farmyard and domestic wastewaters, makes reference to a Discharge Licence Requirement 4.6 (ii). Whereby an Integrated Pollution Prevention and Control (IPPC) licence issued by the Environmental Protection Agency, if the ICW is associated with an activity listed in the First Schedule of the Environmental Protection Agency Acts 1992 to 2007, e.g. Slaughter Plant, Dairy Plant; (Department of the Environment, Heritage and Local Government, 2010 b). Therefore, the discharge licence comes

[6] Valorized waste – the process of converting waste materials into other products including materials, chemicals and fuels.

under the remit of the Environmental Protection agency (EPA), not the Local Authority (LA). Two other Best Available Technology (BAT) papers, located on the EPA websites are (1) IPPC reference document on BAT in the slaughterhouse and Animal By-Products Industries (European Commission Joint Research Centre, 2005), where the report considers slaughterhouse wastewater treatment in the confines of tertiary treatments, such as filtration, e.g. using sand filters, reed beds, coagulation, or precipitation are sometimes used as a final cleaning step for the treated effluent, to reduce the BOD and suspended solid, prior to discharge to a water course.

(2) The Agency BAT[7] Guidance Notes on Best Available Techniques, see Table 1.1, for the Slaughtering Sector (EPA, 2008), where discharges to water are in compliance with the EU Water Framework Directive (2000/60/EC). The BAT and WFD reports relate to discharges to water bodies, not inclusive of constructed wetlands as primary treatments systems, with the European Commission Joint Research Centre (2005), report considers that CWs are a tertiary treatment option not a primary treatment option.

Table 1.1: BAT Associated Emission Limit Values for discharges to water

Constituent Group or Parameter	Emission Level
pH	[6-9]
Number of Toxicity Units (TU)	5
BOD	>90% removal or 20-40mg/l
COD	>75% removal or125-250mg/l
Suspended Solids	60mg/l
Total Ammonia (as N)	10mg/l
Total Nitrogen (as N)	>80% removal or 15-40mg/l
Total Phosphorous (as P)	>80% removal or 2-5mg/l
Oils, Fat and Grease	10-15mg/l
Mineral Oil (from interceptor)	20mg/l
Mineral Oil (from biological treatment)	20mg/l
Other	(-)

Source EU Water Framework Directive (2000/60/EC)

[7] BAT – Best available technology/techniques – is a legal discharge standard for a process approved by authorities. For example; wastewater from a treatment process entering a local stream.

There currently exists no BAT in relation to ICWs treating domestic or industrial waste streams. However, Babatunde *et al*., (2008) has shown that constructed wetlands are now used throughout Ireland as an alternative cost-effective wastewater treatment system for domestic and industrial effluents, see Figure 1.5 and 1.6.

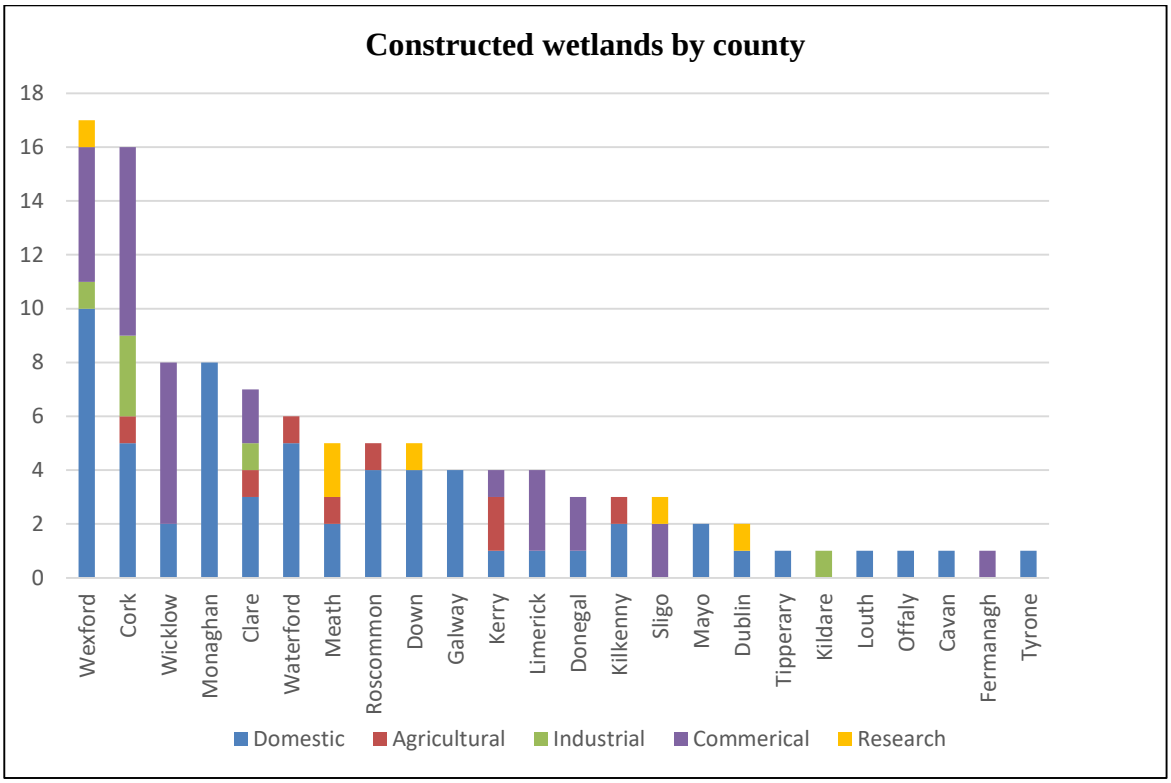


Figure 1.5: Different types of wetlands in Ireland Babatunde *et al*., (2008).

Figure 1.5 shows a breakdown of types of wetlands currently in use in Ireland where surface flow and free water systems dominate the Ireland constructed wetland landscape. It should be noted that surface flow (SF) and free water surface (FWS) wetlands are essentially the same. One possible view on the usage of these terms could depend on the individual county councils deciding on either term within the contents of local authority documents. The reason for the dominance of these FWS and SF wetlands within Ireland are that they are cheaper and easier to design and build and have an aesthetic value and generate a wildlife habitat. The EPA STRIVE Report N0.45 by Zhan *et al.*, (2010), highlights that the constructed wetland use and research in Ireland was increasing in particular between 2003-2009. Another finding from the EPA STRIVE Report N0.45 by

Zhan *et al.*, (2010) is that there has been very little published of either research findings or of treatment performances of operating CWs in Ireland.

The types of CWs and wastewaters treated in Ireland

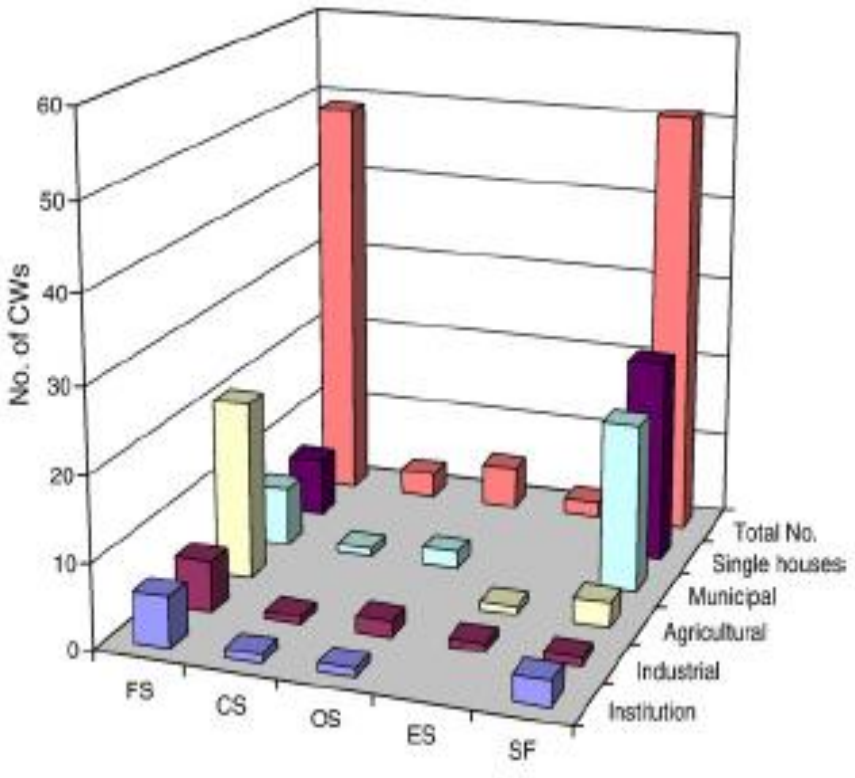


Figure 1.6: The types of CWs and wastewaters treated in Ireland (FS – free water systems; CS- combined systems; OS – other systems such as sludge drying systems and stream filtrations basins; ES- experimental systems; SF –surface flow systems, Babatunde *et al*., (2008).

The Water Framework Directive (WFD, 2015) becomes an important issue due to the fact the wetland surface waters enter may a local stream. The WFD is a key European Union initiative aimed at improving water quality throughout the European member states and applies to rivers, lakes, groundwater and coastal waters. The WFD requires an integrated approach with the aim of maintaining and improving water quality. Within the Irish context, eight River Based Districts (RBDs) have been identified in Ireland for the purpose of implementing the Directive; three of these are shared with Northern Ireland (Shannon, Neagh Bann

and North Western), four RBS are within the state (Eastern, South Eastern, South Western and North Western) and a final RBD is wholly within Northern Ireland (North Eastern). The RB of interest with regard to the ICW within this thesis empties into the South Eastern Region River waters. One of the main variables of interest within the WFD is the monitoring of nitrate concentrations within all the river basin districts. Elevated nitrate concentration in groundwater is an issue, particularly in the South East and South of the country. It may contribute to eutrophication of surface waters and affect drinking waters (McGarrigle *et al*., 2010). One of the main sources of the diffuse release of nitrate to rivers and streams are current intensive agricultural practices.

Many water bodies will require remedial measures to meet the objectives of the Water Framework Directive (WFD). Protection and restoration of high quality waters will also be a significant challenge. Eight water management issues have been identified as being of national importance:

1. Wastewater and industrial discharges
2. Landfills, quarries, mines and contaminated lands
3. Agriculture
4. Waste from unsewered properties
5. Forestry
6. Usage and discharge of dangerous substances
7. Physical modifications to surface waters
8. Abstractions

Table 1.2: WFD- River Quality Index

Ecological class	Number of Water Bodies
High	173 (9%)
Good	738 (40%)
Moderate	509 (28%)
Poor	389 (21%)
Bad	41 (2%)

Table 1.2 Ecological status based on the Water Frame Directive Interim Biological Classification for river water quality (EPA, 2008).

A simple and effective methodology to monitor the constructed wetlands performance, is to adopt the BAT and BATNEEC[8] systems under use for wastewater treatment plants (WWTPs) in respect to wastewater emanating from a WWTP's into a local river of stream.

1.10 Research aim and objectives

The objective of this thesis was to develop multi-level mathematical/statistical methodologies to bridge the knowledge gap that exists in understanding the internal behaviours of a constructed wetland system, treating abattoir wastewater. Traditional wetland modelling has treated these systems as a "black box", which is a system, that can be viewed in terms of its inputs and outputs. The core objective of this thesis is to explore into the internal operations of the constructed wetland. Specifically, to:

1. Find the constructed wetlands dominant variables through the use of statistical analysis.

[8] BATNEEC – best available technology not exceeding costs

2. Improve the traditional wetland models and using the model improvements, to ascertain the hydraulic retention time of the constructed wetland using indicator bacteria.
3. Evaluate the air-borne pathogenic risk within the constructed wetland.
4. Highlight the importance of the wetlands soil in maintaining optimal performance.
5. Investigate whether the constructed wetland can be considered a complex system.
6. Evaluate the wetlands treatment processes i.e. consistency, in treating wastewater using the indicator bacteria as the key variables.
7. Develop hybrid Fuzzy Bayesian models to ascertain the wetlands overall robustness/ resilience in treating abattoir wastewaters.

1.11 Thesis overview

The use of different statistical and mathematical modelling methodologies were utilised in the main study of this thesis, with particular focus on indicator bacteria (total coliforms, *E.coli* and enterococci). Other wetland parameters were analysed such as inorganic ions including nitrate, phosphate, ammonia and sulphate analysed; physical-chemical parameters such as redox potential, dissolved oxygen, turbidity and pH of importance. A brief overview of the statistical and mathematical modelling chapters is provided below:

- Chapter 3, an un-biased statistical modelling approach was investigated using principle component analysis (PCA) in deciding which variables dominate the wetland system. In addition, discriminant analysis (DA) was investigated which assess the relationship between variables, by pre-selecting wetland variables that dominate from the PCA analysis. The analysis performed on PCA in this chapter, was further utilised in the fuzzy indexing analysis section, see chapter 10.
- Chapter 4, the use of traditional wetland modelling, such as plug-flow (PF), dispersed flow model (DFM) and continuous stirred reactors (CSTR) models was explored, using indicator bacteria as tracers within the constructed wetland. A

new dispersed flow model was developed, and the hydraulic retention times (HRTs) for the CW was evaluated using these models.

- Chapter 5 evaluates the air-borne bacteria present around the CW, sample points included the dissolved air floatation (DAF) plant, pond 6 and pond 12. The surface-air-system (SAS) unit was used to monitor the bio-aerosols present.
- Chapter 6, Bayesian belief networks analysis and sensitivity analysis was performed on the wetland soils, recording soil data such as porosity, soil indicator bacteria and nutrients.
- Chapter 7 discusses whether the CW was a complex system through the use of self-organised criticality (SOC) in conjunction with Shannon Entropy, to further elucidate the loading effects i.e. the influent wastewater from the DAF plant, on the wetland system.
- Chapter 8, a combined self-organised maps (SOM) or Kohonen Maps and Artificial Neural Network (ANN) analysis was performed. The SOM provides an unbiased data visualisation and cluster analysis of all the wetland variables and the ANN models are biased due to learning, testing and optimisation of these models.
- Chapter 9, a coefficient of reliability (COR) analysis was carried out. The variables analysed were the indicator bacterial concentrations (total coliforms, *E.coli* and enterococci). The COR method can be utilised by wetland designers and operators in evaluating constructed wetland treatment process based on monthly sampling of ponds over a period of time.
- Chapter 10, combines PCA, fuzzy logic and Bayesian belief networks to understand the robustness/ resilience of the CW in treating the majority of the wetland variables in two fuzzy Bayesian models. A combined fuzzy indexing (FI) and Bayesian belief network (BBN) decentralised based complex model was developed. Dominant variables are extracted using principal components analysis (PCA) are entered into a fuzzy indexing (FI) system to evaluate the performance of six sample points within the wetland (DAF, pond 1, pond 6, pond 9, pond 12 and stream). These performance values (0 = very poor; 100 = Excellent), provide a single value performance for the sample point i.e. an Index value. These values are evaluated for the six sample points over the duration of the sampling period. A probabilistic analysis was then carried out on the sample points fuzzy indexing

and on the wetland variables and a two decentralised based Bayesian belief models were created. This complex model allowed uncertainty and sensitivity analysis to be employed to reveal behaviours within the wetland system.

Chapter 2. Material and Methods

2.1 Constructed wetland analysis

Monthly surface water grab samples were retrieved from February 2007 – February 2008. The six sites sampled were the (input) dissolved air floatation (DAF) tank and ponds 1,6,9 and 12 and the (output) local stream. Approx 500ml of surface water samples were taken from the sample sitesand processed for different physical-chemical readings, bacterial and inorganic ion concentrations.

2.1.1 Visual and odour observations of the water grab sample's

Visual review of the sample sites taken from Goggle Map® and the colour of the liquid samples retrieved can be seen in Figure 2.1 and Figure 2.2 respectively.

- The DAF sample extract was a deep blood red in colour (de-oxygenated abattoir wastewater) with strong hydrogen sulphide (rotten eggs) and thiol ester odours present (sickly sweet smell).
- Pond 1 sample extract was black/grey in colour (anaerobic wastewater) with strong hydrogen sulphide odour.
- Pond 6 sample extract was dark yellow in colour with no odour.
- Pond 9 and Pond 12 extracts were light yellow in colour, with no odour.
- Stream sample had no colour or smell.

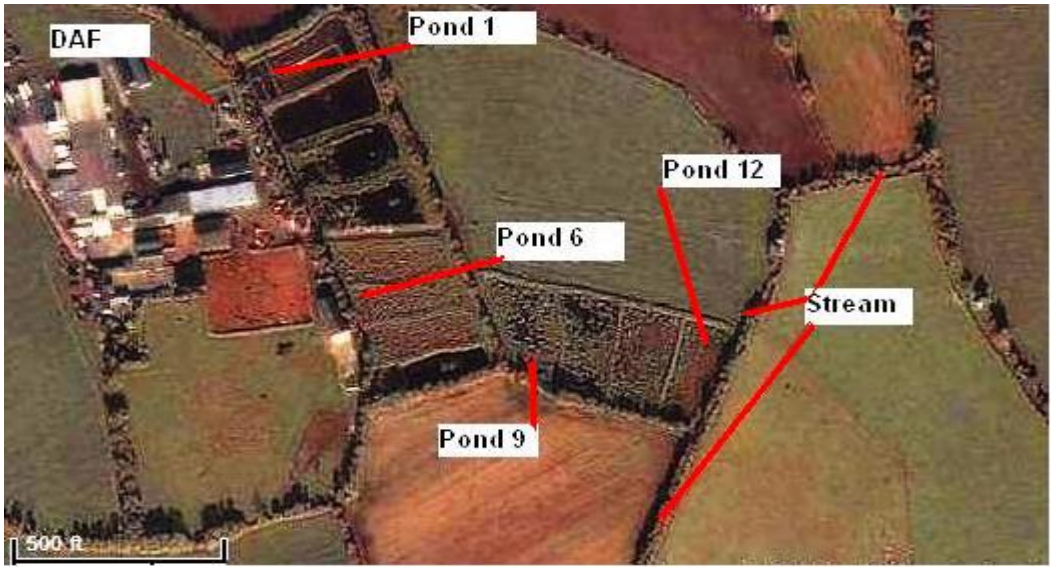


Figure 2.1: The Goggle Map® (April, 2008) of the constructed wetland. The highlighted points on the map, indicate where the surface water samples were taken for laboratory analysis and on-site analysis.

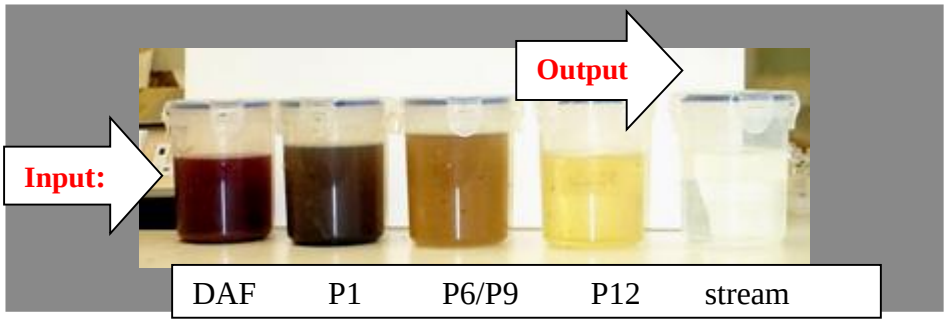


Figure 2.2: The colour of the liquid samples from the constructed wetland. Input (DAF) through to the different sample points pond 1 (P1), pond 6 (P6), pond 9, pond 12 (P12) and the output (stream).

2.2 Analytical equipment used in this study

Several instruments were brought on-site to the constructed wetland and used to record the different variables. These instruments were pre-calibrated in the laboratory prior, to recording surface waters values within the wetland. The instruments were:

1) Hanna Portable Water Meter [Temperature (^{0}C), Redox Potential (mV)]
2) Hanna HI 93703 Turbidity Meter [Turbidity (NTU)]
3) Eutech CyberScan DO 110 Meter records dissolved oxygen (mg/L) but did not require calibration as they were purchased from Eutech® fully calibrated.
4) Surface air Sample (SAS) Super 90 was brought to record bio-aerosols (CFU/L) but did not require calibration as the system had been calibrated (Nov 2007) by Mason Technology, Dublin.

The remaining instruments (off site): Dionex® 120 ion chromatography, Jenway 3510 pH Meter, Jenway 4510 Conductivity Meter were all calibrated prior to running samples as per manufacture's specifications (see Table 2.1).

Table 2.1: Analytical instruments used in this study

Instrument/Method	On site Measurements
Hanna Portable water meter	Temperature
Hanna Portable water meter	Redox Potential
Eutech CyberScan DO 110	Dissolved Oxygen
Hanna HI 93703 Turbidity meter	Turbidity
SAS Super 90	Bio-aerosols

Instrument/Method	Off site/ Laboratory Measurements
Dionex 120	Inorganic Ions
Jenway 3510 Meter	pH
Jenway 4510 Meter	Conductivity
IDEXX Quanti-Tray 2000	Bacterial Concentrations
APHA 1998	BOD

2.22 Preparation of wetland surface water samples for Ion Chromatography analysis

Surface water samples were carefully taken from specific points in the wetland to ensure a low level of particualtes in 1L sterile plastic containers which were placed carefully approximately 2-3cm onto the surface of the water to be sampled. The water was allowed to flow into the container. Once the container was filled, it was removed very slowly, ensuring no disturbance of surface waters and reducing the amount of sediment disturbance.

The containers were brought back to the laboratory within 2 hrs and stored upright at 4^0C in biohazard bags. When perfroming analysis, 10ml of each sample was removed and passed through a Whattman 44 filter (150mm Ø), attached to a funnel and 20ml collection bottle, to remove particulates under gravity. With regard to Dissolved Air Floatation (DAF) samples and Pond 1 (anaerobic settling pond) the Whattman filtering process was repeated a second time due to the high turbidity of both these water samples.

The filtered water samples were then placed in cold storage at 4^0C until processed.

2.23 Bacterial determination of surface water samples using the IDEXX® Quanti-Tray 2000

Monthly grab samples (0.5L) were taken, from the wetland surface waters and transported to the laboratory and processed within 4 hours. Care was taken to remove only the top surface waters from the wetland approx. 1-5cm

Microbial analysis of total coliforms, *E.coli* and enterococci was carried out using the IDEXX MPN Quanti-Tray 2000 method (Technopath, Limerick Ireland) as per the manufacturer's specification. *E.coli* and total coliforms were quantified using the Colilert® Assay for water samples, 100ml of wetland surface water sample was mixed with the assay and incubated at 37^0C in quanti-trays for 24h as per manufacturer's instructions. For the determination of enterococci, the Enterolert® assay was used and the same procedure was repeated for the determination of enterococci bacteria within surface waters.

2.3 Physical and chemical surface water analysis

2.3.1 Determination of inorganic ions of water samples by ion chromatography

2.3.1.1 Power-up and pre-conditioning of the Dionex® 120 IC instrument

The Ion Chromatography system was powered up and left for 30 mins to warm up, during this time the system cycles the eluents (sodium carbonate 3.5mM / sodium bi-carbonate 1.0mM mixture for the anion and sulphuric acid 22mM for the cation) through the single column and de-gasses the system. The "warm up" time is critical in that it allows the system to cycle the eluents and remove trapped air bubbles, thereby ensuring accuracy of results.

After 30 mins double di-ionised water was injected onto the column, via a 2ml syringe through 0.45µm nylon filter to ensure particulates do not enter the column, the Dionex 120 system then takes a fraction (100µL) of the 2ml injected sample that runs through the column. This washing procedure removes any sample material that may have been left on the column. The Dionex 120® then processes the sample (double di-ionised water) through the system and the resultant ion concentration analysed to reveal any negative inorganic residues on the column. When no inorganics such as fluoride, chloride, nitrite, nitrate, phosphate and sulphate are present then the system is ready for multi-point calibration.

2.3.2 Preparation of inorganic ion calibration standards

1. The standards utilised during the calibrations process were purchased from Reagecon® and the standards were stored as per manufacture's specification.
2. There were three sample replicates injected and the mean ascertained for the determination of one sample value.

Negative ion calibration standards

Three 20ml multi-point standards were prepared for calibration purposes. Low calibration standard 0.1-1.0 ppm for fluoride, chloride, nitrite, nitrate, phosphate and sulphate; medium calibration standards 10-20ppm for fluoride, chloride, nitrite, nitrate, phosphate and sulphate and finally high calibration standards 50-70ppm, of chloride, nitrite, nitrate, phosphate and sulphate. All standards were made up using double di-ionised water and vigorously shaken to ensure complete mixing of liquid standard and double di-ionising water, using 50ml volumetric flasks. See Table 2.2.

Positive ion calibration standard

Three 20ml single point standards were prepared for calibration purposes. A low calibration standard of 10 ppm, a medium calibration standard 50 ppm and a high calibration standard of 150 ppm for ammonia. The standards are made up using double di-ionised water and vigorously shaken to ensure complete mixing of liquid standard and double di-ionising water, using 50ml volumetric flask. See Table 2.2

Table 2.2: Ion Chromatography standards

Ion charge	(-ve)	(-ve)	(-ve)	(-ve)	(-ve)	(-ve)	(+ve)
Ion type	Fluoride	Chloride	Nitrite	Nitrate	Phosphate	Sulphate	Ammonia
Low	0.1-1.0	0.1-1.1	0.1-1.2	0.1-1.3	0.1-1.4	0.1-1.5	10
Medium	Oct-20	Oct-20	Oct-20	Oct-20	Oct-20	Oct-20	50
High	50-70	50-70	50-70	50-70	50-70	50-70	150

Table 2.2 shows the ions used and their corresponding (ppm) for calibration on the Dionex® 120 Ion Chromatography.

Using the negative ion multipoint standards; starting with the most dilute standard, the liquid standard was extracted, using a 2ml syringe (Braun Omnifix ®), ensuring no bubbles are present during the extraction process, place a 0.45µm filter on top of the syringe prior to injection. The combined syringe and filter were placed onto the injection port and slowly inject the entire contents of the syringe onto the column. As the standard flows through the column, the user checks the

PC screen to ensure the standard contents are resolving in the correct sequence. The entire run takes 12 mins, starting with the first inorganic, fluoride and ending with the final inorganic sulphate.

After the completion of the sample run, the column was washed with double di-ionised to clean the column prior the next injection.

Repeat the above process for the medium and high concentrated standards and ensure that double di-ionised water injections are placed onto the column, verifying the column is clean prior to calibration runs.

Once all the standard runs were completed, the system is now ready to accept water samples from the wetland.

Note: One can only determine the exact mg/L or ppm value after the entire run was completed and the Dionex 120® prints off the run to reveal the standards' actual values.

2.4 Determination of inorganic ions from wetland surface waters

After the Dionex 120® was calibrated, 2ml aliquots of the wetland surface water samples were removed, ensuring no bubbles were present in the liquid. A 0.45µm nylon filter was placed on top of the syringe and the combined syringe/filter is depressed into the sample injection port.

After each sample run, double di-ionised water was passed through the column to remove the previous injected aliquot and thus reduce cross-contamination between samples. After each wetland surface water sample was processed. A print off reveals the ppm of the injected samples, allowing the user determine the negative inorganic concentrations per injected sample. Each run takes approx.12 mins, including both the sample analysis and the double di-ionised water run the clean the column.

2.4.1 Ion Chromatography Interferences

It should be noted that the Dionex-120 was set-up as a single column system, i.e. only one column is attached; either the negative ion column or the positive ion

column can only be attached at any one time. Therefore, when the negative ions were evaluated for the injected samples, the system was powered down and the positive column then placed on the system to determine positive inorganic ions. All samples must be injected through 0.45μm nylon filter in order to protect the column. On start-up, the system was allowed to warm up and de-gas in order to achieve best results. When the analysis is complete double distilled water injections must be run in order to remove any excess sample residue that may be caught on the column.

2.5 Biological Oxygen Demand (BOD)

The five-day BOD_5 test was carried out as per APHA 1998 specifications.

2.6 pH measurement

The pH of the wastewater was carried out as per APHA 1998 specifications.

2.7 Conductivity

The electrical conductivity of the wastewater carried out as per APHA 1998 specifications.

2.8 Temperature

The water temperature of the wastewater carried out as per APHA 1998 specifications.

2.9 Turbidity

2.9.1 Method of analysis: Turbidity Meter (Hanna HI 93703 Turbidity Meter)
The turbidity of the water is an expression of the amount of particles within the sample, therefore the larger the turbidity the more “cloudy” the sample and conversely the less turbid the sample the smaller the value. Turbidity is measured in NTU (Nephelometric Turbidity Units).

2.9.2 Turbidity Procedure

The turbidity meter was calibrated as per manufacturer's specifications.
The sample was placed in a 10ml glass vial and secured with a locking cap. Excess water was wiped from the outside of the 10ml vial to ensure accurate readings.
The vial was then placed into the meter, which is sitting level on the bench again to ensure accurate readings. The meter was initialised and after 5-8 seconds (depending on the turbidity of the sample) and value was displayed on the LCD screen. The process is repeated three times to ensure accuracy and the average recored.

2.9.3 Turbidity Interferences

Turbidity measurements can be affected in three ways. Calibration standards must be run prior to analysis and adjustments made as per manufacturer's specifications. The sample vial must be clean of excess liquid as this will result in erroneous results and finally the meter itself must be placed on a level surface in order to achieve best results.

2.10 Redox Potential (Hanna Portable Water Meter)

2.10.1 Redox potential

Redox reactions are defined as reactions in which electrons are transferred. The species receiving electrons is reduced and that donating electrons are oxidised. Redox reactions determine the mobility of many inorganic compounds as well as biologically important species such as nitrogen and sulphur. In addition, redox conditions govern the biological degradation of complex hydrocarbon contaminants. In the literature redox potential is reported as Eh, this is the potential generated between a platinum electrode and a standard hydrogen electrode when placed into the water being tested, where hydrogen is considered the reference electrode, and is measured in mV (millivolts). Measurement of redox potential in groundwater can be a difficult task, due to the complex nature

of ground water flow, the natural variability of concentrations of the ground water flow and the problem of obtaining a representative sample from the ground water flow. The common platinum electrode is not responsive to many reactions involving solid phases.

The varying redox potentials of constructed wetland surface waters

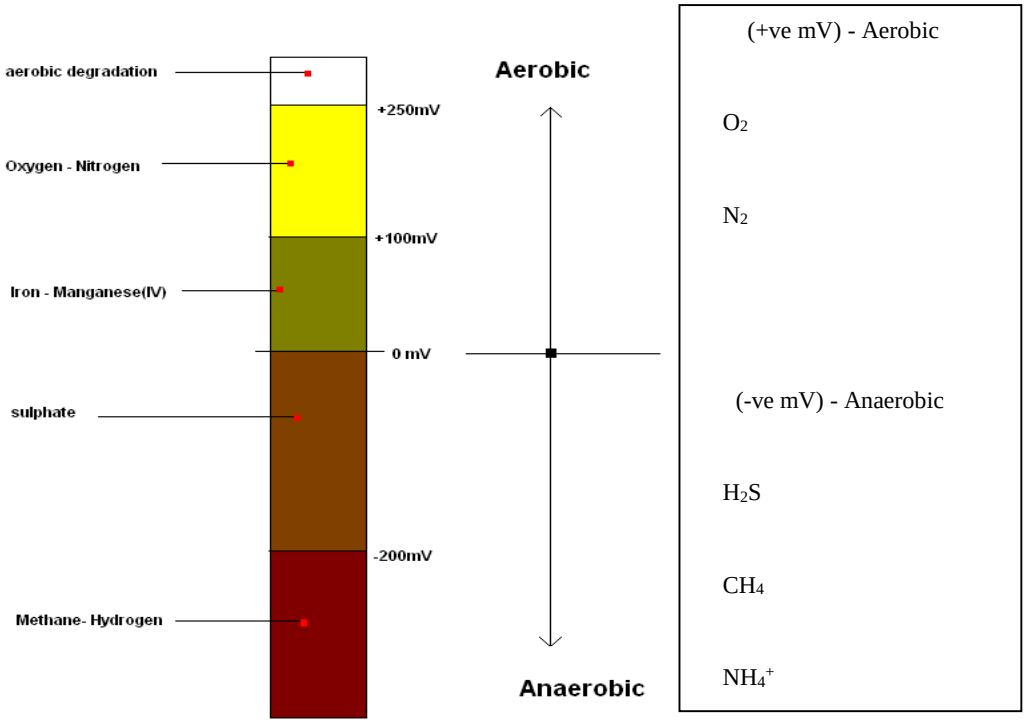


Figure 2.3: The varying redox potentials of constructed wetland surface waters. Also shown are the specific zones and corresponding redox potential values.

The reactions below are listed in order of decreasing Eh conditions as follows:

(1) Aerobic Degradation	+250mv and higher
(2) Oxygen – Nitrogen	+250mv to 100mv
(3 & 4) Iron - Manganese (IV)	+100mv to 0mv
(5) Sulphate	0mv to -200mv
(6) Methane - Hydrogen	-200mv and lower

Water in contact with air will have an Eh in the range of 350mv to 500mv. Microbiological mediated redox processes may decrease the redox potential to values as low as -300mv. These values were observed during the monitoring of the constructed wetland, in particular the DAF and pond 1.

2.10.2 Redox potential interferences

A cautionary note and suggestions for the obtaining the maximum value from the use of redox potential. The caution concerns the field measurement of Eh, it can be difficult to obtain representative samples and Eh measurements themselves may be unrepresentative of true Eh conditions due to the presence of multiple redox couples or general dis-equilibrium. The value of Eh resides primarily as an indicator of contaminant moving through a system (aquifer, stream, river or in this case, through a series of ponds). Evaluation for dissolved species generated from reactions 1 through 6 above provides direct evidence of Eh conditions as well as hydrocarbon degradation.

2.11 Dissolved oxygen (Eutech CyberScan DO 110)

2.11.1 Dissolved oxygen explained

A dissolved oxygen meter is used to measure the amount of oxygen present in a unit volume of water. This information indicates if the water is useful for a specific application like water treatment plants, sewage treatment works, river monitoring and fish farming. The dissolved oxygen sensor consists of a thin teflon diaphragm house inside the probe head.

2.11.2 Dissolved oxygen procedure

The probe was carefully dropped into the wetland surface water, ensure the diaphragm sensor is completely covered with water. The probe is submersed at a depth of 4-5cm below the water surface and the meter initated.
The D.O. meter then evaluates the dissolved oxygen of the water under analysis. The user must allow the meter to stabilise the reading and this may take between 10-25 seconds depending on the wetland water sample.
Once the LCD screen has displayed a constant dissolved oxygen value in mg/L, record the value. Wash the probe sensor with distilled water, especially across the diaphragm, if any particles are present on the sensor head, ensure that these are removed.

2.11.3 Dissolved oxygen interferences

The diaphragm is a thin membrane and is easily damaged if the probe is dropped or comes into contact with solid objects. An erratic reading or the unit fails to calibrate can be an indication that it is damaged. If this does occur follow the manufacturer’s procedure to replace the sensor head.

2.12 Soil Porosity Test

To determine the porosity of the soil, it is necessary to determine the (1) Bulk density and (2) the particle density of the soil under test. The soil testing, including the soil porosity and bulk density testing was carried out using, Thein and Graveel, 1997 as a reference.

2.12.1 Bulk Density Test

The bulk density test indicates how dense the soil is and how tightly it is packed according to the shape of the soil peds (a single unit of soil structure) and the percentage of air space or pores. It is directly related to the compaction level of the soil. The bulk density indicator is measured with the dry mass per volume in g/cm^3 or g/ml.

2.12.2 Bulk Density Test – Method

Weigh a 200ml graduated cylinder. Fill up the cylinder with an air dried soil sample. Record the marking where the soil reaches on graduated cylinder. Measure the weight of soil and the cylinder together.

Equation:

$$\text{Bulk Density} = \frac{\text{Weight of cylinder and soil} - \text{weight of cylinder}}{\text{Volume of soil}} \qquad \text{Eq(2.1)}$$

(Measuring unit: g/cm^3 or g/ml)

2.13 Particle Density (Real Density) Test

The particle density test measures the mass of the soil in a specific volume, which is very similar to the bulk density test. The main difference is that the particle density only measures the density of the soil particle component and excludes the volume of pore spaces, which contains air and water.

2.13.1 Particle Density Test – Method

Take a 200ml graduated glass container and measure its weight. Place 25g of a soil sample inside the container. Measure and record the weight of the soil and container. Add some water, then boil the mixture for 10 minutes to remove all air bubbles. After 10 minutes, allow the sample to cool and place the contents into another glass container. After 24hrs, fill up the container with water to a total volume of 100ml and measure the weight and temperature of the mixture.

2.13.2 Particle Density Equations

Mass of soil = Mass of soil and container – Mass of empty container Eq(2.2)

Mass of water = mass of water, soil and container – Mass of soil and container Eq(2.3)

$$\text{Volume of water} = \frac{\text{Mass of water}}{\text{Density of Water}} \qquad \text{Eq(2.4)}$$

[where density of water =1.0 g/cm^3 or g/ml]

Volume of soil = given volume of mixture (100ml) – Volume of water (cm^3) Eq(2.5)

Soil particle density = Mass of soil/ Volume of soil (g/cm^3 of g/ml) Eq(2.6)

2.14 Soil Porosity Test

The fraction of pore space in the soil is called the soil porosity and is measured in percentage.

Soil Porosity Test – Method

- Bulk Density
- Particle Density

The porosity value always will be between 0-1 of 0-100%

2.14.1 Soil Porosity Equation

Porosity = [1- (Bulk density/ Particle density)] /100% Eq(2.7)

2.15 Biological surface water and bio-aerosol analysis

2.15.1 Bacterial enumeration of surface waters samples using the IDEXX Quanti-tray 2000 with Colisure® and Enterolert® assays

2.15.2 Defined substrate assay enumeration

The Quanti-tray 2000© sealer Model 2X was powered up and allowed to stand for approximately 15 to 20 minutes before the determination of the bacterial counts, this was to ensure the sealer has heated up to the optimum temperature of 180^0 to correctly seal the inserted Quanti-trays. To verify the optimum temperature had been achieved after 20 minutes, a Quanti-tray with no liquid bacteria/assay was passed through the sealer to ensure the correct temperature had been obtained to seal the Quanti-tray.

2.15.3 Determination of total coliforms and *E.coli*

For the determination of total coliforms and *E.coli* counts in the wetland wastewaters, the IDEXX® method was utilised. A total of six water samples were examined for the presence of coliforms and *E.coli* using Colilert® (IDEXX®, Technopath. Limerick, Ireland). Samples (100ml) were poured into pre-sterile Duran 500ml bottles and Colisure® reagent added along with two drops of IDEXX antifoam solution. The bottle was capped shaken thoroughly. The sample was allowed to stand for one minute and shaken thoroughly to make sure there were no large media particles remaining. Repeat the above process until there are

no particulates remaining in the Duran bottles. Then pour the samples in the Quanti-Tray/ 2000 and seal in an IDEXX Quanti-Tray 2000® sealer Model 2X. All Colisure® samples were incubated at 35^0C and were examined after 24h incubation. All samples which were red/magenta in colour were counted as positive for Total Coliforms units/100ml, these sample trays were then placed under an ultraviolet lamp (253.7nm) for the detection of "pink" fluorescence which indicated the presence of *E.coli*. The fluorescence is due to the presence of MUG – Colilert® utilises two substrates: *o*-nitrophenyl-β-D-galactopyranoside (ONPG), which screens for β-D-galactosidase, an enzyme found in lactose-fermenting bacteria and in some coliform bacteria, and 4-methylumbelliferyl-β-D-glucuronide (MUG), which screens for β-D-glucuronidase, an enzyme found in several bacterial species, but predominantly in *E. coli*.

2.15.4 Determination of total enterococci

For the determination of total enterococci counts in the wetland wastewaters, the IDEXX® method was utilised. A total of six water samples were examined for the presence of enterococci using Enterolert® (IDEXX, Technopath. Limerick, Ireland). The same procedure applied for determining total coliforms and *E.coli* except no antifoaming solution was required. These sample trays were then placed under an ultraviolet lamp (253.7nm) for the detection of blue fluorescence, which indicated the presence of enterococci, see Figure 2.4.

Quanti-Tray 2000 with Enterolert® assay (UV 253.7nm)

Figure 2.4: Quanti-Tray 2000 with Entrolert® assay, under UV light (253.7nm).

Note: The Quanti-Tray 2000 has large and small cells. Total cell count = 97. Figure 2.4 is showing positive fluoresecence (blue) cells, indicating the presence of enterococci bacteria.

2.15.5 Determination of Most Probable Number (MPN)

Once all the counts (total coliforms, *E.coli* and enterococci) were determined, the MPN table (provided by IDEXX Quanti-Tray 2000) was used to determine a Most Probable Number 100ml/(MPN) for the samples. For better illustration, the numbers of bacterial communities were converted to $\log_{10}$ values and expressed as $\log_{10}$ CFU/100ml.The above procedure was carried out for each sample site within the wetland.

2.15.6 IDEXX Dilutions

If all the 97 cells within the Quanti-Tray 2000 tray were completely pink/magenta or blue for determining total coliforms, *E.coli* and enterococci bacterial concentrations respectively i.e. the sample was too numerous to count. Serial dilutions were made until an accurate count can be determined to ascertain the Most Probable Number.

2.16 Surface air sample (SAS) Super 90

2.16.1 Surface air sample (SAS) Super 90 procedure

The Surface Air System (SAS) Super-90 (PBI International, Milan, Italy) is a battery operated (rechargeable 8.4-Volt, 1.2 A/hr, nickel-cadmium battery), single stage, sieve type sampler. Air is drawn through a single sieve plate with 487-holes (diameter of each hole=0.1 cm). Particulate matter is collected by inertial impaction and deposited onto an agar medium contained in an 84 mm maxi Replicate Organism Direct Agar Contact (RODAC) plate, (Sarstedt Ltd, Wexford, Ireland). The maximum efficiency of collection is for particulate matter

with a *d50* = 2-4µm. The predetermined flow rate is 90 L/min. The SAS was marketed as a device to enumerate fungi or bacteria. Field calibration of this instrument is not possible however the instrument was factory calibrated on a regular basis. The inlet of the air sampler was at 1.8m above the ground and 100L of air was sampled at a flow rate of 90 L/min. For every sample, the sampler was filled with an 55mm plastic RODAC plate containing the agar medium. Once the required air had been collected the plate was covered and incubated. Three replicates were taken at each sample point, for four different agars. After each sample, the sampler was sterilised by washing with a solution of ethanol (95% v/v). The plates were used for the sampling were incubated at 37 ^{0}C for a total of 72h and then the colonies were counted. The positive-hole correction was used to adjust colony counts as per manufactures specifications.

The SAS Super 90 meter was used to detect bio-aerosols at the constructed wetland site. The surface air sampler (SAS) Super-90 aspirates air at a fixed speed for variable periods through a 219-hole cover onto a 55mm RODAC contact plate, containing four different culture media (Plate count, Trypticase Soy, MacConkey and Membrane Lactose Glucuronide Agars). A distance of 30 meters was observed at each sample site location. Following each air sampling the plates were removed and incubated at 37^{0}C.

After 24hrs at 37^0 C, (day 1) the contact plates were removed and the bacterial colonies recorded, these colony counts are subjected to the probable count (Pr) which is used to calculate the colony forming unit (CFU) per cubic meter of air sample (CFU/m^3).
The plates were placed back into the incubation chamber for another 24hrs at 37^{0}C (day 2) and the plates removed and colonies counted. This process was repeated for third day. Therefore, a three-day colony forming unit per cubic meter of air sample was realised.

2.16.2 SAS Super 90 interferences

If the 55m contact plates are not placed correctly onto the circular platform and locked into position, thus leading to incorrect airflow across the agar. Resulting

from incorrect colony counts on the plates and /or the plates falling off the locking platform. If the battery is not fully charged on the day of sampling, this may result in poor flow across the plates resulting in spurious results.

2.17 Culture Media

Aseptic techniques were adhered to throughout the agar preparation.

2.17.1 Plate count agar (Lab M) 28g/l

The PCA powder was dissolved in 1 litre of deionised water and autoclaved as per manufacturer's instruction.

2.17.2 Trypticase Soy Agar (Lab M) 40g/l

The TSA powder was dissolved in 1 litre of deionised water and autoclaved as per manufacturer's instruction.

2.17.3 MacConkey Agar (Lab M) 52g/l
(Selective media for coliforms)

The MacConkey agar powder was dissolved in 1 litre of deionised water and autoclaved as per manufacturer's instruction.

2.17.4 Membrane Lactose Glucuronide Agar (Lab M) 88g/l
(Selective media for coliforms)

The MLGA powder was dissolved in 1 litre of deionised water and autoclaved as per manufacturer's instruction.

2.18 Bacterial enumeration of sludge and soil samples using the IDEXX Quanti-tray 2000® with Colisure® and Enterolert® assays

2.18.1 Bacterial enumeration of sludge and soil samples procedure

Samples of soil and sludge were taken from the DAF, Pond 1, Pond 6, Pond 9, Pond 12, repeating the water sampling procedure as seen in section 2.15.1, using the IDEXX Quanti-tray 2000 with Enterolert® and Colilert® assays to determine the colony forming unit per gram of soil/ sludge. 10g of sample was added to a sterile, 100ml Duran bottle containing 100ml of sterile deionised water, this mixture was homogenised by vigours shaking for 1min and mechanically shaken for one hour. A series dilution was prepared from the homogenised sample, as described by Brazil and Murphy, (2005). The serial dilutions were processed in a similar manner until the desired final dilution was achieved. The contents of one Colilert® assay are added the 100ml dilutions and thoroughly mixed via hand until the contents of the blister pack was completely dissolved. Each 100ml sample then poured into a Quanti-tray and sealed immediately using the IDEXX Quanti-Tray Sealer. The Quanti-trays are incubated 'wells down' at 37°C for 24 hours, Quanti-Tray results were recorded and multiplied by the appropriate dilution factor.

Once all the counts (total coliforms, *E.coli*) are determined, the MPN table (provided by IDEXX Quanti-Tray 2000) was used to determine a Most Probable Number (MPN) for the samples. The same above procedure was repeated for the determination of enterococci bacteria using the Enterolert® assay.

2.18.2 IDEXX Quanti-tray 2000® with Colisure® and Enterolert® assays interferences

When determining the bacteria counts of the samples (total coliforms and *E.coli* using the Colilert assay), apply two to three drops of the anti-foaming agent to the 100ml liquid sample as provided by the manufacture. If not used, the sample will exhibit increased levels of air bubbles, leading to inaccurate results.

Prior to sealing the Quanti-Tray 2000 with the 100ml liquid sample, remove all air bubbles from the sample tray. Place the sample tray cell side down and carefully using your hand gently press, in small arcing circles, the back of the tray; this action will enable trapped bubbles to be expelled and evenly distribute the liquid sample throughout the tray. If this procedure were not performed, during the high temperature sealing process, trapped air bubbles rapidly expand,

resulting in the ejection of liquid sample as the tray passes through the heated roller. Resulting in loss of liquid sample and inaccurate results.

2.19 Soil and sludge sample preparation

Sludge and soil sampling was carried out in the constructed wetland over a seven-month period from April 2007 to Dec 2007. Sludge samples were retrieved from the top 10cm layer of the sludge contained in pond 1. A soil auger was used to obtain three core soil samples from pond 6, pond 9 and pond 12 wetland soils at a depth of 0.5 to 0.7 metres from different locations within the ponds. All samples were in triplicate. The soil samples were added to a mixing bowl and thoroughly hand mixed for 2 mins and 1g soil extracted for analysis and placed in a 50ml red capped centrifuge tubes (Sarstedt, Wexford, Ireland); 1 g of sample was then placed in a fresh tube and 10ml of double de-ionised water was added and the sample was vortexed for 5 mins and allowed to settle until a clear separation between water and soil matrices was obtained. The supernatant was extracted for bacterial analysis. Approximately 20 g of sludge was retrieved from three different locations from pond 1 and thoroughly mixed using a spatula in a 50ml red capped centrifuge tube. From the 20g sample, 1g of sludge was placed in a fresh tube and 10ml of double de-ionised water was added to the sample tube and the sample vortexed for 5 mins and allow the settle until a clear separation between water and soil matrices was obtained. The supernatant was extracted for bacterial analysis.

2.20 Analytical Profile Index (API) 20E Test

2.20.1 The API 20E test

The API 20E test is an identification system for Enterobacteriaceae and other non-fastidious Gram-negative bacteria (rods), which uses 21 standardised and miniaturesed biochemical tests and a database. It consists of 21 microtubes containing dehydrated substrates. These tubes are inoculated with a bacterial suspension which reconstitutes the media. During incubation, metabolism produces colour changes that are either spontaneous or revealed by the addition

of reagents. The reactions are read according to the table provided and the identification is done using the software provided by the manufacturer on the API website : https://apiweb.biomerieux.com/login [Accessed 08 2008] A seven-digit profile is obtained. API 20E ratings are based on three parameters, including the likelihood of a match between the unknown organism's profile and the computer profile, the relative value between the likelihood of the first and the likelihood of the second choice, and the number of tests against the first choice. Some examples of Enterobacteriaceae bacteria are *E.coli, Klebisella, Salmonella, Shigella, Enterobacter*.

2.20.2 The API 20E Test Kit

The test kit enables the following tests:

ONPG: test for β-galactosidase enzyme by hydrolysis of the substrate o-nitrophenyl-b-D-galactopyranoside
ADH: decarboxylation of the amino acid arginine by arginine dihydrolase
LDC: decarboxylation of the amino acid lysine by lysine decarboxylase
ODC: decarboxylation of the amino acid ornithine by ornithine decarboxylase
CIT: utilization of citrate as only carbon source
H2S: production of hydrogen sulphide
URE: test for the enzyme urease
TDA (Tryptophan deaminase): detection of the enzyme tryptophan deaminase: Reagent- Ferric Chloride.
IND: Indole Test-production of indole from tryptophan by the enzyme tryptophanase . Reagent- Indole is detected by addition of Kovac's reagent.
VP: The Voges-Proskauer test for the detection of acetoin (acetyl methylcarbinol) produced by fermentation of glucose by bacteria utilizing the butylene glycol pathway.
GEL: test for the production of the enzyme gelatinase which liquefies gelatin.
GLU: fermentation of glucose (hexose sugar)
MAN: fermentation of mannose (hexose sugar)
INO: fermentation of inositol (cyclic polyalcohol)
SOR: fermentation of sorbitol (alcohol sugar)

RHA: fermentation of rhamnose (methyl pentose sugar)

SAC: fermentation of sucrose (disaccharide)

MEL: fermentation of melibiose (disaccharide)

AMY: fermentation of amygdalin (glycoside)

ARA: fermentation of arabinose (pentose sugar)

2.20.3 The API 20E Method

1. Confirm the culture is of an Enterobacteriaceae. To test this, a quick oxidase test for cytochrome c oxidase may be performed.
2. Pick a single isolated colony (from a pure culture) and make a suspension of it in sterile distilled water.
3. Take the API 20E Biochemical Test Strip which contains dehydrated bacterial media/bio-chemical reagents in 20 separate compartments.
4. Using a pasteur pipette, fill up (up to the brim) the compartments with the bacterial suspension.
5. Add sterile oil into the ADH, LDC, ODC, H2S and URE compartments.
6. Put some drops of water in the tray and put the API Test strip and close the tray.
7. Mark the tray with identification number (Organism ID), date and your initials.
8. Incubate the tray at 37^{o}C for 18 to 24 hours.

2.20.4 API 20E Results Interpretation

1. For some of the compartments, the colour change can be read straightway after 24 hours but for some reagents must be added to them before interpretation.
2. Add following reagents to these specific compartments:
    - TDA: Put one drop of Ferric Chloride.
    - IND: Put one drop of Kovacs reagent.
    - VP: Put one drop of 40 % KOH (VP reagent 1) & One drop of VP Reagent 2 (α-Naphthol).
3. Get the API Reading Scale (colour chart) by marking each test as positive or negative on the lid of the tray. The wells are marked off into triplets by black triangles, for which scores are allocated.

Triad	I			II			III			IV			V			VI			VII		
Tube	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	oxidase
Reaction	+	+	+	-	-	-	+	+	-	+	-	-	-	+	+	-	-	+	+	-	+
Point	1	2	4	0	0	0	1	2	0	1	0	0	0	2	4	0	0	4	1	0	4
Add	7			0			3			1			6			4			5		
7-digital Code	7 0 3 1 6 4 5																				

4. Add up the scores for the positive wells only in each triplet.
5. Three test reactions are added together at a time to give a 7-digit number, which can then be looked up in the codebook. The highest score possible for a triplet is 7 (the sum of 1, 2 and 4) and the lowest is 0.
6. Identify the organism by using API web: https://apiweb.biomerieux.com/login [Accessed 08 2008]

2.21 Statistical and mathematical methods

Table 2.3 provides a list of statistical and mathematical software packages used in this research.

Table 2.3 List of the statisitcal and mathematical software

	Microsoft Excel	SPSS v.22	Hugin	SoMine	Easy-NN
Chapter 3	Normalised (S)-Distribution	Principal component analysis (PCA), Canonical variate analysis (CVA), Canonical discriminant analysis (CDA)			
Chapter 4	Wetland traditional models and graphs				
Chapter 5	Bio-aerosol graphs				
Chapter 6			Bayesian belief network and Sensitivity analysis		
Chapter 7	Self-organised criticality and Entropy analysis				
Chapter 8				Self-Organised Map Model (unsupervised), "What-if" scenario analysis	Three Artifical Neural Network Models (supervised)

	Microsoft Excel	SPSS v.22	F-IND	UniNet	UniSens	Crystal ball Fusion
Chapter 9	Coefficent of Variation (CV) and Coefficient of Reliability (COR) analysis	Cronbach's Alpha and Coefficient of Reliability (COR) analysis				
Chapter 10 (section 1)		Principal component analysis (PCA)				
Chapter 10 (section 2)			Fuzzy Indexing			
Chapter 10 (section 3)				Construct Bayesian belief network models	Find the Resilience output from the Bayesian models	Resolve the probability ditribution functions (pdfs) for the Bayesian models

Chapter 3. Establishment and multi-variate analysis of a constructed wetland dataset

3.1 Summary

The wetland dataset was compiled taking monthly surface water grab samples over the course of a year from February 2007 to February 2008 from six sample sites in the constructed wetland (CW). The dataset contained bacteria, inorganic ions and physical chemical data. The grab samples were taken from the dissolved floatation air (DAF) tank, pond's 1, 6, 9 and 12 and the local stream.

This chapter utilises several statistical techniques to analyse the wetland data. The normalised S-distribution, ANOVA, canonical variate analysis, principal component analysis and canonical discriminant analysis including discriminant functional analysis are the multi-variate analyses performed. The normalised S-distribution revealed a plateau in the wetland between ponds 9 to 12, which potentially indicates an issue at these sample points.

The two-way ANOVA ($p < 0.05$) without replication indicates that the indicator bacteria have a statistically stronger treatment by pond than by wetland with *p*-values of 0.0000 and 0.044 (row by column) respectively where row = pond and column = wetland. In comparison, the physical-chemical and inorganic data sets which display the opposite behaviour with *p*-values of 0.0037 and 0.0000 (row by column) for physical-chemical data and *p*-values of 0.0023 and 0.0000 (row by column) for the inorganic ion data.

The canonical variate analysis (CVA) explored the interactions of the indictor bacteria and the other wetland variables, with physical chemical parameters pH, redox potential, dissolved oxygen and turbidity display strong negative interactions with indicator bacteria. The CVA revealed that there exists a second potential source of *E.coli* entering the constructed wetland.

Principal component analysis (PCA) shows the dominance of the indicator bacteria in the wetland dataset including water depth, total bacteria, turbidity, and conductivity and dissolved oxygen. The canonical discriminant analysis (CDA) including discriminant function analysis (DFA) elucidated that potentially the

CW displays resilience in the middle of the wetland system i.e pond 6. Stagnant issues may exist in DAF, pond 9 and pond 12.

3.2 Introduction

Wastewater generated by the meat processing industry has a high organic load (both soluble and particulate), dark colour (dark red to brown) and high nutrient and microbial loads (John, 1995). Constructed wetlands for wastewater treatment in Ireland are in their infancy when compared to North America and other European countries (Healy and Cawley, 2002). Yan and Xu (2014) stated that the number of wetlands in Europe is over 50,000 and over 100,000 in North America. Constructed wetlands (CWs) are man-made wastewater treatment systems that encompass a multitude of treatment modules including biological, chemical and physical process, which are similar to the processes that occur in natural wetlands. Initially designed and used for domestic wastewater treatment, CWs have now been successfully used for environmental pollution control, through the treatment of a wide variety of wastewaters including industrial effluents, urban and agricultural stormwater runoff, animal wastewaters, leachates, sludges, mine drainage and river waters (IWA, 2000; Scholtz and Lee, 2005; Brown *et al*., 1999; Jing *et al*., 2001; Reed *et al*., 1995; Shepherd *et al*., 2001). The constructed wetland systems are considered a cost-effective technology for treating wastewaters of small communities and industries where the dominant pollutant load is organic in nature (such as food, the meat processing industry, animal farms) (John, 1995; Reed *et al*., 1995). Within the Irish concept of cost-effectiveness CWs, especially during the 90s, which was the beginning of the Celtic Tiger a period of rapid economic growth, there was a gradual increase in the volume of sewage generated. The need arose for low-cost and effective wastewater treatment by local authorities. In 1993 the first conference on constructed wetlands for wastewater treatment in Ireland was held (Costello, 1993), followed by a second (Babatunde and Zhao), in 2010.

To verify the effectiveness of the CW in treating the abattoir wastewater, statistical techniques were employed to elucidate any potential issues within constructed wetland. DeKeyser *et al*., (2003) used a combined principal component analysis (PCA) along with clustering analysis to develop a plant index

methodology in assessing prairie wetland plant communities. Liao *et al*., (2008) used canonical discriminant analysis (CDA) to identify the source of pollution entering a large lagoon and Sadeghi *et al*., (2014) used advanced data driven methods, to model habitat preference of an aquatic fern in a wetland system, using support vector machines (SVMs) and genetic algorithms (GAs) in conjunction with statistical methods such as paired Student's *t*-test, correctly classified instances (CCI %) and Cohen's Kappa (*k*). The above statistical methods were used to verify model optimization and efficacy. This chapter uses a detailed physico-chemical, inorganic ion and microbiological dataset gathered for 1-year period from monthly grab sampling. A complex dataset was established from the CW to multi-variate analysis on the CWs dataset, in order to ascertain what variables dominate within the CW and what ponds within the CW are any potential issues.

3.3 Constructed wetland used in this study

The abattoir and the associated constructed wetland in this study is regulated by the Department of Agriculture, Ireland with high slaughter rates of bovine, ovine, porcine, equine slaughtered throughout the year, with average of 90 to 130 heads per day and an average flowrate of wastewater between 64m^3/d to 86m^3/d as per 2008.

The abattoir is located in south-east of County Carlow, with a free water surface (FWS) wetland, consisting of a primary treatment dissolved air floatation (DAF) plant and twelve ponds in series, with an average pond surface area of 2,100 m^2 comprising of 2.75 hectares in total. The resultant discharge from the pond system empties into a local stream (Black stream).

There are three characteristics that make constructed wetlands very attractive for agricultural and industrial waste:

- They can physically trap pollutants by absorption on surface soils and organic debris.
- Micro-organisms and plants can be utilised to transforms many compounds.

- Low energy and maintenance is required to attain consistent levels of water quality (Skarda *et al*., 1994).

An eleven-month sampling plan was achieved from February 2007 to February 2008. The only month omitted from the sampling plan was October 2007 due to the unsafe conditions of the wetland system, making sampling dangerous.

3.4 Sample sites within the constructed wetland

Table 3.1 provides an overview of the pond areas' obtained from the owners of the constructed wetland. Figure 3.1 provides a Google map® of the free water constructed wetland that was observed.

Google map® of the constructed wetland, with sample points highlighted.

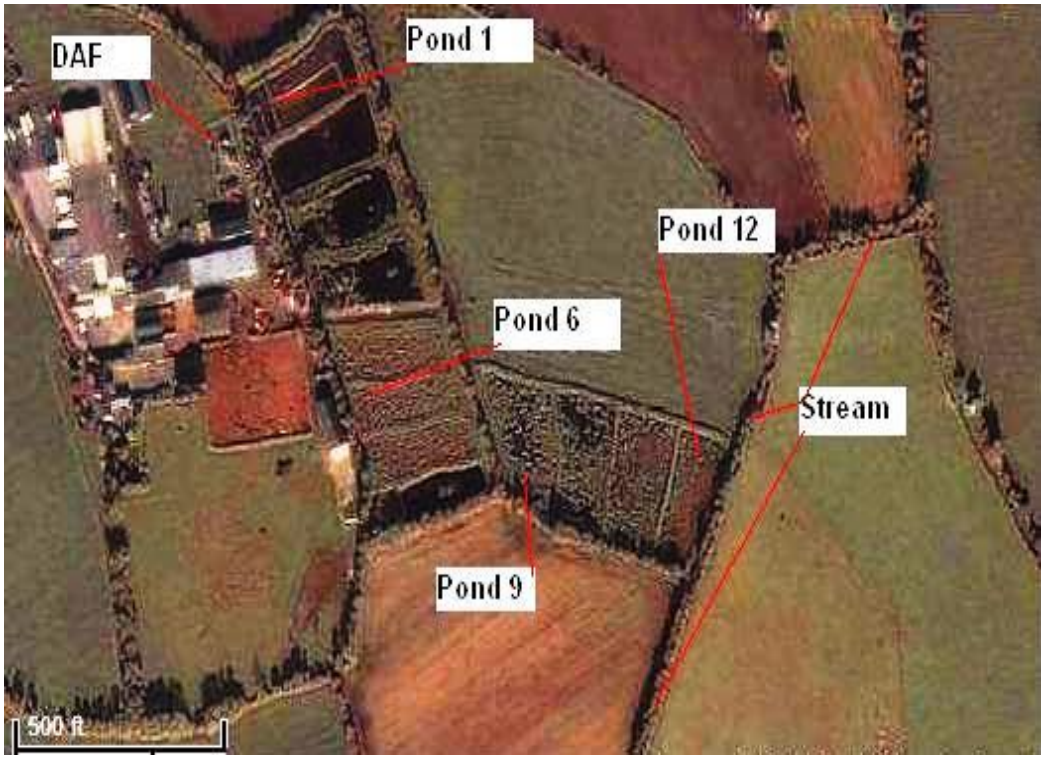


Figure 3.1: The Google map® (April, 2008) of the constructed wetland (the sample points highlighted).

Table 3.1: Area of the wetland ponds.

Pond	1	2	3	4	5	6	7	8	9	10	11	12
Estimated area (m^2)	1898	1757	2330	2619	1508	1804	1867	2140	2813	1936	2845	2428

The wetland influent wastewater comes from an abattoir located 400m to the wetland system. The wastewater first enters the dissolved air floatation (DAF) plant, and then proceeds into the wetland system via pond 1. A method for primary treatment of wastewater is the dissolved air floatation system (DAF), used to reduce effluent load of fats, suspended solids and BOD. In the DAF, air bubbles injected at the bottom of the floatation tank, transport small solids and fat, oils and grease (FOG) to the surface where the scum is consistently skimmed off, see Figure 3.2.

Figure 3.2: The DAF tank removing the white foam containing fat, oils and grease.

Chemicals such as polymers and flocculants are often mixed prior to the DAF process for better performance. Blood coagulants (e.g. aluminium sulphate and ferric chloride) and /or flocculants (e.g. polymers) are added to wastewater to increase protein clumping and fat floatation.

Pond 1 is the first lagoon/pond of the FWS system and the first of the CW's anaerobic ponds. Within pond 1 there are two sedimentation ponds with a combined area of 350m^2, to trap the large quantity of particulates that originates from the DAF effluent. After leaving the sedimentation ponds the wastewater enters the main area of pond 1. There is very little plant life within the pond due to the large BOD >1000mg/L and the high ammonia >300mg/L, but surrounding the circumference of the entire pond, are willow trees. The trees are planted on the soil bank of the pond within close proximity to the ponds black water and digested sludge. The trees also provide stability to the ponds soil bank, see Figure 3.3.

Figure 3.3: Pond 1 with planted willow saplings. Very little plant life present in the pond.

Pond 6 is the midpoint of the CW and is the second pond within the aerobic pond system. Pond 6 is one of the most densely vegetated of the ponds, which a mixture of *Typha latifolia, Typha angustifolia and Phragmities australis.* These aquatic plants cover the entire surface area of the pond 6. There are no willows planted in the vicinity of pond 6, see Figure 3.4.

Figure 3.4: Pond 6 displaying abundant reed growth.

Pond 9 is the first of the polishing ponds of the CW. Pond 9 was reasonably populated with aquatic plants approx. 70% of the surface area covered with a mixture of *Phragmities australis, Iris pseudacorus* and *Carex sp*. The circumference of pond 9 was planted with young willow tree saplings, see Figure 3.5.

Figure 3.5: Pond 9 with sporadic plant life present in the pond.

Pond 12 is the last of the polishing ponds of the CW, with the aquatic species varying in population depending on the time of year, from approximately 50-60% coverage to zero coverage during the summer months. The pond aquatic plant species is similar to pond 9, *Phragmities australis, Iris pseudacorus* and *Carex*

sp. The circumference of pond 12 was planted with young willow tree saplings, see Figure 3.6.

Figure 3.6: Pond 12 with planted willow saplings. Very little plant life present in the pond.

The stream ("Black Stream"), where the output pipe of pond 12 discharges. It is a fast flowing stream. *Phragmities australis* plants are visible within the stream at the output pipe, see Figure 3.7.

Figure 3.7: The output pipe of pond 12, emptying into the local stream.

3.5 General Observations:

Pond 1: is subdivided into smaller internal lagoons. These internal subdivisions serve the following purposes (1) as sediment traps, due to the large amount of particulates emanating from the abattoir process i.e. Fats Oils and Grease (FOG) and (2) to allow the owner to separate the sludge into more manageable compartments when removing the anaerobically digested sludge as a soil bulking agent and fertiliser. Pond 1 contains no vegetation and had strong odours emanating from it, especially during the warm summer months. Pond 1's surface waters are dark red in colour. Large green algal blooms were present in pond 1, especially visible during warm weather summer months varying in size from 4.0 m^2 to 8.0 m^2 in circumference.

Pond 1 to pond 4 have very little plant growth, these ponds are labelled anaerobic ponds due the high BOD, bacteria concentrations and ammonia, making plant growth difficult.

Pond 5 to pond 8 have maximum plant growth and are labelled "aerobic" ponds, due to reduced BOD, bacteria concentrations and ammonia, with increased dissolved oxygen. Pond 8 has very little plant growth and is designed as long channel pond, to improve hydraulic efficiency prior to entering into the final ponds. Pond 6 is the midpoint pond within the constructed wetland, and it contains lots of vegetation (high density cover) with little odour present. The surface water within pond 6 was a strong yellow colour. Pond 6 maintains a good water depth with some variation during the year. No algal blooms were observed during the sampling process.

Pond 9 to pond 12 have medium to no plant growth and are labelled "polishing" ponds, to further treat and polish the wastewaters, reducing BOD, bacteria concentrations and ammonia and increasing dissolved oxygen. The effluent enters pond 9, with varying vegetation cover (low to medium density cover) along with varying water depths. The pond water is medium yellow colour, with no detectable odour. Medium to small green algal blooms have been observed in pond 9 especially during the warm summer months, varying in size from 0.5m^2

to 2.0m^2 in circumference. The final pond, pond 12 has varying amount of vegetative cover or none (barren) (medium to no density grass cover). With a large variation in water depth observed during the season sampling process; the water depth was either extremely high in the late autumn, winter, early spring months or non-existent (dry pond) during the summer/ early autumn months. No algal blooms were observed during the sampling process. The surface water was a pale yellow in colour. No odours were detected within pond 12.

3.6 Key wetland variables

The wetland dataset can be divided into three subsections containing thirteen variables: (1) the physical-chemical data (2) the bacterial populations and (3) the inorganic data. The wetland critical physical-chemical data are: pH, redox potential, conductivity, turbidity, BOD, dissolved oxygen. The wetland critical bacterial population data are: the total coliforms, *E.coli* and enterococci and finally the wetland critical inorganic data are: nitrate, phosphate, sulphate and ammonia, see Appendix A. In order to best visualise the wetland variables, was obtained the mean values for each variables for each samples point within the wetland, see Figures 3.9, 3.10 and 3.11 for the physical-chemical variables, indicator bacteria populations and the inorganic variables respectively.

3.7 Statistical analysis of key wetland data

The analysis of the wetland data was performed using SPSS v.22 and Microsoft Excel 2007. The data analysis tool package (Microsoft Excel 2007) was used to create the ANOVA and normalised S-distribution data results. The SPSS v.22 software was used on the more advanced statistical tests such as canonical variant/variate analysis (CVA), principal component analysis (PCA) and canonical discriminant analysis (CDA). The analysis was performed on the wetland dataset, grouped as follows: group (1) – physical-chemical group; pH, redox potential, dissolved oxygen, biological oxygen demand (BOD), turbidity; group (2) – indicator bacterial group; Log_{10} total coliforms, Log_{10} *E.coli* , Log_{10} enterococci, and group (3) – inorganic group; nitrate, phosphate, ammonia and sulphate. Multiple statistical tests were employed. The normalised (S)-

distribution, the analysis of variance (ANOVA), canonical variant/variate analysis (CVA), principal component analysis (PCA) and canonical discriminant analysis (CDA). The normalised (S)-distribution and the analysis of variance (ANOVA) are simpler statistical techniques in comparison to the CVA, PCA and CDA techniques.

In S-distribution technique, all the data is resolved into proportion with one another, with values between 0 and 1. This technique reduces the effect of large date values or can account for missing data, thereby eliminating bias within the dataset. The data from the sample points were "soft-scaled" or linear transformed. Soft-scaling, delivers transformed variables with zero mean and unit standard deviation. Such a transformation is critical for variables with values spanning over different ranges, to ensure that all variables have common importance, several of these functions have an input range between 0 and 1. Then the soft-scaled data is subjected to a Normalised (S)-distribution and returns the standard normal cumulative distribution function. The distribution has a mean of 0 (zero) and a standard deviation of one.

The analysis of variance (ANOVA) method was used to analyse the differences between groups' means such as the variation between groups (ponds) and the overall group (wetland) at $p < 0.05$ or 95% significance.

The interactions between the individual bacterial concentrations and the physical-chemical and inorganic data in the wetland system, were further studied using canonical variant analysis (CVA) using SPSS® v.22. This test calculates the variable weightings that maximize the differences between the individual bacterial concentrations and the physical-chemical and inorganic data. These weightings (canonical correlations) allow the identification of variables that are more correlated.

Principal component analysis, (PCA) is a dimension reduction technique, reducing a complex dataset to lower dimensions to reveal the hidden and simpler behaviour of the dataset. Leading the PCA algorithm to reveal the dominant or principal variable or components that dominant the majority of the dataset variance, i.e. PCA summarises total variation of the variables.

Canonical discriminant analysis (CDA) is a dimension reduction technique similar to PCA, but CDA derives canonical variables (inter-combinations of the variables) that evaluates between class-variation of the variables. The variables

are clustered into distinct groupings, generally by the classification of a target variable within the dataset. Matthew *et al*., (1994) used CDA and PCA as statistical tools to evaluate differences within pasture lands in New Zealand. Wolhfart *et al*., (2018) used CDA and CVA to analyse satellite radar data from complex wetland and marsh ecosystems, looking at the effect of seasonality with these complex ecosystems.

3.8 The wetland dataset

The wetland dataset can be divided into three subsections containing thirteen variables: (1) the physical-chemical data (2) the bacterial populations and (3) the inorganic data. The wetland critical physical-chemical data are: pH, redox potential, conductivity, turbidity, BOD, dissolved oxygen. The wetland critical bacterial population data are: the total coliforms, *E.coli* and enterococci and finally the wetland critical inorganic data are: nitrate, phosphate, sulphate and ammonia, see Appendix A. In order to best visualise the wetland variables, the mean values were obtained obtained for each variable and for each sample point within the wetland, see Figures 3.10, 3.11, 3.12 and 3.13 for the physical-chemical variables, indicator bacteria populations and the inorganic variables respectively.

3.9 Results Section

Overview of the wetland system:

Figure 3.8. shows the mean accumulated constructed wetland dataset as a graph. Where the wetland dataset was reduced to an accumulated normalised (S)-distribution, to values between 0 and 1 The (S)-Distribution contains a maximum, mean and minimum of all the wetland data. By using the S-distribution technique, the bias of large numbers was removed from the dataset. The normalised (S) distribution resolves the data into an upper, (max) mean and lower (min) data points for each sample point.

Maximum value = winter/ wet months' Minimum value = summer/ dry months'.

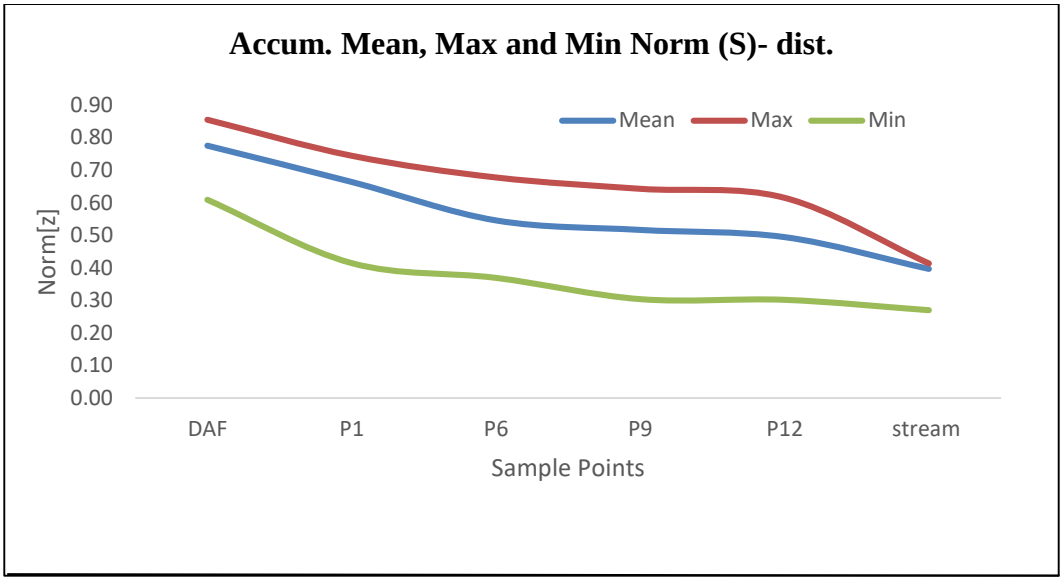


Figure 3.8: The accumulated mean, max and min of the normalised S-distributions of the CW.

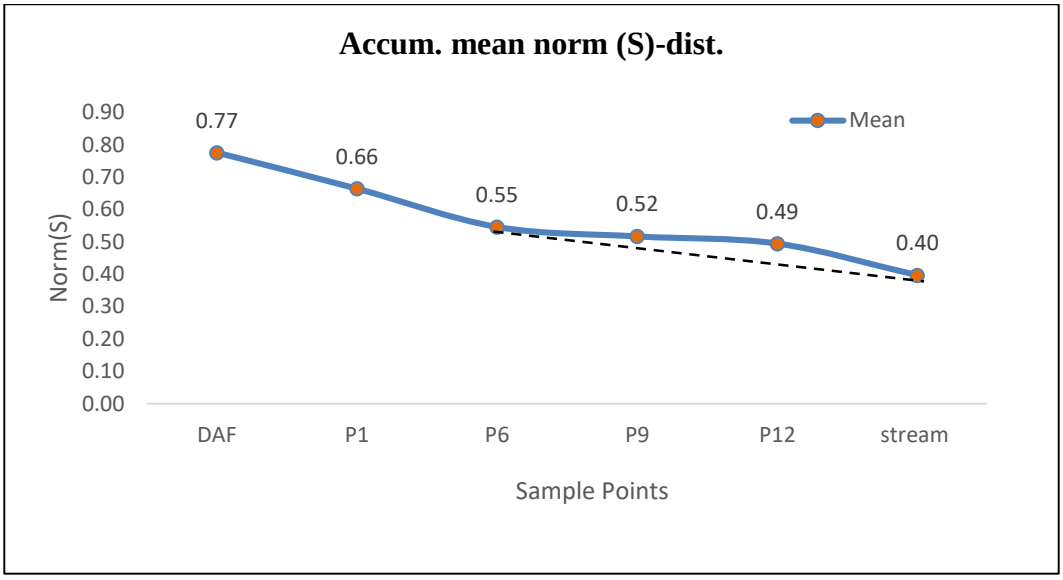


Figure 3.9: The Accum. mean normalised (S)-distribution flow of the key data through the wetland.

Figure 3.9 shows the mean normalised distribution of the wetland. Starting with the input (DAF) with a value of 0.77 and the normalised distribution of data progressing in a decreasing manner through the wetland, with a lower normalised distribution value seen at the output (stream) with a value of 0.40. A black dashed line is superimposed on the graph between pond 6 and the stream. The line indicates the ideal decaying normalised distribution trend for ponds 9 and 12.

Ideally, the normalised values decrease through the constructed wetland from sample point to sample point . Ponds 9 and 12 display a plateau in values of 0.52 and 0.49 repsectively, see Figure 3.9. In Figure 3.8, the maximum trend line displays a more pronounced plateau between pond 9 and pond 12. One of the fundamental hypotheses of this thesis was to examine this plateau and verify whether this plateau is real and complementary of the wetlands performance or is an artefact of an underlying issue with the system.

3.10 Wetland graphs and ANOVA

The data displayed in this section is the mean data from the sampling period (1 year) for each physical-chemical parameter inclusive of standard error bars. Following analysis with ANOVA without replication and considering the data/sample under analysis as a balanced design (the sample sizes are equal, observing the main effects of the data). The above test was evaluated on the wetland dataset, which was divided into three subsets; (1) physical- chemical data (2) bacterial population data and (3) inorganic data. The two-way ANOVA evaluates the data subsets rows by columns (sampling period within the individual sample points and through the wetland). We evaluate two null hypotheses one by row and the other by column. Consider the following null hypothesis *(H_0:)* for the row, where *H_0:* indicates there is no significant difference between the data subsets. For example, consider the p-value ($p < 0.05$) for the rows = $0.0037 < 0.05 = \alpha$ (or $F = 13.548 > 1.534$ = F-critical), thus rejecting the null hypothesis (See Appendix B (ii)). At the 95% level of confidence ($p < 0.05$) in conclusion there is significant difference in the subsets produced. A similar procedure can be applied to the columns. Next consider the following; p-value ($p < 0.05$ or 95% level of confidence) for the rows = $0.1446 > 0.05 = \alpha$ (or $F = 2.63 < 4.76$ = F-critical), the null hypothesis cannot be rejected. Therefore, in conclusion there is no significant difference in the data. In all of the three data subsets analysed the final outcome from the analysis was to reject the null hypothesis for both rows and columns, see Appendix B.

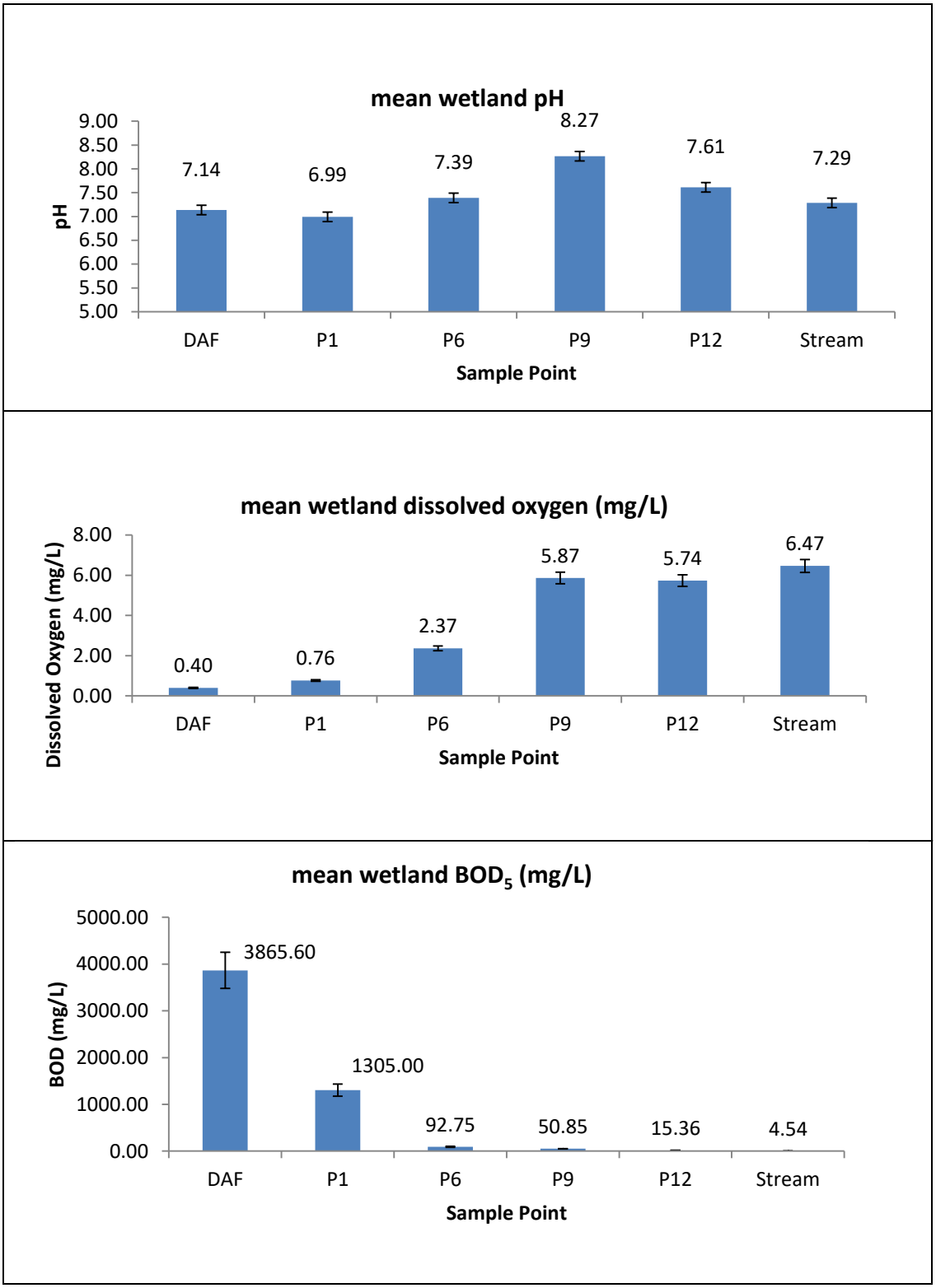


Figure 3.10: The mean wetland physical chemcial pararmeters: pH, dissolved oxygen, redox potential, $BOD_{5\ Day}$ and turbidity, with 5% error bars.

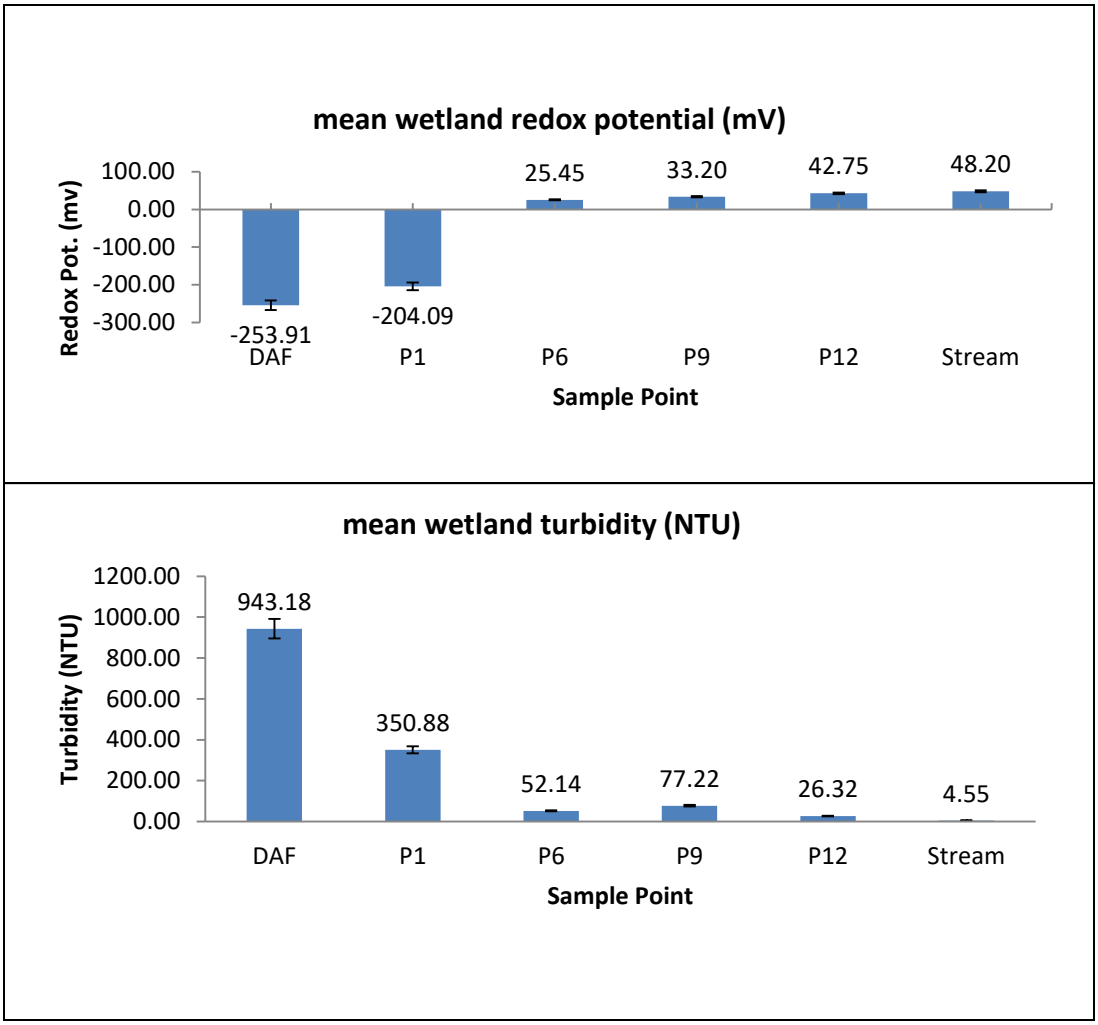


Figure 3.11: The mean wetland physical chemcial pararmeters: redox potential and turbidity,with 5% error bars.

The analysis of variance (ANOVA) of the wetlands physical-chemical data [pH, redox potential, dissolved oxygen, turbidity]. Based on the results of the ANOVA Table for ($p < 0.05$), see Appendix B (i) it can be concluded that:

- Row: There is seasonal difference within the sample points regarding the physical-chemical data.
- Column: There is seasonal difference as the physical-chemical data moves through the wetland.

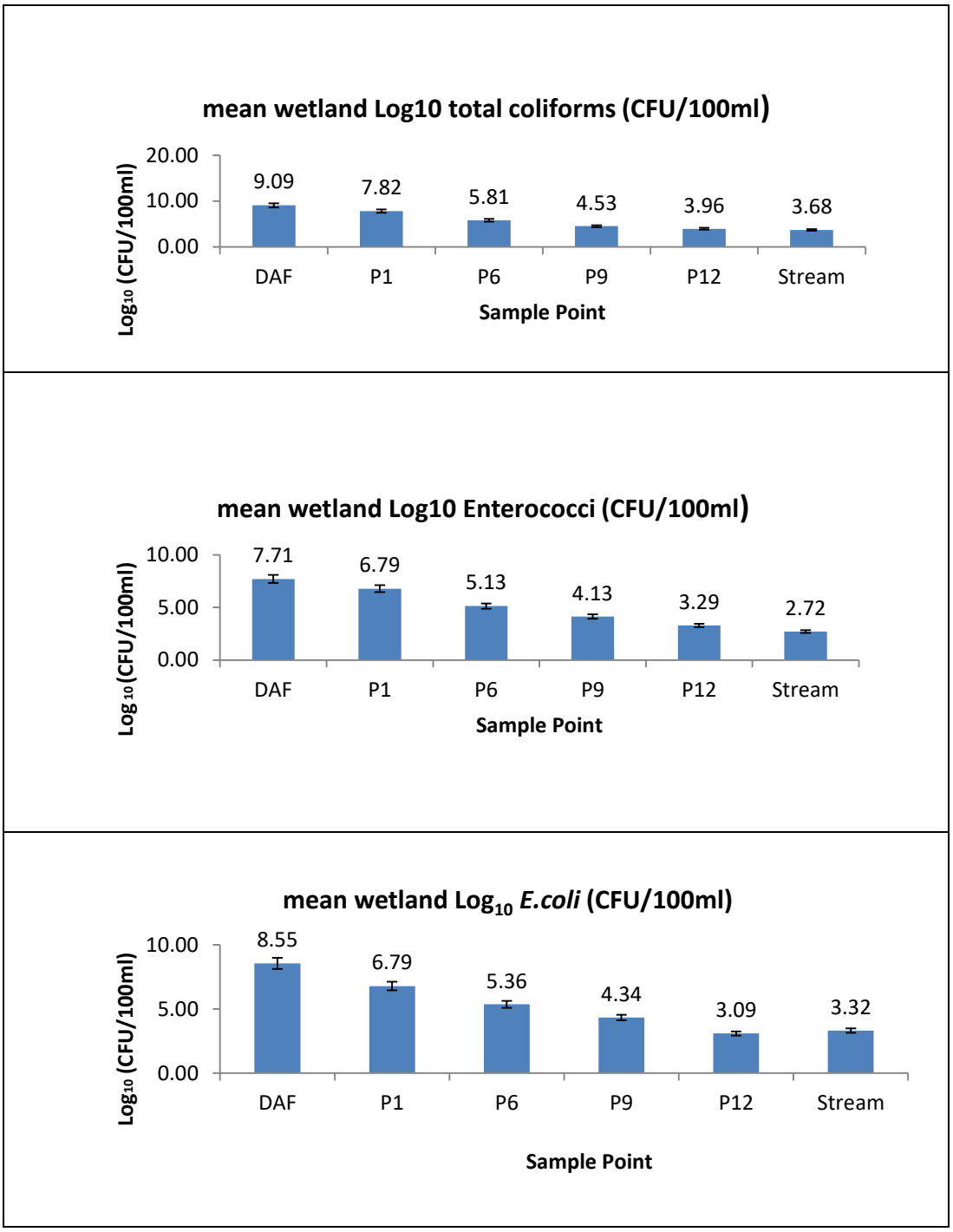
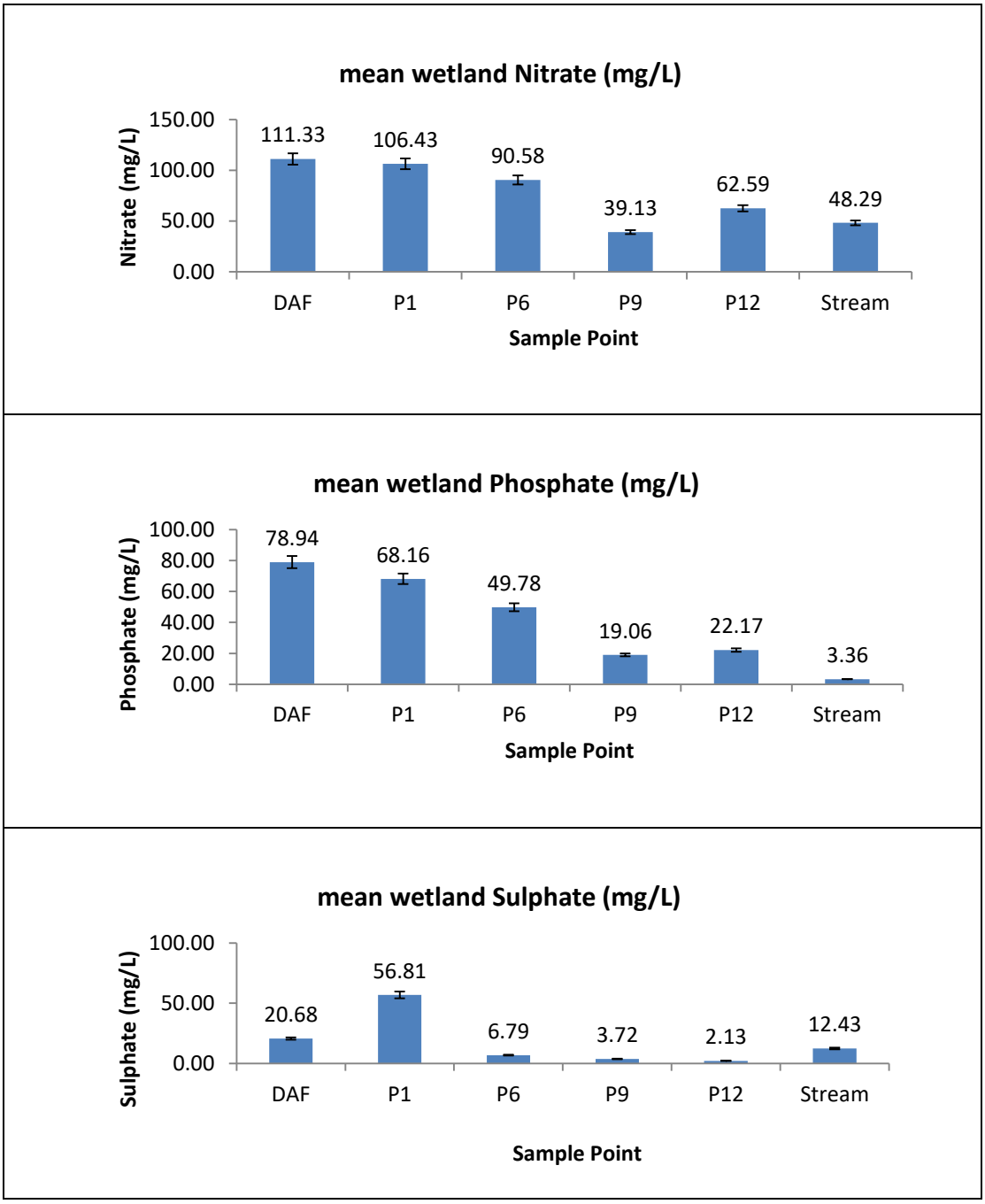


Figure 3.12: The mean wetland indicator bacteria populations: total coliforms, *E.coli* and enterococci, all with 5% error bars.

The analysis of variance (ANOVA) of the wetlands bacterial population data [total coliforms, *E.coli*, enterococci]. Based on the results of the ANOVA table for ($p < 0.05$), see Appendix B (ii) it can be concluded that:

- Row: There is seasonal difference within the sample points regarding the bacterial concentration data
- Column: There is seasonal difference as the bacterial concentrations move through the wetland system.



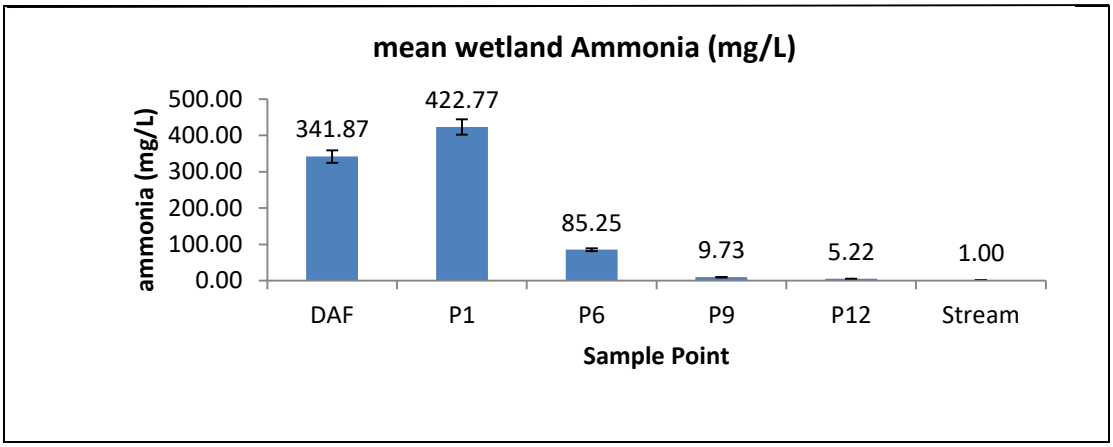


Figure 3.13: The mean wetland inorganic ion concentrations: Nitrate, phosphate, sulphate and ammonia, with 5% error bars.

The analysis of variance (ANOVA) of the wetland's inorganic concentration data [nitrate, nitrite, phosphate, sulphate and ammonia]. Based on the results of the ANOVA Table for ($p < 0.05$), see Appendix B (iii) it can be concluded that:

- Row: There is seasonal difference within the sample points regarding the inorganic ion data.
- Column: There is seasonal difference as the inorganic ion data moves through the wetland.

A review of the three wetland variable groups, see Figures 3.11, 3.12 and 3.13 reveals that all wetland variables display a seasonal difference within the sample points and that there is also a difference as these wetland variables move through the constructed wetland. The most common trends within the graphs, shows an exponential or linear decrease from input (DAF) to the output (stream). The exceptions are the pH, the redox potential and the dissolved oxygen. The pH reveals a wave like pattern with the peak of the wave at pond 9. The redox potential reveals a switch from negative millivolts (anaerobic) to positive millivolts (aerobic) in the vicinity of pond 6. The dissolved oxygen displays an increase in dissolved oxygen content with a noticeable increase at pond 6, see Figure 3.10 mean dissolved oxygen graph; where pond 1 dissolved oxygen value = 0.76 (mg/L) and pond 6 dissolved oxygen value = 2.37 (mg/L) therefore a 3-

fold increase, in comparison to pond 6 = 2.37 (mg/L) and pond 9 = 5.87 (mg/L), therefore a 2.5-fold increase.

Reviewing the ANOVA results more closely for each of the wetland variable groups, see Appendix B. The statistic variable of interest is the row and column *P-value's,* see Table 3.2.

Table 3.2: Wetland data ANOVA *P-value's*

($p < 0.05$)	**Physical-chemical**	**Indicator bacteria**	**Inorganic ions**
Row/pond	0.0037	0.0000	0.0023
Column/wetland	0.0000	0.0446	0.0000

Table 3.2 shows the corresponding *P-values* for the three subsets of the wetland dataset; the physical-chemical, indicator bacteria and the inorganic ions. The rows indicate seasonal difference by sample point i.e. pond, the columns indicate seasonal difference by wetland. Both the physical-chemical and inorganic ions display similar row and column behaviour but the indicator bacteria are the inverse of the latter. In other words, the indicator bacteria are more significant in treatment by pond than by wetland, plus the column value of 0.0446 is near the accepted significant threshold of ($p < 0.05$).

In the next section 3.11, using canonical variant analysis (CVA) and principal component analysis (PCA) determining what wetland variables are dominating the interactions with the indicator bacteria using canonical variant analysis. Using PCA deterimining that indicator bacteria are the dominant variables within the wetland dataset.

3.11 Using canonical variant analysis to explore interactions between bacterial data and other variables within the wetland

The influence of physical-chemical parameters on the distribution of bacterial populations was explored by canonical variant analysis (CVA). It is a generalised

multiple regression that finds linear interactions within each set of variables, that best explain the relationship between them. The analysis gave correlations for individual bacterial populations i.e. $\log_{10}$ total coliforms, $\log_{10}$ *E.coli* and $\log_{10}$ enterococci respectively, see Figures 3.14, 3.15 and 3.16. The correlations suggest how the seasonal conditions in the wetland may influence the bacterial populations throughout the wetland, such as ultraviolet radiation, the effects of the roots of the wetland plants and predation by other microbial populations (Duncan and Groffman, 1994).

From the $\log_{10}$ total coliforms CVA test it can be concluded that there were positive associations between $\log_{10}$ *E.coli*, BOD, $\log_{10}$ enterococci, turbidity, phosphate, sulphate, ammonia and nitrate; and negative associations between $\log_{10}$ total coliforms and redox potential, dissolved oxygen and pH.

From the $\log_{10}$ *E.coli* CVA test it can be concluded that there were positive associations between BOD, ammonia, redox potential, pH, phosphate, nitrate; and negative associations between $\log_{10}$ *E.coli* and dissolved oxygen, turbidity, $\log_{10}$ total coliforms, $\log_{10}$ enterococci, and sulphate.

From the $\log_{10}$ enterococci CVA test it can be concluded that there were positive associations between $\log_{10}$ total coliforms, ammonia, $\log_{10}$ *E.coli*, phosphate, sulphate, turbidity, BOD, redox potential; and negative associations between dissolved oxygen pH and nitrate.

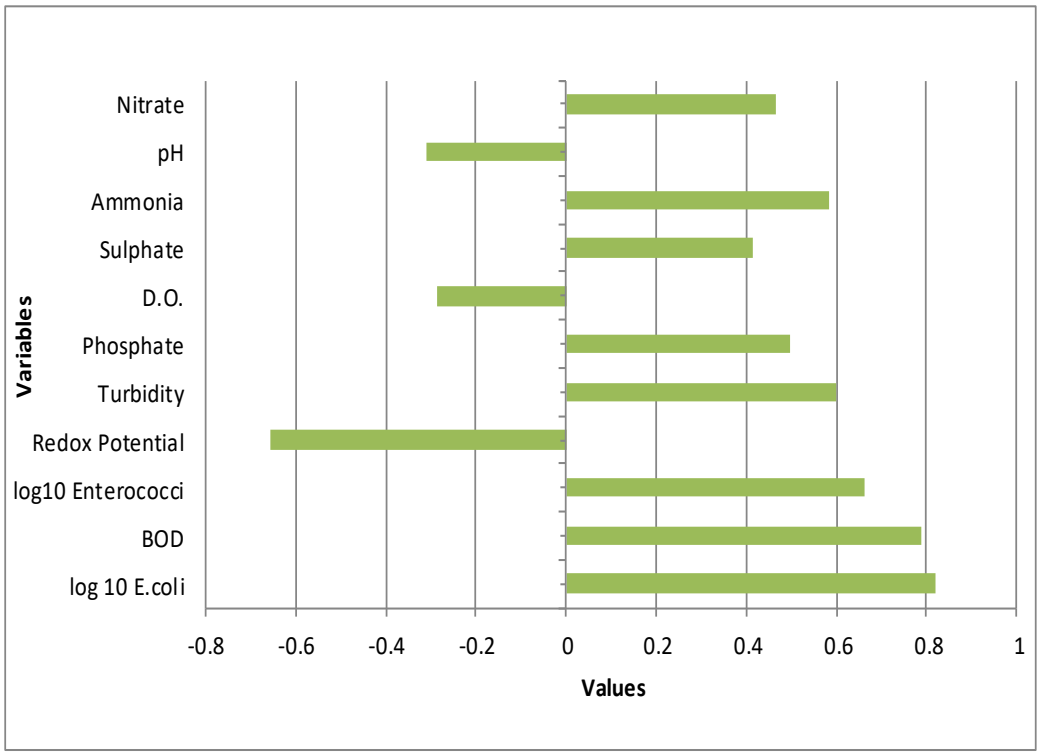


Figure 3.14: The $\log_{10}$ total coliforms canonical correlations. The interaction between different variables in the constructed wetland with $\log_{10}$ total Coliforms versus the other wetland variables using canonical variant analysis. Pooled within-groups correlations between discriminating variables and standardised canonical discriminant functions. Variables ordered by absolute size of correlation within function.

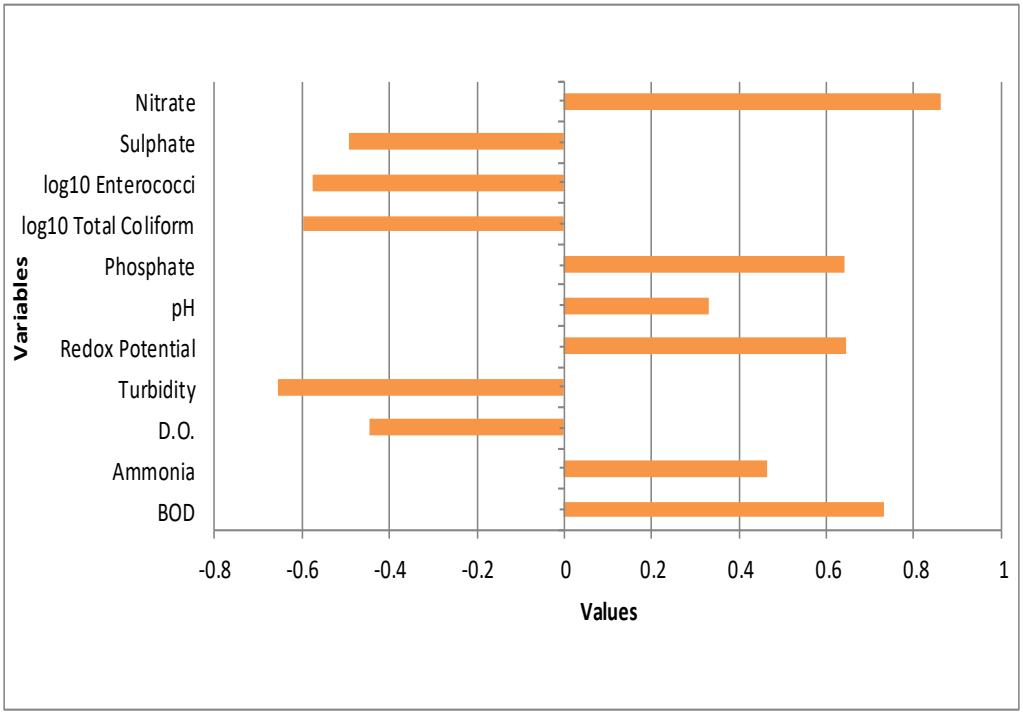


Figure 3.15: The $\log_{10}$ *E.coli* canonical correlations. The interaction between different variables in the constructed wetland with log10 *E.coli* versus the other wetland variables using canonical variant analysis. Pooled within-groups correlations between discriminating variables and standardised canonical discriminant functions. Variables ordered by absolute size of correlation within function.

