Assessing validation and reliability of water and wastewater treatment processes using Bayesian techniques

Author:
Carvajal Ortega, Guido

Publication Date:
2018

DOI:
https://doi.org/10.26190/unsworks/3357

Assessing validation and reliability of water and wastewater treatment processes using Bayesian techniques

Guido Esteban Aquiles Carvajal Ortega

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Civil and Environmental Engineering

Faculty of Engineering

January 2018

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: Carvajal Ortega

First name: Guido Other name/s: Esteban Aquiles

Abbreviation for degree as given in the University calendar: PhD

School: Civil and Environmental Engineering Faculty: Engineering

Title: Assessing validation and reliability of water and wastewater treatment processes using Bayesian techniques

Abstract 350 words maximum: (PLEASE TYPE)

Water and wastewater treatment systems have key roles in protecting public health. To maintain public confidence, techniques are required to quantitatively validate their performance for achieving satisfactory water quality and to characterise system reliability.

Water quality management, especially potable water reuse, is transitioning to a point where validation and reliability assessment will be essential. Quantitative validation and reliability analyses can assist in determining factors affecting performance, calculating system failure probabilities and supporting decision-making. However, current water quality management guidelines and practices do not promote validation and reliability assessment tools or methods.

Research presented in this thesis identifies validation and reliability analysis tools and demonstrates their applicability for water and wastewater treatment. Following a review of commonly used candidate analysis tools, Bayesian Networks (BNs) and methods were identified as very promising for validation and reliability assessment. Applications of BNs and Bayesian methods were investigated in six diverse water and wastewater treatment settings (i.e. activated sludge, ultrafiltration, ozonation, chlorination, UV disinfection and a multiple barrier system). These cases considered a range of issues including evaluation of monitoring parameters for log reduction values (LRVs), optimising LRV calculations, system reliability assessment, and multi-barrier evaluation. These were selected to represent problems and limitations encountered with water reuse.

BNs and Bayesian methods were found to be flexible and applicable to a wide range of cases and problems. The tools were shown to be applicable to systems where high levels of assessment and understanding of water quality are essential. The findings should now serve as templates to facilitate validation and reliability assessment of full scale water reuse applications.

In conclusion, it is proposed that my findings will be of significant value to the water industry and should aid in the further development of potable water reuse, while providing enhanced protection of public health to water supply customers.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

ORIGINALITY STATEMENT

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.’

Signed

Date

Acknowledgements

I would like to thank my dear mother Nancy for her support and unconditional love all these years. You have really being an exemplary mother, I am really proud of being your son. I cannot express but gratefulness to my lovely wife Romina for standing next to me all these years and leaving behind everything in Chile to support me during this process. I truly appreciate all your love and care. You are an awesome person and I am really proud for your achievements and the way you see life. I also need to thank all my family for their love despite of the long distance. I consider important to acknowledge the financial support from the Chilean National Commission for Scientific and Technological Research CONICYT- Becas Chile which funded my Masters and PhD studies at UNSW.

I would like to thank my two joint supervisors Dr. Stuart Khan and Dr. David Roser for their patience, support and often wise advice during my PhD. I would like to especially thank my co-supervisor Dr. Scott Sisson for giving me part of his valuable time to answer my sometimes silly statistical questions.

I really need to thank all my friends from Chilean Student Association (CHISA@UNSW) for their friendship and company. I also want to say thanks to my very good friends Juan, Angeles, Andrea, Rodrigo, Vanessa and Alvaro for their awesome friendship and help during difficult times. Special thanks to Alicia for hosting me in her place during the most difficult part of the PhD. It was awesome to have great Chilean friends here and feel a little bit closer to home. We will have a nice reunion back in Chile!. Many thanks to all my colleagues and friends from the Valentine Annex for their good vibes and willingness to help when needed.

I would also like to thank Dr. Pierre LeClech, Dr. Amos Branch and Dr. Trang Trinh for the opportunity to work with them during the MBR validation project; it was a really growing experience. I would like to especially thank Amos as well for helping me a lot during the development of chapters 5 and 6.

I need to also thank Dr. Louis Wiart from Xylem for the assistance in the UV reliability chapter. I am really grateful for all his time and input. I also have to thank Dr. Jörg Drewes and Mr. Philipp Michel from the Chair of Urban Water Systems Engineering, Technical University of Munich.

I would like to sincerely thank the people from SA Water, Dr. Ben Van Den Akker, Dr. Alexandra Keegan, and Mr. Rudi Regel for sharing valuable data and allocating time to provide feedback and input to my work. The information provided was crucial for the development of the models in my thesis.

Finally, I really need to thank the School staff for their direct or indirect support during all this time at UNSW. Special thanks to Ms Patricia McLaughlin for her patience and help during hard times of my PhD.

Abstract

Water and wastewater treatment systems have key roles in protecting public health. To maintain public confidence, techniques are required to quantitatively validate their performance for achieving satisfactory water quality and to characterise system reliability.

Water quality management, especially potable water reuse, is transitioning to a point where validation and reliability assessment will be essential. Quantitative validation and reliability analyses can assist in determining factors affecting performance, calculating system failure probabilities and supporting decision-making. However, current water quality management guidelines and practices do not promote validation and reliability assessment tools or methods.

Research presented in this thesis identifies validation and reliability analysis tools and demonstrates their applicability for water and wastewater treatment. Following a review of commonly used candidate analysis tools, Bayesian Networks (BNs) and methods were identified as very promising for validation and reliability assessment. Applications of BNs and Bayesian methods were investigated in six diverse water and wastewater treatment settings (i.e. activated sludge, ultrafiltration, ozonation, chlorination, UV disinfection and a multiple barrier system). These cases considered a range of issues including evaluation of monitoring parameters for log reduction values (LRVs), optimising LRV calculations, system reliability assessment, and multi-barrier evaluation. These were selected to represent problems and limitations encountered with water reuse.

BNs and Bayesian methods were found to be flexible and applicable to a wide range of cases and problems. The tools were shown to be applicable to systems where high levels of assessment and understanding of water quality are essential. The findings should now serve as templates to facilitate validation and reliability assessment of full scale water reuse applications.

In conclusion, it is proposed that my findings will be of significant value to the water industry and should aid in the further development of potable water reuse, while providing enhanced protection of public health to water supply customers.

Table of contents

List of figures

List of tables

Abbreviations

AIC	Akaike information criterion
API	Application Programming Interface
AUC	Area under the curve
BAN	Bayesian network augmented naïve Bayes
BMLP	Bayesian multilayer perceptron
BN	Bayesian network
BOD	Biological oxygen demand
C. perfringens	*Clostridium perfringens*
C. parvum	*Criptosporidium parvum*
CFU	Colony forming units
CI	Credible interval
COD	Chemical oxygen demand
CT	Concentration times time
DIC	Deviance information criterion
DOC	Dissolved organic carbon
E. coli	*Escherichia coli*
EEM	Excitation emission matrices
ETA	Event tree analysis
FMEA	Failure modes and effects analysis
FMECA	Failure modes effects and criticality analysis
FNR	False negative rate
FPR	False positive rate
FTA	Fault tree analysis
G. lamblia	*Giardia lamblia*
HACCP	Hazard analysis and critical control points
HBM	Hierarchical Bayesian model

HHRA	Human health risk assessment
HRA	Human reliability analysis
HRT	Hydraulic retention time
KS	Kappa statistic
LOPA	Layers of protection analysis
LRV	Log reduction value
MA	Markov analysis
MCMC	Markov chain Monte Carlo
MCS	Monte Carlo simulation
MLSS	Mixed liquour suspended solids
MLVSS	Mixed liquour volatile suspended solids
MS2	MS2 bacteriophage
MSE	Mean squared error
NB	Naïve Bayes model
NHMB	Non-hiearchical Bayesian model
O_3	Ozone
PARAFAC	Parallel factor analysis
PDF	Probability density function
PFU	Plaque forming units
QMRA	Quantitative microbial risk analysis
RBD	Reliability block diagram
RCM	Reliability centred maintenance
RO	Reverse osmosis
ROC	Receiver operating curve
ROPE	Region of practical equivalence
SIF	Safety instrumented function
SNB	Semi naïve Bayes model
SRT	Solids retention time

SS	Suspended solids
TAN	Tree augmented naïve Bayes model
TF	Total fluorescence
TNR	True negative rate
TOC	Total organic carbon
TPR	True positive rate
UF	Ultrafiltration
UV	Ultraviolet
UVA	UV absorbance
WEKA	Waikato Environment for Knowledge Analysis

Glossary

Concept	Definition
Akaike information criterion	Information-theoretic scoring function, which trades off the model's goodness of fit with its complexity (Kjræulff & Madsen, 2012).
Application Programming Interface	Set of algorithms and functions for a specific application that can be used within a different software application.
Area under the curve	Area Under the Curve for the receiver operating characteristic curve (Witten & Frank, 2005).
Bayesian multilayer perceptron	A multilayer perceptron (see below for a definition) with parameters determined through Bayesian analysis (Vehtari *et al.*, 2000).
Bayesian network	Probabilistic graphical models using direct acyclic graphs and Bayes' rule to perform inference (Korb & Nicholson, 2011).
Bayesian network augmented naïve Bayes	Semi-naïve Bayes model. Two or more arcs between attributes are allowed.
Bow-tie analysis	System reliability technique used to model causes to an undesired event, preventive controls, mitigative controls and consequences (IEC/ISO, 2009).
Copula	Multivariate probability distribution used to describe the dependence between random variables having fixed marginal distributions (Joe, 2014).

Deviance information criterion	Metric used to compare Bayesian models. It combines a measure of goodness of fit (mean deviance) with a measure of complexity or penalty term, (effective number of parameters) (Lunn *et al.*, 2012).
Direct potable reuse	Process used to augment drinking water supply from sewage without the use of an environmental buffer (Khan, 2013).
Event tree analysis	System reliability modelling technique used to model sequential events. It includes an initiating event, pivotal events and consequences (ISO/TR, 2013).
Fluorescence excitation emission matrix	Fluorescence intensities generated for different combinations of excitation and emission light wavelengths (Murphy *et al.*, 2013).
Failure modes and effects analysis	System reliability technique to analyse failure modes, their consequences and controls (IEC/ISO, 2009).
Failure modes effects and criticality analysis	System reliability technique to analyse failure modes, their consequences and controls including quantification of events (IEC/ISO, 2009).
False negative rate	Rate of incorrect negative predictions.
False positive rate	Rate of incorrect positive predictions.
Fault tree analysis	System reliability modelling technique using events and logic gates to find causal events leading to a target event occurrence (IEC/ISO, 2009).

Hierarchical Bayesian model	Type of Bayesian model on which parameters of the likelihood are treated as random variables with their own set of priors.
Indirect potable reuse	Process used to augment drinking water supply from sewage with the use of an environmental buffer (Khan, 2013).
Kappa statistic	Measures the agreement between model predictions and actual values as a metric in the range [-1,1] (Marcot, 2012).
Layers of protection analysis	System reliability modelling technique used to model sequential events. It includes an initiating event, independent barriers and consequences (IEC/ISO, 2009).
Log-likelihood	Measures how well the data fit each model.
LRV	Log reduction value (or Log removal value).
Mapping	Encode or transform a method using a different methodology.
Markov analysis	Stochastic modelling technique used to model dynamic transitions between system states (IEC/ISO, 2009).
Markov chain Monte Carlo	Simulation technique used to calculate complex integrals. Usually used within Bayesian analysis to calculate posterior distributions (Ntzoufras, 2011).

Multilayer perceptron	Type of neural network composed of layers of neurons (elements that generate a transformation of the inputs) with an input layer, at least one hidden layer, and an output layer. In these models the inputs of the neurons in one layer come from the outputs of neurons in a previous layer. Neurons in one layer are connected to the previous layer through weighted connections (Priddy & Keller, 2005).
Naïve Bayes model	Type of Bayesian network used for prediction in which the target node is the parent of all other nodes. This type of model does not assume any type of prior understanding of how the system works.
Bayesian model	Mathematical model using prior distributions for parameters and Bayes' rule to obtain posterior parameter distributions.
Parallel factor analysis	Statistical technique used to decompose tridimensional data and derive independent underlying signals (Murphy *et al.*, 2013).
Receiver operating curve	Plot showing the relationship between the false positive rate (or 1-specificity) and the true positive rate (or sensitivity) for various decision thresholds (Witten & Frank, 2005).
Region of practical equivalence	Small range of parameter values that are deemed to be equivalent to the null value when used in a particular application. Used for hypothesis testing during Bayesian analysis (Kruschke, 2014).

Reliability	Probability for an item to perform a required function under given conditions over a time interval (ISO, 2013a).
Reliability block diagram	System reliability technique in which various system configurations can be analysed. It uses blocks and connections to depict the system (IEC, 2008).
Reliability centred maintenance	System reliability technique used to reduce maintenance costs (IEC/ISO, 2009).
Risk	A set of triplets comprising: 1) a scenario, 2) likelihood of that scenario and 3) consequence or measure of damage (Kaplan & Garrick, 1981).
Semi naïve Bayes model	Type of Naïve Bayes model in which arcs between non-target nodes are allowed.
Soft systems	These elements are used to capture high level interactions. They may incorporate organisation culture, resources, processes, and procedures (Fenton & Neil, 2012).
Tree augmented naïve Bayes model	Type of Naïve Bayes model in which at most one arc between non-target nodes is allowed.
True negative rate	Rate of correct negative predictions.
True positive rate	Rate of correct positive predictions.

Uncertainty	In a risk assessment context uncertainty relates to the lack of data or incomplete understanding of a phenomenon (Aven, 2012). In a Bayesian modelling context uncertainty is reflected in the prior and posterior distributions of model parameters (Ntzoufras, 2011).
Validation	Analysis used to ensure that a treatment system is performing adequately (ISO, 2005).
Variability	Variability relates to heterogeneity of data during an analysis. It cannot be reduced by further observations (Aven, 2012).
Waikato Environment for Knowledge Analysis	A specific software package for Machine learning.
Water reuse	Use of treated sewage for a purpose different from discharge to the environment. Uses can be for potable or non-potable applications (Khan, 2013).

List of publications

Peer reviewed journal articles related to this thesis

Carvajal, G., Roser, D. J., Sisson, S. A., Keegan, A., & Khan, S. J. (2015). Modelling pathogen log_{10} reduction values achieved by activated sludge treatment using naïve and semi naïve Bayes network models. Water research, 85, 304-315. DOI: 10.1016/j.watres.2015.08.035

Carvajal, G., Roser, D.J., Sisson, S.A., Keegan, A. and Khan, S.J. (2017). Bayesian belief network modelling of chlorine disinfection for human pathogenic viruses in municipal wastewater. Water research 109, 144-154. DOI: 10.1016/j.watres.2016.11.008

Carvajal, G., Branch, A., Sisson, S.A., Roser, D.J., van den Akker, B., Monis, P., Reeve, P., Keegan, A., Regel, R. and Khan, S.J. (2017). Virus removal by ultrafiltration: Understanding long-term performance change by application of Bayesian analysis. Water research. 122, 269-279. DOI: 10.1016/j.watres.2017.05.057

Carvajal, G., Branch, A., Michel, P., Sisson, S.A., Roser, D.J., Drewes, J. E. and Khan, S.J. (2017). Robust evaluation of performance monitoring options for ozone disinfection in water recycling using Bayesian analysis. Water research. 124, 605-617. DOI: 10.1016/j.watres.2017.07.079

Peer reviewed journal articles not related to this thesis

Trinh, T., Branch, A., Hambly, A.C., **Carvajal, G.**, Coleman, H.M., Stuetz, R.M., Drewes, J.E., Le-Clech, P. and Khan, S.J. (2016). Hazardous events in membrane bioreactors–Part 1: Impacts on key operational and bulk water quality parameters. Journal of Membrane Science 497, 494-503. DOI: 10.1016/j.memsci.2015.03.003

Branch, A., Trinh, T., **Carvajal, G.**, Leslie, G., Coleman, H.M., Stuetz, R.M., Drewes, J.E., Khan, S.J. and Le-Clech, P. (2016). Hazardous events in membrane bioreactors–Part 3: Impacts on microorganism log removal efficiencies. Journal of Membrane Science 497, 514-523. DOI: 10.1016/j.memsci.2015.10.011

Peer reviewed conference papers

Carvajal, G.E., Roser, D.J., Sisson, S. and Khan, S.J. (2015). Multivariate analysis of activated sludge pathogen removal through Bayesian network modelling. Australian Water Association: OzWater 2015 Conference, Adelaide Convention Centre, Adelaide, 12th – 14th May [Oral Presentation].

Conference proceedings

Carvajal, G.E., Roser, D.J. and Khan, S.J. (2014). Prediction of microbial pathogens and indicators removal during wastewater treatment using Bayesian networks. Australasian Bayesian Network Modelling Society: 2014 Conference, Holiday Inn, Rotorua, New Zealand, 26th – 27th November [Oral Presentation].

Carvajal, G.E., Roser, D.J. and Khan, S.J. (2015). A Bayesian belief network analysis to infer an ultrafiltration membrane condition. Australasian Bayesian Network Modelling Society: 2015 Conference, Monash University, Caulfield Campus, Melbourne, 25th – 26th November [Oral Presentation].

Carvajal, G.E., Roser, D.J. and Khan, S.J. (2016). Removal of microbial indicators and trace organics during ozonation and biological media filtration for wastewater reclamation. Water Reuse Foundation: 2016 Research Conference, Westin Denver Downtown Hotel, Denver, Colorado, USA, 23rd - 24th May [Oral Presentation].

Peer Reviewed Reports

D. Roser, **G. Carvajal**, B. van den Akker, A. Keegan, R. Regel and S. Khan (2015). National Validation Guidelines for Water Recycling: Comprehensive Bayesian Recycled Water Validation, Australian Water Recycling Centre of Excellence, Brisbane, Australia.

Chapter 1: Introduction

1.1 Background

The use of reclaimed water for augmentation of drinking water supplies is associated with potential exposure to water quality hazards including pathogenic subtsances in raw source waters. Potentially increased exposure to such hazards requires assessment and management of associated risks (see Table 1-1 for a detailed definition of the concept). Consequently, risk management has become a fundamental component within a number of international guidelines for water reuse (NHMRC *et al.*, 2006; WHO, 2006; Health Canada, 2007).

A key objective of water reuse treatment systems is to ensure that sufficient treatment is provided, often through the use of multiple barriers, to manage health risks to acceptable levels, particularly in respect to pathogens. Each of the multiple barriers contributes to the overall system removal of the targeted pathogens. The implemented technologies, their organisational functioning and interaction, and the impacts from external sources, such as environmental conditions all play a role in determining the overall treatment performance and reliability (see Table 1-1 for a detailed definition of the concept) (Figure 1-1). Performance and reliability requirements are also dependent upon the location of the system and the final intended use of product water. For example, Californian water reuse standards currently require treatment trains to achieve ≥12 log removal of enteric viruses from raw wastewater to drinking water (California Office of Administrative Law, 2017).

Analysis of reliability in treatment processes can be considered as part of the risk assessment framework and involves understanding the interactions between the key operational parameters in the system. In this way, process performance can be monitored and the impact from system conditions can be evaluated. Although the mechanisms of removal are generally known, they may not always be able to be directly measured due to the lack of suitable (and affordable) analytical technologies. However, indirect measurements are often possible and these can be related to removal efficiencies by inference and statistical analysis. For example, pressure decay tests have been related to the performance of microfiltration membranes for protozoan removal (USEPA, 2005).