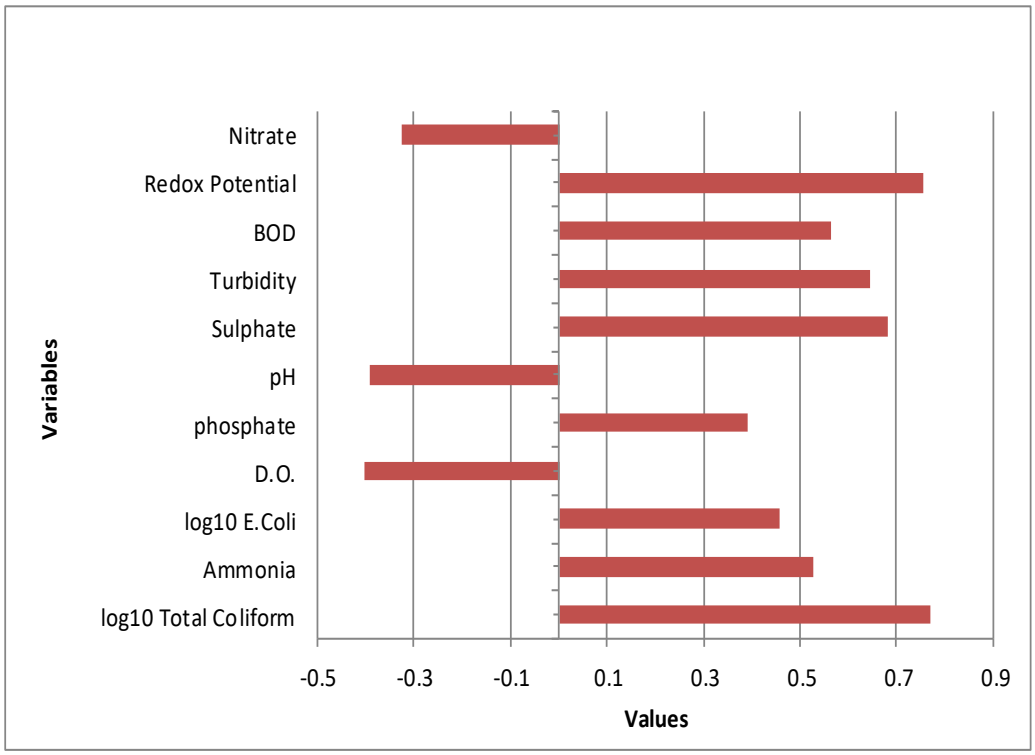


Figure 3.16: The log_{10} enterococci canonical correlations. The interaction between different variables in the constructed wetland with log_{10} Enterococci versus the other wetland variables using canonical variant analysis. Pooled within-groups correlations between discriminating variables and standardised canonical discriminant functions. Variables ordered by absolute size of correlation within function.

The CVA analysis highlights the effects of different variables on the three bacterial populations under analysis; total coliforms, *E.coli* and enterococci (log_{10} CFU/100ml), with redox potential, dissolved oxygen, pH and turbidity the negative trending physical-chemical variables which impact the three bacterial populations, see Table 3.3.

It should be noted that for the *E.coli* CVA, the total coliforms and enterococci bacteria were strong negative associations in the presence of sulphate. The potential exists that there may be a second source of bacteria entering the wetland system, therefore resulting in the dissimilarity between *E.coli* and total coliforms and enterococci.

Table 3.3: Canonical negative correlations

Total coliforms	***E.coli***	**Enterococci**
pH	sulphate	nitrate
dissolved oxygen	total coliforms	pH
redox potential	enterococci	dissolved oxygen
	turbidity	
	dissolved oxygen	

3.12 Principal component analysis of the wetland variables

The complex nature of the wetland system and the multitude of variables recorded from several different sampling points over a period of twelve months leads to issues of how to reduce this complex dataset to a number of key variables that represent an unbiased picture underlying the behaviour of the wetland.

Principal component analysis (PCA) is an approach to reduce the dataset to a lower dimension to reveal the hidden and simpler behaviour of the wetland dataset. The use of PCA within wetland research has been established by Travaini-Lima and Sipaúba-Taveres (2012), Dong and Reddy (2010) and Xu *et*

al (2010). The PCA method incorporates Promax 4 which is a non-orthogonal rotation, which allows correlations between the rotated variables i.e. to provide more meaningful analysis because very few variables are uncorrelated in reality (Hair *et al*, 1998). The eigenvalues of 1.0 of greater was selected within the programming PCA Promax 4 analysis and finally component values of 0.4 or greater were also selected to ensure only the most dominant variables are displayed within the analysis. Statistical software SPSS v22.0 was used to evaluate the data and find the most prominent wetland variables which best represent the entire wetland system for the duration of sampling.

The PCA in this chapter uses thirteen wetland variables: pH, redox potential, air and water temperature, conductivity, dissolved oxygen, BOD, turbidity, total bacteria, total coliforms, *E.coli*, enterococci and water depth as the data for the wetland sample points. The inorganic ion data was not used as it was incomplete (missing data points) and the wetland design data such as wetland volume and area were also not used. However, in chapter 10, all wetland data was used to investigate the wetland system.

Table 3.4: The PCA wetland variables

Descriptive Statistics			
	Mean	**Std. Deviation**	**Analysis N**
pH	7.425	0.617	60
Redox potential	-58.761	134.591	60
Air temp	14.120	6.172	60
Water temp	12.333	5.503	60
Conductivity	1.845	1.7505	60
Dissolved oxygen	3.459	3.158	60
BOD	911.799	1439.624	60
Turbidity	243.074	377.899	60
Total bacteria	6.253	2.263	60
Water depth	0.818	0.3614	60
Total coliforms	5.925	2.297	60
E.coli	5.363	2.103	60
Enterococci	5.077	2.163	60

Table 3.4 shows the PCA with thirteen wetland variables selected by the PCA function, with N = 60 for each variable. Therefore 780 data points were selected for analysis. Table 3.4 was used in conjunction with Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity (BTS) to indicate suitability of the wetland dataset for structure analysis.

Table 3.5: The Kaiser–Meyer-Olkin (KMO) and Bartlett's Test of wetland variables suitability and significance

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.830
Bartlett's Test of Sphericity	Approx. Chi-Square	10760.
	df	78
	Sig.	0.000

The KMO and Bartlett's test whereby the closer the KMO measure of the sampling adequacy is to the value 1.0, indicates the veracity of the factor analysis of the dataset. If less than 0.4 the factor analysis results will not be of any value. A value of 0.83 is a high value, therefore suitable for factor analysis. The significance value of $p < 0.05$ for the Bartlett's Test of Sphericity is a further validation that the dataset that can be used for PCA, the value of $p < 0.000$ i.e. data very significant with respect to factor analysis, see chapter 10, section 1 for more detailed information with respect to the KMO and Bartlett's test. The omission of the inorganic data and wetland design parameters are addressed in chapter 10, section 1.

Table 3.6: The wetland variance and Eigenvalues

Total Variance Explained					
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings	
	Total	% of Variance	Cumulative %	Total	% of Variance
1	7.978	61.367	61.367	7.978	61.367
2	1.828	14.058	75.425	1.828	14.058
3	1.009	7.761	83.186	1.009	7.761
4	0.899	6.918	90.104		
5	0.338	2.601	92.705		
6	0.313	2.404	95.110		
7	0.198	1.522	96.631		
8	0.161	1.238	97.869		
9	0.112	0.860	98.729		
10	0.081	0.620	99.349		
11	0.056	0.433	99.781		
12	0.025	0.190	99.971		
13	0.004	0.029	100.000		

The first three components of Table 3.6 account for 83.186% of the cumulative dataset, with eigenvalues greater than 1. These three components explain 83% of wetland variance. Component 1 accounts for 61%, therefore it is very dominant in explaining the wetland behaviour.

Table 3.7: The wetland component matrix

Component Matrix[a]			
	Component		
	1	2	3
Total bacteria	0.938		
E.coli	0.937		
Total coliforms	0.936		
BOD	0.933		
Redox Pot	-0.917		
Enterococci	0.872		
Conductivity	0.871		
Turbidity	0.839		
Dissolved oxygen	-0.757		0.445
Water Depth	0.651	-0.427	0.531
pH	-0.462		
Air Temp		0.947	
Water Temp	0.434	0.835	
Extraction Method: Principal Component Analysis.[a]			
a. 3 components extracted.			

The wetland component matrix, Table 3.7. Component 1, which dominates 61% of the wetland dataset, was the total bacteria. The wetland variables from the PCA analysis were BOD, redox potential, conductivity, turbidity, dissolved oxygen and water depth as the next dominate components of the wetland system i.e. these ten variables are designated as prominent variables and any further modelling scenarios will involve these variables. Component 2 shows that air temperature and water temperature increase the cumulative percentage to 74.4% and finally the third component pH again increases the cumulative percentage to 83%, its corresponding component value = 0.4. The water depth, dissolved oxygen and water temperature components are shared between either two or three components. In water depths case between three components or between two components for dissolved oxygen and water temperature respectively.

Table 3.8: The wetland pattern matrix

Pattern Matrix[a]			
	Component		
	1	2	3
Dissolved oxygen	**-1.062**		
Enterococci	0.876		
Total bacteria	0.860		
Total coliforms	0.819		
E.coli	0.699		
pH	-0.665		
Redox potential	-0.538	-0.466	
Water depth		**1.173**	
BOD		0.859	
Conductivity		0.749	
Turbidity		0.735	
Air Temp			**.996**
Water Temp			.941
Extraction Method: Principal Component Analysis. Rotation Method: Promax with Kaiser Normalisation.[a]			
a. Rotation converged in 5 iterations.			

The pattern matrix within the wetland variables, which are the regression coefficients for the variables are shown in Table 3.8. The three dominant patterns regression coefficients for Component 1 = dissolved oxygen, for Component 2 = water depth and for Component 3 = air temperature.

Table 3.9: The wetland structure matrix

Structure Matrix			
	Component		
	1	2	3
Total bacteria	0.959	0.759	
Total coliforms	0.949	0.765	
E.coli	0.919	0.816	
Enterococci	0.907	0.698	
Redox Potential	-0.864	-0.848	
Dissolved oxygen	-0.862	-0.498	
pH	-0.528		
BOD	0.787	0.976	
Conductivity	0.735	0.884	0.448
Water Depth	0.432	0.872	
Turbidity	0.716	0.866	
Water Temp			0.959
Air Temp			0.924

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalisation.

The structure matrix of the wetland variables, Table 3.9 shows the correlations between the variables and components. The structure matrix reveals that (1) redox potential and (2) dissolved oxygen variables, display an inverse relationship to the bacterial loadings (total bacteria, total coliforms, *E.coli* and enterococci) and other wetland physical-chemical variables such as BOD, conductivity, water depth and turbidity, with redox potential dominating within component 1 and 2.

3.13 Principal wetland variables

The key variables that best represent wetland behaviour are: water depth (m), conductivity (mS/cm), turbidity (NTU), total bacteria (log_{10} CFU/100ml), BOD (mg/L), turbidity (NTU), dissolved oxygen (mg/L), water temperature (^{0}C) and redox potential (mV). Using PCA as a data reduction technique the wetland dataset was reduced to seven key variables. This technique was of critical importance in chapter 10 in order to perform fuzzy indexing analysis as part of fuzzy Bayesian analysis. Five variables were selected in order to present the wetland system. The variables finally selected were: water depth (m), dissolved oxygen (mg/L), turbidity (NTU), conductivity (mS/cm) and total bacteria (log_{10} CFU/100ml). The five variables were selected on the basis that all five have different units, therefore no inherent bias and they include three distinct types of the wetland data set: physical-chemical, bacteria and wetland design.

3.14 Canonical Discriminant Analysis

Another statistical method employed was canonical discriminant analysis (CDA), where CDA is related to Principal Component Analysis (PCA) in that both are dimension reduction techniques. With PCA the object is to reduce a large set of variables to a smaller set (dimension reduction) that accounts for the majority of variation within the initial dataset. With CDA the object is to separate the variables into distinct groups or clusters (dimension reduction) based on a specific target variable within the dataset. The dataset obtained from the wetland contains multiple different environmental variables over a twelve-month sampling regime and the relationships between these variables are complex. CDA can be used to interpret the spatial distribution of these multiple environmental parameters (Cruz-Castillo et al., 1994; Momen and Zehr, 1998; Comber et al., 2005; Liao and Chang, 2005). The categorical variable of interest within the CDA procedure is the log_{10} total bacteria concentration and its relationship between the wetland dataset. Specifically, to know how many dimensions we would need to express this relationship between total bacteria concentration and the remaining wetland dataset. By using this relationship, the CDA method can predict a classification

based on the continuous variables or assess how well the continuous variables separate the categories in the classification. The initial or category wetland variable to assess the clusters within the CDA was the $\log_{10}$ total bacteria concentration, the variable was used as a tracer function within the wetland system. Using bacteria as tracers within water and soil based systems has been used by Pillai et al., (2003) to observe the transport and survival of bacterial and viral tracers through a submerged flow constructed wetland and Gao *et al*., (2013) used enterococci bacteria to model the importance of sediment effects within the Severn Estuary. The CDA method then reviews the dominant discriminant functions (DF_1 and DF_2) and the CDA graphs reveal the spatial distributions of the tracer (total bacteria) super-imposed by the discriminant functions. In the first part of the analysis discriminant function analysis (DFA), which was used to determine which continuous variables, discriminates between two or more naturally occurring groups. For example, investigate which wetland variables discriminate between the total bacteria concentration (category) and between the sample points within the wetland. The second part of the analysis was the development of CDA graphs indicating the spatial distribution of total bacteria and the sample points within the wetland system. The discriminant analysis could be used to determine which variables are best predictors of total bacteria concentrations within the wetland. DA is very similar to linear regression by predicting an outcome. Discriminant function analysis is broken into a 2-step process: (1) testing significance of a set of discriminant functions, and; (2) classification. The first step is computationally identical to MANOVA. Discriminant function analysis (DFA) is multivariate analysis of variance (MANOVA) reversed. In MANOVA, the independent variables are the groups and the dependent variables are the predictors. In DFA, the independent variables are the predictors and the dependent variables are the groups. As previously mentioned, DFA is usually used to predict membership in naturally occurring groups. It answers the question: can a combination of variables be used to predict group membership? Usually, several variables are included in a study to see which ones contribute to the discrimination between groups. This is the case in this chapter, where multiple variables are involved in “step wise method” DFA. Probably the most common application of discriminant function analysis is to include many protocols in the analysis, in order to determine the protocols that

discriminate between groups. For example, an environmental researcher interested in predicting total bacterial concentration dispersion within the wetland system by entering multiple variables in order to learn which variables offer the best prediction.

Table 3.10: The Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
pH	0.652	2.6	8	39	0.022
Redox Potential	0.209	18.423	8	39	0.000
Water Temp	0.734	1.769	8	39	0.113
Air Temp	0.781	1.368	8	39	0.241
Conductivity	0.275	12.864	8	39	0.000
Dissolved oxygen	0.48	5.278	8	39	0.000
BOD	0.188	21.117	8	39	0.000
Turbidity	0.344	9.305	8	39	0.000
Log10 T.Coliforms	0.052	88.957	8	39	0.000
Log10 *E.coli*	0.176	22.79	8	39	0.000
Log10 Enterococci	0.136	30.962	8	39	0.000
Water Depth	0.47	5.498	8	39	0.000
Wetland area	0.352	8.955	8	39	0.000
Wetland volume	0.698	2.112	8	39	0.058
Fluoride	0.44	6.195	8	39	0.000
Chloride	0.407	7.101	8	39	0.000
Nitrate	0.738	1.726	8	39	0.123
Nitrite	0.524	4.435	8	39	0.001
Phosphate	0.411	6.972	8	39	0.000
Sulphate	0.736	1.748	8	39	0.118
Ammonia	0.345	9.273	8	39	0.000

The Tests of Equality of Group Means, the results of the ANOVA were carried out on each variable with the group under test. There are two dominant groupings which indicate significant difference in the group testing. The first group shown in red indicate a large *F* number and a small Wilks' lambda. These variables show a large difference within the group and are important. The second group shown in green, reveal the same as the red group but to a lesser extent. Critical variables are redox potential, conductivity and Log_{10} (total Coliforms, *E.coli* and enterococci) concentrations with turbidity, wetland area and ammonia the next most important.

Table 3.11: The DFA structure matrix

	Function						
	1	2	3	4	5	6	7
Air Temp(a)	0.125	-0.307(*)	0.069	0.201	-0.14	-0.188	-0.214
Water Temp(a)	0.206	-0.285(*)	0.241	0.144	-0.052	-0.185	-0.058
Conductivity(a)	0.255	-0.427	0.321	.535(*)	0.24	-0.108	0.079
Chloride(a)	0.13	-0.213	0.264	.424(*)	-0.11	0.365	0.255
Nitrate	-0.01	0.127	-0.1	-0.501	.811(*)	0.246	0.075
Ammonia	0.183	-0.171	-0.4	0.234	.735(*)	0.33	-0.272
Redox Pot(a)	-0.19	0.371	-0.18	-0.033	-.522(*)	-0.265	-0.09
pH(a)	-0.04	0.022	-0.03	0.029	-.416(*)	-0.223	0.06
Phosphate(a)	0.104	-0.212	-0.22	0.344	.405(*)	-0.1	0.119
Dissolved oxygen(a)	0.018	-0.025	-0.08	0.056	-.389(*)	-0.126	-0.008
Enterococci	0.371	0.009	-0.02	0.008	-0.418	.800(*)	0.219
T.Coliforms	0.627	-0.257	-0.18	0.021	0.125	-.699(*)	-0.076
Turbidity(a)	0.012	-0.461	0.192	0.086	-0.04	.569(*)	0.216
Wetland volume(a)	-0.2	0.092	-0.16	-0.031	0.121	.261(*)	-0.006
Sulphate	0.073	0.001	0.267	0.177	0.104	0.613	-.711(*)
Water Depth(a)	0.17	-0.217	0.078	0.144	0.294	0.094	.676(*)
BOD	0.255	-0.603	0.233	0.032	0.262	-0.054	.667(*)
Fluoride	0.128	-0.242	0.464	0.389	0.192	-0.343	.636(*)
Wetland area(a)	-0.37	0.235	-0.08	-0.202	-0.273	0.064	-.585(*)
Nitrite(a)	0.368	0.001	0.342	0.141	0.194	-0.212	.583(*)
Log10 E.coli(a)	0.397	-0.374	-0.07	-0.041	0.382	-0.168	.427(*)

Variables ordered by absolute size of correlation within function.

(*) Largest absolute correlation between each variable and any discriminant function. Note: (a) denotes variables omitted from the analysis.

The DFA structure matrix provides another way of indicating the relative importance of the predictors and it can also reveal “grouping” of variables. The variables highlighted in red will be used in the DFA and CDA, they reveal the largest correlations within the Structure matrix.

The following tables (Wilks’ Lambda, Eigen values and standardised canonical discriminant function coefficients) reveal which variables used in the analysis and their importance to the wetland.

Table 3.12: Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 7	0.001	262.721	56	.0001
2 through 7	0.055	112.775	42	.0009
3 through 7	0.250	53.988	30	.005
4 through 7	0.474	29.155	20	.085
5 through 7	0.806	8.422	12	.751
6 through 7	0.934	2.644	6	.852
7	0.972	1.092	2	.579

Wilks'Lambda indicates that seven variables are selected and that the combination of variables one to seven have a Wilks' Lambda value = 0.001, therefore very significant $p < 0.0001$.

Table 3.13: The discriminant analysis Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	45.748(a)	89.600	89.6	0.989
2	3.515(a)	6.900	96.4	0.882
3	0.890(a)	1.700	98.2	0.686
4	0.702(a)	1.400	99.6	0.642
5	0.160(a)	0.300	99.9	0.371
6	0.041(a)	0.100	99.9	0.198
7	0.028(a)	0.100	100.0	0.166

Note: (a) denotes the first d=seven variables used in the analysis.

The discriminant analysis Eigen values evaluates the percentage of variance and the cumulative percentage, with the first variable contributing to 89.6 % variance within the wetland with a very large eigenvalue (45.748). The second variable with 6.9% variance and Eigenvalue 3.515. The discriminant analysis tells the user that the first seven canonical discriminations are used in the analysis, which accounts for Cumulative variance = 1.0.

Table 3.14: The standardised Canonical Discriminant Function coefficients

	Function						
	1	2	3	4	5	6	7
BOD	-0.56	-2.120	0.050	-0.770	-0.210	0.211	-0.08
T.Coliforms	1.175	0.222	0.103	-0.270	-0.010	-0.460	-0.310
Enterococci	0.953	0.734	-0.160	0.125	-0.250	0.530	0.490
Fluoride	0.568	1.456	0.747	1.086	0.533	-0.150	0.563
Nitrate	0.481	0.836	0.495	-0.800	0.638	0.069	0.191
Sulphate	0.266	-0.35	1.101	-0.160	0.11	0.195	-0.82
Ammonia	-0.180	-0.12	-1.030	1.010	0.434	0.239	0.118

The standardised Canonical Discriminant Function coefficients shows the seven variables used and their corresponding functions coefficients.

These standardised coefficients (beta (b)) are used to create the discriminant function (equation). It operates just like a regression equation. The Discriminant Functions 1 and 2 (DF_1) and (DF_2) are shown below:

DF_1= (-0.54*BOD) + (1.175 * Log_{10} total coliforms) + (0.953 * Log_{10} enterococci) + (0.568 * Fluoride) + (0.481 * Nitrate) + (0.266 *Sulphate) + (-0.175 *Ammonia).

DF_2= (-2.12*BOD) + (.222 * Log_{10} total coliforms) + (0.734 * Log_{10} enterococci) + (1.456 * Fluoride) + (0.836 * Nitrate) + (-0.34 *Sulphate) + (-0.118 *Ammonia).

The discriminant function coefficients indicate the partial contribution of each variable to the discriminate function controlling for all other variables in the equation. If there are any missing variables, then evaluated equations won't work. If there are any dummy variables as in regression; individual beta weights cannot be used.

Table 3.15: The seven discriminant variables:

BOD
Log_{10} total coliforms
Log_{10} enterococci
Fluoride
Nitrate
Sulphate
Ammonia

Table 3.15: The seven discriminant variables, show that the wetland system was distinguished by two discriminant functions, DF_1 and DF_2 and was defined by seven discriminant variables as shown above. The CDA graphs were then constructed from these discriminant variables.

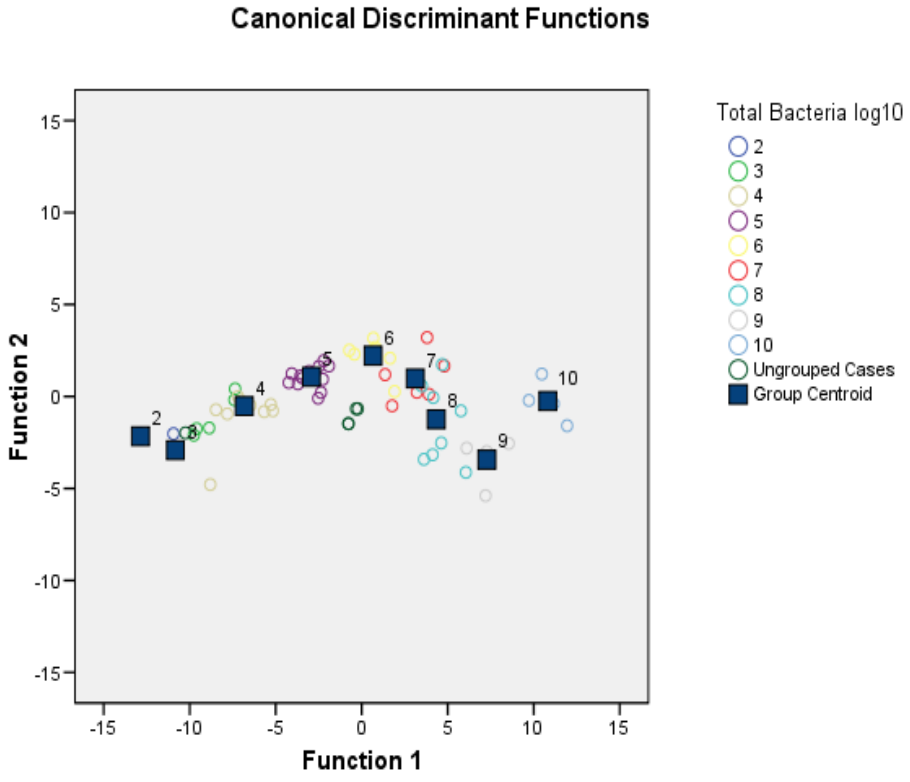


Figure 3.17: The canonical discriminant function or the spatial distribution for total bacteria. For DF_1 = Function 1 and DF_2 = Function 2.

A maximum input of $\log_{10}$ (10) CFU/100ml i.e. total bacteria $\log_{10}$ (CFU/100ml) as seen at sampling input point DAF. DAF to a minimum output of $\log_{10}$ (2) CFU/100ml i.e. total bacteria $\log_{10}$ (CFU/100ml) as seen at sampling output points pond 12 and stream.

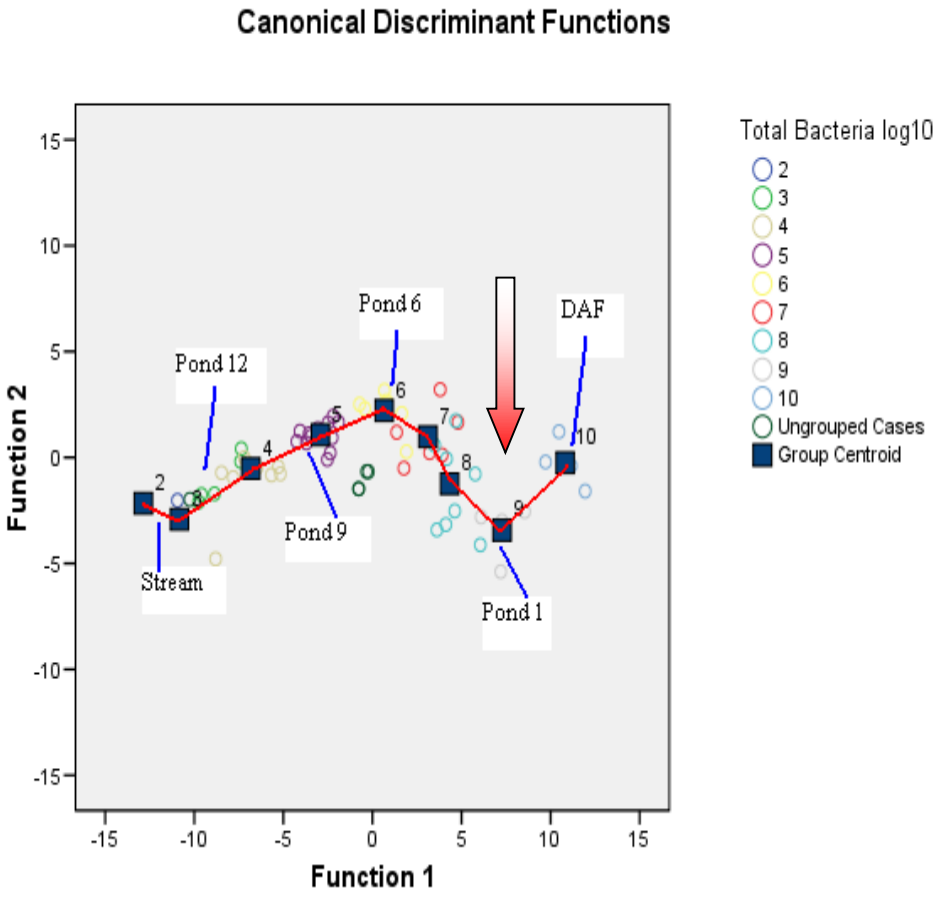


Figure 3.18: The discriminant functions of the total bacteria within the wetland system.

Highlighted are the sample point locations within the wetland. The arrow indicates a potential basin within the system, where the sample points (pond 1-pond 6) was resilient in treating the waste from the abattoir. Where basin or basin of attraction is a term associated with resilience of a complex system. The more basins present the more resilient the system. See chapter 10 for futher explanation.

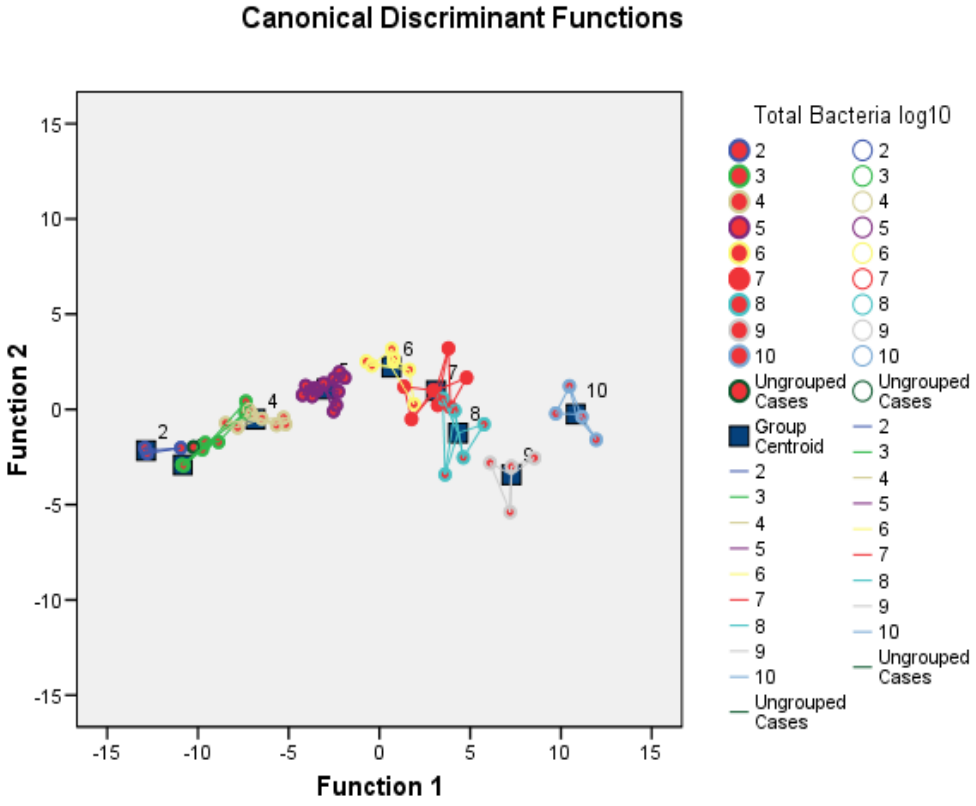


Figure 3.19: The canonical discriminant analysis with centroid clusters of the total bacteria.

The centroid or grouping cases are selected between 10 to 2, where 10 denotes $\log_{10}$ 10 (CFU/100ml) i.e. DAF (input) and 2 denotes $\log_{10}$ 2 (CFU/100ml) i.e. the stream (output). The shared grouping exists between points 8, 7 and 6 and 2, 3 and 4. But 10, 9 and 5 are isolated. Where 10, 9 and 5 denote the wetland sample points DAF, pond 1 and pond 9. The isolated grouping in DAF, pond 1 and pond 9, could indicate potential issues within the wetland, and maybe indicative of stagnant behaviour, i.e. very poor mixing (in chapter four the effect of mixing was explored in greater detail, within the wetland).

3.15 Discussion

Several statistical methods were employed in this chapter to elucidate important variables within the wetland dataset. Along with identifying wetland variables, issues relating to the individual sample points and the overall constructed wetland performance were highlighted. The normalised S-distribution analysis, see Figure 3.9, shows a plateau between ponds 9 and 12. One hypotheses of the thesis was to analyse this effect and verify that this plateau was real and complementary of the wetlands performance or as an artefact of an underlying issue with the system. Reviewing the mean wetland data from Figures 3.10, 3.11, 3.12 and 3.13 for the physical-chemical, indicator bacteria and inorganic ion data respectively the following was found. The dissolved oxygen concentrations (mg/L) at pond 9 and 12 shows a slight decrease between these ponds with values of 5.87 and 5.74 respectively (an increase in dissolved oxygen content was expected, as the wastewater progress through the CW). Similarly, the turbidity values (NTUs) at ponds 9 and 12 show an increase between these ponds, with values of 52.14 and 77.22 respectively, see Figure 3.11. The inorganic ion concentrations (mg/L) of phosphate at pond 9 and 12 show a slight increase between these ponds, with values of 19.06 and 22.17 respectively. Similarly, the nitrate values (mg/L) between pond 9 and 12 show an increase, with values of 39.13 and 62.59 respectively. The mean indicator bacteria data, see Figure 3.12, displays none of the increases, with respect to pond 9 and 12.

When reviewing the two-way ANOVA data without replication, see Table 3.2, the ANOVA shows a 'switch' in the behaviour of the wetland data subsets, (1) physical-chemical, (2) indicator bacteria and (3) inorganic ions. The two-way ANOVA without replication ($p < 0.05$) was performed row by columns, where row represents the ponds and the columns the entire wetland. Both the physical-chemical and the inorganic data behave similarly with row (pond) *p*-values of 0.0037 and 0.0023 respectively i.e. the ponds were very significant in interacting with these wetland variables. When reviewing the column (entire wetland), these two wetland data subsets reveal values of 0.0000 respectively i.e. the wetland system was very significant interacting with these variables. Conversely the indicator bacteria revealed a row (pond) *p*-value of 0.0000 and column (wetland) p-values of 0.0446. The indicator bacteria have a much more significant response

by pond than they do to the constructed wetland. The initial ANOVA *p*-value was set to $p < 0.05$, for rows and columns, the column *p*-value was just under this threshold. Due to the observed ANOVA behaviour of the indicator bacteria, the next statistical test, canonical variant analysis (CVA) was used to explore interactions between bacterial groups and other variables within the wetland. The CVA test was used to define the most probable wetland variables that interact with the indicator bacteria variables. The analysis revealed the dominance of physical-chemical parameters: dissolved oxygen, pH, redox potential, turbidity and inorganic ions parameters nitrate and sulphate. It was observed that *E.coli* was impacted negatively by total coliforms and enterococci bacterial populations. Potential exists that there was a second route of *E.coli* entering the CW, either by another entry of wastewater into the CW or the local wild fowl population (water hens and ducks) may have contributed to the asymmetry in bacteria correlations. Principal component analysis was employed as a more advanced statistical tool to explore the wetland dataset and extract the major principal components. The use of PCA with Promax non-orthogonal rotation which provided more meaningful analysis because very few variables are uncorrelated in reality. The use of internal tests such as KMO and Bartlett's test validated the veracity of the factor analysis of the dataset. The Bartlett's test indicates the significance value of variables selected for PCA, in finding the most suitable variables that best represent the dataset under analysis. The wetland system was best defined by five key variables: - water depth, total bacteria, turbidity, conductivity, and dissolved oxygen. But the PCA reveals the dominance of the indicator bacteria variables in the analysis with the total bacteria, *E.coli* and total coliforms displaying a total variance of 83%, see Table 3.6. This analysis reveals the dominance of the indicator bacteria within the wetland system.

CDA was demonstrated as a useful tool for classifying the wetland system, via discriminant functional analysis (DFA). This method was used to establish a graphical model of the $\log_{10}$ total bacteria concentrations by sample point. Since the PCA confirmed the dominance of total bacteria with an Eigen value of 7.978 and 61.37% variance (see Table 3.6). This variable (total bacteria) was then selected as the categorical variable in the canonical discriminant analysis. When analyising the wetland dataset using CDA analysis, it was decided to pre-select $\log_{10}$ total bacteria as the category variable within the wetland system, due to the

dominance of bacterial populations within previous statistical analysis. The CDA analysis picked seven wetland variables at random to form the discriminant functions (DF1 and DF2). They were BOD, $\log_{10}$ total coliforms, $\log_{10}$ enterococci, fluoride, nitrate, sulphate, and ammonia, see Table 3.15. When DF1 and DF2 were graphed, to form the graphical model, see Figure 3.18, where the categorical variable was super-imposed onto the DF1 and DF2 functions via the seven randomly selected wetland variables. The graph reveals a potential "basin" within the wetland system, where pond 1 to pond 6 is resilient i.e. the wetland shows potential resilience or robustness at these sample points. Where resilience is the ability of the constructed wetland to withstand shocks and disturbances and return its function i.e. treating wastewater. A secondary analysis was carried out in the CDA function, see Figure 3.19, where we analysed for the grouping (cluster) effects of the categorical variable ($\log_{10}$ total bacteria) including the randomly selected variables across the wetland system. The grouping was selected between $\log_{10}10$ (CFU/100ml) to $\log_{10}2$ (CFU/100ml) of the categorical variable i.e. between the DAF and the stream. The graph shows the grouping/centroid effects of the wetland via $\log_{10}$ total bacteria inclusive of the seven randomly selected variables. Issues at the front of the constructed wetland where the DAF (10) and pond 1 (9) display no cluster mixing, these clusters are isolated. A similar behaviour was observed at pond 9 (5). This cluster isolation could pertain to stagnant response at the points (no mixing). The CDA shows important differences within the wetland system using different wetland variables and reducing the dimensionality of the dataset from twenty-two to seven variables, while maintaining the complex interaction among the wetland variables.

3.16 Conclusions

The statistical tests employed in this study revealed that the normalised S-Distribution – a plateau exists from pond 9 to pond 12. Further work in this thesis validates or invalidates the plateau as a natural effect of the CW or that issues exist from pond 9 to pond 12.

The ANOVA (two-way without replication), showed that the inorganic and physical-chemical parameters displayed similar statistical traits but the indicator bacteria was the inverse of these two wetland data subsets. Where individual ponds were more significant in removing bacterial concentrations than the entire wetland system.

The canonical variate analysis (CVA) explored the interactions of the bacterial populations and the other wetland variables. The analysis revealed the dominance of physical-chemical variables such as pH, turbidity, dissolved oxygen and redox potential. The CVA revealed that potentially a second source of *E.coli* was entering the wetland system.

Principal component analysis (PCA) confirmed the dominance of the indicator bacteria parameters within the CW, along with water depth, dissolved oxygen, BOD, turbidity and redox potential.

The canonical discriminant analysis (CDA) inclusive of discriminant functional analysis (DFA) was used to develop graphical models of the behaviour of the wetland. The analysis revealed the wetland exhibits potential resilience in the front –end of the CW, between ponds 1 to 6. The potential also exists that stagnant areas exist within the CW in DAF, pond 1 and pond 9.

Chapter 4. Constructed wetland: Traditional models.

A comparison of different model methodologies using indicator bacteria as tracers, to predict the hydraulic retention time of the constructed wetland treating abattoir wastewater: Introducing the concept of long circuiting.

4.1 Summary

This chapter compares several traditional wetland models used for the prediction of the Hydraulic Retention Time (HRT) using indicator bacteria within a free-water constructed wetland. The two main models used were the Continuous Stirred Reactor (CSTR) model based on work done by Mara and Marais and the Oakley Dispersed Flow (DF) model. In addition, a third model was introduced, the new Oakley Dispersed Flow DF model, which maintains the basic structure of the Oakley DF model but uses different dispersion numbers and die-off coefficient formulae (rate constant k). On the basis of the modelling methodologies undertaken it was found that the new Oakley DF model and the CSTR models were more accurate in predicting the HRTs of the indicator bacteria trends within the constructed wetland. The mean retention times were found during the dry summer months of 128 and 116 days respectively, in comparison to the Oakley DF model which predicted a value of 88 days. Two distinct dynamics were observed in the wetland via the HRTs of the indicator bacteria, short-circuiting and its counterpart long-circuiting. The short-circuiting concept is where large volumes of water flows through the wetland with large dispersion numbers which results in short HRTs.The long-circuiting concept is where poor water flows, low water depths in conjunction with low dispersion numbers give rise to large HRTs, with the potential of increased bacterial loadings.

4.2 Introduction

Pathogen removal from wastewater systems is very important to public health and to the local environment, in particular when the effluent being treated emanates from an abattoir. Constructed wetlands utilising a series of interconnected waste ponds are an alternative "green" treatment system. They are suited for removal of pathogens, nutrients and organic materials from the wastewater. The removal of pathogens is typically assessed by estimating the reduction of indicator organisms such as total coliforms, *E.coli* and enterococci concentrations. The effect of climate and the wetland design parameters has a bearing on the efficacy of the wetlands ability to reduce wastewater loadings.

The design of constructed wetlands for bacterial removal are attributed to methods developed by Marais (1974), who proposed that bacterial removal of Faecal coliforms (FC) can be modelled by first-order kinetics in a completely mixed reactor. This approach was later integrated into design manuals by Mara *et al*., (1992); Mara (1997); Mara and Pearson (1998) and slightly improved upon (Mara *et al*., 2001). These design manuals were for warm temperate climates and took into account flow conditions, for the Middle East, India and the Mediterranean countries. The work highlighted above by Marais (1974) and Mara *et al*., (2001) found recognition in the English and Spanish editions of Wastewater Engineering books by Metcalf & Eddy (1991, 1996), with Latin-American design manuals in particular from Yánez (1992) and Oakley (1997). A comparison of the models was carried out, with the models developed by Mara *et al*., (2001) and Oakley (1997) being of particular interest within this chapter. A potentially new Oakley dispersed flow modelling method was developed as a result of Buchauer (2007), who compares CTSR and DF model regimes. Within the conclusion section of Buchauers' paper (2007); point 6 is of importance to this chapter. It states the following: If the DF model should become a more reliable tool, the calculation of its underlying parameters, particularly of the rate constant k and the dispersion number d, will require more research. It appears that the present state of the art to calculate these parameters is not yet satisfactory. Buchauer (2007), concludes that the CSTR model was better than the DF model, due to fact that the DF model was "too optimistic" in comparison to the CSTR model which matches measured reality regarding faecal coliforms loadings. The aim of this chapter was

to introduce a new Oakley DF model which develops new techniques for both the rate constant k and the dispersion number *d*, using indicator bacteria (total coliforms, *E.coli* and enterococci) concentrations from specific sample points (DAF, pond 1, pond 6, pond 9 and pond 12) within the constructed wetland system in the development of bacterial tracers models. Pillia *et al*., (2003) used viral and pathogenic bacteria (*Salmonella*) as tracers within a submerged flow constructed wetland and Sleytr *et al*., (2007) concluded that "to investigate bacterial fate in detail, bacterial tracer experiments have to be carried out", while investigating bacterial removal in constructed wetlands. Boutilier *et al*., (2009) investigated the adsorption, sedimentation and inactivation of *E.coli* within wastewater emanating from constructed wetlands, where the "bacteria fate and transport within constructed wetlands must be understood if engineered wetlands are to become a reliable form of wastewater treatment." The constructed wetland in this study was a free water surface integrated constructed wetland consisting of a dissolved air floatation (DAF) plant at the wastewater source and twelve ponds in-series treating the abattoir wastewater through a combination of different plants species (integrated), located on approx. 2.8 hectares of land. The wetland system may be divided into different component parts, such as anaerobic, facultative and maturation (polishing) ponds with wetland ponds designs in-series or in parallel (Pescod and Mara, 1988). In many cases anaerobic and facultative ponds are enough for wastewater treatment, but depending on the type of effluent and its final destination, the maturation ponds are provided for further polishing purposes (Phuntsho *et al*., 2008). The wetland system under investigation was designed as follows: Ponds 1 to 4 anaerobic, ponds 5-8 facultative and ponds 9-12 polishing.

The importance of two distinct dynamics are observed in the wetland via the hydraulic retention times (HRTs) of the indicator bacteria, total coliforms, *E.coli* and enterococci (1) short-circuiting, which occurs when a large fraction of water travelling through a system exits well before the resident time, reduces the performance of the constructed treatment wetlands, Lightbody *et al* ., (2009);

(2) A new concept called long-circuiting is introduced, as a juxtaposition to short-circuiting, which occurs when a small fraction of water travelling through a system exits well after the resident time, reduces the performance of the

constructed treatment wetlands. Evidence is presented, that the wetland displays these two dynamics during the sampling period.

4.3 Hydraulic retention time calculations

The company tasked with the design and construction of the wetland, designated a total hydraulic retention time (HRT) = 102 days, or 8.5 days per pond, based on the following equation (8):

$T_n = V/Q$ Eq(4.1)

Where V= water volume of the wetland (m^3), Q = volumetric inflow rate (m^3/day). The wetland holds approximately 27,000 m^3 of water with a T_n = 102, therefore the volumetric inflow rate equals 265m^3/day for the entire wetland system or 265 (m^3day) /12 (ponds) = 22m^3/day per pond. Convert 22m^3/day into litres/sec = 0.25 L/s. Measurements carried out during the sampling process recorded inflow values of 0.21 to 0.27 L/s (summer months) but during the wet months inflow values were much greater with values recorded between 0.40 to 0.45 L/s. With two occurrences of extreme wetland flow values recorded between 0.8 to 0.9 (L/s) in ponds 9 and 12 during the wet months (December and February).

An inflow value of 0.45L/s was recorded and translated back to a nominal retention time T_n:

0.45 L/s * 86.4 = 39 m^3/day per pond. *(86.4 conversion factor for L/s to m^3/day)*

39m^3 /day * 12 (ponds) = 468 m^3/day (total wetland system)

$T_n = V/Q$ = 27,000/ 468 = **57 days**

Repeating the process for 0.21L/s recorded during the summer months:

0.21 L/s * 86.4 = 18 m^3/day per pond

18 m^3/day * 12 (ponds) = 216 m^3/day (total wetland system)

T_n = V/Q = 27,000/ 216 = **125 days**

Repeating the process for 0.9 L/s recorded during the wet months (December and February):

0.9 L/s * 86.4 = 78 m^3/day per pond

78 m^3/day * 12 (ponds) = 936 m^3/day (total wetland system)

T_n = V/Q = 27,000/ 936 = **29 days**

Therefore, during the wet months (late autumn to late spring) the nominal retention time of the system can almost be halved from 102 days to 57 days' HRT. During the dry months (summer) the nominal retention time of the system can be increased from 102 days to 125 days' HRT. During the two extreme events, the nominal retention time of the system decreased to 29 days HRT.

4.4 Mechanisms for indicator bacteria removal

Indicator bacteria removal in constructed wetlands (CWs) is due to a combination of different effects, which depend on the type of CW, with additional factors such as wastewater temperature, pond depth, positioning of inlets and outlets and climate conditions such as wind direction (US EPA 1996, US EPA, 2000). The combination of all of these factors, introduce complex interactions under varying circumstances. The overall effect is the result of combinations of sedimentation, starvation, and high temperature, UV radiation (Pearson *et al*., 1995), where individual variables cannot be assessed qualitatively. Therefore, developing relatively simple modelling formulas based on the wetland dataset would be considered a more prudent approach.

4.5 The Continuous Stirred reactor model

As stated previously, the work of Mara *et al*., (2001) was based on upon Marais (1974). Both models schemes have the same theoretical background; they assume the Faecal Coliform (FC) removal can be modelled by a first order kinetic model in a CSTR (continuous stirred reactor). The resulting formulae for a single pond:

$$FC_e = FC_i / (1 + k_T * R) \qquad \text{Eq(4.2)}$$

Where FC_e = number of FC in the effluent (Colony Forming Units/100ml)
FC_i = number of FC in the influent (Colony Forming Units/100ml)
k_T = first order rate constant or FC removal or die-off rate (d^{-1})
R = Retention Time or Hydraulic Retention Time (HRT) (d)

The difference between Marais (1974) and Mara *et al.*, (2001) concerns the calculation of k_T. Both use a modified Arrhenius equation, but with different coefficients:

Marais (1974) $k_T = 2.6\ (1.19)^{T-20}$ Eq(4.3)

Mara *et al*., (2001) $k_T = 2.6\ (1.15)^{T-20}$ Eq(4.4)

Where T = wastewater temperature (^{0}C)

Equation (4.4) was recommended for shallow, short retention time, with ponds at temperature above 20^{0}C (Mara *et al*., 2001).
The above equations serve to aid the prediction of indicator bacteria removal from the constructed wetlands variables. The basic nature of equation (4.1) caused many designers to think that increasing pond depth and therefore increasing retention time will result in better pond performance, however that is not the case (Pearson *et al*., 1995, Mara *et al*., 2001).

4.6 The Oakley Dispersed Flow Model

The Oakley Dispersed Flow (DF) model (Oakley, 1997) was used in this research since it was considered to be based upon several parameter assumptions (Buchauer, 2007).

The resulting formulae for a first-order equation for dispersed flow:

$$FC_e = 4\ FC_i\ a\ e^{1/2d} / [(1+a)^2\ e^{a/2d} - (1-a)^2\ e^{-a/2d}] \quad \text{Eq(4.5)}$$

Where $a = (1 + 4\ k_T\ R\ d)^{½}$ Eq(4.6)

$d = (L/W) / [-0.261 + 0.254\ L/W + 1.014\ (L/W)^2]$ Eq(4.7)

$k_T = 1.1\ (1.07)^{T-20}$ Eq(4.8)

FC_e, FC_i, R, are the same variables used in Equation (4.2)

L/W = Pond (Length / Width) Ratio

Equation (4.5) can be used for all types of ponds, single or in series. The empirical formulae for the dispersion number d and the rate constant (die-off rate) k_T were developed by Yánez (1992) in a tracer study of 24 separate pond systems in Peru.

4.7 The new Oakley Dispersed Flow Model

The new Oakley dispersed flow model was developed as a result of the conclusion section based on Buchauer, (2007). The Oakley DF model structure remains *in-situ*, but changes to Dispersion number (d) and the rate constant (die-off rate) k_T are changed. The resulting formulae for a first-order equation for dispersed flow:

$$FC_e = 4\ FC_i\ a\ e^{1/2d} / [(1+a)^2\ e^{a/2d} - (1-a)^2\ e^{-a/2d}] \quad \text{Eq(4.9)}$$

Where $a = (1 + 4\ k_T\ R\ d)^{½}$ Eq(4.10)

$d = 2\ (D/UL) - 2\ (D/UL)^2\ (1 - e^{-(UL/D)})$ Eq(4.11)

Peclet No. $= 0.31\ (L/W) + 0.055\ (L/D)$ Eq(4.12)

Dispersion No. $= 1/\ (0.31\ (L/W) + 0.055\ (L/D)$ Eq(4.13)

k_t = Ln (N)/ Ln (N_0) Eq(4.14)

Equation (4.8) based on $N = N_0 e^{-(k_t)}$ Eq(4.15)

FC_e, FC_i, R, are the same variables used in Equation (4.2)

L/W = Pond (Length / Width) Ratio

L/D = Pond (Length/ Depth) Ratio

UL/D = Peclet Number[9]

D/UL = Dispersion Number

N_0 = Initial Concentration

N = Concentration determined after the Initial Concentration times the exponential (rate constant (- k_T))

Equation (4.11), the dispersion number (d) was determined according to Levenspiel, (1972). Equation (4.12), uses the condition to determine the Peclet number using length, width and depth values (Nameche and Vasel, 1998). In computational fluid dynamics methodology using dispersion coefficient, Peclet, number and dispersion number in waste stabilisation ponds was as described by Abbassi *et al*., (2010). The Peclet number is used in the study of transport effects in fluid flows and defined as the ratio of the rate of advection by the flow rate to the rate of diffusion of the same fluid. Dispersion coefficient (D m^2/sec) is used in the study of mass transfer affects and measures the rate of diffusion. Where a large dispersion means fast spreading (turbulent) of tracers (i.e. indicator bacteria concentrations), Small dispersion means slow spreading (laminar) and where dispersion = 0 thus no spreading and hence plug flow conditions (stagnant), see Table 4.1 for example of the dispersion coefficient for the total coliforms population within the wetland.

Kadlec (1994) for FWS wetland systems are characterised by a large amount of apparent dispersion, with $0.07 \leq D \leq 0.35$. The dispersion coefficient describes eddy transport of water elements both upstream, and downstream. In FWS wetlands, such mixing may not occur because flow is often predominantly laminar (Kadlec and Wallace, 2009).

[9] The Peclet number is the inverse of the Dispersion number.

Equations (4.7) and (4.8) relate to the rate constant (die-off rate) $k_{T.}$ Where the rate constant in the new Oakley DF model can be measured from the natural log difference between the sample points (i.e. Ln (Bacteria Pond 1)/ Ln (Bacteria pond 6); the bacteria values are known between the ponds under analysis, thus the rate constant k_T between the ponds can be determined. The rate constant encompasses the true value of the indicator bacteria and not by a modified Arrhenius equation as seen in Equations 4.2, 4.3 and 4.4.

Table 4.1: The dispersion coefficient (D- m^2/sec) values for total coliforms (using the new Oakley DF Model)

	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Nov	Dec	Feb-08
Pond 1	**0.31**	**0.298**	0.272	0.247	0.21	0.21	0.24	0.25	0.275	**0.28**	**0.32**
Pond 6	0.21	0.275	0.21	0.247	0.225	0.229	0.24	0.24	0.27	**0.28**	**0.28**
Pond 9	0.241	0.243	0.243	0.165	**0.097**	0.159	**0.14**	**0.08**	0.22	0.25	0.27
Pond 12	**0.34**	0.266	0.259	Dry	Dry	**0.097**	**0.09**	Dry	0.23	0.26	**0.32**

In Table 4.1 the highlighted values in red indicate trending towards stagnant conditions (D tending towards 0). Highlighted values in blue indicate tending towards turbulent flows (D greater than 0). The total coliforms dispersion values do lie within the apparent dispersion of FWS wetland systems, Kadlec (1994). Similar values were evaluated for *E.coli* and enterococci, not shown.

4.8 Monitoring of key wetland variables

During the sampling regime, the bacteria concentrations (CFU/100ml) were recorded for 11 months, along with the water temperature (^{0}C), the water depths (m), the presence / absence of algal blooms throughout the sampling period and the ponds aspect ratios. The constructed wetlands dimensions were retrieved from local company IE Consulting Engineers, who were tasked with the design of the wetland system, with drawings submitted to Carlow's Local Authority, Planning and Environmental offices. From these engineering drawing the ponds length and widths were recorded. Thus the length/width (L/W) ratios were calculated, see

Table 4.2. Tables 4.3, 4.4 and 4.5 records water depths, the presence of algal blooms and the indicator bacteria (HRTs) for the new Oakley DF model, the Oakley DF model and the Mara CSTR Model respectively.

Table 4.2: Length/ Width Ratios for the wetland sampling ponds

DAF = 1.54
Pond 1 = 2.59
Pond 6 = 2.65
Pond 9 = 2.22
Pond 12 = 2.19

Table 4.2 shows the length/ width aspect ratios for the wetland sampling sites as determined from IE Consulting Engineers. The importance of L/W ratio (aspect ratio) was to minimise short circuiting within the wetland system, where ponds should be longer than wider, i.e. L/W > 1, the larger the aspect ratio, the better the hydraulic behaviour of the wetland system (Champagne *et al.,* 2012). All ponds have L/W aspect ratio greater than 2.0. The water depth was recorded, using a meter stick at the various sample points throughout the sampling period, see Table 4.3. The mean wetland water depth was used in the analysis. During the recording of this variable some water depth values were very low, see Table 4.3, values highlighted in red. The inflow and outflow pipes in the ponds had a fixed height of approx. 0.4m. Therefore, when the water level falls below 0.4m, the pond water cannot move into the next pond, thus the wastewater was isolated (stagnant) within that pond.

In conjunction with the recording of the pond water depth, it was noticed that algal blooms occurred in the wetland and a record of these blooms was tabulated, see Table 4.4. The algal blooms were noticeable in ponds 1, 9 and 12; especially during the summer months.

Table 4.3: The water depth per sample point by sampling period

sample points /month	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Nov	Dec	Feb-08
Pond 1	0.87	0.82	0.72	0.63	0.63	0.6	0.63	0.63	0.73	0.83	0.91
Pond 6	0.72	0.72	0.75	0.62	0.58	0.56	0.59	0.59	0.69	0.75	0.75
Pond 9	0.79	0.67	0.68	0.42	0.23	0.4	0.36	0.19	0.6	0.72	0.79
Pond 12	1.1	0.77	0.74	0	0	0.23	0.22	0	0.65	0.73	1.02
Mean water depth	0.87	0.75	0.72	0.42	0.36	0.45	0.45	0.35	0.67	0.76	0.87
stdev	0.17	0.06	0.03	0.29	0.3	0.17	0.19	0.31	0.06	0.05	0.12

Values highlighted in red, indicated very low water or no water in these ponds.

Table 4.4: Algal blooms per sample point by month

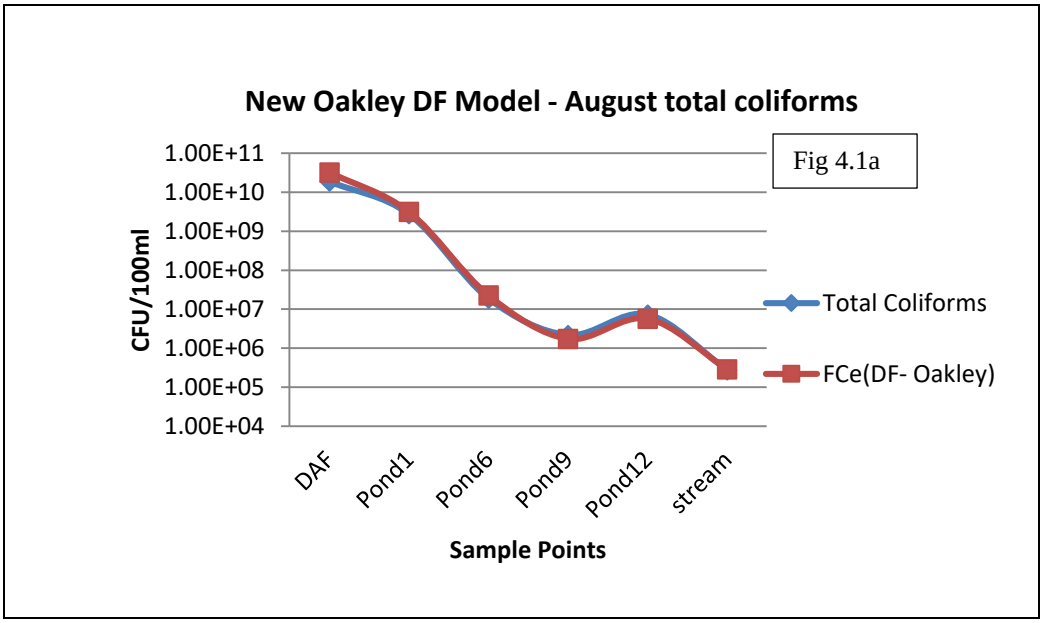
Algal Blooms (pond by month)	Pond 1	Pond 6	Pond 9	Pond 12
Feb	0	0	0	0
Mar	0	0	0	0
Apr	0	0	0	0
May	0	0	0	Dry
Jun	XX	0	X	Dry
Jul	XXX	0	XX	X
Aug	XX	0	X	0
Sept	X	0	X	Dry
Nov	0	0	0	0
Dec	0	0	0	0
Feb(08)	0	0	0	0

X = 0.5 – 1.0m^2 algal bloom
XX = 1.0 – 3.0m^2 algal bloom
XXX >> 3.0m^2 algal bloom

In Table 4.4 the presence / absence of algal blooms was recorded during the analysis of the wetland, 0 denotes no algal blooms and X indicates algal blooms were present. A single X denotes small, and multiple Xs refer to medium to very large algal blooms.

4.9 Comparison of the three models, new Oakley, traditional Oakley and the Continuous Stirred reactor (CSTR) models.

A comparison of the models can be seen in Figures 4.1a, 4.1b and 4.1c, where the new Oakley DF model, trends extremely close to the actual recorded data for the month of August's total coliforms. It should be noted that the month of August data typified the model observations of the sampling period. The Oakley DF model shows a slight deviation between the model and the actual data. In the case of the Mara/Marais CSTR models, the deviation between the actual bacteria data and the models was even larger. This behaviour was applicable across all the models over the sampling period. See Appendix C for further details on the model dataset.



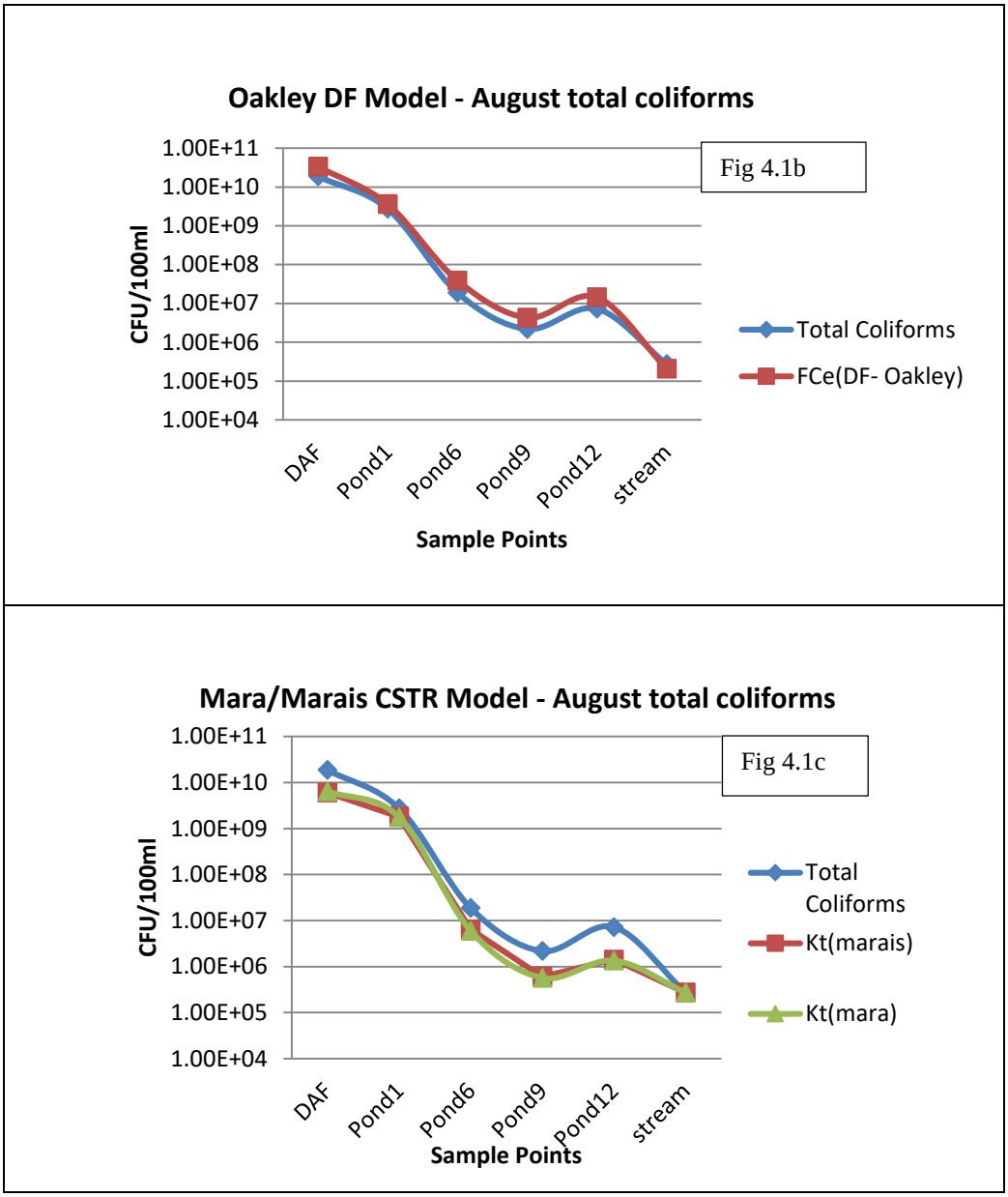


Figure 4.1 a, b, c: The three models for total coliforms for the month of August 2007. The new Oakley DF model 4.1a, the Oakley DF model Fig 4.1b. The Mara/Marais CSTR model Fig 4.1c.

Figures 4.2a and 4.2b shows the mean indicator bacteria $\log_{10}$ (CFU/100ml) versus the mean water depth and the mean water temperature. Which reveal the erratic nature of the bacterial parameter, but highlight the effect of loss of water within the ponds and the increase of water temperature has on the indicator bacteria concentrations. The noticeable dips in June and September are due to pond 12 having no surface water. Therefore, no bacterial loadings which directly correspond to data troughs as seen in Figures 4.2a and 4.2b.

Using the indicator bacteria Figures 4.3a and 4.3b show the new Oakley DF model versus the mean water depth (4.3a) and the mean water temperature (4.3b). The new Oakley DF model shows a good agreement with both the water depth and the water temperature. With very low water depths (dry months) an increase in HRT was observed, this increase also corresponds to the increase in water temperature during the dry months and the indicator bacteria reveal an organised pattern in their behaviour culminating in specific high and low HRTs. The Oakley DF model Figure 4.4a and 4.4b, reveals a more disorganised bacterial HRT pattern, with a spread of high and low HRT responses. The CSTR Mara and Marais models, behave very similar to the new Oakley DF model. Figures 4.5a and 4.5b, reveal an organised pattern, with specific high and low HRTs, but with a skewed response, during the dry months. Table 4.5 records the mean HRT data for all the models, including the original HRT of 102 days as per section 4.3. Figure 4.6 graphs the mean HRTs for all models.

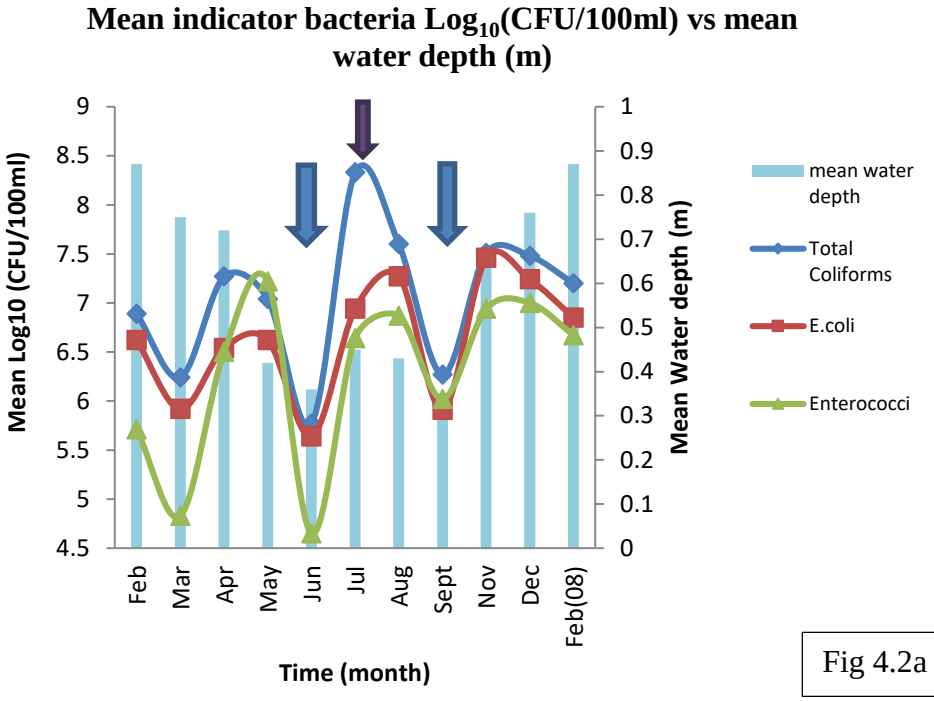


Fig 4.2a

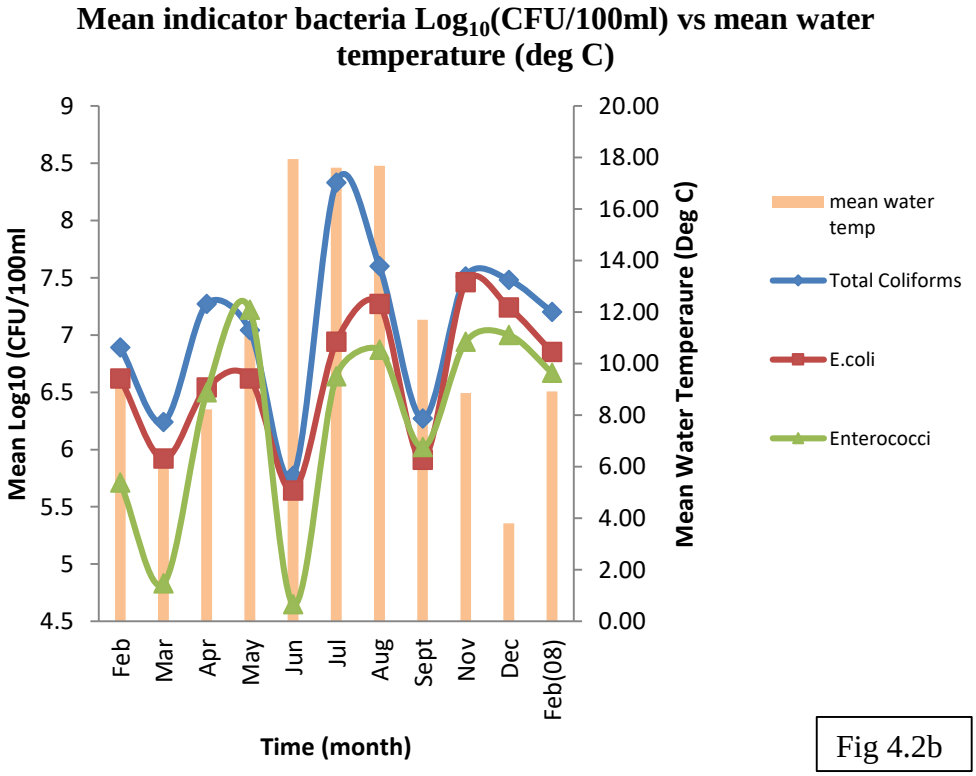


Fig 4.2b

Figure 4.2 a, b: The mean indicator bacteria versus mean water depth and water temperature.

A: The mean indicator bacteria versus the mean water depth.

B: The mean indicator bacteria versus the mean water temperature.

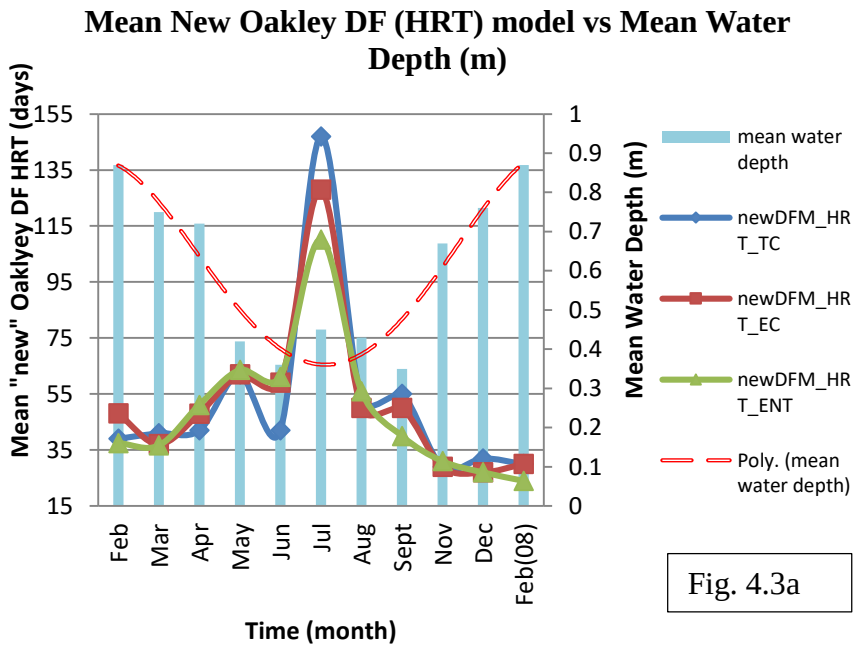


Fig. 4.3a

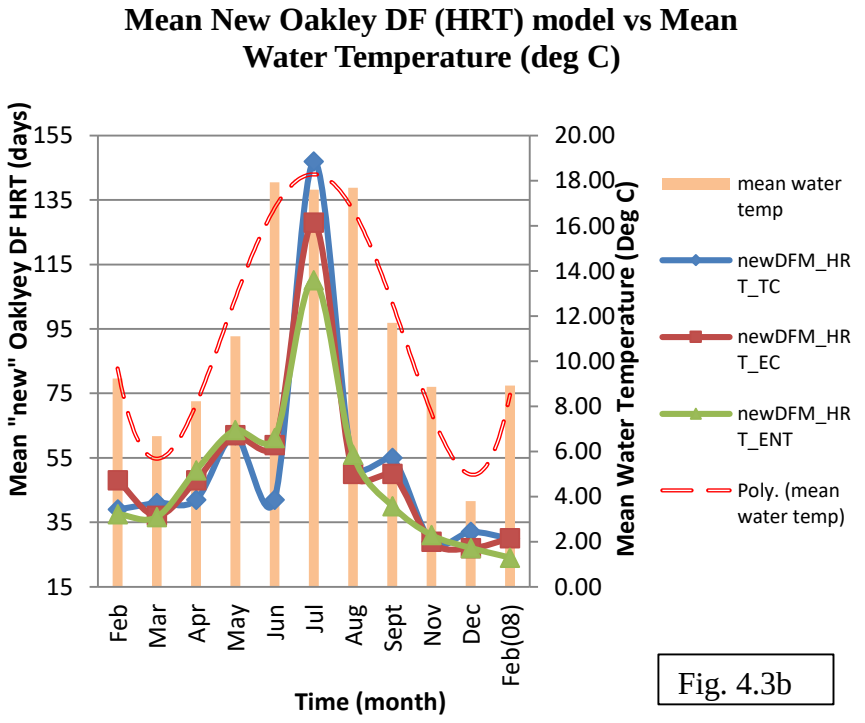


Fig. 4.3b

Figure 4.3 a, b: The new Oakley DF model versus the mean water depth and water temperature**.**

A: The mean new Oakley DF model versus the mean water depth.

B. The mean new Oakley DF model versus the mean water temperature.

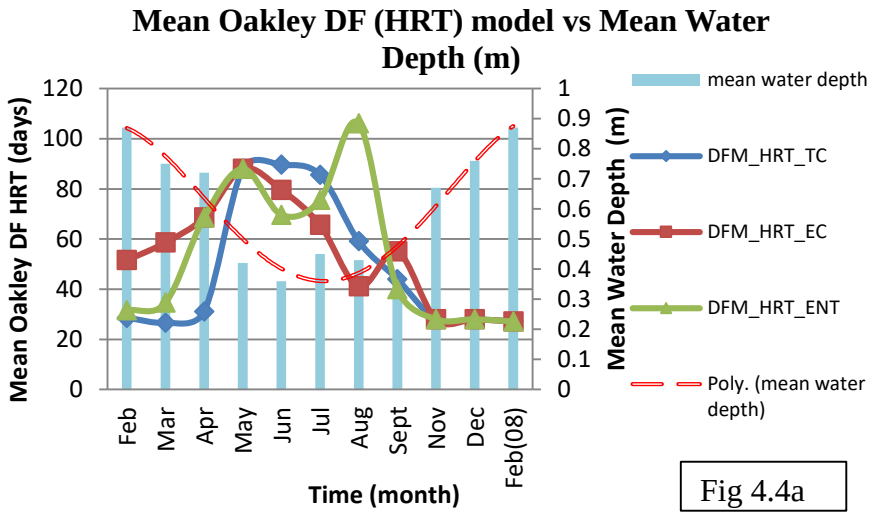


Fig 4.4a

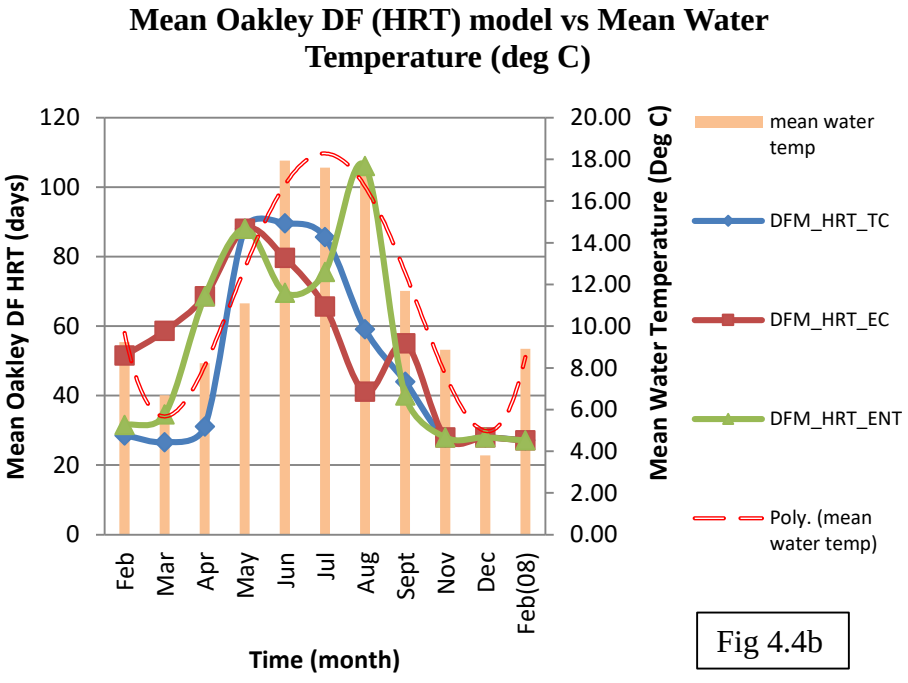


Fig 4.4b

Figure 4.4 a, b: The Oakley DF model versus the mean water depth and water temperature.

A: The mean Oakley DF model versus the mean water depth.

B: The mean Oakley DF model versus the mean water temperature.

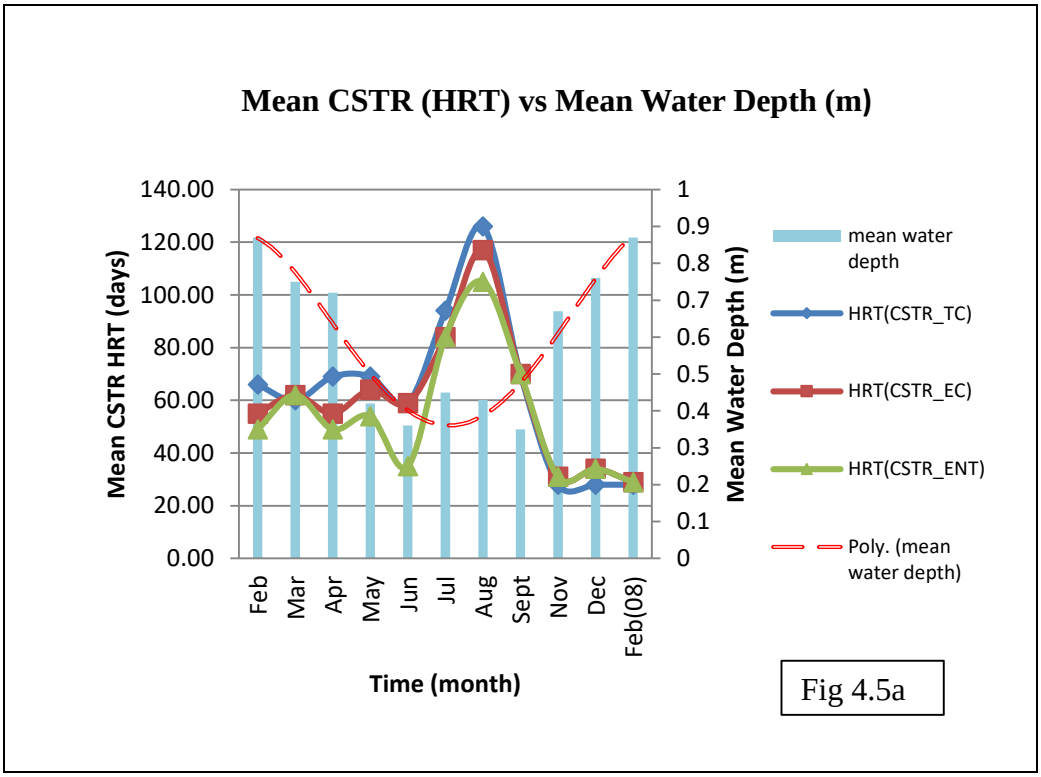


Fig 4.5a

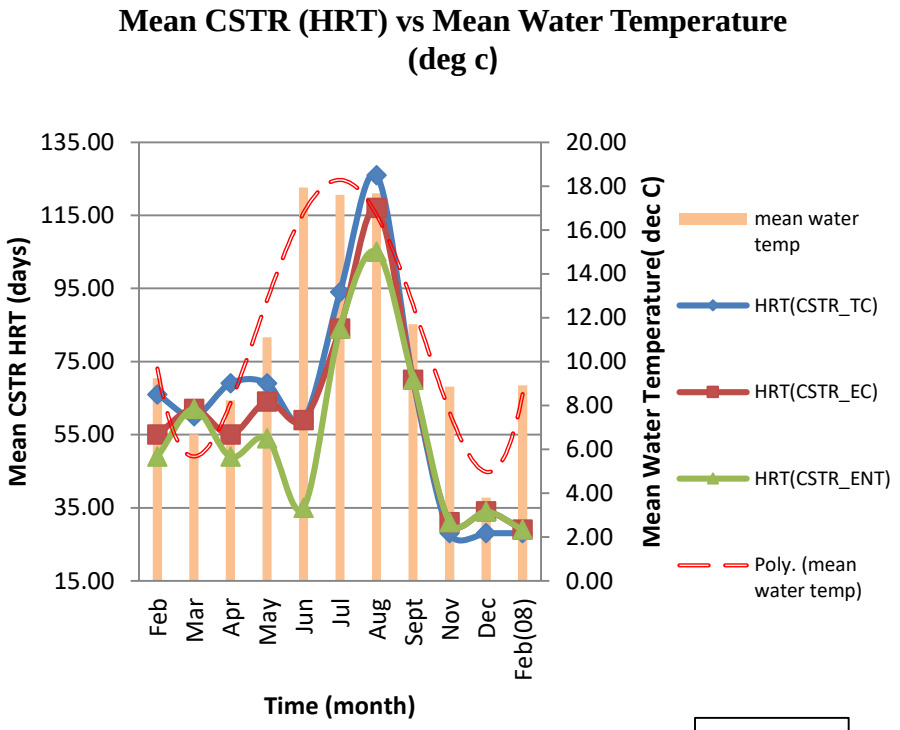


Fig 4.5b

Figure 4.5 a, b: The CSTR model versus the mean water depth and water temperature. A: The mean CSTR model versus the mean water depth.

B: The mean CSTR model versus the mean water temperature.

Table 4.5: The mean hydraulic retention times (HRTs)

	New DFM (Aver)	DFM (Aver)	CSTR (Aver)	Mean of means	Ideal HRT
Feb	42	37	57	54	100
Mar	38	40	61	54	100
Apr	47	56	58	54	100
May	63	**88**	62	54	100
Jun	58	80	51	54	100
Jul	**128**	76	87	54	100
Aug	53	69	**116**	54	100
Sept	48	46	70	54	100
Nov	30	28	30	54	100
Dec	29	28	32	54	100
Feb(08)	**28**	**27**	**29**	54	100
Mean	51.27	52.27	59.36		
Abs Mean	51.00	52.00	59.00	54.00	

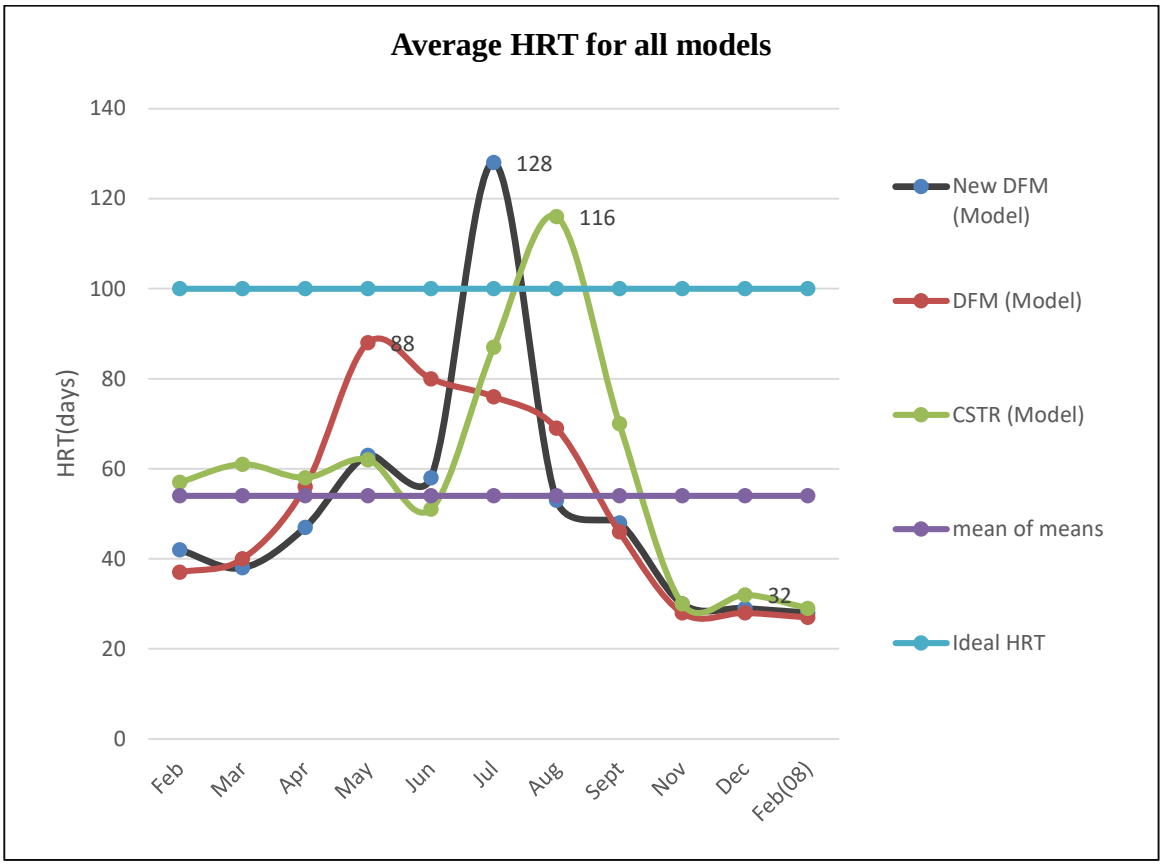


Figure 4.6: The average hydraulic retention times (HRTs) for all models.

Comparison of the three HRT models, the new dispersed flow model, the dispersed flow model (original) and the CSTR model. The ideal HRT is shown approx. 100 days. The mean of means i.e. the overall average HRT of the three models is shown, (54 days).

The action of climate of the wetland is of importance, particularly the monthly rainfall (mm). The rainfall data was taken from Oak Park Carlow weather station. Figure 4.7 shows the erractic/ pulsed nature of rainfall events throughout the sampling period. With rain water entering wetland system during the summer months (June and July) and again in November and December coinciding with increased populations of bacteria.

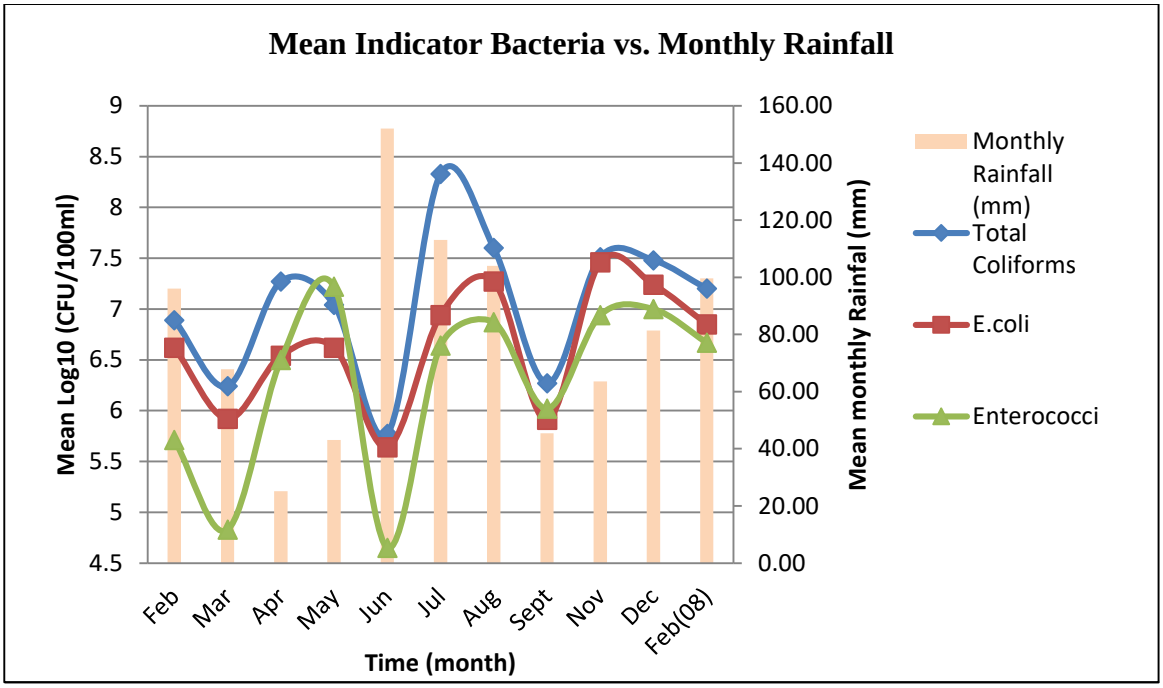


Figure 4.7: The mean indicator bacteria $\log_{10}$ (CFU/100ml) vs monthly rainfall (mm). The mean monthly rainfall (mm) recorded at Oak Park Carlow weather station in conjunction with the mean indicator bacteria.

4.10 Discussion

In this chapter, three models were investigated (1) the new Oakley dispersed flow model, where we introduced a new method to elucidate the first rate constant and the dispersion number (d), utilising the Peclet number. (2) The Oakley dispersed flow model and (3) the Mara and Maris CTSR models. Using the indicator bacteria as tracers in the wetland system as a means to evaluate the hydraulic retention time of the wetland. The new Oakley Model shows that the wetland system, during the dry summer months, exhibits a large hydraulic retention time. For the month of July, the modelled new Oakley model estimated the retention time exceeds the designed HRT of 102 days, with a mean peak modelled value of 128 days, see Table 4.5. The model shows good alignment with all three indicator bacterial values, with increased water temperature and inversely decreased water depth during the dry summer months, see Figures 4.3a and 4.3b. The original Oakley model shows a spread of the three indicator bacteria during the dry months, with no discernible pattern, with a mean peak modelled retention time of

88 days for the month of May 2007, see Table 4.5. The model does not align with peak water temperature and low water depth for the dry summer months, see Figures 4.4a and 4.4b. The Mara and Marais CSTR models display similarities to the new Oakley DF model, with the three indicator bacteria values showing good alignment with a mean peak modelled retention time of 116 days for the month of August 2007, see Figure 4.6. The model shows good alignment with all three indicator bacteria, with a slight skew in relation to increased water temperature and inversely decreased water depth during the dry summer months, see Figures 4.5a and 4.5b. Comparing the three models HRTs versus the hydraulic retention time calculations, the new Oakley DF model and the CTSR model are the closest matching values to these calculations, see Table 4.5. The indicator bacteria with the largest hydraulic retention time were the total coliforms, with all three models indicating this behaviour.

Reviewing other wetland characteristics such as dispersion, algal blooms and water depth, in conjunction with large HRTs and indicator bacterial count, it was found that two distinct extremes are occurring in the wetland system, these are short–circuiting and long-circuiting:

a. Short-circuiting - When large dispersion numbers are present, greater than 0.25, see Table 4.1; including large water depths, greater than 0.75m, see Table 4.3. No algal blooms present, leads to very small HRTs and large bacteria counts, see Table 4.3. The wetland system is very dynamic.

b. Long-circuiting -When low dispersion numbers are present, less than 0.14, see Table 4.1; including low water depths, less than 0.4m, see Table 4.3; with the presence of algal blooms, see Table 4.4, leads to large HRTs and large bacterial counts, see Figures 4.2a and 4.2b. The wetland system is stagnant.

In short-circuiting, a large volume of water, travels through a wetland system exits before the designed hydraulic retention time, thus reducing the performance of the constructed wetland. The wastewater has too little contact time with the

wetland ponds and is flowing through the wetland ponds with very short HRTs of approximately 30 days.

In long-circuiting, a small volume of water, travels through a wetland system exits after the designed hydraulic retention time, thus reducing the performance of the constructed wetland. The wastewater has too much contact time with the wetland ponds and is stagnant within individual ponds with very long HRTs of approximately 128 days.

In both cases the dichotomy effect of these actions leads to the potential of ineffective treatment of effluent entering the wetland system. The wetland has either a very small mean retention (approximately 30 days) time during the winter/wet months, with large water flows entering the system and reduced contact time to effectively treat the abattoir wastewaters or conversely, the mean hydraulic retention time in the summer/dry months is too large (approx. 128 days), with little or no water flows, the constructed wetlands ponds behave independently i.e. not connected to one another, when the water depth gets below 0.4m, with increased bacteria counts recorded.
Figures 4.2 a and 4.2 b, reveal the pulsed nature of the mean indicator bacteria counts with respect to mean water depth and mean water temperature. Similar trends were observed with both nitrate and phosphate concentrations – (chapter 10 section 3). The mean indicator bacteria display peaks and troughs throughout the sampling period, potentially due to different influent loadings emanating from the abattoir plant, or due to the dichotomy of the actions described above. Another aspect to consider with respect to the bacterial loadings within the wetland, is the effect of rainfall throughout the sampling period, see Figure 4.7 when large volumes of water entered the wetland system during the sampling period.

This in turn may increase the dynamic actions of the abattoir wastewater stored in the ponds, resulting in pulses of bacteria released into the wetland system. Hamaamin *et al*., (2014) describes the use of fuzzy interference modelling to describe *E.coli* removal under pulse-loaded conditions in a constructed wetland under two different seasons (winter and summer). The action of pulsing can be of benefit García *et al.,* (2014) describe a hybrid constructed wetland utilising

hydraulic pulses so as to improve oxygen renewal within the wetland thus improving its removal efficiency of organic contaminants. Ji and Jin, (2016) recognised the inflow pulse pattern dictated by the wet season when they modelled the water quality processes in a surface flow CW. Large HRT does not necessarily imply improved removal rates. García *et al.*, (2003) found that microbial inactivation reached saturation point in three days and increasing the HRT above three days did not significantly increase the removal rate. A similar was result was found by Toet *et al*., (2005). When analysing *E.coli*, they concluded that by increasing the HRT above four days shows no significance in *E.coli* removal. As mentioned previously the wetland swings from very large to very small HRTs. Indicator bacteria hydraulic retention times values as low as approximately 30 days, i.e. short-circuiting was evaluated. This reduces the wetlands performance and impacts the water quality of the wetland (Kadlec and Knight, 1996; Dierberg *et al*., 2005). Conversely, the concept of long-circuiting was realised through the modelling methods within the CW, with large HRT, small water flows and volumes and low dispersion numbers (stagnant), with the potential that the wetland ponds are disconnected from one another. The water level is below the inlet/outlet flow pipes, resulting in reduced flows between interconnected ponds within the CW. In both short-circuiting and long-circuiting this can lead to potentially ineffective treatment of the abattoir wastewaters.

Reviewing the three models, (1) the new Oakley dispersed flow model, (2) the Oakley dispersed flow model and (3) the Mara and Marais continuous stirred reactor (CSTR) models, it was found that all three models display similar HRTs in the wet winter months, indicating the presence of short-circuiting within the wetland, with respect to the three indicator bacteria as tracers, with values between 27 -29 days, see Table 4.5. The results are similar to the calculated HRT of 29 days in section 4.3. The main differences occur in the HRTs, when the modelling calculates the tracer bacterial loadings from April to September. Both the new Oakley DF model and the CSTR models display similar HRTs with values greater than the 102 days HRT as per design specfications, with mean HRTs of 128 and 116 days respectively, see Table 4.5. These results compared favourably with the calculated HRT of 125 days.

The Oakley DF model reveals a mean value of 88 days HRT, which is under the 102 days' design HRT, therefore using the Oakley DF model, the system is

performing within specifications. But Buchauer (2007), concludes that the CSTR model was better than the Oakley DF model, due to fact that the DF model was too optimistic in comparison to the CSTR model which matches measured reality regarding faecal coliforms loadings. The new Oakley DF model, in conjunction with the CSTR model reveal a more realistic behaviour of indicator bacteria within the wetland. The wetland moves between CSTR and new Oakley dispersed flow model regimes, with both short-circuiting and long-circuiting dynamics observed.

4.11 Conclusions

The creation of a new Oakley dispersed flow model was to investigate if the model could be improved, by changing two aspects of the formulae; the dispersion number *d* and the rate constant k. The model was tested against the traditional/original Oakley model and the Mara and Marais CSTR models using indicator bacteria as tracers within the constructed wetland. The hydraulic retention times were calculated for all three models with the new Oakley and CSTR models displaying similar trends, with mean HRTs of 128 and 116 days respectively recorded for the dry summer months. This behaviour is the opposite of short-circuiting, a dynamic of long-circuiting was observed.

Where long-circuiting is typified by:

- Large HRTs (greater than the initial design)
- Low water depths (below the inlet/outlet pond piping)
- Low flows (poor inter-connective flows between ponds leading to pond isolation)
- Large dispersion numbers (indicative of stagnation, no mixing)
- The presence of algal blooms within the wetland, especially towards the end point/ outlet ponds of the CW.

The above actions can potentially lead to large bacterial populations in the wetland, impacting the overall efficiency and performance of the CW. The traditional/ original Oakley model was optimistic in its HRT of the wetland

system, with a mean value of 88 days, an adequate HRT, revealing no potential issues with the CW. During the wet winter months all three models displayed similar behaviours, with an average HRT value of 30 days, indicating short-circuiting present in the constructed wetland.

Where short-circuiting is typified by:

- Small HRTs (much smaller than the initial design)
- Large water depths (well above the inlet/outlet pond piping)
- Extreme flows (strong inter-connective flows between ponds)
- Small dispersion numbers (extreme turbulence).
- No plants present (therefore no resistance within the CW).

The above actions can lead to potentially large bacterial populations in the wetland, impacting the overall efficiency and performance of the CW.

Chapter 5. Bio-aerosol analysis of the constructed wetland

The odour impact of constructed wetland treating abattoir wastewater and the levels of airborne microorganisms generated: A preliminary study.

5.1 Summary

Airborne samples were retrieved every month over a 5-month period from a constructed wetland treating abattoir wastewater in order to investigate the bio-aerosol distributions and concentration within the wetland system. Air samples were collected by using a single stage impactor and four different agar types were used during the process to elucidate the airborne bacterial concentrations. In this experiment, the Dissolved Air Floatation (DAF) plant, involving mechanical agitation of the wastewater, generated the largest amount of bio-aerosols (between 1266 – 2286 CFU/m^3) compared to the constructed wetland ponds (between 160-636 CFU/m^3) for Pond 6 and (between 42- 78 CFU/m^3) for Pond 12. It was found that the most significant variables for bioaerosol determination were the Membrane Lactose Glucuronide agar (MLGA) airborne counts along with the actual recorded air temperature at the time of sampling. The 'FIDOL' factor test, which stands for frequency, intensity, duration, relative offensiveness and location, was utilised along with an odour range to determine the equivalent trends by olfactory odours.

The use of MLGA as an agar medium to elucidate bioaerosol (CFU/m^3) as an olfactory acuity test (odour response) was developed.

5.2 Introduction

The primary pathways for transmission of bacteria to humans are: (1) by direct contact with the contaminant source (2) via ingestion and (3) inhalation. Bioaerosols are for the majority, air-borne microorganisms with an aerodynamic diameter ranging from less than 1µm up to 100µm (Cox and Wathes, 1995). Recently, it was found that infectious particles sized less than 10 µm have more serious health implications as they are capable of penetrating into the lower respiratory tract to establish infections (Galton *et al*, 2011). The evaluation of airborne microorganisms on workers exposed to bioaersols while maintaining wastewater treatment plants (WWTPs) with Sánchez-Monedero *et a*l., (2008) recommened the use of air diffusers to minimise the potential biological hazard to wastewater treatment plant workers and Liu *et al*, (2011) state that bioaerosols can have a negative influence on urban air quality as well as human health. Wastewaters contain a large number and diversity of pathogens such as viruses, bacteria, fungi, protozoans and helminths which originate from animal excreta in household, commercial and hospital sewage, (Geradi and Zimmerman, 2005; Fracchia *et al*., 2006). These microorganisms can easily become airborne during normal operations at WWTP, especially during aeration and mechanical agitation of raw wastewater, which is one of the main vectors for bioaerosols, (Pascual *et al*., 2003).

The potential hazards caused by air borne bacteria depends on the specific bacteria released, their corresponding pathogenicity plus environmental factors which govern the survival of these airborne microorganisms (Mohr, 2002). These include the climate conditions, especially wind speed and direction, which dictate bioaerosol dispersion from different emission points and the pathway of microbe into the body.

With reference to WWTP and the personnel working on-site, a significant risk is presented by the microbial allergens and endotoxins (Laitinen *et al*., 1994; Melbostad *et al*., 1994; Rylander, 1999; Thorn *et al*., 2002). A particular form of the illness has been characterised by general malaise, weakness, acute rhinitis and fever that has been called “sewage worker’s syndrome” has been reported by

Rylander *et al*., (1976), Fannin *et al*., (1985), Clark (1987), Douwes *et al*., 2001 and Vantarakis *et al*., 2016.

The use of forced aeration in activated sludge treatments such as splashing and bubble bursting, produce large concentrations of bioaerosols (Fracchia *et al*., 2006). The type of aeration system influences the production of air borne bacteria where (Brandi *et al*., 2000) found remarkable dispersion of airborne bacteria in the mechanical agitation of sludge in the presence of an oxygen supply. Korzeniewska, (2011) found that bioaersols generated during wastewater treatment may pose a hazard to workers or to habitants of their surroundings.

To date only one reference to bioaerosol analysis in a constructed wetland was found (Miao *et al*., 2010), with high bioaersol concentrations of bacteria and fungi present during the summer months June and August.

The main objective of this study was to evaluate the effect of the constructed wetland system used in the biological treatment of the abattoir wastewater on the airborne micro-organisms levels to which people may be exposed, in order to determine the potential health risks, if any, to exposed workers. The main bioaersosol sources and concentrations were evaluated at different sample points on the constructed wetland.

5.3 Surface Air System Super-90 (SAS) operation

See chapter 2 section 2.16

5.4 Air sampling and microbiological analysis

Airborne micro-organisms concentration was monitored with a single stage impactor sampler (SAS Super-90). Sample sites included, DAF plant, Pond 6 and Pond 12 i.e. the input, midpoint and output, respectively. The sample period: was from Feb, March, April, June, July, on one day. Three sample replicates from each sample sites were taken across four agar types (Plate Count Agar, Trypticase Soy Agar, MacConkey agar, Membrane Lactose Glucuronide Agar), therefore sample size n=180. These agar types were used to enumerate the bioaerosol counts because they were readily available in the lab. Preliminary testing was

carried out using these agar types with the SAS unit in a local wastewater treatment plant (WWTP). The Plate Count and Trypticase Soy agars are general bioaerosol enumeration counts. MacConkey and MLGA were more specific for *E.coli*, *Enterobacter* and *Klebsiella* bacteria.

Conformatory testing was completed using the Analytical Profile Index (API®) 20E Test Kit, which validated the following bacterial colonies from the MLGA plates, *E.coli*, *Enterobacter* and *Klebsiella* bacteria, see chapter 2 – section 2.20.

5.5 Surface Air System Super-90 (SAS) method of sampling

See chapter 2, section 2.16

5.6 Olfactory Acuity Test and 'FIDOL' Factors

The FIDOL Factors are used by the Local Authorities in the UK (Defra, 2010), to highlight principles and factors that may be important when assessing a specific odour source that may constitute a statutory nuisance. The FIDOL factors are defined as Frequency, Intensity (i.e. concentration), Duration, relative Offensiveness and Location along with test characteristics, see Table 5.1. Another element was added to the Olfactory Acuity Test which included the odour ranges, whereby a value between 1 and 5 was recorded, where 1 = extremely bad and 5 = very good with respect to Olfactory Acuity Test and the distance (range) to which the odour can be detected. Tables 5.1 and 5.2 displays both the Olfactory Acuity Test plus the odour range value, respectively.

Table 5.1: The Olfactory Acuity Test "FIDOL Factors" by sample point:

		DAF	Pond 6	Pond12
Frequency	*(How often an individual is exposed to odour)*	Often (DAF located close to the abattoir facility)	rarely	very rarely
Intensity	*(Perceived strength of the odour)*	very strong	medium to weak	weak
Duration	*(length of odour event)*	30 mins	30 mins	30 mins
Offensiveness	*(Type of Odour/s)*	Pungent rotten eggs with ammonia, combined with sickly sweet odours	Strong - mild sickly sweet odour	no perceived odours
Location	*(Characteristics of the vicinity of where the odour occurs plus the sensitivity of the complainant)*	DAF located close to abattoir plant and the odours are very specific in comparison to the abattoir facility	Pond 6 located away from the abattoir facility but within 300 metres of DAF plant	Pond 12 is the furthest pond away from abattoir and DAF plant (approx 600 metres).

Table 5.2: The odour ranges values by distance

Odour range		
Value = 1	extremely bad - very bad	Odours detected from 120 metres to 50 metres from source.
Value = 2	very bad- bad	Odours detected from 50 metres to 10 metres from source.
Value = 3	bad- medium	Odours detected from 10 metres to 1 metre from source.
	medium - good	

Value = 4		Odours detected from1 metre from source, too little perceived odour.
Value = 5	good- very good	Little or no perceived odour

5.7 SAS sampling results

The MacConkey Agar revealed three distinct hues, red, white and golden colonies. The red colonies indicate the fermentation of the sugar lactose (Lac +), potentially indicating *E.coli*, *Enterobacter* and *Klebsiella* bacteria; the white and golden colonies are non-lactose fermenting bacteria (Lac-), indicative of *Salmonella*, *Shigella* and *Pseudomonas aeruginosa* bacteria. The majority of the DAF bioaerosol MacConkey agar counts were red colonies but a mixture of white and golden colonies hues was visible. Sampling at the other sites, Pond 6 bioaerosol results the MacConkey agar red colony counts dominated the agar plates the remaining white colonies present to a lesser extent than the DAF counts, no golden hues were detected at Pond 6. In Pond 12 red colony counts dominated with the one white colony detected, see Table 5.3. The MLGA revealed three distinct hues, yellow, green and pink colonies. The green colonies were indicative of *E.coli*, with yellow hues potentially indicating *Klebsiella* and *Enterobacter* bacteria and pink hues potentially indicative of Pseudomonas aeruginosa and *Salmonella*. Approximately half of the DAF bioaerosol MLGA counts were green, with yellow colonies dominating after the green and finally pink colonies present in the lowest counts. At the other sampling sites, a decrease in bacterial populations was observed, Pond 6 bioaerosol MLGA green colony counts greatly dominated the counts with yellow/pink colonies making up a smaller proportion. Pond 12, one green green colony was present, see Table 5.4.

Table 5.3: SAS 90- RODAC mean value of three replicates representing the colony counts per sampling site per month.

DAF					**DAF**				
		Mac					MLGA		
Month	Total	Red	White - Golden		Month	Total	Green	Yellow	Pink
Feb	262	215	47		Feb	15	8	6	1
Mar	247	220	27		Mar	9	4	3	2
Apr	316	255	61		Apr	31	18	11	2
Jun	219	198	21		Jun	22	13	6	3
Jul	369	314	55		Jul	63	38	20	5
Pond 6					**Pond 6**				
		Mac					MLGA		
Month	Total	Red	White		Month	Total	Green	Yellow	Pink
Feb	17	16	1		Feb	3	2	1	1
Mar	13	12	1		Mar	ND	ND	ND	ND
Apr	17	17	0		Apr	ND	ND	ND	ND
Jun	30	29	1		Jun	11	10	1	ND
Jul	16	15	1		Jul	9	7	1	1
Pond 12					**Pond 12**				
		Mac					MLGA		
Month	Total	Red	White		Month	Total	Green	Yellow	Pink
Feb	ND	ND	ND		Feb	ND	ND	ND	ND
Mar	ND	ND	ND		Mar	ND	ND	ND	ND
Apr	2	2	0		Apr	ND	ND	ND	ND
Jun	2	2	0		Jun	ND	ND	ND	ND
Jul	3	2	1		Jul	1	1	ND	ND

* ND= not detected

In Table 5.3, the bioaerosol results were calculated as the geometric mean of three replicates and were expressed as colony forming units per cubic meter of air (CFU/m^3) after positive-hole correction. An Olfactory Acuity Test and odour range values were evaluated during the sampling regime. These values are included in conjunction with the SAS Super-90 bioaerosol sampling, see Table 5.4.

Table 5.4: Bioaerosol SAS Super 90 data plus odour ranges by data and sample point

		PCA	TSA	Mac	MLGA	Total	
Sample Date	Sample Point	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	Olfactory Acuity Test plus Odour ranges.
	DAF	597	781	262	15	1641	very bad (1-2)
Feb	Pond 6	109	118	17	3	244	good (4-5)
	Pond 12	24	54	ND	ND	78	very good (5)
	Total	731	953	279	18	1963	
	Std	351.00	457.09	151.58	8.83	958.91	
		PCA	TSA	Mac	MGLA	Total	
		(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	
	DAF	396	614	247	9	1266	bad (2-3)
Mar	Pond 6	54	93	13	ND	160	very good (5)
	Pond 12	9	34	ND	ND	42	very good (5)
	Total	458	740	260	9	1468	
	Std	230.47	358.73	142.80	5.20	737.14	
		PCA	TSA	Mac	MLGA	Total	
		(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	
	DAF	749	907	316	31	2003	extremely bad (1)
Apr	Pond 6	37	162	17	ND	216	very good (5)
	Pond 12	8	35	2	ND	45	very good (5)
	Total	794	1103	336	31	2264	
	Std	432.99	531.98	183.02	17.90	1163.43	
		PCA	TSA	Mac	MLGA	Total	
		(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	
	DAF	376	718	219	22	1335	very bad (1-2)
Jun	Pond 6	187	408	30	11	636	bad (2-3)
	Pond 12	22	37	2	ND	62	very good (5)
	Total	586	1163	251	33	2033	
	Std	243.18	477.27	127.63	14.20	854.51	
		PCA	TSA	Mac	MLGA	Total	
		(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	(CFU/m^3)	
	DAF	798	1056	369	63	2286	extremely bad (1)
Jul	Pond 6	68	178	16	9	271	bad (2-3)
	Pond 12	10	41	3	1	55	very good (5)
	Total	876	1275	389	73	2612	
	Std	462.43	618.74	213.55	36.75	1329.44	

5.8 Meteorological conditions

The actual recorded air temperature taken on-site during the air sampling; Air temperatures [mean, max and min], humidity, mean wind speed, max gust, solar radiation and total rainfall, were values recorded at Oak park, Carlow, Ireland Meteorological station, see Table 5.5.

Table 5.5: Climate data from Oak park, Carlow, Ireland

Month	Actual recorded Temp (^{0}C)	Max Temp (^{0}C)	Min Temp (^{0}C)	Mean Temp (^{0}C)	Mean Wind speed (m/s)	Max Gust (m/s)	Solar rad (MJ/sec)	Humidity (%Rh)	Total rainfall (mm)
Feb	5.8	10.1	3.3	6.7	5.1	14.5	432.3	89.4	123.9
Mar	2.2	14.3	-0.7	6.8	1.4	3.9	1205.2	85.7	80.5
Apr	8.8	16.4	1.9	9.1	1.8	6	1788	68.8	30.5
June	12.1	28.1	11.4	13.6	4	9.4	1528.7	66.4	167.6
July	15.2	19	10.2	14.6	2.7	9.6	2153.6	77.2	121.5

5.9 Statistical analysis

The experimental data (bacterial population counts and the climate data) were subjected to analysis of variance procedure (SPSS 22.0). The analysis of variance (ANOVA) was performed. Further analysis was performed using principal component analysis (PCA) with Promax Kappa = 4 and Spearman rank correlation for $p < 0.05$ (SPSS 22.0). The PCA was used determine which variables are the dominant components for the dataset analysed. The Promax option provides an extra matrix, called the structure matrix, which provides extra structural information on the dataset. The Spearman's Rank correlation was used to evaluate correlations for 1-tailed significance for $p < 0.05$.

5.10 Discussion of Bioaersol data

Figure. 5.1 shows the concentration of airborne bacteria at different sample points within the constructed wetland. Three replicate control baselines using TSA agar was carried outside the wetland system with values between 30-70 CFU/m^3. These counts represented background levels for airborne bacteria unaffected by the activities of the constructed wetland.

Mechanical activities i.e. DAF plant agitation tank represent potential bio- aerosol source (Laitinen *et al*., 1994; Pascual *et al*., 2003). In this study the DAF plant generated the largest concentrations of bioaerosols. The total bio-aerosol values for DAF plant ranged from 1335 – 2286 CFU/m^3 over the sample period, with lower values recorded within the wetland at Pond 6 the total bioaerosol values ranged from 160 – 636 CFU/m^3 and Pond 12 the total bioaerosol values ranged from 42- 78 CFU/m^3. One possible source of bioaerosol within the constructed wetland sample points was the diffusion of particulates and bacteria from the DAF plant down wind through the wetland ponds. The background levels were between 30-70 CFU/m^3 and Pond 12 values are not that far removed from these background recordings, leading to the possibility that the airborne bacteria emanating from the DAF plant have dissipated with enough dispersive energy that by the time readings taken at pond 12, which is approximately 600 metres from the DAF plant. Pond 12 bioaerosol values are comparable to the background bioaerosol levels. Principal component analysis (PCA) with Promax = 4 for orthogonal view was used, the data sorted in ascending order with component values greater than 0.4 selected. The component matrix shows three variables, the MLGA count, mean temperature and actual recorded temperature highlighted in red. These three variables have eigenvalues >3.0 and explain 90.79% cumulative variance of the dataset, i.e. they are the most dominant variables within the analysed dataset. The structure matrix further elucidates the variable MLGA having a strong relationship with the climate variables in comparison to the other analysed agars, see Fig. 5.2. Spearman's Rank correlation was also employed to evaluate correlation behaviour within the dataset, see Table 5.7. The analysis reveals that MLGA and air temperatures (actual recorded temperature and mean temperature) have the strongest correlations between bioaerosol counts and climate variables within the dataset. The climate data see Table 5.5, records nine

climate variables [actual recorded (air) temperature, mean (air) temperature, min (air) temperature, max (air) temperature, mean wind-speed, max gust, solar radiation, humidity and total rainfall]. The actual recorded temperature was recorded on-site during the sampling, the remainder climate variables were taken from Oak Park, weather station, Carlow, Ireland, see Table 5.5.

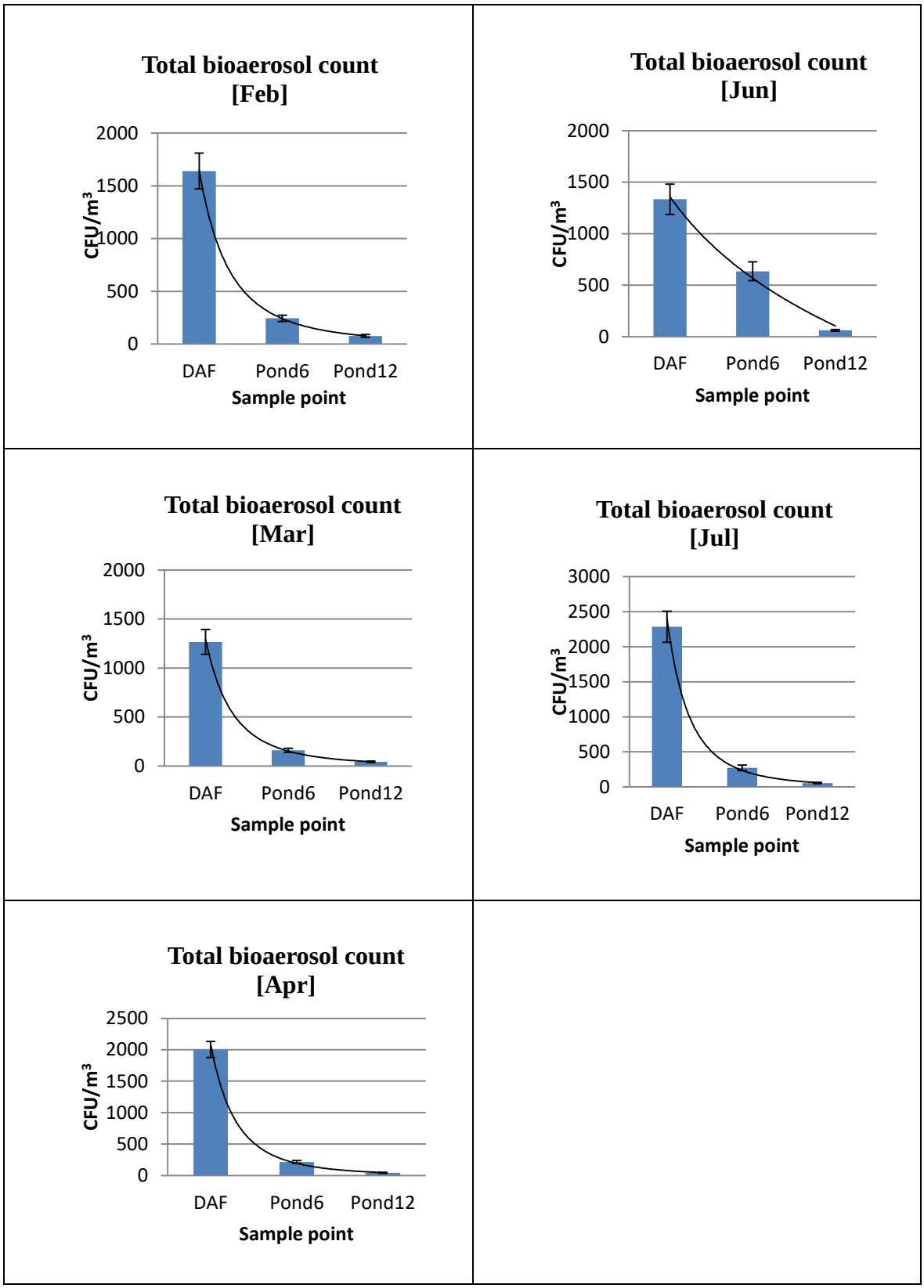


Figure 5.1: The total bio-aerosol bacteria counts from the DAF, pond 6 and pond 12. Vertical bars represent the standard error.

Table 5.6: Principal Component Analysis of the data from Tables 5.4 and 5.5.

Component Matrix[a]				**Structure Matrix**			
	Component				Component		
	1	2	3		1	2	3
MLGA	**0.833**	**0.432**		mean temp	0.978		
mean temp	**0.8**	**-0.56**		actual recorded Temp	0.914		
actual recorded Temp	**0.792**	**-0.49**		min temp	0.889		0.545
min temp	0.757	-0.58		max temp	0.884		
TSA	0.726	0.661		humidity	-0.8		
max temp	0.64	-0.6		solar rad	0.775		-0.47
humidity	-0.57	0.447	0.45	Total		0.99	
Mac	0.635	0.756		PCA		0.99	
PCA	0.657	0.742		TSA		0.98	
Total	0.7	0.706		Mac		0.98	
mean wind speed			0.955	MLGA	**0.471**	**0.9**	
max gust			0.928	mean wind speed			0.97
total rainfall		-0.43	0.688	max gust			0.932
solar rad	0.588		-0.69	total rainfall	0.429		0.826
Extraction Method: Principal Component Analysis.				Extraction Method: Principal Component Analysis.			
a. 3 components extracted.				Rotation Method: Promax with Kaiser Normalisation.			

In Table 5.6, shows the putput of the the Principal Component Analysis (PCA) of the data from Tables 5.4 and 5.5, with output of the PCA matrix identifying three dominant components within the bioaerosol dataset, which are MLGA, mean temperature and actual recorded temperature. The structure matrix further enforces the importance of MLGA when it reveals its shared component values between component 1 and 2, especially component 1, which captures the majority of the climate data.

Table 5.7: Spearman's Rank Correlation for p < 0.05 (1-tailed significance)

	PCA	TSA	Mac	MLGA	Actual recorded Temp	Max Temp	Min Temp	Mean Temp	mean wind speed	max gust	Solar rad	Humidity	Total rainfall
PCA	1.00												
TSA	0.97	1.00											
Mac	0.98	0.96	1.00										
MGLA	0.84	0.85	0.83	1.00									
Total	0.99	0.99	0.98	0.86									
Actual recorded temp	0.16	0.21	0.11	***0.49***	1.00								
max temp	-0.03	0.12	-0.02	0.20	0.64	1.00							
min temp	0.08	0.18	0.04	0.39	0.89	0.77	1.00						
mean temp	0.09	0.18	0.08	***0.45***	0.95	0.80	0.92	1.00					
mean wind speed	0.04	0.05	-0.05	-0.01	0.18	0.03	0.43	0.06	1.00				
max gust	0.10	0.07	0.00	0.09	0.27	-0.15	0.41	0.09	0.94	1.00			
solar rad	0.09	0.14	0.12	0.40	0.71	0.56	0.45	0.76	-0.57	-0.47	1.00		
humidity	-0.04	-0.14	-0.03	-0.20	-0.62	-0.82	-0.53	-0.66	0.18	0.28	-0.69	1.00	
total rainfall	-0.03	0.06	-0.05	0.12	0.39	0.51	0.75	0.50	0.71	0.58	-0.16	-0.01	1.00

Table 5.7 shows the Spearman rank correlation 1-tailed significance for the bioaeraosol dataset.

The most significant correlations are MLGA and the actual recorded (air) temperature and the mean (air) temperature, in comparison to the other bioaerosol samples; Plate Count agar, Trypticase Soy Agar and MacConkey agar.

5.11 The use of MLGA agar as an indicator of air quality

The Principal Component Analysis, see Table 5.6 and Spearman's Rank correlation, see Table 5.7 statistical tests reveal that MLGA agar counts and the actual recorded air temperature are (1) the most dominant variables and (2) the most correlated variables within the dataset. Figure 5.2 shows the MLGA bioaerosol counts and the actual recorded air temperature plotted. The month of February 2007 and June 2007 display inconsistencies between the temperature and MLGA counts, the most likely reason for this was the increased mean wind speed, during the sampling, which were 5.1 (m/s) and 4.0 (m/s), see Table 5.6. MLGA bioaerosol counts are present at pond 6, during February 2007, June and July 2007. During the summer months the air temperature increases as does the MLGA counts. The exception was the month of February 2007, where pond 6 also recorded MLGA counts, one probable reason was the warm weather that was observed during the sampling. MLGA threshold levels as an Olfactory Acuity Test (odour) are as follows:

MLGA bioaerosol CFU/m^3 [(31-23) = extremely bad; (22-15) = very bad; (14-9) = bad; (8-3) = good; (2-0) = very good].

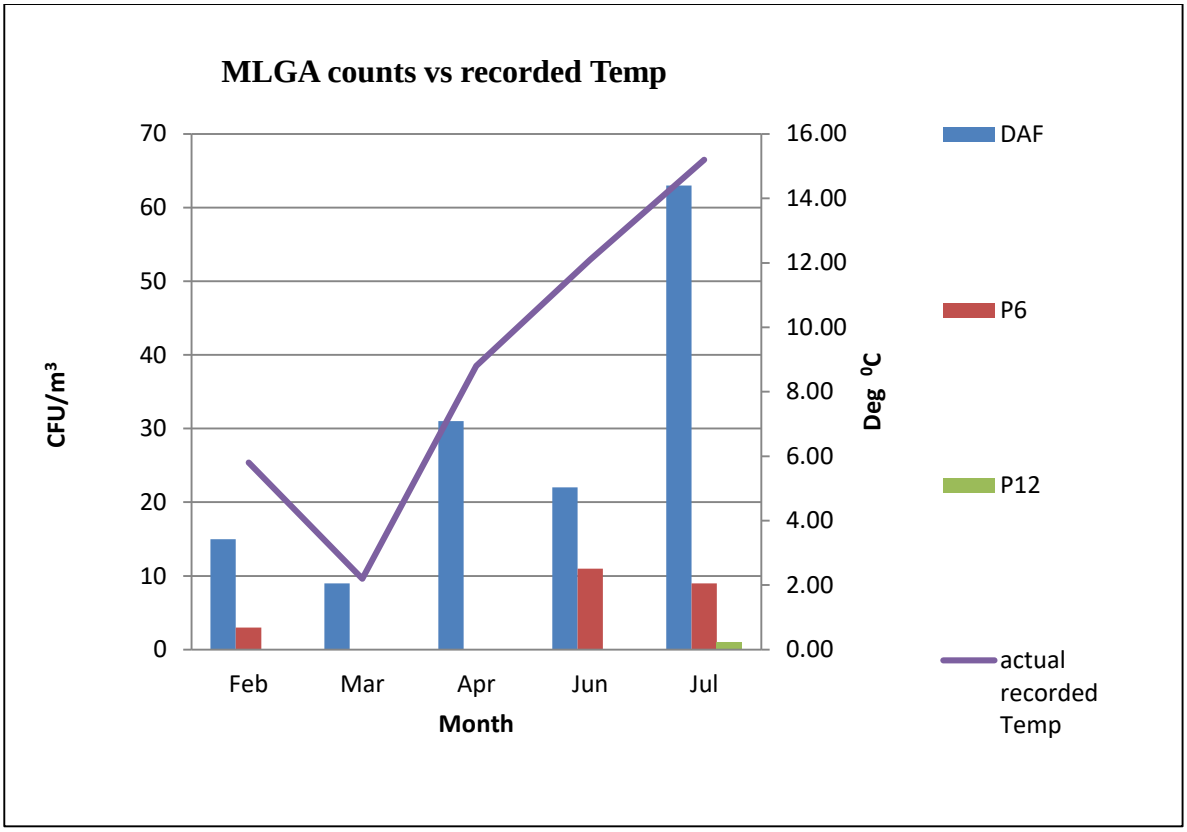


Figure 5.2: The MLGA agar counts (CFU/m^3) versus the actual recorded air temperature (^{0}C), for the sampling period February 2007 – July 2007.

5.12 The Olfactory Acuity Test with Odour range

Figure. 5.3, shows the Olfactory Acuity Test with odour range recorded as per Tables 5.2 and 5.3 and tabulated in Figure 5.3. If the odour range value was between two values i.e. (1-2) or (2-3) the average of these values was used. The graph displays the olfactory response during the sampling period, see Figure. 5.3, where low values of 1-2 (extremely bad dours) were observed in the vicinity of the DAF plant indicating very poor air quality. As the samping progresses into the summer months of June and July the DAF plant maintains its extremely bad odours, with Pond 6 recording values between 2 and 3, indicating bad to medium bad odours in the vicinity of Pond 6.

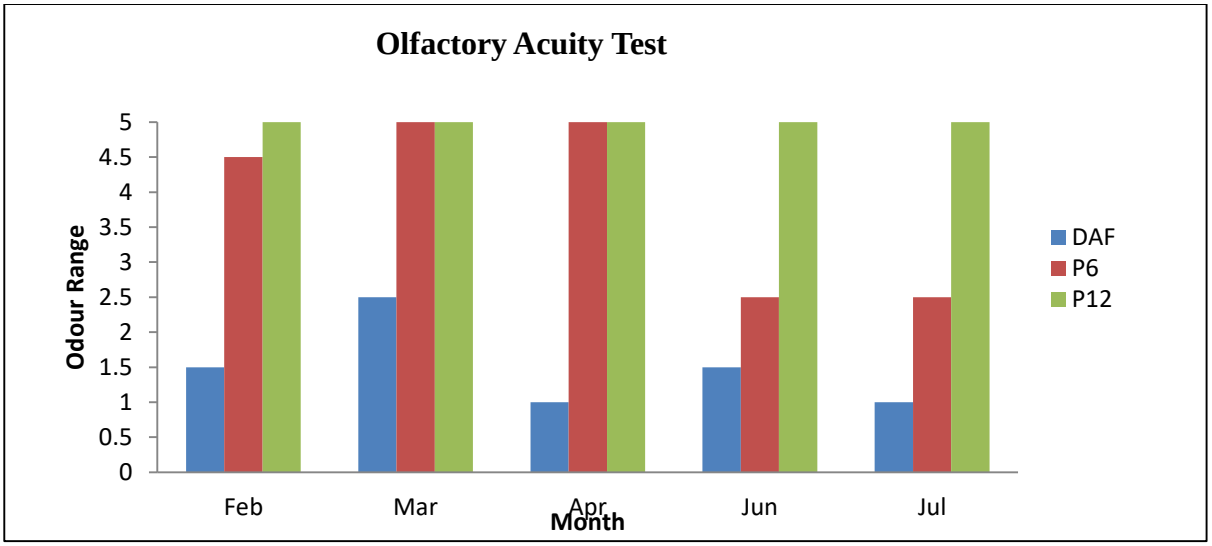


Figure 5.3: The results of the Olfactory Acuity Test. For the DAF, pond 6 and pond 12 by odour range and month.

5.13 Discussion

Meteorological conditions have an impact on bioaerosol levels. Wind speed and direction are important factors governing bioaerosol dispersion also the importance of solar radiation and ambient air temperature influence the levels of airborne bacteria at wastewater treatment plants (Karra and Katsivela 2007).

The DAF plant was the likely source of bio-aerosols during the sampling period, with the combined air diffusion process and the mechanical agitation tank are the likely vectors for the release of bioaerosols. The agitation tank was open to the environment and employing covering around and over the tank would potentially reduce the release of air borne bacteria. Fedorak and Fernado (2005) reported that covering grit and primary settling tanks reduces the release of air borne bacteria. The results obtained show that the dissolved aeration floatation (DAF) plant used in the processing of the abattoir waste greatly affects the amount of bioaersols generated and as a result the levels that people in the vicinity may be exposed to.

5.14 Conclusions

The DAF plant was the main source of the airborne bacteria, in comparison to the constructed wetland. The most probable cause being the air diffusion and mechanical rotors agitating the abattoir wastewaters, thus releasing the bacteria into the surrounding air, resulting in larger bacteria concentrations. The DAF plant, involving mechanical agitation of the wastewater, generated the largest amount of bioaerosols (between 1266 – 2286 CFU/m^3) compared to the constructed wetland ponds (between 160-636 CFU/m^3) for Pond 6 and (between 42- 78 CFU/m^3) for Pond 12. However, meteorological conditions have a bearing on the concentrations and dispersion of bacteria. The constructed wetland has no mechanical components and therefore has a reduced bioaerosol concentration. But a caveat must be included with respect to the build-up of sludge in the wetland ponds, in particular, the front end ponds which must be removed periodically. The removal of sludge from these ponds could therefore constitute a bioaerosol release. In this study no sampling occurred while the sludge removal process occurred. Future work is required to take more samples over a longer period of time, with specific targeting of aerobic, non-aerobic, fungi and endotoxin analysis to reveal a more complete picture of the airborne microflora. The efficacy of the 'FIDOL' factor test was utilised along with an odour range to determine the equivalent trends by olfactory odours. The use of this system establishes a reasonable matrix in determining olfactory issues within treatment systems. The MLGA bioaerosol levels and actual recorded (air) temperature were the dominant variables in determining levels where issues occur. Finally, the use of MLGA as a system to elucidate bioaerosol (CFU/m^3) as an olfactory acuity test (odour response) was developed where MLGA (CFU/m^3) = [(31-23) = extremely bad; (22-15) = very bad; (14-9) = bad; (8-3) = good; (2-0) = very good]. The use of MLGA in combination with the SAS air sampling sytem, has the potential to indicate areas within the constructed wetland treatment processes where high concentrations of potential harmful bacteria are present. Further work would need to be carried out to evaluate the efficacy of MGLA agar in the use of a potential bacteriological risk identification tool.

Chapter 6. A Bayesian network analysis of the wetland's soil and sludge.

The effects of soil and sludge bacterial loadings in a constructed wetland treating abattoir wastewater.

6.1 Summary

An increase in microbial indicator organisms (total coliforms, *E.coli* and enterococci) within the free water constructed wetland system (FWS) treating abattoir wastewater was a cause for concern. It was noticed during the summer months that soil and sludge bacterial concentration increased with noticeable bacterial values recorded throughout the wetland system's sampling points. In particular, pond 12 showed an increased indicator bacterial loadings in total coliforms and Enterococci bacteria, where the Enterococci bacteria values were at the largest within pond 12's soil sediment. This was the only time during the sampling process that Enterococci bacteria were more numerous or comparable to their coliform counterparts.

A combination of primarily climatic and soil conditions with secondary affects such as plant loss may have attributed to these increases. Risk assessment techniques were applied to make practical use of the dataset retrieved, to further our understanding of the dynamics of these complex factors that operate within the FWS. The risk assessment methodology was based on the application of a Bayesian network (BN) used to model the environment. The wetland information retrieved was divided into three sub-parts; (1) the Wetland design parameters (wetland area, plants/m^2, soil porosity and the plants x wetland area) (2) inorganic ions (chloride, fluoride, nitrate, nitrite, phosphate and sulphate) (3) indicator bacteria (total coliforms, *E.coli* and enterococci), thirteen variables in total were used in the BN model. During the model development, sensitivity analysis (SA) was employed as an important model validation and assessment tool.

The outcome of the modelling and analysis revealed that the wetland design parameters (plant and soil porosity variables) dominate soil bacteria concentrations in comparison to the inorganic soluble ions.

6.2 Introduction

Constructed wetlands perform a complex integrated wastewater treatment process. They are integrated and complex due to the fact that water, air, soil, plants, microorganisms and climate all interact to improve water quality. Complexity is further increased by chemical and physical factors included within the process.

Bayesian networks (BN) or Bayesian belief networks (BBN) are powerful modelling tools that deal with uncertainty within systems (Pearl, 1986; Jensen, 1996; Jensen and Nielsen, 2007). In the last twenty-five years, BNs have been utilised across various disciplines from medical analysis (Gurwicz and Lerner 2005), to bankruptcy prediction (Sun and Shenoy 2007), medical prescription fraud (Aral *et al*., 2012), ecological studies (Pollino *et al*., 2007; Spence and Jordan 2013), climate change (Richards *et al*., 2013) and wetland conservation (MacPherson *et al*., 2018). BNs are graphical models that illustrate a network of interactions between variables to account for uncertainty within the dataset under analysis. Other advanced modelling methodologies have been adapted with respect to constructed wetlands, Lee *et al*., (2006a) and Lee *at al*., (2006b) used case based reasoning analysis to understand the performance of constructed wetlands treating storm water run-off. Scholz *et al*., (2008) and Scholz and Lee (2006) have used self-organising maps to assess the performance of constructed wetlands, treating nutrient and heavy metal removal respectively. Overall *et al*., (2009) reported the application of BNs to review storm wastewater treatment using constructed wetland systems, using a risk based modelling. Therefore, using Overall *et al*., (2009) as a template and applying BN's to review abattoir wastewater treatment using a FWS constructed wetland. What factors attribute to increased bacterial loadings in the latter part of the wetland system (pond 9 to pond 12) and the risks associated with these factors such as the design of the wetland, the plant loadings, the nutrients within the soil, the soil porosity or a complex combination of all of these wetland factors. The use of BN analysis combined with sensitivity analysis may help remove some of the uncertainty surrounding these issues.

6.3 Bayesian network

The Bayesian network structure consists of a directed acyclic graph (DAG) network, made up of nodes that represent variables under analysis. The purpose of the DAG is to predict the outcomes of these variables. The nodes interact through linkages of edges that define their relationships, see Figure 6.1.
The relationships are defined probabilistically within a set of conditional probability tables (CPTs). The nodes CPT, will be in a particular state, depending on the given states of the other nodes that are linked to it. BNs take the name and use from the Bayes' theorem, which elucidates the relationship between events and outcomes. The theorem allows the probabilities of variables whose state was unknown to be undated given some new set of observations. Therefore, BNs allow reasoning to proceed in any direction across the Bayesian network of variables under analysis, but it should be noted that any direction is permitted i.e. not limited to the direction of the linkage.

Example: A house alarm (C), was it the result of a fault in the alarm (A) or an actual attempt to break into the house (B). If the result is known that (B) is the most probable outcome, then (C) becomes dependent on (B).

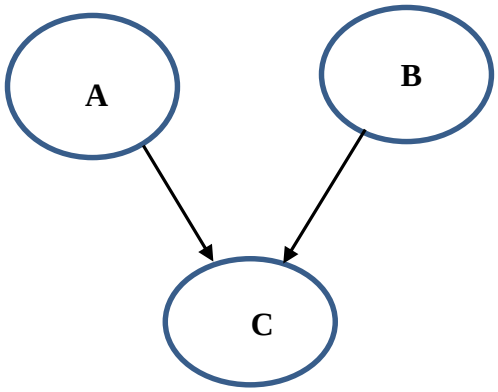


Figure 6.1: A simple BBN. Where p (A, B, C) = p(A) p(B) p(C| A,B). Given C, observing the conditions of A and B. A and B are slightly independent but become dependent once C is known.

6.4 Sensitivity analysis

An important stage during the Bayesian model development was the use of sensitivity analysis (SA). Sensitivity analysis was conducted using the Value of Information function within the Hugin® software to measure the sensitivity of the constructed wetland system probabilities about the variables (i.e. the dominant variables within the wetland dataset). SA measures the degree of influence that one variable has over another variable and can be used to rank the importance of each variable on the model outcome. Sensitivity to findings can be quantified by two methods, (1) entropy and (2) mutual information, within Hugin® the ranking was displayed as mutual information, see Table 6.1, where mutual information was used to measure the effect of one variable (X) on another (Y) (Korb and Nicholson, 2004). The sensitivity ranking process has relative importance within the BN model, as key variables are identified that can be rectified therefore improving the wetlands performance.

Table 6.1: Sensitivity analysis based on Mutual Information (Entropy) for nodes in decreasing order of influence of the CW variables

Node	Mutual information	Wetland Parameter
Wetland area	1.51	Wetland design
Plants x area	0.38	Wetland design
Soil porosity	0.35	Wetland design
Plants/m^2	0.34	Wetland design
Total coliforms	0.22	Wetland bacteria
E.coli	0.18	Wetland bacteria
Enterococci	0.14	Wetland bacteria
Chloride	0.07	Wetland inorganic ions
Phosphate	0.04	Wetland inorganic ions
Fluoride	0.03	Wetland inorganic ions
Sulphate	0.03	Wetland inorganic ions
Nitrate	0.02	Wetland inorganic ions
Nitrite	0.02	Wetland inorganic ions

Mutual Informations is a quantity that measures the relationship between two random variables that are sampled simultaneously. It measures how much information (entropy) is communicated between the two variables.

6.5 Soil and sludge sampling and preparation

See Chapter 2, section 2.19

6.6 Soil bacterial analysis using the IDEXX Method

The supernatant extracts were analysed using the Idexx Method for Colisure® and Enterolert® assays as per manufacturer's specifications (see Chapter 2). The same methodology that was used to analyse the surface water samples, was repeated on the sludge and soil samples.

6.7 Soil Porosity Test

The soil porosity testing was evaluated on the soils all sample sites. Four soil samples were retrieved from May, June, July and Oct. The porosity was evaluated as per the material and methods section (Chapter 2). Table 6.1 shows the soil porosity ranges for sediments, used to classify the wetland soil samples.

Table 6.1 Soil porosity range for sediments (Fetter, 1994)

Material	Porosity (%)
well-sorted sand or gravel	25-50
sand and gravel, mixed	20-35
glacial till	10-20
silt	35-50
clay	33-60

6.8 Wetland Plants

The constructed wetland treating abattoir waste, which is the focus of this study contains several plants species to help remediate the abattoir wastewaters, with an 80% - 20% Common Reed (*Phragmites australis)* and Lesser Bulrush (*Typha latifolia*) dominating pond 6 and a 85% -15% Yellow Flag Iris (*Iris pseudacorus*) and Lesser Bulrush present in ponds 9 and Yellow Flag Iris dominating 99% of pond 12.

6.9 Variables analysed in the Bayesian network

Wetland sludge and soil variables: porosity, plants x area, wetland area, plants per metre squared, total coliforms, enterococci, *E.coli*, nitrate, nitrite, sulphate, chloride, phosphate, fluoride, see Appendix D.

The Hugin® Lite version 7.1 BBN modelling software (Hugin, Aalborg, Denmark, http://www.hugin.com/) [Accessed 03 2014] was used to develop and implement the soil bacteria BN model.

The soil and sludge Bayesian belief network

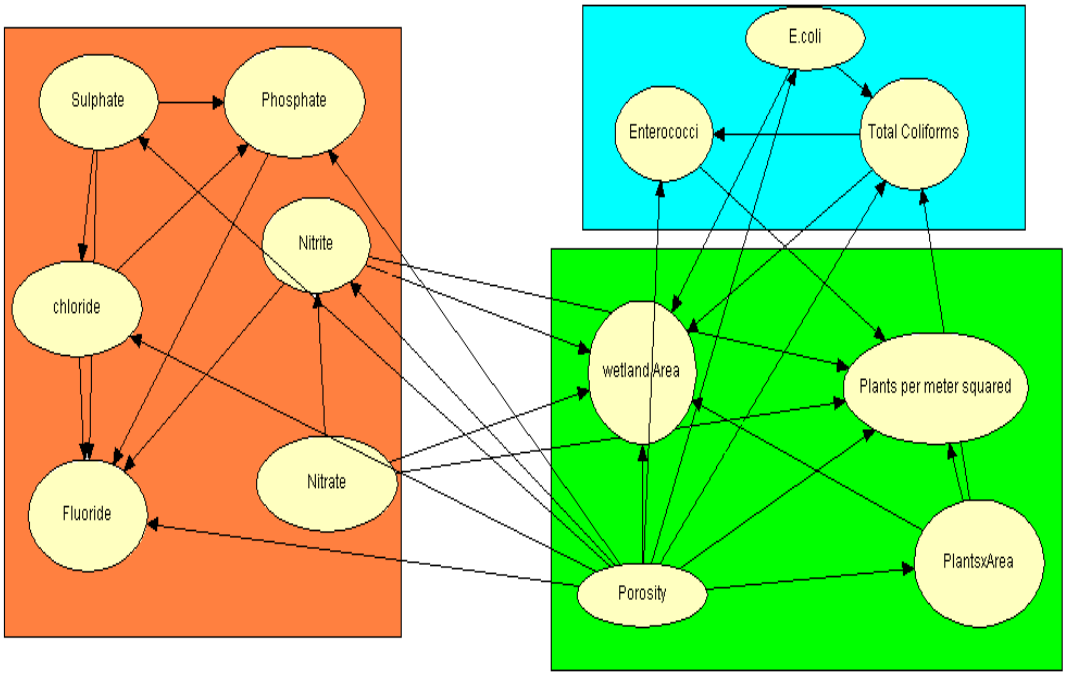


Figure 6.2: The soil and sludge Bayesian belief network. The variables are subdivided into three groups. Group (1: wetland design) [plants x area, plants/m^2, soil porosity and wetland area]; Group (2: inorganic ions) [chloride, fluoride, nitrate, nitrite, phosphate and sulphate]; Group (3: wetland soil/sludge bacteria) [*E.coli*, enterococci and total coliforms].

6.10 Results and discussion

Figure 6.2 shows the BBN that was developed using the Hugin® software the figure highlights the three sub-groups by wetland design, bacteria, and inorganic loadings. Table 6.2 indicates the most sensitive variables in the wetland system, with the wetland area being the most sensitive along with the other wetland design parameters (plant related variables and soil porosity) along with the bacteria loadings. This strongly suggests that the soil and sludge indicator bacteria concentrations are defined by the design parameters of the wetland, inclusive of the plants used to remediate the wastewaters and the soil porosity, with the

inorganic ion loadings the least sensitive. Figure 6.3 shows the variability of the indicator bacterial loadings throughout the seven-month sampling period [April, May, June, July, Sept, Oct, Dec]. All ponds exhibit the increased bacterial loadings during the summer and autumn months. The increased total coliforms and enterococci counts are the most dominant trends within Figure 6.3. In particular, the larger increase in soil bacteria was evident in pond 12 samples, with enterococci becoming dominant. During the summer months' pond 12 was dry with no water retained in the pond and pond 9 water level very low approximately 20cm or less.

Figure 6.4 shows the graph of the mean bacterial concentrations versus the mean plant coverage per metre^2 with Figure 6.5 showing the mean soil porosity versus the mean plant coverage per metre^2. Figure 6.4 shows that there were no plants present, with a very large sludge bacterial concentration range observed in pond 1($\log_{10}$ 8.26-11.96 CFU/g). In Pond 6 the soil bacterial concentration range decreases ($\log_{10}$ 5.17-7.79 CFU/g), with a corresponding increase in plant coverage to an average plant coverage of 6.5 plants/m^2. In pond 9 the soil bacteria concentration range decreases ($\log_{10}$ 3.53- 7.79 CFU/g), with a corresponding decrease in plant coverage to an average plant coverage of 2.64 plants/m^2. In pond 12 the soil bacteria concentration range decreases ($\log_{10}$ 2.11- 6.93 CFU/g), with a corresponding decrease in plant coverage to an average plant coverage of 2.21 plants/m^2.

Figure 6.5 showing the plant coverage per metre^2 in conjunction with the average porosity of the wetland soils reveals a similar trend that was observed in Figure 6.4. Table 6.2 provides a breakdown of pond, soil type, soil porosity and plant coverage

Table 6.2: Soil type, soil porosity and plants coverage by pond

Pond	soil type	soil porosity	plant coverage
1	clay and silt	43.30%	None
6	clay and silt	40.80%	6.5 plants/m^2
9	small sand granules	27.40%	2.64 plants/m^2
12	large and small sand granules	23.54%	2.21 plants/m^2

The Bayesian network model development with sensitivity analysis (i.e. mutual information) highlights that poor soil porosity and loss of plant coverage are contributing factors to increased bacterial soil concentrations within the wetland ponds 9 and 12. Figure 6.3 shows that the largest bacterial loading values within these ponds are comparable with pond 6 bacterial loading values.

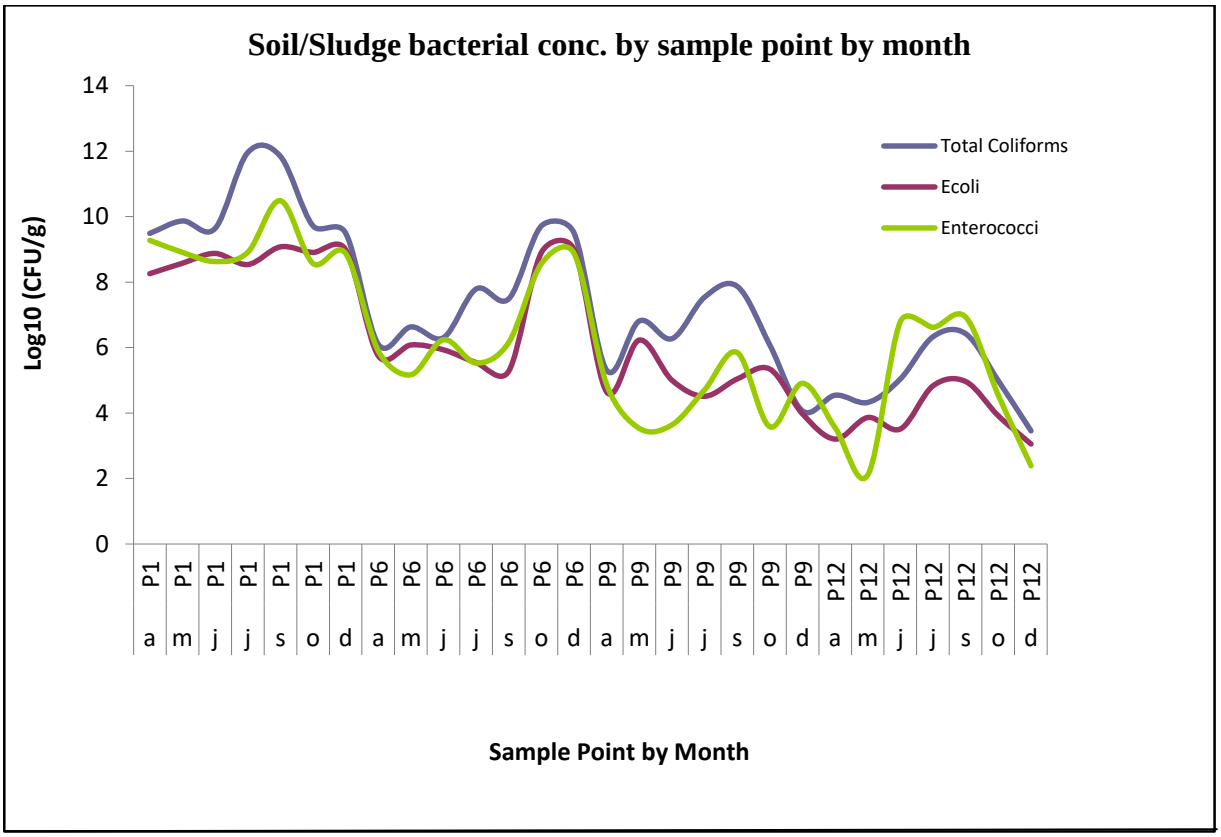


Figure 6.3: The sludge and soil bacteria concentration. Sample points between the months April 2007 to December 2007. [a = April, m = May, j = June, j = July, s = Sept, o = Oct, d = Dec]

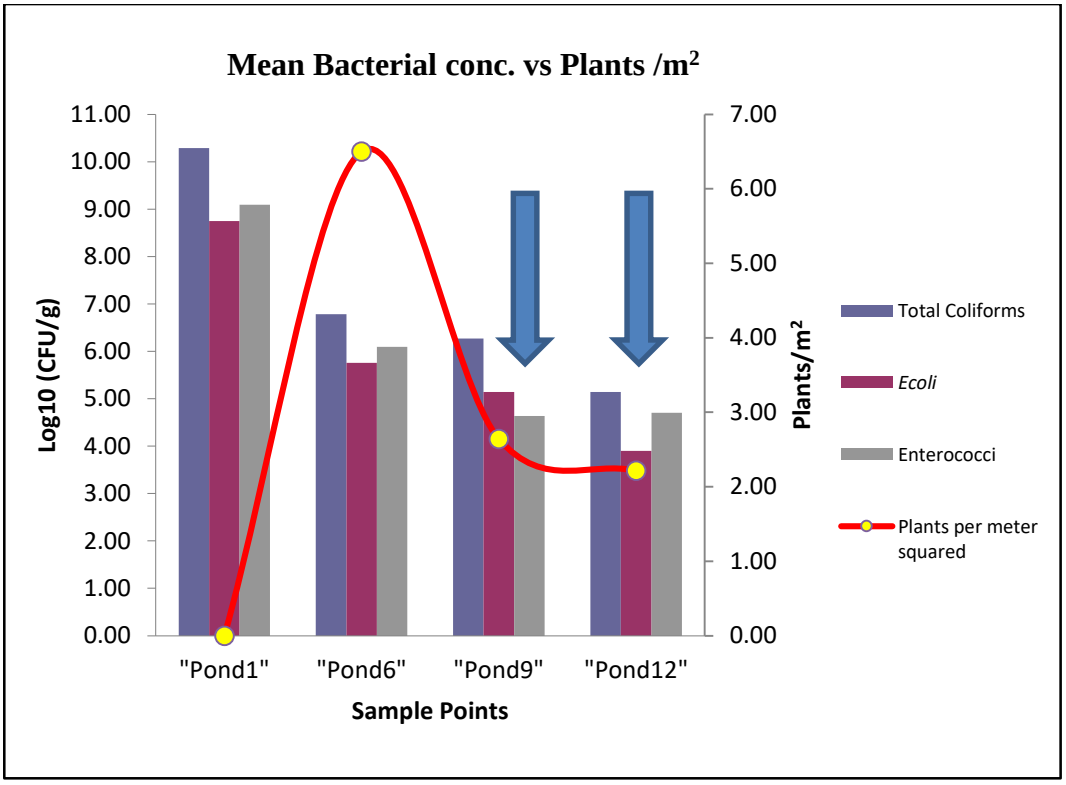


Figure 6.4: The mean bacterial concentrations versus the mean plant coverage per metre^2.

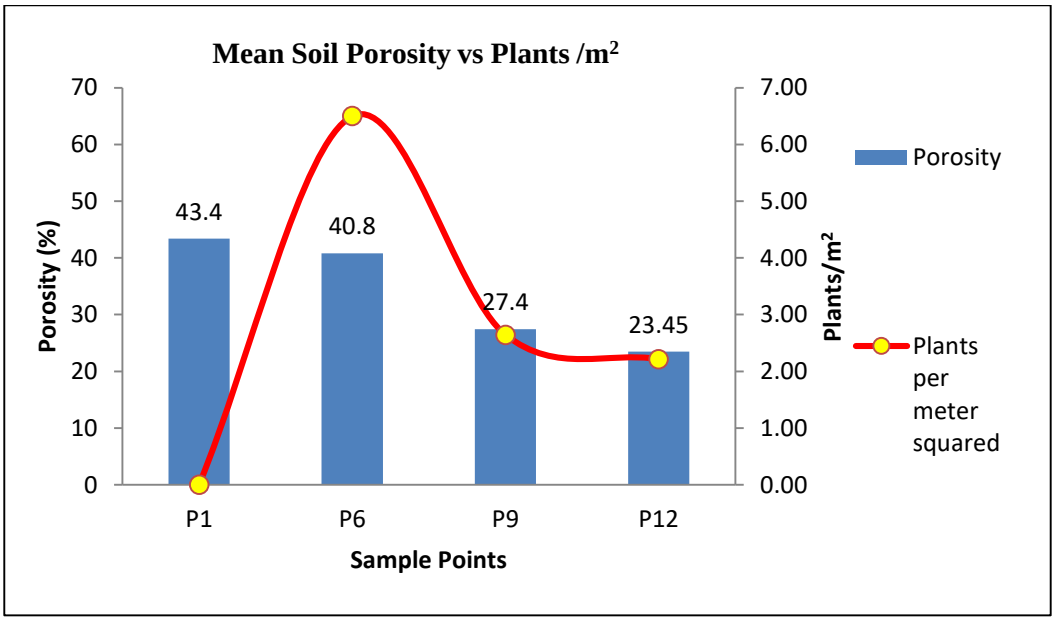


Figure 6.5: The mean soil porosity versus the mean plant coverage per metre^2.

Another factor that may have enabled higher bacterial numbers in the latter part of the wetland system was the presence of nutrients in the soils of pond 9 and pond 12 see Figure 6.6. The presence of these nutrients, no matter how miniscule,

in high concentrations along with increased temperatures during the summer months, could have resulted, in part, to the increased soil bacterial loadings. The most dominant ion present, in pond 12 was chloride with a mean (5 month) value = 14.59 (mg/L), sulphate mean value = 3.51 (mg/L) and nitrate mean value = 1.89 (mg/L).

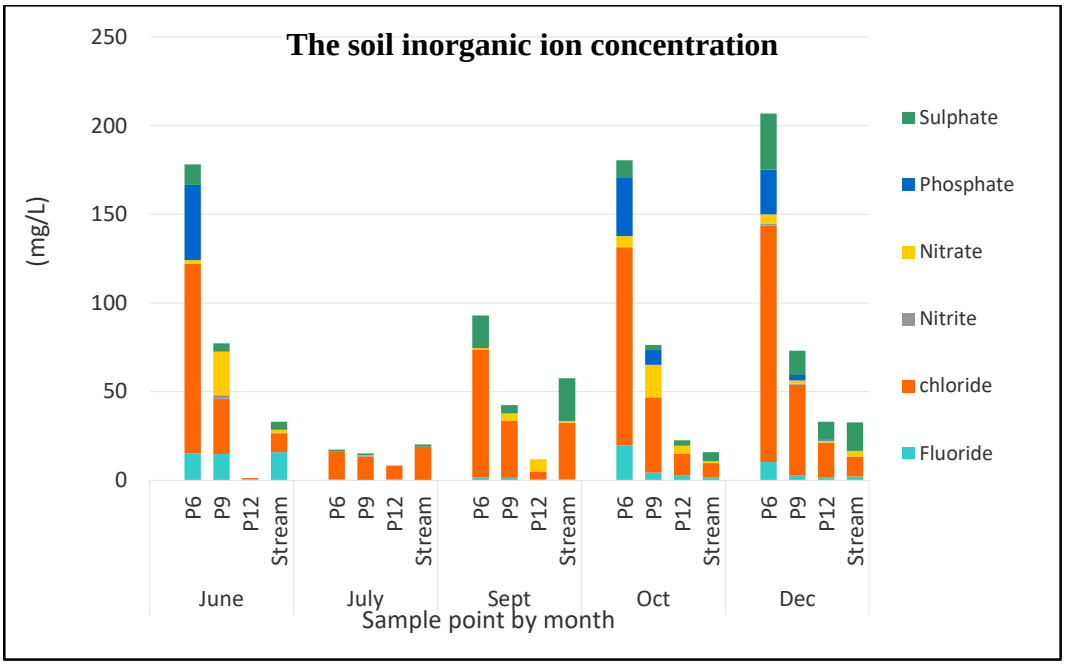


Figure 6.6: The soil inorganic ion concentration (stacked bar chart). Taken from pond 6, pond 9, and pond 12 and the stream.

Note: No inorganic ions were analysed for pond 1 sludge sampling. This was due to previous trial runs which resulted in clogging of the Ion Chromatography column. It was found that a build-up of a bacterial bio-film was present on the column. It was decided not to continue with any pond 1 sludge analysis.

The dominant cause and effect node in the Bayesian network model was the variable porosity (Figure 6.2). The porosity node interacts with all twelve nodes in the BN. Notice the twelve linkages or edges emanating from the porosity node to all other network nodes. Therefore, making this variable (node) an important variable within the Bayesian network. As discussed previously, the sensitivity where the wetland area was the dominant variable with a mutual information value = 1.51 and soil porosity mutual information value = 0.35, see Table 6.1.

Sensitivity and cause and effect should not be misconstrued. Variables (X) and (Y) interacts with variable (Z), with (X) and (Y) causing an effect on (Z), similar to the the interactions observed in Figure 6.1. The house alarm can be trigeerd from a burgular entering the house or a fault with the alarm. Sensitivity analysis reveals the strength associated with these interactions i.e the degree of influence both (X) and (Y) (causes) have over the resultant (Z) (the effect).

6.11 Conclusions

The Bayesian network model developed in conjunction with the sensitivity analysis (SA), provided a useful insight into the complex workings of the wetland system. From a risk assessment perspective, it enables wetland designers to better understand the importance of correct design parameters such has the pond size (wetland area), the type of plants used in the system and the plant coverage required. But of critical importance are the porosity and permeability of the soil when constructing the wetland system. A domino effect can be observed when during increased air/water temperatures and underlying poor soil porosity leads to loss of water, loss of water leads to loss of plants, loss of plants leads to increased soil bacterial indicator concentrations.

One probable cause for increased bacterial loadings in the soil sediment during the summer months was the increased air temperature, loss of plant coverage within the ponds, especially ponds 9 and 12. The plant loss could be attributed to poor soil porosity, which allows water to percolate through the ponds' soil and sediment structures, increasing water loss within these ponds.

Review Figure 6.3 and Appendix D, it can see that pond 1 has very high bacterial loadings in the sludge (almost $\log_{10}$ 12 CFU/g), for both July and September. One possibility is that in September, the increased rainfall releases the sludge sediments (with high bacterial loadings) into the constructed wetland system and as a result the constructed wetland system cannot attenuate or reduce this "signal". In the month of October, pond 1 sludge bacterial loading values were less than $\log_{10}$ 10 CFU/g, with the potential cause related to the a decrease in temperature resulting in reduced soil and sediments bacterial values.

Chapter 7. Self-Organising Criticality and Entropy analysis.

Is the constructed wetland treating abattoir wastewater a complex system?

7.1 Summary

This chapter examines the complexity of a free water constructed wetland system (FWS) treating abattoir wastewater, using key wetland variables to ascertain the mean behaviour of the wetland system over a twelve-month period. One of the methods to deduce complexity is to understand the self-organised criticality of the system under investigation using power laws. Different actions occur within the system, with periods of activity, followed by latency. These are similar to descriptions at a self-organised critical (SOC) point, where a slight disturbance may induce an avalanche of activity. It was observed that some of the wetland variables displayed strong SOC trends, resulting in efficient and very efficient power law distributions i.e. removal of wetland contaminants, other wetland variables displayed poor efficiencies, where exponential and linear trends were observed these constituted poor removal efficiencies. Two uses of Shannon Entropy analysis are demonstrated (1) soft-scaled entropy (SCE) and (2) load based entropy (LBE). SCE was used to reveal the critical state or transition within the wetland system using soft-scaled or normalised entropy, where the wetland variables are normalised and then Shannon entropy realised. The critical state was found to be located between pond 6 and pond 9. The overall loading effects of the wetland variables within the wetland using load based entropy (LBE), where the influent wastewater was overloading pond 1 by inadequately treated abattoir wastewater, resulting in overloading i.e. the DAF was not treating the treating the wetland effectively. The mean overall efficiency of the wetland system treating all key wetland variables over the 12-month period was 83%. Several key wetland variables were found to have the following inadequate removal efficiencies such as the total coliforms, *E.coli* and enterococci exhibiting efficiencies of 68%, 74% and 69% respectively. Other key wetland variables such as nitrate (60%) and phosphate (79%) displayed poor removal efficiencies. It was noted the wetland cannot efficiently maintain a steady water depth (67%).

7.2 Introduction

First described by Bak *et al*, 1998, self-organising criticality (SOC) was introduced as a description of dissipative dynamic systems (DDS). DSS is an open system model extended with a capability to continuously impose a transformation or the spontaneous appearance of symmetry and the formation of complex structures. Systems that obey SOC are characterised by the organisation of a complex system state into a critical state, where avalanches occur, i.e. or similar spontaneous behaviours over a wide range of sizes (i.e. scale free: variations of all sizes are seen). This self-organisation of the complex system into a critical state has been used to describe the behaviour of sand piles (Bak *et al*, 1998), earthquakes (Carlson and Langer, 1989), ecosystems (Jorgensen *et al*., 1998), landslides (Hergarten and Neugebauer, 1998), power outages (Carreras *et al*, 2001), traffic flow (Atzori *et al*., 2006), air pollution (Shi and Liu, 2009), the meandering behaviour of the Ganges river in India (Carling *et al*., 2016) and workplace accidents (Mauro *et al*., 2018). A characteristic feature of critical-point behaviour, the data distribution follows a power law, where the power law tells us about the probability of fluctuations of a given size, the critical state is highly precarious and liable to collapse one way or the other at the slightest provocation (Ball, 2005a) and that Self-organised criticality is a property of non-equilibrium systems (Ball, 2005c). A power law is the dependent variable *y* varies as a power of the independent variable *x* (i.e. $y = Ax^{-\alpha}$, where A is a proportionality constant and α is a scaling exponent). Power laws are linear functions, when the logarithm of both dependent and independent variables is taken i.e. log-log plots (Pascual and Guichard, 2005). In the case of the constructed wetland system it exists in a non-equilibrium situation, the fluctuations within the wetland systems are scale-free, the wetland can be considered to exist in a critical state. The critical state here is a transition state or a phase change, where the wetland system changes from congested to non-congested, similar to traffic in computer networks (Valverde and Solé, 2002). In the case of the wetland system this constitutes a first order change from anaerobic to aerobic state, but it is not an equilibrium state, because abattoir wastewaters are constantly being added. This continuous influent of wastewater was the driving force preventing equilibrium being attained, therefore the wetland was in a state of dynamic dis-equilibrium. The best

method to show the power law is to plot the logarithm of the count size of the variable [Log (N)] versus the logarithm of its rank [Log (r)] and when log-log plots of the data are created, the graph displays a linear trend and a slope (m) of greater than or equal to -1, expresses that the variable has SOC characteristics (Ball, 2005b).

Considering the CW under review as a complex system, which are opaque and uncertain, these systems also exhibit threshold behaviour … a sharp sudden move or flip to a new state (Homer-Dixon, 2011). There is also need to understand the conditions under which thresholds are likely to be crossed and what mechanisms underlie threshold behaviour is critical (Shenton *et al*., 2007).

The aim of the science of self-organised criticality is to yield insight into the fundamental question of why nature in complex, not simple, as the laws of physics imply (Bak, 1996).

A system defined as a group of components functioning as a whole.

One of the key signs of system complexity is Power Law distributions for temporal or spatial correlations, which may be self-similar of fractal structures. An SOC system is self-tuning, no external parameters affect the initial state. These can be classified as open dissipative systems and the components of the systems are governed by simple rules. A threshold exists within the system and forces accumulate in the system until it exceeds the threshold. The system progresses naturally towards the critical state, whereby small agitations in the system can lead to system effects called avalanches and this avalanche effect happens regardless of the initial state of the system. SOC is the intermediate state between order and chaos in which complexity exists (See Figure 7.1).

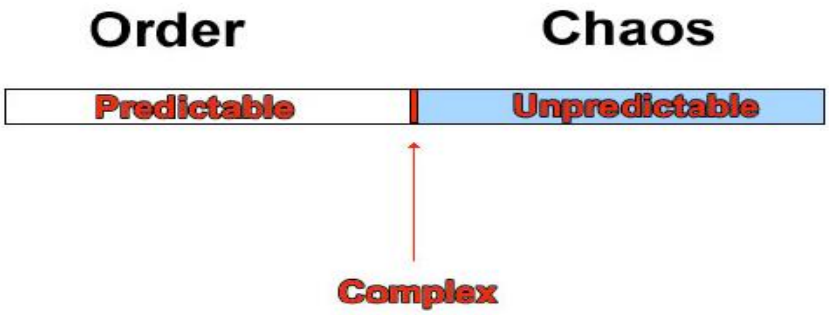


Figure 7.1: The region in which complexity exists, between predictable and unpredictable.

Criticality is mathematically defined as the state of highest efficiency in a complex system. (Yurth and Ayres, 2005*)* or SOC is the most efficient way a dynamic system can organise itself. Another view of criticality is the state the system is said to be, if poised at a phase transition.

For the purposes of this chapter, efficiency is defined by the following equation

$$[(V_{input} - V_{output})/V_{input}] \qquad \text{Eq(7.1)}$$

where V_{input} is the input value of any variable at the DAF and V_{output} is the output value of that variable at pond12. While reviewing the wetland data, it was found that a clear pattern was forming when the collective mean of the data at their respective sampling points was revealing power-law distributions. Table 7.1 shows the dataset used for the SOC verification using seventeen variables taken from the surface waters of the wetland.

The question posed was whether any of the wetland's key variables under analysis exhibit SOC and if so, did this correspond to high efficiency?

7.3 Wetland variables

The key variables analysed from the surface waters of the wetlands were:

1. Indicator bacterial concentrations which consist $\log_{10}$ CFU/100ml (total coliforms, *E.coli*, enterococci and total bacteria).
2. Physical-chemical values which consists of BOD_5 (mg/L), turbidity (NTU), dissolved oxygen (mg/L), conductivity (mS/cm) redox potential (mV)
3. Inorganic ions which consists of ammonia, chloride, fluoride, nitrate, nitrite, phosphate, sulphate [all inorganic ions measured in (mg/L)].
4. Water Depth (m).

The analysis was performed on the mean datasets of each variable i.e. from February 2007 to February 2008, consisting of 11 months of data (n=11) See Table 7.1.

7.4 Entropy

Entropy is typically attributed to the Second law of Thermodynamics, where the entropy (dS) of a system is the infinitesimal transfer of heat (δQ) in a closed system, driving a reversible process, divided by the equilibrium temperature (T) of the system, equation (7.2):

$$dS = \frac{\delta Q}{T} \qquad \text{Eq(7.2)}$$

Figure 7.2 describes the behaviour of equation (7.2), with respect to a glass of ice at temperature (T_1) is placed in a room at temperature (T_2).

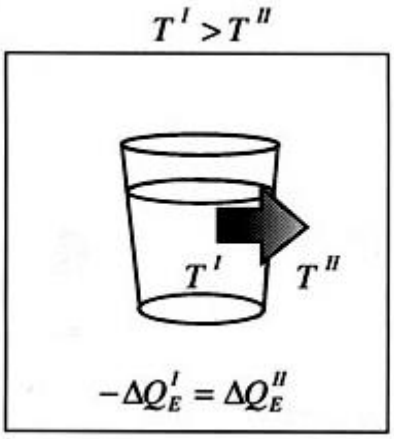


Figure 7.2: Entropy – glass of ice

Figure 7.2: A glass of liquid with ice at temperature T^I is placed in a room at temperature T^{II} such that the disequilibrium produces a field potential that results in a flow of energy in the form of heat from the glass to the room so as to drain the potential until it is minimized (the entropy is maximized) at which time thermodynamic equilibrium is reached and all flows stop. This refers to the conservation of energy in that the flow from the glass equals the flow of heat.

In this chapter entropy is used as a measure of information theory to find the critical transition within the wetland system. Entropy is a measure of the uncertainty associated with a random variable and in this context Shannon entropy is used. Shannon entropy quantifies the probabilistic value of the information contained in a message or in this case the message represents the wetland variables as seen in Table 7.1 and the unit to represent Information Entropy = *bits*. We are further looking to describe the behaviour of the wetland system changing from heterogenous to homogenous (disorder to order) states within the wetland system, by performing two methods of entropy analysis using the data in Table 7.1. Entropy, abbreviated as S, in its discrete form, is expressed as follows:

$$\text{Shannon Entropy (S)} = -\sum_{k=0}^{n} p(\text{i}) \log 2\ p(\text{i}) \qquad \text{Eq(7.3)}$$

Where the given probability distribution has *n* elements and where p_i is the probability of the *i*th value of the distribution (Shannon, 1948). When calculating

entropy, 0 $\log_2$ 0, is defined as zero. The $\log_2$ measure was used in binary classification because this represents the number of bits needed to specify the value in which the probability exists, therefore the largest value at p(i) = 0.5. Using Method 1, the analysis defines the critical state/transition within the constructed wetland (CW), where the critical state is the most efficient state that can be achieved (Paczuski and Bak, 1999); Method 2 defines the loading effects within the CW.

Two methods (Method 1) Soft-scaled Norm(z) Entropy and (Method 2) Load Norm (z) Entropy were used to analyse the wetland variable data set.
Using Method 1, the dataset was first transform into "soft-scaled" values i.e. linear transformation, using equation (7.4)

$$SC = (X_i - X_m)/Std_m \qquad \text{Eq(7.4)}$$

where SC= soft scaled data, X_i = variable (X), X_m = mean value of the group of variable(X), Std_m = standard deviation of the group of variable (X). This transformation leaves the data with a mean of zero and standard deviation of 1.
The "soft scaled" data undergoes a NORMSDIST (S) distribution in Excel 2007®, which returns a probability that the observed value of a standard normal random variable will be less than or equal to z, i.e. a value between zero and one. This type of analysis can also be attributed to Probit analysis (Jones, 1998). The normdist (S) values are entered into entropy (S) equation (7.3).

Using Method 2, the data was transformed into load values, where the total of each sample point was found and the loading was evaluated by dividing the value at the sample point by the sum total.

$$V_L = (X_i/X_T) \qquad \text{Eq(7.5)}$$

where V_L = Load, X_i = variable (X), X_T = sum total of group of variables.
Figure 7.3 lists the progression of the analysis performed by (SCE) and (LBE).

Combining both self-organising criticality and entropy has been used in artificial ecosystems (Yang, 2002), information theory (Dewar, 2003) and in earthquake analysis (Main and Naylor, 2010).

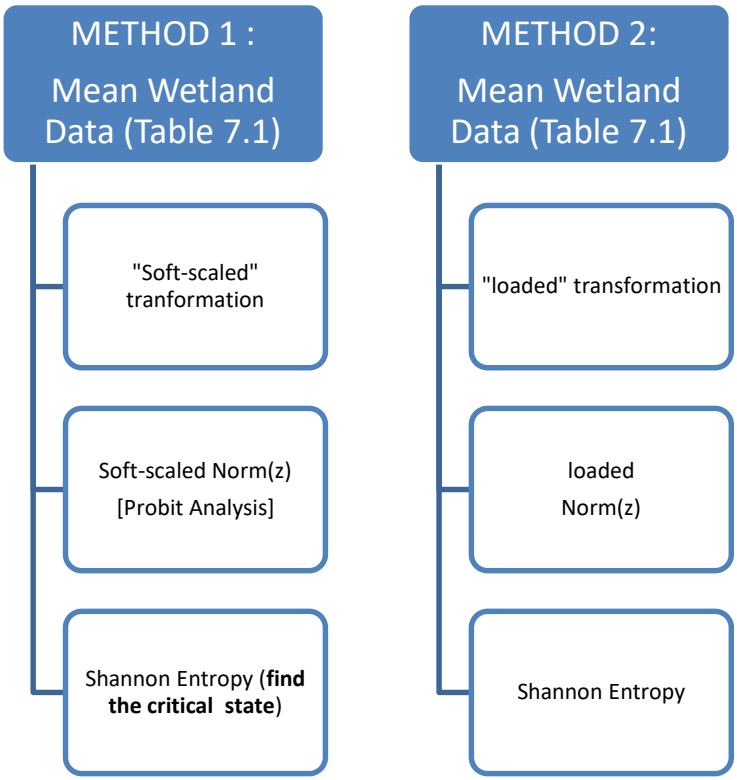


Figure 7.3: The two methods for entropy analysis.

Other tests were completed on the wetland data, such as the use of Spearman Rank Correlations, see Table 7.2. To understand the strong and weak interactions of the wetland variables. Table 7.3 tabulates the wetland variables by their efficiencies from very efficient, efficient and finally inefficient.

Table 7.1: The mean wetland dataset.

	Total Coliforms	***E.coli***	**Enterococci**	**Total Bacteria**			
DAF	9.09	8.55	7.71	9.26			
P1	7.82	6.79	6.79	8.19			
P6	5.81	5.36	5.13	6.16			
P9	4.53	4.34	4.13	5.01			
P12	2.88	2.25	2.4	3.17			
mean	6.03	5.46	5.23	6.36			
Stdev	2.49	2.39	2.11	2.44			
Efficiency	0.68	0.74	0.69	0.66			
	Ammonia	**Fluoride**	**Chloride**	**Nitrate**	**Nitrite**	**Phosphate**	**Sulphate**
DAF	341.87	34.72	463.35	121.33	21.67	87.94	20.68
P1	422.77	4.77	275.84	96.43	1.46	62.16	25.95
P6	75.25	0.38	136.2	74.4	0.24	43.97	6.79
P9	9.73	0.23	106.4	50.8	0.14	27.78	3.72
P12	2.09	0.12	56.18	48.78	0	18.42	1.42
mean	170.34	8.04	207.59	78.35	4.7	48.05	11.71
Stdev	197.67	15.04	164.57	30.92	9.51	27.82	10.92
Efficiency	0.99	1	0.88	0.6	1	0.79	0.93
	Turbidity	**BOD**	**Dissolved Oxygen**	**Redox Potential**	**Conductivity**	**Water depth**	
DAF	943.18	3836.64	0.4	-253.91	4.58	1.48	
P1	350.88	1305	0.76	-204.09	2.67	0.84	
P6	61.14	92.75	2.37	25.45	1.06	0.67	
P9	48.55	50.85	5.91	33.2	0.91	0.51	
P12	47.22	9.96	6.02	42.75	0.73	0.5	
mean	290.19	1059.04	3.09	-71.32	1.99	0.8	
Stdev	387.29	1645.16	2.72	145.15	1.64	0.41	
Efficiency	0.95	1	0.93	0.86	0.84	0.67	

Table 7.1 shows the dataset utilised for SOC and entropy analysis. The bacterial concentration units were $\log_{10}$ (CFU/100ml). Each sample point shows the mean value of that sample point by variable for 11 months.

The next sections (Figure 7.5 - 7.8) display the resulting self-organised critical (SOC) distributions for the wetland variables. The distributions will be either linear, exponential or power laws.

The bacteria trends are shown in section 7.5, with linear distribution observed. No power laws are present, indicating poor interactions in the CW, i.e. poor removal efficiencies.

7.5 Self-organised criticality (SOC bacteria graphs: (Total coliforms, *E.coli*, enterococci and total bacteria)

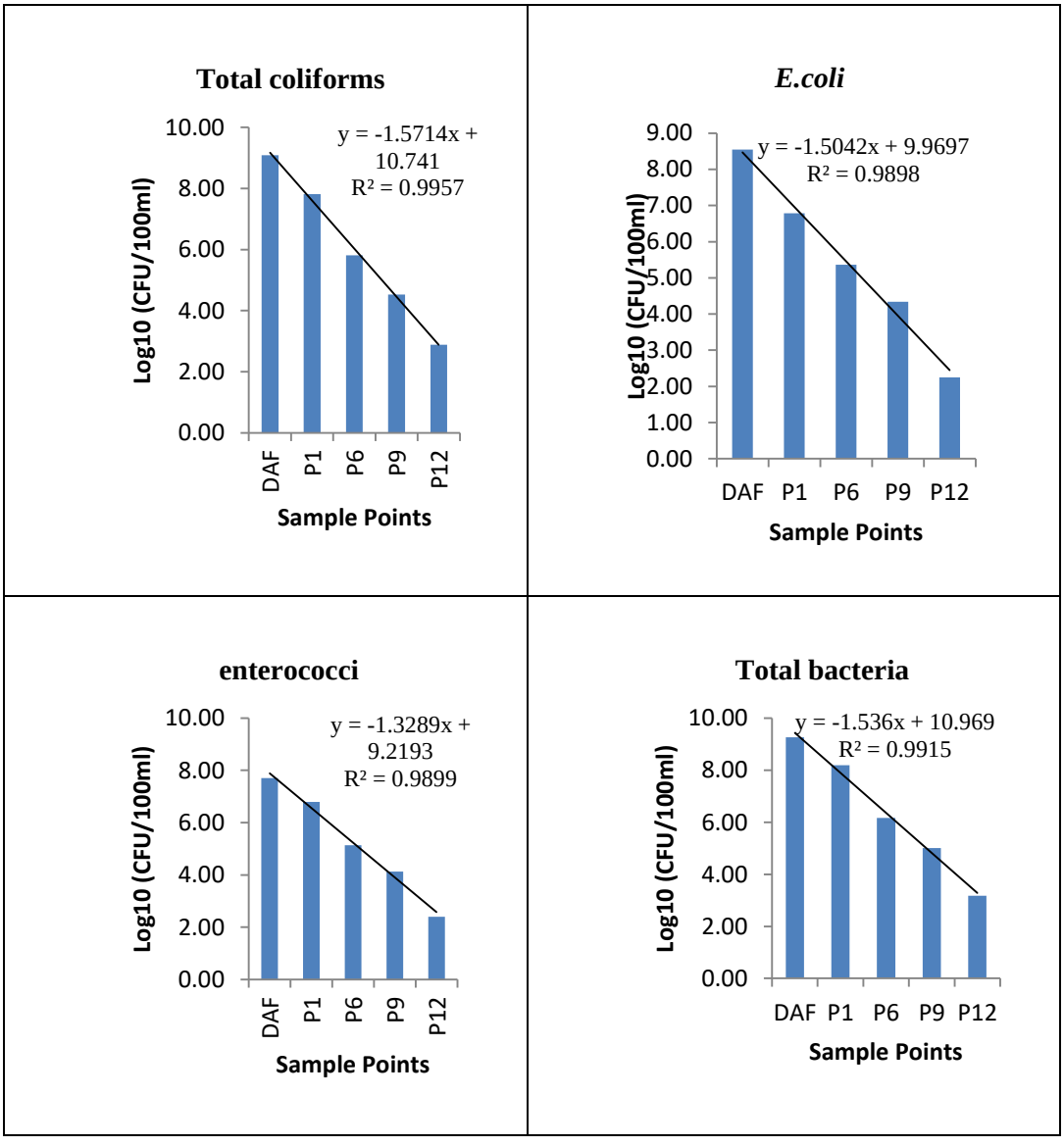


Figure 7.4: The indicator bacteria trends. The bacteria display an ordered linear decrease from input (DAF) to output (pond12).

The physical-chemical variables trends are shown in section 7.6, with power law distributions observed, indicating strong interactions in the CW, i.e. good removal efficiencies.

7.6 Self-organised criticality (SOC) Physical-chemical graphs (Turbidity, BOD_5, Dissolved Oxygen, Conductivity and Redox Potential)

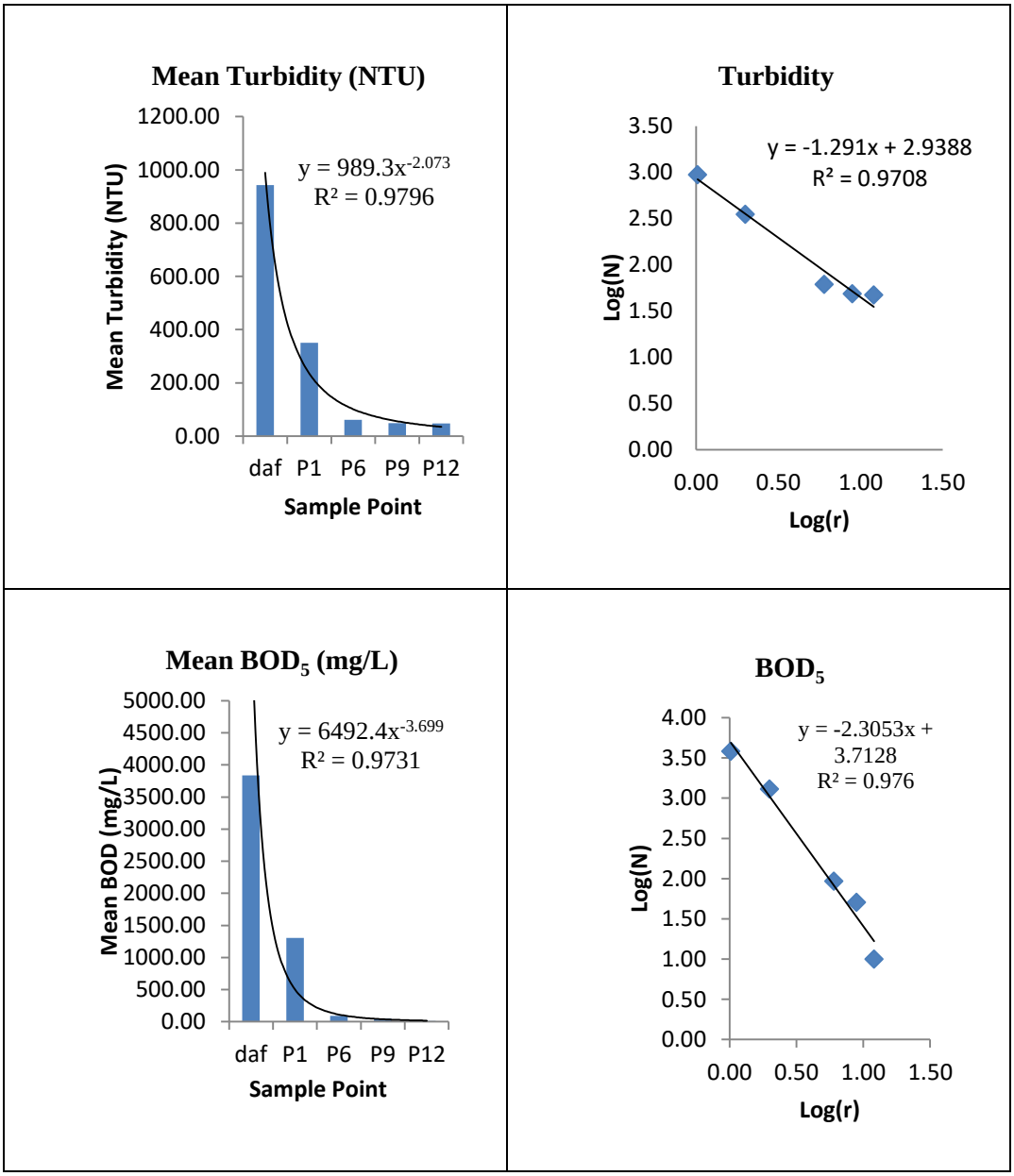


Figure 7.5: The physical-chemical variables trends (1). The power law distributions for the turbidity and BOD_5 data along with their corresponding log-log plot, where the slope of the linear equation shows the self-organising criticality (SOC) of the variable.

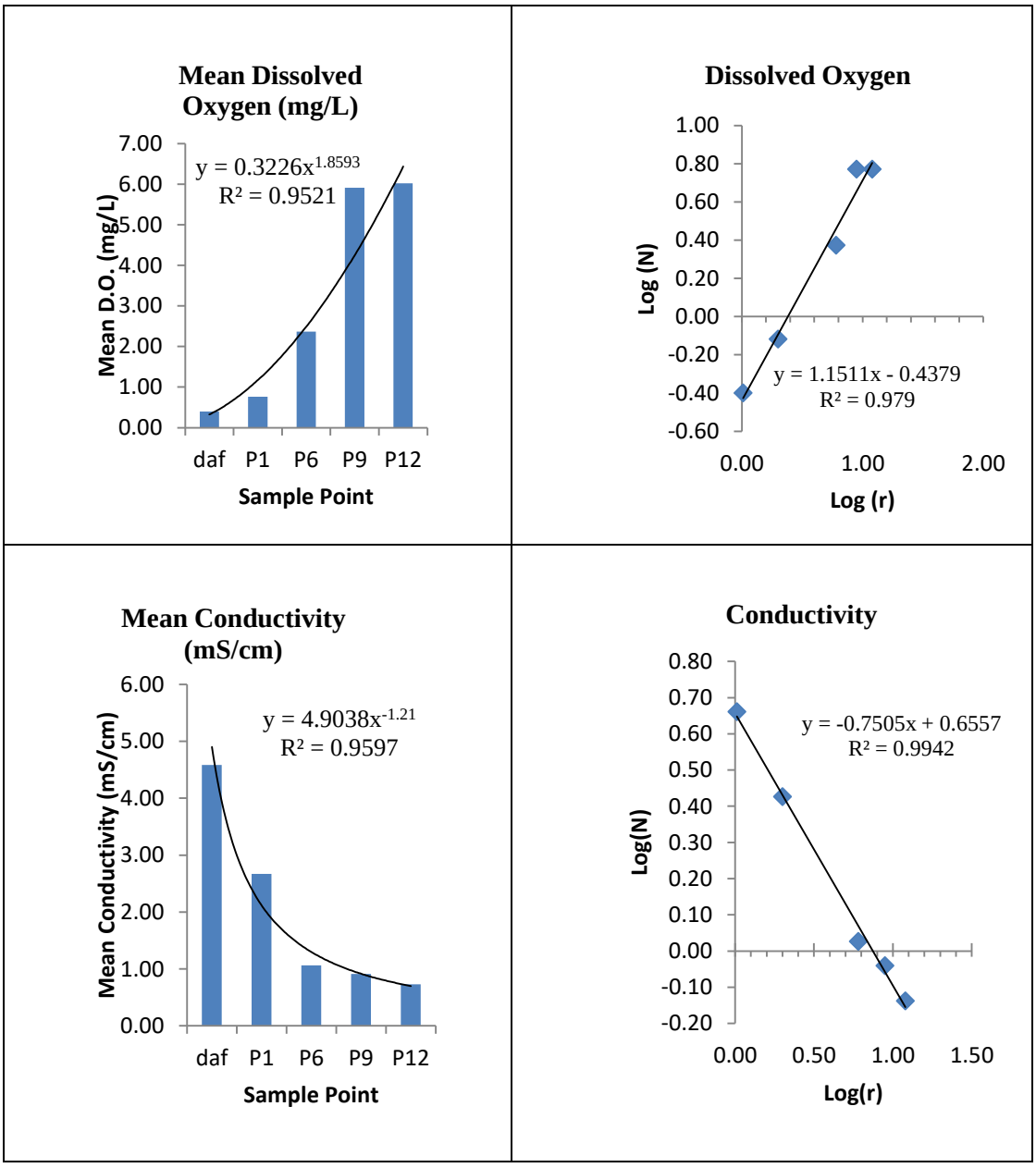


Figure 7.6: The physical-chemical variables trends, the power law distributions for the dissolved oxygen and conductivity data along with their corresponding log-log plots. Where the slope of the linear equation shows the self-organising criticality (SOC) of the variable.

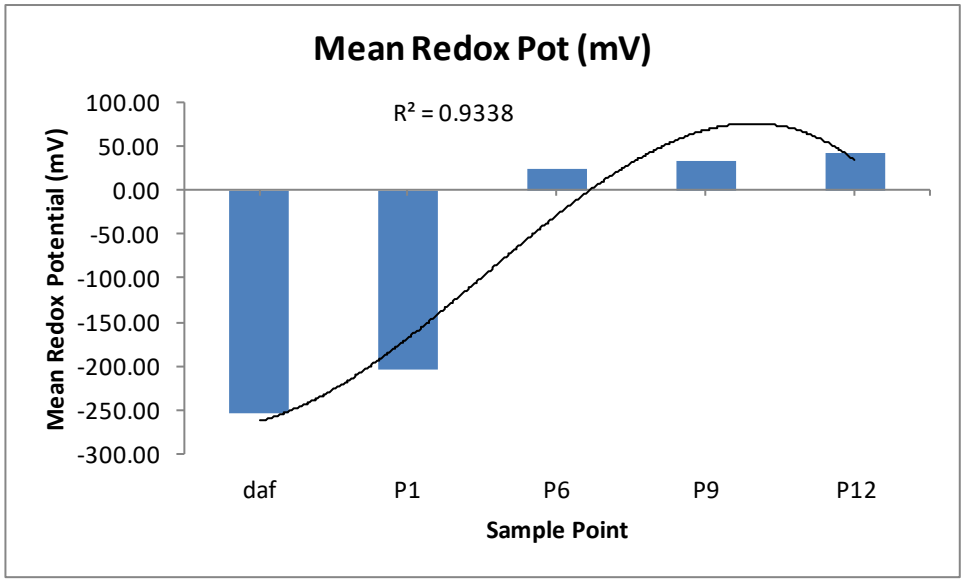


Figure 7.7: The mean redox potential (mV) trend. With a phase transition from anaerobic (negative mV) to aerobic (positive mV) occurring at pond 6.

The inorganic ion (Nitrate and Phosphate) trends are shown in section 7.7, with exponential distributions observed. No power laws are present, indicating medium interactions in the CW, i.e. average removal efficiencies.

7.7 Self-organised criticality (SOC) Inorganic ions graphs (Nitrate and Phosphate)

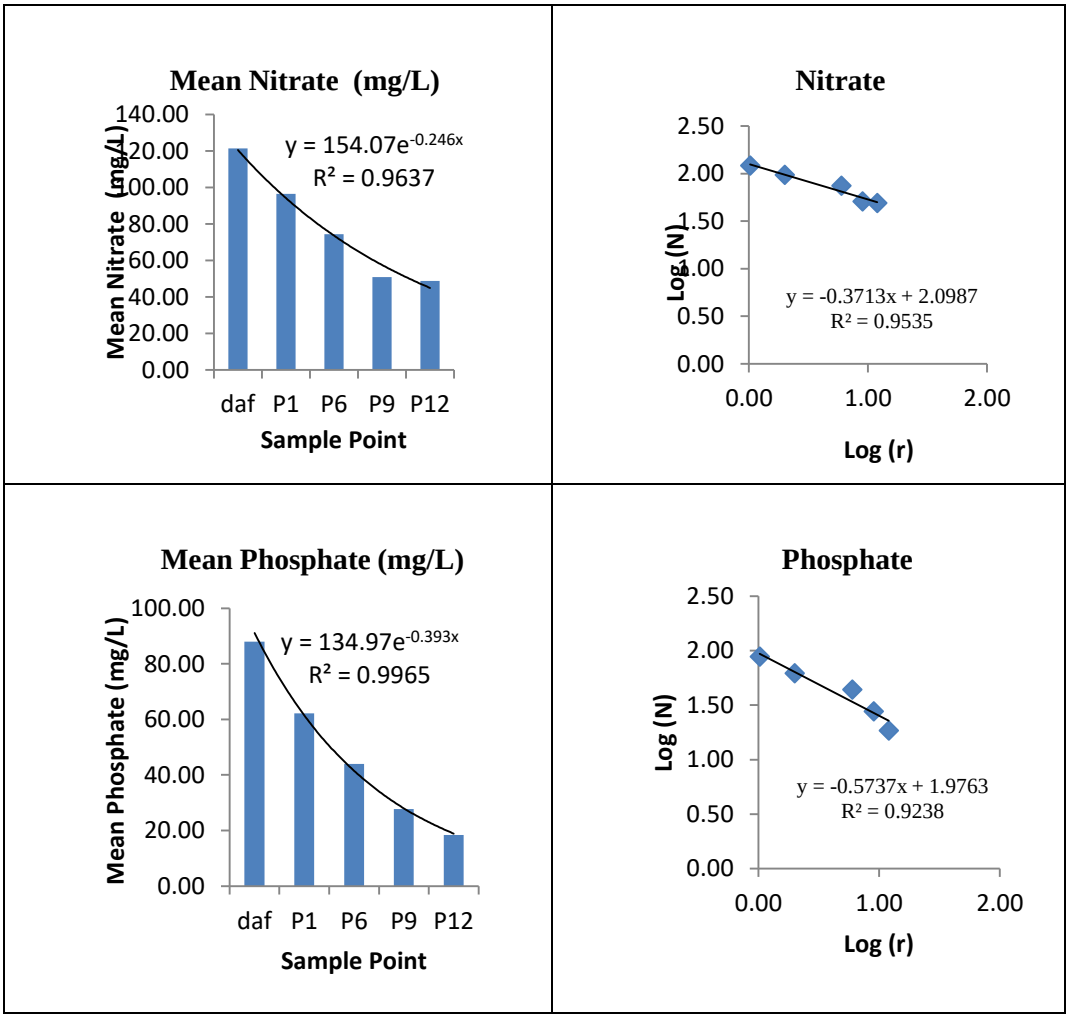


Figure 7.8: The inorganic ion trends shows the exponential distributions for the nitrate and phosphate data along with their corresponding log-log plots, where the slope of the linear equation shows the self-organising criticality (SOC) of the variable.

The inorganic ion (Ammonia, Chloride, Fluoride, Nitrite and Sulphate) trends are shown in section 7.7, with power law distributions observed, indicating strong interactions in the CW, i.e. good removal efficiencies.

7.7 Self-organised criticality (SOC) Inorganic ions graps (Ammonia, Chloride, Fluoride, Nitrite, Sulphate)

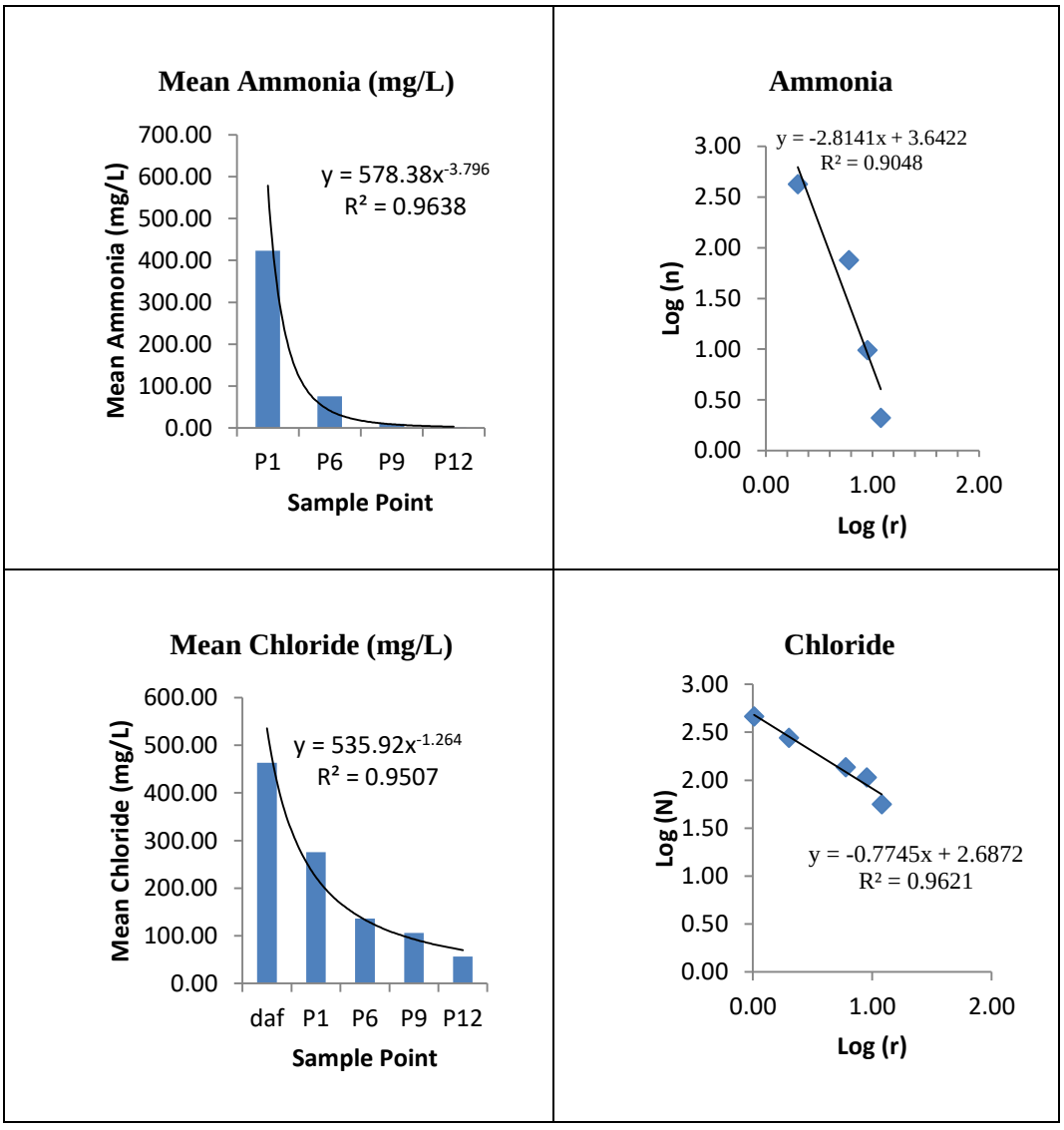


Figure 7.9: The inorganic ion trends the power law distributions for the ammonia and chloride data along with their corresponding log-log plots, where the slope of the linear equation shows the self-organising criticality (SOC) of the variable.

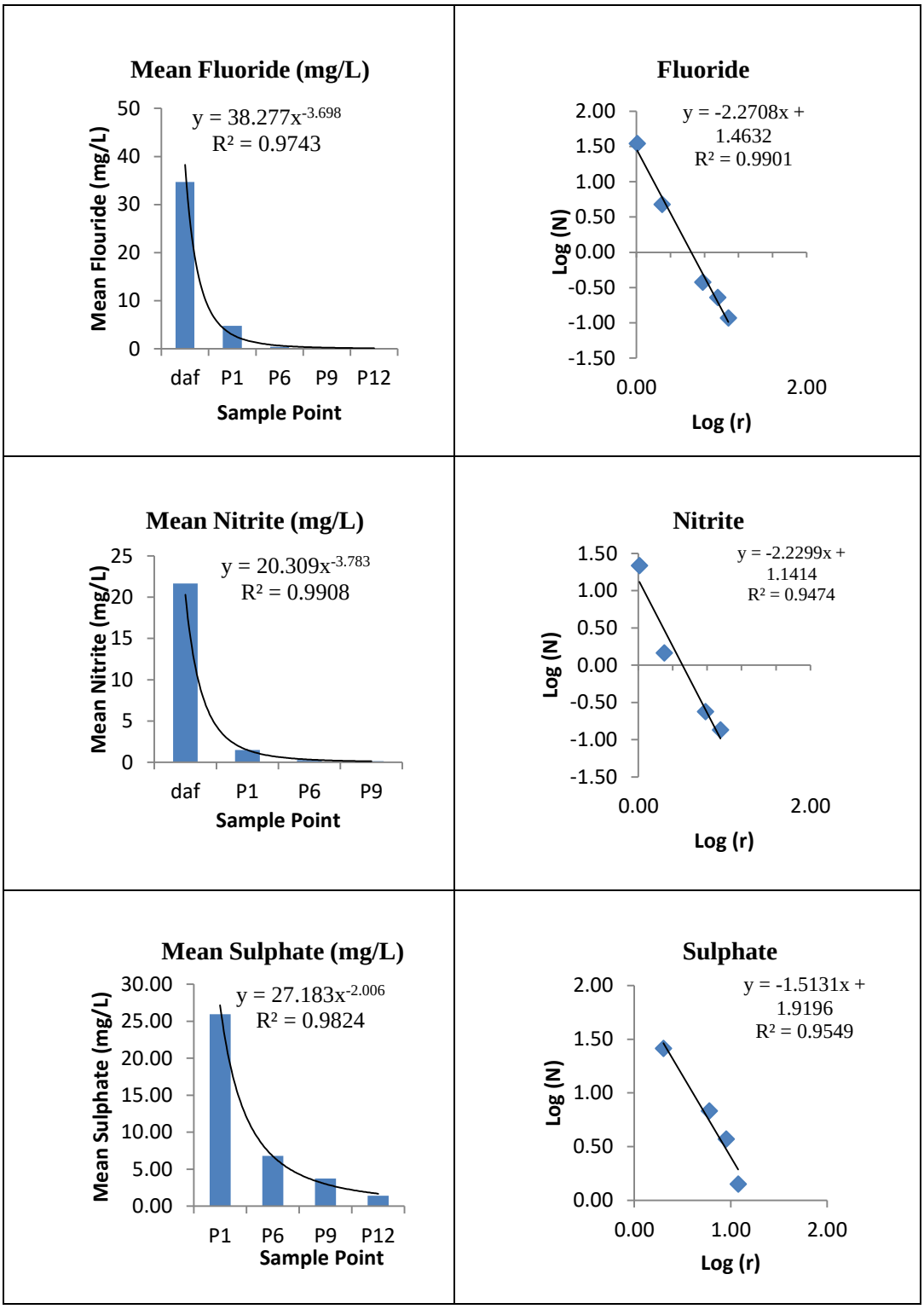


Figure 7.10: The inorganic ion trends, the power law distributions for the fluoride, nitrite and sulphate data along with their corresponding log-log plot, where the slope of the linear equation shows the self-organising criticality (SOC) of the variable.

The water depth trends are shown in section 7.8, with exponential distributions observed. No power laws are present, indicating medium interactions in the CW, i.e. the wetland has issues in maintaining an adequate water depth.

7.8 Self-organised criticality (SOC) Water Depth graphs

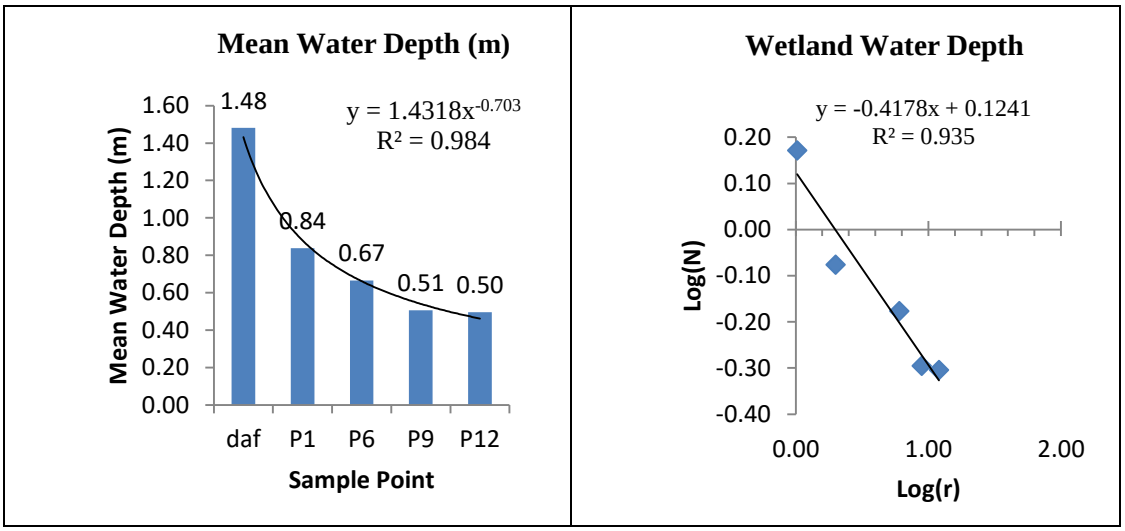


Figure 7.11: The mean wetland water depth (m) trend. Indicating a Power law distribution within the system with a corresponding log-log slope = -0.4178

Table 7.2 Pearson's Rank Correlation analysis of wetland variables

	Total Coliforms	*E.coli*	Enterococci	Total Bacteria	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	Turbidity	BOD	DO	Redox Pot	Conductivity	Water depth
Total Coliforms	1.00																
E.coli	0.99	1.00															
Enterococci	1.00	0.99	1.00														
Total Bacteria	1.00	0.99	1.00	1.00													
ammonia	0.90	0.84	0.89	0.90	1.00												
Fluoride	0.76	0.79	0.73	0.74	0.59	1.00											
chloride	0.95	0.94	0.93	0.93	0.84	0.92	1.00										
Nitrate	*0.97*	*0.96*	*0.95*	*0.96*	0.89	0.84	0.97	1.00									
Nitrite	0.73	0.76	0.70	0.70	0.54	1.00	0.90	0.81	1.00								
Phosphate	*0.98*	*0.98*	*0.97*	*0.97*	0.87	0.86	0.98	0.99	0.83	1.00							
Sulphate	0.90	0.85	0.90	0.91	1.00	0.57	0.83	0.88	0.51	0.86	1.00						
turbidity	0.85	0.86	0.82	0.83	0.75	0.98	0.98	0.92	0.96	0.93	0.73	1.00					
BOD	0.85	0.86	0.82	0.84	0.74	0.98	0.98	0.92	0.96	0.93	0.72	1.00	1.00				
DO	-0.95	-0.92	-0.94	-0.94	-0.89	-0.63	-0.85	-0.95	-0.59	-0.93	-0.89	-0.74	-0.74	1.00			
redox Pot	-0.92	-0.89	-0.90	-0.92	-0.96	-0.79	-0.95	-0.94	-0.75	-0.93	-0.95	-0.90	-0.90	0.86	1.00		
Conductivity	0.91	0.91	0.89	0.90	0.84	0.94	0.99	0.96	0.91	0.96	0.82	0.99	0.99	-0.82	-0.95	1.00	
water depth	0.84	0.86	0.81	0.82	0.66	0.99	0.95	0.91	0.98	0.92	0.64	0.98	0.98	-0.75	-0.82	0.95	1.00

SPSS Version 22.0 correlation table ($p < 0.01$).

*Note: The indicator bacteria (total coliforms, *E.coli*, enterococci and total bacteria) all have a strong correlation to both nitrate and phosphate variables.

There were strong Pearson rank correlations ($p < 0.01$) between the bacterial concentrations and nitrate and phosphate values as can be seen on Table 7.2. It should be noted that both nitrate and phosphate best-fit an Exponential distribution not Power law distributions, and all the indicator bacteria values displayed linear distributions (see Table 7.3)

Table 7.3: Wetland variable distribution (linear, exponential, power law)

Ranking	Variable	Efficiency	Distribution type	Distribution Equation	Goodness of Fit	log-log Slope (SOC)
inefficient	Nitrate	0.60	exp	$y = 154.07e^{-0.246x}$	$R^2 = 0.9637$	-0.371
inefficient	Total Bacteria	0.66	linear	$y = -1.536x + 10.969$	$R^2 = 0.9915$	N/A
inefficient	**water depth	0.67	power	$y = 1.4318x^{-0.703}$	$R^2 = 0.984$	-0.418
inefficient	Total Coliforms	0.68	linear	$y = -1.5714x + 10.741$	$R^2 = 0.9957$	N/A
inefficient	Enterococci	0.69	linear	$y = -1.3289x + 9.2193$	$R^2 = 0.9899$	N/A
inefficient	*E.coli*	0.74	linear	$y = -1.5042x + 9.9697$	$R^2 = 0.9898$	N/A
inefficient	Phosphate	0.79	exp	$y = 134.97e^{-0.393x}$	$R^2 = 0.9965$	-0.574
efficient	Conductivity	0.84	power	$y = 4.9038x^{-1.21}$	$R^2 = 0.9597$	-0.75
efficient	Redox potential	0.86	poly-nominal	$y = -14.827x^3 + 111.85x^2 - 137.28x - 222.58$	$R^2 = 0.9338$	N/A
efficient	Chloride	0.88	power	$y = 535.92x^{-1.264}$	$R^2 = 0.9507$	-0.774
efficient	Sulphate	0.93	power	$y = 27.183x^{-2.006}$	$R^2 = 0.9824$	-1.151
efficient	Dissolved Oxygen	0.93	power	$y = 0.3226x^{1.8593}$	$R^2 = 0.9521$	-1.151
very efficient	Turbidity	0.95	power	$y = 989.3x^{-2.073}$	$R^2 = 0.9796$	-1.29
very efficient	Ammonia	0.99	power	$y = 578.38x^{-3.796}$	$R^2 = 0.9638$	-2.814
very efficient	Fluoride	0.99	power	$y = 38.277x^{-3.698}$	$R^2 = 0.9743$	-2.27
very efficient	Nitrite	0.99	power	$y = 20.309x^{-3.783}$	$R^2 = 0.9908$	-2.229
very efficient	BOD	0.99	power	$y = 6492.4x^{-3.699}$	$R^2 = 0.9731$	-2.3

Table 7.3 shows the descending ordering of the wetland variables from inefficient – efficient – very efficient [0.6 to 0.83 – 0.84 to 0.93 – 0.94 to 0.99] respectively, starting with nitrate and ending with BOD. Generally, exponential and linear regression distributions exhibit low efficiencies and the remaining power distribution are efficient or very efficient. Redox potential displays a polynomial distribution but has a 0.86 efficiency removal. Therefore, based on the previous variables above and below redox potential was efficient in ranking.