Table 1-1: Key definitions for validation, risk and reliability

Term	Definition
Validation	This concept is defined as *"confirmation, through the provision of objective evidence, that the requirement for a specific intended use or application have been fulfilled"* (ISO, 2005). Validation is central to the approval and commissioning of reclaimed water treatment systems. The following definition of this concept relates to water treatment: *"validation of processes is required to show that treatment processes can operate as required*[1]*. It can be undertaken during pilot stage studies, during initial implementation of a new or alternative water treatment system and is a useful tool in the optimisation of existing treatment processes*" (WHO, 2005). From the point of view of pathogen removal performance, validation has also been defined as *"the process of demonstrating that: 1) a treatment system can produce water of the required microbial quality under a defined range of operating conditions and 2) the system can be monitored in real time to provide assurance that the water quality objectives are being continuously met"*(VDoH, 2013).
Risk	Many definitions of risk have been proposed depending on the context. A general definition describes risk as the "*effect of uncertainty on objectives*" (ISO, 2009). Risk has also been defined as a set of triplets comprising: 1) a scenario, 2) likelihood of that scenario occurring and 3) consequence or measure of damage if/when the scenario does occur (Kaplan & Garrick, 1981). In a drinking water management context, risk has been defined as *"the likelihood of identified hazards causing harm in exposed populations in a specified timeframe, including the magnitude of that harm and/or the consequences"* (WHO, 2005).
Reliability	As in the case of risk, reliability has been the subject of multiple definitions. Reliability can be defined as the "*probability for an item to perform a required function under given conditions over a time interval"* (ISO, 2013a). For water treatment, the reliability concept has been defined slightly differently. In this

[1] i.e. reliably

Term	Definition
	context, reliability has been described as *"the percentage of time at which the expected effluent concentrations comply with specified discharge standards or treatment targets"* (Niku *et al.*, 1979; Oliveira & Von Sperling, 2008). In accord with this latter definition, but considering risks, reliability has also been defined as *"the ability of a water treatment system to provide water that consistently meets or exceeds the public health protection standard"* (Pecson et al., 2015).

One of the key current drivers for studying reliability in water treatment processes is an increased interest in direct potable reuse (DPR). In recent times, there has been a growing interest in DPR especially in the United States of America, Australia and South Africa which has led to an increasing research on its application on a larger scale (Khan, 2013). However, because DPR does not incorporate an environmental buffer as in indirect potable reuse, any available "time to respond" to failures is significantly reduced. Therefore, it is of paramount importance to understand and manage the system reliability and levels of risks in general and to validate such systems with these requirements in mind.

Uncertainty and variability are key features of both water treatment performance and reliability monitoring data. Consequently, it is essential for risk assessment to effectively account for them. Validation (see Table 1-1 for a detailed definition of the concept) of systems can be viewed as the assessment of reliability by quantifying the frequency with which a system is performing within its intended operational boundaries and the expected pathogen removal is effectively achieved. Validation assessment is increasingly undertaken using models incorporating a variety of input data and expert opinion.

Correct model representation is crucial for an adequate assignment of credits for pathogen removal. For example, in systems which operate parallel subsystems such as ultrafiltration "skids", individual processes may be deemed as equal, dismissing any differences due to impacts of continuous operation. Alternatively, considering these systems as related but not equal would increase the variability of the predictions needed and increase uncertainty when only limited observations are available for some units.

Reliability might also incorporate hardware, software and human reliability aspects affecting a determined outcome such as water quality related to pathogens. Although these three aspects of

reliability assessment are all influential, hardware, especially sensor performance is particularly crucial for a correct functioning of water treatment systems.

Based on my reading of these issues and concepts a problem framework was developed (Figure 1-1). This scheme shows how the various factors were initially seen as interacting within the context of reliability assessment with factors within a common block most likely to interact with each other. The top block represents the three major influences controlling the reliability of a system (i.e. system characteristics, environment and organisation) (Jackson, 2017). They in turn influence reliability domains within a system including hardware, software and human reliability (Jackson, 2017). These reliability domains further impact performance by establishing the conditions under which a system works.

Securing a high level of process performance for water quality involves the application of a number of operational monitoring and process control measurements (Hamilton *et al.*, 2006). Validation ensures the correct system functioning and removal performance, comprising efficient operational monitoring and process control. The aspects considered in this lower block provide information which is used as feedback for the factors impacting reliability and the overall reliability of the system. At every level of this scheme, appropriate tools for describing, quantifying and assessing the system are needed to support the management framework.

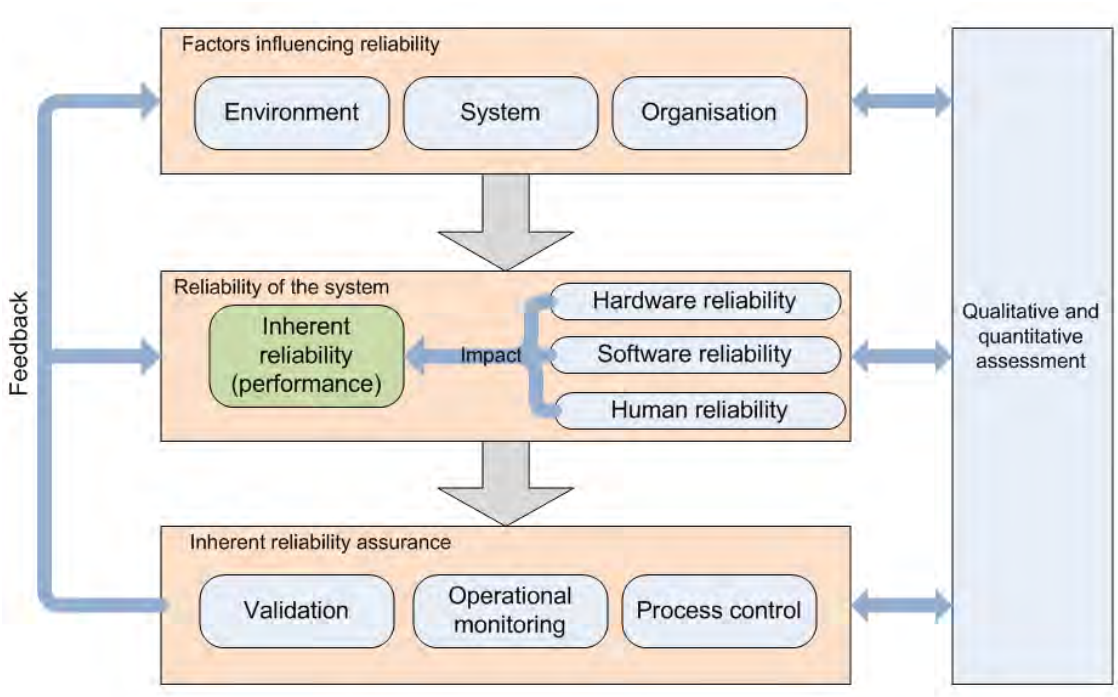

Figure 1-1: Schematic diagram of the factors affecting reliability and their quantification

Statistical assessment, including the application of a variety of machine learning tools have previously been applied for process control and monitoring (Toifl *et al.*, 2010; Haimi *et al.*, 2013). However, only a few of them deal with pathogen removal. Additionally, a number of reliability tools exist to analyse systems and interactions between system components in a probabilistic fashion. These tools have been successfully applied in many other areas such as aircraft and nuclear industries (Dhillon & Balbir, 1999; Keller & Modarres, 2005; Liu *et al.*, 2012a). However, only limited studies exist which assess the association between operational parameters, system state conditions and microbial removal in water reuse systems. Limitations related to water reuse applications include limited data and unknown (or poorly characterised) interactions between system variables. Therefore, analysis of reliability for water reuse systems should appropriately deal with these shortcomings.

This study focused on the evaluation of potential tools for informing the risks and reliability status of water treatment systems increasing used for water reuse. The experimental and modelling work presented in this thesis concentrated on the quantitative assessment of performance and the impacts associated with system conditions. Methodology was developed to apply validation and operational monitoring data to constructing and informing the models. In this thesis a variety of terms are used which may be unfamiliar to some water scientists or engineers. To address potential ambiguity, an extensive Glossary is provided (see page 34). The thesis concludes with an assessment of the major findings and recommendations for future work.

1.2 Formal research objectives

The overall aim of this research project was to identify and select potential tools for reliability assessment. This aim would require an evaluation and demonstration of their utility for water reuse applications related to the validation of pathogen removal performance and process reliability estimation. A central assessment criterion was that these tools should be capable of analysing multivariate systems in a probabilistic fashion. Also, they should be able to include expert information and assess uncertainties explicitly. To address these overall objectives, a literature review was first undertaken to identify candidate tools, followed by diverse validation case studies, focused on a range of important water reuse treatment processes. The performances of these processes were assessed using the identified most promising tools. The following specific aims were addressed in the review and experimental and modelling parts of the thesis:

- Identify, evaluate and compare various risk and reliability techniques for water reuse processes.
    - Literature review (Chapter 2)

- Identify and assess operational parameters affecting the reliability of removal of microorganisms and evaluate potential surrogate parameters.
    - During activated sludge treatment (Chapter 4)
    - During ozonation treatment (Chapter 6)
- Analyse the validation data of systems to improve prediction of performance reliability and develop a method applicable to treatment systems with parallel subunits.
    - For ultrafiltration systems (Chapter 5)
- Explore improving incorporation of uncertainty into pathogen removal efficiency estimation and promote reliability assessment.
    - For chlorination (Chapter 7)
    - For ultrafiltration (Chapter 5)
- Quantify the reliability of a water reuse system using expert knowledge.
    - During UV disinfection (Chapter 8)
- Develop and assess a method to analyse reliability in multi-barrier water reuse systems.
    - During advanced water treatment systems (Chapter 9)
- Synthesise the main findings of this research and propose recommendations for future work.
    - Conclusions and future work (Chapter 10)

1.3 Overview of chapters

This thesis is divided into 11 chapters which were selected based on the areas of study covered in the National Validation Framework for Water Recycling (NatVal) project in Australia (Roser *et al.*, 2015; Robillot *et al.*, 2016). This project focused on the validation of individual and multi-barrier systems using data available locally for water reuse treatment systems. These chapters support and illustrate by example how validation and reliability assessment can be performed with the proposed approaches.

Chapter 1 provides an introduction, the background, and objectives of this research and summarises the structure of the thesis.

Chapter 2 contains the main literature review including a comparison of quantitative reliability tools which have been previously employed in water and wastewater treatment for risk and reliability assessment.

Chapter 3 describes the methods used repeatedly in different case studies and in particular explains the features, development and validation of Bayesian networks and Bayesian statistical models.

Chapter 4 reports the development, validation and use of Bayesian networks for the analysis of potential predictors of pathogen removal during activated sludge treatment of domestic wastewater.

Chapter 5 presents and analyses the use of Bayesian hierarchical models in full-scale ultrafiltration systems operating in parallel for validation using virus removal. A particular focus is exploring how to incorporate and combine the different sources of variability in the units' performance and quantify model uncertainty.

Chapter 6 reports the identification and evaluation of parameters that can be used to monitor the pathogen reduction reliability during ozonation of secondary treated wastewater. Evaluation of potential predictors conducted through Bayesian methods is presented.

Chapter 7 describes the development of methods to interpolate between target virus CT values for given turbidity and pH values during chlorination of secondary treated wastewater as well as inform uncertainty in their estimates to probabilistically estimate process reliability.

Chapter 8 assesses reliability of UV disinfection, in particular the probability of undetected low dose hazardous events with a focus on sensor reliability. Model structure and parameters based on expert elicited data are described.

Chapter 9 investigates a way to improve, representing and assessing pathogen removal performance reliability by a multi-barrier advanced water treatment system using published validation data.

Chapter 10 summarises the major outcomes of this thesis and provides recommendations for future research activities.

Chapter 11 presents the list of references used throughout this thesis.

In addition, the main body of the main body of the thesis is supplemented with Appendices covering additional relevant information and data for:

1. Chapter 4 (Appendix 1)
2. Chapter 5 (Appendix 2)
3. Chapter 6 (Appendix 3)
4. Chapter 7 (Appendix 4)

Chapter 2: Literature Review: Tools for assessing reliability and validation of water and wastewater treatment processes.

2.1 Introduction

Reliability is a concept clearly relevant to incorporation within risk management frameworks for water reuse. Reliability can be analysed from two points of view, the intrinsic variability of a process (how much is it expected to vary?) and mechanical reliability (how likely is it to underperform?). A number of tools have been used in the past to evaluate reliability in water treatment systems. However, due to the specific characteristics of these systems, not all these tools may be equally applicable to efficiently analyse reliability. A rigorous and systematic assessment of the available tools for these applications is needed. The aim of this literature review was to identify and evaluate the current uses and development of reliability tools for assessing water treatment systems.

There are a number of important concepts which are used in the following sections of this review and require proper definition. Other important concepts including reliability, safety and validation are defined in Section 2.2 and reproduced in the Glossary.

- **Hazard**: in the context of water treatment refers to a biological, chemical, physical or radiological agent that has the potential to cause harm (WHO, 2011).
- **Hazardous event**: refers to an incident or situation that can lead to the presence of a hazard (WHO, 2011).
- **Risk:** is defined as the effect of uncertainty on objectives, with objectives covering different aspects (e.g. financial, environmental goals) and levels (e.g. strategic, organisation wide) (ISO, 2009).
- **Failure**: represents the departure of an item from its required or intended operation, function, or behaviour (Ericson, 2005).
- **Repair**: refers to the restoration of a failed component, item, subsystem, or system to an operational state (Ericson, 2005).

2.2 Risk, reliability and validation in water management

Risk and reliability are two closely related concepts. Multiple definitions have been put forward to describe risk (e.g. risk as defined in Section 2.1). In the context of water treatment and water management, risk is typically defined in terms of the likelihood and consequences of a hazardous event leading to exposure of hazards (NHMRC & NRMMC, 2011). Rational decision making requires a clear and quantitative way of assessing and expressing risk. Reliability, on the other hand is defined as the probability that a process or item will perform a required function under

specified conditions, without failure, for a specified period of time (Ericson, 2005). In most circumstances these concepts can be considered as complementary. Risk management and reliability engineering are both employed on strategies to reduce risk (Jackson, 2017). Reliability engineering involves quantitative or probabilistic risk assessment. While probabilistic risk assessment is concerned about scenarios, their probability of occurrence and their consequence, reliability engineering is concerned about failure mechanisms, their probability of occurrence and their consequences. Reliability engineering can be seen as a particular subset of the overall risk management process. Safety, another concept closely related to reliability and risk, deals only with dangerous failure modes. Safety and reliability use common methods and either one can inform the other.

The management of risk refers to the practices, processes, systems, resources and culture embedded in an organisation's management that allow risk to be managed (ISO/TR, 2013). Risk management has become the cornerstone of many international water recycling frameworks (NHMRC *et al.*, 2006; WHO, 2006). Various agencies have developed formalised frameworks for the application of risk management to water supply systems. These include the development of Water Safety Plans (WSP) in the World Health Organization guidelines (WHO, 2005) and the Framework for the Management of Drinking Water Quality in the Australian Drinking Water Guidelines (NHMRC & NRMMC, 2011). There is also the Framework for management of recycled water quality and use (NHMRC *et al.*, 2006). Risk management frameworks for water recycling seek to ensure that sufficient treatment is provided through multiple barriers and that risks are managed to acceptable limits. In the case of direct potable reuse increased risk to human health might be observed because of the elevated concentrations of chemicals and pathogens in source water. The principal difference between chemical and pathogen hazards is that potential risks produced by pathogens are more likely to be acute, severe and widespread (NRMMC *et al.*, 2008), with public health effects potentially occurring from as little as a single exposure (Haas & Trussell, 1998). Consequently, the major public health concern during water reuse is pathogens.

When studying the reliability of a water treatment process there are two points of view to be considered: an inherent reliability (or intrinsic variability) and mechanical reliability (Eisenberg *et al.*, 2001). Inherent reliability of wastewater reclamation and reuse can be defined as the probability of adequate performance for a specified period of time under specified conditions, where performance is determined by the ability to meet regulated water quality treatment objectives (Eisenberg *et al.*, 2001; WHO, 2005). A slightly different definition of reliability relates to the ability of the system to meet or exceed the public health requirements (Pecson *et al.*, 2015). A proposed methodology for computing inherent reliability uses the concept of coefficient of reliability (COR) to help designers to set an adequate effluent concentration when defining

plant performance (Niku *et al.*, 1979). For inherent reliability, the general concepts of failure of compliance and inherent reliability can be defined by Equation 2-1and Equation 2-2 respectively.

Failure = effluent concentration > effluent requirements Equation 2-1

Reliability = 1 – P(failure) Equation 2-2

= 1 – P(effluent conc. > requirements)

Designing a process to meet certain standards requires the treatment process to produce an average effluent quality below the threshold limit considering its variability. Therefore, the issue here is to find a mean value (m_x) that ensures that the effluent concentration is constantly meeting the maximum permitted limit with certain probability. To solve this problem the coefficient of reliability (COR) has been developed to probabilistically relate mean design values to the required standard value (X_s) as shown in Equation 2-3 (Niku *et al.*, 1979).

$$m_x = (COR)X_s \quad \text{Equation 2-3}$$

Where:
m_x: Mean effluent concentration
X_s: A fixed standard
COR: Coefficient of reliability

The COR is a function of the coefficient of variation, which in turn is a function of the mean and standard deviation of the collected data. Although this approach was firstly used to assist on the design of activated sludge plants, it could be equally employed to calculate the reliability of an already operating treatment process (Niku *et al.*, 1979). This tool has been employed to assess inherent reliability in diverse types of processes, such as trickling filters, septic tank + anaerobic filter, facultative pond, anaerobic pond + facultative pond, activated sludge, upflow anaerobic sludge blanket (UASB) reactors alone and UASB reactors followed by post-treatment (Niku *et al.*, 1982; Oliveira & von Sperling, 2007; Oliveira & Von Sperling, 2008). This methodology highlights the importance of process performance variability and its importance of reliable process operation.

Mechanical reliability is a function of how system components impact upon plant availability (interruption of supply) and effluent quality. Hardware configuration requirements to improve reliability include standby power supplies, redundant critical equipment, online monitoring,

flexible piping, and emergency storage or disposal options (Drewes & Khan, 2010). A further aspect of reliability that can be considered separately from these two concepts is human reliability. Investigations of numerous hazardous events have shown human reliability to be crucial for the correct functioning and overall reliability of water treatment plants (Hrudey & Hrudey, 2005).

In general, the inherent reliability of water recycling systems is assessed by a process known as "validation", which is a thorough analysis of the performance of the plant under various environmental conditions to assure their correct functioning and achievement of removal targets (WHO, 2005). Identification of influential factors affecting the efficacy of the treatment and identification of operational parameters as indicators of the performance of the system are also key steps during validation (VDoH, 2013). Validation should be conducted when a new system is implemented or new treatment processes are incorporated (VDoH, 2013). However, validation processes do not always utilise explicit quantitative tools for reliability analysis. This lack of quantitative information limits the ability to undertake any predictive or diagnosis assessment of underperformance.

Reliability assessment and validation of processes for water recycling requires appropriate tools to identify the association between factors which impact upon multi-barrier performance considering multiple sources of information and uncertainty. Such tools should also allow incorporation of various reliability concepts, including inherent and mechanical reliability.

Both qualitative and quantitative tools are fundamental for system reliability evaluation. Understanding the system and detailed assessment of the interactions between components and potential impacts need to be assessed qualitatively. A qualitative analysis is therefore needed as a precursor step for identifying the issues to be analysed and determining whether a detailed analysis is worth doing. Once that determination has been made, there may be a potentially very important role for quantitative risk analysis tools.

2.3 Necessity of improvement of quantitative risk analysis tools and characteristics of an ideal tool to inform reliability

Qualitative and semi-quantitative risk assessment methods such as risk matrices (i.e. heat maps) and risk registers are widely used in the water industry. Major strengths include their simplicity of construction and communication. Their aims are to compile and identify the most significant risks needing control. They are a valuable tool for risk management, especially during the design and operation phases. However, the use of these methods for describing complex incorporating interactions between system components is not adequate (Burgman, 2005). These events might include system component failures, human error or rare external events (e.g. flooding, terrorist

attack, chemical and biological spills). The effect on treated water by some of these events may be difficult to quantify intuitively. An analysis of multiple scenarios is therefore necessary to assess the effect on water quality and risks to consumers.

Hazard Analysis and Critical Control Points (HACCP) is another tool that is often implemented in the water industry when a risk management approach is used (e.g. Dimmock (2000); Mullenger *et al.* (2002); Quinn and Marriott (2002); Hellier (2003); Howard (2003); Jagals and Jagals (2004); WHO (2005); Damikouka *et al.* (2007)). HACCP focuses on critical control points to control risks through system monitoring and corrective action (Havelaar, 1994; Quinn & Marriott, 2002). This method relies heavily on monitoring, design of processes, training and correct procedures among others for correct implementation. HACCP begins by a qualitative assessment of the system, but once developed, it requires quantitative analysis of the collected monitoring data. These data are used to control the process and identify deviations of the critical control points from normal operational boundaries. Statistical process control, soft sensors, early monitoring and reliability assessment are all tools that can be considered as part of the HACCP process. The analysis of process monitoring parameters that correlate with the process performance is not necessarily straightforward, and requires the use of appropriate tools to inform their relationships. Therefore, this task relies upon other approaches such as complex machine learning and statistical methodologies.

When risk assessment requires a broad consideration of factors affecting performance and quality of product water, various aspects of reliability can be considered when implementing a risk analysis, including mechanical, human, and software aspects. Modelling and process analysis are based on quantitative monitoring, uncertainty and probabilities. Evaluation of risks requires tools which aid the assessment of the associations between process controlling factors and performance. The nature and presence of hazards means that process performance is dynamic, and constant conditions are rarely achieved. Thus, we may only assert the *probability* of a system being in a specific state or condition given the observed parameters. A quantitative study permits prediction of desired or non-desired outcomes, and the importance of various factors or variables involved. Quantitative microbial risk assessment (QMRA) and more generally health risk assessment are also commonly applied and promoted by water quality management guidelines (WHO, 2005; NHMRC *et al.*, 2006) to quantify risks and derive performance requirements of the treatment system.

Consequently, it is concluded that the development of better tools for informing on and analysing risks more effectively should be vital for the water industry to improve safety and resilience to the point where direct potable recycling of water can be done with confidence. However, current guidelines and practices do not yet identify or promote a comprehensive range of tools suited to

quantitative reliability assessment. There are a number of characteristics which are likely to be essential for an ideal quantitative tool to be applied for water treatment reliability assessment (Table 2-1). This tool would also need to be applicable despite inherent limitations, such as data availability limitations of the water industry. The criteria used in Table 2-1 were derived from the characteristics of already applied tools in the water treatment sector and general desirable features. Quantitative outcomes and probabilistic assessment are both common features of QMRA and health risk assessment in general (WHO, 2004; NHMRC *et al.*, 2006). Various sources of data is a desired characteristic because data are usually scarce and/or expensive to be obtained (Pollard, 2008; Hokstad *et al.*, 2009). Capability to deal with limited information refers to the ability of incorporating missing or censored values. These types of cases are common in water treatment and should be considered within any model development (Khan, 2010). Models should be simple to develop and communicate to be effective tools for risk management (Fenton & Neil, 2012). In the same line, representing associations between events or system components are important for an effective tool (Fenton & Neil, 2012).

Table 2-1: Characteristics of an ideal tool for reliability analysis

Characteristic	**Practical Description**
Quantitative outcome	Numerical results
Probabilistic assessment	Results that describe the likelihood of various outcomes.
Various sources of data	Ability to incorporate quantitative data and more subjective expert knowledge
Capability to deal with limited information	Ability to manage missing values in dataset, or missing evidence. This also applies to censored values.
Facility of development	Ease of development would encourage its use and improvement

Characteristic	Practical Description
Representation of associations	Causal representation is ideal if the interactions and impacts from variables are to be understood

2.4 ISO 31010 and reliability standards

There are many types of reliability analysis tools. For example Ericson (2005) identified over 100 different reliability analysis techniques in existence. However, only a number of them have been considered and reviewed in current international standards (IEC/ISO, 2009). The limited number of tools reviewed by international standard is because some techniques are rarely practiced and many consist of minor variations of other techniques. To focus on the most relevant and widely employed techniques, the scope of this review considered the techniques identified by ISO-IEC 31010-2009 Standard: Risk Management - Risk Assessment Techniques. This ISO standard is a supporting standard for ISO 31000 and provides guidance on selection and application of systematic techniques for risk assessment (IEC/ISO, 2009). Although the scope of the standard is for risk assessment in general, these tools have widely been used in system reliability analysis as is evidenced in well-known reliability assessment literature (Rausand & Høyland, 2004). Furthermore, the scope of this study was refined to only those techniques providing quantitative outcomes. From the refined list, reliability block diagrams were added. Reliability block diagrams were incorporated because they are commonly used in reliability assessment (Pollard, 2008). A discussion about machine learning tools applied in reliability assessment is also provided. This chapter includes a systematic analysis of the main developments and uses of these tools by the water industry and offers recommendations regarding the most adequate tools for validation of water treatment processes.

Most of the tools evaluated in the following section have also been recommended as risk assessment and management tools for water supply systems within the TECHNEAU project in the European Union (Hokstad *et al.*, 2009). The main focus on that project was risk analysis. However, some aspects of reliability analysis were also discussed. The outcomes from this project revealed that a single tool may not be adequate for the whole risk assessment process, and that application of a method will depend on the problem in question. Also, a single risk management method could not be recommended for all water utilities (Rosén *et al.*, 2007).

2.5 Reliability/Validation and related risk assessment and management tools

A selection of risk analysis tools mentioned in Section 2.4 were described and evaluated considering their advantages, limitations and their previous applications in water treatment processes.

2.5.1 Human reliability assessment

In real systems both hardware reliability and human reliability contribute to the overall system reliability. Human reliability can be defined as the probability that a task executed by personnel within a system's operation is successfully achieved within a specified period of time (Pollard, 2008). Human reliability analysis (HRA) can provide qualitative or quantitative outputs which are then used in probabilistic safety analyses or probabilistic risk assessment (Hollnagel, 2000). The former is used to identify potential erroneous human actions and their causes to reduce the probability of occurrence. Quantitatively, HRA provides data input regarding human failures for other reliability tools including fault tree analysis, event tree analysis, reliability block diagrams and Markov chain analysis (Dhillon & Balbir, 1999; IEC/ISO, 2009).

A common HRA modelling method comprises three stages:

1. the identification of human errors,
2. the prediction of their likelihood, and
3. the reduction of their likelihood, if required.

HRA techniques are usually classified into two generations. The first concerns about the safety assessment of plants, whereas the second applies cognition assessment. First generation tools comprise tools known by their acronyms as THERP (Swain & Guttmann, 1983), HEART (Williams, 1986), SLIM (Embrey, 1984), ASEP (Swain, 1987), TESEO (Bello & Colombari, 1980) and HCR (Hannaman *et al.*, 1984). The second generation tools comprise ATHEANA (Cooper *et al.*, 1996), CREAM (Hollnagel, 1998), or MERMOS (Bieder *et al.*, 1998).

Mathematical modelling of HRA has a number of recognised shortcomings. HRA treats individuals as components in a chain with an associated reliability. However, the models produced cannot predict human behaviour though they may attempt to characterise typical failure rates. Good HRA models incorporate the factors that allow for the operating conditions that may result in an error. These models still have shortcomings however in that they: (a) tend to focus on individual, rather than collective behaviours and the social interactions that inform them; (b) have yet to be fully validated; (c) tend to focus on equipment operators rather than individuals,

organisational policies or company rules higher in the organisational hierarchy (Pollard, 2008). Each specific technique to evaluate human reliability has strengths and limitations. An analysis of these tools and their advantages and disadvantages has previously been reported by Spurgin (2010). An analysis of HRA methods and their application to real accidents in various industries can also be found in Spurgin (2010).

In quantitative HRA, access to relevant human reliability data can be problematic. If specific data are not acquired for the water treatment system, the analyses may be based on generic human reliability data from other industries (Hokstad *et al.*, 2009). Reason (1990) has outlined four levels of human failure, each influencing the next (the Swiss cheese model):

1. organizational influences,
2. unsafe supervision,
3. preconditions for unsafe acts and
4. unsafe acts.

The role of third parties has been proposed in water treatment schemes (Wu *et al.*, 2009), therefore a fifth level can be added:

5. errors of consumers and third parties.

Note that levels 1-3 are considered latent errors (e.g. inadequate equipment, poor design, inadequate supervision, manufacturing defects, maintenance failures, etc.); whereas Level 4, unsafe acts, is regarded as active errors (i.e. human errors). For active errors, consequences can develop within a short period time. These errors can include omission or using the wrong rule and they are likely to be provoked by front-line operators. Consequences of latent errors might begin to appear after a period of time, or when other errors, or specific operational conditions lead to the failure to occur. Level 5, "errors of consumers and third parties", may include both latent errors (pre-existing risk scenarios) and active errors (e.g. an unsafe act by a third party).

Given the high responsibility of process operators, the nuclear energy industry was the first to develop and apply HRA (Kirwan, 1994). Other industry sectors which have now developed their own HRA include aviation and aerospace, rail, air traffic control, automotive, offshore oil and gas, chemical, and parts of the military (Kletz, 2001; Lyons *et al.*, 2004). A review of the role of human error in maintenance across various industries (1981–2003) has been provided by Dhillon and Liu (2006). Limited information is available on the use of HRA to drinking water safety (Wu *et al.*, 2009).

Human errors have been recognised to be among the major contributing factors to drinking water contamination incidents (Hrudey & Hrudey, 2005). However, it has been argued that it has not been sufficiently studied by the water quality management community (Tang *et al.*, 2013). Wu *et al.* (2009) analysed 61 drinking water incident cases and categorised them into various types of human errors according to their circumstances and possible causes. A key outcome of this study was the conclusion that water quality incidents could comprise six "periods", which they described as contamination, sensing, warning, recognition, inspection and recovery. Identification of common human errors could assist in the improvement of countermeasures. Later work conducted by Tang *et al.* (2013) complemented their earlier research with an analysis of 40 drinking water incident cases and human errors in the initiation of each drinking water incident and the distribution of the human errors in each period. They also suggested improving resilience through different approaches such as active consumer complaints channels, cultivating a safety culture, collaboration and the use of preventive risk management tools. Cloete *et al.* (2013) investigated human-machine interface issues in water treatment systems using checklists, semi-structured interviews, questionnaires, and observations of routine work of operators. Lack of human integration in system design and compliance with only 13 of 40 items selected for best practices were found in this study. It was also concluded that operator expertise was underutilised because of a lack of integration and user-centred design. The issue was reflected, for example, in the high allocation of work-hours to routine tasks such as data logging.

2.5.2 Health risk assessment and quantitative microbial risk assessment

Environmental and human health risk assessments (HHRA) are formalised mathematical procedures used to evaluate potential hazards introduced by pollutant emissions to human health and the environment. Formally, this risk assessment process comprises a sequence of well-established steps (NRC, 1983; enHealth, 2012): (1) Hazard identification, (2) Exposure assessment, (3) Dose-response assessment, (4) Risk characterization. The US National Research Council publication, known as the "Red book", represents the seminal source for this framework from which other adaptations have been based (NRC, 1983). Although the NRC document presents a holistic framework, over the last few decades, the approach has been refined by various agencies. The Australian health agency, enHealth has proposed a model with five stages as observed in Figure 2-1 (enHealth, 2012).

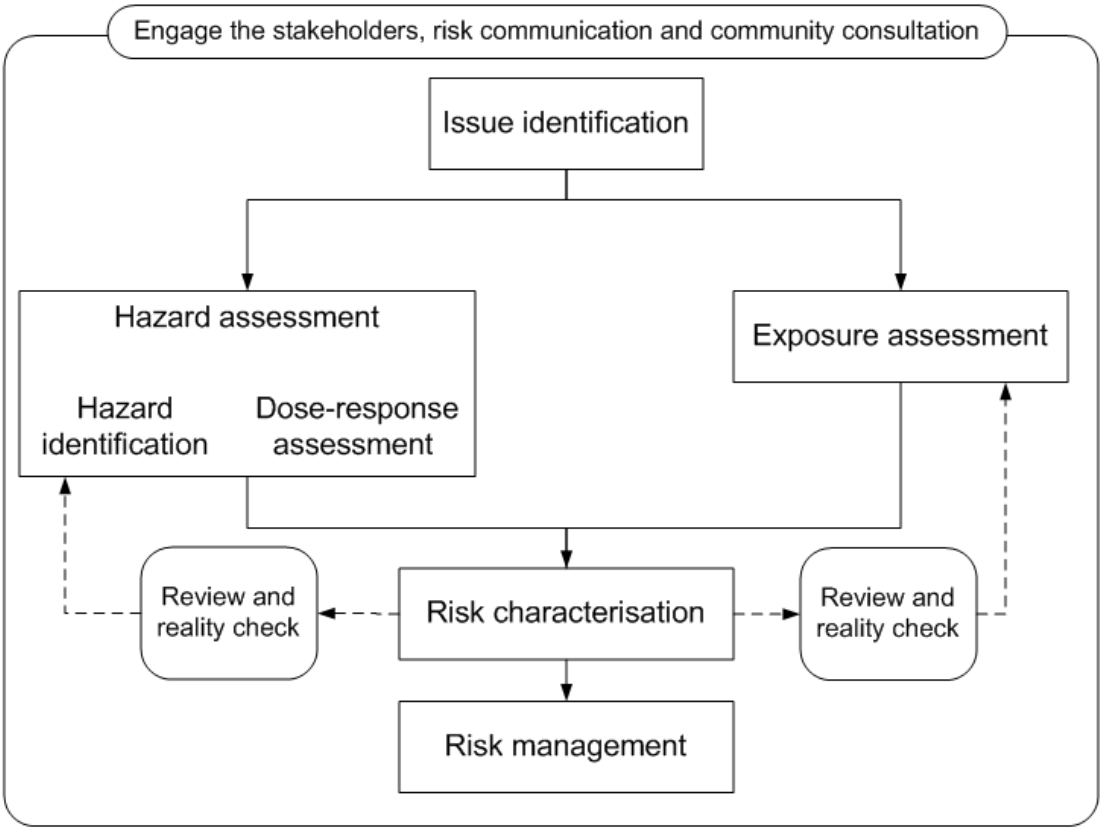

Figure 2-1: Environmental health risk assessment model (enHealth, 2012).

Conventional drinking water guidelines values, such as those published in the Australian Drinking Water Guidelines (NHMRC & NRMMC, 2011) and WHO Guidelines for Drinking-Water Quality (WHO, 2011), have been generally derived based on the environmental health risk assessment concepts defined by the NRC (1983). Derivation of health-based target values for water quality takes into consideration both the exposure (for water consumption is generally assumed as 2 L/person/day) and dose response data. Incorporation of uncertainty factors (UFs) is also necessary for substances in which a threshold effect exists. UFs are used to account for inter-species differences, use of LOAEL (Lowest observed adverse effect level) instead of NOAEL (No observed adverse effect level), intra-species differences, the severity of the adverse effect and quantity and quality of the scientific data (enHealth, 2012).

Derivation of performance targets requires data regarding the source water quality to derive the necessary removal efficiency of the system. The most common application of performance targets is for control of microbial hazards (WHO, 2011).

The application of Disability adjusted life years (DALYs) for defining levels of risk associated with exposure to pathogens was introduced in drinking water guidelines by WHO in the 2004 revision of the Guidelines for Managing Drinking Water Quality (WHO, 2004). For this purpose,

the WHO established a tolerable risk level of 10^{-6} DALYs per person per year, commonly referred to as 1 microDALY (1 μDALY). An important advantage of quantifying health risks using DALYs is that they include a measurement of the severity of the different impacts on human health from the infection and illness. The use of DALYs for pathogen risks was subsequently implemented in the Australian Guidelines for Water Recycling (NHMRC *et al.*, 2006). Discussions regarding the implementation of DALYS in the ADWG are currently ongoing (WSAA, 2015).

Setpoints and critical limits in the Risk Management Framework can be specified through HHRA including safety factors. In consequence, the system will comply with the established health-based targets (without excessive economic implications) even when hazardous events have affected the system. Thus, this technique provides the minimum performance requirements to reduce chemical contaminants or pathogens to below safe limits. In addition, it allows an assessment of the relevance of the different risks and indirectly of the barriers associated to each reduction. Thereby, consequences of hazardous events or system failure can be evaluated in terms of the reduced performance, the change in effluent quality and exposures that may be the consequence of the occurrence of hazardous events (Khan, 2010). This analysis determines the required inherent reliability of the treatment processes and assists in the identification of measures to improve performance. The outcomes may then inform where resources could most effectively be allocated.

The NRC framework approach can be applied in either a deterministic or stochastic manner. Commonly, when a stochastic analysis is conducted, input values and model parameters are considered as probability distributions. The analysis is then typically performed by the aid of Monte Carlo simulation (Section 2.5.11) (Khan, 2010). The outcomes are then obtained in the form of probability distributions instead of point estimates. Incorporating uncertainty and variability within the calculation of risks have the advantage of providing more information. For example, by estimating the likely values a model input or output can take, representing uncertainties, and calculating probabilities of exceeding particular limits. Quantitative microbial risk assessment (QMRA) has been particularly concerned about stochastic analysis for risk assessment (Soller *et al.*, 1999; Teunis & Havelaar, 1999; Haas *et al.*, 2014).

2.5.3 Fault tree analysis

Fault tree analysis (FTA) is a tool used to analyse causal representations of system failures and is based on three main assumptions: (i) events are dichotomous, (ii) events are statistically independent, and (iii) logical Boolean gates (e.g. AND, OR, NOT and Voting) are used to represent the relationship between events. These assumptions also introduce some limitations to the method such as difficulties in modelling complex systems (e.g. sequential events, standby

configurations, redundancy, etc.), partial failure states of components cannot be modelled, and independency between events is not always a valid assumption (Dhillon & Balbir, 1999; Ericson, 2005; Baig *et al.*, 2013). FTA provides a quantitative outcome as well as a qualitative analysis through graphical representation (Figure 2-2) of the different events leading to a target event (Top event on Figure 2-2) (Ericson, 2011). The use of logical gates allows the evaluation of minimal cut sets which is the combination of failure events that can cause the top event to occur (Dhillon, 2004). This step is useful as probabilities are not required for the analysis. One of the most important features of FTA is the identification of combinations of single failures, which by their own do not represent a threat, but together cause the system to fail (e.g. rare event).

In general, different combinations of component failures or events such as human errors, normal events, and environmental factors may cause a system malfunction. The identification of the events leading to a failure provides the basis for the implementation of safety barriers. It is important to add that the type of events in FTA can include not only failures in a particular component but also the occurrence of a breach regarding a parameter in the system. For example, it can be considered an event when the turbidity exceeds a certain limit.

FTA has been used to analyse non-compliance of water quality, mechanical reliability, water supply failures, and conceptual modelling of failures. An important contribution to this topic was provided by Lindhe *et al.* (2009). They created a FTA model to estimate the risk as customer minutes lost, of quality and quantity failures in water production systems. In the study, dynamic gates were created to incorporate a scenario when the system recovers by itself. Failure and repair rates were elicited from experts as probability distributions and they were allowed updating through Bayesian methods (see Section 2.5.12).

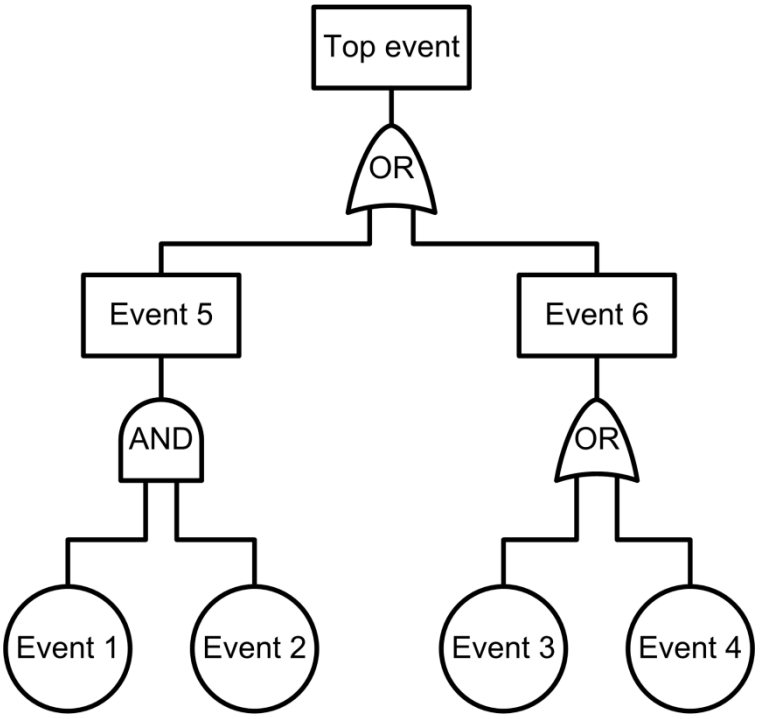

Figure 2-2: Classic representation of a FTA.