Note: N/A denotes not applicable with regard to the indicator bacteria and redox potential values due the fact these values cannot be derived from linear and polynomial equation

7.9 Results - comparison of Method 1 and Method 2

The following section compares and contrasts both Methods 1 and 2, where Method 1 is based on soft-scaled and the Normalised (S) distribution entropy and Method 2 is based on the load based and the Normalised (S) distribution entropy of the constructed wetland.

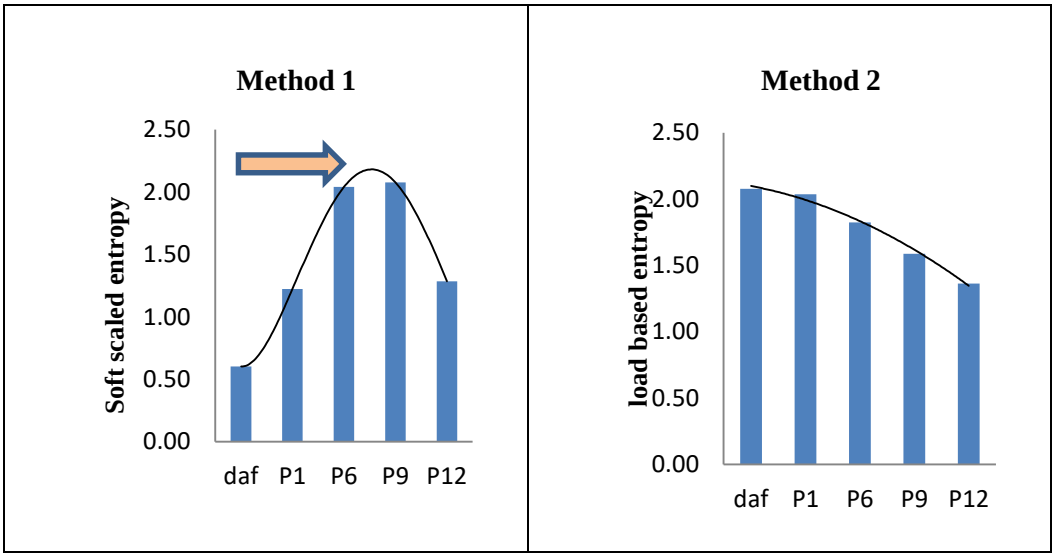


Figure 7.12: Wetland indicator bacteria Method 1 versus Method 2. Method 1 entropy shows the critical transition regions are located between Pond 6 - Pond 9 i.e. the most stable/efficient regions; notice after pond 9 the SCE a large dip indicating instability or disorder with respect to pond 12. Method 2 (load based entropy) shows the linear distribution of the bacteria through the wetland. No overloading of the bacteria within the wetland system.

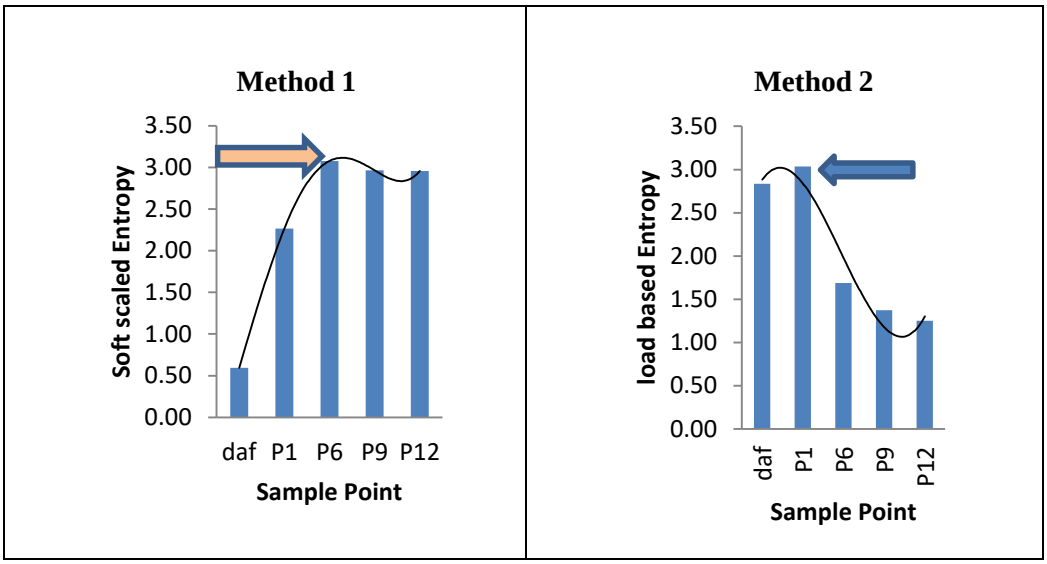


Figure 7.13: Wetland physical-chemical Method 1 vs. Method 2. Method 1 entropy shows the critical transition region was pond 6 i.e. the most stable/efficient region; notice after pond 9 the SCE a small dip indicating instability or disorder with respect to pond 9 and pond 12. Method 2 (load based entropy) shows the polynomial distribution of the physical-chemical variables through the wetland. Highlighting that pond 1 records a large increase in load based entropy. The DAF is potentially overloading pond 1, with respect to physical chemical parameters.

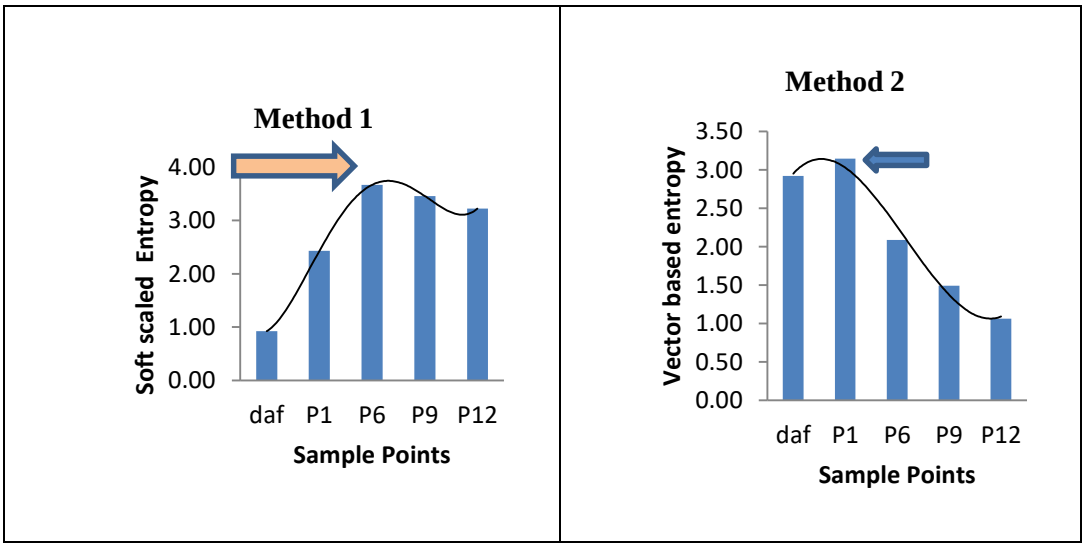


Figure 7.14: Wetland inorganic ions Method 1 vs. Method. Method 1 entropy shows the critical transition region was pond 6 i.e. the most stable/efficient

region; notice after pond 9 the SCE a small dip indicating instability or disorder with respect to pond 9 and pond 12. Method 2 (load based entropy) shows the polynomial distribution of the inorganic ions variables through the wetland. Highlighting that pond 1 records a large increase in load based entropy. The DAF is potentially overloading pond 1, with respect to inorganic ion parameters.

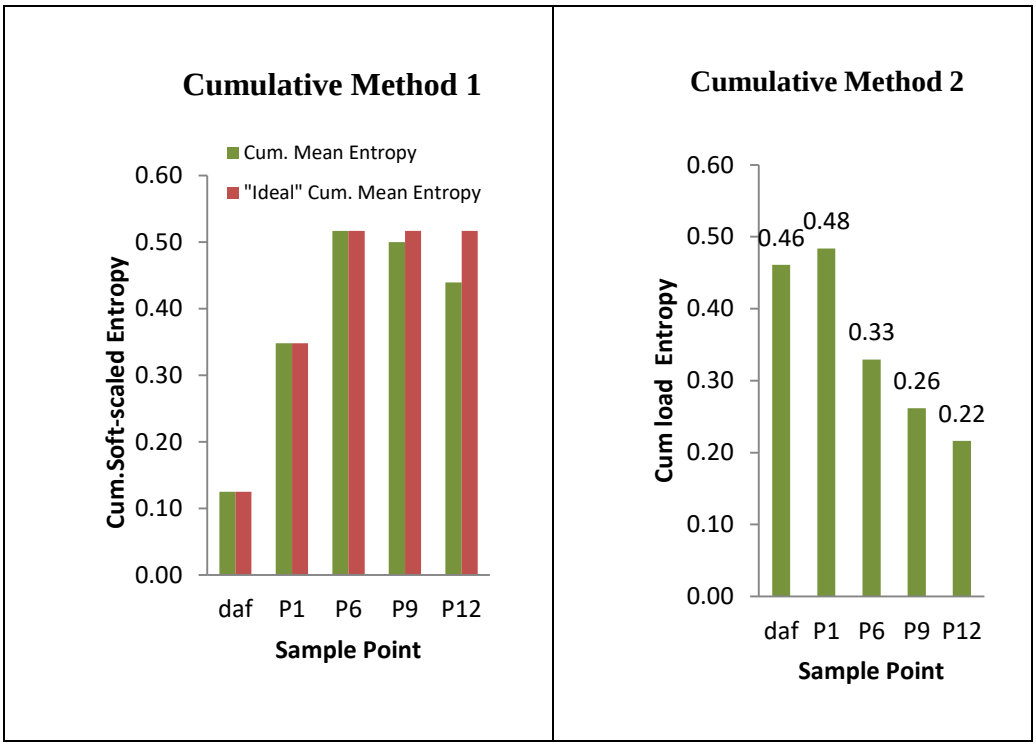


Figure 7.15: Wetland cumulative Method 1 vs. Method 2. Shows the Method 1 soft-scaled entropy and Method 2 load based entropy for cumulative wetland variables (all the wetland variables). Method 1 entropy shows the most stable point was at pond 6 i.e. critical transition, from disorder to order. After pond 6 instability or disorder becomes dominant with respect to pond 9 and pond 12. An ideal data set shows how the wetland should function with no instabilities. Method 2 Entropy, shows that the DAF is overloading pond 1 and between pond 1 to pond 12.

7.10 Discussion

7.10.1 A review of SOC analysis

The indicator bacteria show a linear trend, no power law distribution was observed, therefore each pond was ineffective in treating the bacteria, as a consequence the entire system effect of multiple ponds in series has poor overall efficiency (see Figure 7.4). The inorganic ions reveal exponential distributions (nitrate and phosphate) and the remaining ions (ammonia, nitrite, fluoride, sulphate, chloride) reveal power law distributions, with regard to these ions, (SOC) was observed with good efficiencies, therefore the wetland system can be considered a complex system with self-tuning abilities, (see Table 7.3). The regression equations for each wetland variable and their corresponding efficiencies and inefficiencies can be seen in Table 7.3. The inefficient variables (total bacteria, *E.coli*, enterococci, water depth, nitrate and phosphate) display linear, exponential and polynominal equations, the exception is the water depth variable which does have a power law distribution, the wetland is efficient at losing water. These inefficient variables indicate the wetland requires external tuning, in comparison to the remaining variables which displayed non-linear dynamic power law distributions therefore these variables require no external tuning i.e. complex wetland system. The Pearson's rank correlations analysis re-inforce this behaviour as the indicator bacteria and the nitrate and phosphate ions display strong positive correlations, (see Table 7.2).

7.10.2 A review of entropy analysis:

In Figure 7.15 Method 1 entropy, shows an ideal data trend versus the actual cumulative mean entropy data trend (the soft-scaled entropy of all seventeen wetland variables) of the wetland system, where there was a "dip" or instability (disorder) present in the back-end (pond 9 to pond 12) of the wetland. With regard to Method 2, the cumulative load based entropy indicates that pond 1 was being over-loaded from the DAF plant with values of 0.46 and 0.48 respectively. These exact patterns are seen in Figures 7.13 and 7.14 inorganic ions and physical-

chemical variables respectively. The bacteria analysis shown in Figure 7.12 reveals a stronger instability (dip) within pond 12, indicating poor removal of indicator bacteria in this pond.

The one outlier was the water depth variable, which displayed a power law distribution with a log-log slope of -0.418. The water depth within the wetland was efficiently lost within the wetland system or the water was inefficiently retained within the wetland system. The water depth efficiency was evaluated with a mean wetland water depth retention efficiency of 0.67 or 67%. The overall average efficiency of the wetland system across these seventeen key wetland variables was 0.83 or 83% i.e. the entire wetland system was on the cusp of inefficient to efficient.

When reviewing Method 2 (LBE) data, see Appendix E (i) and E (iii) for the indicator bacteria and inorganic variables respectively. Highlighted in red are entropy bits > 2.00, for indicator bacteria entropy values are 2.23, 2.20, 2.22, and 2.24, for nitrate and phosphate 2.23 and 2.13 respectively, which are treated inefficiently within the wetland, see Table 7.3. Appendix E (ii) shows that the water depth has a value of 2.19. The Method 2 (LBE) therefore indicates that these variables are over-loading the wetland and the complex wetland system cannot reduce these variables. There is a potential that linear and exponential trends and the Method 2 (LBE) with total entropy bits > 2.0 are similar.

These repeating instabilities as seen in the indicator bacteria, physical-chemical variables and inorganic ions are potentially alluding to "overdrive" in the wetland system; where overdrive pertains to processes occurring too quickly within a complex system, therefore inducing destabilising effects; in relation to hydrocarbon production in tectonic plate geometry in the Earth's crust (Crampin, 2000) states that these are only effective when slowly driven. If production is over-driven, so that natural relaxation processes do not apply, the effects are likely to be unpredictable, and Woodward *et al.,* (2007) review of correlation dynamics in self- organising systems state that temporal correlations are found… so long as the system is not overdriven and overdriving the system near self-organised criticality is shown to have a destabilising effect on the SOC state……. has serious implications for the functioning of complex systems (Milovanov, 2011). From a SOC view, the presence of linear and exponential distributions, in relation to indicator bacteria, nitrate and phosphate loadings respectively are not

effectively removed as these variables pass through the wetland system. Potential causes are (1) overloading, (2) very short contact time within the ponds of the wetland (short circuiting) and (3) long-circuiting very long contact time within the wetland, the ponds are isolated from one another.

7.11 Conclusions

In this chapter, scale-free power-law behaviours are found to govern eleven of the key wetland variables (BOD, nitrite, fluoride, ammonia, turbidity, dissolved oxygen, sulphate, chloride, redox potential, conductivity and water depth), and display either sub-criticality (log-log slope -0.75 to -1) or criticality (log-log slope $\geq$ -1) in conjunction with power-law distributions. The remaining variables, such as the indicator bacteria (total coliforms, *E.coli*, enterococci and total bacteria), nitrate and phosphate display linear and exponential distributions respectively, with low log-log slope values $\leq$ -0.7, see Table 7.3. Using Pearson's rank correlation ($p < 0.01$) we can see that the indicator bacteria and the nitrate and phosphate variables share strong correlations to each other with values ≥ 0.95, indicating a synergy between the presence of indicator bacteria and wetland nutrients nitrate and phosphate. The low efficiencies of these variables should be noted, between 0.79 – 0.6.

The findings suggest that several key wetland variables within the constructed wetland resemble self-organising criticality and these variables display global efficiency i.e. efficient removal within a complex wetland system. Where these variables experience avalanches within the wetland system, i.e. the wetland was self-tuning (no external impetus required) but other variables such as the indicator bacteria display linear distributions, see Figure 7.4. Nitrate and phosphate displaying exponential distributions, indicating an ordered and predictable decay, see Figure 7.8, within the wetland, i.e. not complex (local inefficient removals). Therefore, not self-tuning and required external impetus to increase the removal efficiency. The nitrate and phosphate variables display exponential trends, with low log-log slopes which were -0.371 and -0.574, with efficiencies of 60% and 79% respectively.

The entropy analysis highlighted the transitions within the wetland system regarding the indicator bacteria, physical-chemical and inorganic ions, with bacteria showing a noticeable decline towards "instability" towards the back-end of the wetland (ponds 9 and 12), with similar trends observed in both the physical-chemical and inorganic variables but not to the same extent as the indicator bacteria. The overall cumulative effect of the entire wetland variable as seen in Figure 7.15 shows these effects (Method 1) along with the seventeen wetland variables cumulative distribution within the wetlands sample points. Reviewing Method 2, based on the load based entropy, this analysis indicates that the DAF was overloading pond 1. Combining SOC and Entropy shows new insights into the workings of the wetland variables, highlighting what variables are efficiently removed and showing that in this case the wetland has issues in treating/remediating the indicator bacteria and reducing nitrate and phosphate loads. The outlier in the system was the mean water depth showing the wetland displays a power law distribution regarding this variable. The water within the CW was efficiently lost within the wetland system. Both the self-organising criticality (SOC) and the Method 2 (LBE) display similar behaviours with respect to wetland variables. The SOC of wetland variables displaying linear or exponential trends and the LBE with entropy values > 2.0, both express the wetlands' bacteria, nitrate and phosphate and water depth as problematic i.e. the constructed wetland has issues in reducing the indicator bacteria, nitrate and phosphate concentrations and maintaining an optimal water depth, see Appendix E and Table 7.3.

These inefficient variables are potentially (1) over-driven within the system, with hydraulic retention times (HRTs) one potential cause for this effect i.e. the abattoir wastewaters are moving too quickly through the wetland system (short-circuiting), therefore their contact time needs to be increased and conversely the abattoir wastewaters are moving too slowly through the wetland system (long-circuiting) and (2) over-loading the wetland system (hydraulic loading rate) in particular the DAF plant was not treating the abattoir wastewater effectively and releasing poorly treated waste effluent into pond 1, therefore introducing a cascade effect on the entire wetland from pond 1 to pond 12.

Chapter 8. Section 1: The application of artificial neural networks (ANNs) to understand wetland behaviours.

8.1 Summary

Three artificial neural network (ANN) models based on the free water surface (FWS) constructed wetland variables are developed. The analysis classifies the constructed wetlands variables in relation to their interaction with the indicator bacteria (total coliforms, *E.coli*, enterococci and total bacteria concentrations). Three different ANN models were analysed, with varying hidden layers within the three models, to ascertain different relative importance and sensitivities with respect to the wetland indicator bacteria.

8.2 Introduction

Artificial neural networks (ANN) attempt to model the way the brain works, which leads to their appeal, in that they are potentially capable of making decisions as humans, based on complex, noisy and partial information. Artificial neural networks have been found to be powerful methods such as memorising, pattern recognition and analysing non-linear multivariate data. The ANN's have been successfully applied to areas of spectroscopy (Mouwen *et al*., 2006), psychological research (Pourshahriar, 2012), an environmental decision support system (Liao *et al*., 2012) and future energy policy (Sözen, 2009). In previous constructed wetland studies, the dominance of wetland variables such as BOD, COD, ammonia, phosphate and nitrate removal using artificial neural networks are evident (Akratos *et al*., 2009a, Akratos *et al*., 2008 and Akratos *et al*., 2009b). Comparison studies of horizontal subsurface flow (HSF) versus free water surface (FWS) flow constructed wetlands was carried out to determine which system was most effective in treating the above wetland variables. Uyanik *et al*., (2009) used ANN models in a side-by-side comparison of HSF and FWS systems to determine which systems were more applicable to their needs with the HSF providing better

removal efficiencies than the FWS. Chang *et a*l., (2015) compared ANN and genetic programming (GP) algorithms in the use of constructed wetlands treating storm waters and found that ANN performed better than the GP model.

The removal of enteric bacteria from domestic and municipal sewage using different constructed wetlands was described out by Vymazal, (2005), where the author validated the efficiency of sixty CWs treating enteric bacteria such as faecal coliforms and faecal streptococci. Vymazal (2005), concluded that hydraulic retention time (HRT) and the hydraulic loading rate (HLR) in conjunction with emergent macrophytes are important in reducing enteric bacteria populations. It was shown in chapter 7 that indicator bacteria within the wetland are not adequately treated, with overloading (HLR) and overdriving (HRT) as potential causes within the CW. Therefore, performing advanced analysis of the indicator bacteria using ANN could provide added information on potential causes of this observation. Mittal, (2006) undertook a review of wastewaters from abattoirs before land application, whereby different biological processes were assessed (biomechanical and wetlands). The review included pathogen removal by various wastewater treatment process based on Ontario (Canadian) data, which recommends that when using stabilisation ponds, three ponds in series with more than a 25-day hydraulic retention time are required. Rivera *et al*., (1997) reviewed a two-stage system to treat high-strength wastewater from an abattoir in Pachua, Mexico, consisting of an anaerobic digester followed by a horizontal subsurface flow constructed wetland. Analysing the physical-chemical, inorganic ions and bacterial pathogens (total and faecal coliforms, *Salmonella*, *Shigella and Vibrio cholerae*), the authors conclude that Despite the high pathogen removal rates of the Pachua system, the final effluent is clearly above the WHO (1989) guideline value (1,000 faecal coliforms per 100ml) for unrestricted irrigation. Therefore, understanding the potential interactions (sensitivities and importance), by using ANNs may result in understanding the issues impeding the constructed wetlands efficiency.

8.3 A constructed wetland, an Irish context

Scholz *et al*., (2009) reviewed the effect of denitrifying, ammonia-oxidizing bacteria along with total coliforms and *E.coli* concentrations in the Annestown

stream project, using constructed wetlands to remediate farmyard dirty water in Waterford, Ireland; the monitoring of BOD, COD, pH, nitrate, phosphate and ammonia were also included in the study.

The Environmental Protection Agency (EPA) in Ireland reviewed, the Best Available Technologies (BAT) for the disposal of animal carcasses and animal waste (EPA, 2008). However, no reference exists in the document pertinent to the recording of bacterial concentrations for air, water and land.

In one site in Co. Meath, the construction of an ICW was proposed to treat abattoir wastewater including a wastewater treatment facility on site (EPA Ireland, 2009). The variables to be monitored include physical-chemical and inorganic ions, but no reference to any bacterial monitoring. The STRIVE Report N0.45 by (Zhan *et al.*, 2010) reviews novel technologies to treat slaughter waste and classifies the waste from abattoirs as industrial and that constructed wetland are a tertiary (polishing) treatment option. Therefore, critical information regarding the performance of a large scale CW treating slaughter waste and the interactions of climate, physical-chemical, inorganic ions, wetland design and indicator bacteria would be of benefit to understand these complex interactions, especially the role the bacteria play as they migrate through the wetland system. ANN analysis would be of benefit in understanding these interactions and in particular specific complexities that relate to bacterial loadings in the abattoir wastewaters.

This chapter presents the results of exploratory analysis performed in order to evaluate the feasibility of building ANN models to identify the main wetland variables that display importance and sensitivity to indicator bacterial concentrations (total coliforms, *E.coli* and enterococci) within the free water constructed wetland treating abattoir waste.

8.4 Wetland data

In total 27 input variables (wetland design, physical-chemical, inorganic ion and climate values) and 4 output variables representing the total coliforms, *E.coli* and enterococci and total bacterial (Log_{10} CFU/100ml) were used to create the ANNs.

8.5 Artificial neural network analysis

In this section, the wetland data was modelled with artificial neural networks (ANNs) to evaluate the importance and sensitivities of wetland variables to wetland indicator bacteria. ANN is considered a classification method i.e. using a set of parameters/variables to characterise a task, in this case the ANN classification can be sub-divided into (1) training task to configure the dataset (2) validation task to optimise the dataset and (3) testing task to evaluate the performance of the model under scrutiny. This type of classification methodology mimics the behaviour of human brain neurons, based on real world interactions i.e. non-parametric. Van der Wiele *et al*., (2009) used advanced network analysis, both SOMs (Self-Organising Maps) which are unsupervised and a supervised MLP (multilayer perceptron) to analyse land use based on satellite imagery.

A multivariate study was employed for the ANN analysis. The variables employed are called out in the data section and the importance of these variables was carried out using software Easy-NN Plus software version 14.0d www.easynn.com [Accessed 04 2014]. The network software uses a back propagation algorithm and sigmoidal functions for the adjustments of the networks weights and biases values. The ANN uses three layers: an input layer that consists of the passive nodes (variables), one hidden layer consisting of active nodes and an output layer, that has active output nodes, see Figure 8.1. To train the ANN, a set of data containing input and output nodes are actioned. Once training and validation were finalised, the ANN was capable of predicting the output when in response to an input similar to the pattern that it has learned. The ANN was tested for the remaining set of wetland data. The learning rate and momentum value of the network was set to optimise with a target error value of 0.05. The learning rate is the training parameter that controls the weight and bias changes during learning. The momentum parameter is used to prevent the system from converging to a local minimum or 'saddle point'. A high momentum parameter can also help to increase the speed of convergence of the system. However, setting the momentum parameter too high can create a risk of overshooting the minimum, which can cause the system to become unstable. A

momentum coefficient that is too low cannot reliably avoid local minima, and can also slow down the training of the system.

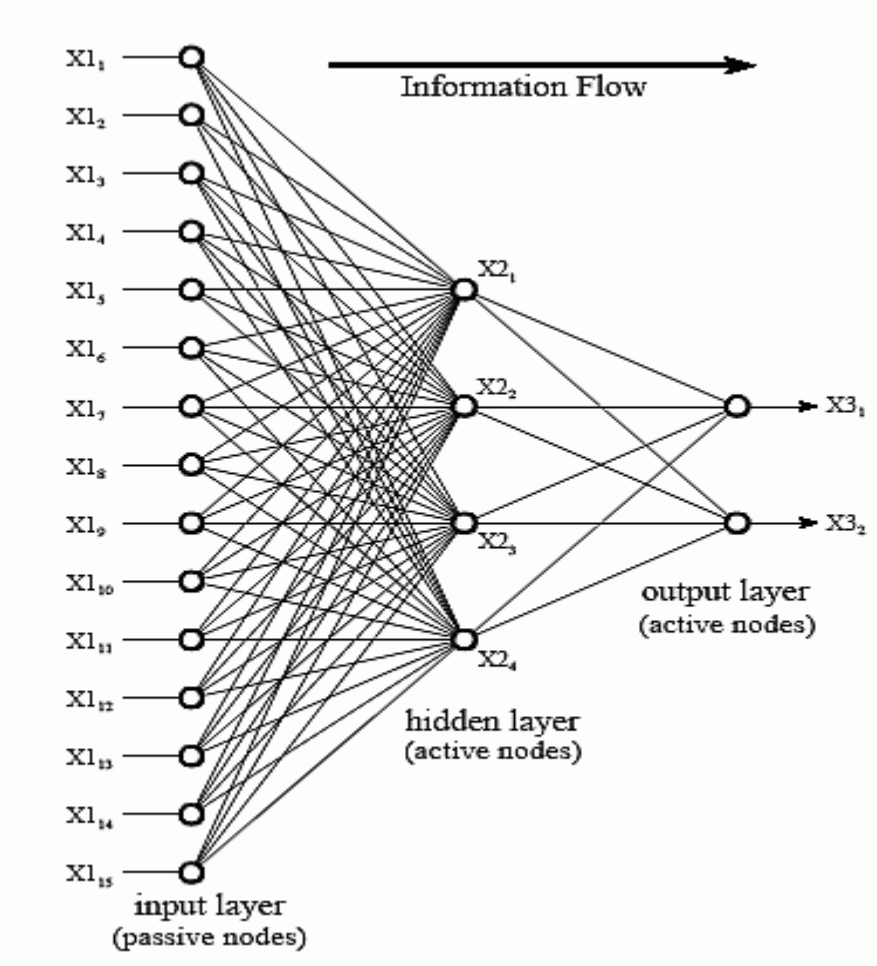


Figure 8.1: The artificial neural network architecture: the figure shows the most common structure for neural networks: three layers with full interconnection. The input layer nodes are passive, doing nothing but relaying the values from their single input to their multiple outputs. In comparison, the nodes of hidden and output layers are active, modifying the signals. The action of the neural network was determined by the weights applied in the hidden and output nodes.

8.6 ANN Modelling

Three models were processed, testing different ANN configurations. All the ANN models have 27 input passive nodes (input variables) and 4 output active nodes (indicator bacteria) with a target error = 0.05 They represent the following characteristics:

ANN Model 1: one hidden layer with 12 nodes, setting the learning rate and momentum to optimise, which produced values of 0.6 and 0.7 respectively. The model ran for 12,724 cycles with 47 training examples and 19 validating examples.

ANN Model 2: two hidden layers with 12 nodes for layer 1 and 7 nodes for layer 2, setting the learning rate and momentum to optimise, which produced values of 0.7 and 0.8 respectively. The model ran for 16,724 cycles with 42 training examples and 24 validating examples.

ANN Model 3: three hidden layers with 12 nodes for layer 1, 8 nodes for layer 2 and 9 nodes for layer 3, setting the learning rate and momentum to optimise, which produced values of 0.7 and 0.8 respectively. The model ran for 11,369 cycles with 48 training examples and 20 validating examples.

An important point to note was the user's visual interpretation of the results. The ANN model not only trains, validates and tests the data but also any inherent noise that may be present, which can produce over-fitting (de Souza and Giunta, 2011). We adopted their approach of early stopping, identifying when the model begins to lose the 'capacity of generalisation' i.e. the user can visualise the exact moment when the target error, in this case 0.05 stops decreasing and begins to increase and therefore the ANN model was halted at this point.

8.7 ANN data and results

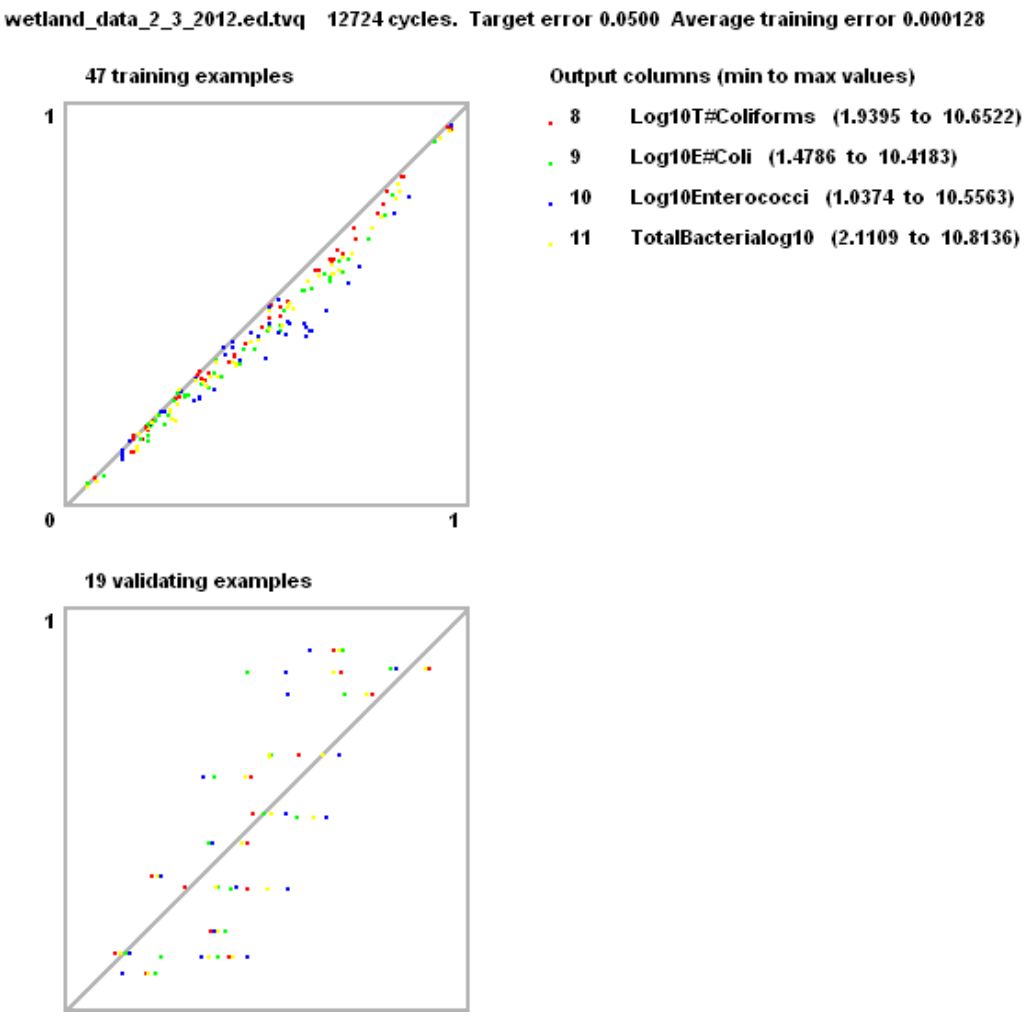


Figure 8.2: The ANN training and validating examples for Model 1, with the output nodes depicting the bacteria under analysis and their minimum to maximum $\log_{10}$ values. The target error was set to 0.05 and the resultant average training error was 0.000128 with 47 training and 19 validating examples.

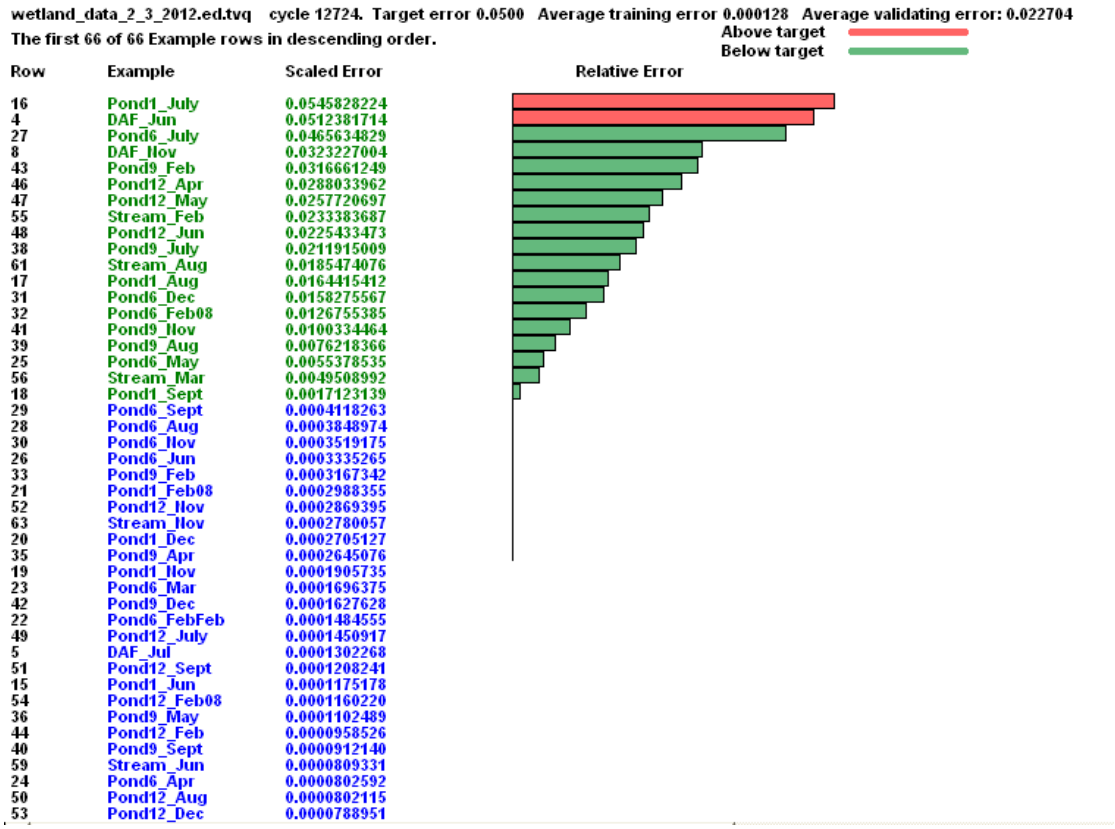


Figure 8.3: The relative error by sample point by month Model 1, indicating that Pond1_July and DAF_June are above the target error = 0.05. This indicates that potential issues existed at these samples points at these months.

wetland_data_2_3_2012.ed.tvq 12724 cycles. Target error 0.0500 Average training error 0.000128

The first 27 of 27 Inputs in descending order.

Column	Input Name	Importance
21	Ammonia	36.5549
1	RedoxPot	36.4920
5	Dissolvedoxygen	35.0530
14	wetlandvolume	33.6222
16	chloride	30.8424
17	Nitrate	29.8599
13	wetlandarea	27.5000
27	RainfallAmount	27.4131
19	Phosphate	27.3234
7	Turbidity	26.5041
4	Conductivity	26.4452
30	Penman	26.1199
0	pH	25.0580
12	WaterDepth	23.7440
6	BOD	23.2601
20	Sulphate	23.1047
15	Flouride	22.1281
29	SolarRadiation	20.4458
2	WaterTemp	19.8071
18	Nitrite	18.5625
23	MinairTemp	17.9141
26	SoilTemperature_300mm	17.8888
28	P#E	17.2599
3	AirTemp	16.4485
22	MaxairTemp	13.9397
24	SoilTemperature_100mm	13.5548
25	SoilTemperature_200mm	13.0658

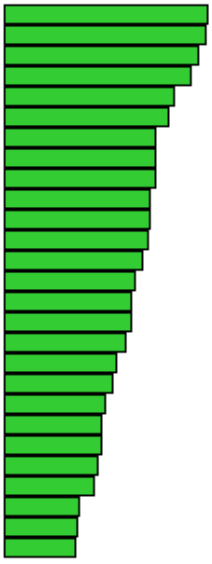


Figure 8.4: The relative importance of wetland variables Model 1, with ammonia, redox potential and dissolved oxygen with respect to Model 1.

wetland_data_2_3_2012.ed.tvq 12724 cycles. Target error 0.0500 Average training error 0.000128

The first 27 of 27 Inputs in descending order. Output column 8 Log10T#Coliforms

Column	Input Name	Change from	to	Sensitivity
4	Conductivity	0.2000	7.8100	0.126642760
28	P#E	14.1000	107.0000	0.083597665
16	chloride	1	819	0.076252529
6	BOD	1.4000	4350.0000	0.075855695
27	RainfallAmount	25.1714	152.0714	0.072671137
0	pH	6.4000	9.5500	0.068357749
29	SolarRadiation	63.2300	556.9650	0.062822963
5	Dissolvedoxygen	0.1400	13.8000	0.062393332
23	MinairTemp	2.8143	11.5429	0.061304632
30	Penman	4.2500	94.2500	0.049894493
7	Turbidity	0.0000	1567.0000	0.048768646
15	Flouride	0	88	0.033022910
20	Sulphate	0	53	0.031011878
22	MaxairTemp	9.7143	18.8000	0.029375980
21	Ammonia	0	477	0.028522819
17	Nitrate	0	241	0.027106951
19	Phosphate	0	106	0.022859330
12	WaterDepth	0.0000	1.5000	0.017902488
25	SoilTemperature_200mm	5.7286	15.9429	0.014510311
24	SoilTemperature_100mm	4.8857	15.7571	0.012417228
14	wetlandvolume	0	2671	0.009578779
26	SoilTemperature_300mm	6.3000	16.6571	0.008128535
1	RedoxPot	-319.0000	85.0000	0.004565740
13	wetlandarea	1804	2813	0.004056182
18	Nitrite	0	51	0.003458698
3	AirTemp	3.2000	24.8000	0.003042519
2	WaterTemp	3.8000	26.2000	0.002636047

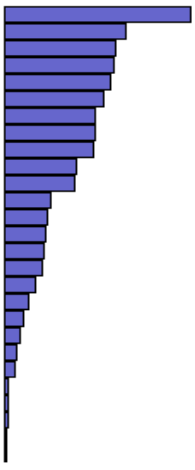


Figure 8.5: The relative sensitivity of wetland variables Model 1, with conductivity, Potential Evapotranspiration (P.E.), chloride and BOD dominant.

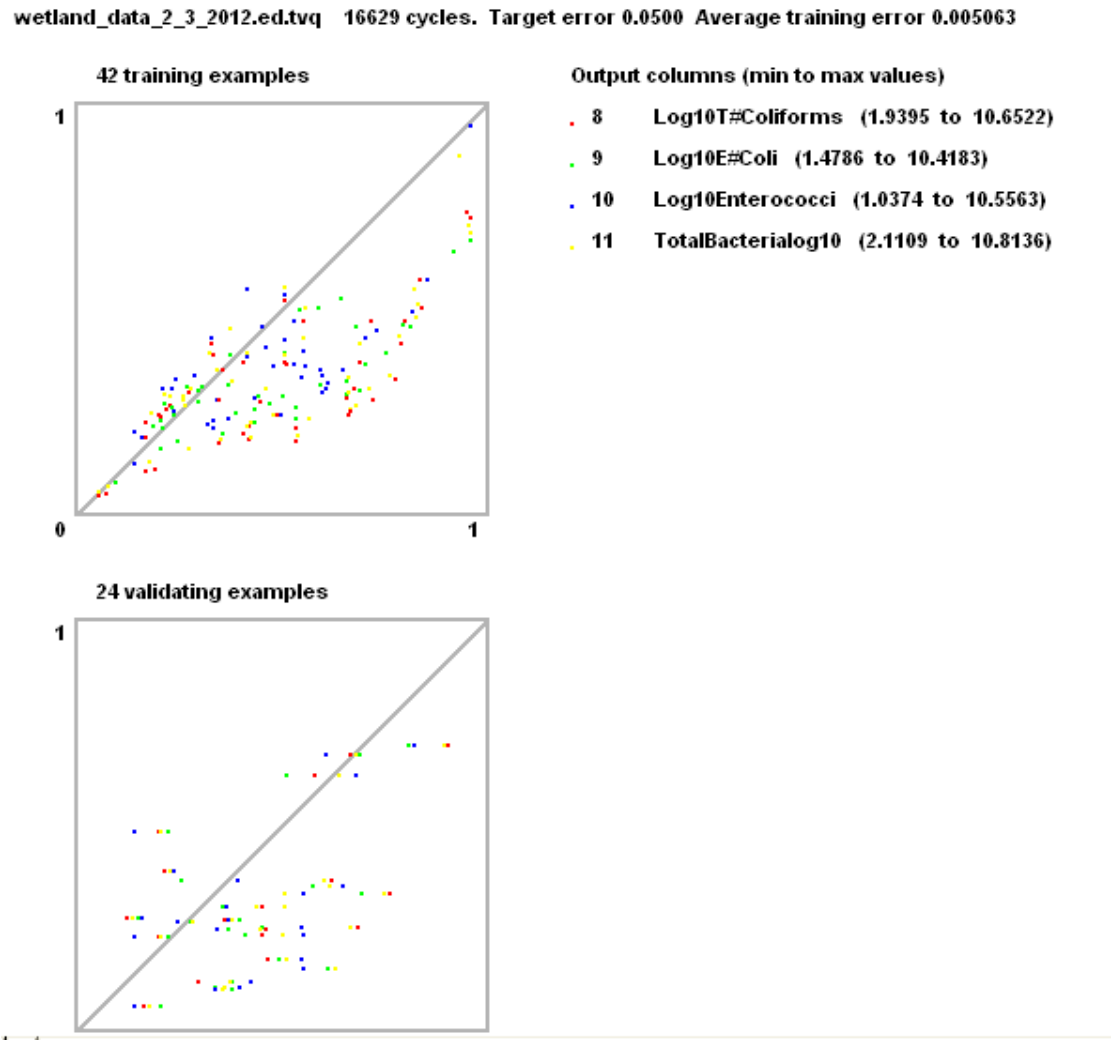


Figure 8.6: The ANN training and validating examples for Model 2, with the output nodes depicting the bacteria under analysis and their minimum to maximum $\log_{10}$ values. The target error was set to 0.05 and the resultant average training error was 0.005063, with 42 training and 24 validating examples.

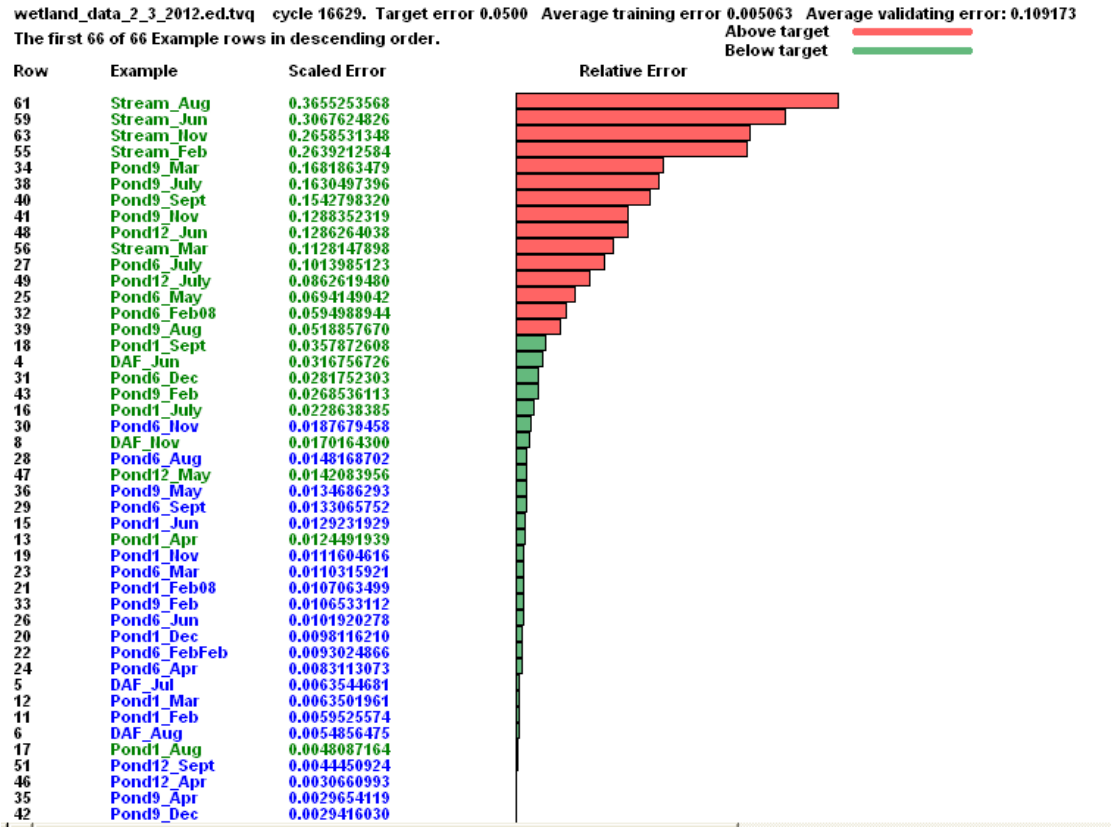


Figure 8.7: The relative error by sample point by month Model 2, indicating that stream, pond 9, pond 12 and pond 6 above the target error = 0.05. This indicates that potential issues existed at these samples points at these months.

wetland_data_2_3_2012.ed.tvq 16629 cycles. Target error 0.0500 Average training error 0.005063

The first 27 of 27 Inputs in descending order.

Column	Input Name	Importance
13	wetlandarea	47.4883
0	pH	41.5619
14	wetlandvolume	41.5287
1	RedoxPot	38.7532
5	Dissolvedoxygen	38.5107
7	Turbidity	35.3873
19	Phosphate	34.1293
27	RainfallAmount	33.6537
17	Nitrate	32.5337
6	BOD	32.0358
3	AirTemp	31.4405
21	Ammonia	30.8966
26	SoilTemperature_300mm	30.6210
23	MinairTemp	30.2599
24	SoilTemperature_100mm	29.0569
25	SoilTemperature_200mm	28.3999
20	Sulphate	28.3902
4	Conductivity	28.2088
29	SolarRadiation	27.8691
12	WaterDepth	27.1463
28	P#E	25.1073
15	Flouride	22.1068
18	Nitrite	21.7503
16	chloride	20.7612
30	Penman	19.9722
22	MaxairTemp	19.8249
2	WaterTemp	14.8697

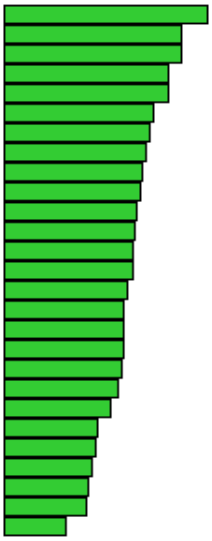


Figure 8.8: The relative importance of wetland variables Model 2, with wetland area, pH and wetland volume with respect to Model 2.

wetland_data_2_3_2012.ed.tvq 16629 cycles. Target error 0.0500 Average training error 0.005063

The first 27 of 27 Inputs in descending order. Output column 8 Log10T#Coliforms

Column	Input Name	Change from	to	Sensitivity
6	BOD	1.4000	4350.0000	0.181022415
14	wetlandvolume	0	2671	0.162545706
17	Nitrate	0	241	0.153183713
19	Phosphate	0	106	0.145881255
16	chloride	1	819	0.143267455
4	Conductivity	0.2000	7.8100	0.128264795
12	WaterDepth	0.0000	1.5000	0.125017600
7	Turbidity	0.0000	1567.0000	0.101772788
5	Dissolvedoxygen	0.1400	13.8000	0.100572516
13	wetlandarea	1804	2813	0.083843824
1	RedoxPot	-319.0000	85.0000	0.068900257
2	WaterTemp	3.8000	26.2000	0.067170134
3	AirTemp	3.2000	24.8000	0.063206375
22	MaxairTemp	9.7143	18.8000	0.053303712
28	P#E	14.1000	107.0000	0.050341973
26	SoilTemperature_300mm	6.3000	16.6571	0.046767006
18	Nitrite	0	51	0.045313638
0	pH	6.4000	9.5500	0.044131814
15	Flouride	0	88	0.035945653
25	SoilTemperature_200mm	5.7286	15.9429	0.035491179
27	RainfallAmount	25.1714	152.0714	0.032786499
21	Ammonia	0	477	0.031048247
29	SolarRadiation	63.2300	556.9650	0.030561076
20	Sulphate	0	53	0.030311506
30	Penman	4.2500	94.2500	0.023148534
23	MinairTemp	2.8143	11.5429	0.008492075
24	SoilTemperature_100mm	4.8857	15.7571	0.007199693

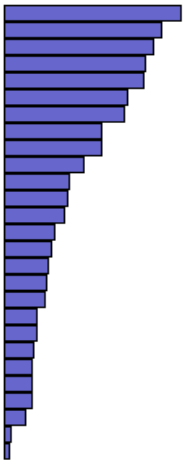


Figure 8.9: The relative sensitivity of wetland variables Model 2, with BOD, wetland volume, nitrate, phosphate and chloride with respect to Model 2.

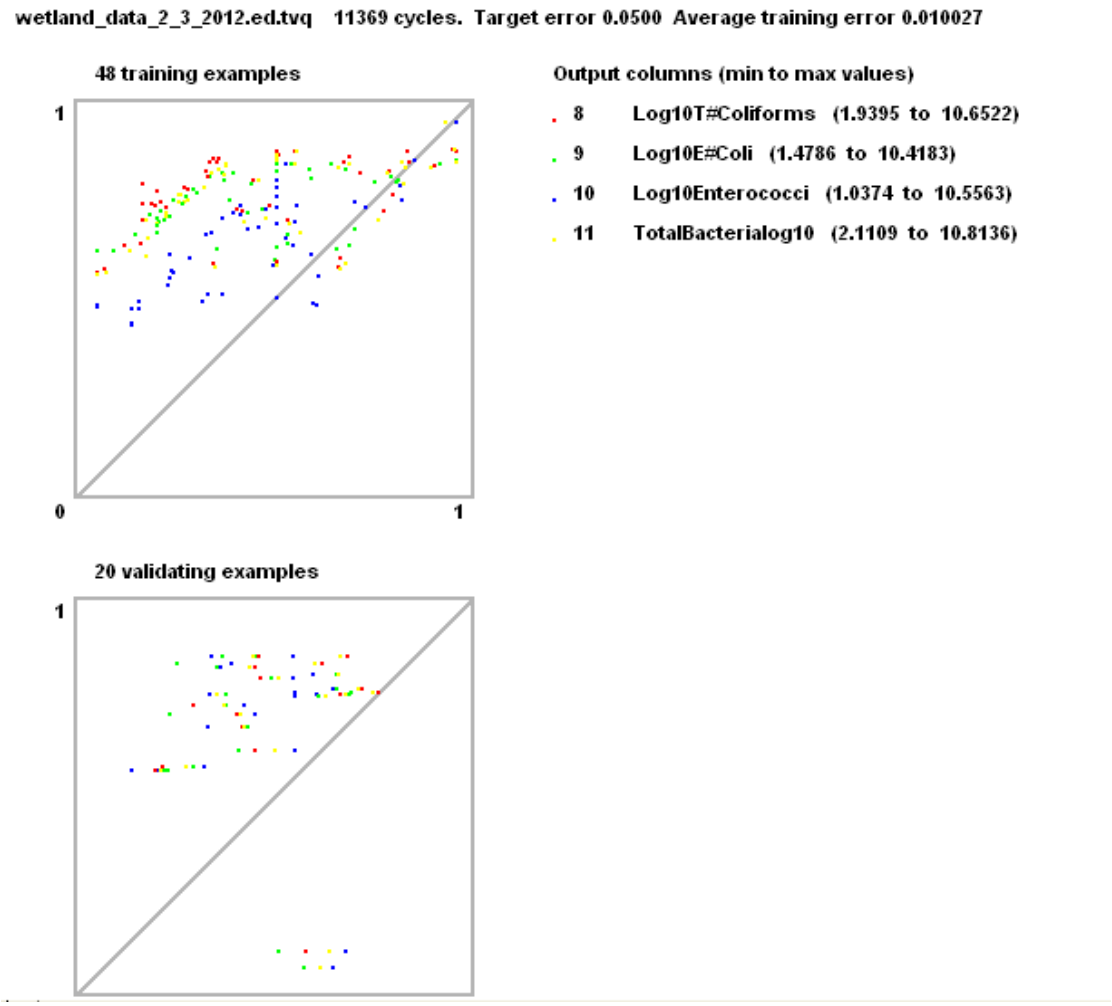


Figure 8.10: The ANN training and validating examples for Model 3, with the output nodes depicting the bacteria under analysis and their minimum to maximum $\log_{10}$ values. The target error was set to 0.05 and the resultant average training error was 0.010027, with 48 training and 20 validating examples.

wetland_data_2_3_2012.ed.tvq 11369 cycles. Target error 0.0500 Average training error 0.010027

The first 27 of 27 Inputs in descending order.

Column	Input Name	Importance	Relative Importance
12	WaterDepth	22.5449	
1	RedoxPot	19.8779	
0	pH	18.2154	
13	wetlandarea	17.3295	
21	Ammonia	15.2314	
27	RainfallAmount	14.4384	
14	wetlandvolume	13.3147	
4	Conductivity	12.3762	
3	AirTemp	12.2129	
15	Flouride	11.6871	
5	Dissolvedoxygen	11.5800	
23	MinairTemp	11.3236	
28	P#E	10.6681	
17	Nitrate	10.6654	
18	Nitrite	10.1335	
26	SoilTemperature_300mm	9.8760	
30	Penman	9.8618	
29	SolarRadiation	9.6919	
7	Turbidity	9.6436	
2	WaterTemp	8.2930	
6	BOD	8.1215	
19	Phosphate	7.9122	
25	SoilTemperature_200mm	7.7434	
24	SoilTemperature_100mm	7.2670	
22	MaxairTemp	5.8219	
16	chloride	4.6171	
20	Sulphate	4.2539	

Figure 8.11: The relative importance of wetland variables Model 3, water depth, redox potential and pH dominant with respect to Model 3.

wetland_data_2_3_2012.ed.tvq 11369 cycles. Target error 0.0500 Average training error 0.010027

The first 27 of 27 Inputs in descending order. Output column 8 Log10T#Coliforms

Column	Input Name	Change from	to	Sensitivity	Relative Sensitivity
12	WaterDepth	0.0000	1.5000	0.078338039	
1	RedoxPot	-319.0000	85.0000	0.077619849	
4	Conductivity	0.2000	7.8100	0.072049973	
28	P#E	14.1000	107.0000	0.065012112	
21	Ammonia	0	477	0.056397711	
13	wetlandarea	1804	2813	0.050568097	
0	pH	6.4000	9.5500	0.048744245	
6	BOD	1.4000	4350.0000	0.048503040	
5	Dissolvedoxygen	0.1400	13.8000	0.045746228	
3	AirTemp	3.2000	24.8000	0.034782545	
23	MinairTemp	2.8143	11.5429	0.029295338	
24	SoilTemperature_100mm	4.8857	15.7571	0.026509273	
30	Penman	4.2500	94.2500	0.024691828	
18	Nitrite	0	51	0.024424101	
27	RainfallAmount	25.1714	152.0714	0.023893247	
29	SolarRadiation	63.2300	556.9650	0.022761496	
25	SoilTemperature_200mm	5.7286	15.9429	0.022276545	
19	Phosphate	0	106	0.021621424	
26	SoilTemperature_300mm	6.3000	16.6571	0.019448455	
16	chloride	1	819	0.018836030	
14	wetlandvolume	0	2671	0.011108496	
17	Nitrate	0	241	0.010992730	
2	WaterTemp	3.8000	26.2000	0.010068233	
15	Flouride	0	88	0.008370550	
20	Sulphate	0	53	0.004964656	
22	MaxairTemp	9.7143	18.8000	0.003017585	
7	Turbidity	0.0000	1567.0000	0.001218310	

Figure 8.12: The relative sensitivity of wetland variables Model 3, with water depth, redox potential and conductivity with respect to Model 3.

8.8 A review of key variables

After the completion of the three models, seven key wetland variables were reviewed. Log_{10} total coliforms (CFU/100ml), water depth (m), ammonia (mg/L), turbidity (NTU), BOD_5 (mg/L), dissolved oxygen (mg/L) and redox potential (mV). Each graph visually depicts the following:

Blue line = Actual Value,
Red Line = Average Value
Green Line = Missing Data
Yellow line = Estimate
Black broken line = Trend.

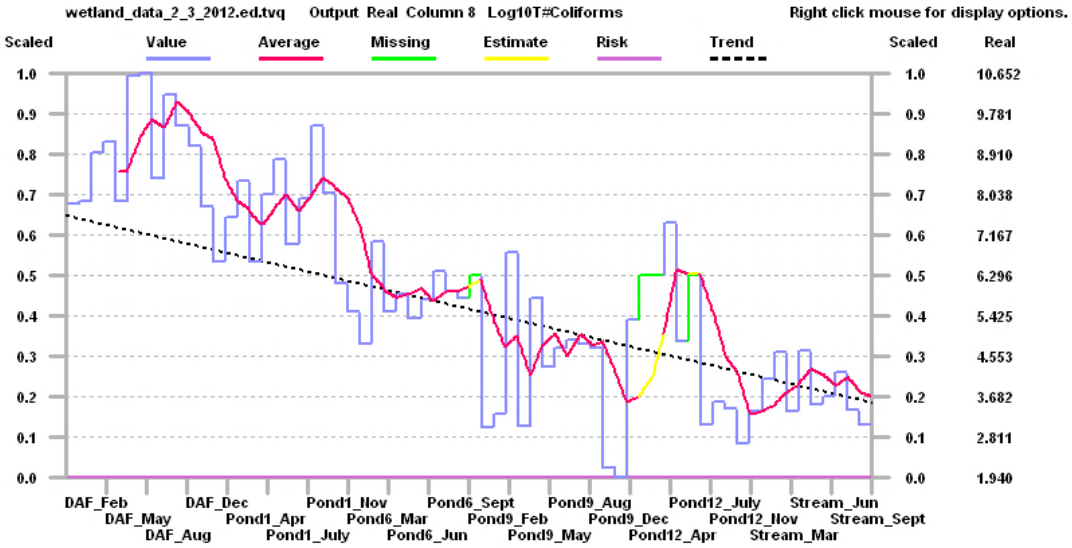


Figure 8.13: The ANN log_{10} total coliforms (CFU/100ml) distribution (by sample point by month).

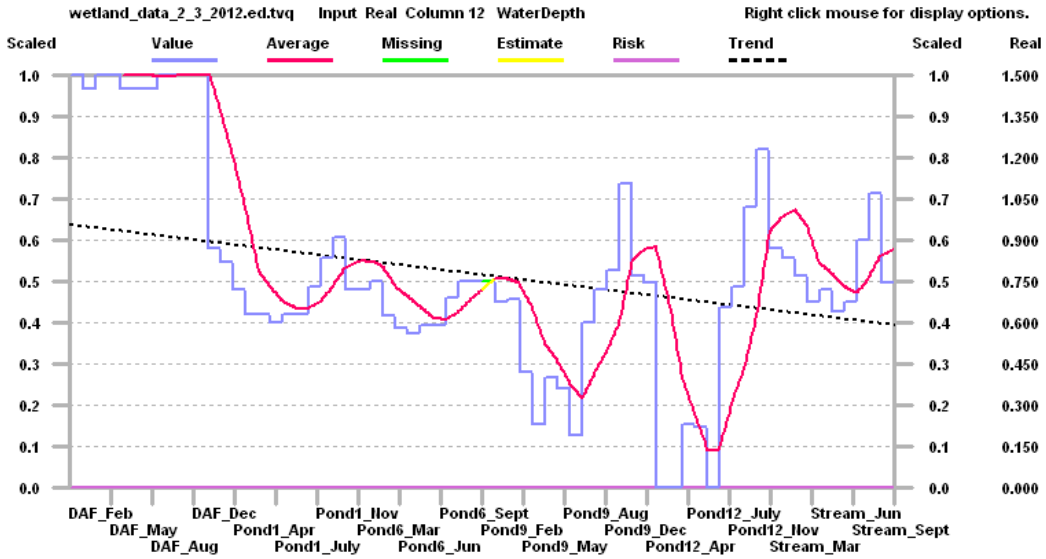


Figure 8.14: The ANN water depth (m) distribution (by sample point by month)

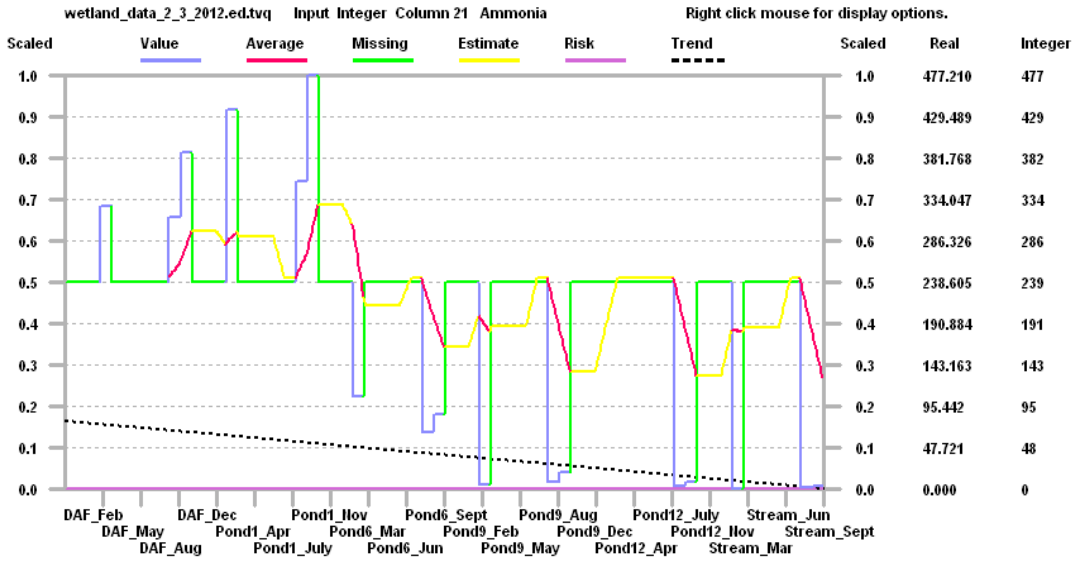


Figure 8.15: The ANN ammonia (mg/L) distribution (by sample point by month).

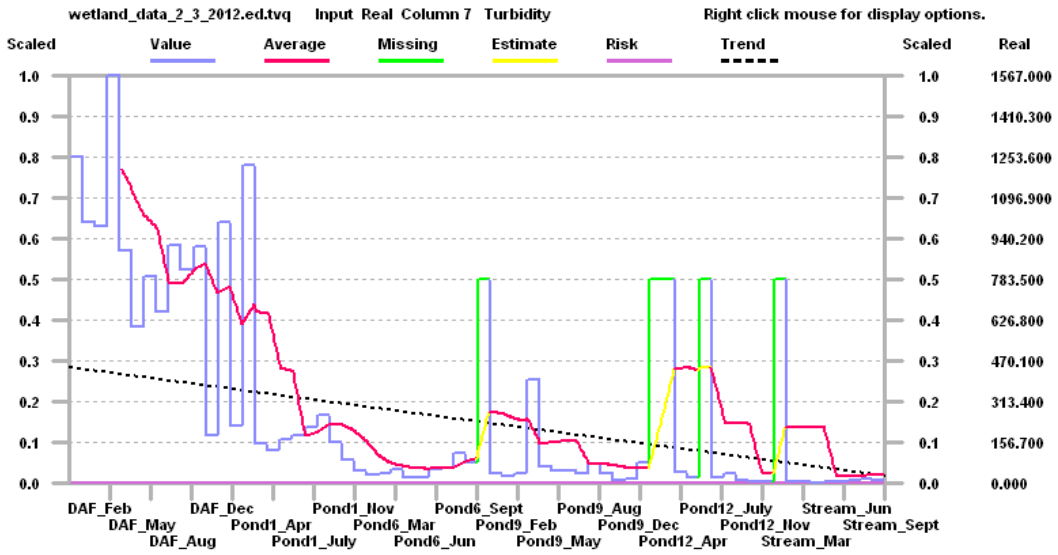


Figure 8.16: The ANN turbidity (NTU) distribution (by sample point by month)

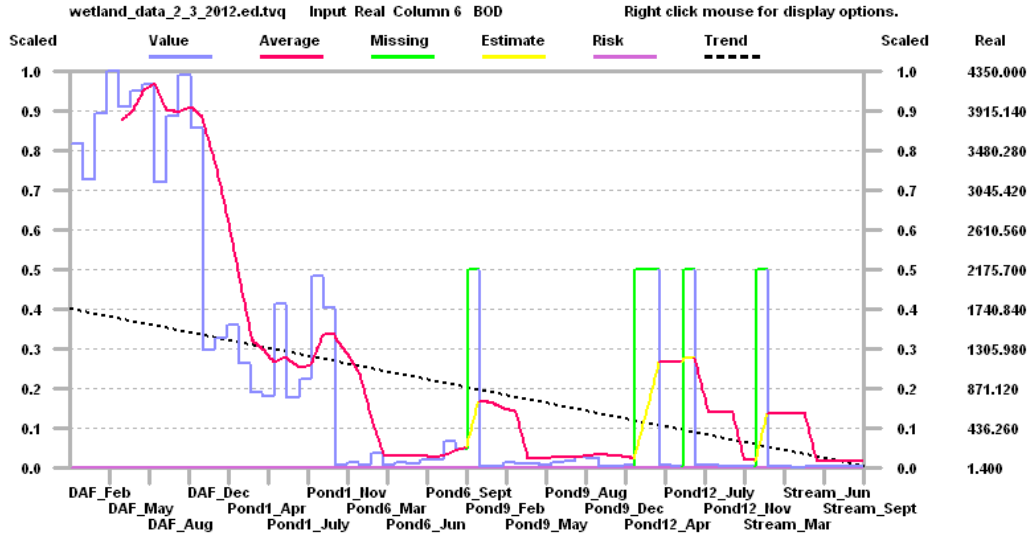


Figure 8.17: The ANN BOD_5 (mg/L) distribution (by sample point by month).

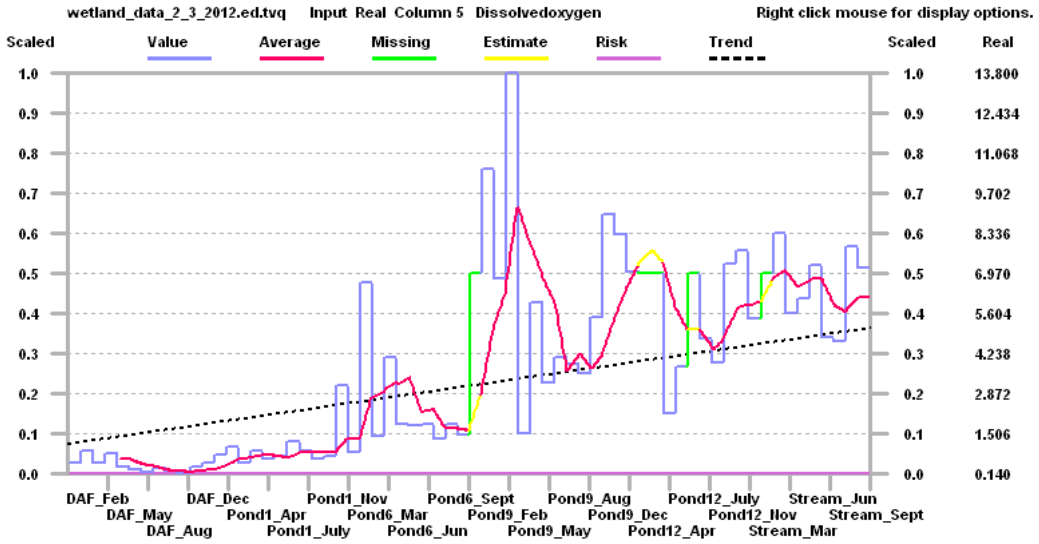


Figure 8.18: The ANN dissolved oxygen (mg/L) distribution (by sample point by month).

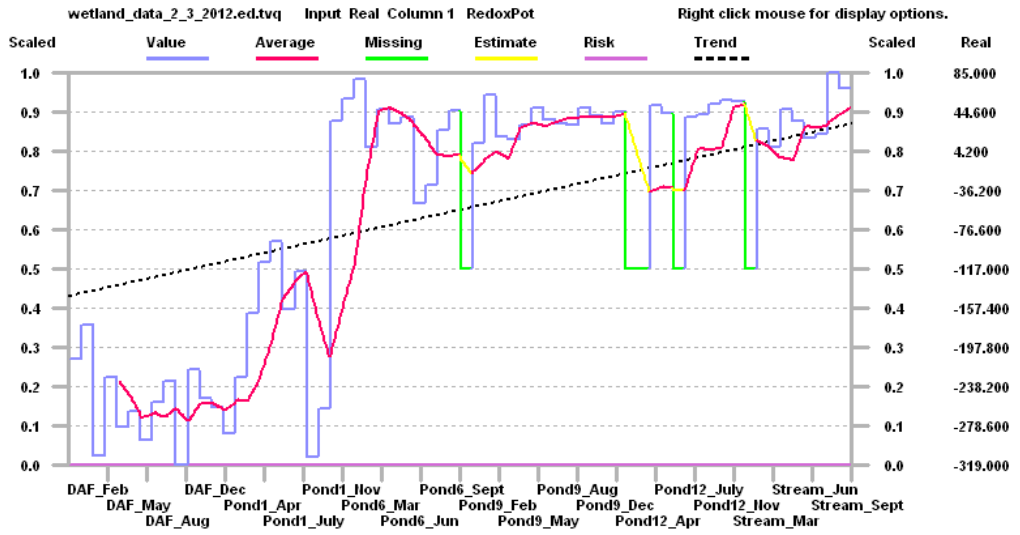


Figure 8.19: The ANN redox potential (mV) distribution (by sample point by month).

8.9 A review of the ANNs Relative importance and Relative sensitivity

It was decided to capture the first seven relative important and sensitive variables from the three models, see Tables 8.1 and 8.2 respectively. The most relative important dominant variables with respect to the output variables [Log_{10} (total Bacteria, total coliforms, *E.coli* and enterococci)] are the wetland volume and area, redox potential, pH and ammonia, see Table 8.2 hits row.

The most relative sensitive dominant variables with respect to the output variables [Log_{10} (total bacteria, total coliforms, *E.coli* and enterococci)] are the conductivity, pH, water depth, BOD and chloride, see Table 8.2 hits row. As a rule of thumb, in this kind of analysis, the variables assuming values lower than 5% should be removed from the models. In this case, there was only one variable pH (Model 3, Sensitivity Table 8.2) which falls outside the criteria, but it was decided to leave the variable in place, as the value was just under the 5% threshold.

Table 8.1: Relative importance of wetland variables as percentage values (%), evaluated from Model 1, 2 and 3.

Model	pH	Redox pot	Turbidity	D.O.	Water depth	Volume	Area	Chloride	Ammonia	Phosphate	Nitrate.	Rainfall
Model 1		36.5		35.05		33.62	27.5	30.84	36.55		29.86	
Model 2	41.56	38.8	35.39	38.51		41.53	47.49			34.13		
Model 3	18.22	19.9			22.54	13.31	17.33		15.23			14.44
Hits	2	3	1	2	1	3	2	1	2	1	1	1

Table 8.2: Relative sensitivity of wetland variables as percentage values (%), evaluated from Model 1, 2 and 3.

Model	pH	Redox pot	Conductivity	BOD	Volume	Area	Chloride	Phosphate	Ammonia	Nitrate	Solar rad.	P.E	Rainfall
Model 1	6.83			12.66			7.62				6.28	8.35	7.26
Model 2			12.501	12.82	16.25		14.32	14.58		15.31			
Model 3	4.87	7.76	7.83	7.2		5.06			6.5			6.5	
Hits	2	1	2	3	1	1	2	1	1	1	1	2	1

8.10 Results and discussion

From reviewing the graphs, the trend line depicts either a decrease in values as in the case with Log_{10} total coliforms (the same trend behaviour can be seen with *E.coli* and enterococci concentrations), Water depth, ammonia, BOD and turbidity, see Figures 8.13-8.17 respectively or an increase in values as depicted by dissolved oxygen and redox potential, see Figures 8.18 and 8.19 respectively. But on closer examination of the total coliforms Figure 8.13, there is an increase in concentration at the later end of the wetland system, located at pond 12. This increase or decrease depending on the parameter under review was replicated in all graphs, where there was missing data, the ANNs' have learned, validated and tested and interpreted their outcome as to what value/s can be inserted into missing data sets.

On closer examination, the water depth graph, Figure 8.14, coincides with the increase of total coliforms, as the bacterial concentration increased, the water depth decreases. Thus as the water depth decreases to below 0.3m - 0.15m, see Figure 8.14, the bacterial count increases, this behaviour was present in *E.coli* and enterococci concentrations (not shown). From reviewing the other graphs (Figures 8.14-8.18) at pond 12's sample location, there was a similar trend as discussed above. The likely cause of the water depth decrease can be attributed to increase solar radiation (see Table 8.3), resulting in an increase in air/water temperatures and increase in microbial activity (Kadlec and Wallace, 2009).

On analysing the ANN models, the prominence of the wetland design characteristics, volume and area was evident in Models 1 and 2, whereas water depth became obvious as a relative important variable in Model 3. Notice that Table 8.2, in relation to relative sensitivity shows that model 2 and 3 highlight water as a sensitive variable.

8.10.1 ANN Model 1 shows that ammonia, redox potential and dissolved oxygen are the most important variables, see Figure 8.4. Conductivity was the most sensitive variable see Figure 8.5 and that pond 1 and DAF are the two locations, see Figure 8.3, that are the sample points of interest (above the relative target error = 0.05) during the summer months.

8.10.2 ANN Model 2 shows that the wetland volume, area and pH are the most important variables, see Figure 8.8. BOD and nutrients (nitrate, phosphate, and chloride) were the most sensitive, see Figure 8.9. The backend of the constructed wetland the stream, pond 9, pond 12 are the locations, see Figure 8.7, that are the sample points of interest (above the relative target error = 0.05) across the sampling period.

8.10.3 ANN Model 3 shows that water depth, redox potential and pH are the most important variables see Figure 8.11. Water depth, redox potential and conductivity were the most sensitive, see Figure 8.12 and that pond 9 and pond 12 are the locations that are of interest (above the relative target error = 0.05).

The importance of pH, redox potential and dissolved oxygen are present in all three models, see Table 8.1, these variables are commonly used in many bacterial analysis and applications, Samper *et al*., (2007) used these variables to create models in relation to microbial consumption in a radioactive waste repository and Pankaj *et al*., (2008) utilises these parameters with respect to nutrient removal in a bioreactor system. The inclusion of wetland design characteristics in the ANN models, such as wetland area, volume and water depth, highlight the importance of wetland design and modelling guidelines (Lizama *et al*., 2011) realised that the wetland design features along with pH, alkalinity, dissolved oxygen, temeperatures and microbial community struture are key in the removal of arsenic

using constructed wetlands. The authors used simple kinetic models not ANN models in their paper.

The presence of ammonia in two of the three ANN models is also critical in the variables networked with bacterial concentrations within the wetland system. Song *et al*., (2013) tested three models (linear regression, mechanistic model and ANN) to model the denitrification process in a mesocosm wetland, monitoring nitrate, pH, dissolved oxygen, water temperature, water depth, dissolved organic carbon and denification enzyme activity (DEA). Particular interest was directed towards the importance of understanding nitrate and dissolved oxygen variations within the modelling scenarios. They found that the ANN model had a greater performance in simulating variations in denification processes (DEA) than the linear regression and the mechanistic models.

In summary, the three ANN models identified different wetland variables, but the importance of several key variables was highlighted with respect to the bacterial concentrations within the wetland, these include as pH, redox potential, dissolved oxygen, conductivity, ammonia, wetland volume, wetland area and water depth. The analysis also reveals other inorganic ions, such as chloride, phosphate and nitrate are of relative importance and sensitivity. From a climate perspective the variables are mainly sensitive in comparison to important with solar radiation and P.E. (Potential Evapotranspiration) dominating with rainfall the only climate variable present in both the relative importance and sensitivity tables.

The final step was to utilise the software's association table, where all the variables used in the models are paired together and ranked in order of the strength that pairing, see Table 8.3.

Table 8.3: ANN Association table

29	SolarRadiation	28	P#E	4.416252
30	Penman	29	SolarRadiation	4.373838
5	Dissolvedoxygen	1	RedoxPot	4.371891
26	SoilTemperature_300mm	25	SoilTemperature_200mm	4.140014
25	SoilTemperature_200mm	22	MaxairTemp	4.131447
24	SoilTemperature_100mm	22	MaxairTemp	4.084553
28	P#E	29	SolarRadiation	4.067841
13	wetlandarea	1	RedoxPot	4.060158
23	MinairTemp	22	MaxairTemp	4.020034
22	MaxairTemp	23	MinairTemp	3.981363
12	WaterDepth	1	RedoxPot	3.974776
1	RedoxPot	29	SolarRadiation	3.891162
8	Log10T#Coliforms	29	SolarRadiation	3.855496
11	TotalBacterialog10	29	SolarRadiation	3.799732
4	Conductivity	29	SolarRadiation	3.632786
9	Log10E#Coli	29	SolarRadiation	3.625541
17	Nitrate	29	SolarRadiation	3.582534
0	pH	29	SolarRadiation	3.535249
21	Ammonia	1	RedoxPot	3.516509
3	AirTemp	28	P#E	3.511455
7	Turbidity	29	SolarRadiation	3.504866
19	Phosphate	6	BOD	3.499414
6	BOD	29	SolarRadiation	3.490973
20	Sulphate	1	RedoxPot	3.465158
10	Log10Enterococci	29	SolarRadiation	3.422227
2	WaterTemp	28	P#E	3.383023
27	RainfallAmount	29	SolarRadiation	3.379865
14	wetlandvolume	29	SolarRadiation	3.332617
18	Nitrite	29	SolarRadiation	3.332315
16	chloride	29	SolarRadiation	3.161049
15	Flouride	29	SolarRadiation	3.128422

The association as shown in Table 8.3, highlights the importance of climate variables, in particular solar radiaton with respect to its strong association with all the key wetland variables, such as the bacterial and inorganic ions concentrations within the wetland system. The association view allows all the columns in the neural network grid to be evaluated for strength of interations or associations between the variables in the grid. All the columns are set to inputs and each column in turn was changed to output and the network created. The association network was trained so the most important input variables can be determined and the output was paired with the most important input variables.

8.11 Conclusions

The approach presented in this chapter demonstrates the importance of using different ANN models to understand varying behaviours within a complex system

such as a constructed wetland. There was no right or wrong model but each model (i.e. additional hidden layers) produced subtle variations with particular wetland variables dominating, such as (1) pH, (2) redox potential, (3) dissolved oxygen, (4) conductivity, (5) wetland area, (6) volume and (7) water depth along with (8) ammonia, these eight variables are critical to understanding bacterial behaviours within a complex wetland system. Therefore, using bacteria to fingerprint a wetland, provides critical insights into wetland/ pond efficiency, such as the increase in bacterial concentration in pond 12 during the summer months, with a corresponding inverse relationship to a decrease in water depth. Due the complex nature of the wetland, the ANN models reveal a dichotomy within the wetland system. DAF, pond 1 (ammonia, redox potential, dissolved oxygen as revealed by Model 1 versus pond 9, pond 12 (water depth, volume, area, pH, redox potential as revealed by Models 2 and 3). i.e. high ammonia concentrations dominant the front of the system, and low water depth/volumes dominate the back of the system.

The inclusion of climate data into the ANN models was also shown to be the cause of issues within the wetland system with the dominance of summer months as a problematic period for the wetland. The subsequent association analysis revealing importance of solar radiation as the dominant variable of interaction across all wetland variables. During the analysis process, the average training error gets larger as the hidden layers' increases, as the complexity of interactions increase, so to do the errors (Model 1 = 0.001, Model 2 = 0.005 and Model 3 = 0.01). As the hidden layers increase in complexity the relative importance and sensitivity also diminish, see Tables 8.1 and 8. 2, which show the values decrease from Model 1 to Model 3, indicating that water depth behaviour was a complex and hidden component of the wetland system.The analysis here indicates the importance of using bacterial values along with a multivariate dataset to understand a complex system and to derive important wetland variables that allow a more comprehensive monitoring solution to these wetland systems. The use of a multi-probe meter, combining pH, temperature, dissolved oxygen, conductivity and redox potential may be useful in understanding wetland behaviour as these variables could provide ample data to asses the overall wetlands performance.

Chapter 8. Section 2: The application of self -organising maps (SOMs) to understand wetland behaviours

8.13 Summary

This section uses self-organising maps (SOM) to examine various variables collected from a free water constructed wetland system (FWS), treating abattoir wastewater. The SOM is a machine learning method that can be used to explore patterns from large and complex datasets for linear and non-linear patterns. The SOM is a type of artificial neural network (ANN) based on unsupervised learning that produces a visual representation of the datset called an input space or map. The SOM can also be used as a what-if scenario analysis tool, based on the fact that the SOM system is based on unsupervised training and learning of datasets. The results show that the wetland can be segregated into five clusters with visual diagrams to reveal complex variable distribution within the wetland system.

The analysis also reveals that a possible scenario exists, where wetland inorganic ions such as ammonia, phosphate, nitrate and sulphate will migrate from high concentrations values in pond 1 (input) to the midpoint pond 6, resulting in increased bacterial concentrations at the pond 12 and stream (output). The SOM model was an appropriate approach to monitor complex behaviours within the constructed wetland and to reveal different scenarios within the system.

The SOM method was evaluated against other clustering algorithm methods such as *K*-means and two-stage hierarchical clustering methods to compare and contrast these three clustering methodologies, with SOM the preferred option.

8.14 Introduction

The SOM, also called the Kohonen map, is a heuristic model (a trial-and-error method of problem solving, used when an algorithmic approach is impractical) for exploring and visualising patterns in high dimensional datasets (Kohonen 1982). SOM can be considered as a clustering technique that identifies clusters in a dataset without rigid assumptions of linearity or normality that is typified by traditional statistical techniques.

The SOM is trained based on an unsupervised training algorithm where no target output is provided and the network evolved until convergence i.e. supervised training algorithms are based on the iterative reduction of a prediction error, but unsupervised training algorithms usually lack this mathematical framework: an unsupervised training algorithm usually does not try explicitly to minimise anything.

The SOM architecture

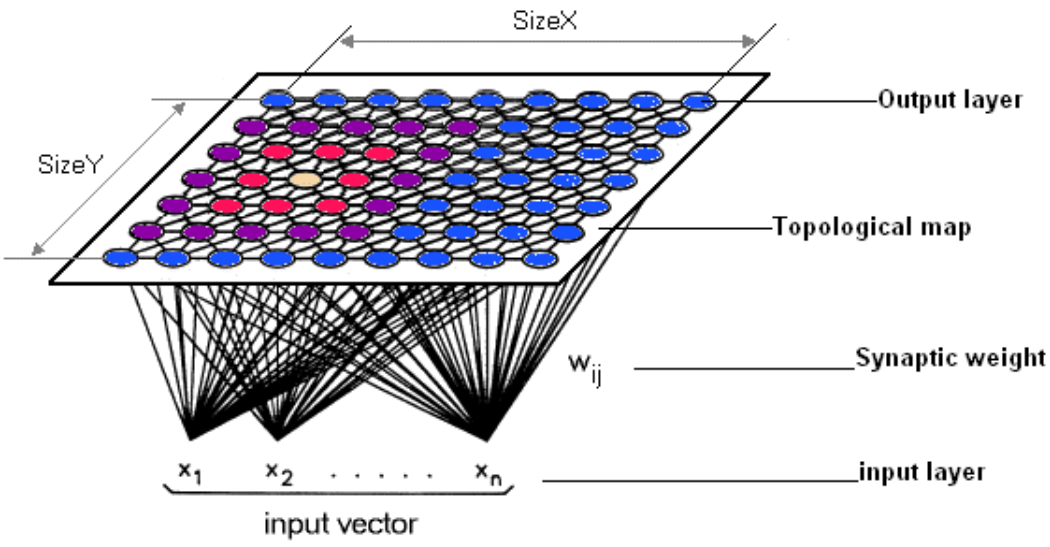


Figure 8.20: The SOM architecture

Ref: http://www.lohninger.com/helpcsuite/kohonen_network_-_background_information.htm>

The SOM consists of two layers: the input layer/vector which classifies the data according to similarity and their similarity and the output layer of radial neurons arranged in a two-dimensional map, see Figure 8.20. The output neurons will self-organise to an ordered map and neurons with similar weights are placed together. They are connected to adjacent neurons (similar to a neighbourhood), dictating the topology of the map (Moreno *et al*., 2006). Since the SOM compresses the dataset while maintaining the most important map relationships of the primary dataset elements on display, it can be used for pattern classification (Silven *et al*, 2003).

As stated previously, the SOM utilises unsupervised training and with the addition of excellent visualisation, self-organised maps have been recently used in a myriad of classification and clustering tasks. Examples include, in the area of environmental analysis; to assess the heavy metal removal performance in constructed wetland (Scholz and Lee, 2006), the use of SOM in water resources analysis modelling and application (Kaltech *et al*., 2008), mutal funds classification (Moreno et al., 2006), traffic management systems (Chen *et al*., 2008), in the analysis of NMR spectra (Dow *et al*., 2004), diversity bacterial gene analysis (Kanaya *et al*., 2001) and wastewater treatment plants (Kern, 2016).

8.15 Self-organising maps (SOM) Data

See Appendix A, for the wetland dataset. Solar radiation and Penman (Climate data) have been included in the self-organising maps analysis to further test the efficacy of the SOM analysis, see Appendix G.

8.16 Self-organising maps (SOM) –based clustering

There are many software packages available for analysing SOM models. SOMine package version 5.2 (Viscovery Software, 2011) was selected. This software applies artificial intelligence techniques to automatically find the efficient SOM clusters. In this study a hierarchical cluster analysis with a Ward linkage method was applied to the SOM to clearly define the edges of each cluster. The number of neurons chosen was 2000, with a tension of 0.75 with accurate mapping, see

Figure 8.21 shows the SOM –Ward converges successfully to a five-cluster solution after the SOM-Ward analysis.

The SOM-Ward clusters for the constructed wetland

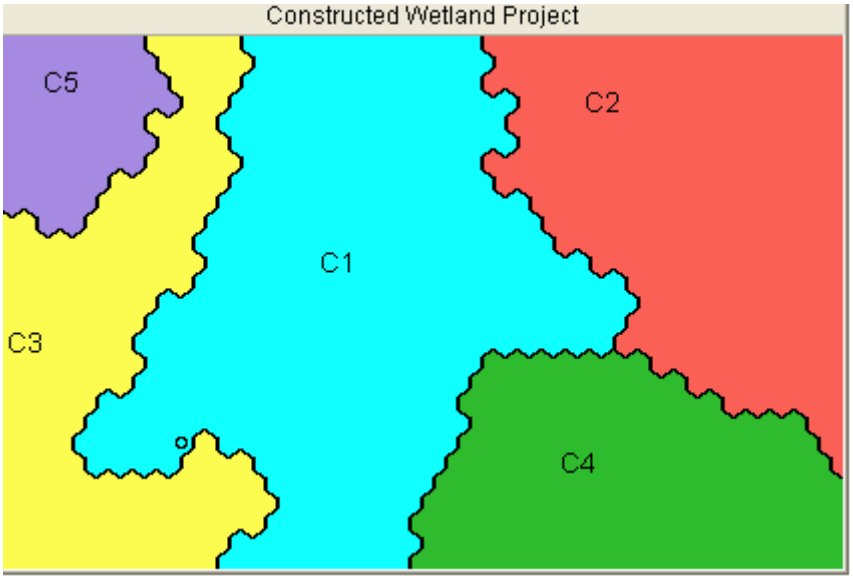


Figure 8.21. The SOM-Ward clusters (Scenario 1)

Where C5 = DAF cluster, C3 = Pond 1 cluster, C1 = Pond 6 cluster, C4 = Pond 9 to Pond 12 Cluster and finally C2 = P12 to Stream Cluster.

Based on the SOM-Ward clusters, components maps can be constructed (Vesanto, 1999). These maps are also known as 'temperature' maps.' (Churilov and Flitman, 2006). On these maps, the nodes which share similar information are organised in close proximity to each other. The two-dimensional hexagonal grid shows clear division of the input pattern into five clusters. Figure 8.21 shows clear delineation between the cluster edges, using a hierarchical cluster analysis inclusive of the SOM-Ward linkage method. There are two learning algorithms for SOM (Kohonen, 2001), the stochastic (sequential) and the batch learning algorithm. In the former, the reference vectors are updated immediately after a single input vector was applied. In the later, the update was done using all input vectors. The "Batch-SOM" method was applied in this analysis as it less likely to

incur convergence (bottleneck) issues. Figure. 8.21 shows the component maps for all 5 clusters and for input variables (vectors). *Note* that five different scenarios were evaluated and we decided to review Scenario 1 (see Table 8.4) due to its cluster frequency, similarity was observed between SOM and the Fuzzy Bayesian reliability analysis as seen in chapter 10 section 3 (fuzzy Bayesian analysis). Each cluster represents a wetland segment. The initial wetland dataset was organised in the following sampling points; DAF, pond 1, pond 6, pond 9, pond 12 and stream i.e. 6 sampling points. The SOM method has re-ordered the dataset into 5 clusters or segments, each now defined cluster as; DAF, pond 1, pond 6, pond 9 to 12, pond 12 to stream (there is cluster spreading occurring in the back-end of the wetland). Cluster 5 (DAF) with the lowest cluster frequency of 8%, Cluster 3 (pond 1) with a cluster frequency of 17%, Cluster 1 (pond 6) with a cluster frequency of 27%, Cluster 2 (pond 9 to pond12) with a cluster frequency of 27% and Cluster 4 (pond 12 to stream) with a cluster frequency of 21%. See Table 6.4 and Figure 6.22.

Table 8.4: Scenario 1 cluster frequency table

Cluster	Description	Frequency
C5	DAF	0.08
C3	P1	0.17
C1	P6	0.27
C2	P9-P12	0.27
C4	P12-Stream	0.21

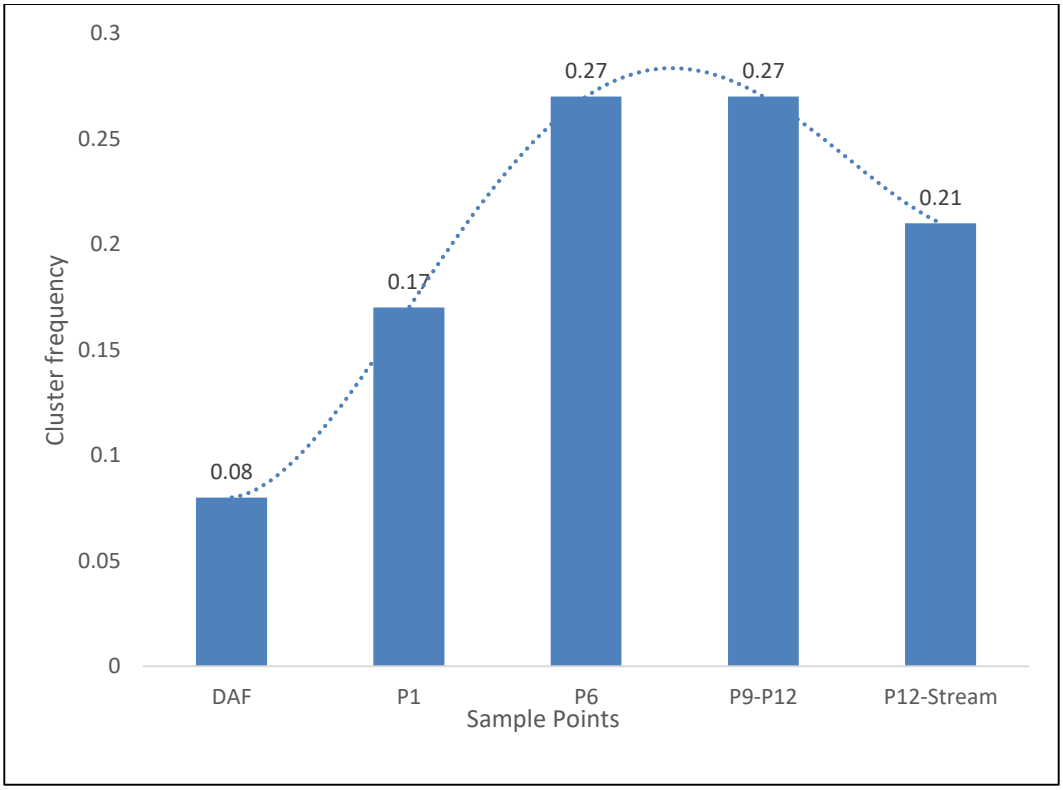


Figure 8.22. The cluster frequency for scenario 1, SOM analysis. Cluster Frequency for each the five clusters within the wetland.

Figure 8.22. shows strong similarities to the fuzzy Bayesian reliability graph displayed in chapter 10 section 3; another similarity that was present with respect to the fuzzy Bayesian analysis was that towards the backend of the wetland parent nodes (pond 9, pond 12, stream) were the strong correlations between ponds, starting with Pond 9. The graph depicts this behaviour, showing the interaction between pond 9-pond 12 and pond 12-stream. Intimating that the SOM results show that the DAF, pond 1 and pond 6 reveal little dispersion of the input abattoir wastewaters; whereas variable (vector) dispersion was more evident from pond 9 through to pond 12, due the vector spreading in the backend of the wetland.

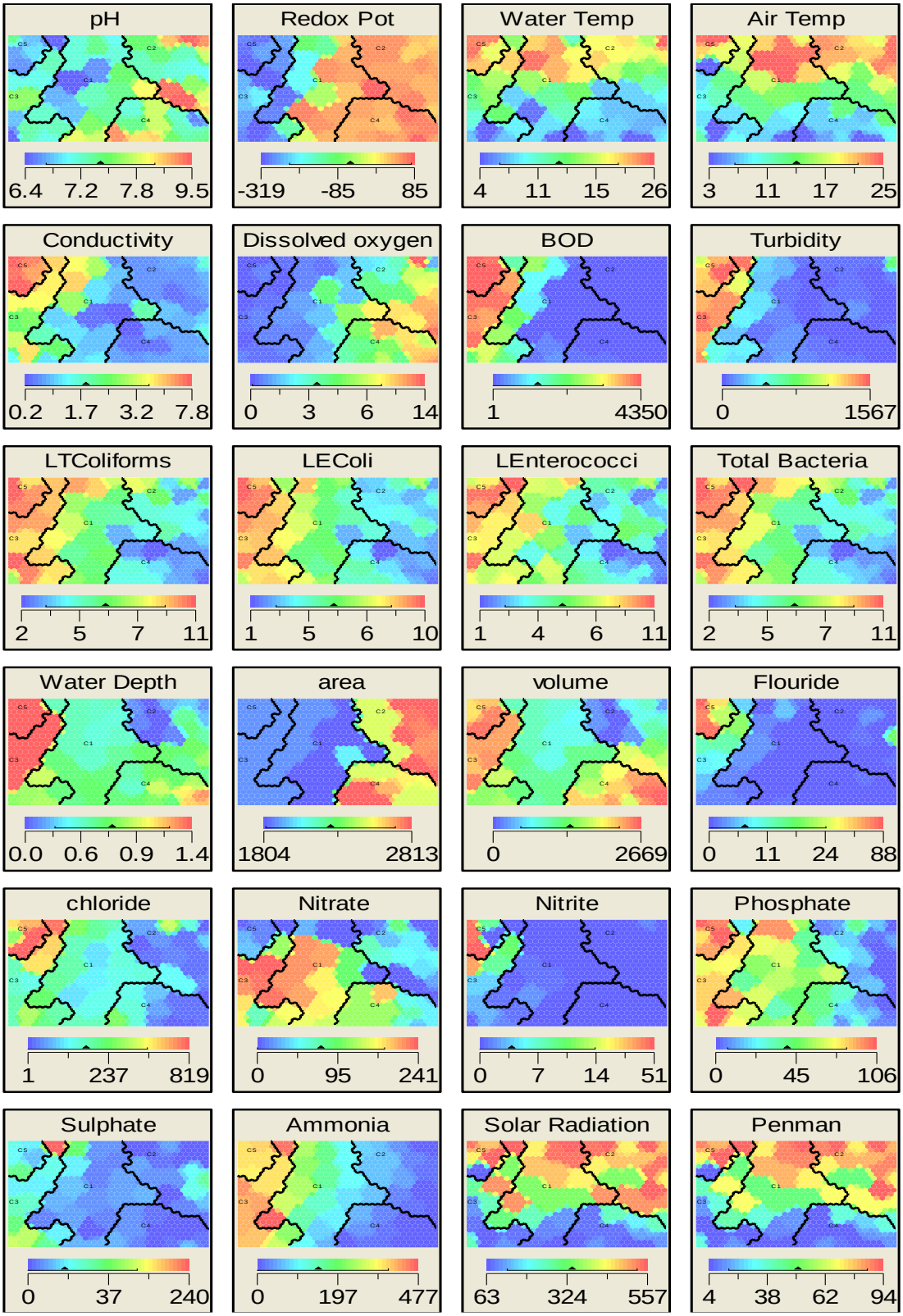


Figure 8.23. The self-organising maps for all the wetland variables. Blue denotes regions of low values, while red denotes regions of high value.

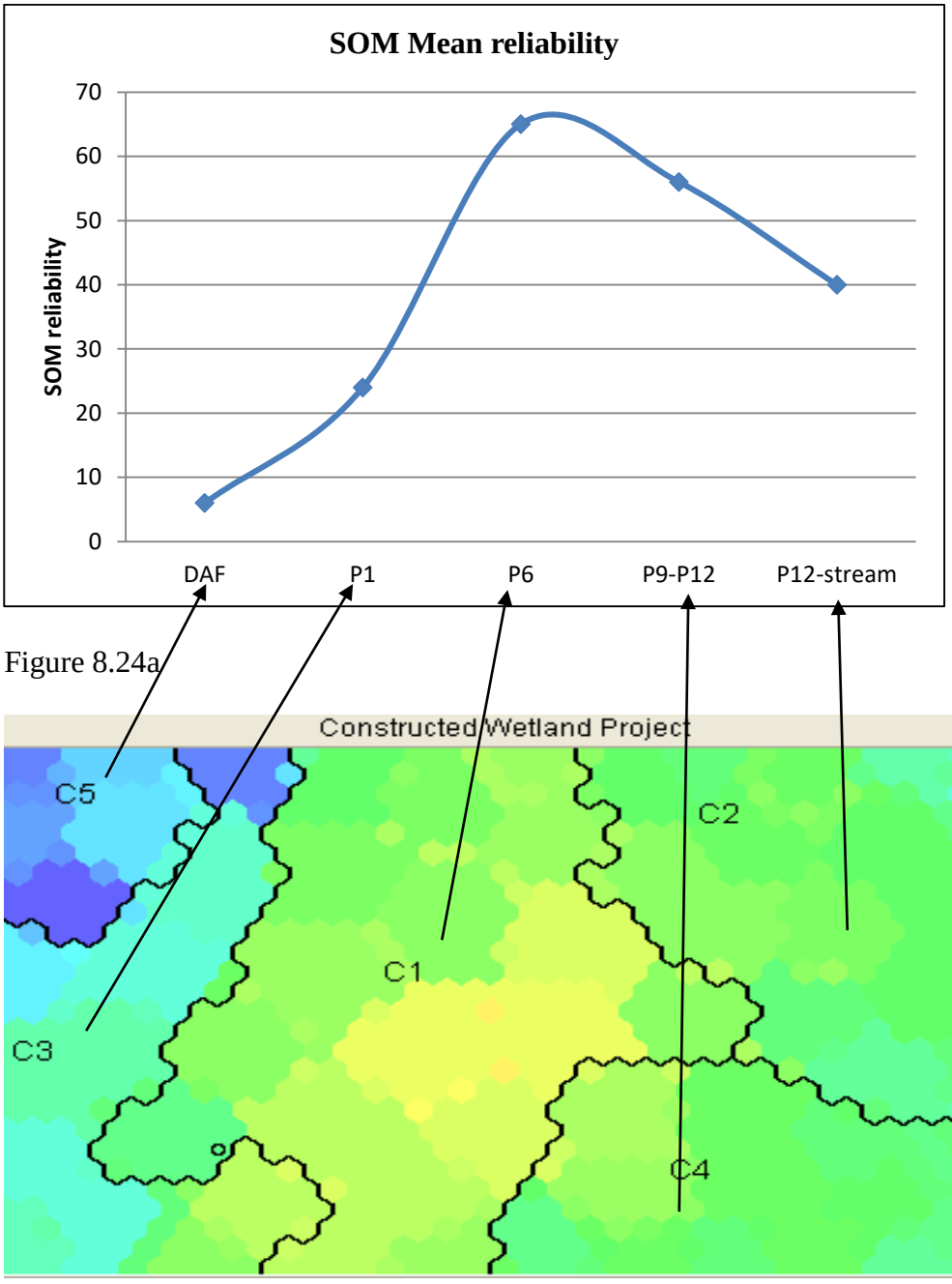


Figure 8.24a

Figure 8.24b shows the overall 2-dimensional map of the global shading colour cluster of the wetland i.e. when all the SOM maps from Figure 8.23 are added together, to reveal the overall wetland colour contour map.

Figure 8.24.a shows the SOM global sharing cluster maps, reduced to a line chart, with the mean line representing Figure 8.24.b, colour contour map.

The importance of the global shading colour cluster is as a visual tool to realise the complete dataset behaviour within a complex system, Zhou and Lopresti (1997) and Park *et al.*, (1998) used global colour clusters with respect to documentation analysis and imaging techniques respectively. The image shows that all the entered variables are transformed into a global 2-D map of the wetland system. Figure 8.24b, shows the resultant transformation of the global shading cluster and Figure 8.24a into a simple line graph. There are similarities between Figure 8.24b and the fuzzy Bayesian reliability graphs as seen in chapter 10 section 3. However, the main difference between the two models, is that SOM reveals subtler variation (visually) within the wetland system, in particular ponds 6 and 9 which are susceptible to ammonia, phosphate and nitrate loadings, see Figure 8.23 for individual ammonia, phosphate and nitrate loadings and trace variable values onto Figure 8.24a global shading cluster map.