FTA has been used to study the availability of full-scale reverse osmosis desalination plants using real operational data (Kutbi *et al.*, 1982; Hajeeh & Chaudhuri, 2000; Bourouni, 2013). A qualitative FTA was employed to analyse the operational failure and *Cryptosporidium parvum* oocysts presence in an UF plant which was constructed through an operators elicitation process (Beauchamp *et al.*, 2010). This tool was also used to analyse the causes required for an outbreak including source of contamination, failure of treatment or distribution system, and inadequate detection or response to the drinking water contamination (Risebro *et al.*, 2007). In a similar fashion, but including supply failures a qualitative FTA was developed for failure scenarios in water production considering quantity and quality failures (Tchórzewska-Cieślak *et al.*, 2011). This tool was also employed to analyse the main factors that affect the not compliance with certain parameter (e.g. BOD, turbidity) to ease the detection of deficiencies (Kelley & Allison, 1979; Kelley & Allison, 1981; Ravikrishna & Mhaisalkar, 1996). Similarly, the probability of non-compliance based on BOD for a wastewater treatment plant has been assessed through a FTA. The basic events probabilities were estimated through operator's consultation (Taheriyoun & Moradinejad, 2015).

2.5.4 Event tree analysis

Event Tree analysis (ETA) is a graphical technique (Figure 2-3) whose purposes are to assess the possible outcomes that can escalate from a particular initiating event (IE), and identify scenarios and their frequencies. It uses an inductive logic and can be applied both qualitatively and

quantitatively (Stapelberg, 2008; IEC/ISO, 2009). The principal component in an ETA is the pivotal event which implies a question/condition with YES/NO (success/failure) outcomes (Stapelberg, 2008). The various possible alternatives are connected by paths, either to other pivotal events or to outcomes (Smith, 2011). The logical sequence can also be expressed with Boolean algebra. The pivotal events can represent barriers which can reduce the consequences of the initiating event. ETA can be used to model a whole system or process, with analysis coverage given to subsystems, components, software, procedures, environment, and human errors (Ericson, 2005). This tool is useful for assessing facilities having engineered accident-mitigating characteristics to identify the events that follow the initiating event and generate given sequences.

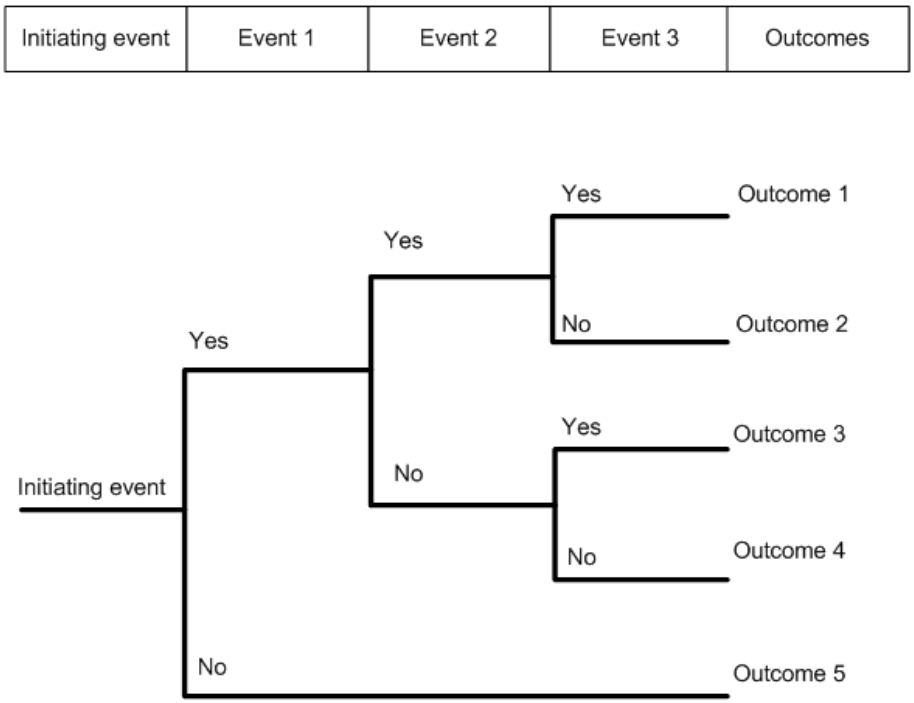

Figure 2-3: Classic event tree representation.

A limitation in basic ETA is the inability to model partial failures. However, this issue can be solved by splitting the tree on more than two branches as within a decision tree. In these cases, the Boolean representation would not be possible (Modarres *et al.*, 1999). It can also overlook subtle system dependencies when modelling the events and it would require multiple ETA when having multiple initiating events. ETA provides a natural language and representation for barriers and the probability of consequence events in case the barriers fail or not. Studies of usage of ETA in water utilities have been scarce. Gale (2002); Gale (2003) used ETA to illustrate the importance of bypasses or minor routes on risk events with high concentrations of microorganisms. ETA was used to model the routes of exposure and which pathways have the highest impacts. Although this

tool may appear different to FTA in the types of cases it can model, the same problems can be modelled using both techniques.

2.5.5 Reliability block diagram

Reliability block diagrams (RBD) are abstract system representations useful for both reliability prediction and maintenance optimization (Barlow, 1998). They depict the dependence (Figure 2-4) of the various elements (pieces of equipment, components or functions of the plants) that contribute to the reliability requirements. Through a set of rules, RBD provides information about the impacts of the component failures on the satisfactory system operation (Ireson *et al.*, 1996). RBD indicates the logical connections of working components required during a particular system function. When the system performs more than one function, separate analyses should be conducted individually for each function. RBDs are suitable for systems with non-repairable components and where the order sequence in which failures occur does not matter. For these cases, Markov methods (see Section 2.5.10) are commonly used (Rausand & Høyland, 2004).

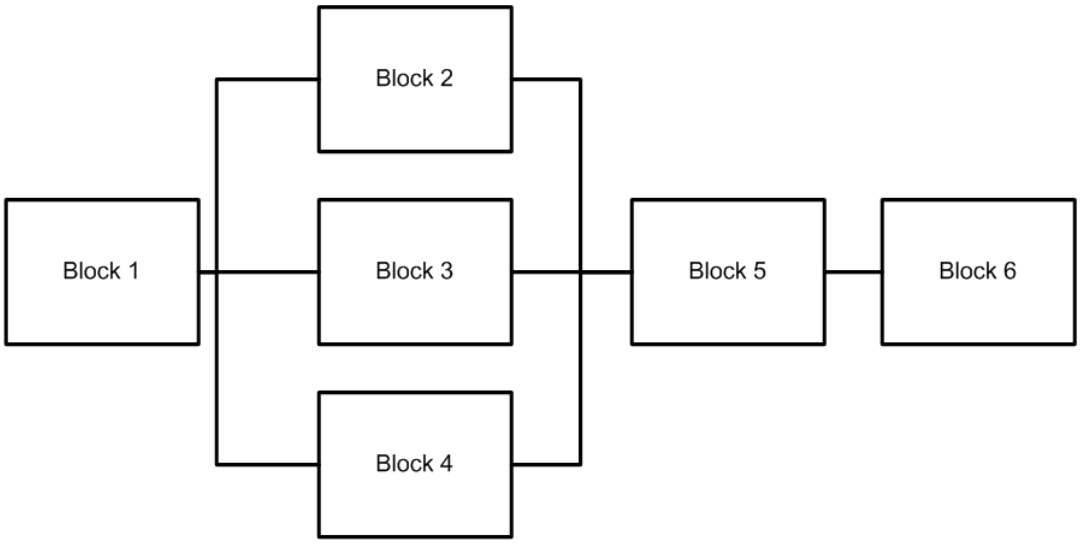

Figure 2-4: Reliability block diagram representation.

Similar to FTA, RBD considers events as independent and only two states are allowed for each component. In addition, minimal cut sets can be defined in RBD. Unlike FTA, this method does not identify causes of failure explicitly through its structure.

RBD has been assessed as a valuable technique for mechanical reliability assessment in water treatment utilities (Hokstad *et al.*, 2009). It has also been shown that it can be used in the process of understanding the system functioning and analyse different configurations. For example, Cyna

(1997) used this methodology to represent a filtration and chlorination system and assess its unavailability. Modifications of the system might be performed to obtain a desirable unavailability. For example, Sydney Water created procedures for RBDs (SydneyWater, 2010b). They have applied these in complex works including wastewater pumping stations, chemical dosing units and non-gravity wastewater systems (SydneyWater, 2014). These procedures should be followed during the conceptual design and planning stage to develop contingency plans and manage risks. During the detailed design stage, they are used to identify monitoring requirements for assessed risks, identify spares to be kept by Sydney Water and develop and provide maintenance plans (SydneyWater, 2011).

Bourouni (2013) assessed the availability of a reverse osmosis desalination plant through RBD and FTA techniques. He evaluated the complete system including pre-treatment, beachwell pumps, RO system and Post-treatment system. Four-year data for model validation and testing were taken from a RO plant in Kuwait. The results showed that the overall plant unavailability was very low and the components having higher unavailability in the plant were the RO modules and high-pressure pumps. To determine the components that influence most the overall plant availability, a sensitivity analysis was performed and the different components were classified in groups according to their influence. This study demonstrated the importance of RBD in assessing availability and their usefulness when considering standby components and r/n redundancies.

A recent study in Australia was conducted to analyse the reliability of advanced water treatment plants from data collected in seven plants around the world (Tng *et al.*, 2015). The aim of that study was to assess the reliability of a typical advanced water treatment system and identify the most important components affecting quantity and quality related failures based on a 10-year period. The analyses were performed using RBDs and Monte Carlo simulations on a proprietary software package. The results indicated that quantity failures were eight times more likely than quality failures. Moreover, from the 427 simulated failures in 10 years, only 5% were related to quality failures. Other scenarios were tested to evaluate maintenance strategies. For example, by reducing the maintenance response time, the number of quality failures per year can be reduced to less than one per year.

2.5.6 Bow-tie analysis

In the context of reliability analysis, a bow-tie is a logical structured technique that is employed to describe and analyse the associations between the causes of an undesired event, the progress of such events, the preventive controls for those events and mitigative controls for the potential consequences (Book, 2012; Khakzad *et al.*, 2013a). Bow-tie analysis focuses specifically on the barriers in place (Figure 2-5). Some authors have considered bow-tie as effectively the same as

cause and consequence analysis in which the causes are investigated through FTA and the consequences are investigated through ETA (De Dianous & Fiévez, 2006; Ferdous *et al.*, 2013; Khakzad *et al.*, 2013a) (Figure 2-6). Bow-ties provide a useful tool to represent the links between risk controls and the management system (Book, 2012).

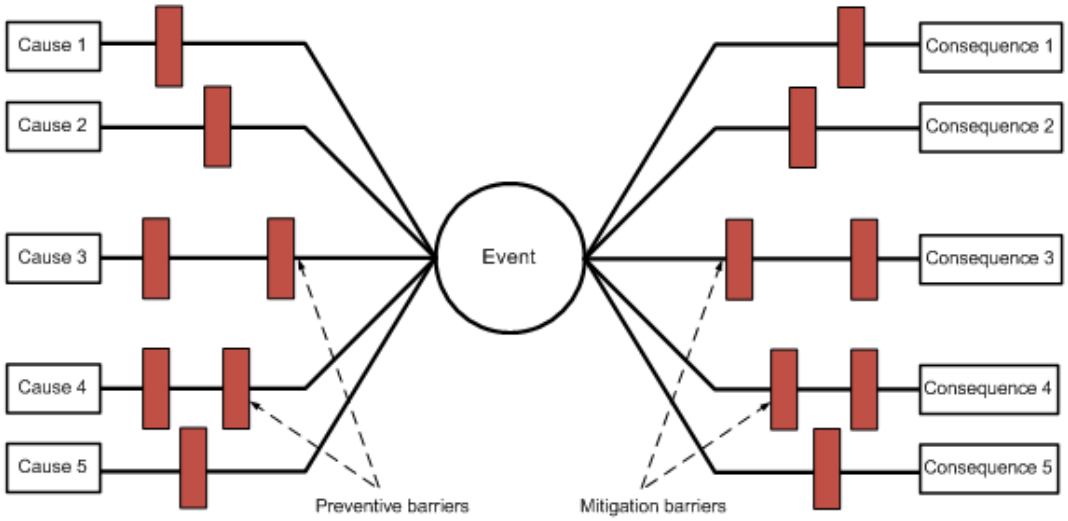

Figure 2-5: Classic bow-tie representation.

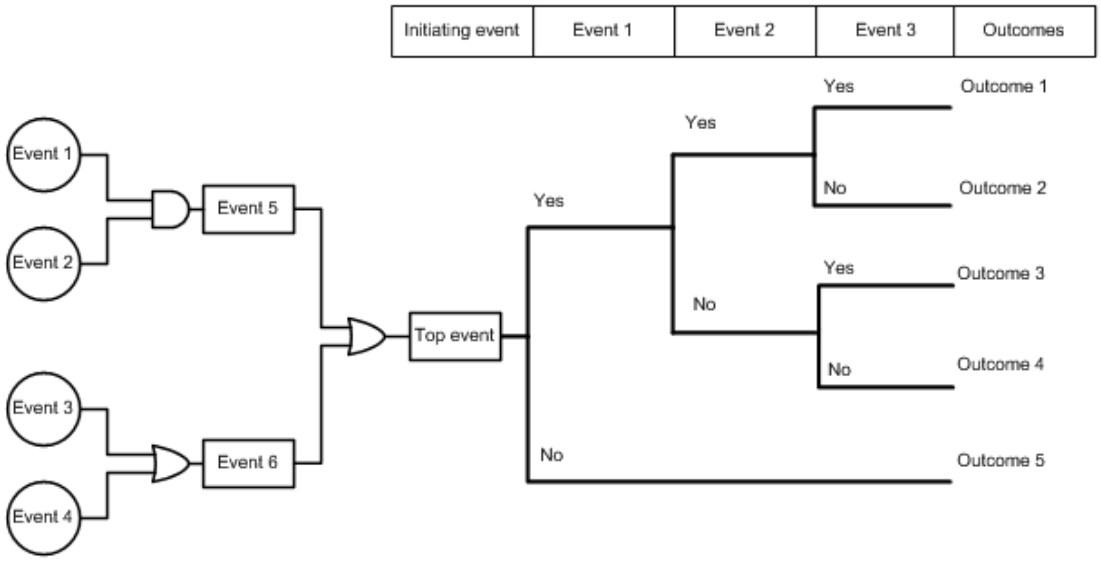

Figure 2-6: Bow-tie representation through FTA and ETA combined

The bow-tie method has been especially designed to show the influence of safety systems (e.g. technical or organisational) on the development of accident scenarios (De Dianous & Fiévez, 2006). Prevention safety systems (to decrease frequency of an event) are found on the fault tree side, while mitigation systems (to decrease frequencies and/or consequences of dangerous phenomena) are found on the event tree side. Visualisation of what safety function (barrier function) applies on a particular scenario is facilitated through bow-tie analysis. The most critical causes or scenarios can be assessed through the use of frequencies of occurrences (De Dianous & Fiévez, 2006). Bow tie analysis can be used for conducting Layers of Protection analysis (See Section 2.5.7). This tool can be used for improving communication with stakeholders regarding risks and measures of prevention and mitigation (Markert *et al.*, 2009). It also helps to target and allocate resources to critical or more relevant barriers.

A safety function ("What is needed to increase safety") represents a technical or organisational action, and not a tangible or a physical system. This action is realised through safety barriers ("how to implement safety functions"). The safety barriers can be either tangible (e.g. engineered systems) or abstract barriers (e.g. human actions based on specific procedures or administrative controls) which are sometimes interchangeable and/or work together (De Dianous & Fiévez, 2006). A related concept is the "barrier system", which is a system that performs one or more barrier functions. A barrier system depicts how a barrier function is performed. Provided that a barrier system is functioning, the barrier function will be as well. Four main categories of safety barriers can be defined which are useful for the evaluation of safety barrier management (De Dianous & Fiévez, 2006; Guldenmund *et al.*, 2006): *Passive barriers, Activated barriers*, *Human actions* and *Symbolic barriers.* In the water industry, the use of this tool has been proposed by a number of authors (Rosén *et al.*, 2007; Blackmore *et al.*, 2008; Maxwell & Franssen, 2012), without the analysis of a real case scenario.

Bow-tie has been applied during the risk assessment process for drinking water treatment systems from groundwater sources in NSW (Contos, 2014). The tool was applied to improve the system in place by identifying three most important hazardous events and their corresponding preventive measures. These events included the inability to maintain chlorine residuals, parameters above the guidelines, and ineffective chlorine operation. The bow-ties were constructed during a workshop through the participation of stakeholders.

Bow-ties have been presented as a valuable tool to represent multiple barriers on water supply systems (Burlingame & Chalker, 2017). The aim of that article was to present this method to improve current practices and clearly visualise the barriers associated to preventive and mitigative controls. They stressed that barriers are not necessarily tangible factors, but also management related.

A modified bow-tie was used to represent the interactions between various risks within water supply systems (Fitzgerald *et al.*, 2017). The authors named this approach "risk pathway method". The developed method was semi-quantitative with risks evaluated through labels high, medium or low. The tool permitted clear visualisation of the interactions between causes, consequences and controls in place.

2.5.7 Layers of protection analysis

Layers of protection analysis (LOPA) is a semi-quantitative tool used to identify barriers that comply with the independent protection layer (IPL) criteria (Summers, 2003). Layers serve the purpose of reducing the risk to an acceptable level. An IPL can be defined as a device, system, or action capable of preventing an initiating causal event to its undesired consequence (Crowl, 2001). These IPLs reduce the frequency and/or the consequence of specific hazardous events. It is relevant to note that LOPA focuses one scenario at a time. LOPA and bow tie analysis present a number of important differences. For instance, they use different criteria for establishing the safety barriers and layers of protection and unlike LOPA, bow tie can incorporate more than one cause and consequence in the analysis.

IPLs can be active or passive systems and they must comply with four different criteria: (1) *specificity*, (2) *independence*, (3) *dependability* and (4) *auditability* (Dowell Iii, 1998; Summers, 2003). The independence criterion has the IPLs be independent of the causal event or any other layer of protection within the analysed scenario. LOPA is used to examine scenarios, often generated by other process hazard analysis tools, such as HAZOP, what-if, checklist or FMEA; as part of the Safety Instrumented Functions (SIF) design; or as part of a design study on a system to classify the various process alternatives and to select the best method (Crowl, 2001). Each scenario can comprise several "cause-consequence" pairs (Gowland, 2006).

Bradshaw *et al.* (2013) investigated organizational and technical barriers (IPLs) and their effectiveness to inhibit incident propagation from over 400 incidents in a large water utility between 1997 and 2006. The identified IPLs contributed on reducing the public health impact of incidents. They observed that the utility effectively improved its organizational and incident management structure to handle customer impacts. A number of organisational attributes were considered IPLs, such as strong organisational culture, staff competencies, and human redundancy among others (Bradshaw *et al.*, 2013).

Current risk management approaches for drinking water and water reuse are largely based on concepts consistent with the application of a LOPA method. Concepts such as independent multiple barriers, operational procedures, staff competence and communication and their impact

over risk reduction are already included in current frameworks of water management. Nevertheless, an analysis of the layers requires meeting the four criteria described above to ensure that they will act as independent protection layers. Defining the layers correctly will prevent an unsafe scenario from developing, even when the initiating event has occurred or another protection layer has failed. LOPA is generally only applicable in specific cases when there are single-cause hazardous events. In cases where multiple causes are relevant, bow tie analysis can be a more appropriate approach. The assessment might also include frequencies of initiating events and probabilities of failure for layers of protection which allow determining the event impact in case IPLs are not successful. However, these may be difficult to determine for a water treatment system. Identification of system failures, consequences and scenarios can be done for instance through hazard identification techniques.

2.5.8 Failure modes and effects (and criticality) analysis

Failure modes and effects analysis (FMEA) is a tool used for the collection and documentation of failure modes, mechanisms of failure and causes, preventive and mitigating mechanisms of failure and related failure effects for product or process functions (Stamatis, 2003). It is a qualitative inductive method, does not require any statistical knowledge and its application is straightforward. FMEA compiles a list of the expected failure modes during the system functioning, elements involved, mode of operation, operation specification, time constraints and the environment (Khan & Abbasi, 1998). Failure modes effects and criticality analysis (FMECA) complements FMEA to provide a failure ranking following their importance or criticality (IEC/ISO, 2009). The criticality analysis can be semi-quantitative or quantitative using actual failure rates.

FMEA has been demonstrated to be a relevant risk assessment tool for the identification, analysis and treatment of potential failures in different systems. Its applicability in improving reliability is well recognized and it has been reported in numerous publications (Liu *et al.*, 2012a). However, the fact that it neither quantifies reliability nor the interaction between failure modes limits its use, requiring criticality analysis to quantify risk or a fault tree analysis to account for multiple failure modes. Usually water or wastewater treatment trains consists of several processes with widely different characteristics, such as operating procedures, operating parameters, chemical requirements, mechanical and electrical needs among others which makes FMEA or FMECA suitable for this purpose.

Applications of FMECA have been found in the water industry. At Sydney Water, for instance, FMECA is used in their facility assets to identify critical assets and maintainable units to develop the most cost-effective maintenance (SydneyWater, 2010a). At Scottish Water, operational level

FMECA studies have been applied (Lifton & Smeaton, 2003). In this application, various components of the water supply system and their failure modes were systematically studied. Dominguez-Chicas and Scrimshaw (2010) described the implementation of water safety plans for a pilot-scale indirect potable reuse (IPR) project following an FMEA approach. In this case, hazards, potential hazardous events and failure modes were described as part of the risk analysis for the IPR system. Such hazards and events can impact the system by challenging the treatment processes or generating operational malfunctions along the water supply chain. Failure modes in this case were analysed considering failures only in the sewage catchment and then classifying their effects into three groups. Six indicators related to the failure modes were presented which could be used for process control. A risk matrix with severity and occurrence was constructed to rank and prioritise the hazards instead of the risk priority number employed in FMEA. Démotier *et al.* (2002); (2003) presented a methodology for assessing the risk of producing non-compliant drinking water (i.e. one of the quality parameters exceeding the maximum values established by water quality standards), considering raw water conditions (i.e. concentrations of water quality parameters) and the process performance of the treatment plant (e.g. technology, different failure modes and corresponding failure rates). The approach consisted of five steps including the use of FMECA for assessing the failure modes, their unavailability and their effect on non-compliant drinking water, and FTA to calculate the probability of non-compliance using the probability of the basic events (i.e. unavailability of treatment steps and probability for the resource to exceed the different thresholds). FMECA was constructed from expert knowledge to determine for each failure mode: the causes, detection means, time to detect the failure, effect on quality and degraded transfer function (i.e. one minus the ratio between output and input concentrations). FMEA has also been used to evaluate the operational risks of retrofitting a water treatment plant with UV disinfection (Passantino & Owen, 2005). Information drawn from the FMEA was used along with engineering-related information to answer a number of questions about implementation issues (e.g. “What are the implementation issues associated with UV disinfection?”, “How are these issues addressed?”). The evaluated risks included pathogen risks because of operation outside of validated range.

FMEA can be used to analyse the reliability of each major process in a treatment plant and identify any instance where a failure (e.g. electrical, instrumentation, mechanical etc.) or water quality condition would prevent the treatment system from meeting its goals. The FMEA risk technique provides a method for determining the most critical component and implementation issues by evaluating the severity, occurrence, and detection of the failure through calculating and comparing the component risk priority numbers.

2.5.9 Reliability centred maintenance

Reliability Centred Maintenance (RCM) is a systematic approach which focusses on the reduction of maintenance costs by the selection of the most relevant components of a system for preventive maintenance by looking at their functions and failure modes, taking into account systems, subsystems and items (Bloom, 2006). The RCM assessment can be performed as a sequence of tasks, with some of them being simultaneously applied (Vatn, 1996): *"(1) study preparation, (2) system selection and definition, (3) functional failure analysis (FFA), (4) critical item selection, (5) data collection and analysis, (6) failure mode effects and criticality analysis (7) selection of maintenance actions, (8) determination of maintenance intervals, (9) preventive maintenance comparison analysis, (10) treatment of non-critical items, (11) implementation, (12) in-service data collection and updating"*.

Some of these activities can be aided by the incorporation of other reliability assessment tools. For example, functional failure analysis could be aided by functional block diagrams, reliability block diagrams, and fault trees (Bloom, 2006). Furthermore, in the critical item selection, importance ranking obtained from tools, such as fault tree analysis, reliability block diagrams, or Monte Carlo simulation may be suitable (Rausand & Høyland, 2004). FMECA is the central tool in this method, because it determines criticalities, failure causes, failure mechanisms, failure characteristics, among other important factors. After this analysis, preventive maintenance actions and their periodicity are selected (Rausand, 1998). An RCM analysis provides a thorough understanding of system functions, functional requirements, functional failures, and causes and consequences of functional failures.

Fynn *et al.* (2007) investigated how water utilities can apply RCM to new and existing plants to assess the costs and benefits of implementing this approach. The outcomes of this research suggest that water utilities can derive extensive benefit (both improved reliability and monetary savings) from utilising this approach to develop optimised maintenance programs for their plants. The benefits were demonstrated by the RCM programs performed on two US water utilities and in the results obtained from pilot studies. The authors assert that benefits may vary depending on whether the project is well structured or not.

2.5.10 Markov chain analysis

Markov chain analysis (MCA) (Figure 2-7) is used to quantitatively model stochastic systems and the transition between their different states over time. This characteristic has made it a key technique for modelling systems with dependent failure and repair modes. Usually the failures and repairs are assumed with constant failure and repair rates (i.e. exponential failure and repair

time distributions) (Durga Rao *et al.*, 2009). In addition, the requirement to solve a set of differential equations for large and complex systems can be problematic (Dhillon & Balbir, 1999). As a result, use of MCA is commonly limited to small system applications (Ericson, 2005). Its principal aim is to model state transitions which allow the comprehension of system operation and calculation of precise failure state probabilities.

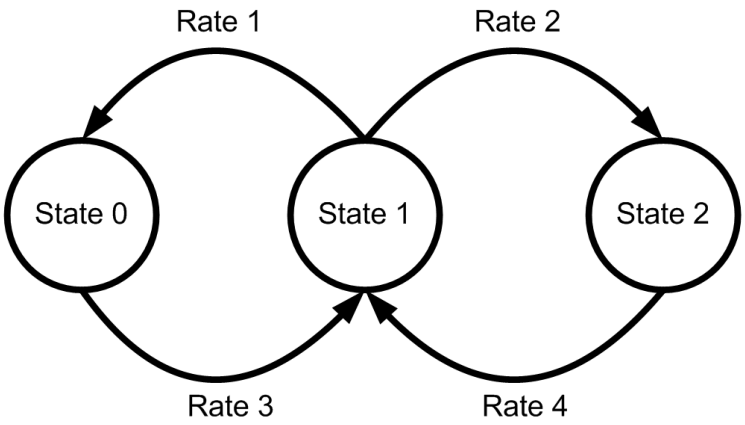

Figure 2-7: Markov chain representation

Weir *et al.* (2011) used MCA to model the fate and reduction of *Cryptosporidium* through a water reuse system for sprayparks (also known as "splash parks") application. They provided a useful and structured way of visualizing the system and the relationship between performance and its configuration. The approach they described also linked the outcomes of an MCA with a QRMA by incorporating Monte Carlo analysis to use the concentration distribution of *Cryptosporidium* oocysts, along with a dose-response distribution, to model risk levels. This methodology can be used as a multiple barrier analysis and it would require obtaining treatment performance and concentration reduction data between the states defined. However, a typical application of this methodology does not consider degraded operating states or effect of other factors on treatment performance which can limit its predictive ability. Lindhe *et al.* (2012) applied MCA to model dynamic gates of fault tree analysis to consider dynamic redundancy when modelling the reliability of a water treatment system. MCA has also been used to simulate sensors with different failure modes such as excessive drift, complete failure, wrong gain and lack of calibration (Rosén *et al.*, 2008). MCA was also used to analyse the performance of a wastewater treatment plant measured by the discharged BOD per day (Assezat, 1989). Each state in the model corresponded to a specific failure mode which is previously identified by FMEA procedure. The probability of the presence of each mode was calculated using mean time between failures, mean time to repair, and fault duration time. Mode effects are quantified by probabilities obtained and the discharged pollution as kg BOD/day during the failure mode.

2.5.11 Monte-Carlo simulation

Monte Carlo Simulation (MCS) performs repeating sampling draws from a probability distribution which are then used as inputs on a defined model (Stapelberg, 2008). The model outputs are obtained in the form of probability distributions. By itself, MCS is not a risk analysis tool, but a technique that is used as an aid for other tools to provide probabilistic outcomes. MCS is computationally intensive because it requires numerous iterations to obtain approximate values of central tendency and confidence limit bounds. The distribution types for input variables are selected based on the conditions surrounding that variable and may include all kinds of probability distributional forms (e.g. normal, triangular, uniform, lognormal, Bernoulli, binomial and Poisson) (Smid *et al.*, 2010). In general, MCS can be used for two distinct purposes (IEC/ISO, 2009): (1) Uncertainty propagation on conventional analytical models and, (2) probabilistic calculations when analytical techniques are not feasible (e.g. complex problems).

MCS has been used in the assessment of availability of complex systems as well as the monetary value of plant operations and maintenances (Durga Rao *et al.*, 2009). MCS has been used extensively to quantify microbial risk by considering water quality parameters, exposure values and parameters for dose-response model stochastic estimation (Teunis & Havelaar, 1999; Barbeau *et al.*, 2000; Fewtrell & Bartram, 2001; Medema *et al.*, 2003; Soller *et al.*, 2006; Lester *et al.*, 2007; Van den Akker *et al.*, 2011; Haas *et al.*, 2014). MCS has also been used for estimation of reliability, although it has not been explicitly referred to as a reliability method (Beauchamp, 2008). Inherent reliability and process performance for multiple barriers in advanced water treatment have been evaluated by MC simulation to account for variability in the performance (Haas & Trussell, 1998; Tanaka *et al.*, 1998; Nokes *et al.*, 1999; Olivieri *et al.*, 1999; Smeets & Medema, 2006; Neumann *et al.*, 2007; Khan & McDonald, 2010; Khan, 2010). For example, the risk posed by *Cryptosporidium* in filtered water during potable water treatment could be assessed through repeated sampling of a distribution of influent concentrations and mathematically combining the results with values drawn from a distribution of filter performance denominated in logs of reduction through physical removal. Such analyses could then be adapted to also simulate disturbances in the removal process to evaluate the risks attributable to process failure (Passantino & Owen, 2005). Mathematical treatment of validation data is usually performed through MCS by fitting parametric distributions to the influent and effluent concentrations, and then calculating the log removal values (LRVs) from the logarithm of the ratio of values obtained from the sampled distributions (van den Akker *et al.*, 2014; Branch *et al.*, 2016). LRVs calculated in this fashion usually assume some degree of correlation between the influent and effluent concentrations ranks (Branch, 2016). In mechanical reliability, MCS is used to analyse the reliability of complex systems (e.g. including stand-by components) and the availability of

repairable systems. Simulation can incorporate preventive maintenance, corrective maintenance, any type of distribution for the failure and repair times, and crew availability.

2.5.12 Bayesian networks and Bayesian methods

Bayesian networks (BNs) are probabilistic graphical models represented by directed acyclic graphs (DAGs) which are formed by nodes (variables) and directed arcs (connections). Two variables are identified as "parent" and "child" nodes if there is an arc from the former to the latter (Korb & Nicholson, 2011). Each node is attached to a conditional probability table defined by the combination of states of its "parents" (Jensen & Nielsen, 2007). The structure of a BN encodes conditional independencies between the variables which allows a factorisation of the joint probability distribution, reducing the amount of required information to define it (Kjræulff & Madsen, 2012). BNs use Bayes theorem (Jensen & Nielsen, 2007) to update the values of target variables given a set of observed variables, using a process is known as "reasoning". Three main types of reasoning arise from BNs (Kjræulff & Madsen, 2012): A deductive or causal analysis in which the occurrence of any node is computed from prior probabilities of root nodes and the conditional dependencies following the direction of the causal links; abductive or diagnostic reasoning which goes against the direction of the causal links and calculates posterior probabilities given some observation (or evidence) of the variables; and inter-causal reasoning which allows decreasing belief in unsupported node(s) by getting evidence that supports another hypothesis or node.

BNs usually comprise discrete variables, continuous Gaussian variables or a mixture of both. However, BNs can be used with any type of continuous distribution by resorting to Monte Carlo simulation, specifically Markov Chain Monte Carlo, which is used to obtain joint and marginal posterior distributions (Scutari & Denis, 2014). Queries and inference are possible with this methodology. Also, uncertainty in queries is possible to be obtained with this technique (Donald & Mengersen, 2014), which is not possible with commonly used inference algorithms.

BNs application has been extensive for some engineered processes, especially in the area of environmental modelling (Aguilera *et al.*, 2011; Phan *et al.*, 2016), and reliability and risk assessment as a decision making aid tool (Weber *et al.*, 2012). However, the application of BNs in water and wastewater treatment processes remains limited. Some of the relevant applications in the water sector comprise inherent reliability assessment including troubleshooting (Chong & Walley, 1996; Sahely & Bagley, 2001; Guo *et al.*, 2015b), and prediction of water quality (Li *et al.*, 2013). BNs for troubleshooting have been employed to identify the causes and predict upsets on biological treatment of wastewater. Li *et al.* (2013) successfully applied BNs for predicting the effluent water quality (total phosphorous, total nitrogen and COD) from operating parameters

and water quality in a modified sequencing batch reactor. Most studies have used expert knowledge for the structure and real data for the parameter learning, or solely expert knowledge on both steps (Chong & Walley, 1996). BNs for QMRA in water treatment processes has also been an area of recent development (Beaudequin *et al.*, 2015a; Beaudequin *et al.*, 2015c; Beaudequin *et al.*, 2017). A BN configuration to assess operational risks and their controlling factors and process indicators is presented in Figure 2-8.

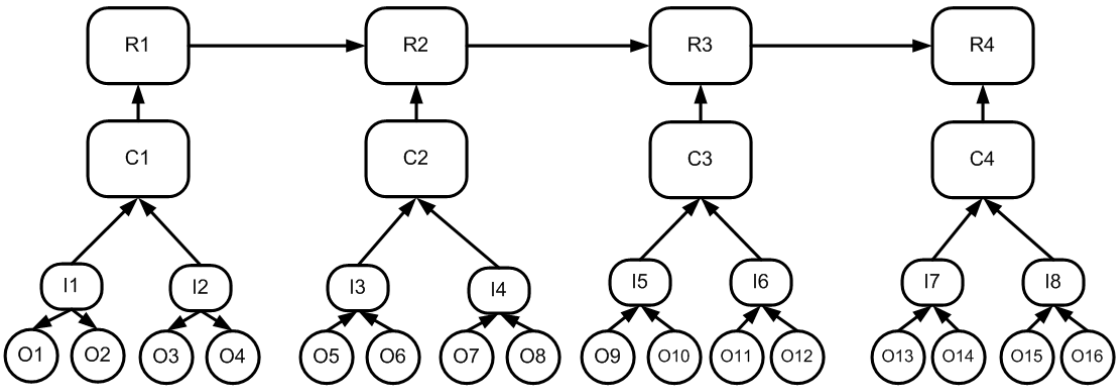

Figure 2-8: Operational risk control model DAG. Rs indicate the residual risks along the system, Cs indicate the risk controls, Is represent the impacts on the controls, and Os are the observations that can be used as indicators or causal factors to the impacts (adapted from Fenton and Neil (2012)).

Toifl *et al.* (2010) reviewed a number of monitoring techniques for wastewater recycling systems. They claimed that BNs may be applicable for process control as they have the ability for structure learning. However, this technique often requires large datasets which can be a limitation. Zhu and McBean (2007) evaluated BNs augmented with decision and utility nodes (i.e. influence diagrams) for deciding what combination of processes (i.e. coagulation, filtration, and disinfection) provided the highest utility considering raw water quality, treatment effectiveness, cost of processes and health risk. They concluded that BNs can successfully model this type of decision problem incorporating uncertainty and decision makers' preferences. The results obtained were consistent with intuition, but they were supported by a defensible decision-theoretic basis.

Cheon *et al.* (2008) developed a BN model for diagnosing and predicting abnormal quality (i.e. total phosphorous, ammonia and nitrates) during a five-stage step-feed enhanced biological phosphorus removal plant. The model's structure was constructed from expert knowledge and the parameters were learned from real data from a full-scale system and then updated with a lab-scale reactor. The models considered water quality (e.g. alkalinity, dissolved oxygen, pH) and operational parameters (e.g. MLSS, SRT, sludge blanket height), as well as failure of mechanical

parts such as a diffuser and blower, as causal factors affecting dissolved oxygen and an inlet pump as a factor affecting ammonia and carbon concentration. The prediction results were satisfactory, allowing the correct prediction of abnormal conditions. With the exception of the study reported by Cheon *et al.* (2008), failure of mechanical parts or integrity issues of critical components in water utilities has generally not been included.

Bayesian methods comprise statistical models which allow expressing variables and parameters as probability distributions and update those probabilities once data become available through the use of Bayes' rule (Gelman *et al.*, 2014). These models are able to incorporate prior information which may come from expert knowledge or previous measurements, define model parameters as random variables, and update information in the models based on new evidence. In the same manner as BNs, Bayesian methods can be graphically represented as a DAG. For instance, a graphical representation of the inference of parameters for a Gaussian distribution is presented in Figure 2-9 (Fenton & Neil, 2012).

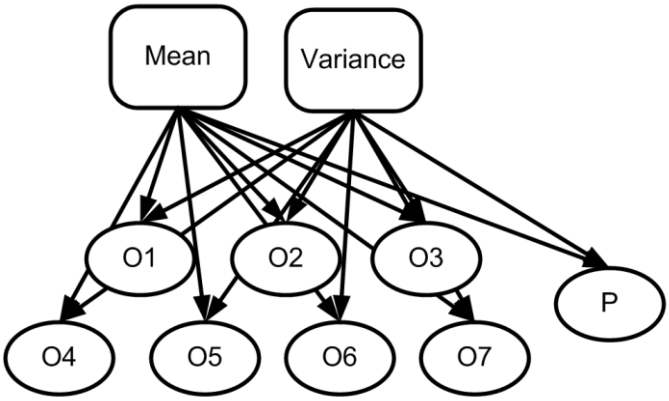

Figure 2-9: BN representation for inference of Gaussian parameters and posterior predictive distribution. Os represent observations and P is the predictive values.