8.17 What if scenarios

SOMine® was used to develop different scenarios that could potentially occur within the wetland, because self-organising maps are unsupervised, they can be utilised to reveal what-if scenarios. Examples include scenario simulation of salinisation responses within different sites contained in freshwater stream habitats (Horrigan, 2005), developing atmospheric scenarios using Peru's seasonal climate data (Gutierrez *et al*, 2005) and regional flood frequency analysis (Srinvas *et al*., 2008). In this scenario, the model was iterated five times to uncover five different behaviours within the wetland system, the Figures 8.25-8.35 show, that the potential exists that Pond 6 could become stressed due to increased inorganic ions (ammonia, nitrate, phosphate and sulphate), with increased BOD, conductivity and turbidity and decreased dissolved oxygen and redox potential, with the potential of an increase in bacterial concentration, not just in pond 6 but through pond 9-12 and pond 12-stream.

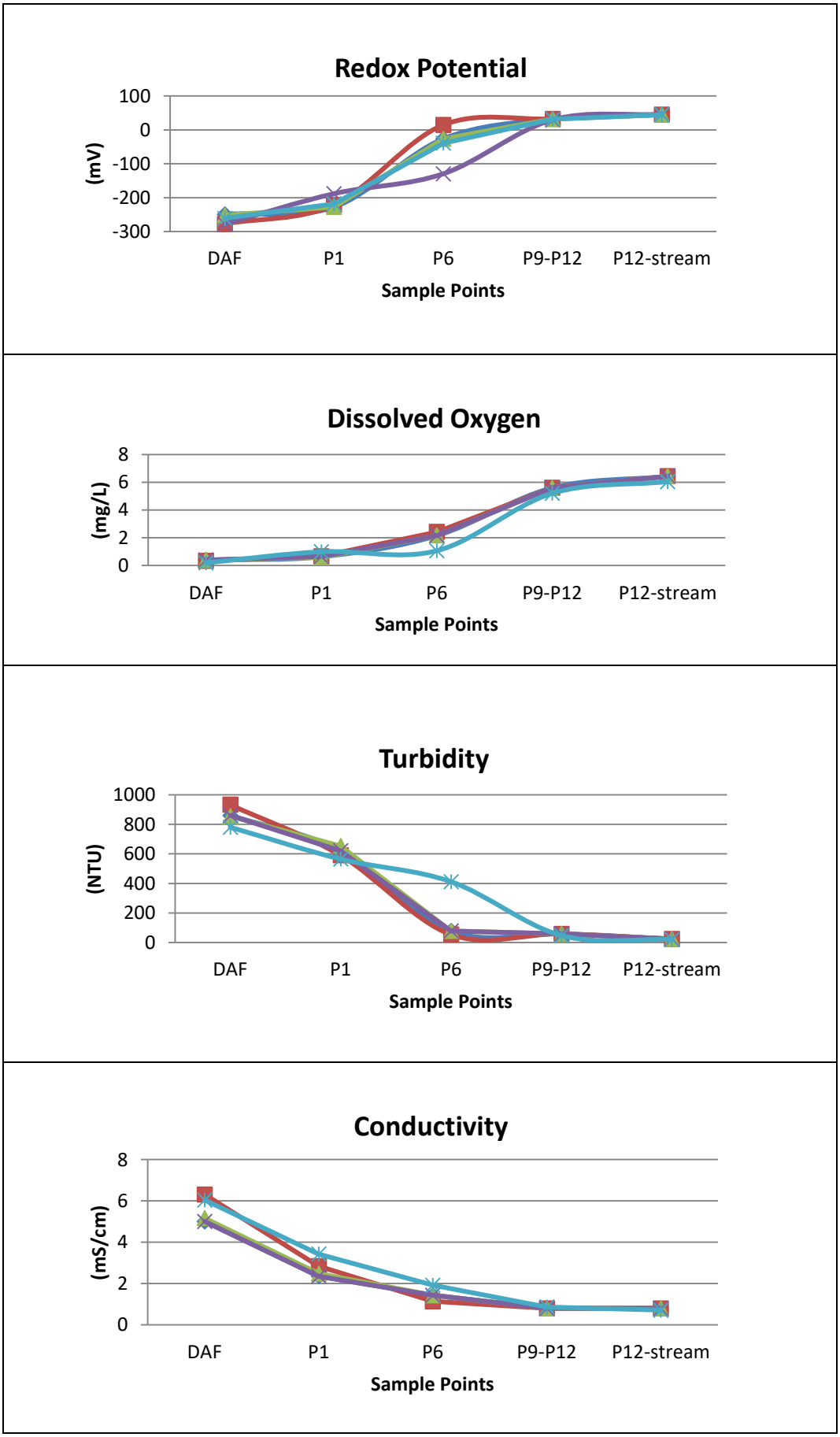

Redox Potential
100
0
-100
-200
-300
(mV)
DAF
P1
P6
P9-P12
P12-stream
Sample Points
Dissolved Oxygen
8
6
4
2
0
(mg/L)
DAF
P1
P6
P9-P12
P12-stream
Sample Points
Turbidity
1000
800
600
400
200
0
(NTU)
DAF
P1
P6
P9-P12
P12-stream
Sample Points
Conductivity
8
6
4
2
0
(mS/cm)
DAF
P1
P6
P9-P12
P12-stream
Sample Points

Fig.8.25. The redox potential, dissolved oxygen, turbidity and conductivity SOMine scenarios. The different colours denote different scenario runs.

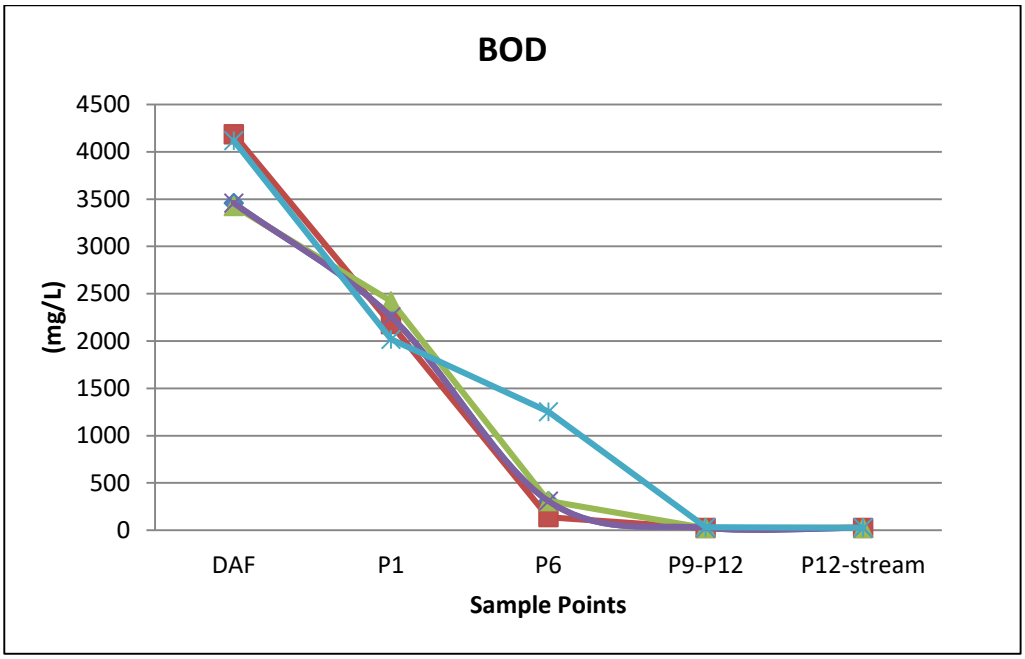


Figure 8.26. The BOD_5 SOMine scenarios

Figures 8.25 and 8.26 depicts five possible scenarios in relation to redox potential, dissolved oxygen, turbidity, conductivity and BOD. One common trend with respect to the wetland physical-chemical variables is the potential that pond 6 may incur changes to these variables. Where a low dissolved oxygen concentration may be recorded or an increase in BOD. This common trend is repeated in the bacterial and inorganic scenarios. The blue colour represents one of scenarios run by the SOMine® software, which indicates the potentail exixts within the analyised wetland dataset of an increase in the vicinity of Pond 6. This scenario is evident in the bacterial and inorganic ion analysis, see Figures 8.19 and 8.20 respectively.

Indicator bacterial concentration scenarios (Log_{10} CFU/100ml – total coliforms, *E.coli* and enterococci)

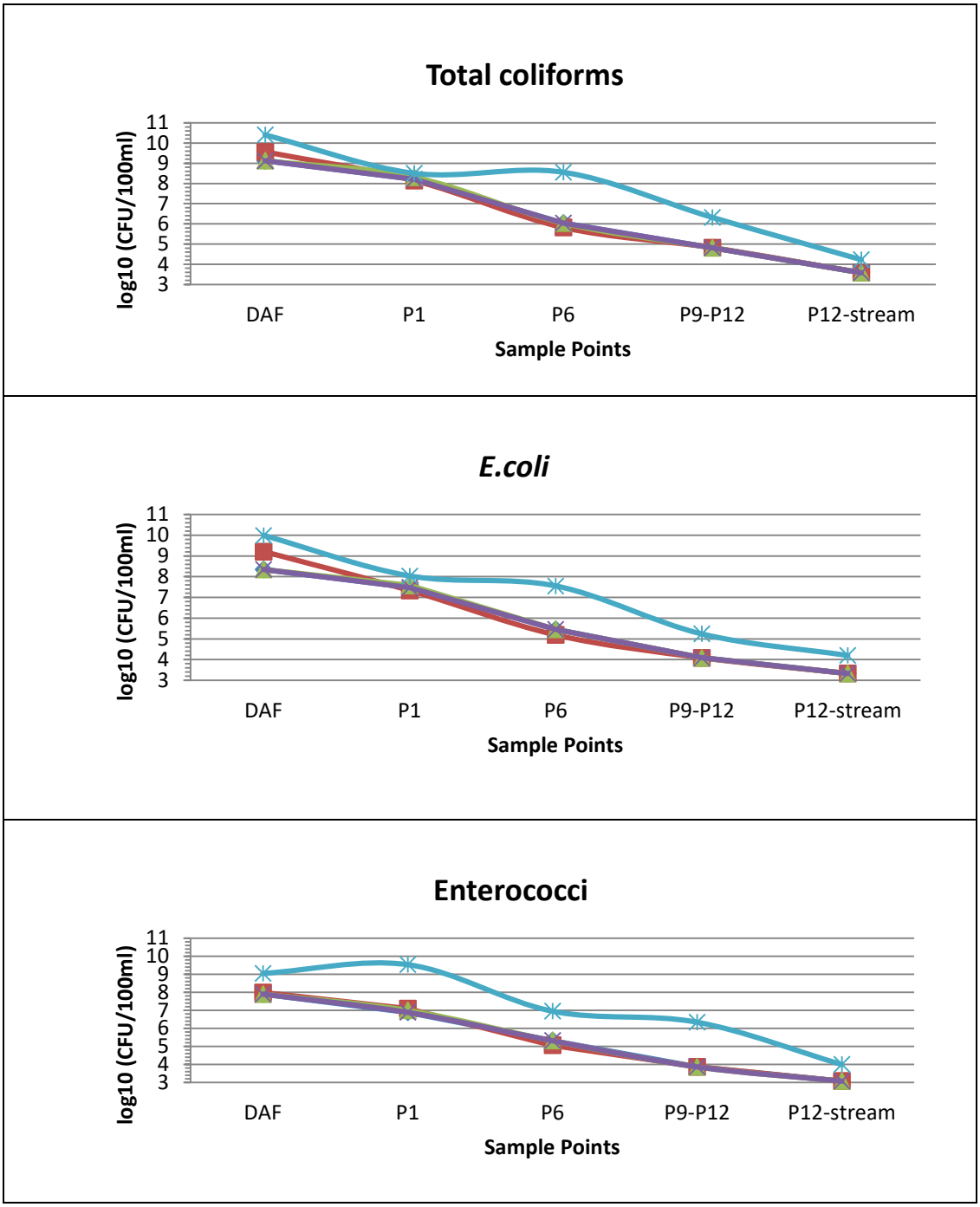


Figure 8.27 Indicator bacteria scenarios

Figures 8.27 depicts five possible scenarios in relation to total coliforms, *E.coli* and enterococci loadings. One common trend with respect to the wetlands bacterial populations, is the potential that pond 6 may incur changes (increases) to these variables, in particular the total coliforms and *E.coli* counts.

Inorganic ion scenarios (nitrate, phosphate, ammonia and sulphate)

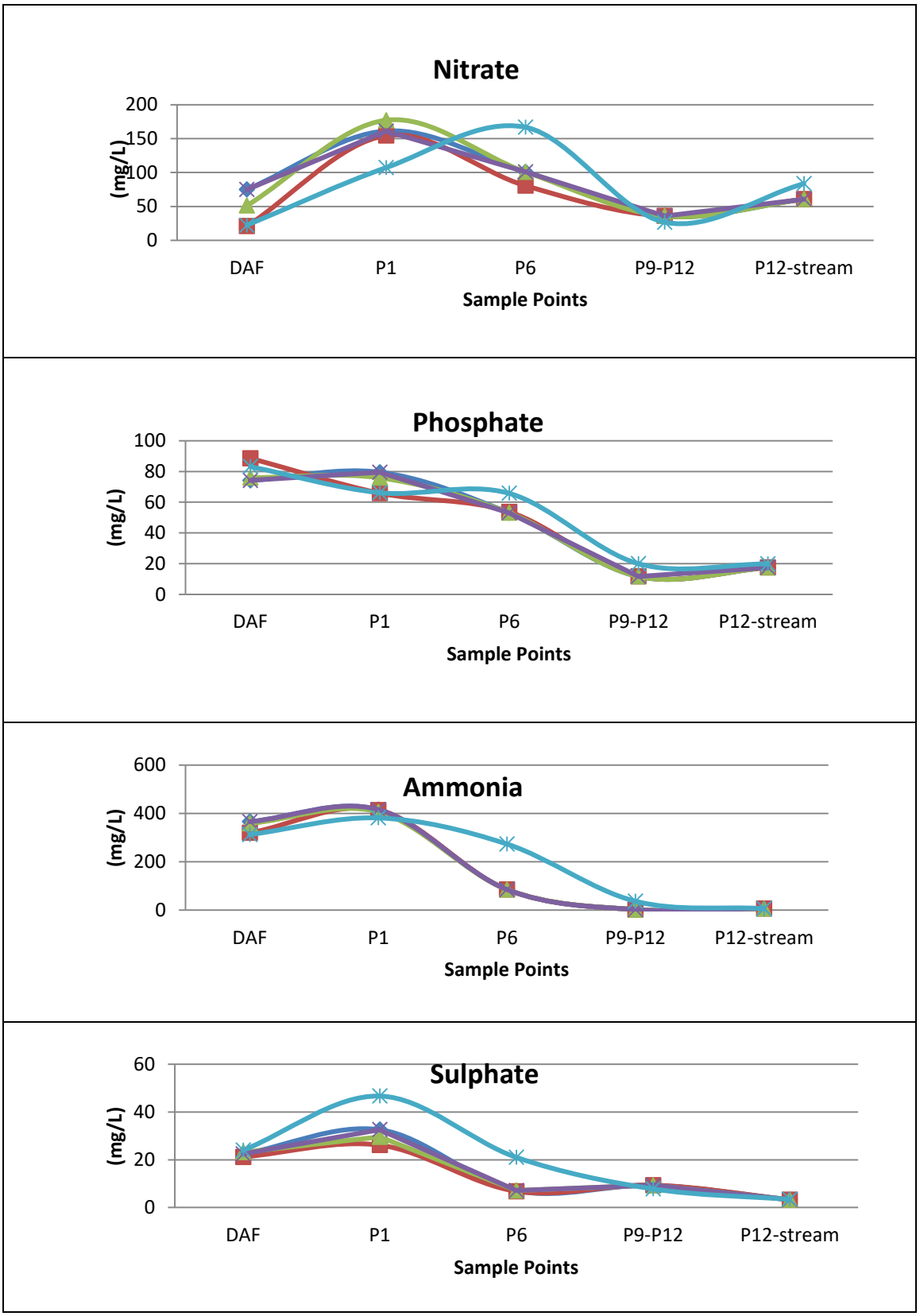


Figure 8.28 depicts five possible scenarios in relation to inorganic ion loadings. Both pond 1 and pond 6 reveal potential increases may occur at these points in the wetland.

8.18 K-means and two-stage cluster analysis

K-means are among the widely used partitioning algorithm (data partitioning using *K*-means reduces the data into specific subsets or in this case clusters i.e. the algorithm partitions *n* observations into *k* clusters in which each observation is a member of the cluster with the nearest mean). It is an unsupervised clustering algorithm that is simple and robust. But one drawback with this algorithm is the operator needs to assign the number of clusters and initial centroids before assigning input data (Canetta *et a*l, 2005). SPSS package version 22.0 was used to conduct the *K*-means clustering application. It was decided to assign five clusters for analysis on the dataset for both *K*-means and two stage hierarchical cluster analysis, as SOMine® retrieved five distinct clusters within the wetland. The *K*-means results are shows in Table 8.5. Based on the results reported in this table, eight variables used were found significant at alpha =0.01.

Table 8.5: ANOVA results comparing cluster means by *K*-means analysis

Variable	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
pH	0.697	4	0.359	12	1.941	0.168
Redox Potential	90,288.71	4	852.557	12	105.9	0.000***
Water Temp	51.706	4	14.446	12	3.579	0.038
Air Temp	22.24	4	36.883	12	0.603	0.668
Conductivity	13.54	4	0.426	12	31.754	0.000***
Dissolved oxygen	25.304	4	11.779	12	2.148	0.137
BOD	10,092,258.83	4	12,596.81	12	801.18	0.000***
Turbidity	891,966.40	4	24,661.38	12	36.169	0.000***
Log10T.Coliforms	16.614	4	1.717	12	9.675	0.001
Log10E.Coli	12.191	4	1.244	12	9.804	0.001
Log10Enterococci	13.726	4	2.399	12	5.722	0.008
Water Depth	0.341	4	0.026	12	13.371	0.000***
Fluoride	634.594	4	43.522	12	14.581	0.000***
Chloride	246,573.94	4	6,558.09	12	37.598	0.000***
Nitrate	11,660.04	4	2,869.30	12	4.064	0.026
Nitrite	388.63	4	110.359	12	3.522	0.04
Phosphate	3,414.95	4	348.052	12	9.812	0.001
Sulphate	300.089	4	137.149	12	2.188	0.132
Ammonia	125,751.38	4	1,893.98	12	66.395	0.000***

*** *p <0.01*

Table 8.5: shows that eight significant variables; redox potential, conductivity, BOD, turbidity, water depth, fluoride, chloride and ammonia.

Table 8.6: *K*-means cluster centre analysis

Initial Cluster Centers

	Cluster				
	1	2	3	4	5
pH	7.0	7.4	7.2	6.5	7.3
RedoxPot	-230.0	85.0	-222.0	-230.0	-312.0
WaterTemp	16.4	3.8	12.8	13.9	3.8
AirTemp	17.5	3.5	9.1	17.5	3.2
Conductivity	4.80	.71	3.51	4.10	3.89
Dissolvedoxygen	0.80	7.87	0.34	0.48	0.64
BOD	4350	4	3725	1140	2101
Turbidity	1567	15	905	1219	256
Log10T.Coliforms	9.15	3.39	9.06	8.33	9.49
Log10E.Coli	8.98	3.26	8.41	5.93	7.34
Log10Enterococci	9.11	2.71	7.91	10.56	6.06
WaterDepth	1.50	1.07	1.50	0.63	0.83
Flouride	22.45	.00	11.90	5.96	2.36
chloride	770.44	14.55	177.20	611.52	288.50
Nitrate	16.72	35.36	241.30	3.29	71.20
Nitrite	0.00	0.00	24.66	0.00	2.65
Phosphate	104.66	8.56	61.50	26.06	105.60
Sulphate	12.60	12.20	17.44	23.79	48.90
Ammonia	325.90	0.90	387.20	436.90	354.20

Two-stage hierarchical cluster analysis was applied to analyse sample point behaviour within the constructed wetland system. Table 8.6 summarises the two-stage analysis results. The results of *K*-means, shows no resemblance to the results depicted by SOM and two-stage hierarchical cluster analysis. The *K*-means data reveals a strong 'skew' towards the front end sampling points [DAF, pond1] and pond 12's data; see Table 8.6, where Cluster 1, 3, 4 and 5 register high wetland variable values consistent with DAF and pond 1 values and Cluster 2 registers with pond12 /stream values.

However, the two-stage hierarchical cluster analysis and SOM results are quite similar, see Table 8.7.

Table 8.7: Two- stage hierarchal clustering

		Redox Pot	Water Temp	Conductivity	Dissolved oxygen	BOD	Turbidity	Total Coliforms	*E.coli*	Enterococci	Water Depth	Nitrate	Phosphate	Sulphate	Ammonia	Air Temp
Cluster	% of combined	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
1	**23.50%**	-278.750	12.075	4.098	0.460	2971.250	534.000	9.032	8.068	7.123	1.185	140.765	81.718	25.680	382.778	6.150
2	**11.80%**	-230.000	15.150	4.450	0.640	2745.000	1393.000	8.737	7.456	9.833	1.065	10.005	65.360	18.195	381.400	17.500
3	**17.60%**	17.667	12.100	0.717	7.833	72.300	25.150	6.132	5.666	6.178	0.603	15.713	23.843	10.120	36.740	17.200
4	**23.50%**	47.500	6.800	1.000	4.405	127.300	51.400	4.549	4.177	4.627	0.815	100.900	28.595	2.863	39.888	7.525
5	**23.50%**	51.250	4.950	0.790	5.185	57.890	37.400	4.107	4.113	3.877	0.828	81.140	24.628	6.188	7.153	4.900
Combined	100.00%	-66.294	9.524	2.035	3.822	1078.390	314.860	6.272	5.726	5.924	0.897	79.904	43.648	12.098	152.488	9.465

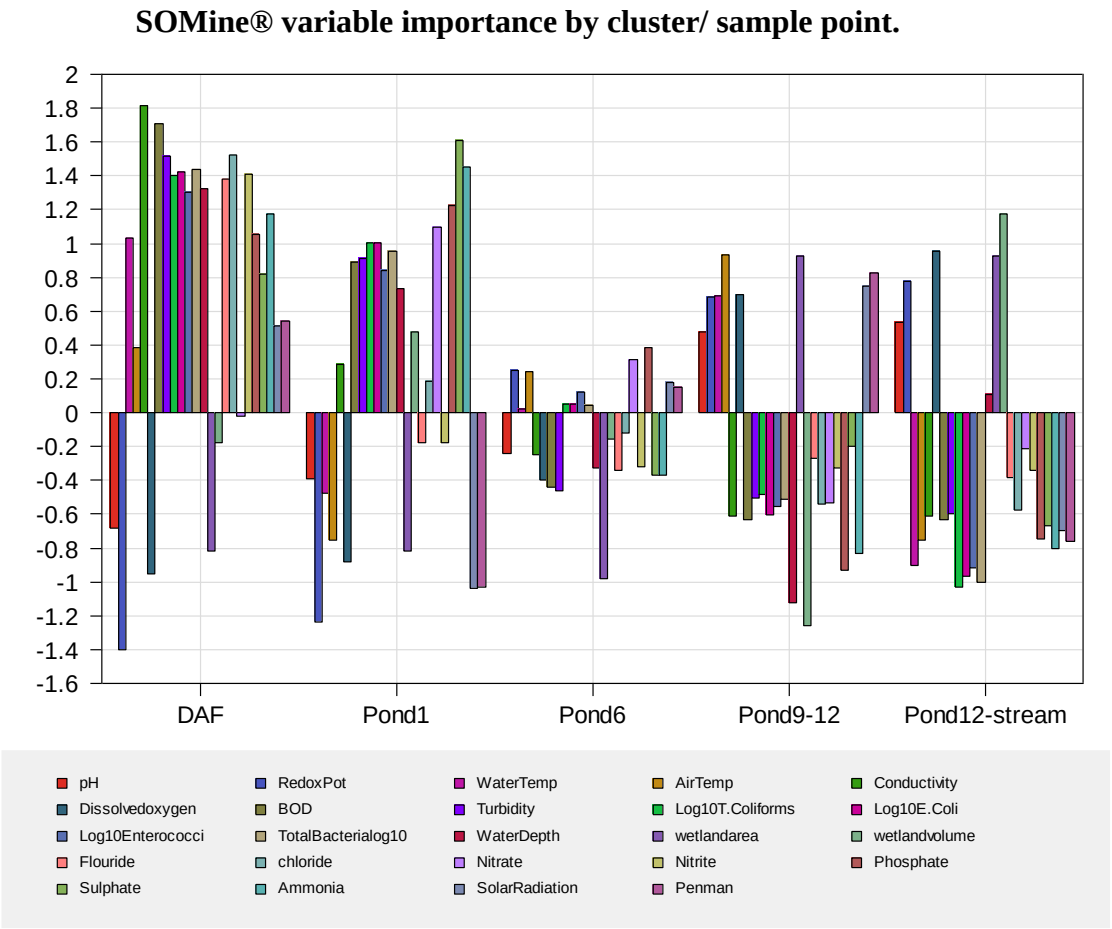


Figure 8.37. SOMine® variable wise importance analysis of each variable within each cluster; it can be seen that large positive variable importance within the DAF and pond 1 clusters and a transition phase in pond 6, leading to a large negative importance within pond 9-12 and pond 12-stream.

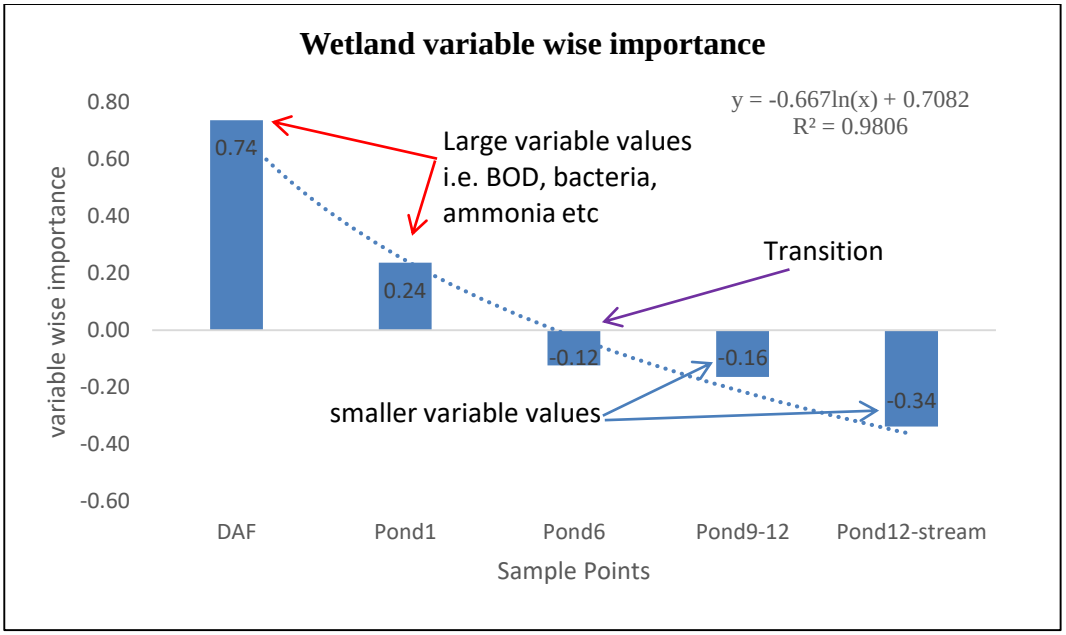


Figure 8.38 shows the wetland variable wise importance. Using the cumulative cluster mean for each sample point. The above graph shows the SOM variable-wise importance analysis by cluster i.e. sample point; representing Scenario 1, see Appendix F, where large values of BOD, bacteria, ammonia etc were seen. As they move through the wetland there is a decreasing of these values. Pond 6 is the transition point and after that point, the values further decrease though pond 9, pond 12 and into the stream.

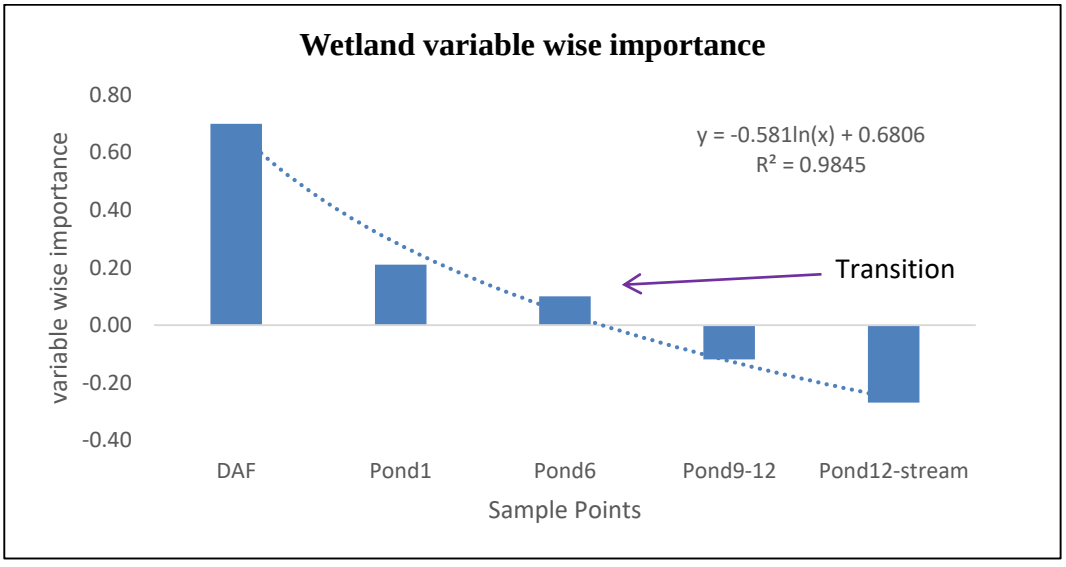


Figure 8.39 shows the wetland variable wise importance. The SOM variable-wise importance analysis by cluster, using Scenario 5 (see Appendix F), where large values of BOD, bacteria, ammonia etc were seen, and as they move through the

wetland there is a decreasing of these values. The transition point between pond 6 and pond 9-12 sample point; the values further decrease though pond 9- 12 and into the stream.

Appendix F shows the five scenarios undertaken using the SOMine software, the variables of interest are:

1. Physical Chemical [BOD, dissolved oxygen, turbidity, redox potential and conductivity]
2. Bacterial concentrations [Log_{10} total coliforms, *E.coli* and enterococci]
3. Inorganic Ions [ammonia, nitrate, phosphate and sulphate.

Table 8.8 shows the data points represented by Figures 8.38 and 8.39. It can be seen that there was a significant change at pond 6 from scenario 1 to scenario 5; where Scenario 1 tells us that the vector (variable) transition from large data variables to smaller data variables occurs at pond 6. Scenario 5 which reveals potentially large inorganic ion concentrations, at pond 6, (see Figure 8.28). This trend was observed when the 5 scenarios were realised for the bacterial loadings, (see Figure 8.27). These potential dynamics result in a delta change (0.32) from -0.12 (Scenario 1) to 0.10 (Scenario 5), therefore the transition from high to low values occurs between pond 6 and pond 9-12.

Note: Similar behaviour was revealed in the fuzzy Bayesian analysis (cause – effect and effect-cause section chapter 10 secton 3)

Table 8.8 SOMine® Scenario 1 versus Scenario 5 accumulative mean of variable wise importance by cluster

Sample Points	Scenario 1	Scenario 5
DAF	0.74	0.70
Pond1	0.24	0.21
Pond6	**-0.12**	**0.10**
Pond9-12	-0.16	-0.12
Pond12-stream	-0.34	-0.27

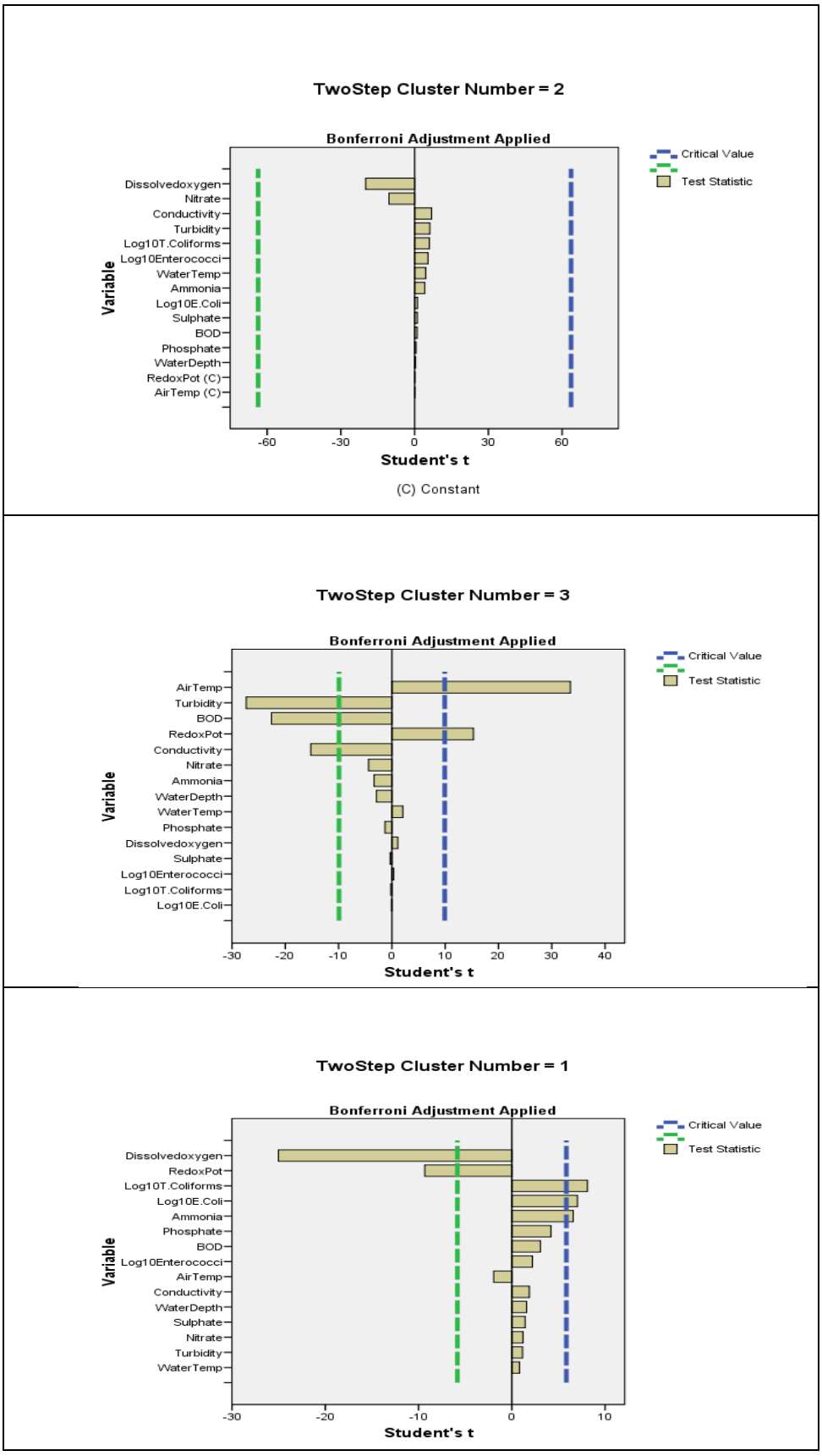


Figure 8.40. Two-step cluster analysis, clusters 2, 1 and 3.

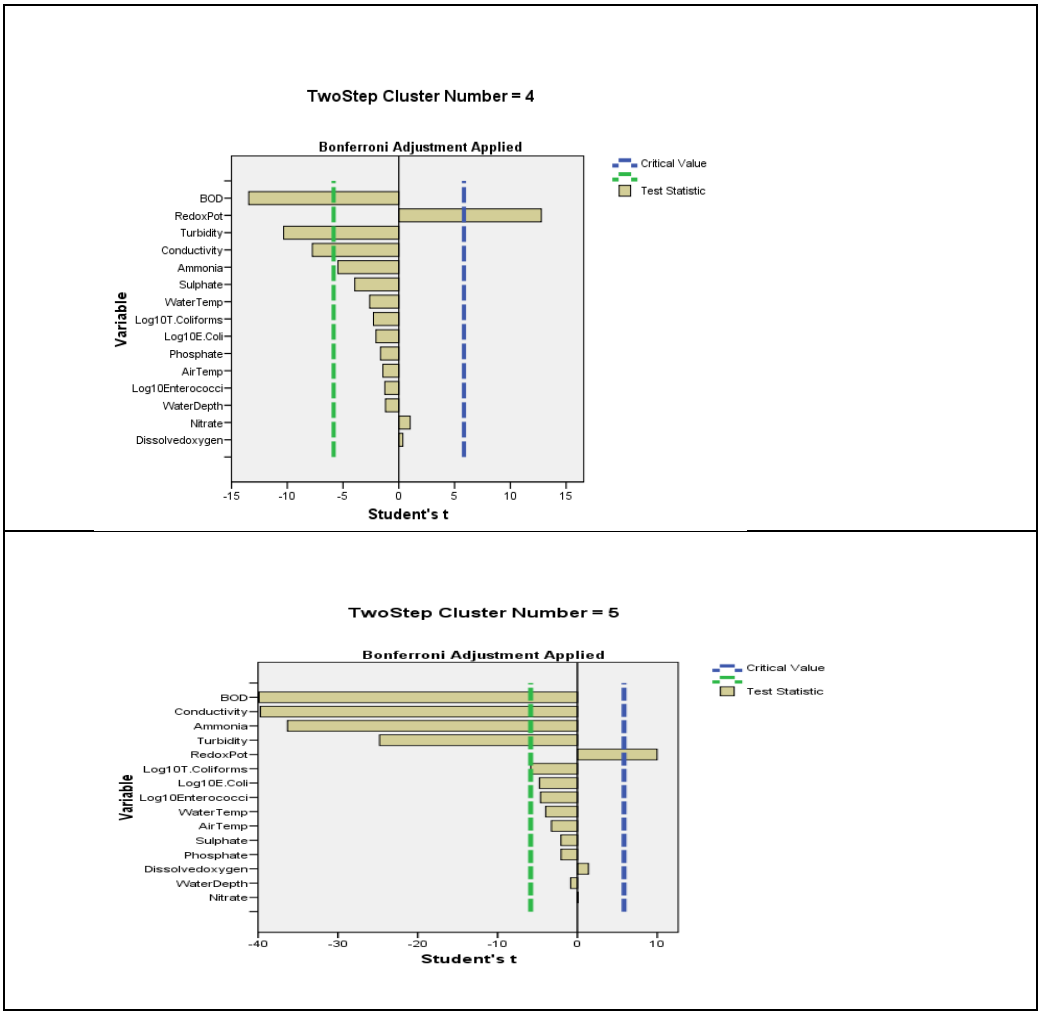


Figure 8.41. Two-step cluster analysis, clusters 4 and 5.

Figures 8.40 and 8.41 shows the two-stage hierarchical clustering analysis (SPSS 22.0) for five clusters; the coloured bars indicate 95% confidence intervals. Therefore, any variable above these coloured bars are of significance within these clusters. Within Cluster 1 variables of importance – [dissolved oxygen, redox potential, $\log_{10}$ total coliforms, $\log_{10}$ *E.coli*, and ammonia] within Cluster 2 there are no variables of importance. Within Cluster 3 – variables of importance [air temperature, BOD, turbidity, redox potential and conductivity]. Within Cluster 4 – variables of importance [BOD, redox potential, turbidity and conductivity] and within Cluster 5- variables of importance [BOD, conductivity, ammonia, turbidity, redox potential and $\log_{10}$ total coliforms].

On reviewing the SOMine® variable wise importance analysis graph as shown in Figure 8.38, which reveals the wetland by cluster and resolve the graph to wetland vectors (variables) – see Figure 8.39 – the SOMine explores the entire wetland variable wise importance.

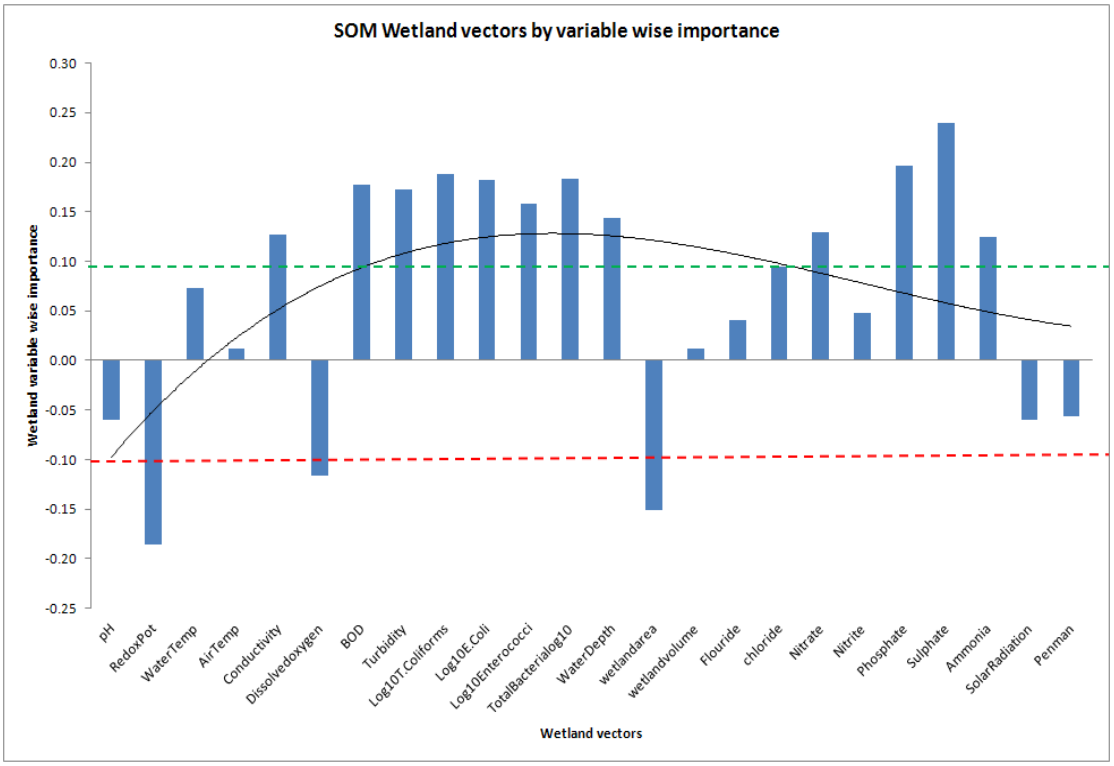


Figure 8.42 shows the wetland variables by importance. The SOMine® wetland variable wise importance of the wetland vectors. The following vectors (variables) that have importance within the wetland are: redox potential, dissolved oxygen and wetland area

(wetland design), with conductivity, BOD, turbidity, all bacterial concentrations, water depth, nitrate, phosphate, sulphate and ammonia are also of critical importance.
The main comparison between the two methods (SOMine® and two stage hierarchal cluster analysis) are that two stage hierarchal cluster analysis only reveals dissolved oxygen as variable wise importance in Cluster 1, but SOMine® reveals a complete system variable wise importance for dissolved oxygen, the importance of wetland area and water depth are also highlighted within the SOMine® wetland vector by variable wise importance. Therefore, the design of the wetland is important but these variables are not evident in the two stage hierarchal cluster analysis.

8.19 Conclusions

The SOM provides excellent visual information to identify relationships between variables, describing complex physical-chemical, biological and biochemical interactions within the wetland system. The SOM models deal very well with the non-linearity relationships that exist within the constructed wetland and provide valuable in-sight into the internal working of specific regions within the wetland system. The global shading cluster map, displaying the entire wetland variables onto a two-dimensional map, reveals the reliability contour of the data within each cluster. This dimension of utilising the global shading cluster map of SOMine® software to visualise the intrinsic reliability of each individual cluster and how each cluster segment resolves to reveal a complete visual map of the entire wetland system i.e. how the individual component parts relate to the entire system.
The SOM global shading cluster graph as shown in Figure 8.24 is similar to the fuzzy Bayesian reliability analysis graphs as shown in chapter 10 section 3, but SOM reveals far more visual detail, as can be viewed from Figure 8.23.

From reviewing the different graphs emanating from the self-organising maps software, the presence of different scenarios becomes evident, because self-organising maps are unsupervised modelling strategies, they can reveal different behaviours within the wetland system; the majority of the graphs reveal a tendency for the wetland variables to shift or drift from pond 1 to pond 6.

Pond 1 has a tendency to be a reservoir for inorganic ions, such as ammonia, nitrate, phosphate and sulphate; from the fuzzy Bayesian analysis, only the ammonia concentration was evident as having a higher value than the DAF, but the SOM analysis reveals by unsupervised cluster analysis that other inorganic salts reside at much higher levels in pond 1 than comparison to the DAF plant.

The comparison of different cluster analysis methodologies has shown that *K*-means algorithm elucidates a front-end skew or bias towards the DAF and pond 1 sample points, such as redox potential and dissolved oxygen, but the algorithm doesn't capture or cluster for all relevant vectors effectively.
Whereas the two-stage hierarchal cluster analysis does effectively capture the vectors behaviours and the cluster frequencies are similar in behaviour to SOMine, but the two-stage hierarchal cluster analysis individual clustering variable wise importance produces specific cluster important variables, whereas SOMine can elucidate the entire importance of that variable within the context of the wetland, not just the cluster.

A limitation of SOM was that every SOM was different each time the initial dataset was challenged using the "Batch-SOM" learning algorithm and one possible solution would be to introduce bootstrapping techniques to formalise the clusters and reduce data drift (Mostafa, 2010). But one could argue that the fact that every data run was different provides valuable information to different scenarios that might not otherwise be realised using conventional supervised techniques.

Future SOM improvements are already *in-situ*, such as integrating fuzzy and genetic algorithms (GA) with SOM Models. Srinvas *et al.*, (2008) used fuzzy clustering algorithms in conjunction with SOM to improve flood frequency analysis and Kuo *et al.*, (2006) used genetic *K*-means algorithm in partnership with SOM to reveal behaviours in electronic commerce.

Chapter 9. Coefficient of Reliability (COR) and Cronbach's Alpha.

The use coefficient of reliability (COR) analysis of the bacterial concentrations in a constructed wetland treating abattoir waste.

9.1 Summary

Reliability analysis based on coefficient of variation (CV) and coefficient of reliability (COR) was further supplemented by utilising Cronbach's alpha from SPSS v.22 as a validation method. These mathematical tools were used to determine the reliability in terms of the bacterial effluent concentrations (*E.coli*, enterococci and total coliforms) within the wetland system with monthly sampling at specific sampling locations (dissolved air floatation, pond 1, pond 6, pond 9, pond 12 and stream) within the wetland. For the purposes of this study we used *E.coli*, enterococci and total coliforms bacterial counts enumerated by the Idexx Quanti-Tray 2000 method with Colilert and Enterolert. The consistency of the components of the system; stream = 0.73, pond 6 = 0.70, pond 1 = 0. 67, pond 9 = 0.62, DAF =0. 60 and finally pond 12 = 0. 47, with ponds 12 COR indicative of 'unacceptable' consistency in treatment. The overall system consistency = 0.64 which indicates a 'questionable' performance of the constructed wetland in removing the indicator bacteria. From the COR analysis, the maximum average bacterial output from the DAF plant should be 10^7 CFU/100ml, for the DAF plant to maintain good internal consistency. Using Pearson's correlations on the CV and COR data, demonstrates the relationships between the individual sample points; with poor correlations between ponds 9 and 12 and pond 12 and the stream even though these ponds, in particular pond 12 wastewaters outputs into a local stream. The DAF plant and pond 1 has very poor correlations, even though the output wastewater from the DAF plant enters the constructed wetland into pond 1, indicating a disconnect in the treatment consistency between these sample points. The overall analysis indicates potential issues exists at the front of the wetland (DAF, pond 1) and towards the back of the wetland (pond 9, pond 12 and the stream).

9.2 Introduction

The reliability of a system can be defined as the probability of achieving adequate performance for a specified period of time under specific conditions. In terms of the performance of a constructed wetland it will be completely reliable if the process performance response has no failure, that is, the limits established by the targets or environmental legislation are not violated. Targets are enforced by either the local authority and/or by the Environmental Protection Agency (EPA) under the Integrated Pollution Prevention Control (IPPC) license for specified industrial and agricultural activities listed in the first schedule to the Environmental Protection Agency Acts, 1991 to 2007. Since the report is dealing with abattoir wastewater the IPPC license will include the BAT (Best Available Technology) Guidance Note for the slaughtering sector EPA (2008). The guidance note reports the BAT Associated Emission Limit Values (ELV) for discharges to water, where BOD_5 ,COD, Total Nitrogen (as N) and Total Phosphorus (as P) are reported as percentage removal values, however there is no reference to bacterial removal within the ELV even though we are dealing with slaughter waste and that removal efficiencies are generally inlet-outlet concentrations, which can lead to biased results as it has been established that efficiencies increase with increasing inlet concentrations (Vymazal, 2002). Those who instigate the building of constructed wetlands and the operators who maintain the wetland systems, mostly do so for economic reasons, not necessarily because they are convinced of the performance and reliability but because of their low operational and maintenance costs and there exists a level of apprehension as to their reliability as reported by Cooper (1990).

The purpose of this chapter was to provide an approach to evaluate the reliability of a FWS constructed wetland using bacterial loadings within the Irish context of wetland design and operations.

A methodology developed by Niku et al, (1979a) was used to determine the coefficients of reliability (COR), for three different bacterial concentrations (*E.coli*, enterococci and total coliforms) within the wetland. It should be noted that the constructed wetland under review is of an active and passive design. The active design component is the dissolved air floatation (DAF) system which clarifies the wastewater by removing suspended matter such as fats, oil and greases (FOG). The DAF effluent then enters the constructed wetland which is a passive treatment i.e. passive systems have been developed, so as they

do not require continuous chemical inputs and take advantage of naturally occurring chemical and biological processes to cleanse contaminated waters.

Because of numerous uncertainties underlying the design and operation of a constructed wetland system, the risk of failure is always present and the constructed wetland should have a designed capability to accept extreme bacterial loadings. The use of the reliability analysis in this chapter can be utilised to locate the best reliability and the corresponding bacterial concentrations for each of the sampling points within the entire constructed wetland and therefore inform as to the optimal design bacterial concentration.

A mean value of bacterial concentration should be used to guarantee an effluent concentration less than a standard with a certain reliability level. Niku *et al*., (1979a) developed a coefficient of reliability (COR) that relates mean constituent values (i.e. design or operational value) to the standards that must be achieved on a probability basis. This methodology has been used and recommended by Metcalf and Eddy (2003) and used or consulted by several authors in the last 25 years (Quek *et al*., 1995; Etnier *et al*., 2005; Gupta and Shrivastava 2006) and Qiang *et al*., 2012.

The key concepts associated with the reliability analysis are coefficient of variation (CV) and COR; coefficient of variation relates to the dispersion of the data in a dataset around the mean, it presents the ratio of the standard deviation to the mean. COR is a test of consistency or accuracy within a data set and is the correlation between responses measured at the different points in time.

A constructed wetland should be designed in order to accommodate the expected variability of its bacterial loadings (based on CV values from a similar wetland, if they exist). Secondly, for an existing CW, operational practices should be such that the resulting bacterial loadings, with its mean concentration and variability (CV value), could lead to the required percentage of compliance with the standard as set out by the IPPC license. Finally, from the regulators view point, if real CV and COR values from constructed wetlands are known, sensible and affordable discharge standards can be specified.

This approach has been taken in this study for the determinations of the COR for the bacterial concentrations for different lagoons within the wetland system i.e. different CV values. Also employed is the use of Cronbachs alpha i.e. COR. Using SPSS® Cronbachs alpha to validate the CV/COR method and finally the use of proximity correlations to understand the reliability relationship between the sample points.

The importance of the study is that there is little or no research covering constructed wetlands treating high strength slaughter wastewaters within Ireland and that uncertainty is present as to the reliability of constructed wetlands within the Irish context.

9.3 Coefficient of reliability (COR) Methods

Numerous papers have been published concerning the distribution of the concentration data from wastewater treatment plants (WWTP) (most of them considering BOD and TSS) report that lognormal distribution gives good overall fit to effluent concentration values (Dean and Forsythe, 1976 a,b; Niku and Schroeder, 1981; Niku *et al*., 1979b, 1981, 1982; Berthouex and Hunter, 1981, 1983; Charles *et al*., 2005; Djeddou *et al*., 2013). With regards to constructed wetlands, reliability analysis and COR in data analysis can be attributed to the following authors (Oliveria and von Sperling., 2007; Oliveria and von Sperling, 2008; Deas and Vaughn, 2006; Oakley *et al*., 2010), in particular Oliveria and von Sperling (2008) have utilised COR and reliability analysis to understand the behaviour of different systems.

COR developed by Niku *et al*., (1979a), based on the assumed lognormality of the data can be used to estimate the reliability of the constructed wetland.

The COR relates the values of the mean design concentrations to the standard to be achieved, on the probability basis:

$$m_x = (\mathrm{COR}) X_s, \qquad \text{Eq (9.1)}$$

where m_x is the mean bacterial concentration (design or operational values) (CFU/100ml), X_s the actual bacterial concentration (CFU/100ml) and COR the coefficient of reliability.

The COR is calculated from the following equation (Niku *et al*., 1979a):

$$COR = \sqrt{CV^2 + 1}\ \exp\{-Z_{1-\alpha}\sqrt{\ln(CV^2 + 1)}\} \qquad \text{Eq (9.2)}$$

where CV is the coefficient of variation of the bacterial data (standard deviation divided by mean); α the probability of failure of meeting the standards and $Z_{1-\alpha}$ the standardised normal variate (obtained from the standard normal variate Tables).

Some selected values of the probability $1-\alpha$ and the associated percentiles $Z_{1-\alpha}$ are shown in Table 9.1. Where 1-α probability of reliability (α is the probability of failure)

Table 9.1: Values of standardised normal distributions

Cumulative probability $(1-\alpha)$ = reliability	$Z_{1-\alpha}$
99	2.326
98	2.054
95	1.645
90	1.282
80	0.842
70	0.525
60	0.253
50	0

Note that COR is expressed based on the properties of the original data and not on the logarithm of the data.

From the COR obtained from the bacterial concentrations, it is possible to determine the bacterial design concentrations. Once these values are obtained, we can evaluate the optimum bacterial concentrations for the wetland.

Niku *et al*., (1979a) developed one approach. They used the relationship between the normal and lognormal distributions with some algebraic manipulations that take into account the CV, to obtain Eq. (31).

$$Z_{(1-\alpha)} = \frac{\ln X_s - [\ln m'_x - \frac{1}{2}\ln(CV^2+1)]}{\sqrt{\ln(CV^2+1}} \qquad \text{Eq (9.3)}$$

where m'_x is the mean effluent (bacterial) concentration (actual values) (CFU/100 ml).

After calculation of the values $(1-\alpha)$, it was necessary to obtain the values corresponding to the cumulative probability of the standardised normal distribution (distribution *S*). These values were calculated by means of the function NORMSDIST in Excel, although

they are easily found in statistics books (e.g. Snedecor and Cochran, 1989; Montgomery and Runger, 1999), corresponding to the area comprised by the standardised normal curve.

9.4 Pearson's Correlation Method

Using the Pearson's correlation's, the following method was utilised within the SPSS v.22 software: Select correlations by distance, label by the sample points, and compute between cases i.e. sample points, measure similarities by Pearson's correlations, standardise the range between -1 to 1 and by case i.e. sample points on the dataset. Therefore, values closest to one, show strong correlation and poor correlation closest to minus 1. This method was employed using the CV and COR data as seen in Table 9.4 with the corresponding correlations displayed in Table 9.10.

9.5 Bacterial sampling analysis

See chapter two, material and methods section 2.2.3.

9.6 Results and Discussion

The mean bacterial concentration of the abattoir wastewater as it moves through the wetland system and the mean removal efficiencies associated with the constructed wetland are presented in Table 9.2.

In general, a large variability (CV) was noticed in the bacterial data and in the removal efficiency over the sampling period, due to the dynamic nature of loadings and the seasonal variation of the climate (temperature, rainfall, sunshine) and other factors such as pH, dissolved oxygen, turbidity, redox potential (data not shown).

The extent of the bacterial variability depends on the behaviour of the influent loading and climate factors, but also the treatment process flow of the wetland. The wetland was designed to have a long hydraulic residence time (HRT) approximately 100 days, providing a buffering capacity into the system.

Table 9.2: Mean bacterial concentration per sample point within the wetland

Sample points	Total Coliforms (CFU/100ml)	*E.coli* (CFU/100ml)	Enterococci (CFU/100ml)	Mean	Stdev	CV
DAF	9.80E+09	3.95E+09	5.67E+08	4.77E+09	4.67E+09	0.98
Pond 1	1.05E+08	8.91E+06	3.61E+08	1.58E+08	1.82E+08	1.15
Pond 6	1.67E+06	8.62E+05	4.49E+06	2.34E+06	1.90E+06	0.81
Pond 9	6.88E+05	2.28E+05	1.33E+05	3.50E+05	2.97E+05	0.85
Pond 12	2.56E+06	2.28E+03	2.38E+04	8.62E+05	1.47E+06	1.71
Stream	1.31E+04	5.84E+03	7.65E+03	8.88E+03	3.80E+03	0.43
removal efficiency*	5.87	5.83	4.87			

*removal efficiency (log unit) = $\log_{10}$ (input/output)

9.7 Interpretation of COR

In order to assist in the interpretation of the COR, Table 9.3 has been prepared for selecting levels of reliability for a wide range of CV values.

Table 9.3: Coefficient of reliability (COR) as a function of CV and % reliability level

% reliability	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	
50%	1.00	1.00	1.02	1.04	1.08	1.12	1.17	1.22	1.28	1.35	
60%	1.00	0.98	0.97	0.97	0.98	0.99	1.01	1.04	1.07	1.11	
70%	1.00	0.95	0.92	0.89	0.88	0.87	0.87	0.88	0.89	0.90	
80%	1.00	0.92	0.86	0.82	0.78	0.75	0.73	0.72	0.71	0.70	
90%	1.00	0.88	0.79	0.72	0.66	0.61	0.57	0.54	0.52	0.50	
95%	1.00	0.85	0.74	0.64	0.57	0.51	0.47	0.43	0.40	0.38	
98%	1.00	0.82	0.68	0.57	0.49	0.42	0.37	0.33	0.30	0.28	
99%	1.00	0.80	0.64	0.53	0.44	0.37	0.32	0.28	0.25	0.22	
% reliability	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**	
50%	1.41	1.56	1.72	1.89	2.06	2.24	2.69	3.16	3.64	4.12	
60%	1.15	1.23	1.32	1.42	1.52	1.62	1.89	2.15	2.42	2.69	
70%	0.91	0.95	1.00	1.04	1.10	1.15	1.29	1.43	1.57	1.70	
80%	0.70	0.71	0.72	0.73	0.75	0.77	0.82	0.88	0.94	1.00	
90%	0.49	0.47	0.45	0.44	0.44	0.44	0.44	0.45	0.46	0.48	
95%	0.36	0.33	0.31	0.30	0.29	0.28	0.27	0.26	0.26	0.26	
98%	0.26	0.22	0.20	0.19	0.17	0.17	0.15	0.14	0.13	0.13	
99%	0.20	0.17	0.15	0.14	0.13	0.12	0.10	0.09	0.09	0.08	

For example 1, using data from Table 9.2; sample point pond 12 for bacterial concentrations with a mean CV value of 1.71 and adopting a reliability level of 95% -α = 0.05 and the percentile $Z_{1-}\alpha$ and reviewing Table 9.1 for 1- α = 0.95 = 1.645, the resulting COR can be calculated using Eq. (2). The resulting COR is 0.289,

$$\text{COR} = \sqrt{1.71^2 + 1}\,.\exp\left\{-1.645\sqrt{\ln(1.71^2 + 1)}\right\} = 0.289.$$

In order to comply with the standards 95% of the time, the mean bacterial concentration should be, using Eq. (1) m_x= (COR)X_s, = 0.289 x Xs. Where X_s equals the mean bacterial concentration for pond 12 = 8.62E+05 CFU/100ml and finally m_x is the mean bacterial concentration (design or operational values) (CFU/100ml) m_x=0.289 * 8.62E+05 = 2.49E+05 CFU/100ml (0.91 log reduction after the COR is applied).

For example 2, using data from Table 9.2; sample point pond 1 for bacterial concentrations with a mean CV value of 1.51 and adopting a reliability level of 95% -α = 0.05 and the percentile $Z_{1-}\alpha$ and reviewing Table 9.1 for 1- α = 0.95 = 1.645, the resulting COR can be calculated using Eq. (2). The resulting COR is 0.302,

$$\text{COR} = \sqrt{1.51^2 + 1}\,.\exp\left\{-1.645\sqrt{\ln(1.51^2 + 1}\right\} = 0.302$$

Therefore, in order to comply with the standards 95% of the time, the mean bacterial concentration should be, using Eq. (1) m_x= (COR)X_s, = 0.302 x Xs. Where X_s equals the mean bacterial concentration for pond 1 = 1.58E+08 CFU/100ml and finally m_x is the mean bacterial concentration (design or operational values) (CFU/100ml) m_x= 0.302* 1.58E+08 = 4.77E+07 CFU/100ml. (0.94 log reduction after the COR is applied).

9.8 COR, CV and reliability values obtained from the bacterial data

The COR and the CV of the bacterial concentrations have been calculated using equation (2) and the data tabulated in Table 9.4. The reliability of the wetland per sample point, using equation 3 with a mean (system) consistency of 0.64 for the sampling period

(February 2007 – December 2007) transforming the data using the NORMDIST function in Excel® into reliability data as shown in Table 9.5.

The order of consistency from the greatest consistency to the least consistency is as follows: stream = 0.73, pond 6 = 0.70, pond 1 = 0.67, pond 9 = 0.62, DAF = 0.60, and finally pond 12 = 0.47.

The overall mean (system) consistency = 0.64, the question arises can we validate this method against another method to verify the efficacy of the CV/COR methodology.

One possible solution is the use of the Cronbach's alpha method incorporated within SPSSv.22. Two methodologies were used in the SPSS validation process (i) the use the raw bacterial data as shown Table 9.6 and (ii) the use of CV and COR data as shown in Table 9.4, the corresponding SPSS (V.22) output are shown in Table 9.8 and 9.9 respectively. The first validation method using the raw data reveals a COR of 0.654 or 65.4% and second validation method using the CV and COR data reveals a COR of 0.664 or 66.4%, which is very close in reliability to the Niku method of 0.64, see Tables 9.7 and 9.8 to review the SPSS outputs for the two validation methods. The values 0.654 and 0.644 from the two Cronbach's methods reveal a questionable internal consistency within the wetland system, for both the indicator bacteria and for the wetland dataset, see Table 9.6.

The CV and COR of the six sampling points of the wetland, Table 9.4 shows the CV and COR of the wetland system over time, with Figure 9.1 and Figure 9.2 showing the CV and COR values respectively in the form of stacked bar charts. It can be seen from Table 9.4 that the mean bacteria CV values above 1.0 are DAF, pond 1 and pond 12 (1.01, 1.07 and 1.12 respectively) indicating a large variability at these locations. The large CV at the input DAF and pond 1 is to be expected as these are the input points accepting the large bacterial loadings from the abattoir, but pond 1 CV is slightly larger than the DAF, the possibility exists that the DAF is over loading pond 1. The CV then decreases through pond 6 with a CV= 0.92, pond 9 CV = 0.78 and then CV increases to 1.22 for pond 12 the final polishing pond with the streams CV = 0.93. Indicating that the towards the back-end of the wetland system the polishing ponds are exhibiting large variations in the bacterial loadings emanating from the aerobic ponds, raising the potential issue that these ponds are having difficulty managing the bacterial concentrations.

Table 9.4: CV and COR values per sample point within the wetland per month

						CV						
	Feb	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Nov**	**Dec**	**Mean CV**	**Stdev**
DAF	0.9	0.75	1.52	0.19	0.75	0.91	1.1	1.28	1.38	1.28	1.01	0.39
Pond 1	0.79	0.83	0.27	1.59	0.83	1.73	1.48	0.87	1.48	0.87	1.07	0.47
Pond 6	1.15	0.91	0.84	1.22	0.91	1.56	0.62	0.46	1.04	0.46	0.92	0.35
Pond 9	1.58	0.75	0.34	0.83	0.75	1.53	0.67	0.43	0.53	0.43	0.78	0.44
Pond 12	0.81	1.61	0.86	**X**	**X**	1.73	1.23	**X**	0.93	0.68	1.12	0.41
stream	0.5	0.86	1.48	0.31	0.86	1.48	0.52	0.91	1.47	0.91	0.93	0.43
Mean	0.96	0.95	0.89	0.83	0.82	1.49	0.94	0.79	1.14	0.77		
Stdev	0.37	0.33	0.54	0.63	0.38	0.3	0.39	0.32	0.38	0.32		
						COR						
	Feb	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Nov**	**Dec**	**Mean COR**	**Stdev**
DAF	0.52	0.4	0.91	0.04	0.4	0.53	0.66	0.77	0.84	0.77	0.58	0.26
Pond 1	0.43	0.46	0.07	0.95	0.46	1.02	0.89	0.49	0.89	0.49	0.62	0.31
Pond 6	0.69	0.52	0.47	0.74	0.52	0.93	0.3	0.18	0.62	0.18	0.52	0.24
Pond 9	0.94	0.4	0.11	0.46	0.4	0.92	0.3	0.16	0.23	0.16	0.41	0.3
Pond 12	0.45	0.96	0.49	**X**	**X**	1.02	0.74	**X**	0.54	0.35	0.65	0.26
stream	0.21	0.48	0.89	0.09	0.48	0.89	0.23	0.52	0.89	0.52	0.52	0.29
Mean	0.54	0.54	0.49	0.46	0.45	0.89	0.52	0.42	0.67	0.41		
Stdev	0.25	0.21	0.37	0.42	0.24	0.18	0.27	0.27	0.26	0.23		

X = denotes dry pond

Table 9.5: Reliability values of sample points per month

			Reliability (1-α)					
	DAF	**Pond1**	**Pond6**	**Pond9**	**Pond12**	**Stream**	**mean (Reliability by month)**	**Stdev**
Feb	0.698	0.753	0.62	0.56	0.743	0.958	0.72	0.14
Mar	0.733	0.696	0.78	0.557	0.718	0.711	0.7	0.08
Apr	0.566	1	0.729	1	0.717	0.569	0.76	0.2
May	0.54	0.559	0.606	0.731	0	1	0.57	0.33
Jun	0.781	0.733	0.696	0.78	0	0.718	0.62	0.3
Jul	0.693	0.549	0.562	0.564	0.549	0.569	0.58	0.06
Aug	0.633	0.569	0.871	0.83	0.605	0.94	0.74	0.16
Sept	0.596	0.713	0.978	0.989	0	0.696	0.66	0.36
Nov	0.581	0.57	0.65	0.937	0.687	0.57	0.67	0.14
Dec	0.596	0.713	0.978	0.989	0.824	0.696	0.8	0.16
Mean (Reliability by sample point)	0.6	0.67	0.7	0.62	0.47	0.73	0.64	
Stdev	0.19	0.13	0.24	0.37	0.38	0.16		

Table 9.6: Cronbach's alpha internal consistency table (SPSS® v.22)

Cronbach's alpha	**Internal consistency**
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

Table 9.7: Bacteria data from the wetland [total coliforms, *E.coli* and enterococci] (based on the IDEXX Quanti-Tray 2000 method and Colilert and Enterolert Assays)

Month	Sample Point	Total Coliforms	*E.coli*	Enterococci	Month	Sample Point	Total Coliforms	*E.coli*	Enterococci
Feb	DAF	6.82E+07	3.04E+07	5.79E+06	**July**	DAF	3.89E+10	2.62E+10	1.73E+06
	Pond 1	5.47E+06	2.15E+06	1.07E+06		Pond 1	1.03E+08	2.59E+05	9.60E+03
	Pond 6	1.32E+06	3.93E+05	2.34E+04		Pond 6	7.49E+05	4.10E+04	1.30E+04
	pond 9	2.30E+05	1.45E+04	3.24E+02		pond 9	6.13E+05	3.00E+04	2.16E+04
	Pond 12	1.31+02	8.21+01	1.09E+01		Pond 12	2.53E+07	3.00E+03	4.14E+04
	Stream	4.38+02	2.84+02	1.50E+02		Stream	4.64E+04	4.00E+03	1.04E+03
March	DAF	7.55E+07	6.02E+07	7.63E+06	**August**	DAF	4.19E+10	1.12E+10	2.93E+09
	Pond 1	3.84E+06	1.73E+06	5.20E+05		Pond 1	6.17E+08	6.49E+07	1.96E+06
	Pond 6	3.05E+05	1.73E+05	1.00E+04		Pond 6	2.19E+05	6.20E+04	1.10E+05
	pond 9	1.00E+03	1.00E+03	9.70E+01		pond 9	2.01E+04	5.10E+04	9.36E+04
	Pond 12	8.70+02	3.01+01	1.20E+01		Pond 12	7.15E+04	8.10E+03	9.24E+03
	Stream	2.28+03	1.49+03	9.60E+01		Stream	3.14E+03	2.10E+03	9.70E+02
April	DAF	6.18E+08	1.72E+07	3.92E+07	**Sept**	DAF	2.33E+08	2.80E+07	2.14E+07
	Pond 1	2.90E+06	1.82E+06	1.95E+06		Pond 1	9.13E+06	4.30E+06	8.76E+05
	Pond 6	6.24E+04	2.95E+04	7.40E+03		Pond 6	5.59E+05	2.22E+05	3.41E+05
	pond 9	2.00E+03	2.50E+03	1.22E+03		pond 9	5.20E+04	2.03E+04	4.57E+04
	Pond 12	2.14E+05	2.00E+03	1.84E+05	***Dry***	Pond 12	0	0	0
	Stream	1.14E+04	1.10E+03	1.50E+02		Stream	4.51E+03	1.22E+03	8.90E+02
May	DAF	1.40E+09	9.50E+08	1.29E+09	**Nov**	DAF	1.45E+10	9.30E+08	1.34E+09
	Pond 1	2.13E+08	8.59E+05	3.60E+09		Pond 1	8.51E+07	7.10E+06	2.26E+06
	Pond 6	1.03E+07	1.07E+06	4.38E+07		Pond 6	2.36E+06	6.23E+06	2.21E+05
	pond 9	5.83E+06	2.07E+06	1.10E+06		pond 9	7.82E+04	7.33E+04	2.27E+04
Dry	Pond 12	0	0	0		Pond 12	1.16E+03	9.53E+03	3.34E+03
	Stream	4.14E+04	4.50E+04	7.10E+04		Stream	1.51E+04	5.22E+02	1.19E+03
June	DAF	7.55E+07	6.02E+07	7.63E+06	**Dec**	DAF	2.33E+08	2.80E+07	2.14E+07
	Pond 1	3.84E+06	1.73E+06	5.20E+05		Pond 1	9.13E+06	4.30E+06	8.76E+05
	Pond 6	3.05E+05	1.73E+05	1.00E+04		Pond 6	5.59E+05	2.22E+05	3.41E+05
	pond 9	1.00E+03	1.00E+03	9.70E+01		pond 9	5.20E+04	2.03E+04	4.57E+04
Dry	Pond 12	0	0	0		Pond 12	1.41E+02	2.37E+01	9.56E+01
	Stream	2.28+03	1.49+03	9.60E+01		Stream	4.51E+03	1.22E+03	8.90E+02

Units (CFU/100ml)

It should be noted that low CV values and high COR values do not necessarily imply a good performance, but simply a more stable operation condition.

Low COR values imply the need to reduce the bacterial loadings into the wetland system, since these will be obtained by multiplying the COR by the actual bacterial values to achieve the recommended mean bacterial concentration (design or operational values) see equation (1) (Oliveira and Von Sperling., 2007; Oliveria and von Sperling., 2008). Reviewing Table 9.6, the Cronbach's alpha metric table, indicates that the alpha values can be evaluated into six categories, from excellent to poor.

The CV and COR values are shown in Figure 9.1 and Figure 9.2 respectively, CV values are high and COR are low, implying unstable conditions within the wetland. This behaviour maybe expected as the system is a passive treatment system (i.e. dependent upon weather/climatic conditions). It should be noted that the DAF which is the only part of the wetland system which actively treats the abattoir wastewater prior to entering the wetland also exhibits this unstable behaviour, giving rise to two potential outcomes (i) the slaughtering of animals is not continuous but a pulsed process, with varying slaughter rates and/or (ii) the DAF plant is having difficulty processing the abattoir wastewater emanating from the slaughter process.

Figure 9.1 shows the collective bacteria CV values, as shown in Table 9.4 for the wetland system from February 2007 to December 2008, with high CV values in July 2007 and October 2007 indicating consistency decrease see Figure 9.3. Superimposed on Figure 9.3 are red (poor) amber (questionable) and green (acceptable) lines to distinguish the different COR zones, as per Table 9.6. The same pattern is observed in Figure 9.2 regarding the COR values, except here the COR values are approximately half those of the CV values. There is a similarity in the graph patterns between the CV and COR values even though the results are very different as can be seen from Table 9.4.

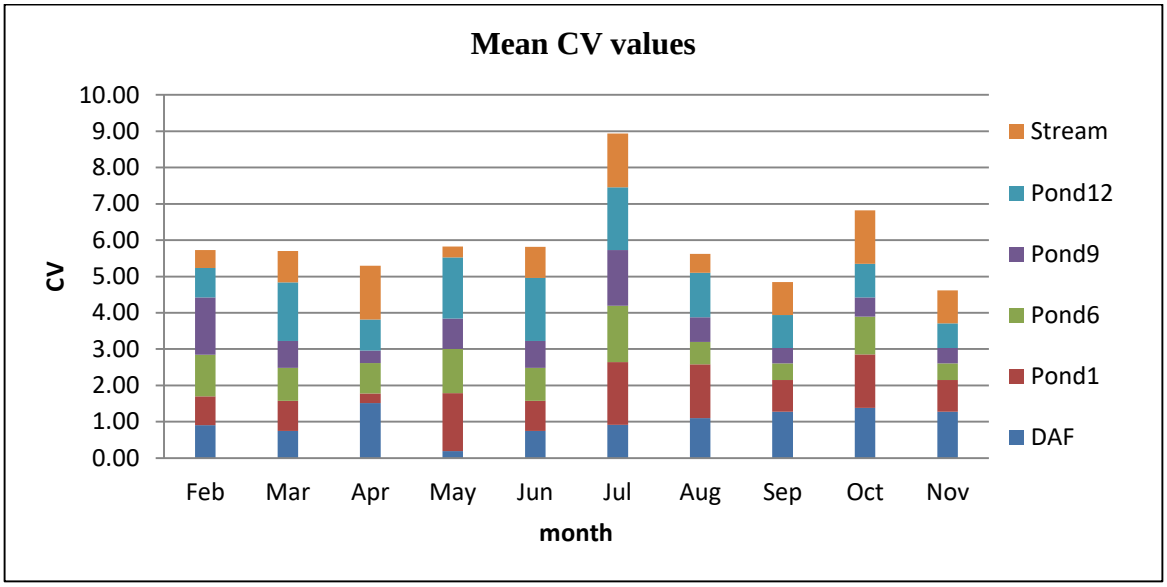


Figure 9.1: The cumulative CV values per month. The collective mean of the CV values of the sample points over the sampling period from February 2007 to November 2007.

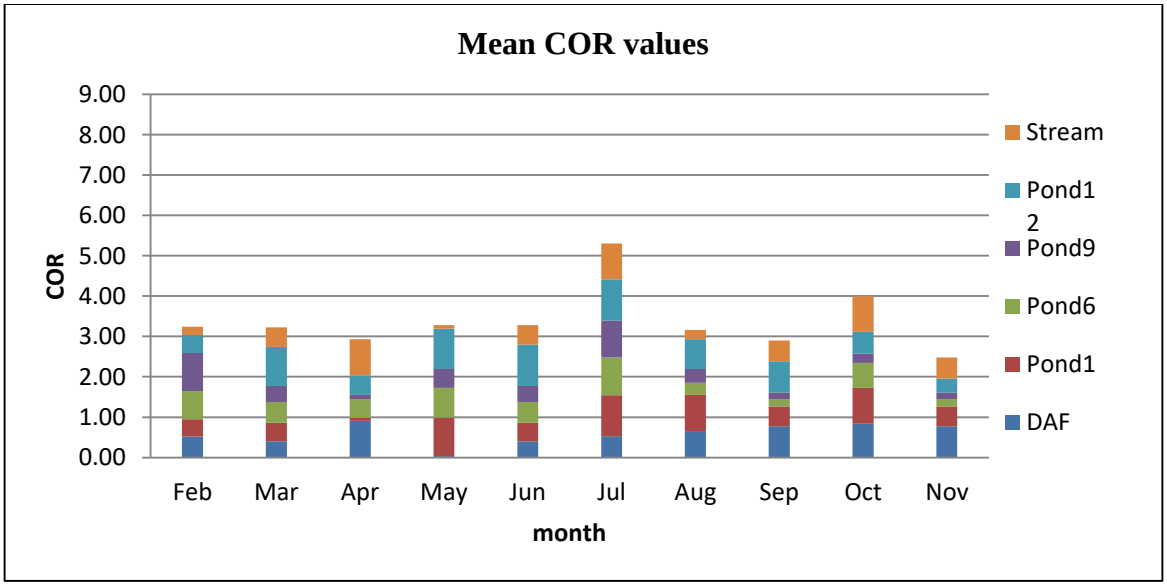


Figure 9.2: The collective mean of the COR values. The COR values of the sample points over the sampling period from February 2007 to November 2007.

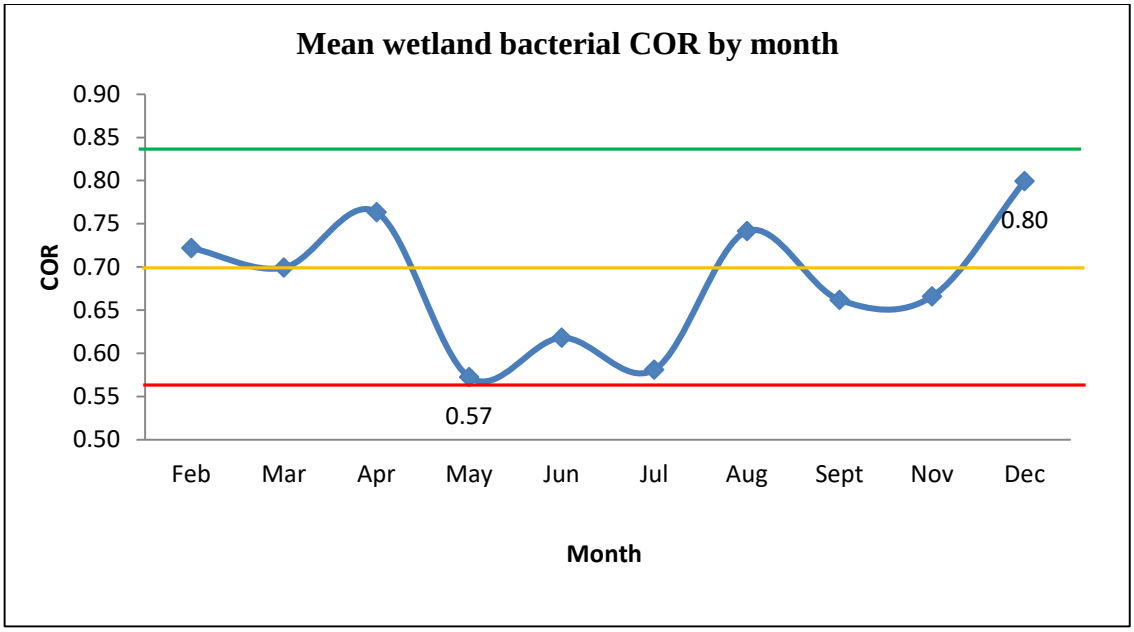


Figure 9.3: The mean wetland bacterial COR by month; the overall wetland reliabilities from February 2007 to December 2007. The wetlands upper and lower reliability ranges are between 0.57 (May 2007 – poor COR) and 0.8 (December 2007 – acceptable COR) respectively.

Table 9.8 Cronbach's alpha indicator bacterial concentrations.

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.654	.766	3

Cronbach's Alpha for the levels of three defined bacterial concentrations (*E.coli*, total coliforms and enterococci as derived from the Idexx Quanti-tray 2000 Method)

Table 9.9 Cronbach's alpha all wetland data.

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.664	.709	20

Cronbach's Alpha for the entire wetland (using the CV and COR derived from the Niku Method).

9.9 Constructed wetland and Pearson's correlation

It was decided to carry out a Pearson's correlation matrix analysis using SPSS v.22 using the coefficient of variation (CV) and coefficient of reliability (COR) data for each of the sample points, see Table 9.10. The analysis revealed the following positive correlations:

DAF: no correlation within the wetland.
Pond 1: medium to strong correlation with ponds 6, 9 and 12.
Pond 6: very strong correlation with Pond 9 and strong correlation with ponds 1 and 12
Pond 9: reciprocates the very strong correlation with pond 6 and medium correlation with ponds 1 and 12.
Pond 12: strong correlations with ponds 1 and 6 and medium correlation with pond 9.
Stream: reciprocates the strong correlation with the DAF.

See Figure 9.2 which shows the Pearson's correlation data represented on a Bar-chart.

Table 9.10: Pearson's correlation of the CV and COR values per sample point

	Correlation between Vectors of Values					
	1:DAF	2:Pond1	3:Pond6	4:Pond9	5:Pond12	6:Stream
1:DAF	1	0.011	-0.02	-0.082	-0.176	0.74
2:Pond1	0.011	1	0.618	0.512	0.61	0.178
3:Pond6	-0.02	0.618	1	0.819	0.673	0.399
4:Pond9	-0.08	0.512	0.819	1	0.502	0.099
5:Pond12	-0.18	0.61	0.673	0.502	1	0.193
6:Stream	0.74	0.178	0.399	0.099	0.193	1

The Pearson correlation analysis reveals correlations between the sample points based on the CV and COR data within the wetland system.

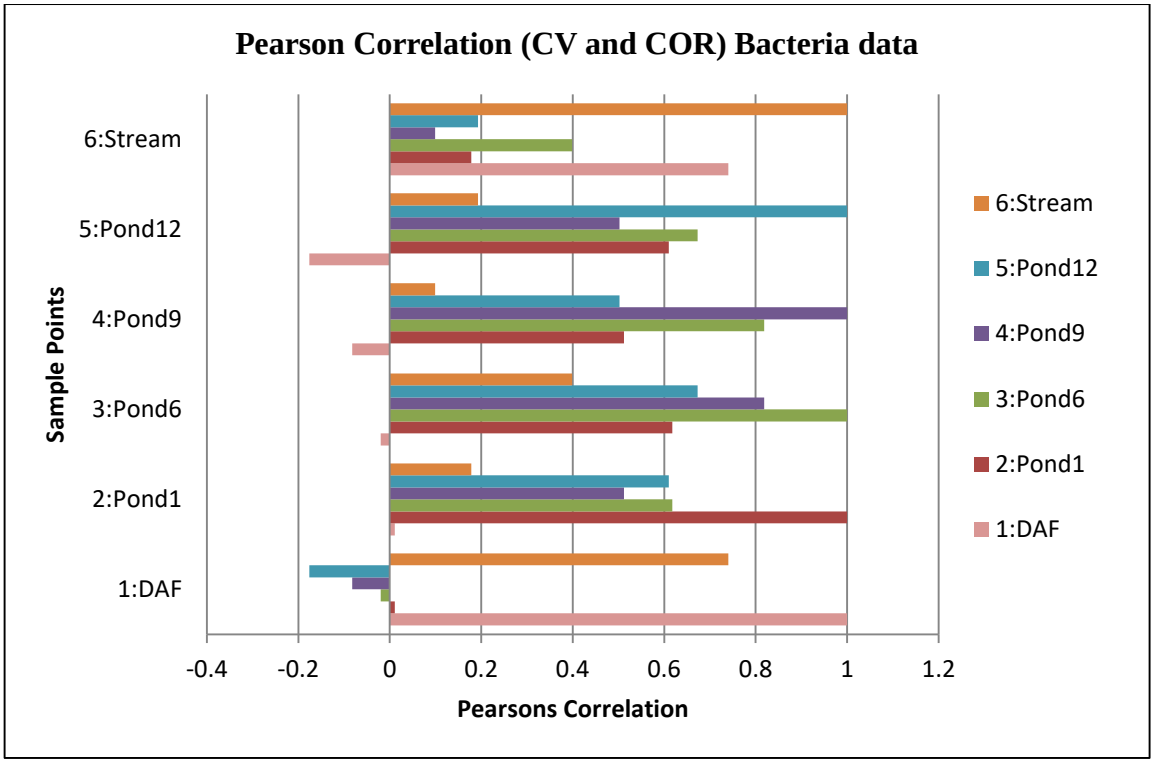


Figure 9.4: The Pearson correlation of the CV and COR of the bacteria data.

Note: The DAF and the stream display strong correlations, see Figure 9.4 (1: DAF). The potential cause for this correlation was that both the DAF and stream employ “mixing” processes. The DAF plant uses diffusers to keep the abbatoir waste in constant motion, to minimise clogging. The stream has natural mixing. The CV and COR bacterial correlations are potential elucidating to the mixing processes, within the DAF and the stream, i.e. the CV and COR for both the DAF and stream are similar.

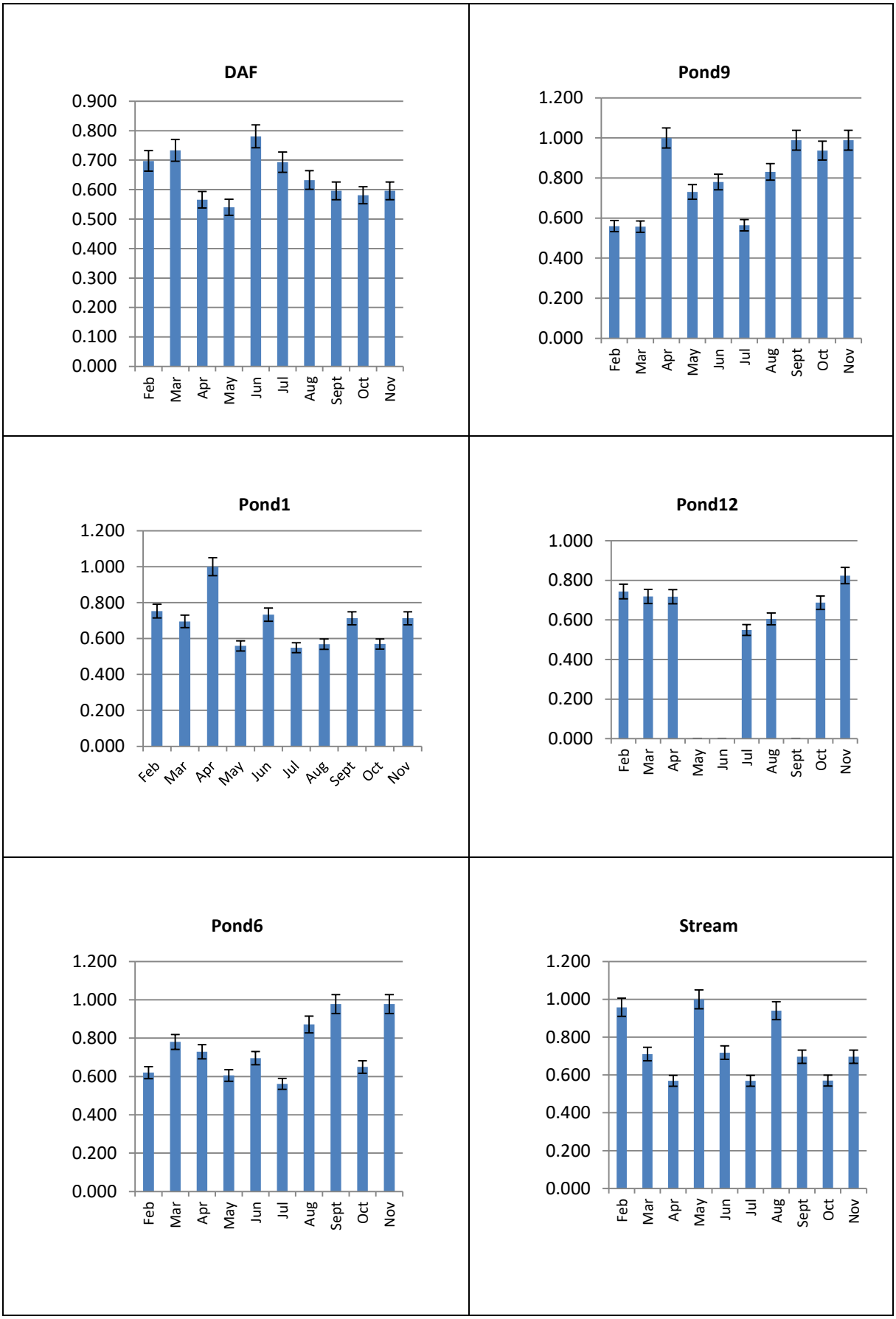


Figure 9.5: The COR for each sample point, with standard error over the sample period, **(**Norm (S) distribution of Equation (9.3) data).

Figure 9.5 shows all six sample point's reliability with 5% error bars. There is varying consistency or COR in the wetland regarding reliability; the system was constantly in flux, due its passive treatment of the waste. Figure 9.5 shows the mean wetland reliability for each sample point over time (months). The graphs indicate the variation of the reliability with little similarity between any sample points except for the summer months. As we progress through the sampling period the performance of the wetland decreases to below 0.6 during the late spring into the summer months and performance rises as we progress into early autumn and into early winter.

Pond 12 shows 3 months (summer) with no reliability due to, little or no water present in the lagoon. The dissolved air flotation plant (DAF); the only part of the effluent treatment that is an active treatment system incorporated into the wetland system exhibits poor reliability from May to November

9.10 Reliability and COR and CV

The highest consistency was achieved in the months of April, August and November, with values of 0.76, 0.74 and 0.80 respectively this corresponds to ponds 1, 6 and 9 and in particular pond 6 having stable consistencies throughout the sampling period. Pond 6 is the mid-point pond of the wetland system and has high macrophyte coverage.

9.11 Conclusions

The purpose of this COR analysis is to generate data that can be utilised by wetland designers and operators in evaluating constructed wetland treatment process based on and monthly sampling of specific ponds over a period of time as stated in the introduction, bias with regard to the reported removal values of the input to output (Vymazal, 2002) and public uncertainty as to the reliability of constructed wetland to treat effluent waste effectively exists (Cooper, 1990). It is important to highlight that current Irish BAT for Slaughtering waste doesn't mention any reference to bacterial counts (EPA, 2008). The Niku (1979a) method verifies the performance of the individual components of the constructed wetland and highlights (using the Pearson's correlation analysis of the CV and COR) where potential issues exist. For instance, the DAF to pond 1 and pond 9 and in particular pond 12 to the stream, identified as having problems, both

using the Pearson's correlation analysis, (see Table 9.8) and using the "mean CV "data, (see Table 9. 4). Optimum consistencies to bacterial loadings for the wetland can also be elucidated by finding the best COR values within each sampling point and cross-checking these values with the corresponding bacterial loadings. The DAF plant which actively treats the waste prior to entering the constructed wetland has good consistencies of 0.733 and 0.781 for March and June respectively (as per Table 9.5) and the corresponding bacterial loadings for these months are between 10^7 and 10^6 CFU/100ml as per Table 9.7, with poor internal consistencies of 0.546 and 0.54 for April and May respectively (as per Table 9.5). The corresponding bacterial loadings for these months are between 10^8 and 10^7 and 10^9 and 10^8 CFU/100ml respectively. For the month of August, the consistencies increases, this occurrence may be attributed to the removal of 200 tonnes of digested sludge from the anaerobic ponds at the end of July. Resulting in an increase in overall consistency in August, leading to an increase in performance of the wetland system, see Figure 9.3 which shows the improvement in thr CW. To improve performance and reliability of the DAF and subsequently the wetland, the operator needs to ensure the DAF plant maintains an average bacterial output value of 10^7 CFU/100ml, this can be achieved by either increasing the air flow into the DAF Plant and/or the introduction of ferric dosing[10] to improve flocculation of the effluent. The above method also highlights a procedure to develop discharge standards for best management practice (BMP) for the constructed wetland. The order of consistency from best to least in treatment of indicator bacteria within the constructed wetland: pond 6 = 0.7, pond 1 = 0.67, pond 9 = 0.62, DAF = 0.6, pond 12 = 0.47, see Table 9.5. The value = 0.47 for pond 12 indicates an unacceptable COR value.

The use of Cronbach's Alpha using SPSS® v.22 to verify the COR and CV values on the raw bacterial loading data adds further credence to the approach used in this study. The evaluation of COR and CV values for independent processes within the wetland, identifies which components are exhibiting seasonal difficulties. The fact that the wetland is also unstable with high CV and low COR values attest to the fact that passive treatment and active treatments system vary in functional behaviours. The potential exists that understanding bacterial processes within wetland systems, may provide an overall wetland system performance as attested to when the Cronbach's Alpha using SPSS®

[10] Ferric dosing is the application of iron based compounds to bind particulates within the treatment process, making their removal easier and improveing treatment efficiency.

v.22, was analysed using only the three bacterial concentrations, versus twenty wetland variables with values of 0.654 and 0.664 respectively, see Tables 9.8 and 9.9.

Reviewing Table 9.6, the overall review of the wetland systems is that it is questionable in its treatment of abattoir wastewater from the point of view of bacterial loadings, with COR values between 0.6 – 0.7. Future work could undertake the effect of seasonal weather factors, in particular the effect of temperature, rainfall and wetland design to identify any issues with reliabilities.

Chapter 10. Hybrid fuzzy logic and Bayesian belief networks.

Section 1. Iterative and reductive Principal Component Analysis.

10.1 Summary

Principal component analysis (PCA) was employed to investigate the dominant variables within a wetland data set comprising of different data types such as bacterial, physical – chemical, inorganic ions and wetland design. The outcome of the analysis was to identify the principal variables that best represent the variance and performance of the wetland system, whilst ensuring that noisy and multi-collinear variables are removed. A process of data reduction and iteration using PCA was used to reduce the original data set from 22 variables to 7. The final 7 variables were dissolved oxygen, conductivity, oxidation reduction (redox), biological oxygen demand (BOD), turbidity, total bacteria and water depth. On the final PCA reduction, BOD accounts the majority of the variance within the wetland system, making BOD the dominant variable within thr wetland system. The PCA technique employed highlights the potential for reasonably reducing the number of wetland variables for the purposes of monitoring, cost and ease of use for stakeholders. The final seven variables were used in section 2 (Fuzzy Indices – F-IND) to create an overall performance metric for each sampled pond within the wetland system.

10.2 Chapter overview

This chapter contain three sections. Involving the construction of a multi-layered complex modelling process. Utilising different statistical and mathematical modelling strategies that are finally combined into an integrated fuzzy Bayesian network, whereby wetland data can be entered into the model to reveal the resilience of the wetland system. Resilience of a complex system is the mutilfaceted capability of the wetland system to absorb, adapt and recover from internal and external shocks.

- **Section 1**: Wetland dataset was diverse with >20 variables of data varying from wetland design, bacteria, climate, physical-chemical etc. A process was therefore required to reduce the variable diversity down but retain the variability and importance of the wetland effects. The main methods employed to reduce and classify the wetland data set was using principal component analysis (a data reduction technique) which elucidates the significant variables within the wetland data set.

- **Section 2**: After the variable selection – A fuzzy logic tool to was utilised to "fuzzify" these iterative and reductive PCA variables from point 1, into environmental indices for each sampling point within the wetland. Different variables are combined, with a single outcome (indices) that represents the culmination of the inputted variables, i.e. the different variables are reduced to one common value, representing the corresponding inputted values. The software tool used in this process was called F-IND.

- **Section 3**: UniNet software was used to construct the BBN. The BBN utilises the uncertainties and sensitivities within the analysis. Two BBNs were developed and constructed to determine the resilience of the wetland system. The resilience was determined using the correlation ratio of the wetland system.

10.3 Introduction

The fundamental idea originated from a telecom's churn paper (Verbeke *et al*., 2012). Within the mobile telecommunication market predicting customers' attrition (churn) to competitors is big business. Verbeke *et al*., (2012) assessed multiple advanced modelling strategies using medium to large data sets and multiple attributes (variables) and concluded that between six to eight variables are all that are required to form a churn prediction model with high accuracy. In relation to social network analysis (Verbraken *et al*., 2014) make similar claims that knowing only a small number of social network connections of potential customers are required and other variables such as age, gender are not required. Using the Verbeke *et al*., (2012) as a template it was decided to use Principal component analysis (PCA) as a data reduction technique to find the six to eight variables that dominate the wetland dataset. Principal component analysis has been used in constructed wetlands to statistically analyse the water quality of integrated constructed wetlands treating farmyard waste (Scholz *et al*., 2007), to design equations for BOD and COD removal prediction (Tsihrintzis *et al*., 2008), to evaluate metabolic diversity of microbial communities (Salmo *et al*., 2009), the verify that wetlands can separate heavy metals from sludge (Peruzzi *et al*., 2011) and to assess the performance of *Phragmites* in treating diazo dye (Ferreira *et al*., 2014).

The aim of this chapter was to utilise principal component analysis (PCA) to reduce and iterate the initial wetland dataset of 22 variables. After the 3rd and final PCA run the remaining 7 variables accounted for the majority of the variance within the wetland system. The remaining variables were then ascribed to monitor the wetland system.

10.4 Data methodology

The data used in this chapter was based on the wetland data obtained through the sampling of the following locations within the wetland system: dissolved air floatation chamber (DAF), pond 1, pond 6, pond 9 and pond 12. The data is multivariate in nature, comprising of; physical-chemical, bacteria counts, wetland design and climate data.

10.5 Theory of principal component analysis (PCA)

Principal component analysis (PCA) is a mathematical method for powerful, unsupervised, patterns recognition technique that can be utilised as a tool for analysing, classifying and reducing the dimensionality of a multivariate numerical data set (Chen and Zhu, 2004). It is also a technique which is widely used to reduce the noise or the dimensionality in a data set, while retaining the most variance (Jolliffe, 2002).

The method uses principal component analysis with Promax rotation, which is a form of oblique rotation which allows factors to correlate, in comparison to the orthogonal rotation where the factors are uncorrelated. The oblique rotation provides slightly improved information versus the standard orthogonal rotation (Costello and Osborne, 2005).

10.6 Wetland data reduction and classification

Principal component analysis will reduce and classify the wetland variables and derive the dominant variables within the wetlands data set using Principal component analysis (PCA). The software of choice for all the analysis was SPSS version.22®.

10.7 Data pre-treatment

Prior to running the PCA, a linear interpolation (LINT) algorithm was activated in the SPSS software, because of the sparseness of the inorganic ion data such as ammonia, sulphate, nitrite etc. By running the protocol, the software populated most of the missing data within the inorganic ions data set. This is a known curve fitting method where missing data can be filled in by interpolating between two known data points and fitting in the missing data point between these two known data points (Oliveria, 2006).

10.8 Statistical analysis

Additional statistical protocols were activated within the principal components algorithm such as Bartlett's sphericity test (BTS) and the Kaiser-Meyer-Olkin (KMO) test, Eigen values greater than 1.0, Promax rotation, and PCA factor loading values were set to greater than or equal to 0.4. Three iterations occur where the original data set of 22 variables was reduced to 12 variables and then finally to 7 variables. Using the following criteria (1) if the wetland variables in the data set display strong inter-correlation between variables potentially indicating multicollinearity effects in the data set and (2) variables that have low principal component values within the first principal component column and have values within other principal component columns. In other words, remove noisy data and reduce the ammount of components (Batina *et al*., 2012). With respect to the multicollinearity contained in the dataset, this is especially true of data that is time-series based, where predictor variables are non-independent tending towards multicollinearity. The following filtering criterion was initiated within the principal components analysis (1) the number of extracted factors with eigenvalues greater than 1 (Kaiser, 1960), also known as the Kaiser rule; and (2) as a rule-of-thumb, a factor loading less than 0.4 is categorized as weak while a loading greater than 0.6 is categorized as strong. Thus, the factor loading was selected at the threshold 0.4 (Boyer *et al*., 1997; Lambrakis *et al*., 2004; Gazzaz *et al*., 2012). Therefore, factor loadings with coefficient values below 0.4 were removed from the analysis.

Two important internal tests were employed to evaluate the wetland data set as to whether the data is of interest regarding the implementation of principal component analysis, i.e. whether the chosen variables are factorable. The indicators are Bartlett's sphericity test (BTS) and the Kaiser-Meyer-Olkin (KMO) test. The KMO is a measure of sampling adequacy, which produces a value between 0 and 1 (Table 10.1). The value of the KMO is associated with the relationship of common variance under analysis, while the value of the BTS determines the correlation matrix is an identify matrix i.e. the null hypothesis (Ocal *et al*., 2007).

Table 10.1: Interpretation of KMO (Kaiser-Meyer-Olkin) value (Kaiser, 1974)

KMO value	Interpretation
0.9 - 1.0	Marvellous
0.8 - 0.9	Meritorious
0.7 - 0.8	Middling
0.6 - 0.7	Mediocre
0.5 - 0.6	Miserable
0.0 - 0.5	Unacceptable

10.9 Results (Twentytwo variables 1st PCA iteration)

The values of BTS and KMO are shown in Table 10.2. The value of KMO is 0.675, which indicates that the degree of common variance is mediocre. Thus the primary components evaluated by the PCA will account for a reasonable amount of variation. In BTS, the chi-square and significance are 1058.113 and 0.000, respectively, therefore, the correlation matrix is not an identity matrix and the analysis of the variables is valid. Another method of viewing the BTS was the wetland data met the sphericity assumption since the chi-square (χ^2) value of 1058.113 ($p < 0.0001$ and $df = 231$), therefore PCA will allow for variability in the data with less than the original number of variables (McNeil *et al*., 2005). The total variance for the 22 variable PCA run is shown in Table 10.3, where ammonia, BOD, total bacteria, total coliforms and redox potential account for 79.67 % of the variance within the wetland, with these five variables all with Eigen values greater than 1 as per the Kaiser rule.

Table 10.2: KMO and Bartlett's Test (22 wetland variables)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.675
Bartlett's Test of Sphericity	Approx. Chi-Square	1058.113
	df	231
	Sig.	0.000

Table 10.2, shows the KMO measure of 0.675, with 231 degrees of freedom and P< 0.0001 for the original 22 wetland variables entered into the PCA analysis.

Table 10.3: Total Variance explained (22 variables)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
Ammonia	9.584	43.565	43.565	9.584	43.565	43.565	8.227
BOD	3.023	13.742	57.307	3.023	13.742	57.307	7.790
Log10 Total Bacteria	2.395	10.884	68.192	2.395	10.884	68.192	3.283
Log10 T_Coliforms	1.409	6.405	74.597	1.409	6.405	74.597	2.863
Redox Pot	1.118	5.083	79.679	1.118	5.083	79.679	3.008
Log10 E_Coli	0.924	4.199	83.879				
Log10 Enterococci	0.806	3.665	87.543				
Phosphate	0.594	2.699	90.242				
Sulphate	0.526	2.391	92.633				
Conductivity	0.452	2.056	94.689				
Fluoride	0.269	1.224	95.914				
Chloride	0.242	1.100	97.013				
wetland area	0.168	0.765	97.778				
Turbidity	0.121	0.551	98.329				
pH	0.097	0.443	98.772				
wetland volume	0.074	0.335	99.107				
Water Depth	0.058	0.262	99.369				
Dissolved oxygen	0.049	0.221	99.590				
Water Temp	0.036	0.165	99.755				
Air Temp	0.031	0.142	99.897				
Nitrate	0.021	0.097	99.994				
Nitrite	0.001	0.006	100.000				

The KMO measure can be improved by removing unwanted variables from the original data set as stated previously (1) where we have strong correlations between variables potentially indicating multicollinearity effects in the data set and (2) variables that have low principal component values in the first principal component column, but have values spread into the remaining 5 columns therefore noisy variables. The KMO value of 0.675 is too low, 5 components in the first PCA run indicate the potential for noisy data, see Table 10.4 and from reviewing Table 10.5a we can see high correlations between common variables such as air temperature and water temperature for example.