Bayesian methods have been shown to be valuable tools for compiling data from different sources when conducting a meta-analysis (Messner *et al.*, 2001; Qian *et al.*, 2004; Pouillot *et al.*, 2015; Brooks & Field, 2016) or estimating uncertainty in model parameters (Sivaganesan *et al.*, 2003; Petterson *et al.*, 2007; Teunis *et al.*, 2009). Multiple barriers have also been studied through Bayesian hierarchical models considering correlations between different genogroups of virus for the influent and removal efficiency (log reduction value, LRV), censored observations, and seasonality influence (Pouillot *et al.*, 2015). Estimation of LRVs from validation data for wastewater treatment processes including left censored values have been studied using Bayesian methods (Kato *et al.*, 2013; Ito *et al.*, 2015; Ito *et al.*, 2016; Kato *et al.*, 2016). These studies have

also analysed the effect of the number of non-censored observations on the accuracy of LRVs and the influence on distribution parameters of pairing of influent to effluent samples.

2.5.13 Summary

To summarise the results presented in this section, the applicability of each tool to validation and reliability was indicated in Table 2-2. All reviewed tools were found to be suitable for reliability analysis. Each tool presented differentiated characteristics that made them useful for particular uses within reliability assessment. From the set of tools analysed, only some of them were shown to be appropriate for validation of water treatment processes. For validation, FTA, ETA, Bow-tie and MCA are methods able to model the association of variables in a systematic fashion which is useful to show the most important factors affecting performance and to understand the key functioning mechanisms. MCS has been widely used in validation to simulate scenarios and enhance through probabilistic assessment the analysis conducted with other methods. HHRA is useful for validation because it provides the means by which the performance and reliability of the multiple barrier system can be benchmarked against. BNs and Bayesian methods were shown to be very applicable for both validation and reliability tasks. They have been used to represent associations in a variety of settings and tasks. They have also shown to be useful because of their flexibility and inference capabilities.

Table 2-2: Applicability of reliability tools for validation and reliability

	Validation	Reliability
HRA		✓
HHRA	✓	✓
FTA	✓	✓
ETA	✓	✓
RBD		✓
Bow-tie	✓	✓
LOPA		✓
FME(C)A		✓
RCM		✓
MCA	✓	✓

	Validation	Reliability
MCS	✓	✓
Bayes	✓	✓

2.6 Evaluation of status

This section presents an initial assessment of the previously reviewed tools for validation and reliability assessment. It presents the most relevant techniques and a general comparison of their characteristics.

2.6.1 Assessment of analysed literature

Tools previously reviewed have demonstrated to be successful in their applications in other industry sectors where safety and reliability are the primary concern (e.g. aircraft, nuclear, aerospace) (Dhillon & Balbir, 1999; Keller & Modarres, 2005; Liu *et al.*, 2012a). The literature analysed showed that all tools have been used or proposed as methods to measure and improve reliability. Each tool offers advantages that make it unique and useful for a particular purpose.

The literature reviewed above indicates that while some reliability techniques are increasingly commonly applied in the assessment of water and wastewater treatment systems, others have not been fully explored. For example, none of the reviewed studies have quantitatively assessed human reliability in water treatment systems. Software reliability has also not being addressed in these studies. Given the importance of human decision on these systems and the reliance on multiple sensors along the process, they are crucial components within the reliability estimation of the process.

FTA was revealed to have been one of the most common quantitative approach for studying reliability, availability, compliance scenarios and hazardous events. Some features that make this tool attractive include its explicit representation of causal associations and the generation of minimal cut-sets. But few studies have incorporated real failure data into their models, suggesting that data, such as failure times and repair times may not be readily available. Expert elicitation has been successfully employed for estimating failure rates and uncertainties when available data have been limited (Lindhe *et al.*, 2009). In addition, a Bayes' theorem has been shown to be useful for updating distribution parameters with real operational data when subjective data have been initially used. ETA is useful for determining the consequences that require further analysis,

treatment and devotion of resources. In addition, a valuable advantage of ETA is the sequential study of events, which is appropriate for water treatment and management processes.

Bow-tie and LOPA appear to be highly applicable to water processes for analysis of all types of risks. Their focus on the barriers between the causes and the risk, and the risk and consequences make them very useful for a multiple barrier system. They can also be used to assess accident scenarios. When applying any of these two techniques, attention should be paid to not over-simplify complex systems and as a result failing to consider interactions between multiple events. The attributes required to evaluate each barrier can be relevant for determining whether sufficient preventive and mitigating measures are in place in a water treatment and supply system. The main limitations of FTA, ETA, Bow-tie, LOPA and RBD are the difficulties of expressing multistate and dynamic outcomes. Special care should be taken when common cause failures affect the system as they increase the complexity in the model calculation.

FMEA and FMECA are valuable tools for reliability assessment, especially for analysing several potential hazardous events. However, the analysis of rare events from the combination of different events occurring at the same time is not possible, nor is the analysis of dependencies. Because of this limitation, FMEA/FMECA should be complemented with quantitative system analysis tools such as FTA or ETA.

Important outcomes from HHRA and QMRA are the derivation of water quality health based targets and performance targets for water treatment processes. MCS has been the preferred tool when analysing validation data as well as for enhancing HHRA and QMRA. In general, MCS serves as an add-on statistical method to enhance the analysis of other techniques by incorporating stochastic results and multiple scenarios through simulation.

BNs and Bayesian methods were revealed to be more versatile tools than ISO 31010 suggested and hence appeared to have potential for application to a broad range of real problems and to encode other tools (Section 2.7). These conclusions have also drawn from studies evaluating the ability of BNs to provide a broad framework for reliability and risk analysis (Fenton & Neil, 2012; Roser *et al.*, 2015). Their use in the water utility sector has focused mainly on biological systems because they present multivariate problems with complex relationships which make BNs suitable to be applied. Water treatment processes are dynamic and the monitoring incorporates variability and uncertainty. BNs and Bayesian methods are intrinsically probabilistic rather than deterministic which make them more convenient for decision support in this case. Because of the explicit graphical representation of BNs (and availability of user-friendly commercial software packages), they are able to summarise large amount of data and facilitate interpretation, requiring little statistical training. They also do not require detailed modelling for effective use.

The tools reviewed above can be assessed in terms of their ability to meet the needs of the proposed ideal tool as shown in Table 2-1. Although a number of tools incorporate many of the important characteristics (Table 2-3), BNs and Bayesian methods support all the features required by the ideal tool. These characteristics make these tools versatile and adequate for a large number of problems in reliability and decision reasoning. FTA, ETA and Bow-tie also possess a number of characteristics which make them attractive for a broad range of cases in reliability assessment (Table 2-3). HRA, RCM and MCS methods were not included because they are tools that need to be used in conjunction with other tools to be useful.

Table 2-3: Strongly applicable reliability tools

Tool class analysed	Quantitative outcome	Probabilistic assessment	Different sources of data	Continuous or multistate outcomes	Easy construction	Representation of associations	Causal representation	Inference
QMRA	✓	✓		✓				
FTA	✓	✓	✓		✓	✓	✓	
ETA	✓	✓	✓		✓	✓	✓	
RBD	✓	✓	✓		✓	✓		
Bow-tie, CCA	✓	✓	✓		✓	✓	✓	
LOPA	✓	✓	✓			✓	✓	
FMEA/FMECA			✓		✓			
MCA	✓	✓		✓		✓		
Bayesian	✓	✓	✓	✓	✓	✓	✓	✓

2.6.2 Machine learning tools

Machine learning tools have acquired prominence in the recent time because of their effectiveness in dealing with a broad range of decision making and prediction tasks. Such prediction tasks can also include reliability analysis. Machine learning tools within the water treatment domain have been largely used for prediction of water quality and performance. Various tools have been

proposed to predict the quality of water given operating and water quality parameters. Another topic which has brought significant attention is the use of statistical tools for process control and monitoring for fault detection in water treatment processes (Toifl *et al.*, 2010; Haimi *et al.*, 2013). Process control refers to the activities conducted to maintain a system working within normal conditions and avoid significant performance deviations. Process monitoring for fault detection is responsible for the diagnosis of the process and identifying failures of system components. In general, these tools have been shown to be problem-specific, meaning that one tool may be appropriate for a particular application, but not necessarily for another. They have also been applied in combination to extend their performance.

Artificial Neural Networks (ANNs) are so far, one of the most popular tools for both water and wastewater treatment prediction tasks (Hadjimichael *et al.*, 2016). ANNs are computational tools composed of a set of interconnected units called neurons. They are based on three types of neuron layers called input layer (i.e. independent variables), hidden layer (which acts as a feature detector) and an output layer (i.e. dependent variables). Khataee and Kasiri (2011) reviewed the application of ANNs for biological water and wastewater treatment processes. These models often fit biological systems' data well because they can model complex interactions without the need of additional mechanistic information.

In second place, following neural networks, clustering techniques have also been widely used within water and wastewater treatment (Hadjimichael *et al.*, 2016). Other examples of applied methodologies include regression trees (Barca *et al.*, 2016), support vector machines (Huang *et al.*, 2009; Fang *et al.*, 2011; Verma *et al.*, 2013; Guo *et al.*, 2015a; Mohammadpour *et al.*, 2015), partial least squares (Lee *et al.*, 2006; Singh *et al.*, 2010; Qin *et al.*, 2012; Flyborg *et al.*, 2017), PCA (Tao *et al.*, 2013; Wang *et al.*, 2017), and multiple linear regression (Uyak *et al.*, 2005; Singh *et al.*, 2010; Dotan *et al.*, 2017).

BNs and Bayesian methods can be benchmarked against the previously mentioned machine learning tools. Some of these tools are potentially applicable especially for the validation of treatment process. However, the characteristics that make Bayesian methods unique also include the capabilities of representing causal relationships between variables, their transparency in showing explicitly the associations between variables, naturally handling missing and censored values, and the potential inclusion of both expert knowledge and empirical data.

2.6.3 Interactions between tools

It appeared that some of the tools reviewed in this chapter can be combined to work in conjunction with one another (Table 2-4), meaning that certain reliability techniques together can result in

more informative and complete outcomes (Khakzad *et al.*, 2011; Khakzad *et al.*, 2013a; Khakzad, 2015). In some instances, combining tools may even provide new characteristics or features to the overall approach. An example of these new features is the use of MCS in conjunction with FTA, ETA, and Bow-Tie analysis. MCS provides the incorporation of uncertainty in parameters and facilitates the analysis of more potential scenarios. A brief description of each interaction is explained in Table 2-5.

Table 2-4: Interactions between risk assessment tools

	HRA	HHRA	FTA	ETA	Bow-tie	FME(C)A	RCM	LOPA	MCA	MCS	RBD	Bayes
HRA		-	✓	✓	✓	-	-	✓	-	-	-	✓
HHRA			-	-	-	-	-	-	-	✓	-	✓
FTA				✓	✓	-	-	-	✓	✓	-	✓
ETA					✓	-	-	✓	-	✓	-	✓
Bow-tie						-	-	-	-	-	-	✓
FME(C)A							✓	-	-	-	-	-
RCM								-	-	✓	✓	✓
LOPA									-	-	-	✓
MCA											-	✓
MCS											✓	✓
RBD												✓
Bayes												

Table 2-5: Rationale for perceived overlap/interactions between risk assessment tools

Abbreviated names	Description
FTA-HRA	Fault trees can be used in conjunction with human reliability analysis. Probabilities obtained from human reliability analysis can be incorporated into fault trees.
ETA-HRA	Event trees can be used in conjunction with human reliability analysis. Human reliability is used as one of the inputs for event tree analysis (e.g. initiating event, failure of barrier).
ETA-FTA	Fault trees can be used to estimate the probability of failure of a barrier in event tree analysis.
Bow-tie-HRA	Cause consequence has the same inputs as fault tree analysis and event tree analysis.
Bow-tie-FTA	Cause and consequence analysis incorporates both fault tree analysis and event tree analysis.
Bow-tie-ETA	Event trees are incorporated in cause and consequence analysis.
RCM- FME(C)A	FMEA/FMECA provides input information for reliability centred maintenance.
LOPA-HRA	Layers of protection analysis concerns the analysis of barriers, human reliability is important to keep barriers working.
LOPA-ETA	Layers of protection analysis uses event trees to depict the barriers and calculate the probability of undesired events.
MCA-FTA	Markov analysis is used to create the dynamic (time inclusion) version of the fault tree analysis or reliability block diagram.

Abbreviated names	Description
MCS- HHRA	Monte Carlo simulation is used to include uncertainty and variability into health risk assessment.
MCS- FTA	Monte Carlo simulation is useful when common causes are present in the fault tree analysis and when complex configurations are present. It is used to sample from probability distributions of failure and repair time.
MCS-ETA	Monte Carlo simulation can be used to include uncertainty and variability in the calculation of undesired events during event tree analysis.
MCS-RCM	Monte Carlo simulation is used for analysis of reliability within specific tools that provide input to reliability centred maintenance.
RBD-RCM	Reliability block diagram provides input information for reliability centred maintenance.
RBD-MCS	It is used when complex configurations are present. It is used to sample from probability distributions of failure and repair times.
Bayes-HRA	BNs have been used to encode human reliability models and model dependencies between factors affecting human failure events. Bayesian methods can be used to incorporate subjective information during human reliability.
Bayes- HHRA	BNs and Bayesian methods give the same options as Monte Carlo for health risk assessment but also incorporating inference.
Bayes-FTA	BNs can model fault trees. Fault trees can be considered as a special case of BNs. Bayesian methods have been used to incorporate uncertainty into the fault tree analysis methodology.
Bayes-ETA	Event trees can be considered as a special case of BNs. Bayesian methods have been used to incorporate uncertainty into the methodology.

Abbreviated names	Description
Bayes-Bow-tie	Cause and consequence analysis can be considered as a special case of BNs. Bayesian methods have been used to incorporate uncertainty into the methodology.
Bayes-RCM	BNs and Bayesian methods can be used to stochastically assess different maintenance strategies.
Bayes-LOPA	Layers of protection analysis (event tree part) can be performed through BNs. Bayesian methods have been used to incorporate uncertainty into the methodology.
Bayes-MCA	BNs can model Markov chains. Bayesian networks generalise them.
Bayes-MCS	Monte Carlo simulation can be used to perform inference in BNs. Useful for large and complex models.
Bayes-RBD	Reliability block diagrams can be modelled by BNs. Bayesian methods have been used to incorporate uncertainty into the methodology.

2.7 Mapping other tools with Bayesian networks

This section presents additional evidence of the potential of BNs as a versatile tool that can be applied to validation and reliability assessment and management in different settings. BNs have been used for different types of assessments even for mapping of classic reliability techniques such as FTA, ETA, bow-tie and RBD. BNs make the analyses more powerful and flexible, providing more information to the modeller. Tools which have been encoded into BNs are presented in Table 2-6.

Table 2-6: Tools mapped into Bayesian networks

Mapped tool	**References**
HRA	Mkrtchyan *et al.* (2015)
FTA	Bobbio *et al.* (2001); Montani *et al.* (2008); Fenton and Neil (2012)
ETA	Bearfield and Marsh (2005); Marsh and Bearfield (2008); Fenton and Neil (2012)
Bow-tie, CCA	Khakzad *et al.* (2013a)
RBD	Torres-Toledano and Sucar (1998); Solano-Soto and Sucar (2001)
LOPA	Pasman and Rogers (2012); Pasman and Rogers (2013)
MCA	Weber and Jouffe (2003); Weber *et al.* (2004); (Montani *et al.*, 2005); Koller and Friedman (2009); (Cai *et al.*, 2013)
QMRA, ERA	Smid *et al.* (2010); Liu *et al.* (2012b); Beaudequin *et al.* (2015b)

Human reliability analysis (HRA) tools have been encoded through BNs to improve different aspect of the HRA methodology. Some of the applications of BNs into HRA include modelling of organizational factors, analysis of the relationships among failure influencing factors, BN-based extensions of existing HRA methods, dependency assessment among human failure events, assessment of situation awareness (Mkrtchyan *et al.*, 2015).

When FTA and ETA methods are employed to assess system reliability, the focus is in the whole system reliability rather than individual components which make impossible to update individual component performance statistics (Martz & Waller, 1982). By the use of BNs, individual component reliability can be targeted, and evidence on any node in the network can provide information on all other nodes (Mahadevan *et al.*, 2001).

Fault trees can be considered special cases of BNs. Bobbio *et al.* (2001) published an algorithm that can be used to convert FTA to BNs. In a similar fashion event trees (Bearfield & Marsh,

2005) and Bow-tie (Khakzad *et al.*, 2013a) can be modelled by BNs. Reliability block diagrams has also been mapped into BNs including complex structures (Torres-Toledano & Sucar, 1998). Pasman and Rogers (2012); Pasman and Rogers (2013) studied the advantages of encoding a Layers of Protection Analysis models through BNs.

The advantages of mapping to BNs have been recognised by many authors and applied to case studies in different fields (Liu *et al.*, 2008; Lampis & Andrews, 2009; Duan *et al.*, 2010; Khakzad *et al.*, 2013b; Khakzad *et al.*, 2013a; Leu & Chang, 2013). BNs have provided more transparency, offering multi-states, and explicitly representing event dependencies when encoding FTAs and ETAs. BNs can overcome some of the shortcomings common for these tools such as binary states, independent events assumption, redundancies, sequential and common causal events. In general, BNs require more computational effort than FTA, ETA or Bow-tie approaches. However, the additional effort permits the incorporation of backward propagation once evidence becomes available on any of the BN nodes (Mahadevan *et al.*, 2001). An important application is also "probability adapting" which is used to update probabilities of consequences by belief updating on safety analysis (Khakzad *et al.*, 2013a). The fact that BNs can be flexible enough to model fault or event trees is a proof that this tool can be used in several applications. This feature would allow taking advantage of the benefits of both tools in one application.

Health risk assessment, specifically quantitative microbial risk assessment, has been reviewed by Beaudequin *et al.* (2015b) across different domains including water and food. The review comprised literature about BN modelling and hierarchical Bayesian analysis. Hierarchical Bayesian analysis is often expressed as DAGs in the same manner as BNs do. The advantage of Bayesian analysis is the use of either discrete or continuous distributions without limitation of on the distribution type. The models constructed in this fashion can incorporate prior knowledge and data. In addition, they also easily incorporate parameters explicitly in the model. However, they require defining the function and relationships between the variables in advance. BNs are particularly suitable for QMRA because of the explicit causal representation, incorporation of evidence and backwards reasoning.

Markov chains allow modelling reliability in a dynamic fashion. They have been mapped into BNs using Dynamic Bayesian Networks (DBNs). Incorporation of the time variable permits modelling the change of a particular system variable through time, which is important when degradation of a system condition and repair need to be modelled including causal factors (e.g. failure of a mechanical component).

2.8 Conclusions

In this review, reliability aspects have been considered in respect to water reuse systems including inherent reliability and mechanical reliability. Inherent reliability relates to the variability in system performance and through validation this aspect can objectively be evaluated. However, validation entails a number of tasks that need to be addressed also including the analysis of potential performance predictors and incorporation of uncertainties as well as censored or missing values into the reliability/risk estimations. Validation of treatment systems has traditionally been performed through Monte Carlo simulation. Although Monte Carlo simulation is a powerful technique, Bayesian methods offer the opportunity to incorporate priors, compile information from different sources, determine parameter uncertainties explicitly, and deal with missing and censored values. Few studies were found to assess or report applications of Bayesian analysis for water or wastewater treatment process validation. But there were sufficient studies to indicate the analysis was possible in principle and many potential opportunities from the application of Bayesian tools.

In the case of mechanical reliability, the review of current literature revealed that the water reuse industry, including especially the potable water reuse sector, is currently exploring how to ensure reliability as a matter of priority. The study of reliability can be conducted at various levels of detail depending on the required scope. For example, a particular process could be evaluated in terms of its specific mechanical components to calculate the process reliability. For potable reuse systems, the analysis of the multiple barriers and their reliability appears a particularly important assessment task. However, few options exist to model reliability while effectively incorporating the inter-relationships between individual barriers of the multiple barrier system. Because of limited data availability and the large human integration within these systems, a participatory process, in which information from experts and real data could be incorporated, appears necessary. The reviewed literature indicated that these requirements could be well captured and modelled using BNs.

A number of tools exist for performing risk and reliability analysis. From the variety of tools reviewed, it was found that only a few of them offer the flexibility required for the comprehensive assessment of reliability at diverse levels of abstraction and complexity. It was found that FTA was one of the most commonly applied reliability tool in water and wastewater treatment. An important characteristic of this tool is the capability for providing both cause-effect analysis and probabilistic analysis. The literature revealed that BNs and Bayesian methods could encode and enhance the features of FTA and other traditional reliability techniques.