The following 10 wetland variables were removed prior to the second PCA analysis, thereby reducing the data set. Air and water temperature, total coliforms, *E.coli* and enterococci due to high correlations values (> 0.9), pH, wetland volume, nitrate and nitrite were removed due to noisy effects, as they have no value within PC 1, but have values spread across the other remaining principal components, 2,3, 4 and 5. Finally the wetland area was removed because the area of the wetland remains static throughout the sampling period, even though the wetland area is of importance within the scope of the analysis, the individual pond areas' will not change, there is no variance in this variable. The dissolved oxygen and water depth remain in the second round of the analysis because they have reasonable factor values in PC1 (>0.4) and PC2 (>0.6) and no high correlation (multi-collinearity) evident in the correlation matrix.

Table 10.4: Principal Component Matrix (22 wetland variables)

	Component				
	1	2	3	4	5
Ammonia	0.910				
BOD	0.865				
Log10 Total Bacteria	0.842				
Log10 T Coliforms	0.841				
Redox Pot	-0.840				
Log10 E.coli	0.836				
Log10 Enterococci	0.794				
Phosphate	0.789				
Sulphate	0.774				
Conductivity	0.764				
Fluoride	0.760				
Chloride	0.729				
wetland area	-0.656			0.457	
Turbidity	0.594			0.569	
pH		0.910			
wetland volume		0.798	-0.496		
Water Depth	0.418	0.690	-0.506		
Dissolved oxygen	-0.522	0.662			
Water Temp			0.805		
Air Temp		0.409	0.795		
Nitrate			-0.438	-0.467	
Nitrite			-0.411		0.631

Extraction Method: Principal Component Analysis.

Table 10.4: The original 22 wetland data set in a 5 component matrix table (PC1, PC2, PC3, PC4 and PC5). The variables removed or reduced from the data set are highlighted in red above. In total 10 variables are removed from the original data set containing 22 variables.

Table 10.5a Correlation Matrix of wetland data (22 variables)

	pH	Redox Pot	Water Temp	Air Temp	Conductivity	Dissolved oxygen	BOD	Turbidity	Log10 T_Coliforms	Log10 E_Coli	Log10 Enterococci	Log10 Total Bacteria	Water Depth	Area	volume
pH	1.000	0.084	0.446	0.466	0.236	0.458	0.026	0.020	0.087	0.179	0.106	0.527	0.601	0.065	0.647
RedoxPot	0.084	1.000	-0.044	0.009	-0.613	0.490	-0.822	-0.602	-0.628	-0.609	-0.579	-0.595	-0.242	0.450	-0.024
WaterTemp	0.446	-0.044	1.000	0.922	0.391	0.010	0.079	0.142	0.172	0.177	0.150	0.369	-0.074	-0.055	-0.112
AirTemp	0.466	0.009	0.922	1.000	0.342	0.105	0.010	0.137	0.143	0.143	0.162	0.365	-0.025	-0.107	-0.072
Conductivity	0.236	-0.613	0.391	0.342	1.000	-0.265	0.695	0.516	0.585	0.522	0.534	0.673	0.306	-0.425	0.123
Dissolved oxygen	0.458	0.490	0.010	0.105	-0.265	1.000	-0.458	-0.311	-0.530	-0.492	-0.482	-0.256	0.207	0.567	0.481
BOD	0.026	-0.822	0.079	0.010	0.695	-0.458	1.000	0.526	0.669	0.653	0.554	0.672	0.327	-0.458	0.107
Turbidity	0.020	-0.602	0.142	0.137	0.516	-0.311	0.526	1.000	0.391	0.326	0.542	0.465	0.131	-0.241	0.008
Log10 T_Coliforms	0.087	-0.628	0.172	0.143	0.585	-0.530	0.669	0.391	1.000	0.925	0.919	0.615	0.222	-0.546	-0.036
Log10 E.coli	0.179	-0.609	0.177	0.143	0.522	-0.492	0.653	0.326	0.925	1.000	0.859	0.665	0.283	-0.541	0.023
Log10 Enterococci	0.106	-0.579	0.150	0.162	0.534	-0.482	0.554	0.542	0.919	0.859	1.000	0.611	0.210	-0.488	-0.017
Log10 Total Bacteria	0.527	-0.595	0.369	0.365	0.673	-0.256	0.672	0.465	0.615	0.665	0.611	1.000	0.475	-0.524	0.243
Water Depth	0.601	-0.242	-0.074	-0.025	0.306	0.207	0.327	0.131	0.222	0.283	0.210	0.475	1.000	-0.342	0.892
wetland area	0.065	0.450	-0.055	-0.107	-0.425	0.567	-0.458	-0.241	-0.546	-0.541	-0.488	-0.524	-0.342	1.000	0.085
wetland volume	0.647	-0.024	-0.112	-0.072	0.123	0.481	0.107	0.008	-0.036	0.023	-0.017	0.243	0.892	0.085	1.000
Fluoride	0.023	-0.764	0.094	0.110	0.523	-0.386	0.666	0.552	0.560	0.490	0.584	0.589	0.282	-0.441	0.080
Chloride	0.244	-0.555	0.182	0.180	0.540	-0.292	0.567	0.550	0.592	0.587	0.661	0.743	0.216	-0.309	0.067
Nitrate	0.180	-0.274	-0.006	-0.068	0.016	-0.176	0.261	0.008	0.204	0.313	0.129	0.227	0.471	-0.346	0.318
Nitrite	0.034	-0.408	-0.116	-0.106	0.089	-0.160	0.352	0.050	0.261	0.305	0.184	0.261	0.214	-0.192	0.105
Phosphate	0.245	-0.526	0.186	0.148	0.604	-0.410	0.639	0.159	0.662	0.692	0.504	0.715	0.413	-0.641	0.127
Sulphate	0.060	-0.708	0.020	0.006	0.613	-0.402	0.746	0.557	0.574	0.584	0.522	0.593	0.268	-0.425	0.067
Ammonia	0.043	-0.815	0.194	0.178	0.734	-0.499	0.830	0.525	0.701	0.695	0.612	0.700	0.366	-0.672	0.062

** water temperature, air temperature, total coliforms, *E.coli* and enterococci removed due to high correlations (multi-collinearity).

Table 10.5b Correlation Matrix of wetland data (22 variables continued)

	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	Ammonia
pH	0.023	0.244	0.180	0.034	0.245	0.060	0.043
RedoxPot	-0.764	-0.555	-0.274	-0.408	-0.526	-0.708	-0.815
WaterTemp	0.094	0.182	-0.006	-0.116	0.186	0.020	0.194
AirTemp	0.110	0.180	-0.068	-0.106	0.148	0.006	0.178
Conductivity	0.523	0.540	0.016	0.089	0.604	0.613	0.734
Dissolved oxygen	-0.386	-0.292	-0.176	-0.160	-0.410	-0.402	-0.499
BOD	0.666	0.567	0.261	0.352	0.639	0.746	0.830
Turbidity	0.552	0.550	0.008	0.050	0.159	0.557	0.525
Log10 T_Coliforms	0.560	0.592	0.204	0.261	0.662	0.574	0.701
Log10 *E.coli*	0.490	0.587	0.313	0.305	0.692	0.584	0.695
Log10 Enterococci	0.584	0.661	0.129	0.184	0.504	0.522	0.612
Log10 Total Bacteria	0.589	0.743	0.227	0.261	0.715	0.593	0.700
Water Depth	0.282	0.216	0.471	0.214	0.413	0.268	0.366
wetland area	-0.441	-0.309	-0.346	-0.192	-0.641	-0.425	-0.672
wetland volume	0.080	0.067	0.318	0.105	0.127	0.067	0.062
Fluoride	1.000	0.568	0.223	0.597	0.405	0.470	0.735
Chloride	0.568	1.000	0.165	0.094	0.517	0.498	0.577
Nitrate	0.223	0.165	1.000	0.284	0.381	0.127	0.362
Nitrite	0.597	0.094	0.284	1.000	0.268	0.404	0.348
Phosphate	0.405	0.517	0.381	0.268	1.000	0.710	0.740
Sulphate	0.470	0.498	0.127	0.404	0.710	1.000	0.677
Ammonia	0.735	0.577	0.362	0.348	0.740	0.677	1.000

10.10 Results (Twelve variables 2nd PCA iteration)

In the second PCA run, the original data set of 22 variables was reduced to 12 of the original wetland variables. They are in order of Eigen value importance (BOD, redox potential, conductivity, total bacteria, chloride, turbidity, ammonia, phosphate, fluoride, water depth, sulphate and dissolved oxygen). The same PCA process was undertaken as before. The KMO value = 0.858, improving the previous score from mediocre to meritorious, see Table 10.6.

Table 10.6: KMO and Bartlett's Test (12 wetland variables)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.858
Bartlett's Test of Sphericity	Approx. Chi-Square	751.658
	df	66
	Sig.	0.000

Table 10.6. The KMO measure of 0.858 for the reduced 12 wetland variables entered into the PCA analysis.

In Table 10.7, the Eigen values on the twelve remaining components are shown with BOD and redox potential with Eigen values >1.0. These wetland variables resolve to indicate that cumulative variance of 76.29%. BOD and redox potential accounts for 76.29% of the variance within the wetland system.

Table 10.7: Total Variance explained (12 variables)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
BOD	7.883	65.689	65.689	7.883	65.689	65.689	6.932
RedoxPot	1.271	10.595	76.285	1.271	10.595	76.285	6.790
Conductivity	0.715	5.956	82.240				
Total Bacteria	0.516	4.298	86.538				
chloride	0.438	3.650	90.188				
Turbidity	0.350	2.914	93.102				
Ammonia	0.252	2.098	95.199				
Phosphate	0.205	1.706	96.906				
Fluoride	0.160	1.337	98.243				
Water Depth	0.096	0.798	99.041				
Sulphate	0.082	0.685	99.726				
Dissolved oxygen	0.033	0.274	100.000				

Table 10.8: Principal Component Matrix (12 variables)

	Component	
	1	2
BOD	0.941	
Redox Pot	-0.914	
Conductivity	0.901	
Total Bacteria	0.858	
Chloride	0.852	
Turbidity	0.837	
Ammonia	0.831	
Phosphate	0.828	
Fluoride	0.751	0.472
WaterDepth	0.694	0.539
Sulphate	0.625	
Dissolved oxygen	-0.613	0.519

Table 10.8: The second reduction and iteration PCA run with 12 wetland variables. The principal component matrix displays 2 component orders PC1 and PC2.

Prior to the 3rd and final PCA run, the inorganic ions were removed (chloride, ammonia, phosphate, fluoride and sulphate). The reason to remove or reduce these variables from the final PCA run was due the sparseness of these variables. The process of elucidating these variables from the wetland surface waters, using ion chromatography was time consuming and costly. It was previously established (Data pre-treament) to utilise a linear interpolation (LINT) algorithm to populate the missing data. Therefore, the final seven wetland variables selected are BOD, redox potential, conductivity, total bacteria, turbidity, water depth and dissolved oxygen. The resultant seven variables represent physical-chemistry (BOD, redox potential, conductivity, turbidity and dissolved oxygen), bacteria (total bacteria) and wetland design (water depth).

10.11 Results (Seven variables 3rd PCA iteration)

In the third PCA run, 7 variables remain from the original 22 wetland variables remaining. The same PCA process was repeated for the final iteration. The KMO value was found to be 0.851, (Table 10.9), indicating no discernible loss from the previous 2nd PCA run KMO value of 0.858, (Table 10.6).

Table 10.9: KMO and Bartlett's Test (7 wetland variables)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.851
Bartlett's Test of Sphericity	Approx. Chi-Square	422.198
	df	21
	Sig.	0.000

The KMO measure of 0.851 for the reduced 7 wetland variables entered into the PCA analysis.

In Table 10.10, the Eigen values on the seven remaining components are shown with only BOD with Eigen values >1.0. This wetland variable resolved to indicate that cumulative variance of 71.75%. Therefore, BOD accounts for 71.75% of the variance within the wetland system.

Table 10.10: Total Variance (7 variables)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
BOD	5.023	71.753	71.753	5.023	71.753	71.753
Redox Pot	0.883	12.620	84.373			
Conductivity	0.391	5.587	89.960			
Turbidity	0.251	3.583	93.543			
Total Bacteria	0.229	3.269	96.812			
Water Depth	0.173	2.474	99.286			
Dissolved oxygen	0.050	0.714	100.000			

In Table 10.11, after the 3rd and final PCA reduction and iteration, seven of the original wetland variables are shown, with only one principal component. The seven remaining variables are in order of importance: BOD, redox potential, conductivity, turbidity, total bacteria, water depth and dissolved oxygen.

Table 10.11: Principal Component Matrix (7 variables)

	Component
	1
BOD	0.967
Redox Potential	-0.918
Conductivity	0.889
Turbidity	0.880
Total bacteria	0.846
Water depth	0.762
Dissolved oxygen	-0.619

10.12 Results and Discussion

The fundamental idea for this method originated from Verbeke *et al*., 2012. Using mobile phone datasets to reveal between six and eight attributes (variables) are sufficient to predict customer attrition (churn). The same methodology was used in this chapter using only principal component analysis with addition of oblique testing (Promax) in conjunction with KMO and BTS. Variables that displayed multi-collinearity effects and unwanted noisy variables in larger dimensions were removed. The purpose was to reduce and classify the wetland dataset into approximately six to eight variables which could reasonably account for the wetland variance. Three reduction and iterations/classifications were implemented. Starting with the original twenty-two wetland variables, with an intermediate (2^{nd} run) twelve wetland variables and finally after the 3^{rd} PCA run, seven wetland variables remained. It should be noted the importance of four inorganic ions; chloride, ammonia, phosphate and sulphate. In the overall picture on the wetland's performance, these variables have an impact. However, it was decided to remove these inorganic ions from the remaining PCA analysis due to sporadic sampling of these variables. When the inorganic variables were removed and the PCA re-run with the remaining seven wetland variables there was no significant difference between the KMO values from the 2^{nd} PCA run (KMO = 0.858) and the 3^{rd} PCA run (KMO = 0.851). The inorganic ions are therefore not as statistically dominant in comparison to the remaining seven wetland variables.

10.13 Conclusions

The initial PCA run with all twenty-two variables was classified as mediocre with a KMO value of 0.656. With ten variables removed due to (1) multi-collinearity (air and water temperature, total coliforms, enterococci and *E.coli*) (2) static variable (wetland area) and (3) noisy and high dimensionality (pH, nitrate, nitrite and wetland volume). These ten variables were removed from the initial dataset, leaving twelve wetland variables in the analysis. The PCA was re-run and remaining variables were classified as meritorious with a KMO value of 0.858. The reduction and iteration/classification PCA process was used to define the most dominant wetland variables, from an initial data set of twenty-two wetland variables, with the 3rd and final PCA iteration revealed seven dominant wetland variables. The dominant variable resulting in approximately 72% of the wetlands' variance was BOD.

Chapter 10. (Section 2) F-IND: Fuzzy indices of the environmental variables of the constructed wetlands.

10.14 Summary

Using the variables selected in section 1, as part of the reductive and iterative principal components analysis method. F-IND (fuzzy indices) was utilised to further classify wetland environmental variables into fuzzy indices that combines different environmental variables into one resultant variable that fully represents the multivariate inputs. Fuzzy logic is well adapted in dealing with a multitude of problems such as a group of variables that are qualitative and quantitative with differing units and with variables with incomplete or ambiguous relationships. The resulting wetland indices creates a wetland efficiency index (WEI). Providing the user with a unique but common temporal-spatial index for each sample point within the wetland over the sampling period. Values of less than or equal to 50 indicate a poor wetland performance in treating the effluent, 50 to 70 will be classified as medium efficiency and values equal to or greater than 75 will be classified as high treatment efficiency. The fuzzy indexing software provides a linguistic tool in comparison to a mathematical one, to analyse multivariate complex systems and rationalise these interacting variables into a single fuzzy index value. The resulting analysis shows that ponds 6 was the most reliable treatment pond in the system throughout the analysis. The final polishing ponds 9 and 12 were the least reliable. The mean wetland efficiency value was 46.5%, indicating a poor wetland performance in treating effluent.

10.15 Introduction to fuzzy logic

Fuzzy logic is a powerful technique with which to treat heterogeneous information and can be used to develop environmental indices and decision support tools in ecosystem management (Silvert, 1997, 2000). The technique is adept at dealing with information that is affected by uncertainty and inaccuracy (Rodríguez and Peche, 2012). Multiple authors have successfully applied fuzzy logic systems to derive fuzzy indices for different ecosystems based on their corresponding environmental attributes. The fuzzy models have been used in diverse applications such as, to evaluate macro-invertebrate habitat suitability in running waters (Van Broekhoven *et al*., 2006); species distribution (Mouton *et al.*, 2009); river health (Yang and Zhao, 2009); in activated sludge control process (Zhang *et al*., 2014) and in developing a fuzzy identification model for predicting meat spoilage (Kodogiannis and Alshejari, 2014) and wetland habitat vulnerability (Talukdar and Pal, 2018).

Typically, the information required to design and implement a fuzzy environmental quality index (EQI) is provided by an expert panel. The panel is comprised of experts from all different aspects related to the site or ecosystem under consideration (Rodríguez and Peche, 2012). When dealing with these multi-disciplinary panels all aspects of individual experts must converge to an outcome that fully represents the expert panel (Gallopín, 1996). Within the context of this chapter, principal component analysis (PCA) was used to reduce and iterative and to realise the optimal wetland variables for the purposes of fuzzy indexing analysis. Ouyang (2005) used principal component analysis in the evaluation of river quality monitoring stations. Within the context of constructed wetlands Mohammadpour *et al*., (2014) used hierarchical cluster and principal component analysis to define water quality in a free-surface constructed wetland and Liu *et al*., (2015) used PCA in their analysis of storm water treatment performance of a constructed wetland.

The fuzzy Index software F-IND® was developed by Marchini *et al*., (2009) from the University of Pavia, Italy - http://robot.unipv.it/F-IND/ [Accessed 05 2014]. The analysis performed on the selected wetland data, utilises the F-IND framework software. This chapter aims to develop a wetland efficiency index (WEI), which is a versatile methodology that evaluates the wetland performance

based on the selected variables. An indices table for each sampled point within the wetland over a sampling period was produced. Therefore, the fuzzy index can be used as a black box which takes multivariable data and produces corresponding index values. A wetland efficiency index (WEI) method was developed using the F-IND framework with three categories to distinguish the grouping of the wetland efficiencies based on the wetland hydraulic efficiency index (λ) developed by Persson *et al.*, (1999). Once the method has been developed, validated and tested, environmental stakeholders such as technicians, engineers and managers have a useful technique to monitor the wetland performance.

10.16 Fuzzy indices (F-IND) - the basics

A fuzzy set F represents a category of a variable; for example, the variable 'hot' can be described in the following categories 'low' and 'high', each of them represented by a fuzzy set. In comparison to classical set theory, that allows an element either to belong or not to a set, in fuzzy set theory the membership of an element to a set F can be partial. Fuzzy sets can overlap; thus the element F can have partial membership for more than one set. The degree of overlapping between two fuzzy sets is proportional to the amount of uncertainty included in the description of a variable; in a fuzzy model, it also determines the uncertainty in the definition of the model output, see Figure 10.1.

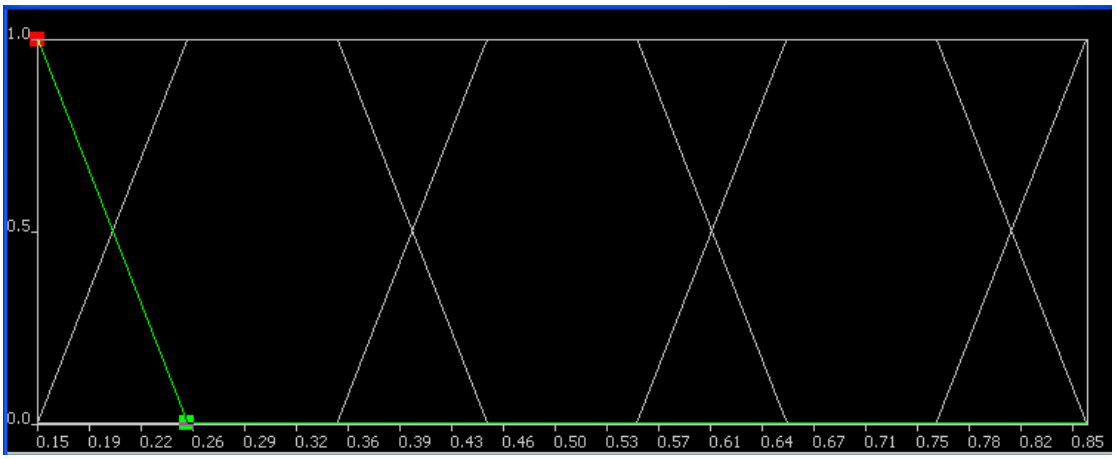


Figure 10.1: Overlapping trapezoidal fuzzy index set function

The mathematical definition of a fuzzy set F is a membership function $\mu_F(x)$ that associates any value of x of the variable X to a membership grade between 0 and 1 ($0 \leq \mu_F(x) \leq 1$). Shape and position along the x-axis are parameters of a membership function, to be chosen to be the fuzzy index creator.

Membership functions can be derived from expert knowledge from one or more experts and/or from statistical analysis of the datasets (Marchini *et al*., 2009). The membership functions from both expert and statistical processes was derived.

The membership function can exist in three shapes, bell-shaped, triangular and trapezoidal; with triangular and trapezoidal the most frequently used shapes, where bell-shaped represents non-linear interactions and both triangular and trapezoidal shapes represent linear interactions. The trapezoidal shape in this analysis using the F-IND® software was utilised. The shapes represent membership from zero (0) to full membership (1), see the y-axis in Figure 10.1.

There are two main types of fuzzy rules: Takagi-Sugeno type, where the rule conclusion is a crisp value and Mamdani type, where the rule conclusion is a fuzzy set. For the development of environmental indices, the latter is the most useful type. A Mamdani rule is in the form 'if antecedent then consequent', i.e.: 'if variable is category then output class'. In a multivariable system, for example a constructed wetland system, several variables are combined with the '*and'* operator in the antecedent condition of the 'if …then' fuzzy analysis. The majority of the fuzzy analysis was performed using the Mamdani rule with trapezoidal shape within the fuzzy index creator. All the wetland variables were subjected to the trapezoidal shape.

Fuzzy models take into account all the possible combinations of categories of different variables, for this reason, the number of variables and fuzzy rules increase exponentially with the number of variables. As per section 1 of this chapter, using the reduce and iterate PCA technique only seven variables were selected.

In F-IND the following mathematical operators are used in fuzzy indexing rule associations:

- *and*: product
- *if… then*: product
- *or*: sum

These operators ensure a smooth response to changes in the input variables by the user, thereby removing decisions about the fuzzy indexing rule associations. Therefore, the user just needs to assess the upper and lower limits of each variable i.e. the best-case and worst-case scenarios.

10.17 F-IND analysis

Initially the user selects the relevant variables that accurately describe the environmental interactions under investigation. Using the classification and reduction PCA techniques from section 1, it was found that the following wetland variables represented 100 % of the wetland's variance, with BOD contributing to approx. 72% of variance. The variables and their corresponding measurement units are BOD_5 (mg/L), redox potential (mV), dissolved oxygen (mg/L), water depth (m), turbidity (NTU), Log_{10} total bacteria (CFU/100ml) and conductivity (µS/cm). The variables all contain different units, with the exception of BOD and dissolved oxygen (mg/L), see Table 10.12. It should be noted that redox potential was not used in the fuzzy indexing analysis as the software cannot resolve negative values such as -300mV.

Table 10.12: Fuzzy index variables and their corresponding Best Case and Worst-Case thresholds

F-IND Variables	Categories of the variable [b]	Best - case	Worst-case
BOD (mg/L)	very low; low; medium; high; very high	low	high
Redox Potential (mV)[a]	very low; low; medium; high; very high	high (+ve)	low (-ve)
Conductivity (µS/cm)	very low; low; medium; high; very high	low	high
Total Bacteria[c]	very low; low; medium; high; very high	low	high
Turbidity (NTU)	very low; low; medium; high; very high	low	high
Water Depth (m)	very low, low; medium; high; very high	medium	low
Dissolved Oxygen (mg/L)	very low; low; medium; high; very high	high	low

[a] = Variable not used in the analysis

[b] = Five rule Fuzzy set

[c] = Log_{10} (CFU/100ml)

The method to generate fuzzy multivariable indices of environmental variables from the constructed wetland has been applied to evaluate the performance of the wetland over a year (February 2007 to February 2008). Using the wetland variables from Table 10.12, omitting redox potential as described previously. F-IND generates a GUI (graphical user interface) that calculates the fuzzy index value that corresponds to the input variables, see Figure 10.2. Each input variable has a range of values as determined by the user, represented on sliders. Five variable wetland fuzzy indices were generated using, water depth, dissolved oxygen, turbidity, total bacteria (bacteria log) and BOD, see Figure 10.2. The GUI will display an index value on the bottom of the GUI (grey column) and beside the index value column, five other columns represent colour changes as the sliders are activated. The index value ranges are normalised into values from 0 to 100, therefore when all the variables are selected to their minimum (worst –case) value, the resulting index value equals 0 and vice versa when all the variables sliders are selected into the maximum (best-case) value, the index value equals 100. Above the index value column, the variable sliders are located with the high and low value windows for each variable i.e. best-case and worst-case.

The user then proceeds to use the sliders to set the closest value for each variable across each sample point in the constructed wetland, therefore deriving an index value for five input variables.

10.18 Application of F-IND to constructed wetland data

The method proposed in this chapter was to generate fuzzy multivariate indices to environmental constructed wetland data, in order to describe a fuzzy index metric performance metric to each sample point. Using a set of statistically predetermined wetland variables (PCA) that were significant to all of the wetlands sampling points.

Each of the wetland variables shown in Figure 10.2, were constructed using a five rule trapezoidal fuzzy set, see Figure 10.3, representing the variable fuzzy outputs. The fuzzy index (F-IND) value was calculated as:

$$\text{F-IND}_{5\text{ classes}} = 0\mu_{\text{Very Low}} + 25\mu_{\text{Low}} + 50\mu_{\text{Medium}} + 75\mu_{\text{High}} + 100\mu_{\text{Very High}} \quad \text{Eq (10.1)}$$

Figure 10.2: Graphical user interface (GUI) the Fuzzy Index generator, five key variables within the wetland, water depth, D.O, Turbidity, Bacteria, and BOD.

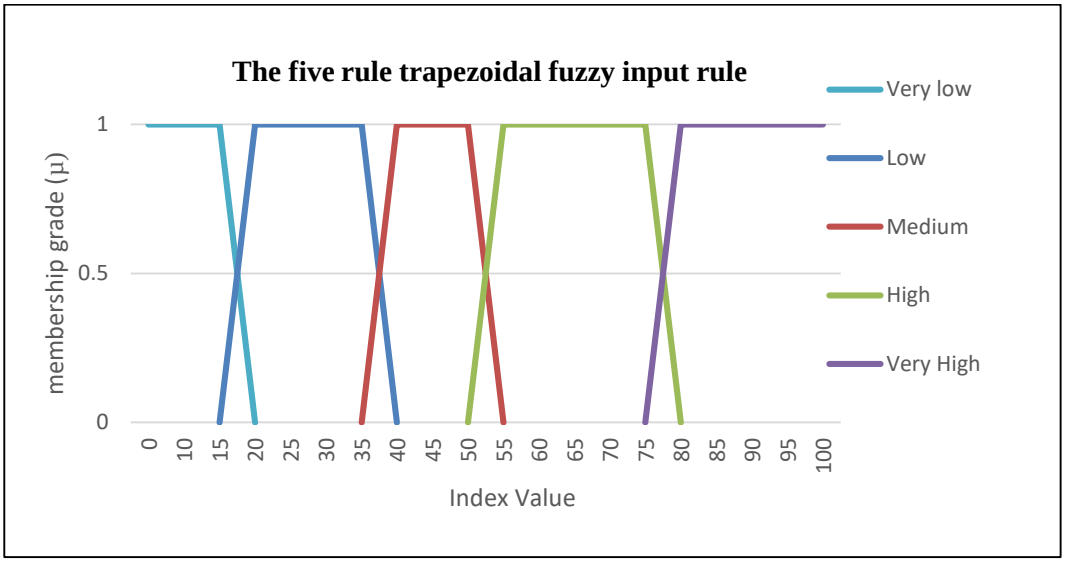


Figure 10.3: The five rule trapezoidal fuzzy input rule used to represent one of the input variables to determine the index of the wetland's performance, ranging from 0 to 100. Membership grades (µ) on the y-axis versus the fuzzy index (F-IND) value on the x-axis. The figure graphically depicts the 'five categories of the variable' from Table 10.12.

The F-IND software was applied to wetland variables and the output fuzzy Index values are shown in Table 10.13 for each sample point. The normalised index values (0 to 100) represent a performance metric for the sample points for the sample period using the preselected wetland variables.

Table 10.13: Fuzzy index values per sample point

Sampling Month / Point	DAF	Pond1	Pond6	Pond9	Pond12	Stream
Feb	10.1	25	74.69	64.2	100	100
Mar	10.1	3	75	77.6	91	70
Apr	10.1	25	68	79.2	81	71
May	10.1	21	55.55	72.5	0	67
Jun	7.6	10	75	61.6	0	67
Jul	5	50	74.05	72.6	54	72
Aug	5	25	73.83	70.8	75	67
Sept	7.6	49	65.28	65.2	0	74
Nov	5	48	59.25	75	74	72
Dec	5	25	60.33	74.6	75	99
Feb-08	10.1	25	59.58	75	94	95
Mean	7.79	27.27	67.32	71.66	58.55	77.64
Stdev	2.41	16.28	7.6	5.69	39.46	13.18
Upper (95%) CI	9.21	36.89	71.81	75.03	81.86	85.43
Mean	7.79	27.27	67.32	71.66	58.55	77.64
Lower (95%) CI	6.37	17.65	62.83	68.3	35.23	69.85

Table 10.13 shows the resultant fuzzy indices values for each of the wetland's six sampling using the F-IND software.

The fuzzy indices are shown for the eleven sampling times. The sampling period can be further subdivided into three events the mean fuzzy index value, the summer season from May 2007 to September 2007 which corresponds the lower 95% CI as seen in Figure 9.4 and the winter season from November 2007 to February 2008 which corresponds to the upper 95% CI as see in Figure 9.4, with zero performance recorded on three occasions. The three zero performance events occurred in pond 12 in the months of May 2007, June 2007 and September 2007 due to dry pond conditions in the pond at this time. Figure 10.4 displays the mean, upper and lower 95% confidence intervals values of the six sampling points.

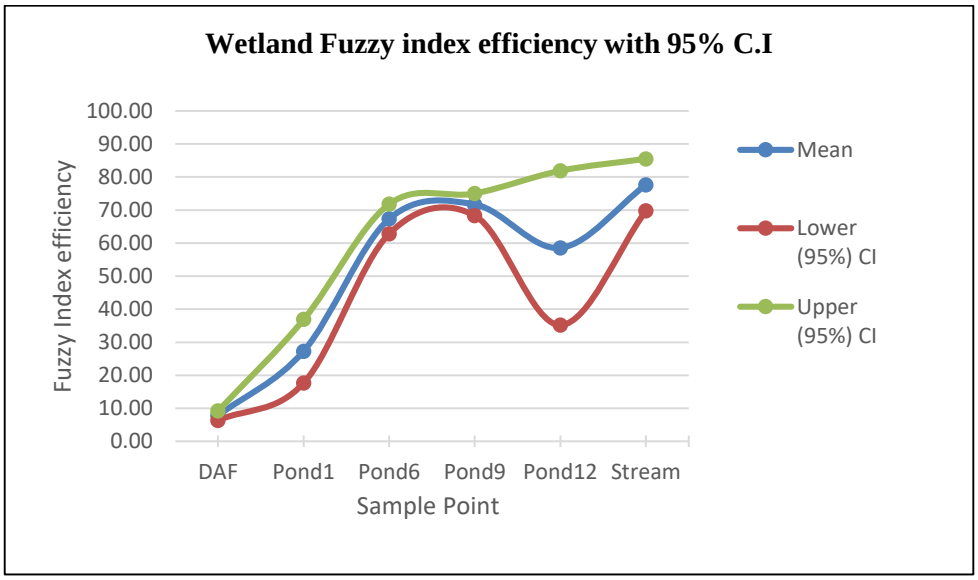


Figure 10.4: The wetland fuzzy index efficiency using the F-IND. Included are the 95% confidence intervals for each sample point.

As the effluent passes through the wetland system from the input (DAF) to the output (pond 12) into the stream, the mean performance of the system increases until ponds 6 and 9. These ponds reveal the largest mean fuzzy indices performance metric within the wetland at 67.32 and 71.66 respectively. Figure 10.4 depicts that the DAF, pond 6 and pond 9 show the close fitting of their respective confident intervals, indicating that these points are best at resolving internal and/or external perturbations during the sampling period. With respect to the DAF it has the lowest performance value with a mean fuzzy index value of 7.79, whereas both pond 6 and pond 9 have a mean fuzzy indices of 67.32 and 71.66 respectively. Therefore, from an ecosystem point of view, the wetland's reliability in treating abattoir wastewater over varying internal and /or external influences lies between these two ponds. Reliability functions has been used by Vidal *et al.*, (2011) in the performance evaluation of constructed wetlands treating waste from small communities. The authors used Weibull analysis, a probability density function (pdf) used primarily in engineering failure models. Wahl *et al.*, (2012) used reliability functions to quantify the performance of wetlands using hazard indices.

The largest deviations with respect to confidence intervals was present in ponds 1 and 12, with pond 12 showing the largest deviation within the wetland system, with an upper CI = 81.90 and a lower CI = 35.19 and a mean fuzzy index value of 58.55 (lower than both ponds 6 and 9). A performance difference of 46.71 exists in the last "polishing" pond of the wetland (pond 12). In comparison to pond 1 with an upper CI =36.89 and a lower CI = 17.65 and a mean fuzzy index value of 27.27, leaving a performance difference of 19.24. But in the case of pond 1, this pond is the first receiving pond within the wetland system. Thus the influent is anaerobic wastewater emanating from the DAF plant, therefore these low performance values are expected. Pond 12's large deviation is of concern, as it's the last receiving pond prior to the treated effluent leaving the wetland system into a local stream. The mean fuzzy index value for pond 12 was 58.55 in comparison to the local stream with a value of 77.65, as per Table 10.13. Therefore, pond 12 was of a lower performance than the stream in which the effluent was being emptied into. An ideal scenario of an 'improved' pond 12 can is shown in Figure 10.5, which a reduced deviation with respect to 95% confidence intervals and an overall improve fuzzy index performance.

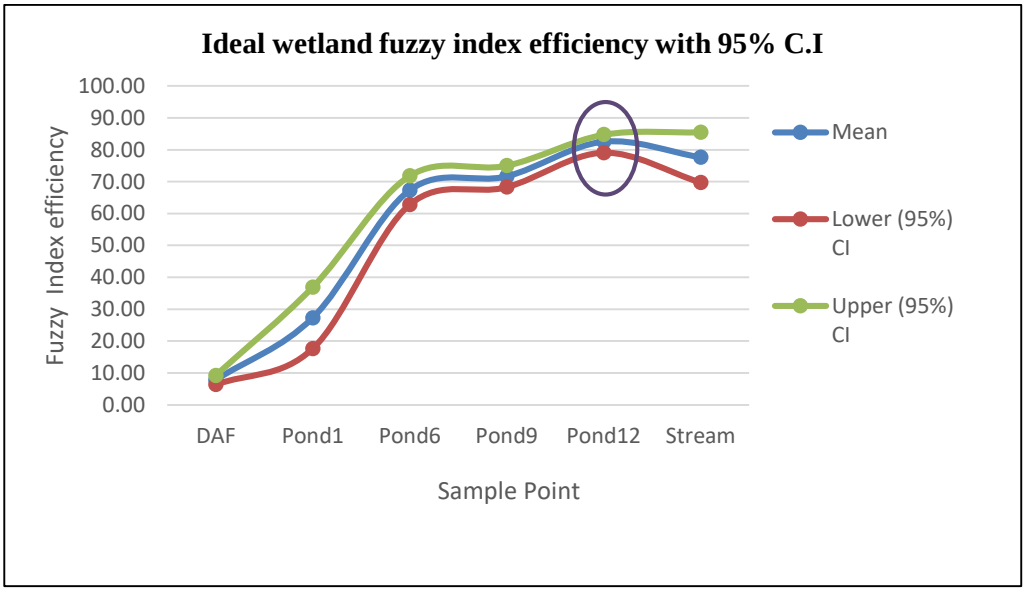


Figure 10.5: An ideal wetland fuzzy index efficiency. Graphically depicts a small confidence interval deviation, typifying the behaviour of ponds 6 and 9.

Note: To achieve an ideal fuzzy index efficiency, would be to pre-line the ponds with high density polyethylene (HDPE) within CW, in order to minimise water loss within the ponds.

10.19 Classifying the wetland efficiency index (WEI)

The most commonly used wetland classification index was developed by Persson *et al*., (1999) which was based on the hydraulic efficiency index of the constructed wetlands. The index (λ) is primarily based on the pond shape, the locations of inlet/outlets and the overall hydrodynamics of the pond systems. The key wetland hydrodynamic measures that are calculated using tracer dispersion experiments are the effective volume ratio and the amount of mixing present within the ponds. The Persson wetland hydraulic efficiency may be categorised into the following groups; (i) good hydraulic efficiency $\lambda > 75$; (ii) satisfactory hydraulic efficiency with $50 < \lambda \leq 75$; and (iii) poor hydraulic efficiency $\lambda \leq 50$. Using this information and reviewing the data present in Table 10.13, the wetland was classified into the three Persson groups. Thus, overlaying the wetland efficiency index (WEI) developed using the fuzzy index framework (F-IND) onto an existing wetland hydraulic index method (λ). The three data points were used; the upper 95% CI, Mean and lower 95% CI for each of the sample points in the wetland system. The results can be seen in Table 10.14.

Table 10.14: Wetland efficiency classification

(efficiency) (λ)	Poor	Satisfactory	Good
DAF	xxx		
Pond 1	xxx		
Pond 6		xxx	
Pond 9		xx	x
Pond 12	x	x	x
Stream		x	xx

Using the Persson efficiency (λ) method as a grouping template, it can be seen that both DAF on pond 1 have poor efficiencies ($\lambda \leq 50$), ponds 6 and 9 have satisfactory to satisfactory to good efficiencies respectively ($50 < \lambda \leq 75$; $\lambda > 75$).

Pond 12 has poor, satisfactory and good efficiencies and finally the stream has satisfactory to good efficiencies. Taking the mean WEI values for the DAF, pond 1, pond 6, pond 9 and pond 12 from Table 10.13, 7.79, 27.27, 67.32, 71.66 and 58.55 respectively the mean value = 46.5%, indicating the mean overall performance of the wetland ($\lambda \leq 50$), indicative of a poor efficiency.

10.20 Conclusions

The F-IND software developed fuzzy indices from wetland environmental multivariate data that can be used to classify the performance of the sample points within the wetland system for the purposes of determining issues within the constructed wetland. The potential exists to use this framework to evaluate a wetland efficiency index (WEI) of individual points within the system based upon selected wetland variables. The grouping category was based on the Persson wetland hydraulic efficiency (λ) template. The WEI analysis using the F-IND framework identified weaknesses within the wetland system, namely pond 12 and also potential loading issues with respect to the DAF and pond 1. Once a common template can be created to review the main sampling points within a wetland system; the wetland efficiency index (WEI) based on the fuzzy indexing (F-IND) framework provides insights to issues within the system not previously known by the main stakeholders.

F-IND has many advantages within the scope of ecological and environmental analysis of complex and differing variables. The framework allows the development of a quality status indexing such as performance, vulnerability, suitability using expert knowledge. The fuzzy indexing system uses a linguistic rationale rather than a mathematical system, which is one of its strengths. Another important concept was the interaction of wetland multi-variables; the software evaluates the weights of the variables interaction resulting in the computation of the index value.

But several disadvantages remain in that the software cannot use variables that contain negative values such as redox potential, where the DAF and pond 1 data values are negative millivolt values (mV), thus oxidation redox potential was not included in the analysis. The user must interpret the results subjectively as it was

observed that the data from the wet/winter season (upper 95% CI) fuzzy index values, which provided the highest performance values indicating a very good treatment performance.

Pond 12 displays the largest deviation in the fuzzy index values (delta =50.2 lower 95% C.I. WEI = 35.23 and the higher 95% C.I WEI = 85.43, see Table 10.13), thus the user must question the overall validity of this pond's performance with values. Pond 12 has a lower mean WEI than the stream, with WEI values of 58.55 and 77.64. Potentially the output wastewaters from pond 12 maybe impacting the local stream. The mean overall efficiency of the wetland system from the DAF to pond 12 has a WEI value = 46.5%. In order for the CW to achieve a satisfactory hydraulic efficiency with $50 < \lambda \leq 75$ i.e. greater than 50, Pond 12 must achieve a WEI value great then 80%, see Figure 10.5 (with a small deviation in fuzzy index values).

It will be shown in the final part of this chapter (section 3), using the BBN model, comprising of the statistical and fuzzy logic analysis undertaken in parts 1 and 2, that the wet season data does not conform to good wetland treatment performance but was in fact the opposite.

Chapter 10. (Section 3) Creation of two fuzzy logic decentralised Bayesian belief networks.

10.21 Summary (section 3)

The creation of causal and diagnostic decentralised BBN were used to understand the issues within the constructed wetland system. The analysis and data from the reduce and iterate process taken from principal component analysis (section 1) and the fuzzy indexing (section 2) were used in the creation of the Bayesian networks. The BBNs were constructed in two phases. Phase 1 created the main sample points (parents) taking the data from section 2 (fuzzy indexing) to describe these sample points in terms of a Weibull probability distribution, thus a reliability metric to the sample points. In phase 2, decentralised causal and diagnostic BBN models were undertaken, using best of fit probability distributions and correlation values for multiple wetland variables, to each of the corresponding sample points. The resulting resilience analysis using Pearson's, Spearman's and global sensitivity analysis metric and correlation ratio reveals the area of greatest resilience/ robustness was the front end of the wetland from pond 1 through to pond 6. The back end of the wetland system, pond 9 to pond 12, shows poor performance and efficiency in treating the effluent.

10.22 Introduction

The wetland efficiency index (WEI) created for each of the wetland sample points (section 2) was further utilised to create two Bayesian belief network (BBN) models. The hybrid model containing both fuzzy logic data and the creation of BBN allows the user to (1) integrate the sample points into a wetland hybrid fuzzy Bayesian model and (2) then create a radial based fuzzy Bayesian model using the probability distributions from the wetland variables.

Bayesian networks or BBNs are also known as the following; Bayes Nets, Belief Nets, Causal Nets and Probability Nets. The dominant naming convention within this chapter will be BBNs which are an efficient data structure graphical modelling technique for encoding information using joint probability distributions for a set of variables to be tested, which defines a domain. BBNs are powerful mathematical tools for modelling complex ecosystems and can be utilised as decision systems in the management of natural resources (Cain, 2001; Nash *et al*., 2010; Ticehurst *et al*., 2011; Keshtkar *et al*., 2013; Spence and Jordan, 2013; Meyer *et al*., 2014) The BBNs are graphical models to illustrate a system of variables, as a network of interactions between the variables using probability to account for the uncertainty within the system. These BBN models use arrows (arcs) to link nodes (variable data described as probabilities) into a network of directed acyclic graphs (DAGs) of (usually) causal influences containing arcs and nodes. Nodes are then evaluated into two main groups, 'parent 'and 'child' nodes. An explanation of parent and child nodes is described in the BBN section. The use of Bayesian networks therefore provides a useful tool in evaluating the functionality and efficiency of the constructed wetland treating the abattoir wastewater. Three BBN models were constructed.

Model (1) was the base model for models 2 and 3, where the Fuzzy index data from Part (2) was taken and analysed into a Weibull probability distribution function. Therefore, the sample points which were previously expressed as wetland efficiency Index data were now expressed as fuzzy Bayesian Weibull distributions. The fuzzy index (F-IND) data has been realised as a Weibull probability distribution function (PDF) through a Bayesian inference program (UniNet). Each sample point (node) was then linked via arrows (arcs) depending on the correlation of the individual nodes. The correlation was elucidated using a

SPSS (v.22) program. Model (2) which was developed as a causal or predictive model, expands Model (1) methodology, by introducing the wetland variables that were dependent to the sample location i.e. DAF fuzzy Bayesian Weibull distribution linked to DAF BOD probability distribution. The corresponding correlation value was extrapolated using a statistic program. Model (3) was developed as a Diagnostic model. Model (3) uses the same methodology as Model (2) except the links (arcs) are in reverse i.e. the DAF BOD probability was linked to the DAF fuzzy Bayesian Weibull distribution. A more detailed discussion on the model methodologies will be further explained in the chapter. These complex modelling methods are quite common in environmental ecosystem analysis. The use of Bayesian analysis has also been applied to ecological risk assessment in complex catchment systems, where the model needed to characterise complex ecological interactions with high uncertainty associated with lack of information (Pollino *et al*., 2007). Lynam *et al*., (2010) used BBN adaptive modelling with major stakeholders (managers, researchers, scientists) to reduce uncertainty and provide guidance to where, when and what interventions are required in water quality scenarios with respect to water quality in the Great Barrier Region of Australia. Spence and Jordan (2013) used BBNs to analyse 30 freshwater wetlands with respect to nitrogen removal.

Using Bayesian analysis allows for two important modelling outcomes uncertainty and sensitivity analysis. Uncertainty analysis answers what is the uncertainty in the outputs given the uncertainties of the inputs. Sensitivity analysis resolves how important are the individual elements of the inputs with respect to the uncertainty of the outputs (Zheng *et al*., 2015). The use of global sensitivity analysis is to define the resilience of the complex wetland system. The term resilience (robustness) defined in terms of 'system states' where disturbances can cause shifts among these states (Holling, 1996; Carpenter and Brock, 2004).

This chapter attempts to understand the complex dynamics and behaviours of the wetland system through the use a hybrid fuzzy Bayesian networks combined with sensitivity analysis. Both uncertainty and sensitivity are combined to understand and visualise the efficiency and resilience of the constructed wetland system.

10.23 Bayesian belief network

A Bayesian network is a graphical representation of the joint probability distributions for a set of variables (Baras *et al*., 2009). In a BBN analysis, for (n) number of mutually exclusive parameters B_k (i=1,2,. . .,n), and a given observed data N ,the updated probability is computed by:

$$P(B_k \mid A) = \frac{P(A \mid B_k) \cdot P(B_k)}{\sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)} \qquad \text{Eq (10.2)}$$

where p(B) denotes the prior occurrence probability of B, p(A) denotes the marginal (total) occurrence probability of N and is most or less constant since the obtained data is available, p(A | B) refers to the conditional occurrence probability of A given that B occurs too and p(B | A) represents the 'posterior' occurrence probability of B given the condition that A occurs. Bayesian networks are directed acyclic graph (DAG) which means there are no cycles.

Assume there is a causal link between A and B (A → B), we state that B is a child of A and A is a parent of B. This is not implying causality. The implication is a direct influence of A over B and the probability B is conditioned on the value of A (Changliang and Zhanfeng, 2009). The creation of a DAG represents the causal (A → B) or diagnostic (A ← B) dependence between the nodes and gives qualitative reasoning in Bayesian networks (Kayakutla and Cinar, 2010).

The quantitative part of Bayesian networks is represented by the Conditional Probability Table (CPT), with conditional probabilities of nodes. The structure of BBNs relies on the chain rule which is the joint probability distribution of the variables within the BBN. The joint probability of X_i is then:

$$P(X_1 \ldots \ldots X_d) = \prod_{i=1}^{d} P\,(Xi \mid parents(Xi)) \qquad \text{Eq (10.3)}$$

The BBN can be used to represent the relationships between the variables via two methods; (1) the top-down inference also known as causal or predictive analysis. It is called causal or predictive analysis because the prediction is in the same direction as the causal arc i.e. from evidence (E) to a query (Q) or cause and effect (A → B). The other method is the reverse of the causal analysis called (2) the bottom-up inference also known as diagnostic analysis, with the arc in the

opposite direction i.e. from a query (Q) to evidence (E) or from effect and cause (A ← B). The diagnostic analysis method has been applied to medical decision support systems (Pople *et al*., 2005) in cognitive sciences (Meder *et al*., 2009) in earthquake disaster analysis (Cockburn and Tesfamariam, 2012) and in identifying root causes resulting in human errors within organisational systems using both causal and diagnostic methods using fuzzy Bayesian analysis (Peng-Cheng *at al*., 2012).

The BBN was designed as a decentralised network. There are three main network designs they are (A) centralised, (B) decentralised and (C) distributed, see Figure 10.6. Where decentralised networks are ubiquitous with complex systems. In organisational psychology for the completion of complex tasks, decentralised networks are more efficient. When multiple operations have to be performed, decentralised networks were more effective (Stohl, 1995). In the real world, decentralised networks are the rule, whereas centralised and distributed networks are rare (Hoste, 2013). Gajduk *et al*., (2014), developed game-theoretic approach in decentralised wireless networks to review their energy efficiency based on insights from evolutionary biology.

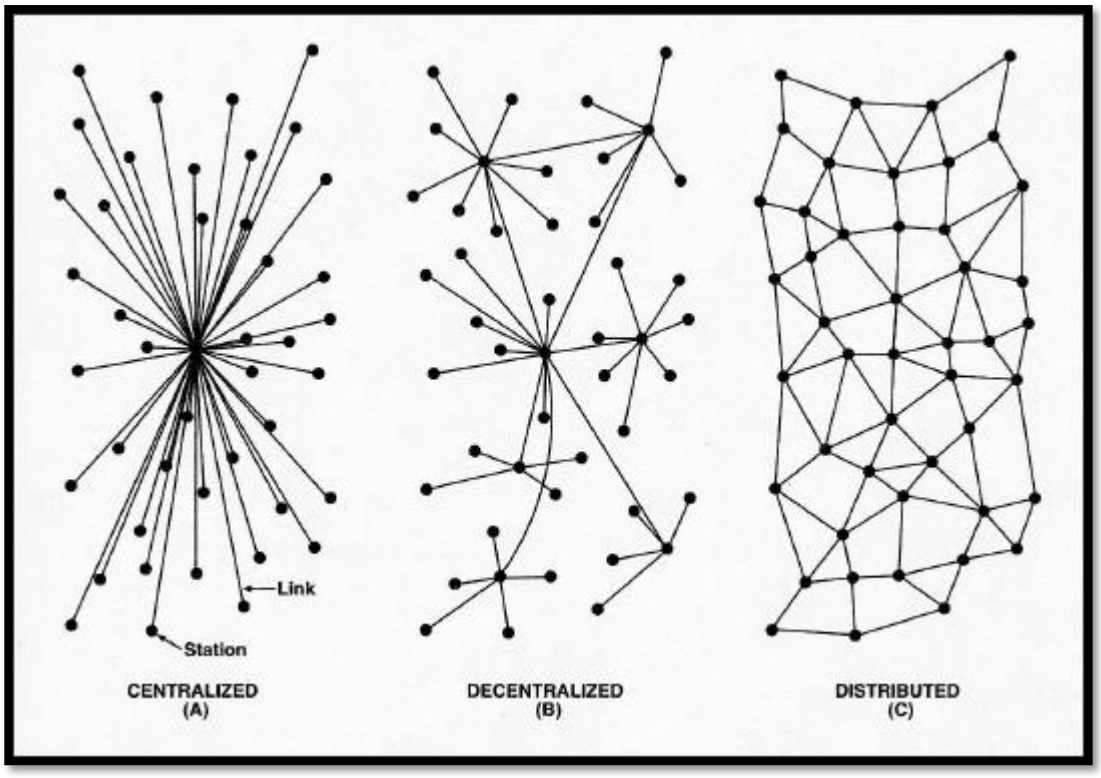


Figure 10.6: Three possible architectures for networks. Paul Baran's networks. Baran suggested three possible architectures for a network such as the Internet: Centralised (A), Decentralised (B) and Distributed. Reproduced from Barabási (2002).

10.24 Fuzzy Bayesian Hybrid

The BBNs employed in this chapter are a hybrid containing wetland fuzzy indices values from section 2 inserted into a BBN. The models incorporate both fuzziness in the model (section 2) and uncertainty through the creation of the BBN (section 3).

The uses of combining different "soft computing" strategies such as fuzzy logic and BBN in understanding complex systems across different disciplines has been studied in Fault tree analysis (Contini, 1995) and modelling the sea-level along the east coast of Britain (Randon *et al*., 2008). Eleye-Datubo *et al*., (2008) explained maritime offshore safety risk modelling in a fuzzy-Bayesian network. They proposed incorporating a flexible fuzzy and uncertainty modelling strategy taking into account the effects of human performance within this complex

domain. Ferreira and Borenstein (2012) used hybrid fuzzy Bayesian methodologies in supply chain and supplier selection functionalities in organisational structures. In this chapter the integratation of the sample point wetland efficient index (WEI) data from section 2, based on fuzzy indices, into a Bayesian inference program was realised; incorporating a Weibull probability distribution function (PDF) for each sample point.

10.25 Statistical analysis

Pearson's non-parametric product moment correlation (PNPPMC) was used to calculate the causal influences between the variables/nodes for both parent and child interactions. The correlations were a measure of between -1 to +1 between two nodes. The resulting correlations formed the following nodal interactions, parent to child for causal model, child to parent for diagnostic model or parent and parent for causal model. The analysis was carried out using SPSS v.22. The correlation analysis was carried out for Model (1), the interaction of the parent to parent nodes and then repeated with all the wetland data, both parent and children and children and parent interactions.

Oracle Crystal Ball fusion© was used to define the best fit probability distributions for the Wetland variables. Along with Weibull probability distribution functions; beta, gamma, uniform, triangular, log uniform and normal probability distribution functions were detected. BBN model's 2 and 3 contain 66 nodes, both parent and children. As stated previously, the parent nodes i.e. the six sample points (DAF, pond 1, pond 6, pond 9, pond 12 and the stream) were analysed using Weibull probability distribution. Therefore, all the parent nodes retained the same probability distribution i.e. Weibull probability distribution (reliability-failure analysis). The Weibull PDF for the parent nodes was calculated using the Fusion Crystal Ball software. The remaining sixty child nodes continuous probability distributions were calculated as best fit by the Crystal Ball software. Each parent had 10 child nodes attached. The best fit continuous probability distribution p-value calculations for the sixty child nodes was defined by the Crystal ball software using both Kolmogorov-Smirnov test (K-S test) Anderson-Darling (A-D Test). The Anderson-Darling test (Anderson and Darling, 1952) was the predominant hypothesis test used in the determination

of the goodness-of-fit of the PDFs on the wetland variables. The Anderson-Darling test (Stephens, 1974) is used to test if a dataset came from a population with a specific distribution function. It can be considered a modification of the Kolmogorov-Smirnov test. The A-D test provides more information to tailing effects in datasets than the K-S test is therefore more sensitive in calculating critical values (NIST, 2012). The A-D test compares the multiple probability distribution functions (PDFs) and identifies which probability distribution function minimised the Anderson-Darling statistic (for a more detailed discussion on goodness-of-fit statistics in distribution fitting, see Vose, 2000). Marcot (2012) used the Kolmogorov-Smirnov test in his metrics for evaluating performance and uncertainty of Bayesian network models. But Greenwood *et al*., (2014) used both Kolmogorov-Smirnov test and Anderson-Darling tests, in their Bayesian analysis of rainfall runoff models in forest water management; stating the KS tests are most sensitive around the median values and less so around distribution extremes whereas the A-D test is more powerful in distinguishing more dispersed differences in the data, i.e A-D tests are more applicable to skewed, tailing or data with extreme distributions, as is the case with the wetland variables.

10.26 UniNet -Bayesian belief network

UniNet allows for the quantification of mixed (discrete and non-parametric) BBNs. The theory for non-parametric continuous BBNs is extended in Hanea and Kurowicka (2008). The software has been used in several research endeavours, showing its ease of use and flexibility. Hanea and Ale (2009), demonstrated its usage as decision tool is assessing the risk of human fatality in building fires. UniNet was used in the development of an air transport safety model using fault trees and BBNs (Ale *et al*., 2009). Hanea *et al*., (2010) used the UniNet software to analyse pollutant emissions and fine particulate concentrations from electricity generating stations over the United States. The UniNet software allows for multiple probability functions to be analysed. The software can analyse nine probability distribution functions; discrete, uniform, log uniform, triangular, normal, log normal, exponential, Weibull and gamma. The user can utilise the fact that the PDFs also contains multiple parameter entries on the PDFs, for example the Weibull PDF contains a type III Weibull parameter entry, where the

location, shape and scale are accepted, in comparison to the standard type II Weibull PDF where shape and scale are typical. The extra degree of freedom on the parameter allows for better accuracy in the model creation. All models underwent 400,000 iterations in the Bayesian analysis. For each of the parent and child nodes, the mean, 5% quantile and 95% quantile are shown. Quantiles are specific values allocated at timed intervals from the inverse function of the probability distribution under analysis. Therefore, the variables under analysis can be subdivided into subsets called ***q***-quantiles i.e. 5% and 95% quantiles. A second part of the UniNet software suite was the UniSens package, which quantified the correlations (interactions) between the nodes.

10.27 UniSens -Sensitivity Analysis

The UniSens software, which is a subset package of UniNet, analyses the correlations between all the nodes present in the BBN. It can be considered one of the most useful analysis as it formulates as a sensitivity profile between the nodes in the network. UniSens records the sensitivities of the dynamic variable interactions, such as the correlation of a child node to the parent node. This enhances our understanding on the complex interactions that exist in the wetland system. The purpose of sensitivity analysis was to determine the overall effect of the children (variables) on the parents (sample points) and therefore provide an overview on what points in the wetland system are resilient and where deficiencies exist. The concept of resilience will be further explained in the next section. Using UniSens (which calculates various statistics and sensitivity measures based on samples imported from BBN files) the sensitivity of the sample point fuzzy Bayesian metrics and their corresponding sensitivities with their specific data inputs (children) were evaluated for the three BBN models (1) the sample point (parent) interaction model (2) cause–effect (causal/top down) network and (3) effect-cause (diagnostic/bottom-up) network. All models underwent 400,000 iterations in the sensitivity analysis.

UniNet is a stand-alone program using BBNs for stochastic modelling and for multivariate ordinal data mining. It is designed by the Risk and Environmental Modelling group at the Department of Applied Mathematics of the Delft

University of Technology and can be downloaded from http://dutiosa.twi.tudelft.nl/_risk [Accesssed 03 2015]

10.28 The concept of resilience (ball-and-cup heuristic)

Resilience is the capacity of a system to respond, adapt and tolerate disturbances and/or future changes without collapsing in an alternative state, resulting in a different set of processes controlling the system. A resilient ecosystem can withstand shocks and rebuild itself when necessary. Other views of resilience are in the concepts and associations with (1) self-organisation, (2) adaptive capacity and (3) redundancy (Carpenter *et al.*, 2001; Low *et al*., 2003, Tompkins and Adger, 2004). Where self-organisation resilient systems can withstand and diffuse external perturbations, i.e. shocks. Adaptive capacity of a system is associated with dynamics of the system's environment to both internal and external effects and learning from these changes (Carpenter and Brock, 2008). Redundancy can be viewed as an associate of uncertainty. When a system is disturbed, the energy within the system can become uncertain relative to its state of dynamic balance i.e. 'basin of attraction'. This increased uncertainty is associated with redundancy (Arreguín-Sánchez and Ruiz-Barreiro, 2014). Liao, (2012) referred to redundancy as the provision for 'insurance against total system failure' and 'the extent to which systems components are substitutable'. They are two modes of resilience, engineering and ecological. The difference between engineering and ecological resilience can be illustrated by the ball-and-cup diagram (Scheffer *et al*., 1993, Walker *et al*., 2004). The cup represents the region in the state space or basin of attraction, in which the system tends to remain. The basin of attraction includes all possible values of system variables of interest. The ball represents the state of the system at any given time. The engineering resilience concept assumes only one regime, thus only one possible basin of attraction with the very bottom of the basin represents the ideal stable state. The ecological resilience concept assumes multiple regimes, therefore more than one basin of attraction. The system may oscillate within the basin, never settling at the bottom. The system may also cross a threshold and settle in a new basin of attraction. The notion of engineering resilience is concerned with whether the system can remain at the bottom of the basin; while the notion of ecological

resilience is concerned with whether the system can remain within the current basin (Holling 1996) see Figure 10.7.

Two resilience concepts: engineering versus ecological concepts (ball-and-cup)

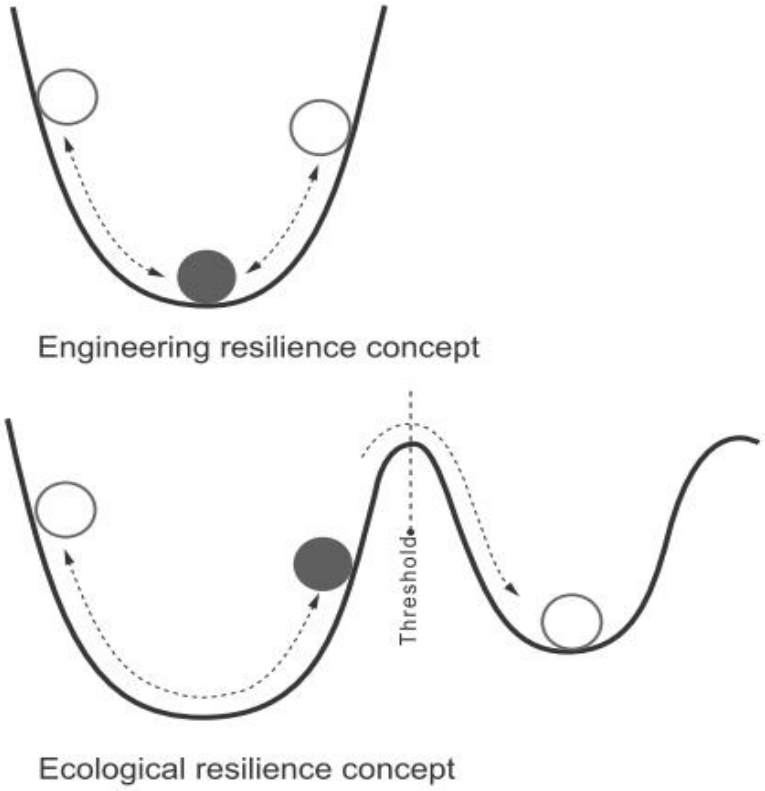


Figure 10.7: Two resilience concepts: engineering versus ecological resilience; single versus mulitple basins of attraction (taken from Liao, 2012).

An extension of the ecological resilience concept can be futher explored. We can subvide the ecological concept into Case A and Case B, See Figure 10.8, where the state changes within the basin of attractions can have different shapes and sizes. If a system displays a narrow or shallow basin, Case A, its considered an undesirable system. If a system displayes a wide or deep basin, Case B, it's considered a desirable system.

Ecological resilience (further concepts)

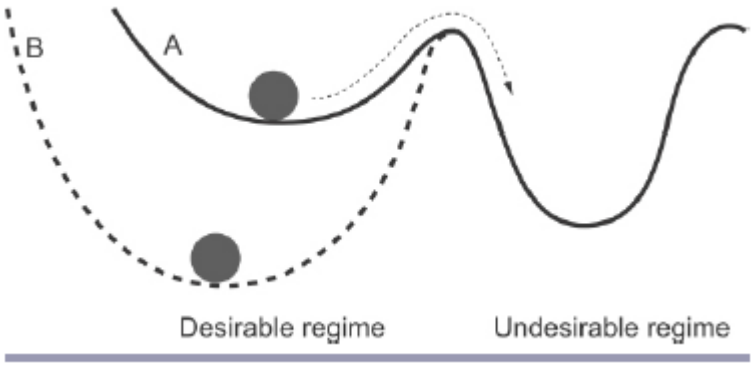
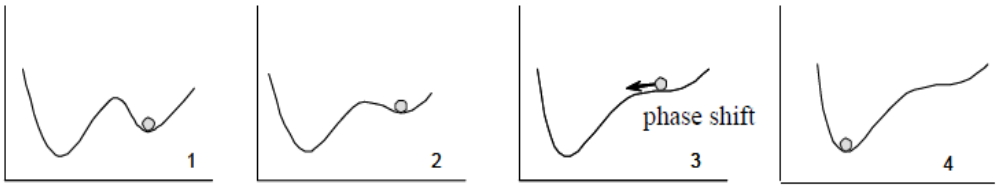


Figure 10.8: Ecological resilience (further concepts) (taken from Liao, 2012)

A further expansion of the ecological resilience concept is shown in Figure 10.9 from (Deutsch *et al*. 2003), which depicts four different states/phase shifts in coral reefs. As previously stated, ecological systems have more than one basin of attraction. The wetland system can be treated as a complex ecosystem and Scheffer *et al*., (2001) proposed that these complex ecosystems tend to have more than one state and can slide between alternate states.



coral dominance	overfishing, coastal eutrophication	disease, bleaching hurricane,	algal dominance
clear water	phosphorous accum-ulation in soil and mud	flooding, warming, overexploitation of predators	turbid water
grassland	fire prevention	heavy rainfall and intense grazing	shrub-bushland

Figure 10.9: Resilience phase shifts in a complex ecosystem. The phase shifts in four different coral reef ecosystems. A desirable ecosystem state with valuable ecosystem services (1) the coral reef in a state dominated by coral. The stability domain is affected by various management practices that reduce the resilience of the system the cup becomes shallower; the reef is still dominated by coral, but human activity is reducing the ecosystems resilience (2). A disturbance that previously could be absorbed pushes the system into a new stability state; the coral reef becomes more vulnerable (3). The shift moves the system into an undesirable state with a loss of ecosystem functionality, where the reef ecosystem is dominated by algae (4) (Deutsch *et al.*, 2003).

10.29 Global sensitivity analysis (GSA)

Taking the results from the UniSens, the sensitivity analysis package allows the user to delve into the wetland data. Using the resilience 'cap-and-ball' diagrams as a visual representation of the sensitivities expressed from the causal decentralised fuzzy Bayesian model (2) and the diagnostic decentralised fuzzy Bayesian Model (3). Sensitivity analysis (SA) can be divided into two classes which are 'local' and 'global'. A local analysis addresses sensitivity relative to point estimates of parameters values while global analysis examines sensitivity with regard to entire parameter distribution (Hamby, 1995), therefore global sensitivity analysis (GSA) considers the whole variation range of inputs (Saltelli *et al*., 2000a). One technique to determine the GSA is using the correlation ratio (CR). The correlation ratio is part of a group of global quantitative measures of importance of input variables for a given model. It is a variance-based non-parametric method and closely related to Sobol' indices (Sobol 1993, Chan *et al*., 2000). The Sobol's method is a global sensitivity analysis (GSA) technique which determines the contribution of each input (or group of inputs) to the variance of the output (Isaacs and Glen, 2012). As a measure of sensitivity, the names 'importance measure',' correlation ratio', or 'sensitivity index', are used (Saltelli *et al*., 1999, Saltellli *et al*., 2000a, Saltelli *et al*., 2000b). For the purposes of this chapter correlation ratio (η^2) was used. The sensitivity metrics (UniSens) of interest are the correlation ratios of the resultant model's outcomes; where Karl Pearson (1905), introduced correlation ratio η^2 as a measure for the non-linear effect of random input vector X on a random output variable Y. Global sensitivity analysis has been used in ecological resilience analysis (Perz *et al*., 2013) to vehicle tyre model analysis (Anstett-Collin *et al*., 2014) and in hydrological modelling (Song *et al*., 2015). Other correlations recorded during the sensitivity analysis, were the product moment (Pearson's) and rank correlations (Spearman's). The difference between Pearson's Product Moment correlations and Spearmans' Rank Correlations is as follows; Spearman's rank correlation (ρ) is used to test the association between two ranked variables, or one ranked variable and one measurement variable, where Spearman's is evaluated on ranks and depicts increasing or decreasing (monotonic) relationships. Whereas

Pearson's product moment correlation (r) is a measure how well data are related and is used on true values and depicts linear relationships. There is a difference between Pearson's Product Moment correlations and Spearmans' Rank Correlations, where Spearman's produces better discrimination in complex data sets especially facial recognition (Ayinde and Yang, 2002) and colour imaging analysis (Muselet and Trémeau, 2008). The relationship pertains to the computational power required to derive the Spearman's rank correlation and Pearson product moment correlation. Spearman's rank requires the order of ($n.log_2\ n$) comparisons whereas Pearson's requires the order of (n); therefore, when dealing with nonlinear data Spearman's rank correlation produces better results, but at a cost of computational power. For more detailed information of Spearman's (ρ), Pearson's (r) and Correlation ratio refer to Goshtasby (2012). All three correlations will be used in this chapter.

10.30 Model methodology

Model 1: The first BBN links the sample points within the wetland. The sample points each have a series of fuzzy indexing data derived from section 2. This data was resolved into a type III (location, scale and shape) Weibull probability distribution for each six sample points (DAF, pond1, pond 6, pond 9, pond 12 and steam). Weibull analysis is commonly employed in reliability in engineering failure models over time. This methodology has been adopted, in that the wetland was sampled over time at different locations. Vidal *et al.*, (2011) used five probability distributions with Weibull being one of the dominant probability distributions in their assessment of constructed wetlands treating small communites's wastewater, where Weibull was selected when there is non-compliance of standards i.e. failures. The sample fuzzy index values are now represented as a Weibull distribution; thus, each sample point represents a fuzzy index Weibull distribution (FIWD). The values for the arcs (links) between the sample points were evaluated using Pearson's non-parametric correlation using SPSS v.22, See Figure 10.10a and 10.10b.

Model 2: (Causal/Predictive/top –down): The second part of the BBN describes a causal model which links the sample point (parent) to their corresponding variables within that sample point individual, i.e. DAF (Fuzzy index Weibull distribution (FIWD) (parent) will link to DAF_BOD (child) [(DAF → DAF_BOD), see Figure 10.11. This process was repeated for all of the individual wetland variables (children). The DAF fuzzy Bayesian Weibull distribution parent node was linked to all its corresponding DAF variables (Children), See Figure 10.11.

Model 3: (Diagnostic/ Bottom-up): The third part of the BBN describes a diagnostic model which links the individual variables to their corresponding sample point, i.e. DAF_BOD (child) will link onto the DAF Fuzzy index Weibull distribution (FIWD) (parent) [(DAF ← DAF_BOD. This process was repeated for all of the individual wetland variables (children) linking to their corresponding DAF Fuzzy index Weibull distribution parent node, see Figure 10.12.

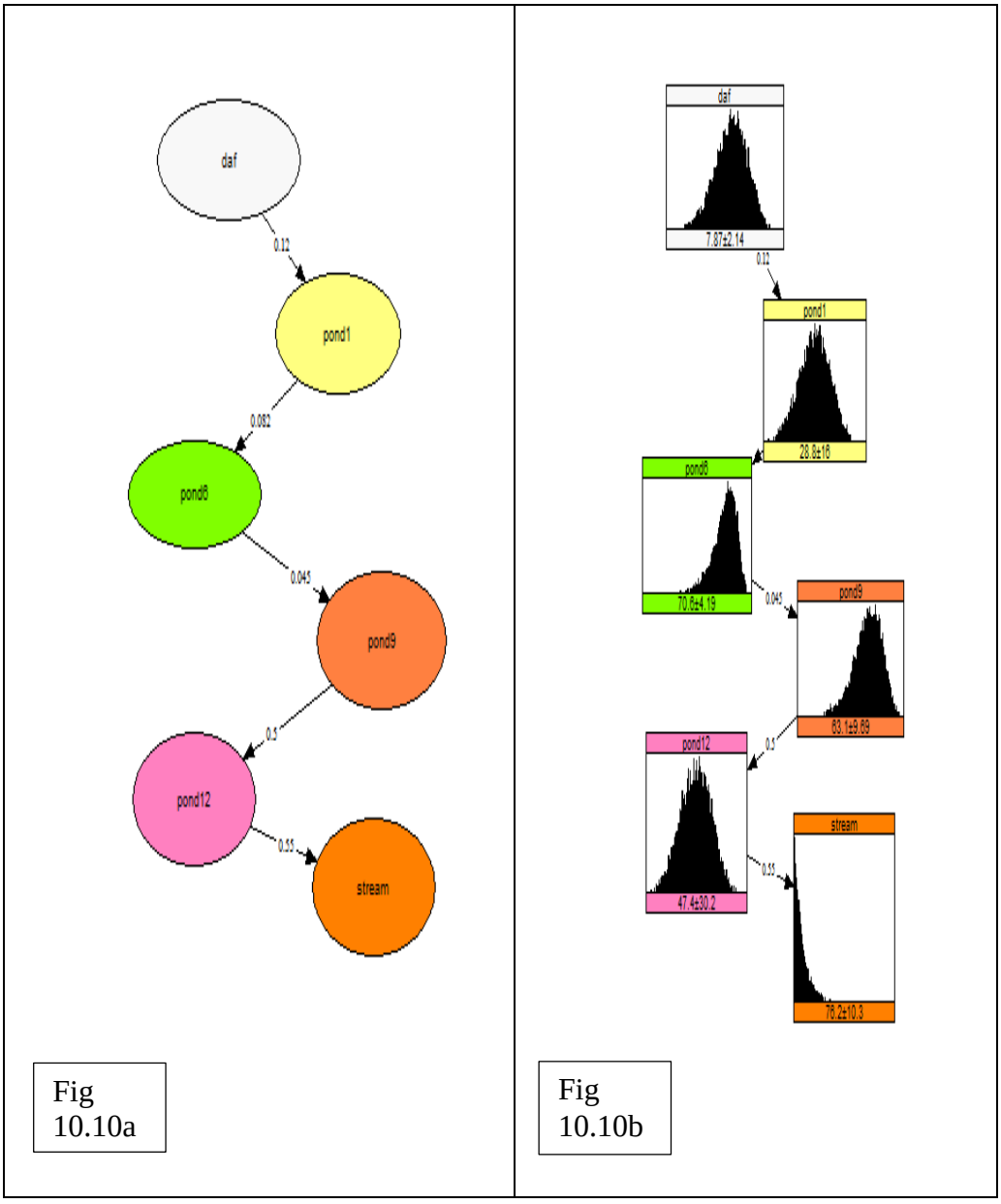


Figure 10.10 a, b: Bayesian wetland parent nodes (ellipses versus Bayesian Weibull distribution). Figure 10.10 a representing Model (1) shows the central hubs (parents) of the decentralised causal based BBN, which are DAF, pond 1, pond 6, pond 9, pond 12 and stream. The nodes are depicted as ellipses. In Figure 10.10b, each sample point (parent) was described by an individual probability distribution showing the mean fuzzy Bayesian Weibull probability distribution, shown are the corresponding fuzzy Bayesian (FuzBayes) Weibull probability for each parent (sample point) with ± standard deviation after 400,000 iterations.

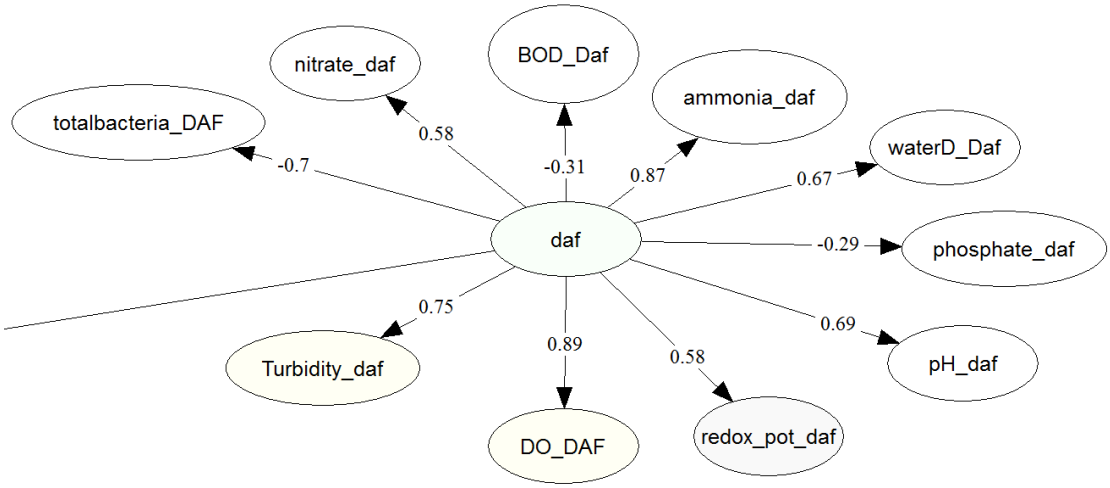


Figure 10.11: The DAF causal/predictive decentralised Bayesian belief network (cause – effect configuration).

The parent node (DAF fuzzy index value) at the centre and radiating around this core are the ten child nodes (pH, redox potential, DO, turbidity, BOD, water depth, ammonia, nitrate, phosphate and $\log_{10}$ total bacteria). These child nodes only relate to the parent node i.e DAF fuzzy Bayesian Weibull probability distribution (reliability) with respect to DAF variables. The correlations are non-parametric rank correlations, which relates to the non-parametric measure of statistical dependence between the variables.

Notice the direction of the arrows (arcs) (A → B).

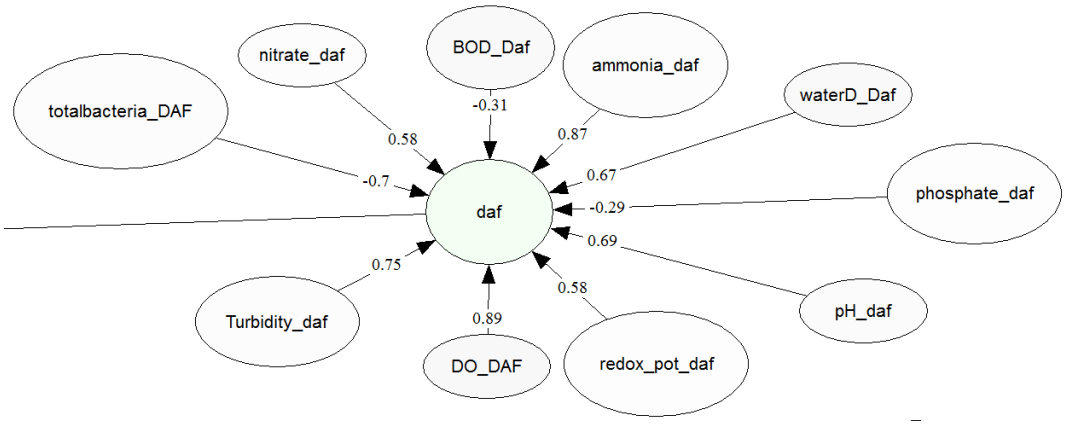


Figure 10.12: The DAF diagnostic decentralised BBN (effect – cause configuration). The parent node (DAF fuzzy Bayesian Weibull distribution) at the centre and radiating around the parent node are the ten child nodes (pH, redox potential, DO, turbidity, BOD, water depth, ammonia, nitrate, phosphate and $\log_{10}$ total bacteria). These child nodes only relate to the parent node i.e DAF fuzzy reliability with respect to DAF variables. The correlations are non-parametric rank correlations, which relates to the non-parametric (data doesn't rely on any particular probability distribution) measure of statistical dependence between the variables.

Notice the direction of the arrows (arcs) (A ← B).

10.31 Model process flow

Step 1: Using Principal Component analysis (section 1), in conjunction with a reduce and iterate procedure, the initial twenty-two variables were reduced to a final seven wetland variables after the third PCA iteration, the remaining wetland variables were BOD, redox potential, conductivity, turbidity, $\log_{10}$ total bacteria, water depth and dissolved oxygen. Redox potential was omitted from the next stage – fuzzy index analysis, because the fuzzy tool (F-IND) cannot accommodate negative redox potential readings (-mV). The conductivity readings were also not included because the model displayed no extra value with the inclusion of this wetland metric.

Step 2: Using the combined wetland fuzzy indices (section 2) to evaluate the reliability of each of the sample points, using 5 key variables (water depth, BOD, turbidity, $\log_{10}$ total bacteria, dissolved oxygen); the resultant fuzzy index values for each sample point over the sampling period was evaluated and recorded. These index values are subjected to Weibull analysis probability distribution using Bayesian inference (UniNet) to determine the reliability metrics for the sample points.

Step 3: Determine the best fit probability distribution for each of the variables in relation to the variable values at each sample point, using the Oracle Fusion Crystal Ball® software.

Step 4: Determine the non-parametric Pearson's rho correlations using SPSS v.22 software for the Weibull probability distributions of the sample points (DAF, pond 1, pond 6, pond 9, pond 12 and stream and their corresponding sample point variables.

Step 5: (Combining Steps 1, 2 and 3). A decentralised BBN was then established, where variable specific data for each sample point and the resultant fuzzy Bayesian Weibull probability distribution i.e. DAF Weibull probability distribution as determined by fuzzy indices for the 5 key variables as mentioned above, and the DAF related variables such as DAF pH, BOD, total bacteria, ammonia, sulphate., with their corresponding non-parametric correlations.

Step 6: Two decentralised based BBNs were created (1) cause –effect network and (2) effect-cause network.

1. Cause – Effect network: also known as a "top down", "causal" network was used to predict the effects of a given set of inputs/variables.
2. Effect-Cause network: also known as a "bottom up, "diagnostic" network was used to find hidden causes within the network.

In both these network analysis, the main points of interest were the resulting sensitivity parameters, which will be discussed in greater detail later in the chapter. See section 10.36.

Step 7: Using UniSens (sensitivity analysis) to review the subsequent correlations (local and global) within the fuzzy Bayesian networks and determine significant behaviours of the wetland system, in relation to (1) causal (predictive) effects and (2) diagnostic (hidden) effects.

Figure 10.13, depicts the model flow diagram encapsulating the above steps. Figure 10.14a shows the decentralised diagnostic model (Model 3) for the DAF, Pond 1, Pond 6 and pond 9. The sample points (parents) are shown with their corresponding sample pointy specific data (children) connected to them via arrows (arcs). Both parents and children are depicted as elliptical nodes. Figure 10.14b, shows the decentralised diagnostic model based on the nodes individual probability distributions.

Figure 10.13: The fuzzy Bayesian model flow diagram. Depicts graphical layout of the chapter ten, section 3.

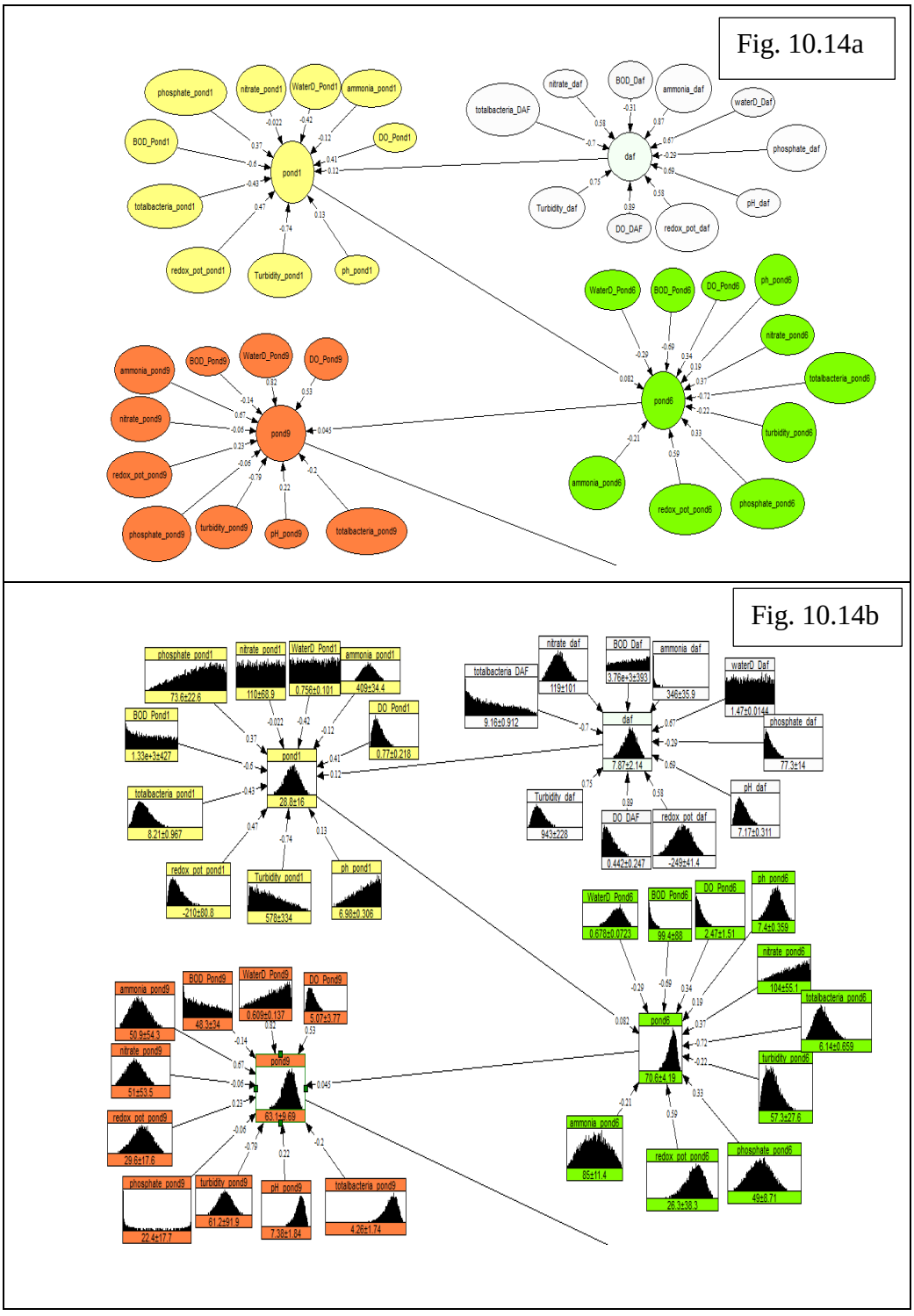


Figure 10.14 a, b: The wetland decentralised diagnostic BBN, showing DAF, pond 1, pond 6 and pond 9; Figure 10.14a depicts the BBN nodes as elliptical. Figure 10.14b depicts the nodes as probability distributions (pond 12 and stream nodes not shown).

10.32 Results and Discussion

10.32.1 Fuzzy Bayesian review - wetland sample points (Parent Nodes)

The sample points are divided into three parts, the mean, 5%-quantile and 95%-quantile. The quantiles represent the extreme top and bottom subsets of the inverse function of the variables probability distributions taken at specific intervals during the 400,000 iterations of the model sampling. A compare and contrast of the wetland fuzzy indexing efficiency with 95% confidence intervals was performed, see Figure 10.15. The graph displays the mean fuzzy index efficiency for each of the sample points (DAF, pond 1, pond 6, pond 9, pond 12 and stream) along with their corresponding confidence intervals (95% C.I.)

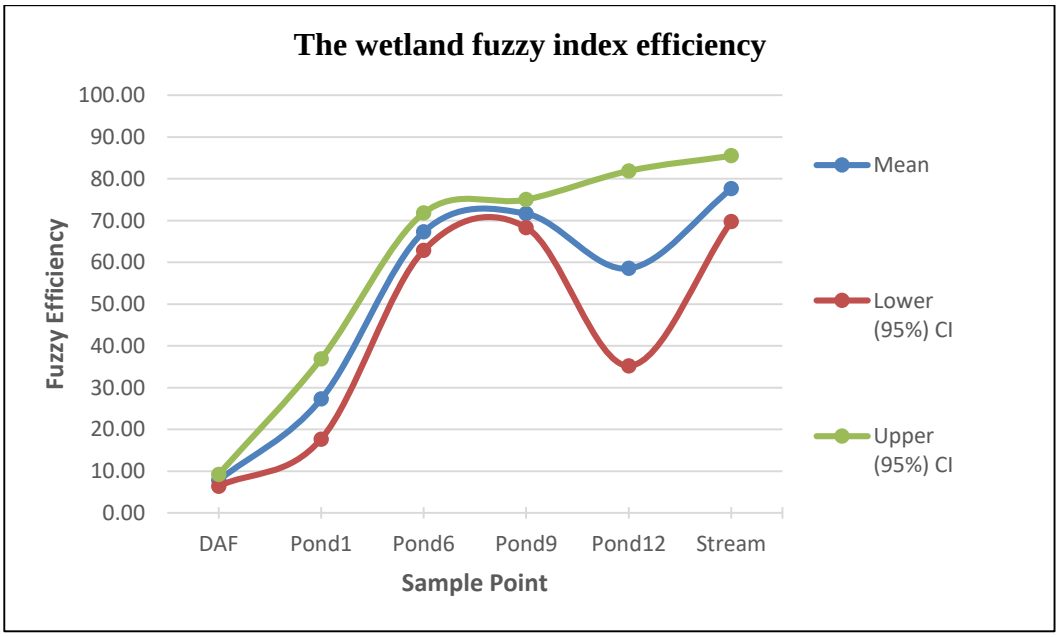


Figure 10.15: The wetland fuzzy index efficiency, shows the wetland fuzzy index using the F-IND software (section 2). Included are the 95% confidence intervals for each sample point.

The causal fuzzy Bayesian Weibull distribution using the sample points (parent nodes) was run i.e. Model (1).

Model (1) depicts the causal fuzzy Bayesian Weibull distribution efficiency of each of the sample points (parent nodes). Figure 10.16, shows the graphical representation of the Bayesian analysis of Model (1) after 400,000 iterations.

Model (1) differs considerably from the fuzzy index efficiency model, with a greater variation in the upper and lower extremes of the model.

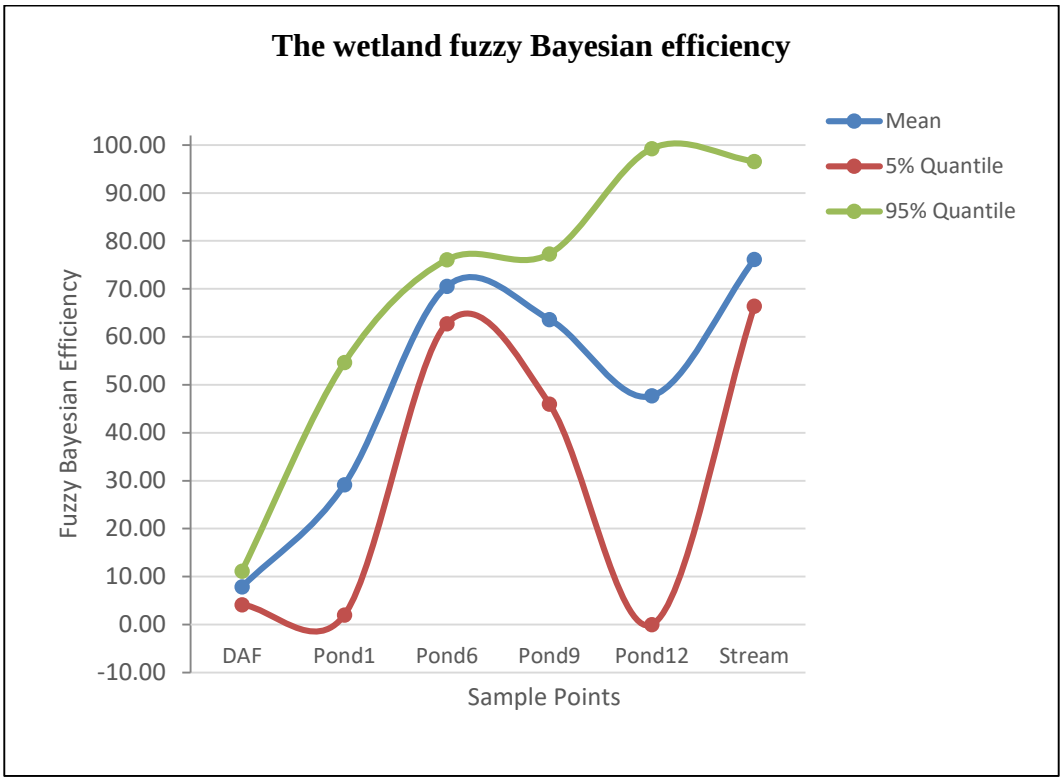


Figure 10.16: The wetland fuzzy Bayesian efficiency shows the use of the fuzzy indices variables inputted into the Bayesian analysis software (UniNet©). The mean, 5% quantile and 95% quantile lines are displayed to reveal the upper and lower space of the wetlands' fuzzy Bayesian efficiency.

A comparison of the mean fuzzy index efficiency model (F-IND) data and the mean fuzzy Bayesian efficiency model (FuzBayes) data was performed, see Figure 10.17; revealing that the mean data points were similar for DAF, pond 1, pond 6 and the stream, but variations were observed between the models at ponds 9 and 12. Pond 9's mean fuzzy indexing efficiency (F-IND) value was 71.66%, in comparison to the mean fuzzy Bayesian efficiency (FuzBayes) value was 63.63%, a decrease of 8% efficiency. Pond12's mean fuzzy indexing efficiency (F-IND) value was 58.55%, in comparison to the mean fuzzy Bayesian efficiency (FuzBayes) value was 47.70%, a decrease of 10.85%.

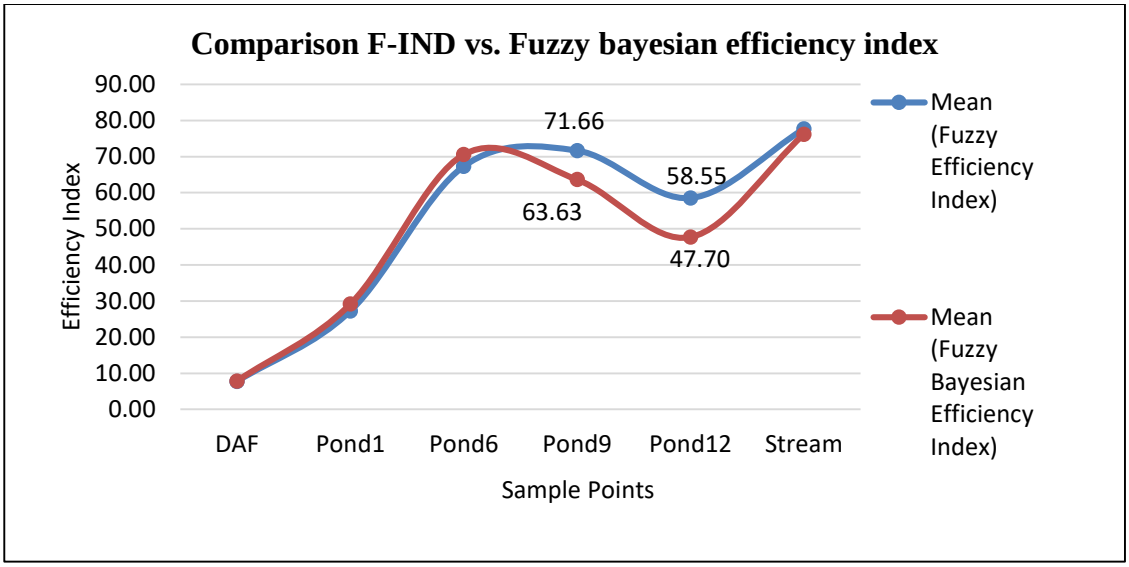


Figure 10.17: A comparison between mean (F-IND) and mean fuzzy Bayesian efficiency index.

The difference between the fuzzy based wetland efficiency index (F-IND) and the fuzzy Bayesian efficiency index after 400,000 iterations. The graph depicts the sample points on the x-axis and an efficiency index on the y-axis.
NOTE: With respect to the fuzzy Bayesian analysis, the quantile boundaries of 5% and 95% are effectively the warm and wet weather conditions within the constructed wetland where the;

- 5% quantile represents the warm months from May 2007 to September 2007
- 95% quantile represents the wet months from October 2007 to February 2008

A similar analogy applies to the fuzzy index data, see figure 10.15, concerning the 95% C.I. upper and lower trends. The mean trend represents the average of these two extremes within the constructed wetland.

10.32.2 Fuzzy Bayesian review - wetland variables (child nodes)

In this section, we review the outcomes of the causal fuzzy Bayesian analysis Model (2) on the ten key wetland variables. Therefore the model consists of the sample points and their corresponding wetland variables (cause and effect/ predictive/ A → B). The wetland variables are redox potential, water depth,

ammonia, total bacteria, dissolved oxygen, pH, nitrate, phosphate, BOD and turbidity. A scatter plot of the wetland variable with the moving average trend associated with that variable is displayed. The next graph then reveals the fuzzy Bayesian plot of the variable, with the mean, 5% quantile and 95% quantile plotted.

The 66 scatter plot sample points can be subdivided into the following six sample points: Scatter plots data points:

1. Points 1- 11 (DAF)
2. Points 11-22 (pond 1)
3. Points 22-33 (pond 6)
4. Points 33- 44 (pond 9)
5. Points 44-55 (pond 12)
6. Points 55-66 (stream).

The exception to the rule was the ammonia data, which was sparse due to the limited sampling regime for this variable.

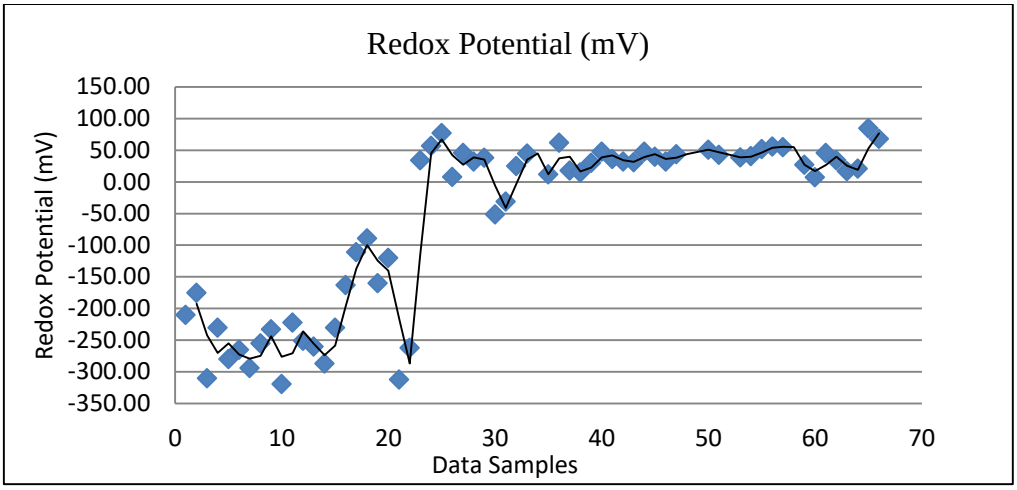


Figure 10.18: The redox potential (mV) of the wetland system (no modelling) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account the upper and lower boundary conditions of the data. The data shifts from negative readings (anaerobic) to positive readings (aerobic) with data points associated with pond 6 (points 22 -33).

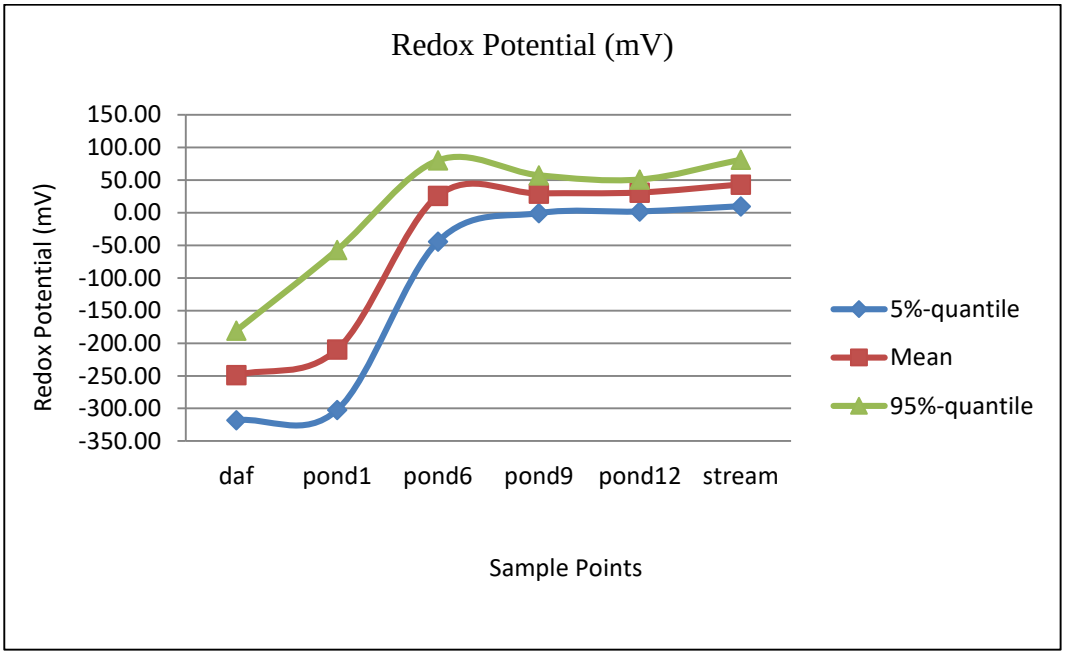


Figure 10.19: The redox potential (mV) of the wetland system (fuzzy Bayesian model) shows the fuzzy Bayesian analysis of the redox potential (mV) data of the entire wetland system for the duration of the sampling.

The data has been analysed with respect to the correlations between the redox potential data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the redox potential. It can be seen that

there is a transition between anaerobic to aerobic behaviour within the wetland, occurring at pond 6, where DAF, pond1 → pond 6 display anaerobic values and pond 6 → pond 9 → pond12 → stream display aerobic in their values. Pond 6 sits on the division between these extremes, displaying both anaerobic and aerobic values.

This process and the subsequent behaviour were revealed for all relevant wetland system variables.

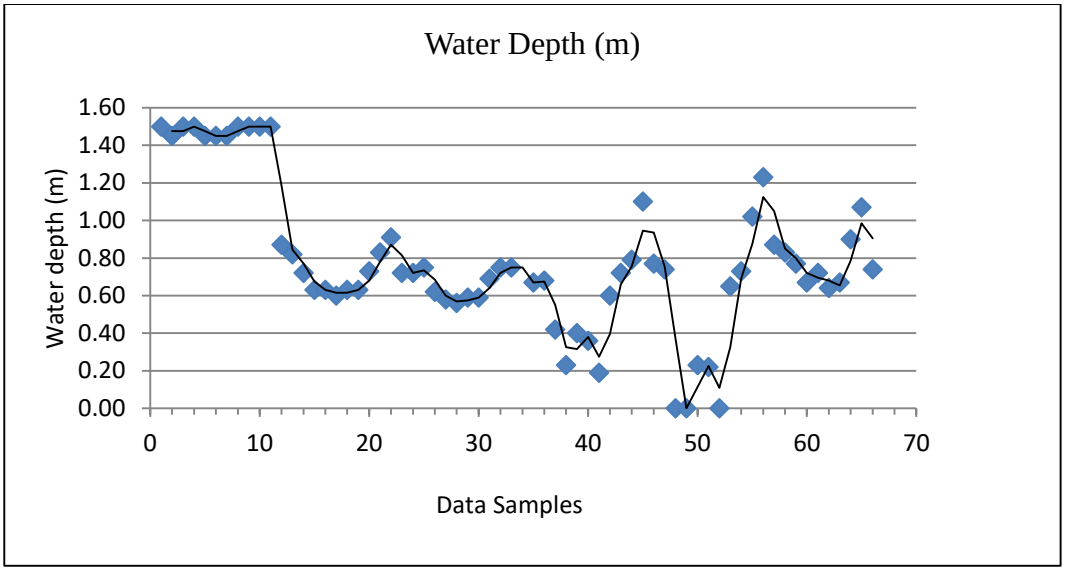


Figure 10.20: The water depth (m) of the wetland system (no modelling) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account the upper and lower boundary conditions of the data. Large variations within the water depth data can be seen between data points' associated with pond 9 and pond 12 (points 33-44 and 44-55 respectively). Pond 12 displays extreme water depth behaviour.

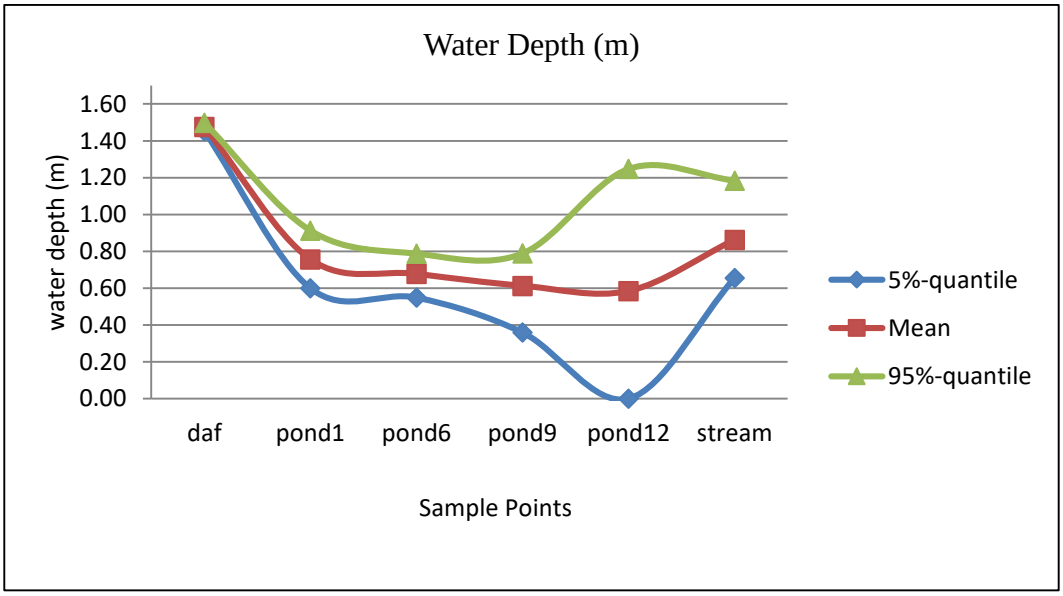


Figure 10.21: The water depth (m) of the wetland system (fuzzy Bayesian model) of the water depth (m) data of the entire wetland system for the duration of the sampling, but the data has been analysed with respect to the correlations between the water depth data for each sample point (parent node). The analysis shows the

5% quantile and 95% quantile and the mean of the water depth. Extreme water depth behaviour is evident from pond 9 to pond 12.

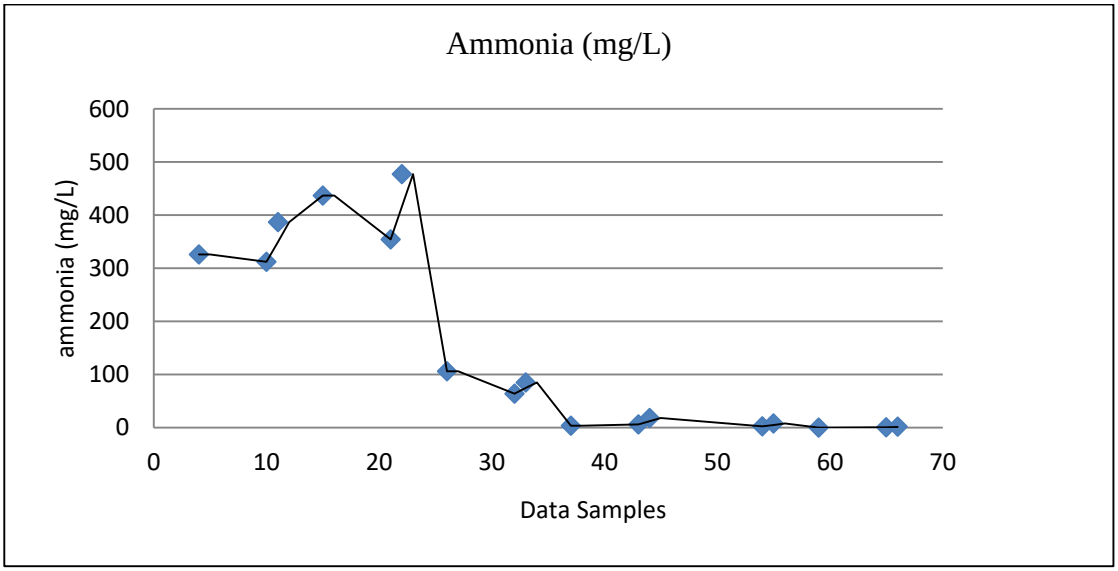


Figure 10.22: The ammonia concentration (mg/L) of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account the upper and lower boundary conditions of the data.

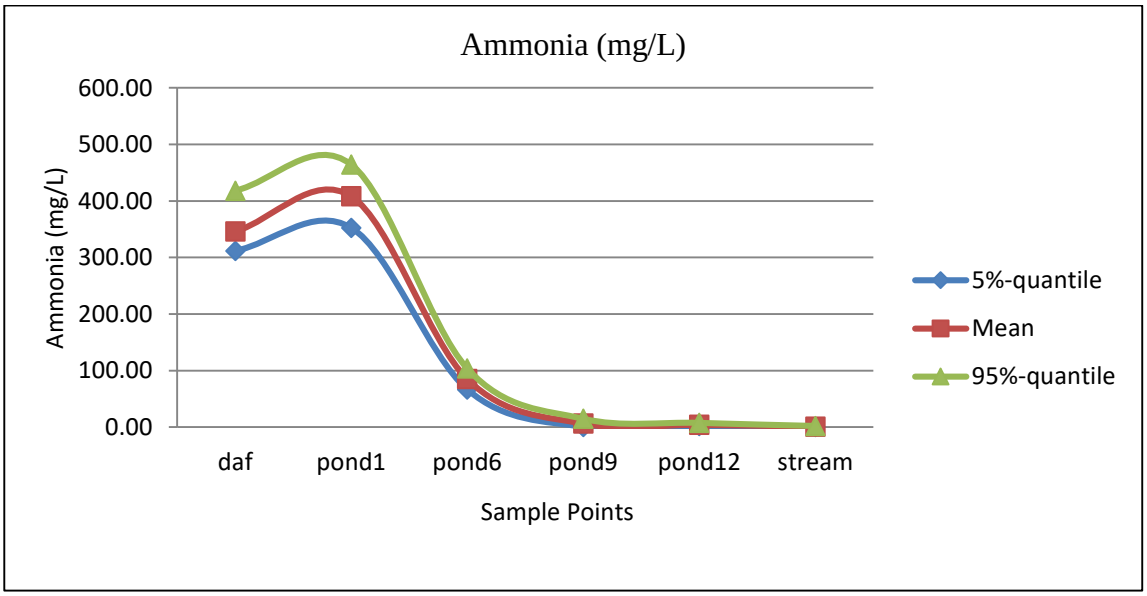


Figure 10.23: The ammonia concentration (mg/L) of the wetland system (fuzzy Bayesian model) of the ammonia (mg/L) of the wetland system for the duration of the sampling, but the data has been analysed with respect to the correlations between the ammonia data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the water depth. Large

ammonia concentrations were evident towards the front of the wetland system, from DAF and pond1; with pond 1 exhibiting larger ammonia values, than the DAF plant. With the potential that pond1 is a reservoir for ammonia, due to anaerobic conditions present.

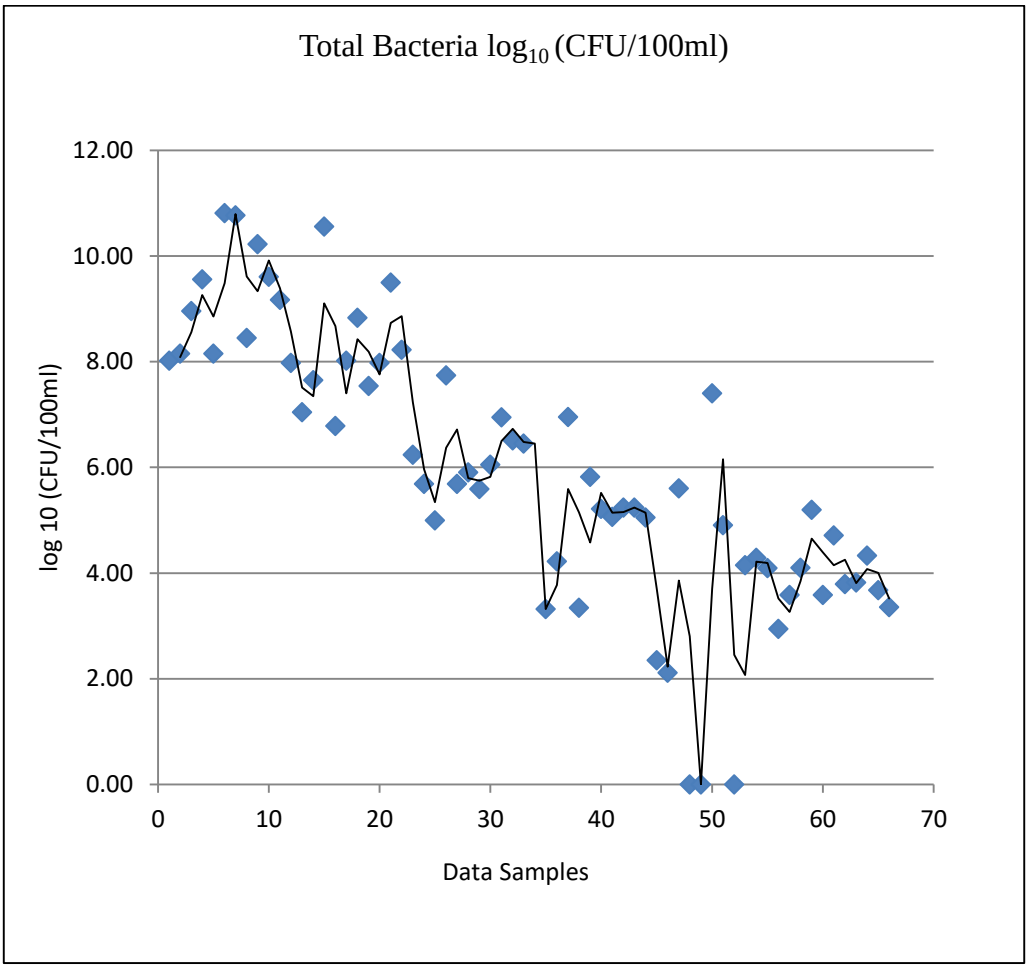


Figure 10.24: The $\log_{10}$ total bacteria (CFU/100ml) of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account the upper and lower boundary conditions of the data. The bacterial data displays erratic behaviour but shows an overall decrease in total bacteria concentration within the wetland system.

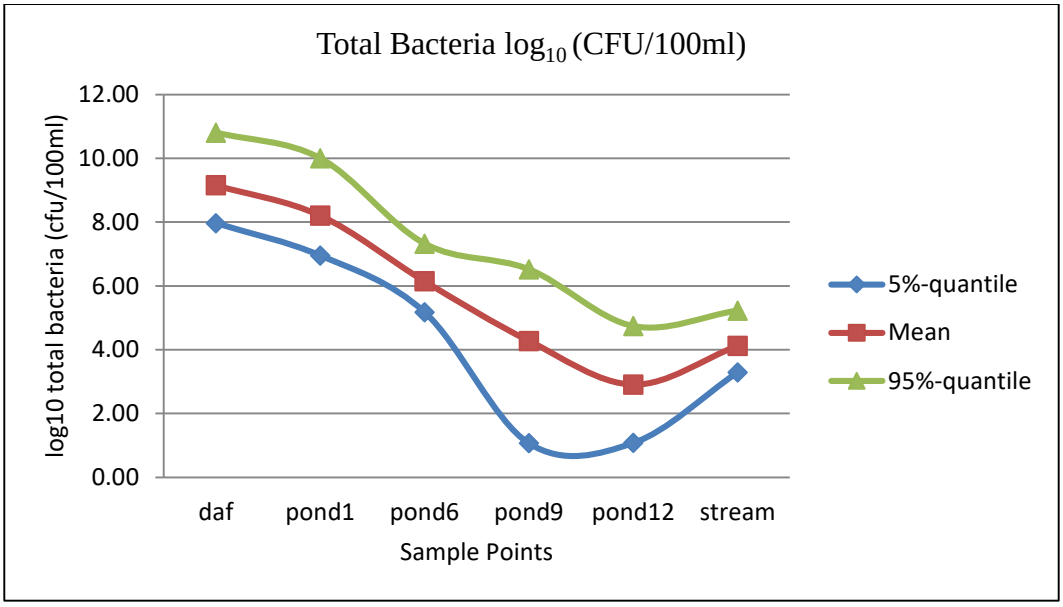


Figure 10.25: The $\log_{10}$ total bacterial (CFU/100ml) concentration of the wetland (fuzzy Bayesian model) for the duration of the sampling. The data has been analysed with respect to the correlations between the ammonia data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the $\log_{10}$ total bacteria (CFU/100ml). Large bacterial concentrations were evident towards the front of the wetland system, from DAF and pond 1; with the DAF exhibiting larger bacterial concentrations values. Pond 9 and pond12 both exhibit greater fluctuations in bacterial concentrations towards the back-end on the wetland system, with pond 12 having on average a lower bacterial concentration than the local stream.

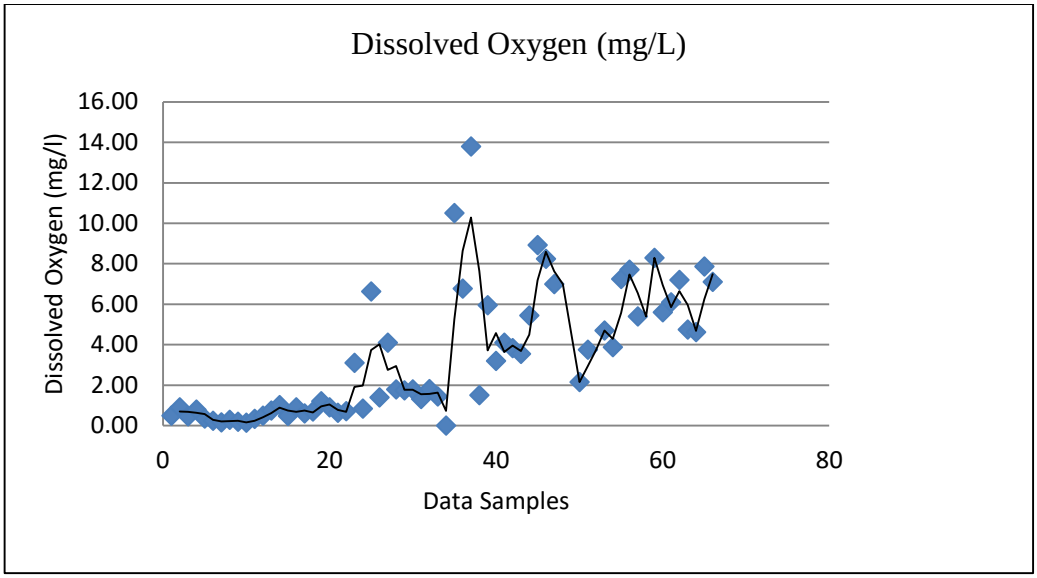


Figure 10.26: The dissolved oxygen (mg/L) of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account upper and lower boundary conditions of the data. The dissolved oxygen data remains reasonably constant until ponds 6 and 9 (data points 22-33 and 33-44 respectively).

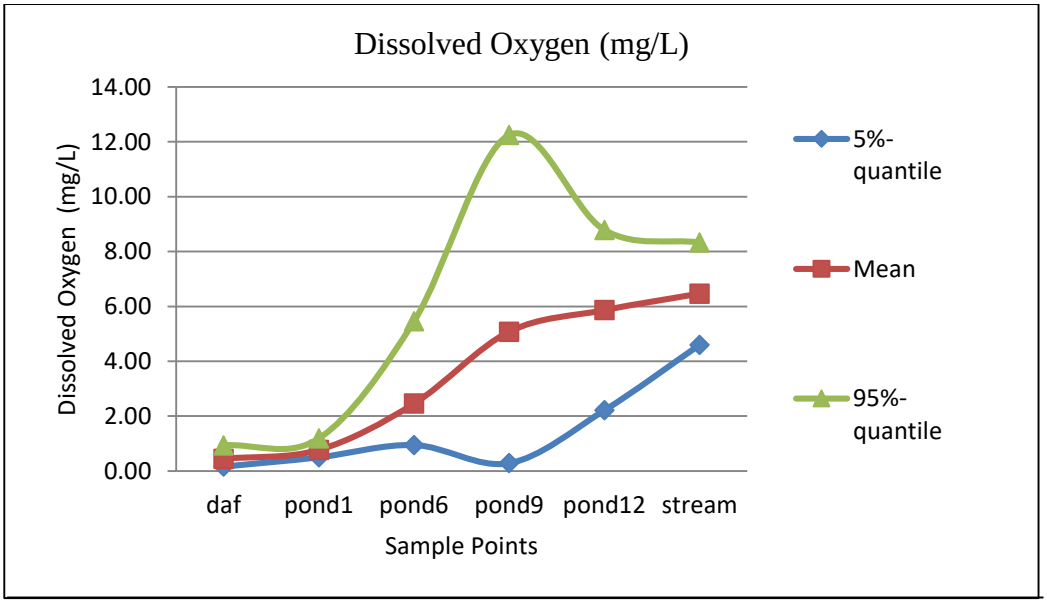


Figure 10.27: The dissolved oxygen (mg/L) of the wetland system (fuzzy Bayesian model) for the duration of the sampling, but the data has been analysed with respect to the correlations between the dissolved oxygen data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the dissolved oxygen. Increased dissolved oxygen behaviour was evident towards the back-end of the wetland system, from pond 9

and pond 12. The DAF and pond 1 exhibiting very low concentrations values with low variation. The analysis depicts large variations occurring within the back-end of the system.

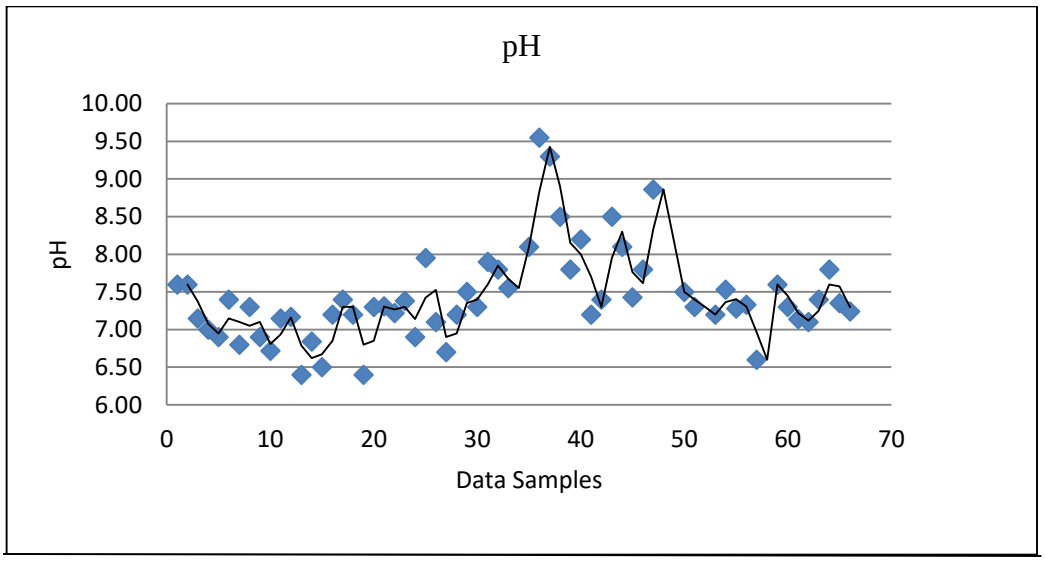


Figure 10.26: The pH of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account the upper and lower boundary conditions of the data.

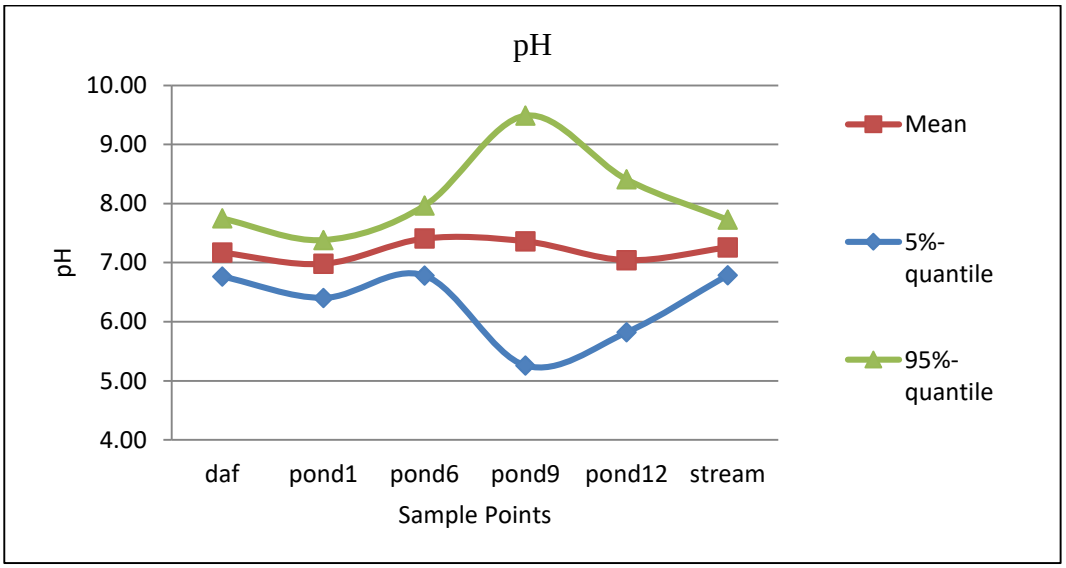


Figure 10.27: The pH of the wetland system (fuzzy Bayesian model) for the duration of the sampling, but the data has been analysed with respect to the correlations between the pH data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the pH. Increased pH behaviour was evident towards the back-end of the wetland

system, from pond 9 and pond 12; with DAF and pond 1 exhibiting low variation. The analysis depicts large variations occurring within the back-end of the system, in particular pond 9.

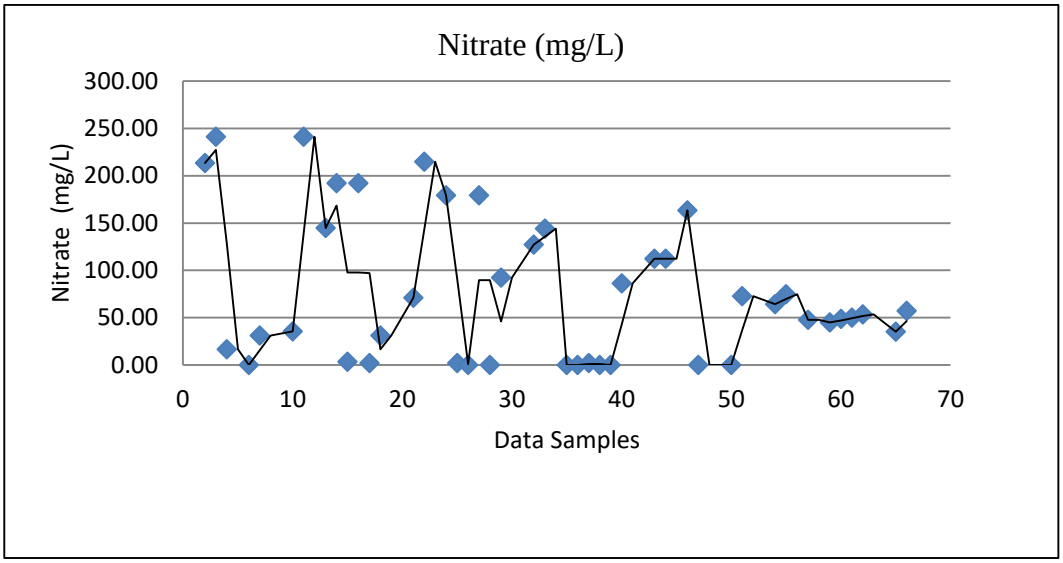


Figure 10.28: The nitrate concentration (mg/L) of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account upper and lower boundary conditions of the data. Noticed the "pulsed" behaviour of the nitrate as it migrates through the wetland system.

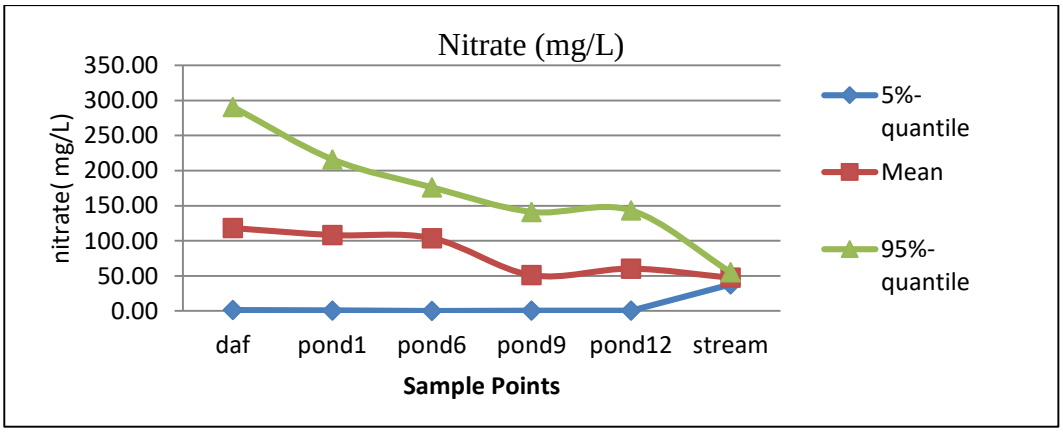


Figure 10.29: The nitrate concentration (mg/L) of the wetland system (fuzzy Bayesian model) for the duration of the sampling, but the data has been analysed with respect to the correlations between the nitrate data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the nitrate. Large pulsed variation in nitrate was

observed through-out the wetland, with zero nitrate values (limit of detection on the IC was 0.05 mg/L) observed in all the sampling points except for the stream. Pond 9 shows a decreasing shift in mean nitrate values, from Pond 6 to Pond 9.

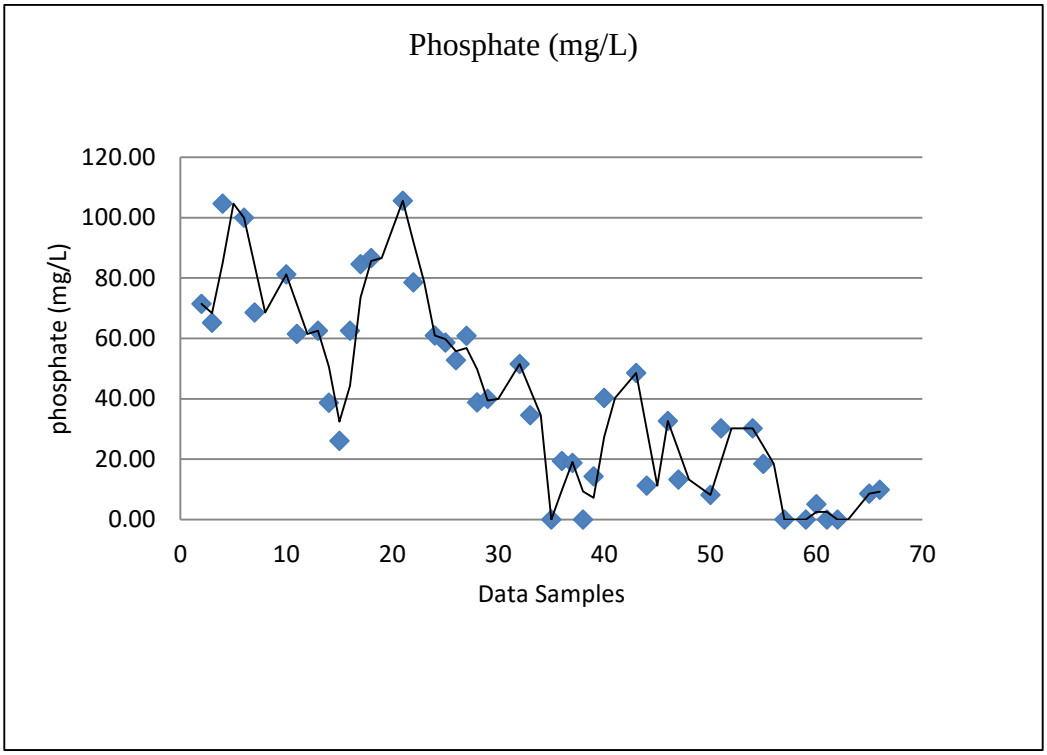


Figure 10.30: The phosphate concentration (mg/L) of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account upper and lower boundary conditions of the data. Notice the pulsed behaviour of the phosphate as it migrates through the wetland system.

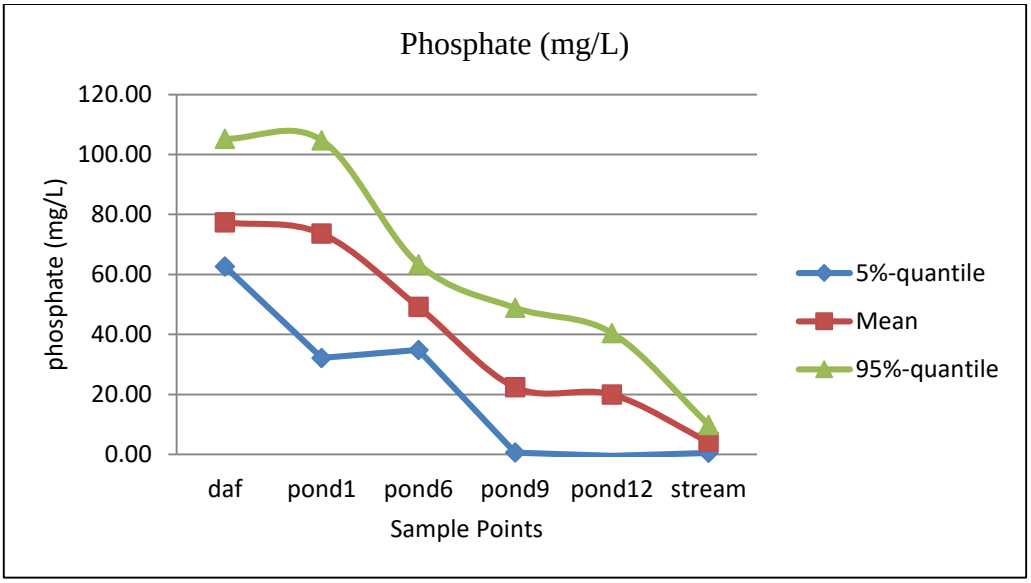


Figure 10.31: The phosphate concentration (mg/L) of the wetland system (fuzzy Bayesian model) for the duration of the sampling, but the data has been analysed with respect to the correlations between the phosphate data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the phosphate. A medium pulsed variation in phosphate was observed through-out the wetland system.

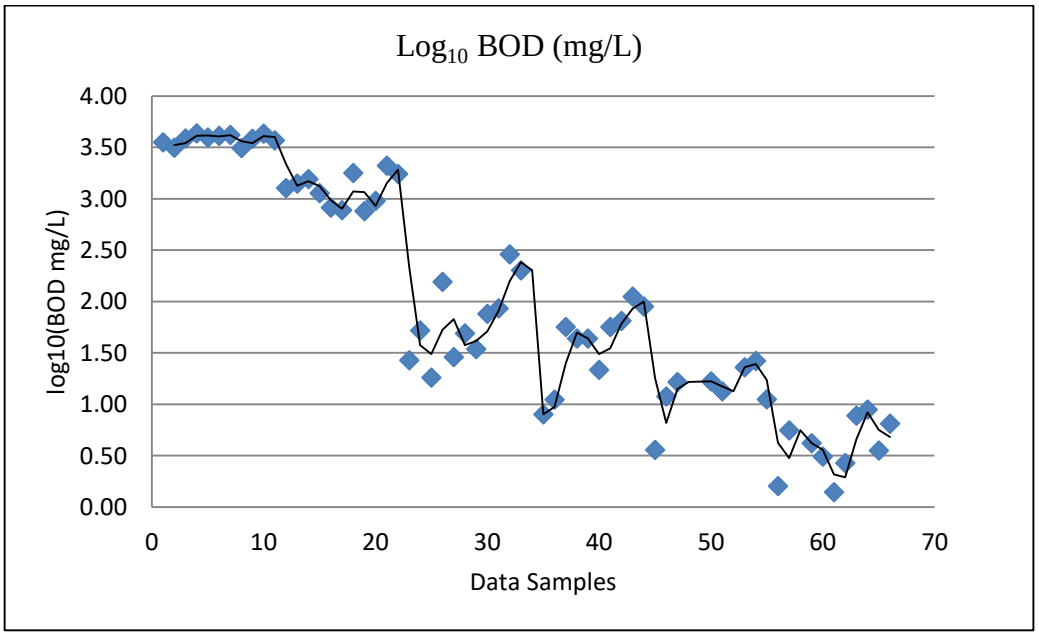


Figure 10.32: The $\log_{10}$ BOD (mg/L) of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data gathered, but does not take into account the upper and lower boundary conditions of the data.

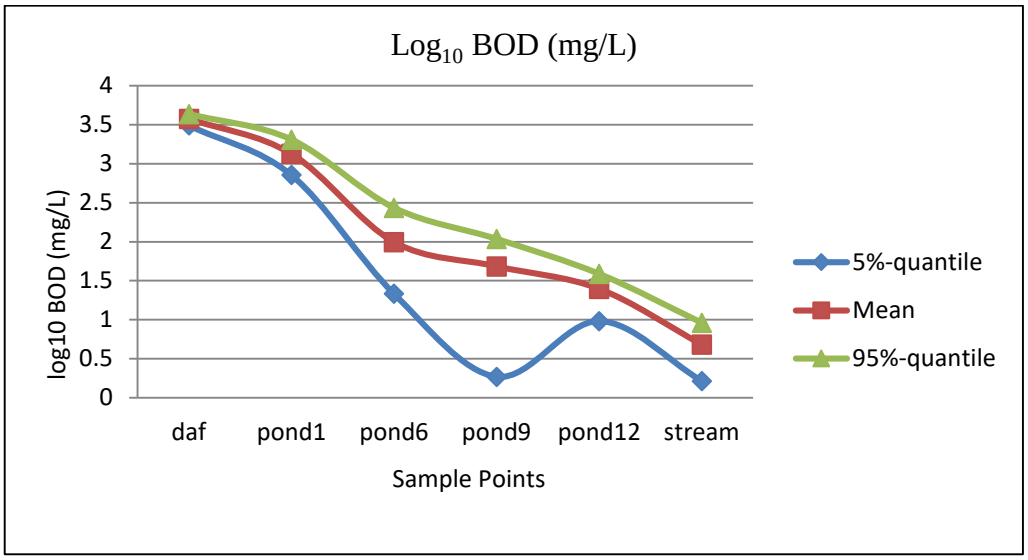


Figure 10.33: The $\log_{10}$ BOD of the wetland system (fuzzy Bayesian model) for the duration of the sampling, but the data has been analysed with respect to the correlations between the $\log_{10}$ BOD data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the BOD. The $\log_{10}$ function was invoked to better reveal the BOD behaviour within the wetland system pond 12 shows a slight increase/ variability in BOD within the wetland.

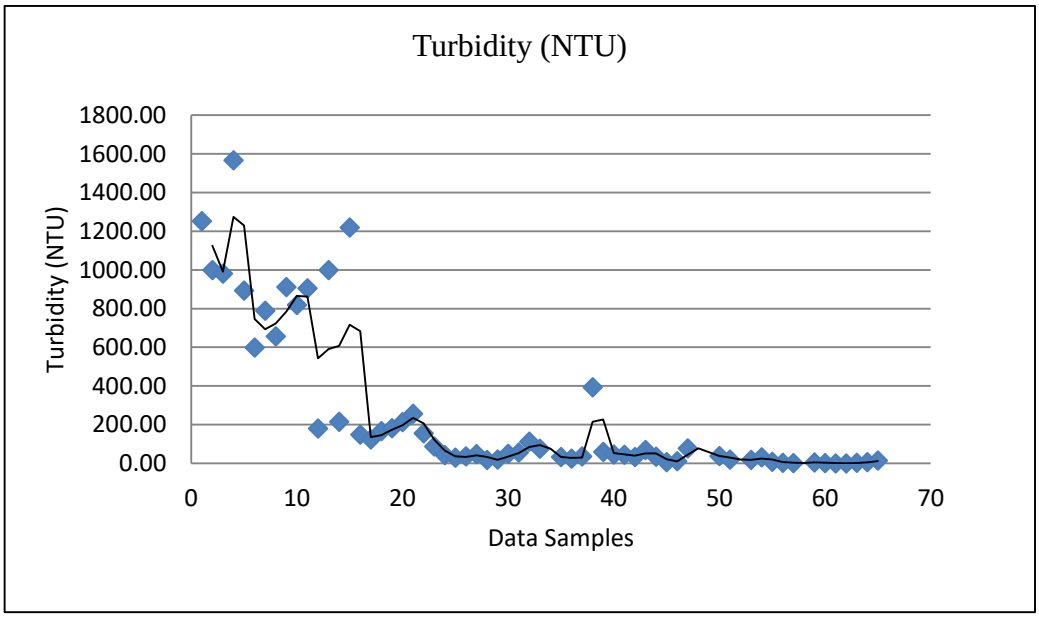


Figure 10.34: The turbidity (NTU) of the wetland system (no model) for the duration of the sampling. The line indicates the optimum trend of the data

gathered, but does not take into account the upper and lower boundary conditions of the data.

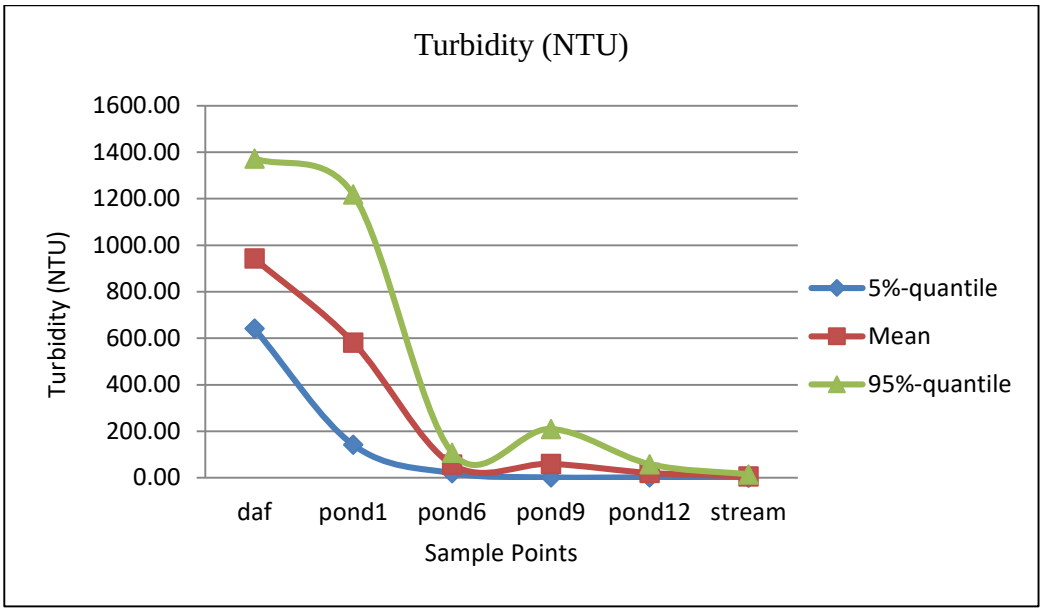


Figure 10.35: The turbidity (NTU) of the wetland system (fuzzy Bayesin model) for the duration of the sampling, but the data has been analysed with respect to the correlations between the phosphate data for each sample point (parent node). The analysis shows the 5% quantile and 95% quantile and the mean of the turbidity. The DAF and pond 1 show large variability, with pond 9 displaying the largest variability towards the back end of the wetland.

10.33 Observations: Fuzzy Bayesian review of wetland variables (Child Nodes)

The graphical displays of the fuzzy index Bayesian analysis indicate upper, lower and mean trends can be evaluated for each variable, in comparison to their corresponding raw data graphs. The analysis reveals the variation within the wetland system, but in particular pond 9 has shown a lot of variation with respect to the following variables; dissolved oxygen, pH, BOD, $\log_{10}$ total bacteria concentration, with strong to medium pulsed behaviour evident through the wetland system. These pulsed events were especially evident with respect to nitrate and phosphate concentrations within the wetland. The nitrate sampling reveals a regular pattern within the wetland and within each sampling point

location. The graphs display a visual distinction between the moving average trend data and the resultant Fuzzy Bayesian (FuzBayes) analysis. The fuzzy Bayesian resolves the wetland variables (children) to resemble their parents (sample points), in that three trend lines are evident, mean, 5% quantile (warm weather) and 95% quantile (wet weather), see Figure 10.16. Therefore, from the UniNet BBN, all the graphs displayed as a result of the causal and diagnostic models (66 variables/nodes – 6 parent nodes and 60 child nodes) are all resolved to be similar i.e. All of the fuzzy Bayesian graphs display three trend lines, 5% quantile, mean and 95% quantile for the parents (sample points) and the children (wetland variables) within the models.

10.34 Sensitivity analysis

The data interaction with model 1 and model 2 was previously reviewed, in relation to probability distribution functions of the wetland variables (children) and the sample points (parents). One of the main purposes of modelling within complex ecosystems is to understand the sensitivities and resilience components of the ecosystem. In this section the sensitivities of causal and diagnostic fuzzy Bayesian models 2 and 3 respectively, are reviewed. Using UniSens we have reviewed the correlations of all the interactions in both cause-and–effect (causal) and effect-and-cause (diagnostic). Table 10.15 and 10.16 shows the sensitivity correlation outputs (1) The correlation analysis using product moment correlation and rank correlation (2) the global correlation analysis using correlation ratio (CR). The tables record the causal, see Table 10.15 and diagnostic, see Table 10.16 correlations from models (2) and (3).

Table 10.15: Causal Correlations

Sample Points	Product Moment Correlation	Rank Correlation	Correlation Ratio
DAF	0.360	0.373	0.441
Pond 1	-0.097	-0.095	0.186
Pond 6	-0.031	-0.031	0.144
Pond 9	0.120	0.122	0.225
Pond 12	0.277	0.296	0.482
Stream	0.281	0.291	0.246

Table 10.16: Diagnostic Correlations

Sample Points	Product Moment Correlation	Rank Correlation	Correlation Ratio
DAF	0.144	0.146	0.096
Pond 1	-0.037	-0.034	0.085
Pond 6	-0.004	-0.004	0.086
Pond 9	0.128	0.129	0.094
Pond 12	0.097	0.124	0.047
Stream	0.130	0.138	0.061

The above Tables 10.15 and 10.16 show the causal and diagnostic correlation results respectively from the UniSens analysis. The Product Moment Correlation (Pearson's), the Rank Correlation and the Correlation ratio. The Product Moment Correlation and Rank Correlation have similar results for both causal and diagnostic analyses. The difference between the two correlations is evident in the Input (DAF) and at the output (pond 12-Stream). See Table 10.15, causal Product moment correlation and the corresponding rank correlations and Table 10.16, diagnostic Product moment correlation and the corresponding rank correlations
We review the causal Spearman's rank correlations for all the sample points within the wetland, see Figure 10.36.

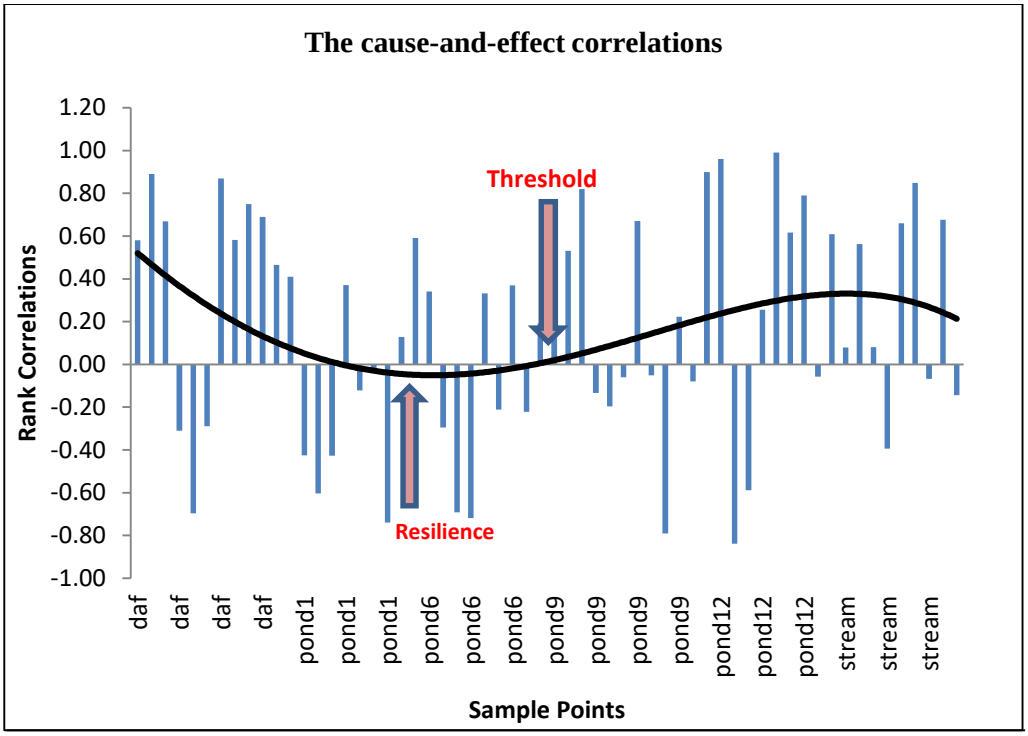


Figure 10.36: The cause-and-effect (causal/top down) rank correlations (Model 2) network of the wetland system. The positive and negative Spearman's rank correlations for the data variables as seen in decentralised causal based Bayesian network.

The black polynomial trend-line shows the mean rank correlation for each sample point across the specific wetland variables for the sample point, revealing a wave-form trend within the wetland. Highlighted are the resilience area (ball-and-cup), and the threshold point on the diagram at the interface between pond 6 and pond 9.

This wave form was also viewed when discriminant analysis was performed- See discriminant analysis, see Figure 3.18 - chapter 3.

To better reveal, understand and simplify the behaviour of the wetland's resilience through causal rank correlations, we super-impose the sample points mean fuzzy Bayesian Weibull efficiencies onto the mean causal rank correlations, see Table 10.15. Figure 10.37 provides a graphical representation.

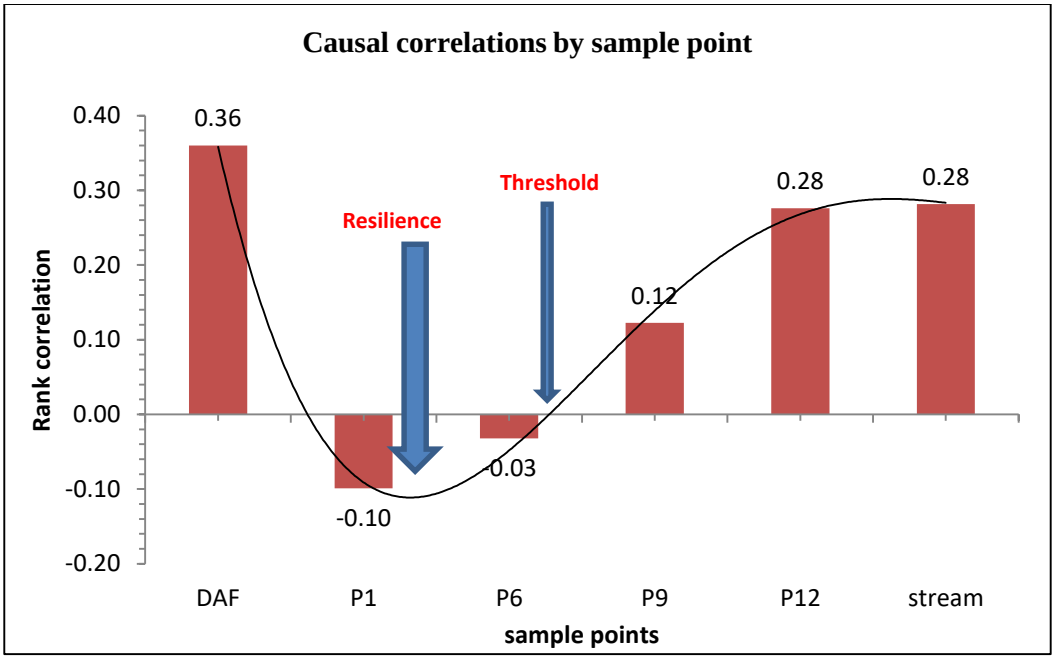


Figure 10.37: The wetlands rank causal correlations imposed on the sample points. The mean sensitivities (rank correlations) of each sample point, with their corresponding mean fuzzy Bayesian Weibull efficiencies. The resilience (cup-and-ball) area is shown, plus the threshold point within the wetland system.

The wetland causal sensitivity using rank correlations indicates a "basin of attraction" within the system as shown on the Figures 10.38 and 10.39, was located between pond 1 and pond 6. At pond 6 the threshold is shown. The fuzzy Bayesian wetland efficiency values are also shown (DAF = 8%, pond 1= 29%, pond 6 = 70%, pond 9 = 63%, pond 12 = 48%, stream = 76%), indicating each sample points mean efficiency. But a second basin of attraction is not evident, as proposed by Holling (1996) and Liao (2012) based on ecological resilience concept, see Figure 10.7.

The diagnostic Spearman's rank correlations for all the sample points within the wetland was reviewed, see Figure 10.38.

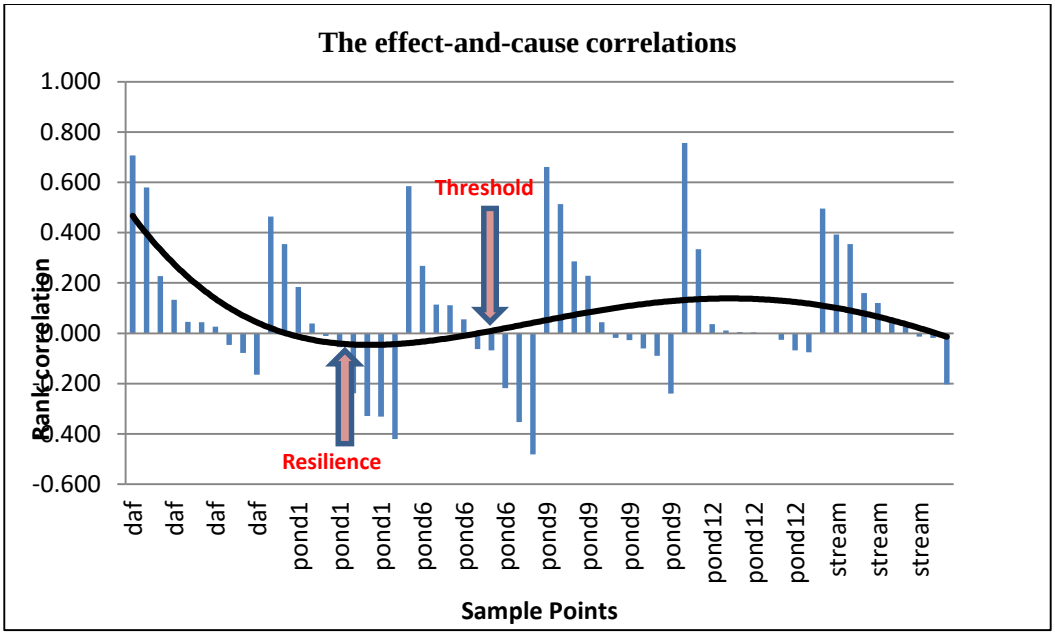


Figure 10.38: The effect-and-cause (diagnostic/bottom up) rank correlations (Model 3) network of the wetland system. The positive and negative Spearman's rank correlations for the data variables as seen in decentralised diagnostic based Bayesian network see Figure 10.14a and 10.14b.

The black polynomial trend-line shows the mean rank correlation for each sample point across the specific wetland variables for the sample point, revealing a wave-form trend within the wetland. Highlighted are the resilience area (ball-and-cup), and the threshold point on the diagram at the interface between pond 6 and pond 9. The trend line is similar in its behaviour to Figure 10.37.

To better reveal, understand and simplify the behaviour of the wetland's resilience through diagnostic rank correlations, we super-imposed the sample points mean fuzzy Bayesian Weibull efficiencies onto the mean diagnostic rank correlations, see Table 10.16. Figure 10.39 provides a graphical representation.

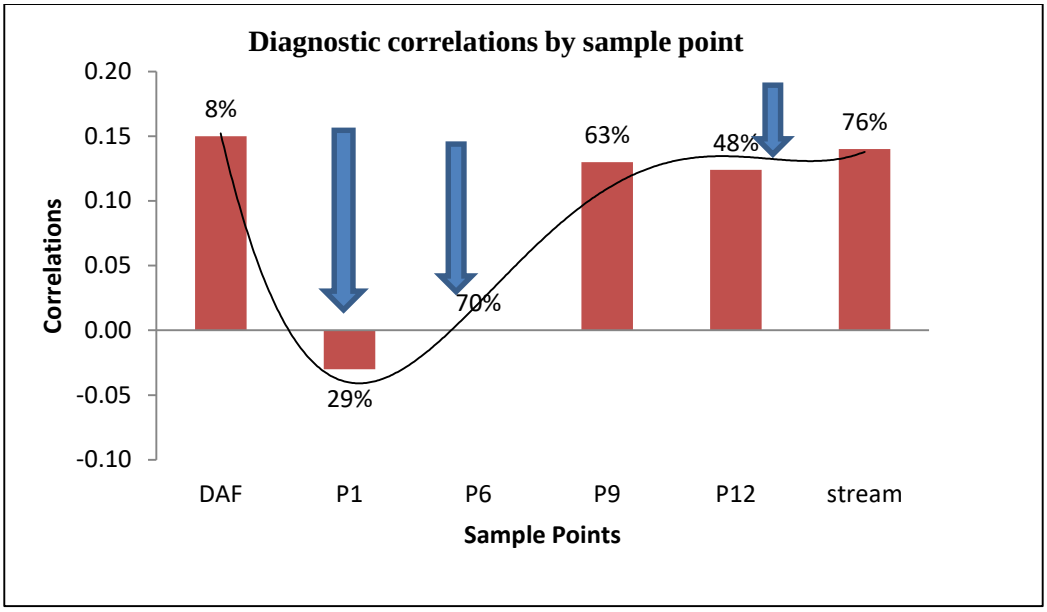


Figure 10.39: The wetland diagnostic correlations imposed on the sample points. The mean sensitivities (rank correlations) of each sample point, with their corresponding mean fuzzy Bayesian Weibull efficiencies. The mean sensitivities (rank correlations) of each sample point, with their corresponding mean reliabilities. The first basin of attraction located between pond 1 and pond 6. A second (smaller) basin of attraction exists and located in the vicinity of pond 12. The threshold occurs at pond 6. See Figure 10.40 where the basins of attraction are more prevalent.

The wetland diagnostic sensitivity using rank correlations reveals not one "basin of attraction" but a second smaller subtler "basin attraction" towards the back-end of the wetland system at pond 12, revealing a hidden effect within the wetland. There was a system wide decrease in sensitivity across all points, except between pond 6 and pond 9, which shows a minor difference between the top down (0.13) network and bottom up (0.12) network. Therefore, within the wetland system between pond 6 and pond 9, there exhibits a stable region regardless of top down or bottom up decentralised BBN, i.e. between pond 6 and pond 9 exhibits stability in variation, regardless of the two decentralised BBN models. It should be noted that both pond 1 and pond 12 show the greatest variation in fuzzy Bayesian efficiency, see Figure 10.40, which highlights the large variations at the input (DAF-pond1) and the output (pond 12-stream). Both the causal and diagnostic graphs see Figures 10.37 and 10.39 respectively;

indicate a threshold event in the region of pond 6. This can be explained by the mean fuzzy Bayesian efficiency readings 70% for this pond. The pond has the largest fuzzy Bayesian efficiency within the wetland sample points. After pond 6, pond 9's fuzzy Bayesian efficiency decreases to a 63% and a further decline was present in pond 12 with 48%.

Both the causal and diagnostic sensitivities along with the fuzzy Bayesian efficiency of the constructed wetland, were reviewed, see Figure 10.40.

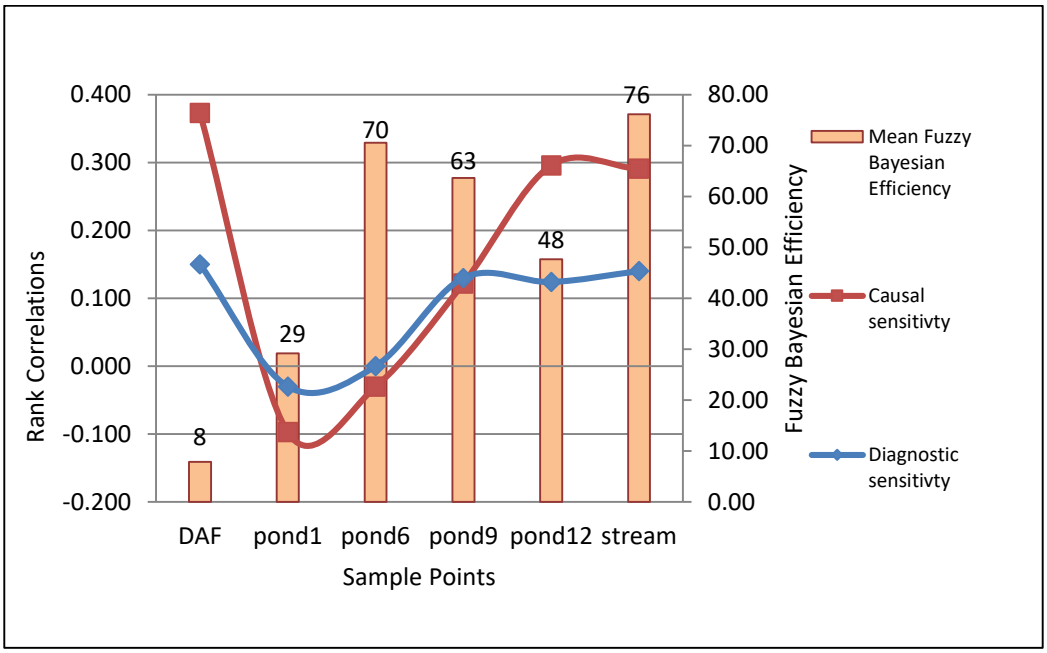


Figure 10.40: The rank correlations for causal and diagnostic sensitivities (fuzzy Bayesian efficiency per sample point) shows the mean fuzzy Bayesian efficiency and the causal and diagnostic sensitivities (cup-and –ball diagrams) based on Spearman's rank correlations; for all the analysis, 400,000 iterations were undertaken.

The graph in Figure 10.40, depicts the mean cause and effect (causal/predictive) sensitivity and its inverse the mean effect and cause (diagnostic) sensitivity. A large deviation exists between the sensitivities, at the input (DAF) and beginning from pond 9 to the output (pond 12 into the stream). The DAF has a very low mean fuzzy Bayesian efficiency reading at 8%. As we progress into pond 1 the efficiency increases to 29% and the attractor basin for the cup-and-ball diagram builds. Pond 6 has a high efficiency with a value of 70%. Both the causal and diagnostic sensitivities trend close together. After pond 6 the efficiency decreases

to 63% for pond 9. The causal and diagnostic sensitivities cross at this point and another large deviation can be viewed. From pond 9 to pond 12 a further decrease in efficiency occurs, with pond 12 recording a value of 48% and the stream records' an increase in efficiency of 76%. Potentially the possibility exists that on average pond 12 has a poorer efficiency in polishing the wastewater. The causal (cause and effect) rank correlation sensitivity displays similar traits to Figure 10.7, the engineering resilience concept, indicating that no second basin of attraction exists within the constructed wetland. On reviewing the diagnostic (effect and cause) model to further ascertain any hidden effects within the wetland, we find that the resilience profile resembles the ecological resilience as depicted in Figure 10.8. Indicating that the wetland system is an undesirable regime and vulnerable and not functioning correctly as a constructed wetland. The main agents to the vulnerability and loss of functionality are ponds 9 to 12.

A further sensitivity study was performed using the correlation ratio metric i.e. Global Sensitivity Analysis (GSA). The global sensitivity analysis considers the whole variation range of inputs (Saltelli *et al*., 2000a) and one of the main methods to review GSA was to record the correlation ratio (CR). The correlation ratio data was taken from Table 10.15 and 10.16 for the causal and diagnostic sensitivities and plotted, see Figure 10.41.

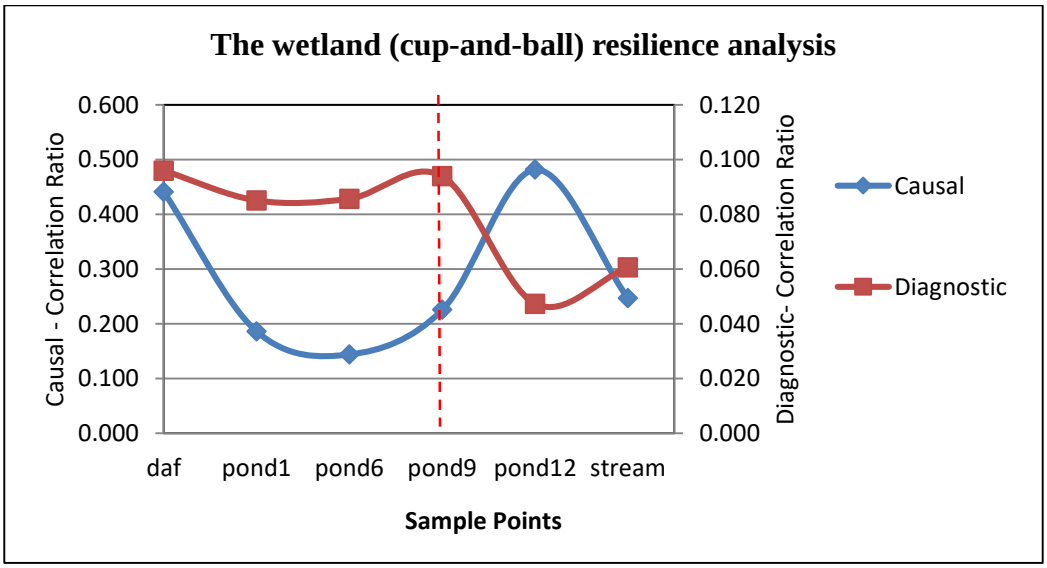


Figure 10.41: The wetland global sensitivity analysis of both decentralised BBNs. Using the correlation ratio for causal and diagnostic models i.e. The wetland "cup-

and-ball" resilience analysis. The red line defines the threshold point within the Diagnostic GSA.

10.35 Discussion

Using Global Sensitivity Analysis (GSA) as per the correlation ratio data, see Tables 10.15 and 10.16. Figure 10.41 shows the causal (predictive) GSA and the diagnostic GSA i.e. correlation ratio. The causal graph is similar to the engineering resilience diagram, see Figure 10.7 and the causal Spearman's rank correlation graph, and see Figure 10.40. There is no threshold within the system when taking pond 12 as the end point of the wetland. The wetland does not display the expected ecological resilience diagram of two basins of attraction. The diagnostic (bottom-up) correlation ratio graph is similar to the Case A diagram, See Figure 10.8. Where the two basins of attraction exits, but in the wrong configuration, leading to an undesirable system i.e. system displays a narrow or shallow basin. The red dotted line denotes the threshold point, based on the diagnostic GSA.

The causal 'cup-and-ball' diagram does not define any threshold within the wetland system i.e. between the DAF and pond 12 as defining an ecological threshold concept. When reviewing the diagnostic GSA, the threshold point within the wetland system, see Figure 10.41 (dotted red line = pond 9) where a change in the basin of attraction between the front part of the wetland i.e. Daf to pond 6 and the back-end on the wetland system pond 9 to stream is present. The cause and effect (causal/predictive) based on global sensitivity analysis and its inverse, the effect and cause (diagnostic) global sensitivity analysis have described the duality of complex interactions within an ecosystem i.e. constructed wetland.

The fuzzy Bayesian analysis from model (1) reveals that the upper and lower limits of the initial fuzzy index data from section 2 (F-IND) has changed. When the data from the F-IND analysis and fuzzy Bayesian Weibull distribution analysis are compared, see Figure 10.17. The difference in the efficiencies exists between pond 9 and pond 12. Pond 9 initial fuzzy index efficiency value was 71.66% after the fuzzy Bayesian analysis decreased to 63.63%. Similarly pond 12 initial fuzzy index efficiency value was 58.55% which decreased to 47.70%.

Figure 10.16 indicated the extremes that exist in the constructed wetland in particular at pond 9 and pond 12. The analysis does reveal that pond 12's fuzzy Bayesian efficiency was lower than the stream, therefore the analysis strongly suggests that pond 12 would be inefficient in treating the abattoir wastewater.

The UniNet software resolved the parents (sample points) and child nodes (wetland variables) to display the same three trends lines, the 5% quantile (warm weather data – May 2007 to September 2007), the 95% quantile (wet weather data – October 2007 to February 2008) and the resultant mean overall data. The initial resilience analysis involving predominantly the Spearman's rank correlation of both the causal (predictive/cause and effect) and diagnostic (effect and cause), indicates that using the causal rank correlations, the wetland system behaves more like an engineering resilience system rather than the expected ecological resilience concept i.e. multiple basins of attraction. Figure 10.40, potentially reveals the large variations in the input DAF between the causal and diagnostic sensitivities and from pond 9 to pond 12, suggesting that issues exists at both the input of treatment process (DAF plant) and towards the back-end of the wetland system (pond 9 to pond 12). Reviewing its inverse sensitivity, the diagnostic resilience graph, highlights that the system is vulnerable and cannot maintain its functionality. The final resilience analysis using the global sensitivity (GSA) correlation ratio (CR) data, re-enforces the idea that the constructed wetland has no second basin of attraction, as seen by the causal (GSA) graph, see Figure 10.41. Reviewing the causal sensitivity inverse function and the diagnostic profile, the resultant profile resembles the undesirable regime as seen in Figure 10.8.

10.36 Conclusions

The BBNs were constructed using information from the 'reduction, iteration and classification' principal components analysis PCA technique (section 1). The subsequent information gathered from the PCA method was utilised to create a fuzzy indices framework for each of the individual sampling points within the wetland. The fuzzy indices for each sample point was evaluated (section 2) and recorded. The fuzzy indices data was used to create the backbone of the BBN model (1); using the Weibull probability distribution to determine the individual

efficiencies for each sample point (parent). This model became the foundation model for the remaining decentralised causal and diagnostic models. These models were then utilised to analyse the wetland variables. The sample points (parents) and their children (wetland variables) were characterised with the same three trend lines; 5% quantile (warm weather- lower trend) the 95% quantile (wet weather – upper trend) and the overall mean, therefore no bias was invoked within the models, the parents and children displayed the same trend behaviours. The Bayesian model outputs via the UniNet and UniSens provided ample information on the wetland variable interactions and their subsequent sensitivities, thus revealing the constructed wetlands resilience. The use of resilience visualised by the cup-and-ball heuristic diagrams indicates several important behaviours within the constructed wetland. The front end of the system from pond 1 to pond 6 shows strong resilience to perturbations that occur through the seasons, in comparison to the back end of the system from pond 9 to pond 12, which overall exhibits poor resilience to both the rank correlations and the diagnostic global sensitivity analysis (correlation ratio). The rank correlation analysis reveals the constructed wetland behaves similar to an engineered regime with one basin of attraction. The ideal ecological regime expects two or more basins of attraction. The diagnostic rank correlation analysis suggests a resilience system that is vulnerable with the likelihood of loss of functionality. To elaborate and test the rank correlation hypotheses further, the global sensitivity analysis was employed through the use of correlation ratio metrics. The causal GSA was similar to the rank correlation analysis with the outome that the wetland system was an engineered regime not an ecological regime. The diagnostic GSA confirmed the system was overall an unsuitable or undesirable regime, confirming the constructed wetland was not optimum for its designed function of treating abattoir wastewater. Potential issues are present in ponds 9 to pond 12. Ponds 9 through 12 show poor resilience and efficiency in treating the effluent waste from the abattoir. A number of factors come into play here, the potential causes for this, as revealed from chapter 6 are the poor soil porosity identified in ponds 9 through 12 in comparison to the soils of pond 1 and 6. There was no protection barrier placed into the wetland system to minimise percolation effects thus maintaining water within the wetland ponds. In contrast high density polyethylene (HDPE) barriers are used in landfills to protect the environment from contaminants

entering the local aquifers. One of the main aspects within resilience analysis is to identify the threshold point or the critical threshold within the wetland system, where a noticeable phase change has occurred. The most likely region within the wetland exists between pond 6 and pond 9 thus either pond 7 or 8 the most likely regions i.e. the pond/s with the largest efficiency.

Disadvantages exist within the construction of these complex models. The time and effort to create these complex models can deter environmental scientists to undertake these hybrid complex models. After the models' creation the user must dedicate time to sift through the data such as the sensitivity analysis results to locate the relevant result. To enable the user to visualise these results and therefore comprehend the complex interactions of different wetland variables (parents and children), time must be allocated to create the cup-and-ball diagrams. But the major advantages to these complex hydrid models is that once they are created they can re-used over time, as the data increases. Because the model methodology is data centric based i.e. we are only utilising the data from wetland. These models could be applicable to other complex system analysis to establish risks within financial, banking and power grid systems, to verify how resilient these systems are to shocks and hidden signals.

Chapter 11. Conclusions and recommendations.

11.1 Conclusions

In Ireland, the application of constructed wetlands to treat high strength wastewaters is not common and abattoir wastewater effluent is considered one of the most extreme of industrial wastewaters for treatment. A rigorous statistical and mathematical undertaking was employed to review the processes underlying the performance of a constructed wetland system treating abattoir wastewater. Constructed wetlands are emerging as a cost effective eco-friendly alternative to treat wastewaters, with multiple processes occurring in unison, such as physical, chemical, biological and all of these processes under the influence of varying weather conditions. This research attempts to understand these conditions in a multi-faceted modelling approach. Different statistical and mathematical approaches were used to delve into the wetlands "black-box" to try and comprehend the complex internal operations within a constructed wetland, enabling stakeholders to understand and identify these issues and to address them from this research.

The constructed wetland (CW) has changed when this research was undertaken. Ponds 9 to 12 were removed and replaced with a willow tree plantation. Therefore, the CW is a zero output system, no wastewater leaves the wetland system and enters the local stream. Instead the willow tree planatation bio-remediates the wastewater through phyto-remediation; the use of plants to reduce, remove, stabilise and collect contaminants into the plant tissue.

The conclusion is that the constructed wetland under study has many potential difficulties in effectively, consistently and reliably treating the abattoir effluent wastewater. A synopsis of each chapter provides insights into the issues affecting the constructed wetland.

11.2 A synopsis of the thesis by chapter

Each chapter was reviewed and a brief synopsis provided under the following five headings: (1) Model/s used (2) Model/s explained (3) Variables of interest (4) Sample points of interest and (5) Learning outcomes.

11.3 Chapter 3. Multi-variate analysis of a constructed wetland (CW)

Models used: Five statistical modelling techniques were applied. They were the analysis of variance (ANOVA), canonical variate analysis (CVA), principal component analysis (PCA) and canonical discriminant analysis (CDA) used in conjunction with discriminant functional analysis (DFA).

Model/s explained: Using simple and advanced statistical techniques (multi-variate analysis) to highlight dominant wetland variables and sample points of interest within the CW. The potential of resilience or robustness was introduced using the DFA technique. This concept appears in greater detail in chapter 10, section 3.

Variables of interest: Indicator bacteria (total coliforms, *E.coli*, enterococci), dissolved oxygen, redox potential and turbidity.

Sample points of interest: Individual sample points and the entire wetland system.

Findings: Two-way ANOVA without replication revealed a 'switch' between individual ponds and the complete wetland system with respect to bacteria. CVA revealed an asymmetry between the indicator bacteria, potentially highlighting a second source of *E.coli* entering the CW. PCA confirmed the dominance of indicator bacteria and DFA graphically depicted the trend of wetland variables in the CW, with potential resilience found between pond 1 - pond 6.

11.4 Chapter 4. Constructed wetland: Traditional models

Models used: The Mara and Marais continuous stirred reactor models (CSTR), Oakley Dispersed flow model (DFM), New Oakley Dispersed flow model (new DFM).

Model/s explained: Using indicator bacteria as tracers to understand the hydraulic retention times (HRTs) in the CW. Development of a new Oakley DFM model.

Variables of interest: HRT. The concept that indicator bacteria (total coliforms, *E.coli*, enterococci) can be successfully used as tracers in CWs.

Sample points of interest: Pond 9 and pond 12.

Findings: The wetland moves between two model scenarios, the CSTR and the new Oakley DFM. The CW displayes two distinct dichotomies, long-circuiting and short-circuiting, within the sampling period. The HRTs combined with turbulence (short-circuiting) and stagnant (long-circuiting) flows. Long-circuiting displays very long retention times, especially during the summer dry months, where pond 9 and 12 have little or no water, thus increased HRTs (mean HRT 128 days), well above the initial design HRT of approx. 100 days. Short-circuiting displayed were short retention times, especially during the wet winter months (mean HRT 54 days), below the design of HRT of approx. 100 days.

11.5 Chapter 5. Bio-aerosol analysis of the constructed wetland using the SAS-90 (SAS) meter

Models used: Population counts of bio-aerosol bacteria using surface air system Super-90 (SAS) across different agar media.

Analysis explained: The effects of odour and air borne bacteria in the vicinity of the constructed wetland.

Variables of interest: Air-borne bacteria: *E.coli*, *Klebsiella*, *Enterobacter*, *Pseudomonas aeruginosa* and *Salmonella* and climate variable air temperature.

Sample points of interest: DAF and pond 1.

Findings: The use of Membrane Lactose Glucuronide agar (MLGA) to potentially identify air-borne bacteria. With green colonies were indicative of *E.coli*, with yellow hues potentially indicating *Klebsiella* and *Enterobacter* bacteria and pink hues potentially indicative of *Pseudomonas aeruginosa* and *Salmonella*. The DAF plant potentially the source of the pathogenic air-borne bacteria.

11.6 Chapter 6. A Bayesian network analysis of the wetlands soil and sludge

Models used: Soil and sludge by Bayesian analysis including sensitivity analysis.

Model/s explained: Using Bayesian and sensitivity analysis (Mutual Information = Entropy) on the wetlands soil on the following parameters: bacteria, inorganic ions and wetland design and plant coverage.

Variables of interest: Wetland area and soil porosity/ permeability.

Sample points of interest: pond 9 and pond 12.

Findings: Using both Bayesian and sensitivity analysis (Mutual information = Entropy) provides a two-tier modelling system. The Bayesian belief network finds that soil porosity dominates the network with 12 linkages, while the sensitivity analysis highlights the dominance of wetland area (wetland design). The wetland porosity in ponds 9 to 12 is approx. 27%, thus susceptible to leakage through the soils.

11.7 Chapter 7. Self-Organising Criticality and Entropy analysis

Models used: Self organised criticality (SOC) via the use of power laws and entropy analysis, reviewing Shannon (information) entropy.

Model/s explained: Self-organised criticality (SOC) is a property of non-equilibrium systems. The system is in a critical state and Power laws are indicative of these critical states. The presence of Power laws may typify the presence of complex systems.

Findings: Indicator bacteria (total coliforms, *E.coli,* enterococci), nitrate, phosphate and water depth.

Sample points of interest: pond 1, pond 9 and pond 12.

Learning outcomes: SOC shows that indicator bacteria and nutrients such as nitrate and phosphate are not effectively treated within the CW. These parameters display linear (indicator bacteria) and exponential (nitrate and phosphate) trends. The entropy analysis reveals two effects within the wetland (1) Over-driven (unstable = ponds 9 and 12) and (2) over-loading, the DAF plant not treating influent waste prior to entering the CW. There is an issue with water depth, the constructed wetland was losing water.

11.8 Chapter 8. (Section 1 and 2). The application of artificial neural networks (ANNs) and self-organising maps (SOMs) to understand wetland behaviours

Models used: Artificial neural networks (supervised modelling) and self-organised maps (SOMs) (unsupervised modelling).

Model/s explained: Three supervised ANN models were developed to understand the complex interactions of the wetland dataset on the indicator bacterial populations (target variable). SOM is an unsupervised clustering technique that identifies clusters in a dataset without rigid assumptions of

linearity or normality that is typified by traditional statistical techniques, no target variables.

Variables of interest: ANN: - pH, redox potential, dissolved oxygen, conductivity, ammonia, wetland volume, area and water depth and climate conditions.

SOM: - similar data profile seen in fuzzy Bayesian analysis - see chapter 10 Part III, dominant variables are redox potential, dissolved oxygen and wetland area.

Sample points of interest: Entire wetland (issues observed in ponds 1, 9 and 12).

Findings: The application of three supervised ANN models and one unsupervised SOM model on the wetland dataset. The ANNs were created to elucidate the important and sensitive variables from the wetland dataset on the indicator bacteria. In the SOM model, all wetland variables were accessed together on a graphical map that defines the clustering effects of these applied wetland dataset. The ANN's required training and validation to implement the modelling strategies, whereas the SOM doesn't. SOM can evaluate "what-if scenarios", reviewing potential effects that could happen. The user can design for worst case scenarios. The SOM provides an excellent graphical overview of the wetland complex dataset.

11.9 Chapter 9. Coefficient of reliability (COR) or Cronbach's Alpha

Models used: Coefficient of reliability (COR), Cronbach's alpha.

Model/s explained: Using reliability analysis, based on coefficient of variation (CV) and coefficient of reliability (COR). Finally, utilising Cronbach's alpha as a control test, to verify the overall COR of the wetland based on (1) the three indicator bacteria and (2) the entire wetland dataset.

Variables of interest: Indicator bacteria (total coliforms, *E.coli*, enterococci).

Sample points of interest: Pond 12.

Findings: The potential exists that understanding bacterial processes within wetland systems, may provide an overall wetland system performance as attested to when the Cronbach's *Alpha* using SPSS®V.22. Using only the three bacterial concentrations, versus twenty wetland variables with values of 0.654 and 0.664 respectively i.e. using three indicator bacteria with an *Alpha* = 0.654 provided a close match to an *Alpha* = 0.664 for all the wetland data.

11.10 Chapter 10. Hybrid fuzzy logic and de-centralised Bayesian belief networks

(Section 1, Section 2 and Section 3)

Chapter 10. (Section1). Iterative and reductive principal component analysis

Models used: Iterative and reductive principal component analysis (PCA) used in conjunction linear interpolation with Kaiser–Meyer-Olkin (KMO) and Bartlett's Test methods to verify wetland variables of interest.

Model/s explained: Using principal component analysis (PCA), plus linear interpolation (LINT) to populate missing data. Two important internal tests were employed to evaluate the wetland data set as to whether the data is of interest regarding the implementation of principal component analysis, i.e. whether the chosen variables are factorable. The indicators are Bartlett's sphericity test (BTS) and the Kaiser-Meyer-Olkin (KMO) test. The initial dataset was reduced from 22 to 7 variables (i.e. the 3rd PCA run) with a KMO value = 0.851.

Variables of interest: BOD, redox potential, conductivity, turbidity, total bacteria, water depth, dissolved oxygen.

Sample points of interest: N/A

Findings: Using methods based on mobile telephony advanced modelling analysis, to resolve the telephony data under analysis to between six and eight

attributes (variables). The methods developed were sufficient to predict customer attrition (churn). The same methodology was used in this chapter using principal component analysis with addition of linear interpolation and oblique testing (Promax) in conjunction with KMO and BTS. To reduce the wetland dataset from 22 variables to 7 dominant variables, which were dissolved oxygen, turbidity, total bacteria, redox potential, BOD, conductivity and water depth.

Chapter 10. (Section 2). F-IND: Fuzzy indices of the environmental variables of the constructed wetlands.

Models used: Fuzzy indices (F-IND)

Model/s explained: The F-IND software developed fuzzy indices from Chapter 10 section 1 (key wetland parameters) that can be used to classify the performance of the sample points within the wetland system for the purposes of determining a wetland efficiency index (WEI). The variables of interest from Chapter 10 section1, are used in the fuzzy logic analysis.

Variables of interest: BOD, conductivity, turbidity, total bacteria, water depth, dissolved oxygen.

Sample points of interest: (DAF- pond 1) and (pond 9 -pond 12).

Findings: The potential exists to use the F-IND method to evaluate a wetland efficiency index (WEI) of individual points within the system based upon selected wetland variables. The grouping category was based on the Persson wetland hydraulic efficiency (λ) template. The combined PCA and fuzzy logic method may be a data alternative the Persson wetland design method, to evaluate wetland efficiency.

Chapter 10. (Section 3). Creation of two fuzzy logic decentralised Bayesian belief networks.

Models used: Using the modelled data gathered in section 2. Two hybrid fuzzy Bayesian decentralised networks were developed.

Model/s explained: Using hybrid fuzzy Bayesian networks to realise the robustness/ resilience of the constructed wetland (complex system). The Bayesian networks are created with parent and child nodes. The parents signifying the sample points via fuzzy indices, the child nodes are the wetland parameters for each sample point. The concept of ball-and-cup resilience model was introduced. The mode locates regions of resilience within the wetland system via global sensitivity analysis.

Variables of interest: Parent nodes: - BOD, conductivity, turbidity, total bacteria, water depth, dissolved oxygen.

Child nodes: - All wetland data.

Sample points of interest: Pond 9 and pond 12.

Findings: The Bayesian model provides critical information on the wetland variable interactions and their subsequent sensitivities, thus revealing the constructed wetlands resilience. Ponds 9 through 12 shows poor resilience and efficiency in treating the effluent waste from the abattoir.

11.11 Actual outcomes arising from this study

- All constructed wetlands treating high strength industrial waste should be lined, with effective high density polyethylene (HDPE 3.0 - 5.0 mm thickness) barrier to minimise leakage into the surrounding soils or local aquifers.

- The use of physical-chemical multi-probe instruments to monitor the constructed wetland ponds as determined by chapter 10, section 1 using principal compontent analysis on the wetland dataset. The resultant output of the PCA shows that a combined pH, redox potential, conductivity, dissolved oxygen instrument will provide critical information for the overall CW performance, so that individual ponds performance and efficacy can be cost-effectively evaluated, not just the input and output.

- The importance of monitoring bacteria within the constructed wetland, as determined by chapter 4, using traditional wetland models and the potential that indicator bactertia can determine the hydraulic retention time. In chapter 9, using Cronbach's alpha tests (coefficient of reliability) verified that the three indicator bacteria data provides very similar COR results to using the entire wetland datset. The research has identified, the critical role that bacteria play in the behaviour of the CW.

- Environmental monitoring plays acritical role in analysing the health of the constructed wetland. A regular monitoring regime for new and establishing constructed wetlands is recommended.

- An effective Irish database of all constructed wetlands, lagoons, ponds is regulated by the Department of the Environment. Including data entry on a monthly, quarterly, yearly basis depending the type and age of the wetland system.

- The creation of a best available technique (BAT) or best available technique not entailing excessive cost (BATNEEC) for constructed wetlands within the Irish

wastewater landscape, inclusive of the points above, especially for those treating high strength industrial wastewaters.

11.12 Future recommendations

The following areas of future research are recommended based on the findings of this thesis:

- The importance of correctly using statistical methods to elucidate important signals within the data under analysis. Methods such a data classification and clustering that was employed in chapter 3 and chapter 8 section 2 respectively.

- Pilot field studies in wetland design research, with large aspect ratio's greater than 3, where the ponds within the pilot studies would be long and narrow. The wetland ponds in this study have aspect ratios > 2, but still the HRT models show short-circuiting occurring, see chapter 4.

- Further research on the new Oakley dispersed flow model (chapter 4), to confirm the model's accuracy and to verify the long-circuiting events occur in constructed wetlands. Constructed wetlands shall have a minimum and maxiumum acceptable HRT.

- To further increase the research in the areas of entropy and self-organised criticality (SOC) and power laws, see chapter 7. The aim would be to develop more advanced mathematical and statistical modelling techniques such as percolation theory, combined with highly optimised tolerance (HOT) to understand thresholds and critical transitions and fragility within these complex systems.

- The use of power laws and self-organised criticality (SOC) in future wetland research will be of crucial importance with regard to increasing uncertainty in climatic events. Natural and constructed wetlands will need to adapt to an uncertain climate conditions and using these mathemtcial tools will enable

reseachers to understand complex interactions within these complex wetland systems.

- Further investigate the potential of using principal component analysis (PCA) with iterative and reduction methods combined with fuzzy logic indices, see chapters 10 section 1 and 2; as a data centric method to establish a wetland efficiency index (WEI). This could be developed as an alternative to using Persson's efficiency model (λ), which is based on the wetland design characteristics, by utilising the data retrieved from the wetland to realise its efficiency.

11.13 Key Findings

The constructed wetland was a data gathering exercise in complex system analysis. An in-depth multi-layered process into complex system modelling and analysis was developed. In particular chapter 10 (fuzzy Bayesian models) are a unique and new method into realising resilience in a complex system, such as a constructed wetland and detailed work was done on visualizing the cause-and–effect and effect-and-cause modelling strategies employed in this new modelling methodology.

The fuzzy Bayesian models developed in ths thesis can be utilised in other complex systems. These models are data driven and are not focused on a precise system structure or design, such as the traditional wetland models in chapter 4, but are purely based on data and therefore are adaptable and transferable to other complex systems such as:

- Environmental systems (understanding process issues within WWTP)
- Banking (reviewing a banks overall resilience/vulnerabilities)
- Macro-economics (reviewing countries overall resilience/vulnerabilities)
- Socio-environmental economics (understanding complex interactions, what are the main components are these interactions)
- Cancer research (reviewing cancer trials in different countries for similar cancers, which trial is more effective in treating a particular cancer.

References

Abbassi, R., Khan, F., Hawboldt, K., (2010). A methodology of finding dispersion coefficient using computation fluid dynamics (CFDs). *International Journal of Water Resources and Environmental Engineering* 2 (5), 114-120.

Akratos, C.S., Papaspyros, J.N.E., Tsihrintzis, V.A., (2008). An artificial neural network model and design equations for BOD and COD removal prediction in horizontal subsurface flow constructed wetlands. *Journal of Chemical Engineering* 143, 96-110.

Akratos, C.S., Papaspyros, J.N.E., Tsihrintzis, V.A., (2009a). Total nitrogen and ammonia removal prediction in horizontal subsurface flow constructed wetland: use of artificial neural networks and development of a design equation. *Biosource Technology* 100, 1656-1663.

Akratos, C.S., Papaspyros, J.N.E., Tsihrintzis, V.A., (2009b). Artificial neural network use in ortho-phosphate and total phosphorus removal prediction in horizontal subsurface flow constructed wetlands. *Biosystems Engineering* 102 (2), 190-201.

Ale, B.J.M., Bellamy, L.J., van der Boom, R., Cooper, J., Cooke, R.M., Goossens, L.H.J., Hale, A.R., Kurowicka, D., Morales, O., Roelen, A.L.C., Spouge, J., (2009). Further development of a Causal model for Air Transport Safety (CATS): Building the mathematical Heart. *Reliability Engineering and System Safety* (94), 1433-1441.

Almeida, C.M.R., Santos, F., Ribeiro, I., Mucha, A.P., (2019). Potential of constructed wetland for the removal of antibiotics and antibiotic resistant bacteria from livestock wastewater. *Ecological Engineering* 129, 45-53.

Anderson, T.W., Darling, D.A., (1952). Asymptotic theory of certain ‘goodness of fit’ criteria on stochastic processes. *Annals of Mathematical Statistics* 23, 193-212.

Anstett-Collin, F., Mara, T., Basset, M., (2014). Application of global sensitivity analysis to a tire model with correlated inputs. *Simulation Modelling Practice and Theory* 44, 54-62.

APHA, (1989). American Journal of Public Health: June 1989, Vol. 79, No. 6

Aral, D.K., Güvenir, A.H., Sabuncuoğlu, I., Akar, R.A., (2012). A prescription fraud detection model. *Computer Methods and Programs in Biomedicine,* 106, 37-46.

Arreguín-Sánchez, F., Ruiz-Barreiro, T.M., (2014). Approaching a functional measure of vulnerability in marine ecosystems. *Ecological Indicators* 45, 130-138.

Atzori, L., Aste, N., Isola, M., (2006). Estimation of multifractal parameters in traffic measurements: an accuracy-based real-time approach. *Computer Communications* 29, 1879-1888.

Ayinde, O., Tang, Y-H., (2002). Face recognition approach based in rank correlation of gabor-filtered images. *Pattern Recognition* 35, 1275-1289.

Babatunde, A., Zhao, Y., (2010): 2nd International Conference on Constructed wetlands for wastewater treatment and Environmental Pollution Control. University College Dublin, Dublin. October 1-2.

Babatunde, O.A., Miranda-Casoluengo, R., Imitiaz, M., Zhao, Y.Q., Meijer, W.G., (2016). Performance assessment and microbial diversity of two pilot scale multi-stage sub-surface flow constructed wetlands systems. *Journal of Environmental Sciences* 46, 38-46.

Babatunde, O.A., Zhao, Q. Y., O'Neill, M., O'Sullivan, B., (2008). Constructed wetlands for environmental pollution control: A review of developments, research and practise in Ireland. *Environment International*, 34 (1), 116-126.

Bak, P., (1996). How nature works, Springer. https://www.springer.com/gp/book/9780387947914 [Accessed 05 2016]

Bak, P., Tang, C., Wisenfeld, K., (1998). Self-organizing criticality, *Physical Review A* 38, 364-374.

Ball, P., (2005a). Critical Mass: how one thing leads to another? p 295. Arrow Books, UK.

Ball, P., (2005b). Critical Mass: how one thing leads to another? p 297. Arrow Books, UK.

Ball, P., (2005c). Critical Mass: how one thing leads to another? p 298. Arrow Books, UK.

Barabási, A.L., (2002). The new science of networks. Cambridge (MA), Perseus Publishing. 144.

Baras, D., Fournier, L., Ziv, A., (2009). Hardware and Software: Verification and Testing: 4th International Haifa Verification Conference, HVC 2008, Haifa, Israel, October 2008, Revised Selected Papers. Vol. 5394, 53-67.

Batina L., Hogenboom J., Woudenberg J. G. J., (2012). Getting more from PCA: First results of using principal component analysis for extensive power analysis. Topics in Cryptology – CT-RSA 2012. *Lecture Notes in Computer Science* 7178, 382-397.

Berthouex, P.M., Hunter, W.G., (1981). Simple statistics for interpreting environmental data. *Journal of Water Pollution Control Federation* 53 (2), 167-175.

Berthouex, P.M., Hunter, W.G., (1983). How to construct reference distribution to evaluate treatment plant effluent quality. *Journal of Water Pollution Control Federation* 55 (12), 1417-1424.

Boutilier, L., Jamieson, R., Gordon, R., Lake, C., Hart, W., (2009). Adsorption, sedimentation, and inactivation of *E.coli* within wastewater treatment wetlands. *Water Research* 43, 4370-4380.

Boyer, J.N., Fourqurean, J.W., Jones, R.D., (1997). Spatial chacterization of water quality in Florida Bay and Whitewater Bay by multivariate analyses: zones of similar influence. *Estuaries and Coasts* 20 (94), 743-758.

Boyer, T., Polasky, S., (2004). Valuing urban wetlands: a review of non-market valuation studies. *Wetlands* 24, 290-295.

Bradley. C., Grapes, T.R., Petts, G.E., (2006). Hydrodynamics of floodplain wetlands in a chalk catchment: The River Lambourn, UK. *Journal of Hydrology* 320, 324-341.

Brandi, G., Sisti, M., Amagliani, G., (2000). Evaluation of the environmental impact of microbial aerosols generated by wastewater treatment plants utilizing different aeration systems. *Journal of Appied Microbiology* 88, 845-852.

Brazil, D., Murphy, G., (2005). Microbiological Laboratory Manual. Department of Science and Health, Institute of Technology, Carlow.

Britto, D.T., Kronzuker, H.J., (2002). NH_4^+ toxicity in higher plants: a critical review. *Journal of Plant Physiology* 159, 567-584.

Brix, H., Sorrell, B.K., Lorenzen, B., (2001). Are phragmites-dominated wetlands a net source or sink of greenhouse gases? *Aquatic Botany* 69, 313-324.

Brown, DS; Kreissi, J. F.; Gearhart, RA., Kruzic, A. P.; Boyle, W.C.; Otis, R.J. (1999) Manual of Constructed Wetlands Treatment of Municipal Wastewaters; EPA-625/R-99/010; US Environmental Protection Agency, National Risk Management Research Laboratory, Office of Research and Development; Cincinnati, Ohio.

Buchauer, K., (2007). Comparison of Model Approaches for Predicting Coliform Removal in Waste Stabilization Ponds. *Water and Environmental Journal* 21, 108-113.

Bullock, A., Acreman, M., (2003). The role of wetlands in the hydrological cycle. *Hydrology and Earth systems Science* 7, 358-389.

Bustillo-Lecompte, C., Mehrvar, M., (2017). Slaughterhouse Wastewater: Treatment Management and Resource Recovery. Chapter 8. *Physico-Chemical Wastewater Treatment and Resource Recovery*. http://dx.doi.org/10.5772/65499

Cain, J., (2001). Planning improvements in Natural Resources Management Guidelines for using Bayesian Networks to support the planning and management of development programmes in the water sector and beyond. Centre for Ecology and Hydrology, Wallingford, Oxon.

Cameron, K., Madramootoo, C., Crolle, A., Kinsley, C., (2003). Pollutant removal from municipal sewage lagoon effluents with a free-surface wetland. *Water Research* 37, 2803-2812.

Cantrell, K.B., Ducey, T., Ro., K.S., Hunt, P.G., (2008). Livestock waste-to-bioenergy generation opportunities. *Bioresource Technology* 99, 7941-7953.

Carling, P.A., Gupta, N., Atkinson, P.M., Huang Qing He., (2016). Criticality in the planform behavior of the Ganges River meanders. *Geology*, 44 (10), 859-862.

Carlson, J.M., Langer, J.S., (1989). Mechanical model of an earthquake fault. *Physical Review A* 40, 6470-6484.

Carpenter, S.R., Brock, W.A., (2004). Spatial complexity, resilience and policy diversity: fishing on lake-rich landscapes. Ecology and Society 9 (1), Article 8 http://www.ecologyandsociety.org/vol9/iss1/art8/ [Accessed 04 2015]

Carpenter, S.R., Brock, W.A., (2008). Adaptive capacity and traps. *Ecology and Society* 13(2), Article 40.

Carpenter, S.R., Walker, B., Anderies, J.M., Abel, N., (2001). From metaphor to measurement: resilience of what to what. *Ecosystem* 4 (8), 765-781.

Carreras, B.A., Newman, D.E., Dobson, I., Lynch, V.E., (2001). Evidence for self-organized criticality in electric power system blackouts. In: Presented at the 34th Hawaii international conference on system sciences, Maui, Hawaii, USA.

Carty, A., Scholz, M., Heal, K., Gouriveau, F., Mustafa, A., (2008). The universal design operation and maintenance guidelines for Farm Constructed Wetlands (FCW) in temperate climates. *Bioresource Technoloy* 99, 6780-6792

Chadwick, D.R., Fish, R.D., Oliver, D.M., Heathwaite, A.L., Hodgson, C.J., Winter, M., (2008). Management of livestock and their manure to reduce the risk of microbial transfers to water – the case for and interdisciplinary approach. *Trends in Food Science and Technology* 19,240-247.

Champagne, P., Rodrígues-Alcocer, J.D., Giácoman-Vallejos, G., (2012). Assessment of the plug flow and dead volume ratios in a sub-surface horizontal-flow packed-bed reactor as a representative model of a sub-surface horizontal constructed wetland. *Ecological Engineering* 40, 18-26.

Chan, K., Saltelli, A.., Tarantola, S. (2000). Winding stairs: A sampling tool to compute sensitivity indices. *Statistics and Computing* (1), 187-196.

Chang, N-B., Mohiuddin, G., Crawford, A. J., Kaixu, B., Jin, K-R., (2015). Diagnosis of the artificial intelligence-based predictions of flow regime in a constructed wetland for stormwater pollution control. *Ecological Informatics* 28, 42-60.

Chang, N-B., Xuan, Z., Marimon, Z.A., (2013). System dynamic modelling with sensitivity analysis for floating treatment wetlands in a stormwater wet pond. *Ecological Modelling* 267, 66-79.

Changling, X., Zhangfeng, S., (2009). Wind energy in China: current scenario and future perspectives. *Renewable and Sustainable Energy Reviews* 8, 1966-1974.

Charles, K.J., Ashbolt, N.J., McGuinness, R., Deere, D.A., (2005). Effluent quality from 200 on-site sewage systems: design values for guidelines. *Water Science and Technology* 51 (10), 163-169.

Chen, S. C., Zhu, Y. L., (2004). Sub-pattern-based principle component analysis. *Pattern Recognition* 37(5), 1081–1083.

Chen. Y., Zhang. Y., Jianming. H., (2008). Multi-dimensional traffic flow time series analysis with Self- Organizing Maps. *Tsinghua Science and Technology* 2 (13), 220-228.

Churilov, L., and Flitman, A. (2006). Towards fair ranking of Olympics achievements: The case of Sydney 2000. *Computers and Operational Research* 33, 2057- 2082.

Clark, C.S., (1987). Potential and actual biological related health risk of wastewater industry employment. *Journal of Water Polution and Control Federation* 59, 999-1008.

Cockburn, G., Tesfamarian, S., (2012). Earthquake disaster risk index for Canadian cites using Bayesian belief networks, *Georisk Assessment and Management of Risk for Engineered Systems and Geohazards* 6 (2). 128-140.

Cole, S. (1998). The Emergence of Treatment Wetlands. *Environmental Science and Technology*, 32(9), 218A-223A.

Comber, A., Fiaher, P., Wadsworht, R., (2005). Comparing statistical and sematic approaches for identifying change from land cover datasets. *Journal of Environmental Management* 77, 47-55.

Commission of the European Communities, (2000). *Directive 2000/60/EC of the European Parliament and of the Council Establishing a Framework for the Community Action in the Field of Water Policy*. European Commission, Brussels.

Contini, S., (1995). A new hybrid method for fault tree analysis. *Reliability Engineering and System Safety* 49 (1), 13-21.

Cooper, P.F., (1990). European Design and operation guidelines for Reed Bed Treatment systems. Report prepared for the European Community/European Water pollution control association Emergent Hydrophyte Treatment system expert Contact Group. Published by WRc as report UI 17, December 1990.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo. J., Raskin, R.G., Sutton, P., Van Den Belt, M., (1997). The value of the world's ecosystem services and natural capital. *Nature* 387, 252-259.

Costello B. A., Osborne W. J., (2005). Best Practices in Exploratory Factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research and Evaluation* 10 (7), 1-9.

Costello, C.J., (1993): Costello CJ, Editor. Proceedings of conference on constructed wetlands for wastewater treatment in Ireland. Kinsale. Maxpro.
Cox, C.S., Wathes, C., (1995). Bioaerosols Handbook. Lewis Publishers, New York, USA, 621.

Crampin, S., (2000). Shear-wave splitting in a critical self-organized crust: The New Geophysics. 70th Annual International SEG meeting, Calgary, Canada.

Cressie, N., Majure, J.J., (1997). Spatio-temporal statistical modelling of livestock waste in streams. *Journal of Agriculture, Biological and Environmental Statistics* 2, 24-27.

Cronk, J.K., (1996). Constructed wetlands to treat wastewaters from dairy and swine operations: a review. *Agriculture Ecosystems and Environment* 58, 97-114.

Cruz-Castillo, J., Ganeshanandam, S., Mackay, B.R., Lawes, G.S., Wolley, D.J., (1994). Application of Canonical Discriminant analysis in horticulture research. *HortScience* 29,1115-1119.

de Souza, L.C.L., Giunta, M.B., (2011). Urban indices as environmental noise indicators. *Computer, Environment and Urban Systems* 35, 421-430.

Dean, R.B., Forsythe, S.L., (1976 a). Estimating the reliability of advanced waste treatment. Part 1. *Water Sewage Works*, 87-89.

Dean, R.B., Forsythe, S.L., (1976 b). Estimating the reliability of advanced waste treatment. Part 1. *Water Sewage Works*, 87-89.

Deas, M., Vaughn, J., (2006). U.S Bureau of Reclamation, Klamark Area office, Water Course Engineering, Inc.

DeKeyser, E.S., Kirby, D.R., Ell, M.J., (2003). An index of plant community integrity: development of the methodology for assessing prairie wetland plant communities. *Ecological Indicators* 3, 119-133.

Department for Environment Food and Rural Affairs (Defra), UK. (2010). Odour Guidance for Local Authorities.
http://archive.defra.gov.uk/environment/quality/local/nuisance/odour/ [Accessed 02 2015]

Department of the Environment, Heritage and Local Government, Ireland (2010 a). *Integrated Constructed Wetlands: Guidance Document for farmyard Soiled Water and Domestic Wastewater Applications*. http://www.environ.ie/en/Publications/Environment/Water/FileDownLoad,24931,en.pdf [Accessed 07 2016]

Department of the Environment, Heritage and Local Government, Ireland (2010 b). *Integrated Constructed Wetlands: Guidance Document for farmyard Soiled Water and Domestic Wastewater Applications*, 50-51. http://www.environ.ie/en/Publications/Environment/Water/FileDownLoad,24931,en.pdf [Accessed 07 2016]

Deutsch, L., Folke, C., Skånberg, K., (2003). The critical natural capital of ecosystem performance as insurance for human well-being. *Ecological Economics* (44), 205-217.

Dewar, R., (2003). Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *Journal of Physics A: Mathematical and General* 36 (3).

Dierberg, F.E., Juston, J.J., DeBusk, T.A., Pietro, K., Gu, B., (2005). Relationship between hydraulic efficiency and phosphorus removal in a submerged aquatic vegetation-dominated treatment wetland. *Ecological Engineering* 25, 9–23.

Djeddou, M., Achour, B., Martaud, M., (2013). Determination and analysis of daily reliability of municipal wastewater treatment plant. *Courrier du Savoir* , N0 17. 39-46

Dong, X., Reddy, G.B., (2010). Nutrient removal and bacterial communities in swine wastewater lagoon and constructed wetlands. *Journal of Environmental Science Health Part, A Toxic/Hazard Substances Environmental Engineering.* 45(12), 1526-1535.

Douwes, J., Mannetje, A., Heederik, D., (2001). Work-related symptoms in sewage treatment workers. Annals of agricultural and environmental medicine: AAEM 8, 39-45.

Dow. L. K., Sandepp. K., Dow. E. R., (2004). Self-organizing maps for the analysis of NMR spectra. *Biosilico* 2(4), 157-163.

Duncan, C.P., Groffman, P.M., (1994). Comparing microbial parameters in natural and constructed wetlands. *Journal of Environmental Quality*, 23, 298-305.

Eleye-Datubo, A.G., Wall, A., Wang, J., (2008). Marine and offshore safety assessment by incorporative risk modelling in a Fuzzy-Bayesian network of an induced mass assignment paradigm. *Risk Analysis* 28 (1), 95-112.

Ellis, J., Shutes, R.B.E, Revitt, D.M., (2003). Guidance Manual for Constructed England and Wales, Bristol, UK.

Entropy Law, (2001)
http://www.entropylaw.com/entropy2ndlaw.html [Accessed 02 2015]

EPA, (2010). STRIVE Report Series N0.57. Environmental Technology Development of an Alum Sludge-Based Constructed wetland system for improving organic matter and nutrients removal in high-strength wastewater. STRIVE; Environmental Protection Agency Programme 2007-2013, Ireland.

EPA, (2009). Environmental Indicator dashboard.
http://www.epa.ie/irelandsenvironment/environmentalindicatorsdashboard/groundwaterquality/ [Accessed 06 2014]

EPA, Ireland (2008). Integrated Pollution Prevention; BAT Guidance note on Best Available Techniques for the Disposal or Recycling of Animal Carcasses and Animal Wastes (1st Edition). ISBN: 1-84095-279-2

EPA, Ireland (2009). IPPC Licence P0811-01. Abattoir using an integrated constructed wetland (ICW) and on-site Waste Water Treatment Plant (WWTP) to treat waste water from the installation. http://www.epa.ie/licences/lic_eDMS/090151b280335e41.pdf [Accessed 06 2014]

EPA, USA (2000). Generation, management and regulation of biosolids in the US under the EPA part 503 biosolids rule. Environmental Protection Agency, Washington, DC.

EPA.,(2008).ISBN:1-84095-276 http://www.epa.ie/downloads/advice/bat/BATGuidanceNoteforAnimalSlaughter.pdf [Accessed 02 2014].

Etnier, C., Willets, J., Mitchell, C., Fane, S., Johnstone, S., (2005). Decentralized Wastewater treatment system. Reliability Analysis. EcoEng-Newsletter, no. 11. http://www.iees.ch/EcoEng051 [Accessed 03 2014]

European Commission Joint Research Centre, (2005). http://eippcb.jrc.ec.europa.eu/reference/sa.html [Accessed 06 2014]

Fannin, K.F., Vana, S.C., Jakubowsky, W., (1985). Effect of an activated sludge wastewater plant on ambient air densities of aerosols containing bacteria and viruses. *Applied Environmental Microbiology* 49, 1191-1196.

Faulwetter, J.L., Gagnon, V., Sundberg, C., Chararene F., Burr, M.D., Brisson J., Camper, A.K., Stein, O.R., (2009). Microbial processes influencing performance of treatment wetlands: a review. *Ecological Engineering* 35, 987-1004.

Fedorak, M.P., Fernando, L.N., (2005). Changes at an activated sludge treatment plant alter the numbers of airborne aerobic microorganisms. *Water Research* 39, 4597-4608.

Ferreira, L., Borenstein, D., (2012). A Fuzzy-Bayesian model for supplier selection. *Expert Systems with Applications* 39, 7834-7844.