Rigorous risk and reliability assessment is required to ensure high levels of water quality and compliance with risk management requirements imposed by regulatory agencies. Current practices and guidelines do not generally promote the use of quantitative reliability tools which makes the task difficult to manage. Reliability has an inherent mathematical meaning, therefore a quantitative assessment appears required if it is to be characterised, controlled, predicted and enhanced. In this sense, conventional qualitative tools, although useful, they are limited by the potential outcomes they can provide.

BNs applications in the water management field have mainly been limited to biological systems and to the prediction of bulk water quality parameters. However, the characteristics of these tools suggest they are suitable for a much wider range of cases, including reliability analysis in water treatment systems. BNs in general provide prediction capabilities, representation of causality and uncertainty assessment which appear suited to undertaking most relevant aspects for assessing reliability and risk in water treatment systems.

The following chapters of this thesis describe activities that were undertaken to confirm this promise in a practical and experimental way and so assess the operational usefulness of BNs and Bayesian methods in a number of water reuse systems applications. Chapters 4 to 7 present case studies for the problem of validation during activated sludge (Chapter 4), ultrafiltration (Chapter 5), ozonation (Chapter 6) and chlorination (Chapter 7). Chapter 8 presents a reliability assessment for UV disinfection and Chapter 9 shows how a multi-barrier system can be modelled through BNs.

Chapter 3: Materials and methods

This chapter describes the analysis techniques used to develop the Bayesian networks (BNs) and Bayesian methods in this study. These tools were common to different chapters. Experimental settings, modelling and data inputs specific for each case study are explained within each particular chapter.

There were multiple times in which terminology was repeated referring to different concepts depending on the context. These terms included: parameters and variables. To clarify their meaning and use they were defined in here. Parameters in a BN refer to the conditional (or marginal) probabilities associated to each node. In Bayesian analysis, parameters are the values defining the characteristics of a probability distribution. For example, for a normal distribution, the mean and standard deviation correspond to its parameters. In water treatment processes monitoring parameters refer to measurements made by instruments used to observe some characteristic of the process or water quality. Variable in the context of BNs is a synonym of node, so they can be used indistinctly. In Bayesian analysis, variables are the inputs and outputs of interest in a model. For example, in a model that is constructed to analyse system performance, the measured performance would be the output variable, whereas the values used to explain the performance would be the input variables.

3.1 Bayesian networks

As described in the literature review (Chapter 2, Section 2.5.12), BNs are probabilistic graphical models represented by "Directed Acyclic Graphs" (DAGs), which can model non-recursive causal relationships in complex systems and facilitate inferential reasoning. A BN structure is defined by directional connections, known as "arcs", which specify the dependence and independence assumptions between random variables, termed "nodes" (Figure 3-1). These interdependencies determine what information is required to specify the probability distribution among the random variables of a network. Two variables are identified as related "parent" and "child" nodes if there is an arc from the former to the latter (Korb & Nicholson, 2011). When a variable has parents, a set of conditional probabilities must be defined in the child node for each combination of parent node "states" which may be categories, values or value ranges. Nodes without parents (root nodes) only require marginal probabilities. BNs reduce the quantity of information required to define a joint probability distribution through factorisation conducted using the chain rule as shown in Equation 3-1.

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^{n} P\left(X_i | X_{pa[i]}\right) \qquad \text{Equation 3-1}$$

Where $P(X_1,X_2,\dots,X_n)$ is the joint probability distribution of variables (X_1, $X_2,\dots,X_n$), X_i corresponds to a random variable represented by the node i in {1,…,n} and $pa[i]$ denotes the

parents of node i, $X_{pa[i]}$ indicates a set of random variables associated with $pa[i]$. Each node from a BN has mutually exclusive discrete states which are associated to marginal probabilities (indicated by the red filled rectangle in Figure 3-1). These marginal probabilities are obtained by marginalising the joint probability distribution. More detailed information about marginalisation can be found in Koller and Friedman (2009).

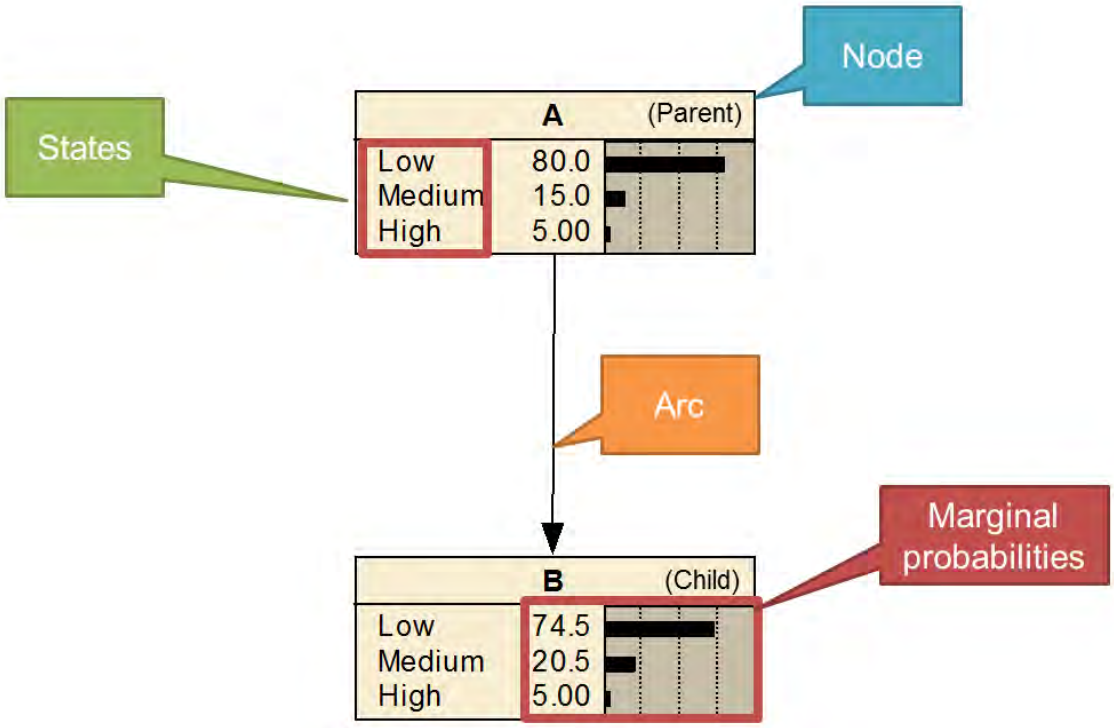

Figure 3-1: Structure and main components of a Bayesian network

3.1.1 Model construction approaches considered

3.1.1.1 Types of connections

Three types of connections exist in BNs including serial (Figure 3-2 a), converging (Figure 3-2 b), and diverging (Figure 3-2 c) connections. These three types of structures have also been known as causal chain, common effect and common cause, respectively (Korb & Nicholson, 2011). The types of connections in a BN determine the conditional dependencies and the flow of information or influence between different nodes when evidence is provided. This characteristic is relevant as the network should represent the behaviour and influences of the modelled system. In the serial structure knowing that A has occurred does not provide information to C if B has already been observed. The conditional independence in this case can be written as [2] $A \perp\!\!\!\perp C|B$. In the converging structure, a node has n number of parent nodes which are marginally independent, but they

[2] This reads A is conditionally independent of C given B

become dependent once an observation is provided for the common effect. Therefore, for this type of connection, the parents are conditionally dependent given B, which can be written as [3]A ⫽C|B. The diverging structure gives rise to the same conditional independence structure found in the serial connection, so the relationship can be represented as A⫫C|B.

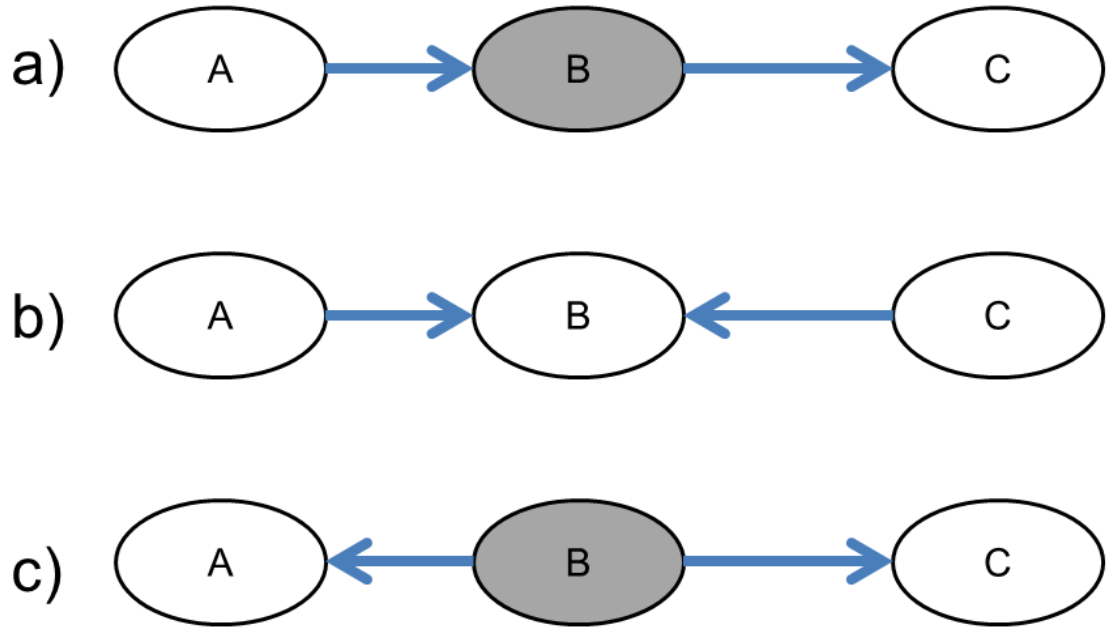

Figure 3-2: Types of connections in Bayesian networks. a) serial, b) converging and c) diverging connections. Grey nodes represent evidence that blocks information.

3.1.1.2 Construction options employed

A BN structure can be determined based on expert information, data or a combination of both. Although there is no universally accepted formal procedure for BN construction, some methods have been proposed for this task. For example, the use of expert information can be facilitated by the utilisation of idioms which are pre-defined BN structures modelling a particular system behaviour (Fenton & Neil, 2012). These idioms can be divided into five types including definitional, cause-consequence, measurement, induction, and reconciliation idioms. More detailed information about the use of idioms and their structure can be found in Neil *et al.* (2000). Another option for structuring models is using other traditional reasoning methods such as fault tree analysis, event tree analysis, Bow-tie to generate the basic structure of the network and then map such models into a BN including additional structures or connections (Leu & Chang, 2013).

[3] This reads A is conditionally dependent of C given B

If the construction of the network is based on data or a combination of data and expert information, various options can be selected. The appropriate choice of options depends on the purpose of the model and, on the algorithms used for their construction. With regards to the purpose, if the model seeks to predict the outcomes of a particular variable, non-causal or naïve BNs can be used (Korb & Nicholson, 2011). Naïve BNs present a pre-defined structure in which the target node is the parent of all the attributes or explaining variables. Some modifications to this method have been incorporated, such as the inclusion of arcs between the attributes to improve prediction capabilities. Such models are known as semi-naïve BNs. When a "causal" representation or understanding of the conditional dependencies of the system is required, a "causal" type of network can be employed. In this case, connections are not forced in a particular configuration. However, an expert can provide input in the form of forbidden connections and forced connections between nodes which are known to exist in reality. Appendix 1 presents an example to illustrate the difference between naïve Bayes models and causal BNs.

The issue of the second choice relates to the algorithms used to achieve the final model structure. A number of algorithms have been developed in which the data are used to generate the connections between the nodes. Two main types of algorithms exist for this purpose including score-based and constraint-based algorithms for structure learning. The first types of algorithms use a search procedure and a score function (Section 3.1.2) which is maximised. The second type uses conditional independence tests such as the mutual information test and χ^2 test for conditional independence (Scutari & Denis, 2014). A disadvantage of constraint-based algorithms is that arcs without direction can be produced by this method, which is not the case of score-based algorithms. More information about the different types of algorithms used for structure learning can be found in Jensen and Nielsen (2007); Koller and Friedman (2009); Korb and Nicholson (2011); Kjræulff and Madsen (2012); Scutari and Denis (2014).

3.1.2 Model parameters

Bayesian networks use marginal probabilities (for nodes without parents) and conditional probabilities (for nodes with parents) to define the parameters of the model. In the case of the model presented in Figure 3-1, two sets of parameters need to be obtained, one for node "A" (Table 3-1) and another one for node "B" (Table 3-2). Parameters for node "B" are conditional probabilities because there is an arc from "A" to "B". It is important to note that each row in the parameters table must sum to 1. When data are used to estimate model parameters, each cell is populated by counting the number of occurrences for a particular condition. This algorithm is known as maximum likelihood estimation (MLE). When missing values are present, more sophisticated algorithms are required, in which inference is required as an intermediate step to

obtain an estimation of the probabilities. Algorithms for this type of computation include expectation maximisation and gradient descend (Koller & Friedman, 2009).

Table 3-1: Parameters for parent node A

P(A=Low)	P(A=Medium)	P(A=High)
0.8	0.15	0.05

Table 3-2: Parameters for child node B

A	P(B=Low|A)	P(B=Medium|A)	P(B=High|A)
Low	0.90	0.10	0
Medium	0.15	0.80	0.05
High	0.05	0.10	0.85

3.1.3 Network scoring for assessing BN quality

In a BN context, likelihood is defined as the probability of a set of data given the parameters in the model (which is defined by the model structure G) (Equation 3-2) (Kjræulff & Madsen, 2012). Likelihood can be used as a goodness of fit metric for a BN. Because calculations of likelihoods imply a multiplication of probabilities, the logarithm of likelihoods (log-likelihood) is used to change the multiplication to summation and facilitate calculations. Therefore, the log-likelihood of the data $D=\{c^1,\ldots,c^N\}$ given structure G is the sum of the contributions from each variable for each case.

$$l(G:D) = \sum_{l=1}^{N} logP(c^l)$$ Equation 3-2

$$= \sum_{l=1}^{N} \sum_{i=1}^{|V|} logP(X_i = x_i^l | pa(X_i) = x_{pa(X_i)}^l, c^l)$$

Where $(x_i^l, x_{pa(x_i)}^l)$ are the values of $\left(X_i, pa(X_i)\right)$ in the l^{th} case c^l of D. The number of nodes is represented by V.

The structure of a network can usually be evaluated using score metrics which trade-off goodness of fit and complexity. Therefore, a more complex network (more connections) will require a high goodness of fit to be selected. Goodness of fit in this case is measured through the log-likelihood score whereas the complexity can be measured in different ways including the number of independent parameters required by the model and the size of the dataset. A common score used during network comparison is the Akaike Information Criterion (AIC) (Equation 3-3). AIC uses the independent number of parameters (k) for penalising the complexity in the network. In general, lower score values indicate better fit.

$$AIC = l(D; \theta, G) - k \qquad \text{Equation 3-3}$$

3.1.4 Model evaluation tools

Two main types of evaluations can be conducted after a model is developed including hold-out testing and cross-validation. In hold-out testing the data are usually divided into two parts, one for fitting or training the model and the other one for testing its performance. Hold-out testing is usually performed when the amount of data is large so the size of the data to train the model does not affect its performance. When the data are scarce, cross-validation can be used to test the performance of the model. Cross-validation was used in this study.

3.1.4.1 Cross-validation

Typically, the type of cross-validation used on many BN studies corresponds to 10-fold cross-validation (Koller & Friedman, 2009). This approach randomizes and partitions the data into 10 equally sized sets and then 10 validations are made using $9/10^{th}$ and $1/10^{th}$ of the data for training and testing, respectively (Figure 3-3), every time with each portion of the data (Koller & Friedman, 2009). Cross-validation is performed to mitigate any bias produced by a particular sample chosen for training and testing. Cross-validations are usually used along with stratification randomisation. Randomised stratification means that the proportions within the classes of the class node are approximately the same in each fold. Another commonly used approach for

validation of models using cross-validation is the leave-one-out cross-validation (LOOCV). In this case, LOOCV consists of removing a single observation from the dataset (of size: n) and fitting the model using n-1 observations, and then using the removed observation to validate the model's prediction of these data. The process is repeated n times, so all observations are used for model construction and testing. In this case, no stratification is required.

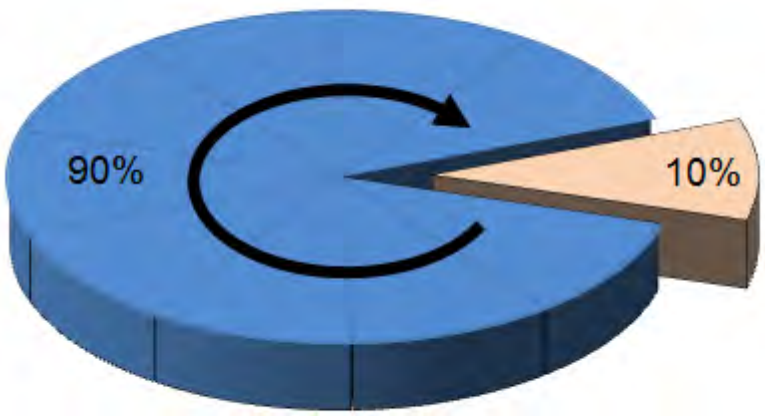

Figure 3-3: Schematic representation of the 10-fold cross-validation.

3.1.4.2 Performance metrics

A number of metrics have been developed to estimate the performance of a BN after testing it against real data. Common metrics were employed in this study including prediction accuracy (Equation 3-4), Kappa statistic (Ks) (Equation 3-6), rates (i.e. false/true-negatives/positives) (Equation 3-11 to Equation 3-14), receiver operating curve (ROC) and its area under the curve (AUC). In general, using the metrics on their own is not sufficient and a combination of them is recommended to assess the real performance of a model (Marcot, 2012). The results of the performance of a BN for a two-outcome node can be represented by a contingency table as shown in Table 3-3. As observed, the real and predicted outcomes can be true or false. When the predicted outcome is true and the real outcome is also true, a true positive count is obtained, whereas when both the predicted and real outcomes are negative, a true negative count is obtained. A false positive is obtained when the predicted outcome is true, but the real outcome is false. Conversely, a false negative is found when the predicted outcome is false, but the real outcome is true. In Table 3-3, TP, FN, FP and FN represent the number of instances or counts which were found to be true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

Table 3-3: Contingency table for two outcomes

		Real outcomes	
		True	False
Predicted	True	TP	FP
	False	FN	TN

Prediction accuracy measures the number of correctly estimated outcomes divided by the total number of outcomes (Equation 3-4) (Witten & Frank, 2005). This metric is commonly used when evaluating the BN prediction performance. However, one of its main limitations is providing high accuracies when only few cases are available for one of the real outcomes compared to the other(s) (this is known as an unbalanced outcome). Calculation of prediction accuracy generalised for g number of outcomes is presented in Equation 3-5.

$$Prediction\ accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Equation 3-4}$$

$$Prediction\ accuracy = \frac{1}{n}\sum_{i=1}^{g} f_{ii} \quad \text{Equation 3-5}$$

Where n is the total number of occurrences and f_{ii} is the number of occurrences in the diagonal of a square contingency matrix with g distinct outcomes.