Ferreira, R.A, Duarte, J.G, Vergine, P, Antunes, C.D, Freire, F, Martins-Dias, S., (2014). Phragmites sp. Physiological changes in a constructed wetland treating an effluent contaminated with a diazo dye (DR81). *Environmental Science and Pollution Research International* 21 (16), 9626-43.

Fetter, C.W., (1994). Applied Hydrogeology, 3rd Ed. Upper Saddle River, NJ: Prentice Hall, Inc.

Flood, S., (2012). Climate Change and Potential Economic Impacts in Ireland: The Case for Adaptation. PhD Thesis, National University of Ireland Maynooth.

Fonder, N., Headley, T., (2013). The taxonomy of treatment wetlands; a proposed classification and nomenclature system. *Ecological Engineering* 51, 203-211.

Forrester, J.W., (1961). Industrial Dynamics. The MIT Press, Cambridge, MA, USA.

Forrester, J.W., (1968). Principles of System. Productivity Press, Cambridge, MA, USA.

Fracchia, L., Pietronave, S., Rinaldi, M., Martinotti, G. M., (2006). Site-related airborne biological hazard and seasonal variation in two wastewater treatment plants. *Water Research* 40, 1985-1994.

Gajduk, A., Utkovski, Z., Basnarkov, L., Kocarev, L., (2014). Energy-efficiency in decentralized wireless networks: A Game-theoretic approach inspired by evolutionary biology. 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt). IEEE Conference Publications. DOI: 10.1109/WIOPT.2014.6850358, 636-643.

Gallopín, G.C., (1996). Environmental and sustainability indicators and the concept of situational indicators. A system approach. *Environmental Modelling. and Assessment* (1), 101-117.

Galton, J., Tovey, E., McLaws, L.M., Rawlinson, D.W., (2011). The role of particle size in aerosolised pathogen transmission: A review. *Journal of Infection*, 62, 1-13.

Gao, G., Falconer, R.A., Lin, B., (2013). Modelling importance of sediment effects on fate and transport of enterococci in the Severn Estuary, UK. Marine Pollution Bulletin 67, 45-54.

García, J., Ávila, C., Matamoros, V., Reyes-Contreras, C., Piña, B., Casado, M., Garcia, J., Vivar, J., Aromir, M., Mujeriego, R. (2003). Role of hydraulic retention time and granular medium in microbial removal in tertiary treatment reed beds. *Water Research* 37 (11),2645-2653.

Gazzaz, N.M., Yosoff, M.K., Ramil, M.F., Aris, A.Z., Juahir, H., (2012). Characterization of spatial patterns in river quality using chemometric patterns recognition techniques. *Marine Pollution Bulletin* 64, 688-698.

Gerardi, M.H., Zimmerman, M.C., (2005). Wastewater Pathogens. Wiley-Interscience, John Wiley and Sons, inc., New Jersey, USA, 179.

Goshtasby, A.A., (2012). "Similarity and Dissimilarity Measures", in "Image Registration: Principle, Tools and Methods". Springer Chapter 2.

Greenwood, A.J.B., Schuops, G., Campbell, E.P., Lane, P.N.J., (2014). Bayesian scruntiny of simple rainfall-runoff models used in forest water management. *Journal of Hydrology* (512), 344-365.

Guimaraes, G.P., Jorge, M.F., Alves, D.G., Nascentes, A.L., de Pinto, C.F., da Silva, L.D.B., de Melo, A.C.F., (2016). Removal of nitrogen and phosphorous from cattle farming wastewater using constructed wetland systems. *African Journal of Agricultural Research.* 11(44),4542-4550.

Gupta, A.K., Shrivastava, K., (2006). Uncertainty analysis of conventional water treatment plant design for suspended solids removal. *Journal of Environmental Engineering* 132 (11),1413-1421.

Gurwicz, Y., Lerner, B., (2005). Bayesian network classification using spline-approximated kernel density estimation. *Pattern Recognition Letters*, 26,1761-1771.

Gutiérrez, J.M., Cano, R., Cofiño, A.S., Sordo, C., (2005). Analysis and downscaling multi-model seasonal forecasts in Peru using self-organizing maps. *Tellus A: Dynamic Meteorology and Oceanography* 57A, 435-447.

Haberl, R., Grego, S., Langergraber, G., Kadlec, R.H., Cicalini, A.R., Martins Dias S (2003). Constructed wetlands for the treatment of organic pollutants. *Journal of Soil & Sediments.* 3 (2), 09-124.

Hair, J., Tatham, R.L., Anderson, R.E., Black, W. C. (1998). Multivariate data analysis. NJ: Prentice Hall.

Hamaamin, Y.A., Adhikari, U., Nejadhashemi, A.P., Harrigan, T., Reinhold, D.M., (2014). Modelling Escherichia coli removal in constructed wetlands under pulse loading. *Water Research* 50, 441-454.

Hamby, D.M., (1995). A comparison of sensitivity analysis techniques. *Health Physics* 68 (2), 195-204.

Hanea, A.M., Kurowicka, D., (2008). Mixed non-parametric continuous and discrete Bayesian belief nets. Advances in Mathematical modelling for reliability. ISBN 978-1-58603-865-6. IOS Press.

Hanea, A.M., Kurowicka, D., Cooke, R.M., Ababei, D.A., (2010). Mining and visualising ordinal data with non-parametric continuous BBNs. *Computational Statistics and Data Analysis* (54), 668-687.

Hanea, D., Ale, B., (2009). Risk of human fatality in building fires: A decision tool using Bayesian networks. *Fire Safety Journal* (44), 704-710.

Hansson, L.-A., Broenmark, C., Anders Nilsson, P., AabjoernssonmK., (2005). Conflicting demands on wetland ecosystem services: nutrient retention, biodiversity or both? *Freshwater Biology* 50,705-714.

Harrington, R., (2005). The relationship between plant vigour and ammonium concentrations in surface waters of constructed wetlands used to treat meat industry wastewaters in Ireland. In: Dunne, E.J., Reddy, K.R., Caron, O.T. (Eds), *Nutrient management in Agricultural Watersheds: A Wetlands Solution.*Wageningen Academic Publishers, Wageningen,179-188.

Harrington, R., Dunne, E.J., Carroll, P., Keohane, J. Ryder, C. (2005).' The concept, design and performance of integrated constructed wetlands for the treatment of farmyard dirty water' in Dunn, E.J., Reddy, K.R and Carton, O.T. (Eds): *Nutrient management in Agricultural Watersheds: A Wetlands Solution,* Wageningen Academic Publishers, the Netherlands, 179-188.

Harrington, R., McInnes, R., (2009). Inegrated Constructed Wetlands (ICW) for livestock wastewater management. *Bioresource Techology* 100, 5498-5505.

Harrington, R., Ryder, C., (2002). The use of integrated constructed wetlands in the management of farmyard runoff and waste water. In: *Proceedings of the National Hydrology Seminar on Water Resource Management*: Sustainable Supply and Demand, 19th November, Tullamore Offaly. The Irish National Committees of the IHP and ICID, Ireland, 55-63.

Hart, B.T., Pollino, C.A., (2009). Bayesian modelling for risk-based environmental water allocation. *Waterlines report*. National Water Commission, Canberra, Australia.

Healy M, Cawley AM (2002) Nutrient process capacity of a constructed wetland in Western Ireland. *Environmental Quality* 31(5),1739-47.

Healy, M. Cawley, A. M. (2002). Nutrient processing capacity of a constructed wetland in Western Ireland. *Journal of Environmental Quality* 31 (5), 739-1747.

Henze, M., Gujer, W., Mino, T., Loosdrecht, M. (Eds, 2000). Activated Sludge Models ASM1, ASM2, ASM2d and ASM3. In: *IWA Publishing*, London, UK.

Hergarten, S., Neugebauer, H.J., (1998). Self-organized criticality in a landslide model. *Geophysical Research Letters* 25, 801-804.

Hill, D.T., Payne, V.W.E., Rogers, J.W., Kown, S.R., (1997). Ammonia effects on the biomass production of five constructed wetland plant species. *Bioresource Technology* 62,109–113.

Holling, C.S., (1996). Engineering resilience versus ecological resilience. Engineering within ecological constraints. P.C. Schulze, Editor. National Academy Press, Washington D.C., USA, 31-44.

Homer-Dixon, T., (2011). Shifting the Trajectory of Civilisation. Oxford Leadership Journal, 2 (1), 7.

Horrigan. N., (2005). Neural network based decision support framework for the assessment and management of freshwater stream habitats. Ph.D. thesis, School of Earth and Environmental Sciences. University of Adelaide, Australia.

Hoste, I., (2013). Ornamentals and invasive plants: if you choose one you get the other as well. Mini symposium of aliens and invasive species. National Botanic Garden of Belgium, Bouchout Castle. December 2, 2011. Published online February, 2013.

Irish Water (2019) https://www.water.ie/wastewater/wetlands/dunhill/[Accessed 04 2018]

Isaacs, K., Glen, G., (2012). Estimating Sobol sensitivity indices using correlations. *Environmental modelling and Software*. (37), 157-166.

ITRC, (2003 a). Technical and regulatory guidance document for constructed treatment wetlands. *The Interstate Technology and Regulatory Council Wetland Team,* p105.

ITRC, (2003 b). Technical and regulatory guidance document for constructed treatment wetlands. *The Interstate Technology and Regulatory Council Wetland Team,* p43.

IWA. Constructed wetlands for pollution control: process, performance, design and operation (2000). In: Kadlec RL. Vymazal J, Brix H, Cooper P, Haberl R, editors. IWA scientific and Technical report, vol. 8. IWA publishing.

Jahangir, M. M. R., Richards, K. G., Healy, M. G., Gill, L., Müller, C., Johnston, P., and Fenton, O., (2016). Carbon and nitrogen dynamics and greenhouse gas emissions in constructed wetlands treating wastewater: a review, *Hydrology Earth System Science* 20, 109-123.

Javaid, M.S., Shahid, M., Gabriel, H.F., Nabi. A., (2015) Irrigation and drainage—sustainable strategies and systems. *InTech* p.53

Javaid, M.S., Shahid, M., Gabriel, H.F., Nabi. A., (2015) Irrigation and drainage—sustainable strategies and systems. *InTech* p.63

Jensen, F.V., (1996). Introduction to Bayesian Networks. Springer-Verlag, New York, Inc.

Jensen, F.V., Nielsen, T.D., (2007). Bayesian Networks and Decision Graphs. Springer- Verlag, New York, Inc.

Ji, Z.G., Jin, K.R., (2016). An integrated environmental model for a surface flow constructed wetland: Water quality processes. *Ecological Engineering* 86, 247-261.

Jing, S.-R., Lin, Y-F., Less, D-Y., Wang, T-W., (2001) Nutrient Removal from Polluted River Valley Water by using Constructed Wetland. *Bioresource Technology* 76 (2),131.

John, M. R. (1995) Developments in Wastewater Treatment in the Meat Processing Industry: A review. *Bioresource Technology* 54(3), 204.

Jolliffe, I.T (2002): Principal Component Analysis, 2nd edn. Springer Series in Statistics. Springer, New York.

Jones, D.G., (1998). The Epidemiology of Plant Diseases. Kluwer Academic Publishers, The Netherlands. p 397.

Jorgensen, S.E., Mejer, H., Nielsen, S.N., (1998). Ecosystem as self-organizing critical systems. *Ecological Modelling* 111, 261-268.

Kadlec, R.H., (1994). Detention and mixing in free water wetlands. Ecological Engineering 3(4); 345-380.

Kadlec, R.H., Knight, R. L., Vymazal, J., Brix, H., Cooper, P., Haberl, R., (2000). Constructed wetlands for pollution control. *IWA scientific and technical report 8*. London, UK. IWA Publishing.

Kadlec, R.H., Knight, R.L., (1996). *Treatment Wetlands*. CRC Press Inc., Boca Raton, FL.

Kadlec, R.H., Wallace, S.D., (2009). *Treatment Wetlands* (2nd Ed). CRC Press Taylor Francis Group, Boca Raton, London, NewYork.

Kaiser, H. F., (1960). The application of electronic computers to factor analysis. *Education and Psychological Measurement* 20 (1), 141-151.

Kaiser, H. F., (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36.

Kaltech. A.M., Hjorth. P., Berndtsson. R., (2008) Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling and Software* 23 (7), 835-845.

Kanaya, S., Kinouchi, M., Abe, T., Kudo, Y., Yamada, Y., Nishi, T., Mori, H., Ikemura, T., (2001). Analysis of codon usage diversity of bacterial genes with a self-organizing map (SOM): characterization of horizontally transferred genes with emphasis on the *E. coli* O157 genome. *Gene* 276 (1-2), 89-99.

Karra, S., Katsivela, E., (2007). Microorganisms in bioaerosol emissions from waste water treatment plants during summer at a Mediterranean site. *Water Research* 41, 1355-1365.

Kayakutlu, G., Cinar, D., (2010). Scenario analysis using Bayesian networks: A Case study in energy sector. *Knowledge-Based Systems* 23 (3), 267-276.

Kern, P., (2016). Computational Intelligence Techniques for Control and Optimization of Wastwater Treatment Plants. Ph.D. thesis, Department of Electronic Engineering, National University of Ireland, Maynooth.

Keshtkar, A.R., Salajegheh, A., Sadoddin, A., Allan, M.G., (2013). Application of Bayesian networks for sustainability assessment in catchment modeling and management (Case study: The Hablehrood river catchment). *Ecological Modelling* (268), 48-54.

Kinyua, M.N., Rowse, L.E., Ergas., S.J., (2016). Revew of small-scale tublar anerobic digesters treating livestock waste in developing world. *Renewable and Sustainable Energy Reviews* 58, 896-910.

Knight, R. L., Kadlec, R.H., Ohlendorf, H.M., (1999). The use of treatment wetland for petroleum industry effluents. *Environmental Science and Technology* 33 (7),973-980.

Knight, R.L., (1992). Ancillary benefits and potential problems with the use of wetlands for nonpoint source pollution control. *Ecological Engineering* 1(1-2), 97-113.

Knight, R.L., Payne, V.M.E., Borer, R.E., Clarke, R.A., Pries, J.H., (2000). Constructed wetlands for livestock wastewater management., *Ecological Engineering* 15,41-55.

Kodogiannis, V., Alshejari, A., (2014). An adaptive neuro-fuzzy identification model for the detection of meat spoilage. *Applied Soft Computing* (23), 483-497.

Kohonen, T. (2001). Self- organizing maps (3rd ed.) Springer.

Kohonen, T., Oja, E., Simula, O., Visa, A., Kangas, J., (1996). Engineeering applications of the self-organizing maps. Proc. IEEE 84 (10), 1358-1384.

Korb, K.B., Nicholson, A.E., (2004). Bayesian Artificial Intelligence. Chapman and Hall/ CRC Press, London, p364.

Korzeniewska, K., (20110. Emission of bacteria and fungi in the air from wastewater treatment plants – a review. Front Biosci (SchoolEd). (1:3), 393-407.

Koumanova, B., Lavrova, S., (2010). Influence of recirculation in a lab-scale vertical flow constructed wetland on the treatment efficiency of landfill leachate. *Bioresource Technology* 101, 1756-1761.

Kumwimba, M.N., Zhu, B., Suanon, F., Muyembe, D.K., Dzakpasu, M., (2018). Long-term impact of primary domestic sewage on metal/loid accumulation in drainage ditch sediments, plants and water: Implications for phytoremediation and restoration. *Science of the Total Environment* (581-582), 773-781.

Kuo, R.J., An, Y.L., Wang, H.S., Chung, W.J., (2006). Integration of self-organizing feature maps neural network and genetic *K*-means algorithm for market segmentation. *Expert systems with Applications.* (30), 313-324.

Kuschk, P., Wiessner, A., Seeger, E.M., Kästner, M., Kappelmeyer, U., Paredes, D., Shtemenko, N.I., (2012). The Status of Research on Constructed Wetlands. In: Vitale K. (eds) Environmental and Food Safety and Security for South-East Europe and Ukraine. NATO Science for Peace and Security Series C: Environmental Security. Springer, Dordrecht.

Laitinen, S., Kangas, J., Kotimaa, M., Liesivuori, J., Martikainen, P.J., Nevalainen, A., Sarantile, R., Husman, K., (1994). Workers' exposure to airborne bacteria and endotoxins at industrial wastewater treatment plants. *American Industrial Hygiene Association Journal* 55, 1055-1060.

Lambrakis, N., Antonakos, A., Panagopoulos, G., (2004). The use of multicomponent statistical analysis in hydrogeological environment research. *Water Research* 38 (7),1862-1872.

Langerbergraber, G., Rousseau, D., García, J., Mena, J., (2009a). CWM1- a general model to describe biokinetic processes in subsurface flow wetlands. *Water Science Technology* 59 (9),1687-1697.

Langerbergraber, G., Šimunek, J., (2005). Modelling variably-saturated water flow and multi-component reactive transport in constructed wetlands. *Vadose Zone Journal* (4), 924-938.

Langergraber, G., Giraldi, D., Mena, J., Meyer, D., Peña, M., Toscano, A., Brovelli, A., Korkusuz, E. A., (2009b). Recent developments in numerical modelling of subsurface flow constructed wetlands. *Science of the Total Environment* 407, 3931-3943.

Lee, H.B., Scholz, M., Horn, A., (2006a). Constructed Wetlands: treatment of concentrated storm water runoff (Part A). *Environmental Engineering Science* 23, 320–331.

Lee, H.B., Scholz, M., Horn, A., Fuber, M.A., (2006b). Constructed Wetlands: prediction of performance with case-based reasoning (Part B). *Environmental Engineering Science* 23, 332–340.

Leon, A.S., Yun, T., Duan, C., Yolcu, A., Glennie, C., Pennings, S.C., (2018). *Water* (10), 3, 325. doi:10.3390/w10030325

Levenspiel, O., (1972). Chemical Reaction Engineering. Department of Chemical Engineering, Oregon State University, USA, ISBN 0-471-53016-6.

Li, C., Zhang, B., Zhang, J., Wu, H., Xie, J., Qi, P., (2011). Physiological responses of three plant species exposed to excess ammonia in constructed wetland. *Desalination Water Treatment* 32, 271-276.

Liao, L., (2012). A theory on urban resilience to floods – a basis for alternative planning practices. *Ecology and Society* 17(4) Article 48.

Liao, S-W., Chang, W.L., (2005). Interpretation and discrimination of marshy wetlands by soil factors in the Kuan-Tu Natural Park, Taiwan. *Environmental Monitoring and Assessment* 107, 181-202.

Liao, S-W., Gau, H-S., Lai, W-L., Chen, J-J, Le, C-G., (2008). Identification of pollution of Tapeng Lagoon from neighbouring rivers using multivariate statistical method. *Journal of Environmental Management* 88, 286-292.

Liao, Z., Wang, B., Xia, X., Hannam. M. P., (2012). Environmental emergency decision support system based on Artificial Neural Network. *Safety Science* 47 (6), 843-852.

Lightbody, A.F., Nepf. H.M., Bays, J.S., (2009). Modelling the hydraulic effect of transverse deep zones on the performance of short-circuiting constructed treatment wetlands. *Ecological Engineering* 35, 754-768.

Liu, A., Egodawatta, P., Goonetilleke, A., Mangangka, R.I., (2015). Sectional analysis of stormwater treatment performance of a constructed wetland. *Ecological Engineering* (77), 172-179.

Liu, J., Gao, M., Li, L., (2011). Distribution characterization of microbial aerosols emitted from a wastewater treatment plant using Orbal oxidation ditch process. *Process Biochemistry* 46, 910-915.

Lizama A, K., Fletcher, D.T., Sun, G., (2011). Removal processes for arsenic in constructed wetlands. *Chemosphere* 84, 1032-1043.

Low, B., Ostrom, E., Simon, C., Wilson, J., (2003). Redundancy and diversity: do they influence optimal management? In F. Berkes, J. Colding and C. Folke. Navigating social-ecological systems: building resilience for complexity and change. Cambridge University Press, Cambridge, UK, 83-114.

Lynam, T., Drewry, J., Higham, W., Mitchell, C., (2010). Adaptive modelling for adaptive water quality management in the Great Barrier Reef region, Australia. *Environmental Modelling and Software* (25), 1291-1301.

MacPherson, M., Webb, E.B., Raedeke, A., Mengel, D., Nelson, F., (2018). A review of Bayesian belief network models as decision-support tools for wetland conservation: Are water birds potential umbrella taxa? *Biological Conservation*, 226, 215-223

Magesan, G.N., Williamson, J.C., Yeates, G.W., Lloyd-Jones, A.R., (2000). Wastewater C:N ratio effects on soil hydraulic conductivity and potential mechanisms for recovery. *Bioresource Technology* 71, 21-27.

Mai, T., Smith, C., (2018). Scenario-based planning for tourism development using system dynamic modelling: A case study of Cat Ba island, Vietnam, *Tourism Management* 68, 336-354.

Main, I.G., Naylor, M., (2010). Entropy production and self-organized (sub) criticality in earthquake dynamics. *Philosophical Transactions of the Royal Society A* 368, 131-144.

Mara, D., (1997). Design manual for Waste Stabilization ponds in India. Lagoon Technology International Ltd, Leeds, England, 1-125.

Mara, D., Alabaster, G., Pearson, H., Mills, S. (1992). Waste Stabilization ponds. A Design manual for Eastern Africa. Lagoon Technology International Ltd, Leeds, England.

Mara, D., Pearson, H., (1998). Design Manual for Waste Stabilization Ponds in Mediterranean Countries. Lagoon Technology International Ltd, Leeds, England, 1-112.

Mara, D., Pearson, H., Oragui, J., Arridge, H., Silva, S.A., (2001). Development of a new approach to waste stabilization pond design. Research Monograph No. 5, University of Leeds,1-56.

Marais, G.V.R., (1974). Faecal bacterial kinetics in waste stabilization ponds. *Journal of Environmental Engineering Division*, American Society of Civil Engineers, 100 (1), 119-139.

Marchini, A., Facchinetti, T., Mistri, M., (2009). F-IND: A framework to design fuzzy indices of environmental conditions. *Ecological Indicators* 9, 485-496.

Marcot, B., (2012). Metrics for evaluating performance and uncertainty of Bayesian network models. *Ecological Modelling* 230, 50-62.

Matthew, C., Lawoko, C.R.O., Korte, C.J., Smith, D., (1994). Application of canonical discriminant analysis, and canonical correlations analysis as tools for evaluating differences in pasture botanical composition. *New Zealalnd Journal of Agricultural Research*, 37 (4) ,509-520.

Mauro, J.C., Diehl, B., Marcellin Jr, R.F., Vaughn, D.J., (2018). Workplace accidents and sel-organized criticality. *Physica A*, 506, 284-289.

McGarrigle, M., Lucey, J., Ó Cinnéide, M., (Eds.) (2010). Water Quality in Ireland (2007-2009). Environmental Protection Agency, Johnstown Castle, Wexford.

McInnes, R., Alexander, S., (2013). The benefits of wetland restoration. 10.13140/RG.2.1.3024.1124.

McNeil, V.H., Cox, M.E., Preda, M., (2005). Assessment of chemical water types and their spatial variation using multi-stage cluster analysis, Queensland, Australia. *Journal of Hydrology* 310 (1-4), 81-100.

Meder, B., Mayrhofer, R., Waldmann, M.R., (2009). A rational model of elemental diagnostic inference. In N.N. Taatgen and H. van Rijn (Eds.). Proceeding of the 31st annual conference of Cognitive Science Society, Austin: Cognitive Science Society, 2176-2181.

Melbostad, E., Eduard, W., Skogstad, A., Sandven, P., Lassen, J., Søstrand, P., Heldal, K., (1994). Exposure to bacterial aerosols and work-related symptoms in sewage workers. *American Industrial Hygiene Association Journal* 25, 59-63.

Mendieta-Pino., C.A., Ramos-Martin, A., Perez-Baez., A.O., Brito-Espino., S., (2019). Management of slurry in Gran Canaria Island with full-scale natural treatment systems for waster(NTSW). One-year experience in livestock farms. *Journal of Environmental Management* 232, 666-678.

Metcalf & Eddy (1991). WasteWater Engineering. Treatment, Disposal and Resus. Mcgraw-hill inc., 3rd Edition, 641-662.

Metcalf & Eddy (1996). Ingeriería de aguas residuals. Tratamiento, vertido y reutilización. McGraw-Hill inc., Tercera Edición, 729-752.

Metcalf & Eddy, (2003). Wastewater Engineering: Treatment, and Reuse, Fourth Ed. Metcalf & Eddy, inc., New York, 1819.

Meyer, S.R., Johnson, M.L., Lilieholm, R.J., (2014). Development of a stakeholder-driven spatial modeling framework for strategic landscape planning using Bayesian networks across two urban-rural gradients in Maine, USA. *Ecological Modelling* 291, 42-57.

Miao, Y., Qun, S., Zhiwen, S., Lei, W., (2010). An investigation of Microbial Aerosol in a full-Scale Constructed wetland. Environmental Pollution and Control http://en.cnki.com.cn/Article_en/CJFDTOTAL-HJWR201002002.htm [Accessed 03 2019]

Millennium Ecosystem Assessment, (2005). Ecosystems and Human Health: Water and Wetlands Synthesis. World Resources Institute, Washington, DC, USA.

Milovanov, V.A., (2011). Dynamic polarization random walk model and fish-bone like instability for self-organized critical systems. *New Journal of Physics*, 13.

Mita, L., Rivetti, C., Barate, C., Maria Bayona, J. (2014). Attenuation of emerging organic contaminants in a hybrid constructed wetland system under different hydraulic loading rates and their associated toxicological effects in wastewater. *Science of the total Environment* 470-471, 1272-1280.

Mitsch, W.J and Wise, K. M., (1998). Water Quality, Fate of metals and predictive model validation of a constructed wetland treating acid mine drainage *Water Research* 32 (6),1888-1900.

Mitsch, W.J., Blanca, B., Hernandez, M.E., (2015). Ecosystem services of wetlands, International Journal of Biodiversity Science, *Ecosystem Services and Management,* 11(1),1-4.

Mittal, G.S., (2006). Treatment of wastewater from abattoirs before land application – a review. *Bioresource Technology* 97, 1119-1135.

Mohammadpour, R., Shaharuddin S., Chang, C.K., Zakaria, N.A., Ab Ghani, A., Chan, N.W., (2015). Prediction of water quality index in constructed wetlands using support vector machine. *Environmental Science Pollution Research International* 22(8),6208-6219.

Mohammadpour, R., Shaharuddin, S., Chang, C.K., Zakaria, N.A., Ab Ghani, A., (2014). Spatial pattern analysis for water quality in free-surface constructed wetland. *Water Science Technology* 70 (7), 1161-1167.

Mohr, A.J., (2002). Fate and transport of microorganisms in air. In: Hurst, C.J. (Ed.), Manual of Environmental Microbiology, second ed. ASM Press, Washington, 827-838.

Momen, B., Zehr, J.P., (1998). Watershed classification by discriminant analyses of lake water-chemistry and terrestrial characteristics. *Ecological Applications* 8, 497-507.

Montgomery, D.C., Runger, G.C., (1999). Applied Statistics and Probability for Engineers, second ed. Wiley, New York, p463.

Moomaw, W.R., Chmura, G.L., Davies, G.T., Finlayson, C.M., Middleton, B.A., Natali, S.M., Perry, J.E., Roulet, N., Sutton-Grier, A.E., (2018). Wetlands in a Changing Climate: Science, Policy and Management. *Wetlands* 38,183-205.

Moreno, D., Marco, P., Olmeda, I., (2006). Self-organizing maps could improve the classification of Spanish mutual funds *European Journal of Operational Research* 172 (2), 1039-1054.

Mostafa, M. M., (2010). Clustering the ecological footprint of nations using Kohonen's self-organizing maps. *Expert system with Applications* 7, 2747-2755.

Mouton, M.A., De Baets, B., Van Broekhoven, E., Goethals, M.L.P., (2009). Prevalence-adjusted optimisation of fuzzy models for species distribution. *Ecological Modelling* 220, 1776-1786.

Mouwen, M. J. D., Capita, R., Alonso-Calleja, C., Prieta-Gómez J., Prieto, M., (2006). Artificial neural network based identification of *Campylobacter* species by Fourier transform infrared spectroscopy. *Journal of Microbiological Methods* 67 (1), 131-140.

Murray-Gulde, C.L., Bridges, W.C., Rodgers Jr, J.H., (2008). Evaluating performance of a constructed wetland treatment system designed to decrease bioavailable copper in a waste stream. *Environmental Geosciences*,15 (1) 21-38.

Muselet, D., Tremeau, A., (2008). Rank correlation as illumination invariant descriptor for color object recognition. ICIP 2008. 15th IEEE International Conference on Image Processing. San Diego, CA. ISSN 1522-4880, 157-160.

Mustafa, A., (2009). Nutrient removal with integrated constructed wetlands. Ph.D thesis. School of Engineering. University of Edinburgh, UK.

Mustafa, A., Scholz, M., Harrington, R., Carroll, P., (2009). Long-term performance of a Representative Integrated Constructed Wetland treating Farmyard Runoff. *Ecological Engineering*, 35 (5), p 779-790.

Nameche, T., Vasel, J.L., (1998). Hydrodynamic studies and modelization for aerated lagoons and waste stabilization ponds. *Water Research* 32(10), 3039-3045.

Narayan, S., Beck, M.W., Wilson, P., Thomas, C.J., Guerrero, A., Shepard, C.C., Reguero, B.G., Franco, G., Ingram, J.C., Trespalacios, D., (2017). *Nature*: Scientific Reports 7, Article Number: 9463.

Nash, D., Hannah, M., Robertson, F., Rifkin, P., (2010). A Bayesian network for comparing dissolved nitrogen exports from high rainfall cropping in south-eastern Australia. *Journal of Environmental Quality* 39, 1699-1710.

Niku, S., Schroeder, E.D., (1981). Factors affecting effluent variability from activated sludge processes. *Journal of Water Pollution Control Association* 53 (5), 546-559.

Niku, S., Schroeder, E.D., Haugh, R.S., (1982). Reliability and stability of trickling filter processes. *Journal of Water Pollution Control Association* 54 (2), 192-134

Niku, S., Schroeder, E.D., Samaniego, F.J., (1979 a). Performance of activated sludge process and reliability-based design. *Journal of Water Pollution Control Association* 51 (12), 2841-2857.

Niku, S., Schroeder, E.D., Tchobanoglous, G., Samaniego, F.J., (1979 b). Performance of activated sludge process: reliability, stability and variability. Environmental Protection Agency, EPA grant no. R805097-01 p 1-24.

NIST, (2012). US Department of Commerce. NIST/SEMATECH e-Handbook of Statistical Methods. http://www.itl.nist.gov/div898/handbook/ [Accessed 08 2017]

Nixon, S., Trent, C., Marcuello, C., Lallana, C., (2003). Europe's Water: An indicator-based assessment. European Environmental Agency ISBN: 92-9167-581-4.

O'Luanaigh, N., Gill, L.W., (2012). The attenuation capacity of constructed wetlands to treat domestic wastewater in Ireland. Paper presented at the international symposium on domestic wastewater treatment and disposal systems: Trinity College, Dublin, Ireland, September 10–11. Environmental Protection Agency, Johnstown Castle Estate, Wexford

O'Regan, B., Moles, R., (2001a). A system Dynamics model of Mining Industry Investment Decisions within the Context of Environmental Policy. *Journal of Environmental Planning and Management*, 44(2), 245-262.

O'Regan, B., Moles, R., (2001b). Modelling the effect of information feedback on the spread of the Ebola virus. *Environmental Management Health* (12) 415-427.

Oakley, S.M., (1997). Manual de Diseño, Operación y Mantenimiento para Lagunas de Estabilizatión en Honduras. Fondo Hondureño de Inversión Social (FHIS) & Cooperación de la Agencia de los Estados Unidos parael Desarrollo Internacional (USAID),1-105.

Oakley, S.M., Gold, A.J., Oczkowski, A.J., (2010). Nitrogen control through decentralized wastewater treatment: Process performance and alternative management strategies. *Ecological Engineering* 36 (11), 1520-1531.

Ocal, M. E., Oral, E. L., Erdis, E., Vural, G., (2007). Industry financial ratios-application of factor analysis in Turkish construction industry. *Build Environment* 42(1), 385–392.

Oliveira S.C., von Sperling M., (2007). Reliability analysis of stabilisations pond systems. *Water Science Technology* 55(11), 127-34.

Oliveira S.C., Von Sperling M., (2008). Reliability analysis of wastewater treatment plants. *Water Research* (42), 1182-1194.

Oliveria, P., (2006). Interpolation of signals with missing data using PCA. Acoustics, Speech and Signal Processing, ICASSP Proceedings. IEEE International Conference Vol (3).

Ouyang, Y., (2005). Evaluation of river quality monitoring stations by principal component analysis. *Water Research* 39, 2621-2635.

Overall, A.R., Grace, R.M., Pollino, A.C., Hart, T.B., (2009). A risk-based approach to the improved understanding and management of denitrification in urban stormwater treatment wetlands. 18th World IMACS/MOSIM Congress, Cairns, Australia. http://mssanz.org.au/modsim09 [Accessed 09 2015]

Paczuski, M., Bak, P., (1999). "Self-organization of complex systems." *arXiv preprint cond-mat/9906077*.

Pankaj, T., Tapas, N., Pallave, U., Pravin, M., (2008). Correlating on-line monitoring parameters, pH, DO and OPR with nutrient removal in an intermittent cyclic process bioreactor system. *Bioresource Technology* 99 (16), 7630-7635.

Pant, H.K., Rechcigl, J.E., Adjei, M.B., (2003). Carbon sequestration in wetlands: concept and estimation, *Food, Agriculture and Environment* 1, 308-313.

Park, S., Yun, I., Lee, S. (1998): Colour image segmentation based on 3-D clustering: Morphological approach. *Pattern Recognition* 31(8), 1061-1076.

Pascual, L., Pérez-Luz, S., Yáñez, M.A., Santamaría, A., Gibert, K., Salgot, M., Apraiz, D., Catalán, V., (2003). Bioaerosol emission from wastewater treatment plants. *Aerobiologia* 17, 99-108.

Pascual, M., Guichard, F., (2005). Criticality and Disturbance in spatial ecological systems. *Trends in Ecology and Evolution.* 20 (2), 88-95.

Pearl, J., (1986). Fusion, propagation and structuring in belief networks. *Artificial Intelligence* 29 (3), 241-288.

Pearson, H.W., Mara, D.D., Arridge, H.A., (1995). The influence of pond geometry and configuration on facultative and mature waste stabilisation pond performance and efficiency. *Water Science and Technology* 31 (12), 129-139.

Pearson, K., (1905). On the general theory of skew correlation and non-linear regression. In: Mathematical contributions to the theory of evolution, Drapers' Company Research Memoirs, vol. XIV. London: Dulau & Co.

Peng-cheng, L., Guo-hua, C., Li-coa, D., Li, Z., (2012). A fuzzy Bayesian network approach to improve the quantification of organizational influences in HRS frameworks. *Safety Science* (50), 1569-1583.

Persson, J., Somes, N., Wong, T., (1999). Hydraulic efficiency of constructed wetlands and ponds. Water Science and Technology 40 (3), 291-300.

Peruzzi, E., Masciandaro, G., Macci, C., Doni, S., Mora Ravelo, S., Peruzzi, P., Ceccanti, B., (2011). Heavy metal fractionation and organic matter stabilization in sewage sludge treatment wetlands. *Ecological Engineering* 37, 771-778.

Perz, S.G., Muñoz-Carpena, R., Kiker, G., Holt, D.R., (2013). Evaluating ecological resilience with global sensitivity and uncertainty analysis. *Ecological Modelling* 263, 174-186.

Pescod, M.B., Mara, D.D., (1988). Design, Operation and Maintenance of Wastewater Stabilization Ponds, in Treatment and Use of Sewage Effluent for Irrigation, Butterworths London, 93-115.

Petersen, S.O., Sommer, S.G., Beline, F., Burton, C., Dach, J., Dourmad., J.Y., Leip, A., Misselbrook., T.H., Nicholson, F., Poulsen, H.D., Provolo, G., Soresen, P., Vinneras, B., Weiske, A., Bernal, M.-P., Bohm, R., Juhasz, C., Mihelic, R., (2007). Recycling of livestock manure in a whole-farm perspective. *Livestock Science* 112,180-191.

Phuntsho, S., Shon, K.H., Vigneswaran, S., Kandasamy, J., (2008). Wastewater stabilization ponds (WSP) for wastewater treatment, edited by S. Vigneswaran, in Encyclopaedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford, UK. http://www.eolss.net [Accessed 10 2016]

Pillai, S D., Lesikar, B., Vega, E., (2003). Transport and survival of bacteria and viral tracers through submerged-flow constructed wetland and sand-filter system. *Bioresource Technology* 89, 49-56.

Pillia, S.D., Lesikar, B., Vega, E., (2003). Transport and survival of bacterial and viral tracers through submerged-flow constructed wetland and sand-filter system. *Bioresource Technology* 89, 49-56.

Platzer, C., (2000). "Development of Reed Bed systems – A European Perspective." In: Proceedings of the IWA 7th International Conference of wetland system for water Pollution Control, Lake Buena Vista, Florida, USA.

Poggi-Varaldo, M. H., Gutierrez-Saravia, A., Fernandez-Villagomez, F., Martinez-Pereda, P., Rinderknecht-Seijas, N., (2002). A full-scale system with wetlands for slaughterhouse wastewater Treatment. In: Nehring KW, Brauning SE, editors. Wetlands and remediation II. Columbus, Ohio: Battelle Press; 2002, 213-23.

Pollino, C.A., Woodberry, O., Nicholson, A., Korb, K., Hart, B.T., (2007). Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environmental Modelling and Software* 22, 1140-1152.

Pollino, C.A., Woodberry, O., Nicholson, A., Korb, K., Hart, T.B., (2007). Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environmental Modelling and Software* 22, 1140-1152.

Pople, H.E., Myers, J.D., Miller, R.A., (2005). Dialog Internist: A model of Diagnostic Logic for Internal Medicine; Proc, IJCAI-75, 849-855.

Pourshahriar, H., (2012). Correct vs. accurate prediction: A comparison between prediction power of artificial neural networks and logistic regression in psychological researchers. *Procedia- Social and Behavioural Sciences* 32, 97-103.

Qiang, Z-M., Dong, H-Y., Wang, W-D., Jin, H., (2012). Evaluation of rural wastewater treatment processed in a county of eastern China. *Journal of Environmental Monitoring*, 14, 1906-1913.

Quek, S.T., And, K.K., Ong, S.L. (1995). Reliability of domestic-waste biofilm reactors. *Journal of Environmental Engineering* 121 (11), 785-790.

Ragusa, R.R., de Zoysa, D.S., Regasamy, P., (1994). The effect of microorganisms, salinity and turbidity on hydraulic conductivity of irrigation channel soil. *Irrigation Science* 15 (4),159-166.

Randon, N.J., Lawry, J., Horsburgh, K., Cluckie, I.D., (2008). Fuzzy Bayesian Modeling of Sea-Level Along the East Coast of Britain. *IEEE Transactions of Fuzzy Systems* 16(3), 725-738.

Reddy, K.R., Kadlec, R.H., Flaig, E., Gale. P.M., (1999). Phosphorous retention in streams and wetlands: a review. Critical Reviews in Environmental Science and technology 29, p 83-146.

Reed, S. C., Crites, R. W., Middlebrooks, J.E. (1995) Natural Systems for Waste Management and Treatment, 2nd ed., McGraw-Hill: New York.

Richards, R., Sanó, M., Roiko, A., Carter, W.R., Bussey, M., Matthews, J., Smith, F.T., (2013). Bayesian belief modelling of climate change impacts for informing regional adaption options. *Environmental Modelling and Software*, 44, 113-121.

Rivera, F., Warren, A., Curds, R.C., Robles, E., Gutierrez, A., Gallegos, E., Calderón, A., (1997). The application of the root zone method for the treatment and reuse of high-strength abattoir waste in Mexico. *Water Science Technology* (5), 271-278.

Rodríguez, E., Peche, R., (2012). Development of environmental quality indexes based on fuzzy logic. A case study. *Ecological Indicators* 23, 555-565.

Rousseau, D.P.L., Vanrolleghem, P.A., De Pauw, N., (2004). Model-based design of horizontal subsurface flow constructed treatment wetlands: a review. *Water Research* 38 (6),1484-1493.

Rylander R., Anderson, K., Belin, L., Berglund, G., Bergstrom, R., Manson, L.A., Lundholm, M., Mattsby, I., (1976). Sewage worker's syndrome. *Lancet* 8, 478.

Rylander, R., (1999). Health effects among workers in sewage treatment plants. *Occupational and Environmental Medicine* 56,354-357.

Sadeghi, R., Zarkami, R., Van Damme, P., (2014). Modelling habitat preference of an alien aquatic fern, *Azolla filiculoides* (Lam.), in Anzali wetland (Iran) using data-driven methods. *Ecological Modelling*, 284,1-9.

Salmo, S., Münch, C., Röske, I., (2009). Evaluation of the metabolic diversity of microbial communities in four different filter layers of a constructed wetland with vertical flow by Biolog™ analysis. *Water Research* 43, 4569-4578.

Saltelli, A., Chan, K., Scott. E.M., editors (2000a). Sensitivity analysis. Wiley Series in Probability and Statistics. Wiley, 2000.

Saltelli, A., Tarantola, S., Campolongo, F., (2000b). Sensitivity analysis as an ingredient of Modelling. *Statistical Science*. 15(4), 377-395.

Saltelli, A., Tarantola, S., Chan, K., (1999). A quantitative model independent method for global sensitivity and analysis of model output. *Techometrics* 41 (1), 39-56.

Samper, J., Yang, C., Molinero, J., Bonilla, M., (2007). Modelling geochemical and microbial consumption of dissolved oxygen after backfilling a high level radioactive waste repository. *Journal of Contaminant Hydrology* 93,130-148.

Sánchez-Monedero, A.M., Aguilar, I.M., Fenoll, R., Roig, A., (2008). Effect of the aeration system on the levels of airborne microorganisms generated at wastewater treatment plants. *Water Research* 42, 3739-3744.

Sandec/Eawag (2009 a). Free water surface wetland.png. www.sandec.ch [Accessed 02 2014]

Sandec/Eawag(2009b). Horizontal subsurface flow constructed wetland.png. www.sandec.ch. [Accessed 02 2014]

Sandec/Eawag(2009c). Vertical subsurface flow constructed wetland.png. www.sandec.ch. [Accessed 02 2014]

Sani, A., Scholz, M., Babatunde, A., Wang, Y., (2013). Impact if clogging on the outflow water quality of vertical-flow constructed wetlands treating urban wastewater. *Water Air Soil Pollution* 224 (3), 1488-1506.

Sather, J.H., Smith, R.D., (1984). An Overview of Major Wetland Functions. US Fish and Wildlife Services. FWS/OBS-84/18.

Seo, D-C., Park, J-H., Kim, S-H., Delaune, R.D., Kang, B-H., Kang, S-W., Cho, J-S., Ok, Y.S., (2016). Enhancement of phosphorous removal with near-neutral pH utilizing stell and ferronickel slags for application of constructed wetlands. *Ecological Engineering*, 95, 612-621.

Scheffer, M., Carpenter, S.R., Foley, J., Folke, C., Walker. B., (2001). Catastrophic shifts in ecosystems. *Nature* 413, 591-696.

Scheffer, M., Hosper, S.H., Meijer, M-L., Moss, B., Jeppesen, E., (1993). Alternative equilibra in shallow lakes. *Trends in Ecology and Revolution*. 8(8), 275-279.

Scholtz, M., (2008). Classification of flood retention basins: The Kaiserstuhl case study. *Environmental and Engineering Geoscience* 24 (2), 61-80.

Scholtz, M., Harrington, R., Mustafa, A., Zhang, L., (2008). Assessment of the nutrient removal performance in integrated constructed wetlands with the self-organizing map. *Water Research* 42, 3519-3527.

Scholtz, M., Lee, B-H., (2005) Constructed Wetlands: a review. *International Journal of Environmental Studies* 62 (4), 421-47.

Scholtz, M., Lee, B-H., (2006). Application of the self-organizing map (SOM) to assess the heavy metal removal performance in experimental constructed wetlands. *Water Research* 40(10), 3367-3374.

Scholtz, M., Mustafa, A., Harrington, R., Carroll, P., (2009). Long-term performance of a representative integrated constructed wetland treating farmyard runoff. *Ecological Engineering* (35), 5, 779-790.

Scholtz, M., Sadowski, A. J., Harrington, R., Carroll, P., (2007b). Integrated Constructed Assessment and Design for Phosphate Removal. *Biosystems Engineering*, 97 (3), 415-423.

Scholz, M., Harrington, R., Carroll, P., Mustafa, A., (2007a). The Integrated Constructed Wetlands (ICW) concept. *Wetlands* 27, 337-354.

Scholz, M., Lee, H.B., (2006). Application of the self-organizing map (SOM) to assess the heavy removal performance in experimental constructed wetland. *Water Research* 40, 3367-3374.

Scholz, M., Sadowski, A., Harrington, R., Carroll, P., (2007). Integrated constructed wetlands assessment and design for phosphate removal. *Biosystems Engineering* 97, 415-423.

Scholz, M., Zhang, L., Mustafa, A., Harrington, R., (2008). Assessment of the nutrient removal performance in integrated constructed wetlands with the self-organizing map. *Water Research* 42, 3519-3527.

Scholz, M., Zhang, L., Mustafa, A., Harrington, R., (2009). Application of the self-organizing map as a prediction tool for an integrated constructed wetland agroecosystem treating agricultural runoff. *Bioresourse Technology* 100 (2), 559-565.

Scott, C.A., Faruqui, N.I., Raschid-Sally, L., (2004). Wastewater sue in irrigated agriculture: management challenges in developing countries. In: Scott, CA., Faraqui, N.I., Raschid-Sally, L., (Eds), Wastewater use in Irrigated Agriculture: Confronting the Livelihood and Environmental Realities. CAB International/IWMI/ IDRC,1-10.

Shannan, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal,* 27, 379-423, 623-656.

Sharrock, P., Fiallo, M., Nzihou, A., Chkir, M., (2009). Hazardous animal waste carcasses transformation into slow release fertilizers. *Journal of Hazardous Material* 167, 119-123.

Shenton, W., Grace, M., Hart, B., (2007). Environmental Flow Bayesian Network Decision – Making Framework – Progress Report 1.LWA Project UMO49. Water Studied Centre, Monash University, Australia.

Shepherd, H. L., Grismer, M. E., Tchnobanoglous, G. (2001) Treatment of High-Strength Winery Wastewater using a Subsurface Constructed Wetland. *Water Environmental Research* 73, 394.

Shi, K., Liu, C.Q., (2009). Self-organized criticality of air pollution. *Atmospheric Environment,* 43, 3301-3304.

Silva, J.P., Phillips, L., Jones, W., Eldridge, J., O'Hara, E., (2007). LIFE and Europe's Wetlands: Restoring a Vital Ecosystem. Office for Official Publications of the European Communities, Luxembourg.

Silven, O., Niskanen, M., and Kauppinen, H. (2003). Wood inspection with non-supervised clustering. *Machine vision and Applications* 13, 275-285.

Silvert, W., (2000). Fuzzy indices of environmental conditions. *Ecological Modelling* 130, 111-119.

Skarda, S, M., Moore, J.A., Niswander S.F. and Gamroth, M.J., (1994) Preliminary results of wetlands for treatment of dairy farm wastewater. Proceeding: Constructed Wetlands for Animal Waste Management. p34-42. Lafayette, IN.

Sleytr, K., Tietz, A., Langergraber, G., Haberl, R., (2007). Investigation of bacterial removal during the filtration process in constructed wetlands. *Science of Total Environment* 380, 173-180.

Slivert, W., (1997). Ecological impact classification with fuzzy sets. *Ecological Modelling* 96, 1-10.

Snedecor, G.W., Cochran, W.G., (1989). Statistical Methods, eight ed. Iowa State University Press, Ames, IA, 503.

Sobol, I. (1993). Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling and Computational Experiment* 1, 407-414.

Song, K., Park, Y-S., Zheng, F., Kang, H., (2013). The application of Artificial Neural Network (ANN) model to the simulation of denitrification rates in mesocosm-scale wetlands. *Ecological Informatics* 16,10-16.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., Xu, C., (2015). Global sensitivity analysis in hydrological modelling: Review of concepts, methods, theoretical framework and applications. *Journal of Hydrology* 523, 739-757.

Sözen, A., (2009). Future projection of the energy dependency of turkey using artificial neural network. *Energy Policy* 37(11), 4827-4833.

Spence, L.P., Jordan, J.S., (2013). Effects of nitrogen on freshwater wetland ecosystem services – A Bayesian network analysis. *Journal of Environmental Management* 124, 91-99.

Spence, P.L., Jordan, S.J., (2013). Effects of nitrogen inputs on freshwater wetland ecosystems services – A Bayesian network analysis. *Journal of Environmental Managemen*t 124, 91-99.

Srinvas, V.V., Shivam. T., Ramachandra Rao, A, Govindaraju., R S., (2008). Regional flood frequency analysis by combining self-organizing feature map and fuzzy clustering. *Journal of Hydrology* 348 (1-2), 148-166.

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons, *Journal of the American Statistical Association* 69, 730-737.

Stern, N., (2006). Review on the Economics of Climate Change. H.M. Treasury, UK.
https://webarchive.nationalarchives.gov.uk/20100407172811/http://www.hm-treasury.gov.uk/stern_review_report.htm [Accessed at August 2016]

Stohl, C., (1995). Organisational communication- connectedness in action. London: Sage.

Sun, L., Shenov, P.S., (2007). Using Bayesian networks for bankruptcy prediction: Some methodological issues. *European Journal of Operational Research* 180, 738-753.

Thein, S.J., Graveel, J.G., (1997). Laboratory manual for Soil Science. Agricultural and Environmental Principles. 7th Ed.

Talukdar, S., Pal. S., (2018). Wetland habitat vulnerability of lower Punarbhaba river basin of the uplifted Barind region of Indo-Bangladesh. *Geocarto International*. 10.1080/10106049.2018.1533594

Thorn, J., Beijer, L., Jonsson, T., Rylander, R., (2002). Measurement strategies for the determination of airborne bacterial endotoxin in sewage treatment plants. *Annals of Occupational Hygiene* 46, 549-554.

Ticehurst, J.L., Curtis, A., Merritt, W.S., (2011). Using Bayesian networks to complement conventional analyses to explore landholder management of native vegetation. *Environmental Modelling and Software* 26, 52-65.

Toet, S., Van Logtestijn, R.S.P., Kampf, R., Schreijer, M., Verhoeven, J.T.A., (2005). The effect of hydraulic retention time on the removal of pollutants from sewage treatment plant effluent in a surface-flow wetland system. *Wetlands*, 25 (2), 375-391.

Tompkins, E.L., Adger, W.N., (2004). Does adaptive management of natural resources enhance resilience to climate change? *Ecology and Society* 9(2): Article 10.

Travaini-Lima, F., Sipaúba-Tavares, L. H., (2012). Efficiency of a constructed wetland for wastewaters treatment. *Acta Limnolgica Brasiliensia*. 24(3), 255-265.

Tsihrintzis, A.V., Sylaios, K.G., Kotti, P. I., (2013). Fuzzy logic models for BOD prediction in free-water surface constructed wetlands. *Ecological Engineering* (51), 66-74.

Tsihrintzis, V.A., Akratos, C.S., Papaspyros, J,N.E., (2008). An artificial neural network model and design equations for BOD and COD removal prediction in horizontal subsurface flow constructed wetlands. *Chemical Engineering Journal* 143 (1-3), 96-110.

United States Environmental Protection Agency (1996). Protecting Natural Wetlands: A Guide to Stormwater Best Management Practices. Office of Water, Washington, DC. EPA -843-B-96-001.

United States Environmental Protection Agency (2000). Constructed Wetlands: Treatment of Municipal Wastewaters. Office of Research and Development, Cincinnati, Ohio. EPA/625/R-99/010.

USEPA., (2008). Handbook for Developing Watershed Plans to restore and Protect Our waters. US Environmental Protection Agency. Office of Water. Washington, DC. EPA 841-B-08-002.

Uyanik, S., Naz, M., Yesilnacar, I.M., Sahinkaya, E., (2009). Side-by-side comparison of horizontal subsurface flow and free water surface flow constructed wetlands and artificial neural network (ANN) modelling approach. *Ecological Engineering* 35, 1255-1263.

Valverde, S., Solé, R.V., (2002). Self-organized critical traffic in parallel computer networks. *Physics A 312*, 636-648.

Van Broekhoven, E., Adriaenssens, V., De Baets, B., Verdonschot, P., (2006). Fuzzy rule-based macro-invertebrate suitability models for running waters. *Ecological Modelling* 198, 71-84.

Van der Valk, A.G., (2006). The Biology of Freshwater Wetlands. Oxford university Press, Oxford, UK. p 173.

Van Der Wiele, C.F., Yuan, H., Khorram, S., (2009). An Automated Artificial Neural Network System for Land use/Land Cover Classification from Landsat TM Imagery. *Remote Sensing* 1, 243-265.

Van Deun, R., Van Dyck, M., (1994) "Katholieke Hogeschool Kempen" Department Agro and Beiotechnology, Geel, Beligium. http://www.constructedwetlands.net/ [Accessed 04 2016]

Vantarakis, A., Paparrodopoulos, S., Kokkinos, P., Vantarakis, G., Fragou, K., Detorakis, I., (2016) Impact on the Quality of Life When Living Close to a

Municipal Wastewater Treatment Plant. *Journal of Environmental and Public Health*. Research Article. ID 8467023

Verbeke, W., Dejaeger, K., Martens, D., Hur, J., Baesens, B., (2012). New insights into churn prediction in the telecommunications sector: A profit driven data mining approach. *European Journal of Operational Research* 218, 211-229.

Verbraken, T., Goethals, F., Verbeke, W., Baesens, B., (2014). Predicting online channel acceptance with social network data. *Decision Support Systems* 63, 104-114.

Verhoeven, J.A.T., Tromp, K., Lima, A. T., Barendregt, A., (2012). Retention of heavy metals and poly-aromatic hydrocarbons from road water in a constructed wetland and the effect of de-icing. *Journal of Hazardous Materials* (203-204) 290-298.

Vesanto, J. (1999). SOM-based data visualisation methods. *Intelligent Data Analysis* 3, 111-126.

Vidal, G., Moragas, L., Sáez, K., García, J., Vera, I., (2011). Performance evaluation of eight years' experience of constructed wetland systems in Catalonia as alternative treatment for small communities. *Ecological Engineering* 37(2), 364-371.

Vose, D., (2000). Risk Analysis, A Quantitative Guide., 2nd Ed. John Wiley & Sons, p 148.

Vymazal, J., (2002). The use of sub-surface constructed wetlands for wastewater treatment in the Czech Republic: 10 years experience. *Ecological Engineering* 18(5), 633-46.

Vymazal, J., (2005). Horizontal sub-surface flow and hybrid constructed wetlands system for wastewater treatment. *Ecological Engineering* 25, 478-490.

Vymazal, J., (2005). Removal of enteric bacteria in Constructed Treatment Wetlands with Emergent Macrophytes: A Review. *Journal of Environmental Science and Health*, 40,1355-1367.

Vymazal, J., Brix, H., Cooper, P.F., Haberl, R., Perfler, R. Laber, J., (1998). Removal mechanisms and types of constructed wetlands, In: Vymazal, J., Brix, H., Cooper, P. F., Green, M. B., Haberl, R., (Eds) (1998). Constructed wetlands for wastewater treatment in Europe. Backuys Publishers, Leiden, p 366.

Wahl, M., Brown, L., Soboyejo, A., Dong, B., (2012). Quantifying the hydraulic performance of treatment wetlands using reliability functions. *Ecological Engineering* 47, 120-125.

Walker, B., Westley, F., (2011). Perspectives on resilience to disasters across sectors and cultures. *Ecology and Society* 16(2): Article 4

Wallace, S. D., Knight, R. L., (2006). Small scale constructed wetlands treatment systems – feasibility, design criteria and O&M requirements. Report 01-CTS-5, Water Environment Research Foundation (WERF), Alexandria.

Water Framework Directive (2015).
http://www.epa.ie/water/watmg/wfd/ [Accessed at August 2016]

Wetlands. R&D Tech. Report P2–159/TR2. Environment Agency for
WETPOL, (2013). 5th International Symposium on Wetland Pollutant Dynamics and Control.
http://web.imt-atlantique.fr/x-ener/wetpol2013/ [Accessed at August 2016]

WHO (1989). *Health Guidelines for the use of Wastewater in Agriculture and Aquaculture.* Technical Report Series No. 778. World Health Organisation, Geneva.

Wohlfart, C., Winkler, K., Wendleder, A., Roth, A., (2018). TerraSAR-X and Wetlands: A review. *Remote Sensing*, (10) 916.

Woodward, R., Newman, D.E., Sánchez, R., Carreras, B.A., (2007). Persistent dynamic correlations in self-organized critical system away from their critical point. *Physica A* 373, 215-230.

World Bank (2009). Global Study of Livestock Markets, Slaughterhouses, and Related Waste Management Systems. Final report, by Nippon Koei in association with ProAnd Associates Australia, Contract # 7142400.Available at: siteresources.worldbank.org/INTUSWM/Resources/4636171205264154387/Cointreau_SWMStudyFinalReportFeb2009.pdf [Accessed at 09 2016]

Xiong, J., Guo, G., Mahood, Q., Yue, M., (2011). Nitrogen removal from secondary effluent by using integrated constructed wetland system. *Ecological Engineering*, 37(4), 659-662.

Xu, D., Zhang, S. Y., Zhou, Q, H., He, F., Cheng, S.P., Lian, W., Du, C., Wu, Z.B., (2010). Vertical-flow constructed wetlands applied in a recirculating aquaculture system for channel catfish culture. *Polish Journal of Environmental Studies* 19(5), 1063-1070.

Yánez, F., (1992). Lagunas de Estabilizatión: Teoría, Diseño. Evaluación y Mantenimiento. Instituto Ecuatoriano de Obras Sanitarias, Ministerio de Salud Pública, Quito, Ecuador.

Yan, Y., Xu, J., (2014). Improving winter performance of constructed wetlands for wastewater treatment in Northern China: A Review. Wetlands 34(2), 243 – 253.

Yang, F.Z., Zhao, W.Y., (2009). Integrative fuzzy hierarchical model for river health assessment: A Case study of Yong River in Ningbo City, China. *Communications in nonlinear Science and numerical Simulations* 14,1729-1736.

Yang, X.S., (2002). Characterization of multispecies Living Ecosystems with Cellular Automata. Artificial Life VIII, Standish, Abbass, Bedau (eds) (MIT Press),138-141.

Yazdi, S. K., Scholtz M., (2010). Assessing Storm Water detention systems treating road runoff with an artificial neural network predicting fecal indicator organisms. *Water, Air, & Soil Pollution* 204 (1-4),35-47.

Yurth, D.G., Ayres, D., (2005). Y-Bias and Angularity: The Dynamics of Self Organizing Criticality from the Zero Point to Infinity. p15. Utah, USA. http://NovaInstituteofTechnology.com [Accessed 02 2015]

Zeng, M., Soric, A., Ferrasse, J.H., Roche, M., (2013 a). Interpreting hydrodynamic behaviour by the model of stirred tanks in series with exchange zones: preliminary study in lab-scale tricking filters. *Environmental Technology* 34 (18), 2571-2578.

Zeng, M., Soric, A., Roche, M., (2013 b). Calibrating of hydrodynamic behaviour and biokinetics for TOC removal modelling in biofilm reactors under different hydraulic conditions. *Bioresearch Technology* 144, 202-209.

Zhao, X., Wang, Y., Wang, J., Song, X., Gong, J., (2016). The inhibition and adaptability of four wetland plant species to high concentration of ammonia wastewater and nitrogen removal efficiency in conctructed wetlands. *Bioresource Technology*, 202, 198-205.

Zhao, Y., Yang, Y., Liu, R., Morgan, D., (2018). Global development of various emerged sunstrates utilized on contructed wetlands. *Bioresource Technology* 262, 441-452.

Zhan, A., Healy, G.M., Li, J., Norton, D., (2010). STRIVE report 45- Enhanced Nitrogen Removal for Slaughterhouse Wastewater using Novel Technologies.

Zhang, L., Yang, T., Qiu, W., Ma, Y., Chadli, M., (2014). Fuzzy model-based predictive control of dissolved oxygen in activated sludge processes. *Neurocomputing* 136, 88-95.

Zheng, Z., Itoh, H., Kawaguchi, K., Tamaki, H., Maruyama, Y., (2015). Application of Bayesian nonparametric models to the uncertainty and sensitivity of source terms in a BWR severe accident. *Reliability Engineering and System Safety* 138, 253-262.

Zhou, J., Lopresti, D., (1997): Extracting test from WWW images. *International Conference on Document Analysis and Recognition*, 248-252.

Appendices

Appendix A

Key data variables within the DAF:

Months	pH (pH units)	Redox Pot (mV)	Dissolved oxygen (mg/L)	BOD (mg/L)	Turbidity (NTU)	Total Coliforms A	*E.coli* A	Enterococci A	Nitrate (mg/L)	Phosphate (mg/L)	Sulphate (mg/L)	Ammonia (mg/L)
Feb	7.6	-210	0.49	nd	1252	7.83	7.48	6.76	nd	nd	nd	nd
Mar	7.6	-175	0.91	3150	1000	7.88	7.78	6.88	213.4	71.5	24.5	nd
Apr	7.15	-310	0.47	3872	982	8.92	7.89	6.91	241.2	65.2	18.2	nd
May	7	-230	0.8	4350	1567	9.15	8.98	9.11	16.72	104.66	12.6	325.9
Jun	6.9	-280	0.35	3952	893	7.88	7.78	6.88	nd	nd	Nd	nd
Jul	7.4	-265	0.24	4120	599	10.59	10.42	6.24	0	99.93	21.44	nd
Aug	6.8	-294	0.17	4203	789	10.65	10.05	9.47	31.06	68.56	18.75	nd
Sept	7.3	-255	0.29	3130	657	8.37	7.45	7.33	nd	nd	nd	nd
Nov	6.9	-233	0.19	3850	912	10.16	8.97	9.13	nd	nd	nd	nd
Dec	6.72	-319	0.14	4304	819	9.51	8.82	8.19	35.66	81.25	31.8	312.5
Feb08	7.15	-222	0.34	3725	905	9.06	8.41	7.91	241.3	61.5	17.44	387.2
Mean	7.14	-253.91	0.40	3865.60	943.18	9.09	8.55	7.71	111.33	78.94	20.68	341.87
SD	0.31	44.70	0.25	432.83	270.88	1.05	1.01	1.12	113.79	17.14	6.12	39.83
Max	7.60	-175.00	0.91	4350.00	1567.00	10.65	10.42	9.47	241.30	104.66	31.80	387.20
Min	6.72	-319.00	0.14	3130.00	599.00	7.83	7.45	6.24	0.00	61.50	12.60	312.50

A = $\log_{10}$ (CFU/100ml)
nd= not detected

Key data variables within Pond 1:

Months	pH (pH units)	Redox Pot (mV)	Dissolved oxygen (mg/L)	BOD (mg/L)	Turbidity (NTU)	Total Coliforms A	*E.coli* A	Enterococci A	Nitrate (mg/L)	Phosphate (mg/L)	Sulphate (mg/L)	Ammonia (mg/L)
Feb	7.17	-251	0.49	1280	180.7	7.76	7.48	6.76	nd	nd	nd	nd
Mar	6.4	-260	0.75	1410	1000	6.58	6.24	6.75	144.7	62.54	52.5	nd
Apr	6.84	-287	1.02	1562	215	7.52	6.99	6.32	192.11	38.7	9.5	nd
May	6.5	-230	0.48	1140	1219	8.33	5.93	10.56	3.29	26.06	239.79	436.9
Jun	7.2	-163	0.9	823	149	6.58	6.24	5.72	192.11	62.54	Nd	nd
Jul	7.4	-111	0.6	775	123	8.01	5.41	6.24	2.07	84.64	6.87	nd
Aug	7.2	-89	0.7	1789	167	8.79	7.81	6.29	31.06	86.64	35.52	nd
Sept	6.4	-160	1.2	760	181	6.96	6.63	7.33	nd	nd	nd	nd
Nov	7.3	-120	0.91	960	213	7.93	6.85	6.35	nd	nd	nd	nd
Dec	7.31	-312	0.64	2101	256	9.49	7.34	6.06	71.2	105.6	48.9	354.2
Feb08	7.22	-262	0.72	1755	156	8.06	7.71	6.32	214.9	78.52	4.58	477.21
Mean	6.99	-204.09	0.76	1305.00	350.88	7.82	6.79	6.79	106.43	68.16	56.81	422.77
SD	0.39	77.77	0.22	458.76	379.97	0.89	0.77	1.32	89.73	26.26	83.14	62.71
Max	7.40	-89.00	1.20	2101.00	1219.00	9.49	7.81	10.56	214.90	105.60	239.79	477.21
Min	6.40	-312.00	0.48	760.00	123.00	6.58	5.41	5.72	2.07	26.06	4.58	354.20

A = $\log_{10}$ (CFU/100ml)
nd= not detected

Key data variables within Pond 6:

Months	pH (pH units)	Redox Pot (mV)	Dissolved oxygen mg/L	BOD mg/L	Turbidity NTU	Total Coliforms A	E.coli A	Enterococci A	Nitrate (mg/L)	Phosphate (mg/L)	Sulphate (mg/L)	Ammonia (mg/L)
Feb	7.38	34	3.11	27	88	6.12	5.59	4.37	nd	nd	nd	nd
Mar	6.9	57	0.85	52.5	43.3	5.48	5.24	4.00	179.44	60.97	8.44	nd
Apr	7.95	77	6.63	18.2	30.2	4.80	4.47	3.87	1.96	58.67	10.42	nd
May	7.1	8	1.4	156	35.76	7.01	6.03	7.64	0	52.82	6.49	106.2
Jun	6.7	46	4.1	29	47.17	5.48	5.24	4.00	179.44	60.9	4.19	nd
Jul	7.2	32	1.8	49.3	17.6	5.87	4.61	4.11	0.13	38.81	5.62	nd
Aug	7.5	38	1.75	34.5	19.3	5.34	4.79	5.04	92.29	40	7.83	nd
Sept	7.3	-51	1.8	76	48.55	5.75	5.35	5.53	nd	nd	nd	nd
Nov	7.9	-31	1.32	86.2	56.12	6.37	6.79	5.34	nd	nd	nd	nd
Dec	7.8	25	1.81	289.5	111.5	5.92	5.86	6.23	127.2	51.5	9.8	64.2
Feb08	7.55	45	1.45	202	76	5.79	5.04	6.32	144.2	34.55	1.55	85.36
Mean	7.39	25.45	2.37	92.75	52.14	5.81	5.36	5.13	90.58	49.78	6.79	85.25
SD	0.40	37.51	1.68	87.07	29.21	0.58	0.68	1.22	79.51	10.61	2.97	21.00
Max	7.95	77.00	6.63	289.50	111.50	7.01	6.79	7.64	179.44	60.97	10.42	106.20
Min	6.70	-51.00	0.85	18.20	17.60	4.80	4.47	3.87	0.00	34.55	1.55	64.20

A = $\log_{10}$ (CFU/100ml)
nd= not detected

Key data variables within Pond 9:

Months	pH (pH units)	Redox Pot (mV)	Dissolved oxygen mg/L	BOD mg/L	Turbidity NTU	Total Coliforms A	*E.coli* A	Enterococci A	Nitrate (mg/L)	Phosphate (mg/L)	Sulphate (mg/L)	Ammonia (mg/L)
Feb	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mar	8.1	12	10.5	8	32	3.00	3.00	1.99	0	0	3.11	nd
Apr	9.55	62	6.78	11.1	24.2	3.30	3.40	4.09	0	19.43	5.49	nd
May	9.3	18	13.8	56.7	35.76	6.77	6.32	6.04	2.02	18.71	1.8	4.02
Jun	8.5	15	1.5	43.8	39.4	3.04	3.00	1.99	0	0	1.5	nd
Jul	7.8	30	5.95	43.8	60	5.79	4.48	4.33	0	14.34	0	nd
Aug	8.2	48	3.2	21.7	45.6	4.30	4.71	4.97	86.18	40.32	7.83	nd
Sept	7.2	36	4.1	56.7	45.09	4.70	4.31	4.66	nd	nd	nd	nd
Nov	7.4	32	3.83	65.2	32.56	4.89	4.87	4.36	nd	nd	nd	nd
Dec	8.5	31	3.54	112	68.5	4.79	4.60	4.85	112.5	48.5	9.8	6.54
Feb08	8.1	48	5.45	89.5	34.5	4.70	4.71	4.04	112.3	11.2	0.25	18.62
Mean	8.27	33.20	5.87	50.85	41.76	4.53	4.34	4.13	39.13	19.06	3.72	9.73
SD	0.74	16.00	3.72	33.12	13.57	1.20	1.00	1.27	54.05	17.44	3.62	7.80
Max	9.55	62.00	13.80	112.00	68.5	6.77	6.32	6.04	112.50	48.50	9.80	18.62
Min	7.20	12.00	1.50	8.00	24.20	3.00	3.00	1.99	0.00	0.00	0.00	4.02

A = $\log_{10}$ (CFU/100ml)
nd= not detected

Key data variables within Pond 12:

Months	pH (pH units)	Redox Pot (mV)	Dissolved oxygen (mg/L)	BOD (mg/L)	Turbidity (NTU)	Total Coliforms A	*E.coli* A	Enterococci A	Nitrate (mg/L)	Phosphate (mg/L)	Sulphate (mg/L)	Ammonia (mg/L)
Feb	7.43	40	8.93	3.6	6.71	2.12	1.91	1.04	nd	nd	nd	nd
Mar	7.8	32	8.25	12	11.25	1.94	1.48	1.08	163.35	32.64	2.58	nd
Apr	8.86	44	7	16.5	77	5.33	3.30	5.26	0	13.27	0	nd
May	D	D	D	D	D	D	D	D	D	D	D	D
Jun	D	D	D	D	D	D	D	D	D	D	D	D
Jul	7.5	51	2.15	16.7	38	7.40	3.48	4.62	0	8.13	0	nd
Aug	7.3	43	3.75	13.4	19.3	4.85	3.91	2.97	72.88	30.25	7.69	nd
Sept	D	D	D	D	D	D	D	D	D	D	D	D
Nov	7.2	39	4.7	23	18.39	3.06	3.98	3.52	nd	nd	nd	nd
Dec	7.53	41	3.88	26.5	31.4	3.55	3.89	3.91	64.4	30.25	2.5	2.55
Feb08	7.28	52	7.25	11.2	8.5	3.42	2.81	3.96	74.9	18.45	0	7.88
Mean	7.61	42.75	5.74	15.36	26.32	3.96	3.09	3.29	62.59	22.17	2.13	5.22
SD	0.54	6.50	2.44	7.13	23.21	1.82	0.95	1.54	60.41	10.30	3.00	3.77
Max	8.86	52.00	8.93	26.50	77.00	7.40	3.98	5.26	163.35	32.64	7.69	7.88
Min	7.20	32.00	2.15	3.60	6.71	1.94	1.48	1.04	0.00	8.13	0.00	2.55

A = $\log_{10}$ (CFU/100ml)
nd = not detected
D= Dry Pond

Key data variables within Stream:

Months	pH (pH units)	Redox Pot (mV)	Dissolved oxygen (mg/L)	BOD (mg/L)	Turbidity (NTU)	Total Coliforms A	*E.coli* A	Enterococci A	Nitrate (mg/L)	Phosphate (mg/L)	Sulphate (mg/L)	Ammonia (mg/L)
Feb	7.33	56	7.7	1.6	2.84	2.64	2.45	2.18	nd	nd	nd	nd
Mar	6.6	55	5.4	5.6	1.87	3.36	3.17	1.98	47.65	0	0	nd
Apr	nd	nd	nd	nd	nd	4.06	3.04	2.18	nd	nd	nd	nd
May	7.6	27	8.3	4.2	3.94	4.62	4.65	4.85	45.12	0	22.07	0
Jun	7.3	73	5.6	3.1	1.86	3.36	3.17	1.98	48.9	5.1	1.91	nd
Jul	7.14	46	6.1	1.4	0	4.67	3.60	3.02	50.18	0	20.24	nd
Aug	7.1	34	7.2	2.7	0.1	3.50	3.32	2.99	53.53	0	30.47	nd
Sept	7.4	17	4.75	7.8	3.29	3.65	3.09	2.95	nd	nd	nd	nd
Nov	7.8	21	4.62	8.9	6.78	4.18	3.72	3.08	nd	nd	nd	nd
Dec	7.35	85	7.87	3.56	15.2	3.39	3.26	2.71	35.36	8.56	12.2	0.9
Feb08	7.24	68	7.11	6.5	9.6	3.06	3.00	1.99	57.3	9.88	0.1	2.11
Mean	7.29	48.20	6.47	4.54	4.55	3.68	3.32	2.72	48.29	3.36	12.43	1.00
SD	0.32	23.18	1.34	2.57	4.76	0.63	0.55	0.84	6.95	4.43	12.22	1.06
Max	7.80	85.00	8.30	8.90	15.20	4.67	4.65	4.85	57.30	9.88	30.47	2.11
Min	6.60	17.00	4.62	1.40	0.00	2.64	2.45	1.98	35.36	0.00	0.00	0.00

A = $\log_{10}$ (CFU/100ml)
nd= not detected

Appendix B

Appendix B (i)

Analysis of variance (ANOVA) for Physical-chemical Data [pH, redox potential, BOD, dissolved oxygen, turbidity]

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	51580476	65	793545.8	1.674	***0.0037***	1.375
Columns	34693578	4	11564526	24.392	***0***	2.651
Error	92449906	195	474102.1			
Total	1.79E+08	264				

Appendix B (ii)

Analysis of variance (ANOVA) for Bacterial Concentration Data [total Coliforms, *E.coli* and Enterococci]

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	512.528	60	8.542	13.548	***0***	1.534
Columns	2.658	3	2.658	4.21	***0.044***	4.001
Error	37.828	60	0.63			
Total	553.015	121				

Appendix B (iii)

Analysis of variance for Inorganic Data [nitrate, phosphate, sulphate and ammonia]

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	559212.1	61	9167.411	1.713	***0.0023***	1.371
Columns	727468.8	4	181867.2	33.978	***0***	2.409
Error	1306026	244	5352.564			
Total	2592706	309				

Appendix C

Appendix C displays the $\log_{10}$ (CFU/100ml) indicator bacteria values for the total coliforms, *E.coli*, enterococci and the total bacteria. The bacteria data point for each month was comprised of the average total coliforms for each sample point, i.e. the month of February 2007 total coliforms for the DAF, pond 1, pond 6, pond 9 and pond 12 averaged as a single data point. This data point represents the entire wetland for that month. This process was repeated for *E.coli* and enterococci bacterial loadings. The total bacteria concentration is the sum of the three indicator bacteria; the mean water depth (m) and the mean water temperature (^{0}C). The bottom three-quarters of the table contains the HRT data for the three indicator bacteria using the new Oakley, the traditional Oakley and the CSTR models.

	total coliforms	*E.coli*	enterococci	total bacteria	mean water depth	mean water temp
Feb	6.89	6.62	5.71	7.12	0.87	9.23
Mar	6.24	5.92	4.83	6.53	0.75	6.68
Apr	7.27	6.54	6.5	7.54	0.72	8.23
May	7.04	6.62	7.22	7.42	0.42	11.1
Jun	5.77	5.64	4.65	5.96	0.36	17.93
Jul	8.33	6.94	6.64	8.36	0.45	17.6
Aug	7.6	7.27	6.87	7.86	0.43	17.68
Sept	6.27	5.91	6.02	6.45	0.35	11.7
Nov	7.51	7.46	6.94	7.88	0.67	8.86
Dec	7.48	7.24	7	7.9	0.76	3.8
Feb(08)	7.2	6.85	6.67	7.55	0.87	8.92
Month/days	New DFM (TC)	New DFM (EC)	New DFM (ENT)	New DFM (Aver)		
Feb	39	48	38	42		
Mar	41	37	37	38		
Apr	42	48	51	47		
May	62	62	64	63		
Jun	42	59	61	58		
Jul	147	128	110	128		
Aug	53	50	56	53		
Sept	55	50	40	48		
Nov	29	29	31	30		
Dec	32	27	27	29		
Feb(08)	30	30	24	28		
Month/days	DFM (TC)	DFM (EC)	DFM (ENT)	DFM (Aver)		
Feb	29	52	32	37		
Mar	27	59	35	40		
Apr	31	69	69	56		
May	88	88	88	88		
Jun	90	80	70	80		
Jul	86	66	76	76		
Aug	59	41	106	69		
Sept	44	55	40	46		
Nov	28	28	28	28		
Dec	28	28	28	28		
Feb(08)	27	27	27	27		

Month/days	CSTR (TC)	CSTR (EC)	CSTR (ENT)	CSTR (Aver)	
Feb	66	55	49	57	
Mar	60	62	62	61	
Apr	69	55	49	58	
May	69	64	54	62	
Jun	59	59	35	51	
Jul	94	84	84	87	
Aug	126	117	105	116	
Sept	70	70	70	70	
Nov	28	31	31	30	
Dec	28	34	34	32	
Feb(08)	28	29	29	29	

Where the indicator bacteria: total coliforms, *E.coli* Enterococci and total bacteria are measure = $\log_{10}$ (CFU/100ml). Total bacteria are the sum total of the three indicator bacteria *prior* to the logarithm application. Then after the addition of the three bacteria, the $\log_{10}$ is applied, see example below.

Example:

	Total coliforms	*E.coli*	Enterococci	Total bacteria
(CFU/100ml)	4.90E+06	3.80E+06	3.10E+06	1.18E+07
$\log_{10}$ (CFU/100ml)	6.69	6.58	6.49	7.07

The mean water depth was measured in metres (m). The mean water temperature was measured in (^{0}C). The hydraulic retention time (HRT) was measured in days.

New DFM = new Oakley Dispersed Flow model

DFM= Oakley Dispersed Flow model

CSTR = Mara/Maris continuous stirred reactor models

Appendix D

Month	sampling point	total coliforms	*E.coli*	Enterococci	Fluoride	Chloride	Nitrite	Nitrate	Phosphate	Sulphate
April	P1	9.49	8.26	9.28	X	X	X	X	X	X
	P6	6.1	5.75	5.9	X	X	X	X	X	X
	P9	5.3	4.64	4.87	X	X	X	X	X	X
	P12	4.54	3.2	3.55	X	X	X	X	X	X
May	P1	9.87	8.58	8.91	X	X	X	X	X	X
	P6	6.63	6.08	5.17	X	X	X	X	X	X
	P9	6.81	6.23	3.53	X	X	X	X	X	X
	P12	4.32	3.86	2.11	X	X	X	X	X	X
	Stream	6.2	4.86	3.94	X	X	X	X	X	X
June	P1	9.64	8.88	8.63	X	X	X	X	X	X
	P6	6.3	5.93	6.24	15.09	107.07	0	2.03	42.36	11.68
	P9	6.27	5	3.64	14.77	31.21	1.96	24.5	0	4.78
	P12	5.04	3.51	*6.8*	0.46	0.72	0	0	0	0
	Stream	8.47	6.81	7.27	15.89	10.62	0.33	1.67	0	4.39
July	P1	**11.96**	8.54	8.91	0	0	0	0	0	0
	P6	7.79	5.54	5.53	0.65	15.74	0	0	0	0.95
	P9	7.54	4.51	4.71	0	13.31	0.3	0.44	0	1.11
	P12	*6.73*	4.83	*6.62*	0.29	7.9	0	0	0	0
	Stream	6.96	6.33	7.66	0	18.81	0	0	0	1.34
Sept	P1	**11.85**	9.08	10.49	0	0	0	0	0	0
	P6	7.49	5.28	6.16	1.56	72.2	0	0.74	0	18.38
	P9	7.87	5.04	5.85	1.33	32.2	0	4.15	0	4.61
	P12	6.92	4.96	6.93	0.73	3.98	0	7.12	0	0
	Stream	6.86	5.83	6.8	0.38	32.1	0	0.92	0	24.09
Oct	P1	9.72	8.91	8.57	0	0	0	0	0	0
	P6	6.5	5.81	6.35	19.66	111.5	0.5	6.05	33.18	9.55
	P9	6.07	5.34	3.58	4.2	42.5	0.21	18.2	8.35	2.9
	P12	4.98	3.91	4.54	2.6	12.54	0	4.27	0	3.11
	Stream	8.5	6.64	6.71	1.57	7.52	0.54	1.13	0	5.01
Dec	P1	9.5	9	8.88	0	0	0	0	0	0
	P6	6.69	5.93	7.33	10.18	133.54	1.05	5.23	25.33	31.52
	P9	4.06	3.97	4.91	2.88	51.06	0.46	1.95	3.08	13.71
	P12	3.45	3.05	2.39	1.66	19.58	0	1.02	0.84	9.81
	Stream	4.5	3.82	7.06	2.11	11.18	0	3.25	0	16.02

Month	Wetland Area (m^2)	Mean Plants /m^2	Plants x Area	Porosity (%)
April	1898	0	0	
	1804	6.7	12087	
	2813	4.5	12659	
	2428	5.1	12383	
May	1898	0	0	47.8
	1804	7.1	12808	41.1
	2813	2.4	6751	27.9
	2428	2.3	5584	23.8
	0	0	0	
June	1898	0	0	39.7
	1804	6.8	12267	42.8
	2813	1.8	5063	25.7
	2428	0	0	22.7
	0	0	0	
July	1898	0	0	
	1804	6.5	11726	
	2813	1.1	3094	
	2428	0	0	
	0	0	0	
Sept	1898	0	0	43.2
	1804	6.9	12448	38.9
	2813	3.4	9564	28.8
	2428	2.8	6798	24.2
	0	0	0	
Oct	1898	0	0	42.9
	1804	6.3	11365	40.4
	2813	3.1	8720	27.2
	2428	3.2	7770	23.1
	0	0	0	
Dec	1898	0	0	
	1804	5.2	9381	
	2813	2.2	6189	
	2428	2.1	5099	
	0	0	0	

X = denotes no sampling performed. The bacterial concentrations (total coliforms, *E.coli*, and Enterococci) are recorded as Log_{10} (CFU/g). Inorganic ion concentrations are recorded as (mg/L). The mean plants per metre squared (m^2) was determined by constructing a 1-meter square frame using wooden dowel rods and throwing the frame into the ponds and counting the plants contained within the frame, this was repeated three times and an average plant/ m^2 was deduced. The variable "Plants x Area" was calculated by multiplying the wetland area by the mean plants /m^2.

Appendix E (i) (Indicator bacteria concentrations)

Method 1 Data	**total coliforms**	***E.Coli***	**enterococci**	**total bacteria**	
DAF	9.09	8.55	7.71	9.26	
P1	7.82	6.79	6.79	8.19	
P6	5.81	5.36	5.13	6.16	
P9	4.53	4.34	4.13	5.01	
P12	2.88	2.25	2.4	3.17	
mean	6.03	5.46	5.23	6.36	
std	2.49	2.39	2.11	2.44	
Soft-scaling	**total coliforms**	***E.Coli***	**enterococci**	**total bacteria**	
DAF	1.23	1.29	1.17	1.19	
P1	0.72	0.56	0.74	0.75	
P6	-0.09	-0.04	-0.05	-0.08	
P9	-0.6	-0.47	-0.52	-0.55	
P12	-1.26	-1.34	-1.34	-1.31	
Norm(S)	**total coliforms**	***E.Coli***	**enterococci**	**total bacteria**	
DAF	0.89	0.9	0.88	0.88	
P1	0.76	0.71	0.77	0.77	
P6	0.47	0.48	0.48	0.47	
P9	0.27	0.32	0.3	0.29	
P12	0.1	0.09	0.09	0.1	
Entropy	**total coliforms**	***E.Coli***	**enterococci**	**total bacteria**	**SUM**
DAF	0.15	0.13	0.16	0.16	**0.6**
P1	0.3	0.35	0.29	0.29	**1.22**
P6	0.51	0.51	0.51	0.51	**2.04**
P9	0.51	0.53	0.52	0.52	**2.08**
P12	0.34	0.31	0.31	0.32	**1.29**
Method 2 Data	**total coliforms**	***E.Coli***	**enterococci**	**total bacteria**	
DAF	8.21	7.72	6.85	8.38	
P1	7.29	6.9	6.61	7.56	
P6	5.47	5.1	4.08	5.64	
P9	3.61	3.58	3.4	4.19	
P12	3.13	2.23	2.46	3.35	
Total	27.71	25.53	23.4	29.12	
Entropy	**total coliforms**	***E.Coli***	**enterococci**	**total bacteria**	
DAF	0.3	0.3	0.29	0.29	
P1	0.26	0.27	0.28	0.26	
P6	0.2	0.2	0.17	0.19	
P9	0.13	0.14	0.15	0.14	
P12	0.11	0.09	0.11	0.12	
Entropy	**total coliforms**	***E.Coli***	**enterococci**	**total bacteria**	**SUM**
DAF	0.52	0.52	0.52	0.52	**2.08**
P1	0.51	0.51	0.52	0.51	**2.04**
P6	0.46	0.46	0.44	0.46	**1.82**
P9	0.38	0.4	0.4	0.4	**1.59**
P12	0.36	0.31	0.34	0.36	**1.36**
Total Entropy	2.23	2.2	2.22	2.24	

Appendix E (ii) (Physical-Chemical incl. water depth)

Method 1 Data	turbidity	BOD	dissolved oxygen	redox potential	conductivity	water depth	
DAF	943.18	3836.64	0.4	253.91	4.58	1.48	
P1	350.88	1305	0.76	204.09	2.67	0.84	
P6	61.14	92.75	2.37	25.45	1.06	0.67	
P9	48.55	50.85	5.91	33.2	0.91	0.51	
P12	47.22	9.96	6.02	42.75	0.73	0.5	
Soft-scaling	turbidity	BOD	dissolved oxygen	redox potential	conductivity	water depth	
DAF	1.69	1.69	-0.99	1.31	1.58	1.68	
P1	0.16	0.15	-0.85	0.85	0.41	0.1	
P6	-0.59	-0.59	-0.27	-0.8	-0.57	-0.32	
P9	-0.62	-0.61	1.03	-0.72	-0.66	-0.72	
P12	-0.63	-0.64	1.07	-0.64	-0.77	-0.74	
Norm(S)	turbidity	BOD	dissolved oxygen	redox potential	conductivity	water depth	
DAF	0.95	0.95	0.16	0.9	0.94	0.95	
P1	0.56	0.56	0.2	0.8	0.66	0.54	
P6	0.28	0.28	0.39	0.21	0.29	0.37	
P9	0.27	0.27	0.85	0.23	0.26	0.24	
P12	0.27	0.26	0.86	0.26	0.22	0.23	
Entropy	turbidity	BOD	dissolved oxygen	redox potential	conductivity	water depth	**SUM**
DAF	0.06	0.06	0.19	0.13	0.08	0.07	**0.6**
P1	0.47	0.47	0.2	0.26	0.4	0.48	**2.27**
P6	0.51	0.51	0.53	0.48	0.52	0.53	**3.08**
P9	0.51	0.51	0.46	0.49	0.5	0.49	**2.96**
P12	0.51	0.51	0.47	0.51	0.48	0.49	**2.96**
Entropy	2.06	2.06	1.85	1.86	1.98	2.06	**2.96**
Method 2 Data	turbidity	BOD	dissolved oxygen	redox potential	conductivity	water depth	
DAF	943.18	3836.64	0.4	253.91	4.58	1.48	
P1	350.88	1305	0.76	204.09	2.67	0.84	
P6	61.14	92.75	2.37	25.45	1.06	0.67	
P9	48.55	50.85	5.91	33.2	0.91	0.51	
P12	47.22	9.96	6.02	42.75	0.73	0.5	
Total	1450.97	5295.2	15.46	559.4	9.96	3.99	
load	turbidity	BOD	dissolved oxygen	redox potential	conductivity	water depth	
DAF	0.65	0.72	0.03	0.45	0.46	0.37	
P1	0.24	0.25	0.05	0.36	0.27	0.21	
P6	0.04	0.02	0.15	0.05	0.11	0.17	
P9	0.03	0.01	0.38	0.06	0.09	0.13	
P12	0.03	0	0.39	0.08	0.07	0.13	
Entropy	turbidity	BOD	dissolved oxygen	redox potential	conductivity	water depth	**SUM**
DAF	0.4	0.34	0.53	0.52	0.52	0.53	**2.83**
P1	0.5	0.5	0.53	0.53	0.51	0.47	**3.04**
P6	0.19	0.1	0.41	0.2	0.34	0.43	**1.69**
P9	0.16	0.06	0.21	0.24	0.32	0.38	**1.37**
P12	0.16	0.02	0.14	0.28	0.28	0.38	**1.25**
Total Entropy	1.42	1.02	1.82	1.78	1.96	2.19	

Appendix E (iii) (inorganic ions)

Method 1 Data	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	
DAF	341.87	34.72	463.35	121.33	21.67	87.94	20.68	
P1	422.77	4.77	275.84	96.43	1.46	62.16	25.95	
P6	75.25	0.38	136.2	74.4	0.24	43.97	6.79	
P9	9.73	0.23	106.4	50.8	0.14	27.78	3.72	
P12	2.09	0.12	56.18	48.78	0	18.42	1.42	
Soft-scaling	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	
DAF	0.87	1.77	1.55	1.39	1.79	1.43	0.82	
P1	1.28	-0.22	0.41	0.58	-0.34	0.51	1.3	
P6	-0.48	-0.51	-0.43	-0.13	-0.47	-0.15	-0.45	
P9	-0.81	-0.52	-0.61	-0.89	-0.48	-0.73	-0.73	
P12	-0.85	-0.53	-0.92	-0.96	-0.49	-1.07	-0.94	
Norm(S)	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	
DAF	0.81	0.96	0.94	0.92	0.96	0.92	0.79	
P1	0.9	0.41	0.66	0.72	0.37	0.69	0.9	
P6	0.32	0.31	0.33	0.45	0.32	0.44	0.33	
P9	0.21	0.3	0.27	0.19	0.32	0.23	0.23	
P12	0.2	0.3	0.18	0.17	0.31	0.14	0.17	
Entropy	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	**SUM**
DAF	0.25	0.05	0.08	0.11	0.05	0.11	0.26	**0.92**
P1	0.14	0.53	0.39	0.34	0.53	0.37	0.13	**2.43**
P6	0.53	0.52	0.53	0.52	0.53	0.52	0.53	**3.67**
P9	0.47	0.52	0.51	0.45	0.53	0.49	0.49	**3.46**
P12	0.46	0.52	0.44	0.43	0.52	0.4	0.44	**3.22**
Method 2 Data	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	
DAF	341.87	34.72	463.35	121.33	21.67	87.94	20.68	
P1	422.77	4.77	275.84	96.43	1.46	62.16	25.95	
P6	75.25	0.38	136.2	74.4	0.24	43.97	6.79	
P9	9.73	0.23	106.4	50.8	0.14	27.78	3.72	
P12	2.09	0.12	56.18	48.78	0	18.42	1.42	
Total	851.7	40.21	1037.96	391.74	23.51	240.27	58.56	
Load	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	
DAF	0.4	0.86	0.45	0.31	0.92	0.37	0.35	
P1	0.5	0.12	0.27	0.25	0.06	0.26	0.44	
P6	0.09	0.01	0.13	0.19	0.01	0.18	0.12	
P9	0.01	0.01	0.1	0.13	0.01	0.12	0.06	
P12	0	0	0.05	0.12	0	0.08	0.02	
Entropy	Ammonia	Fluoride	Chloride	Nitrate	Nitrite	Phosphate	Sulphate	**SUM**
DAF	0.53	0.18	0.52	0.52	0.11	0.53	0.53	**2.92**
P1	0.5	0.36	0.51	0.5	0.25	0.5	0.52	**3.15**
P6	0.31	0.06	0.38	0.46	0.07	0.45	0.36	**2.09**
P9	0.07	0.04	0.34	0.38	0.04	0.36	0.25	**1.49**
P12	0.02	0.02	0.23	0.37	0	0.28	0.13	**1.06**
Total Entropy	1.43	0.68	1.98	2.23	0.47	2.13	1.79	

Appendix F

Scenario 1	Frequency	Redox Potential	Water Temp	air temp	conductivity	Dissolved oxygen	BOD	Turbidity	total coliforms	E.coli	Enterococci
DAF	**8.10%**	-249.4	17.96	16.44	5.006	0.4	3459	860	9.12	8.35	7.89
P1	**16.89%**	-227.9	9.72	9.44	2.351	0.64	2257	621	8.21	7.48	6.88
P6	**27.10%**	-26.5	12.44	15.54	1.426	2.16	309	78	6.04	5.46	5.31
P9-P12	**27.50%**	32.3	16.1	19.8	0.793	5.6	24	59	4.82	4.09	3.86
P12-stream	**21.10%**	44.9	7.39	9.45	0.792	6.43	25	24	3.58	3.33	3.07
Scenario 2	Frequency	Redox Potential	Water Temp	air temp	conductivity	Dissolved oxygen	BOD	Turbidity	total coliforms	E.coli	Enterococci
DAF	**7.58%**	-277.6	20.42	17.7	6.312	0.34	4186	933	9.56	9.21	7.89
P1	**24.24%**	-221.2	12.59	12.87	2.835	0.66	2178	590	8.14	7.33	7.09
P6	**19.70%**	14.3	11.77	14.93	1.136	2.41	139	54	5.82	5.19	5.05
P9-P12	**21.21%**	32.3	16.1	19.8	0.793	5.6	24	59	4.82	4.09	3.86
P12-stream	**27.27%**	44.9	7.39	9.45	0.792	6.43	25	24	3.58	3.33	3.07
Scenario 3	Frequency	Redox Potential	Water Temp	air temp	conductivity	Dissolved oxygen	BOD	Turbidity	total coliforms	E.coli	Enterococci
DAF	**13.64%**	-252.4	18.53	17.26	5.172	0.4	3429	855	9.13	8.35	7.89
P1	**13.64%**	-227.2	10.07	9.4	2.48	0.6	2420	653	8.31	7.58	6.99
P6	**24.24%**	-26.5	12.44	15.54	1.426	2.16	309	78	6.04	5.46	5.31
P9-P12	**21.21%**	32.3	16.1	19.8	0.793	5.6	24	59	4.82	4.09	3.86
P12-stream	**27.27%**	44.9	7.39	9.45	0.792	6.43	25	24	3.58	3.33	3.07
Scenario 4	Frequency	Redox Potential	Water Temp	air temp	conductivity	Dissolved oxygen	BOD	Turbidity	total coliforms	E.coli	Enterococci
DAF	**15.15%**	-249.4	17.96	16.44	5.006	0.4	3459	860	9.12	8.35	7.89
P1	**12.12%**	-227.9	9.72	9.44	2.351	0.64	2257	621	8.21	7.48	6.88
P6	**24.24%**	-26.5	12.44	15.54	1.426	2.16	309	78	6.04	5.46	5.31
P9-P12	**21.21%**	32.3	16.1	19.8	0.793	5.6	24	59	4.82	4.09	3.86
P12-stream	**27.27%**	44.9	7.39	9.45	0.792	6.43	25	24	3.58	3.33	3.07
Scenario 5	Frequency	Redox Potential	Water Temp	air temp	conductivity	Dissolved oxygen	BOD	Turbidity	total coliforms	E.coli	Enterococci
DAF	**6.06%**	-277.8	18.38	13.82	6.04	0.19	4119	780	10.4	9.99	9.04
P1	**16.67%**	-188.1	17.15	19.22	3.422	0.99	2017	564	8.5	8.04	9.53
P6	**19.70%**	-129.4	9.05	9.22	1.918	1.06	1253	411	8.56	7.55	6.95
P9-P12	**36.36%**	29.7	13.93	17.98	0.858	5.2	32	47	6.32	5.24	6.34
P12-stream	**21.21%**	46.2	7.19	8.58	0.707	6.04	28	19	4.23	4.2	3.99

Appendix F (contd)

Scenario 1	water depth	Fluoride	Chloride	Nitrate	Nitrite	phosphate	sulphate	ammonia	Solar radiation	Penman
DAF	1.311	30.14	460.2	75	18.52	74.2	22.44	365.6	427.5	68.11
P1	1.076	3.65	211.2	161.1	1.69	79.5	32.62	415.7	150.7	17.75
P6	0.661	0.97	153.7	101	0.19	53.1	7.07	85.3	367.9	55.68
P9-P12	0.346	2.1	75.1	35.9	0.11	11.7	9.35	2	469.7	77.27
P12-stream	0.83	0.2	68.5	60.7	0.01	17.5	3.28	6.4	211.7	26.45
Scenario 2	water depth	Fluoride	Chloride	Nitrate	Nitrite	phosphate	sulphate	ammonia	Solar radiation	Penman
DAF	1.47	52.58	562.4	20.9	28.85	88.6	21.15	319.2	406.4	66.7
P1	1.022	5.24	272.2	154.5	4	65.9	26.1	413.9	297.5	42.64
P6	0.661	0.89	151.6	80.7	0.18	53.7	6.8	85.3	352.1	53.5
P9-P12	0.346	2.1	75.1	35.9	0.11	11.7	9.35	2	469.7	77.27
P12-stream	0.83	0.2	68.5	60.7	0.01	17.5	3.28	6.4	211.7	26.45
Scenario 3	water depth	Fluoride	Chloride	Nitrate	Nitrite	phosphate	sulphate	ammonia	Solar radiation	Penman
DAF	1.29	32.75	500.7	51.3	17.64	76	23.16	358.4	427.5	68.11
P1	1.123	5.3	204.4	177.1	6.28	75.9	29.58	406.2	150.7	17.75
P6	0.661	0.97	153.7	101	0.19	53.1	7.07	85.3	367.9	55.68
P9-P12	0.346	2.1	75.1	35.9	0.11	11.7	9.35	2	469.7	77.27
P12-stream	0.83	0.2	68.5	60.7	0.01	17.5	3.28	6.4	211.7	26.45
Scenario 4	water depth	Fluoride	Chloride	Nitrate	Nitrite	phosphate	sulphate	ammonia	Solar radiation	Penman
DAF	1.311	30.14	460.2	75	18.52	74.2	22.44	365.6	427.5	68.11
P1	1.076	3.65	211.2	161.1	1.69	79.5	32.62	415.7	150.7	17.75
P6	0.661	0.97	153.7	101	0.19	53.1	7.07	85.3	367.9	55.68
P9-P12	0.346	2.1	75.1	35.9	0.11	11.7	9.35	2	469.7	77.27
P12-stream	0.83	0.2	68.5	60.7	0.01	17.5	3.28	6.4	211.7	26.45
Scenario 5	water depth	Fluoride	Chloride	Nitrate	Nitrite	phosphate	sulphate	ammonia	Solar radiation	Penman
DAF	1.38	62.62	493	22.2	38.47	83.2	24	312.5	276.2	41.75
P1	0.631	6.92	357.4	107.3	1.08	66.2	46.7	381.4	464.8	73.64
P6	0.779	3.48	181.7	167	4.13	65.8	21	273.6	152.5	18.25
P9-P12	0.481	1.33	97.2	26.7	0.07	20	7.8	36.7	459.4	72.33
P12-stream	0.861	0.25	60.4	83.5	0	19.9	3.4	6.4	127.6	14.14

Appendix F, shows the 5 scenario's undertaken using the SOMine software, the variables of interest are:

1. Physical Chemical [BOD, dissolved oxygen, turbidity, redox potential and conductivity]
2. Bacterial concentrations [Log_{10} total coliforms, *E.coli* and enterococci]
3. Inorganic Ions [ammonia, nitrate, phosphate and sulphate]

Appendix G

sample_points	Water Depth(m)	area (m2)	volume (m3)	Rainfall Amount(mm)	Penman (mm)	Solar Radiation (MJ/sqm)
DAF_Feb	1.50	N/A	N/A	96.00	16.75	133.07
DAF_Mar	1.45	N/A	N/A	67.84	37.25	273.76
DAF_Apr	1.50	N/A	N/A	25.17	68.00	499.21
DAF_May	1.50	N/A	N/A	43.11	94.25	556.97
DAF_Jun	1.45	N/A	N/A	152.07	78.25	459.38
DAF_Jul	1.45	N/A	N/A	113.09	86.50	527.60
DAF_Aug	1.45	N/A	N/A	103.97	70.25	424.78
DAF_Sept	1.50	N/A	N/A	45.40	47.00	334.95
DAF_Nov	1.50	N/A	N/A	63.61	6.00	89.17
DAF_Dec	1.50	N/A	N/A	81.39	4.25	63.23
Pond1_Feb	0.87	1898.00	1,651.26	96.00	16.75	133.07
Pond1_Mar	0.82	1,898.00	1,556.36	67.84	37.25	273.76
Pond1_Apr	0.72	1,898.00	1,366.56	25.17	68.00	499.21
Pond1_May	0.63	1,898.00	1,195.74	43.11	94.25	556.97
Pond1_Jun	0.63	1,898.00	1,195.74	152.07	78.25	459.38
Pond1_July	0.60	1,898.00	1,138.80	113.09	86.50	527.60
Pond1_Aug	0.63	1,898.00	1,195.74	103.97	70.25	424.78
Pond1_Sept	0.63	1,898.00	1,195.74	45.40	47.00	334.95
Pond1_Nov	0.73	1,898.00	1,385.54	63.61	6.00	89.17
Pond1_Dec	0.83	1,898.00	1,575.34	81.39	4.25	63.23
Pond6_Feb	0.72	1,804.00	1,298.88	96.00	16.75	133.07
Pond6_Mar	0.72	1,804.00	1,298.88	67.84	37.25	273.76
Pond6_Apr	0.75	1,804.00	1,353.00	25.17	68.00	499.21
Pond6_May	0.62	1,804.00	1,118.48	43.11	94.25	556.97
Pond6_Jun	0.58	1,804.00	1,046.32	152.07	78.25	459.38
Pond6_July	0.56	1,804.00	1,010.24	113.09	86.50	527.60
Pond6_Aug	0.59	1,804.00	1,064.36	103.97	70.25	424.78
Pond6_Sept	0.59	1,804.00	1,064.36	45.40	47.00	334.95
Pond6_Nov	0.69	1,804.00	1,244.76	63.61	6.00	89.17
Pond6_Dec	0.75	1,804.00	1,353.00	81.39	4.25	63.23
Pond9_Feb	0.85	2,813.00	1,884.71	96.00	16.75	133.07
Pond9_Mar	0.67	2,813.00	1,884.71	67.84	37.25	273.76
Pond9_Apr	0.68	2,813.00	1,912.84	25.17	68.00	499.21
Pond9_May	0.42	2,813.00	1,181.46	43.11	94.25	556.97
Pond9_Jun	0.23	2,813.00	646.99	152.07	78.25	459.38
Pond9_July	0.40	2,813.00	1,125.20	113.09	86.50	527.60
Pond9_Aug	0.36	2,813.00	1,012.68	103.97	70.25	424.78
Pond9_Sept	0.19	2,813.00	534.47	45.40	47.00	334.95
Pond9_Nov	0.60	2,813.00	1,687.80	63.61	6.00	89.17
Pond9_Dec	0.72	2,813.00	2,025.36	81.39	4.25	63.23
Pond12_Feb	1.10	2,428.00	2,670.80	96.00	16.75	133.07

Pond12_Mar	0.77	2,428.00	1,869.56	67.84	37.25	273.76
Pond12_Apr	0.74	2,428.00	1,796.72	25.17	68.00	499.21
Pond12_May	0.00	2,428.00	0.00	43.11	94.25	556.97
Pond12_Jun	0.00	2,428.00	0.00	152.07	78.25	459.38
Pond12_July	0.23	2,428.00	558.44	113.09	86.50	527.60
Pond12_Aug	0.22	2,428.00	534.16	103.97	70.25	424.78
Pond12_Sept	0.00	2,428.00	0.00	45.40	47.00	334.95
Pond12_Nov	0.65	2,428.00	1,578.20	63.61	6.00	89.17
Pond12_Dec	0.73	2,428.00	1,772.44	81.39	4.25	63.23
Stream_Feb	1.23	N/A	N/A	96.00	16.75	133.07
Stream_Mar	0.87	N/A	N/A	67.84	37.25	273.76
Stream_Apr	0.83	N/A	N/A	25.17	68.00	499.21
Stream_May	0.77	N/A	N/A	43.11	94.25	556.97
Stream_Jun	0.67	N/A	N/A	152.07	78.25	459.38
Stream_July	0.72	N/A	N/A	113.09	86.50	527.60
Stream_Aug	0.64	N/A	N/A	103.97	70.25	424.78
Stream_Sept	0.67	N/A	N/A	45.40	47.00	334.95
Stream_Nov	0.90	N/A	N/A	63.61	6.00	89.17
Stream_Dec	1.07	N/A	N/A	81.39	4.25	63.23