The Kappa statistic (Ks) metric (Equation 3-6) uses the prediction accuracy (Equation 3-7) but removes the agreement obtained by chance (Equation 3-8). Ks measures the agreement between model predictions and actual values as a metric in the range [-1,1]. KS = 1 means perfect agreement, KS = 0 means that agreement is equal to chance, and KS = −1 means "perfect" disagreement (Marcot, 2012).

$$Ks = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$ Equation 3-6

$$\Pr(a) = \frac{TP + TN}{TP + TN + FP + FN}$$ Equation 3-7

$$\Pr(e) = \frac{(TP + FP) \cdot (TP + FN) + (FN + TN) \cdot (FP + TN)}{(TP + TN + FP + FN)^2}$$ Equation 3-8

Where Pr(a) is the proportional agreement (same as prediction accuracy) and Pr(e) is the expected agreement. For any number of outcomes the Pr(a) is defined by Equation 3-9, while Pr(e) is defined by Equation 3-10.

$$\Pr(a) = \frac{1}{n}\sum_{i=1}^{g} f_{ii}$$ Equation 3-9

$$\Pr(e) = \frac{1}{n^2}\sum_{i=1}^{g} f_{i+} \cdot f_{+i}$$ Equation 3-10

Where f_{i+} is the total number of occurrences for the i^{th} row and f_{+i} is the total number of occurrences for the i^{th} column. The true positive rate (TPR) is defined by Equation 3-11, false negative rate (FNR) by Equation 3-12, false positive rate (FPR) by Equation 3-13 and true negative rate (TNR) by Equation 3-14. When more than two outcomes are used, rates can be calculated by collapsing the remaining states into a single one.

$$TPR = \frac{TP}{(TP + FN)}$$ Equation 3-11

$$FNR = \frac{FN}{(TP + FN)}$$ Equation 3-12

$$FPR = \frac{FP}{(FP + TN)}$$ Equation 3-13

$$TNR = \frac{TN}{(FP + TN)}$$ Equation 3-14

The receiver operating characteristic curve is used to analyse the relationship between the false positive rate (or 1-specificity) and the true positive rate (or sensitivity) for various decision thresholds (Figure 3-4). In a two-outcome case, this threshold is commonly assumed at 0.5. From this curve, the most adequate threshold can be selected considering the required false positive rate and true positive rate. The curve is also used to calculate the area under its curve or AUC. AUC ranges between 0 and 1, where 1 represents perfect matching, 0.5 reflects totally random models, and <0.5 indicates models generating predominantly inaccurate predictions (Korb & Nicholson, 2011). In the case of more than two outcomes, ROC and AUC can be computed by collapsing the remaining states in a single one.

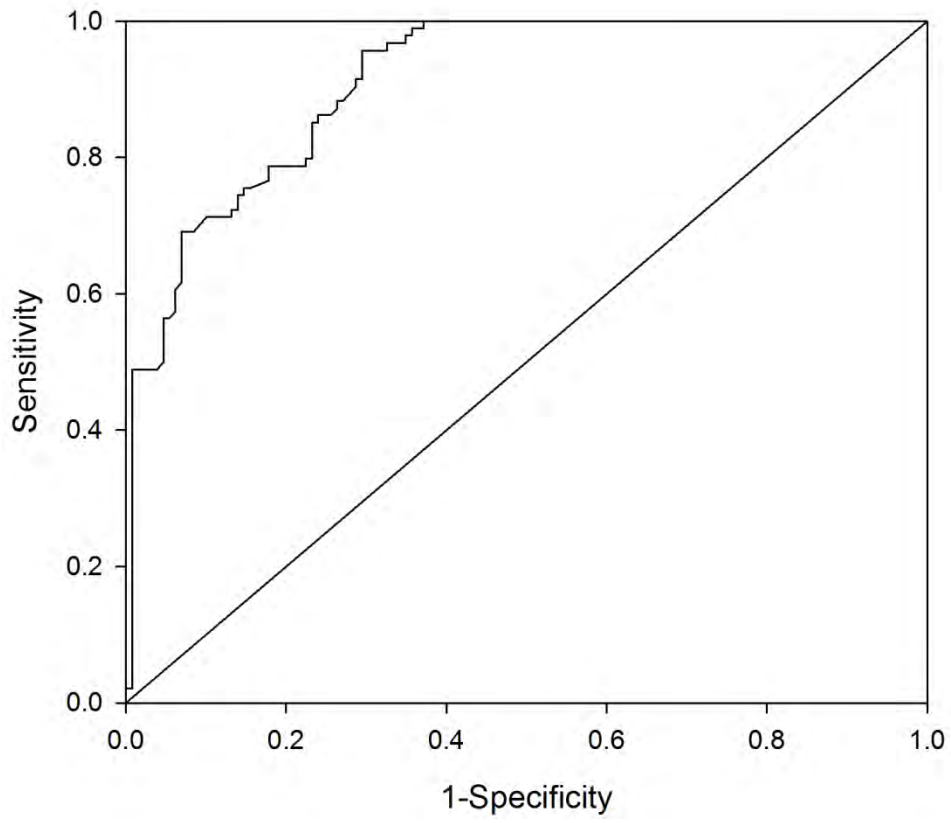

Figure 3-4: receiver operating curve. Diagonal line represents complete uncertainty.

3.1.4.3 Inference

Inference algorithms are very important because inference reasoning is one of the main advantages of BNs (Korb & Nicholson, 2011). Bayesian networks, as their name indicates, make use of the Bayes rule (Equation 3-15) to update the information in the network once evidence is provided. A number of inference algorithms have been developed to facilitate this process including Junction tree, loopy belief propagation, and likelihood weighting, with the most common being the Junction tree algorithm (Korb & Nicholson, 2011). Junction tree algorithm is an exact algorithm, whereas the other two are approximate. The advantage of using approximate methods is their capability for managing larger models for which Junction tree would require greater computational capabilities. These algorithms are commonly integrated in commercially available software packages. Furthermore, inference algorithms are freely available for their use in programming languages such as R, MatLAB and Python, therefore their development is not generally required. Explanation of the functioning of these algorithms can be found in Jensen and Nielsen (2007); Koller and Friedman (2009).

In a BN context, the marginal distributions observed in a BN when there is no incorporation of evidence are called prior probabilities. Once evidence is input to the model, the variables which were not observed will potentially have a change in their marginal probabilities. This new set of marginal probabilities is called posterior probabilities. Variables for which there is an interest for evaluating its results after some evidence is entered are called query nodes.

Using the example in Figure 3-1, inference in a BN can be illustrated. In this case, the Bayesian inference is used to determine the posterior distribution for "A" given an observation in "B" (B=High) (Equation 3-15). The posterior distribution for "A" P(A|B=High) given the observation B=High is obtained using the prior distribution for "A" P(A) and the conditional probability distribution P(B=High|A) specified in the model.

$$P(A|B = High) = \frac{P(B = High|A)P(A)}{P(B = High)} \qquad \text{Equation 3-15}$$

3.1.4.4 Entropy and mutual information

In BNs, the strength of association between a pair of variables can be quantified non-parametrically by mutual information. Mutual information (Equation 3-17) measures how much uncertainty or *entropy* (Equation 3-16) is reduced in a target variable when information is known about a second variable. Mutual information is usually used to compare which variables are more

important for a particular target variable given some evidence. Mutual information between two variables depends on both the structure and the parameters of the network.

$$H(X) = -\sum_{x} P(x) \cdot logP(x) \qquad \text{Equation 3-16}$$

$$I(X,Y) = -\sum_{x}\sum_{y} P(x,y) \cdot log\frac{P(x,y)}{P(x)P(y)} \qquad \text{Equation 3-17}$$

Where H(X) is the *entropy* of variable X, P(x) is the probability of the variable x. I(X,Y) represents the mutual information between X and Y, P(x,y) is the joint probability of the variables x and y, P(x) is the probability of x and P(y) is the probability of y.

Usually, mutual information is normalised to make results comparable between model alternatives. Relative mutual information for a target node is defined as the ratio between the mutual information of the target node and any other node, and the marginal *entropy* of the target node.

3.1.4.5 Difference and lift

Difference and *lift* are used to evaluate the change in the marginal probabilities of a node given some evidence. The *difference* is defined as the subtraction between the posterior and prior probabilities for a particular node state (Equation 3-18). *Lift* or normalised likelihood is defined as the ratio between the posterior and prior probabilities for a particular node state (Kjræulff & Madsen, 2012) (Equation 3-19).

$$Difference = P(x|\varepsilon) - P(x) \qquad \text{Equation 3-18}$$

$$Lift = \frac{P(x|\varepsilon)}{P(x)} \qquad \text{Equation 3-19}$$

Where x is the probability of the state for the target node and ε is the evidence.

3.1.4.6 Sensitivity analysis

Two main types of sensitivity analysis can be performed during the BN development including sensitivity to evidence (or sensitivity to findings) and sensitivity to parameters. Sensitivity to evidence can be conducted through the use of mutual information (or relative mutual information) (Section 3.1.4.4). An alternative to the use of mutual information is evaluating the marginal probabilities of a target node when evidence is input into the model. One or a group of nodes can be manipulated to analyse the effect on the target node (Korb & Nicholson, 2011). "Sensitivity to parameters" measures how sensitive the results of propagation of evidence are to the influence of changes in parameters (i.e. marginal or conditional probabilities associated to a node). Although it is possible to perform a sensitivity analysis to any number of parameters, it has been recommended for practical purposes that only one-way or at most two-way parameter sensitivity analysis is considered (Kjræulff & Madsen, 2012). In this section one-way and two-way sensitivity analysis are explained. Both types of sensitivity analysis are based on the observation that the probability of the evidence changes linearly to one or two parameters in the model. For the one-way analysis Equation 3-20 is used, while for two-way sensitivity analysis Equation 3-21 is used. In the one-way sensitivity equation, the single variable "t" is analysed, whereas for the two-way sensitivity equation, t_1 and t_2 are analysed. The parameters of the functions (i.e. α, β, γ, δ, α_1, α_2, β_1, $\beta 2$, $\gamma 1$, $\gamma 2$, $\delta 1$, $\delta 2$) are determined by propagating the evidence under different values of the parameter to be analysed. Further information about the derivation and use of sensitivity to parameter functions can be found in Jensen *et al.* (2002); Kjræulff and Madsen (2012)

$$f(t) = P(h|\varepsilon)(t) = \frac{P(h,\varepsilon)(t)}{P(\varepsilon)(t)} = \frac{\alpha \cdot t + \beta}{\gamma \cdot t + \delta} \qquad \text{Equation 3-20}$$

$$\begin{aligned} f(t_1, t_2) &= P(h|\varepsilon)(t_1, t_2) \\ &= \frac{P(h,\varepsilon)(t_1, t_2)}{P(\varepsilon)(t_1, t_2)} \\ &= \frac{\alpha_1 \cdot t_1 \cdot t_2 + \beta_1 \cdot t_1 + \gamma_1 \cdot t_2 + \delta_1}{\alpha_2 \cdot t_1 \cdot t_2 + \beta_2 \cdot t_1 + \gamma_2 \cdot t_2 + \delta_2} \end{aligned} \qquad \text{Equation 3-21}$$

Where "f()" is the sensitivity function, t_n represent the parameter(s) to be analysed for n parameters, h is the target node, and ε is the evidence. α, β, γ, and δ are the parameters of the one-way sensitivity function, while α_1, α_2, β_1, β_2, γ_1, γ_2, δ_1, and δ_2 are the parameters of the two-way sensitivity function.

3.1.5 Modelling software for BNs

A number of commercially available BN software packages exist as shown in Table 3-4. The main differences between them are the incorporation of continuous nodes, structure learning capabilities and the number of inference engines they have available. In this study four BN software packages were used, including Netica™ (Norsys, 2015), Bayes Server (Bayes Server, 2017), WEKA (Hall *et al.*, 2009), and bnlearn (Scutari, 2010). Netica™ provides a simple graphical interface to construct and evaluate BNs. Inference is performed through *junction tree* algorithm. Bayes Server provides a graphical interface to construct and analyse BNs having a number of options for sensitivity analysis. WEKA includes machine learning algorithms for data mining and provides various tools for data processing and evaluation of algorithm optimality (Witten & Frank, 2005). WEKA also has a module for BN construction and evaluation. However, WEKA is limited in terms of the algorithms for parameter learning (only MLE available) and flexibility in model construction. Bnlearn is a package available for the programming language R and can be used to construct, infer and evaluate BNs. This package contains a number of structure learning algorithms and also permits exportation of files compatible with Netica™.

3.1.6 Other types of Bayesian networks

Alternative BN modelling techniques have been developed to deal with some of the more challenging issues of BNs including continuous variables and explicit definition of correlations between variables. An approach that incorporates these characteristics is the non-parametric BNs. This technique uses conditional rank correlations to measure the associations between the variables and normal copulas to define the joint probability distributions. A software package, Uninet has been developed to construct and evaluate these types of models (Cooke *et al.*, 2007). Uninet offers a user interface and an Application Programming Interface (API) to model non-parametric BNs. The program supports both probabilistic and functional nodes. Probabilistic nodes can be defined from a collection of nine common parametric distributions or determined directly from data. Functional nodes can be the child of probabilistic or other functional nodes, but they cannot have probabilistic children. The user then specifies the DAG and the rank correlations either from data or expert knowledge. Stochastic modelling in Uninet is performed through Monte Carlo tools which samples the entire joint distributions. More information about this technique and software package can be found in Hanea *et al.* (2015).This model technique was used in Chapter 9 to model a multi-barrier system.

Table 3-4: Comparison of available features of a number of commonly used BN software packages which were used or provisionally trialled

Package	Continuous nodes	Parameter learning	Number of structure learning algorithms	Dynamic BNs	Number of engines	[a]Visualisation	Price (USD)	License period	[a]Support	Website
Netica		✓	1	✓	1	✓✓✓	685	Perpetual	✓	www.norsys.com
GeNIe		✓	>1	✓	>1	✓✓	Free	[b]NA	✓	www.bayesfusion.com
Hugin	✓	✓	>1	✓	1	✓✓	4,240	Perpetual	✓✓✓	www.hugin.com
BayesiaLab		✓	>1	✓	>1	✓✓	[c]24,620	Perpetual	✓✓	www.bayesia.com
AgenaRisk		✓			1	✓✓	3,000	1 year	✓✓✓	www.agenarisk.com
Bayes Server	✓	✓	>1	✓	>1	✓✓	623	Perpetual	✓✓✓	www.bayesserver.com
Weka	✓	✓	>1		1	✓	Free	NA	✓	www.cs.waikato.ac.nz/ml/weka/
Uninet	✓	✓			1	✓✓	Free	NA	✓✓	www.lighttwist.net/wp/uninet

[a] Based on own experience.

[b] Not applicable.

[c]: Other price options available depending on the period of time.

3.2 Bayesian models

Bayesian models are statistical models which allow expressing variables and parameters as probability distributions and update those probabilities once data become available. The fundamental concept behind Bayesian models is Bayesian inference.

3.2.1 Bayesian inference

Bayesian inference in Bayesian models is commonly used to learn about probabilities of unobservable parameters. In this case Bayes' theorem is employed in a slightly different fashion as in Bayesian networks as shown in Equation 3-22. **P(θ)** is the prior distribution for **θ** (i.e. vector of parameters) and represents the uncertainty about the values of **θ** before incorporating the observed data. **P(θ|y)** indicates the posterior distribution for **θ** and represents the uncertainty of **θ** after conditioning on the data **y**. The term **p(y|θ)** indicates how the data depend on the parameter values. **P(y)** is used as normalising factor and ensures that the posterior probability integrates to 1. Usually, this normalising factor is not required to be calculated, so the Bayes' theorem can be written as shown in Equation 3-23. Because **p(y|θ)** is actually a function of **θ** for a fixed y so this term can be expressed as **L(θ;y)** (Equation 3-24). **p(y|θ)** or **L(θ;y)** are generally known as the likelihood. Therefore, the Bayes' theorem can be stated as posterior ∝ likelihood x prior.

$$p(\theta|y) = \frac{p(y|\theta) \cdot p(\theta)}{p(y)} \qquad \text{Equation 3-22}$$

$$p(\theta|y) \propto p(y|\theta) \cdot p(\theta) \qquad \text{Equation 3-23}$$

$$p(\theta|y) \propto L(\theta; y) \cdot p(\theta) \qquad \text{Equation 3-24}$$

The updated posterior parameters can be used to make predictions about future outcomes ($\tilde{y}$) which are called posterior predictive distributions. Posterior predictive distributions are calculated using Equation 3-25.

$$p(\tilde{y}|y) = \int p(\tilde{y}|\,\theta) \cdot p(\theta|y) d\theta \qquad \text{Equation 3-25}$$

3.2.2 Markov Chain Monte Carlo

Calculation of posterior probability distributions in Bayesian models commonly requires solving complicated integrals. When solving these integrals, common numerical methods such as Monte Carlo integration are usually not effective and other more advanced techniques are required. One of these techniques is Markov Chain Monte Carlo (MCMC). In a Bayesian context, MCMC algorithms are a general class of computational methods used to generate samples from the posterior distribution. The objective of MCMC is to simulate values from a posterior distribution of a parameter vector. MCMC produces random walks over a probability distribution. When sufficient steps from the random walks are taken, various regions of the distribution are visited in proportion to their posterior probabilities. Once the samples obtained approximate to the target distribution, it is said that the sampled distribution has converged. Various MCMC algorithms have been created including the Metropolis algorithm, the Gibbs sampler, Hamiltonian method, and Metropolis-Hastings algorithm (Gelman *et al.*, 2014). From these algorithms, the Gibbs sampler is one of the most widely used to simulate Markov Chains (Lunn *et al.*, 2012).

3.2.2.1 Gibbs sampler

The success of MCMC algorithms usually depends on the choice of the proposal distribution which is an important step during the fine-tuning of the sampler. This proposal distribution is problem-specific and can become difficult to define. The Gibbs sampler is a particular case of the Metropolis-Hastings algorithm and facilitates the issue of the proposal distribution by avoiding its use (Ntzoufras, 2011). The Gibbs sampler splits the vector of multiple parameters into subvectors for each individual parameter, conditional on the remaining parameters (Lunn *et al.*, 2012). The advantage of Gibbs sampling is that simulation from a complex, high-dimensional joint posterior distribution is reduced to a sequence of algorithms for sampling from one- or low-dimensional distributions. This algorithm can be represented as follows (Lunn *et al.*, 2012).

"*Given a vector of unknown parameters* $\theta = (\theta_1, \theta_2, \dots, \theta_k)$*:*

1. *Choose arbitrary starting values* $\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)}$ *for each component, where subscripts denote sub-components of* θ *and superscripts denote the iteration number (iteration zero being the initial state of the Markov chain).*
2. *Sample new values for each element of* θ *by cycling through the following steps:*
    a. *Sample a new value for* θ_l*, from the full conditional distribution of* θ_l *given the most recent values of all other elements of* θ *and the data.*

$$\theta_1^{(1)} \sim p\left(\theta_1 \middle| \theta_2^{(0)}, \theta_3^{(0)}, \ldots, \theta_k^{(0)}, y\right)$$

b. *Sample a new value $\theta_2^{(1)}$ for the second component of θ, from its full conditional distribution $p\left(\theta_2 \middle| \theta_1^{(1)}, \theta_3^{(0)}, \ldots, \theta_k^{(0)}, y\right)$. Note that as a new value for θ_1 has already been sampled, it is this "most recent" value that is conditioned upon, together with the starting values for all other elements of θ."*

3.2.3 Modelling software for Bayesian models

A number of software packages have been created to facilitate Bayesian analysis. One of the most popular is WinBUGS which uses the Gibbs sampler to perform MCMC analysis (Lunn *et al.*, 2000). Other programs include OpenBUGS (Spiegelhalter *et al.*, 2007) and *JAGS*[4] (Plummer, 2013) which have been developed using similar algorithms and language syntax (i.e. BUGS[5] language). Stan (Stan-Development-Team, 2016) is another package which uses a different type of language and performs Hamiltonian MCMC. Additional improvements include the creation of packages to run the models through programming languages such as R and Python. Some of these packages include R2jags, rjags, R2Winbugs. This study used R2jags which provides an R software interface to *JAGS* including the same functionalities as the *JAGS* toolbox. In this work *JAGS* was employed because of the simplicity it provides for model definition and the capability it has for dealing with variable data formats.

3.2.4 Model construction in BUGS

By the use of *JAGS*, the model construction and evaluation is facilitated. BUGS language allows definition of likelihood functions and priors using a number of available parametric discrete and continuous probability distributions (Plummer, 2015). In the case of having to define a new type of probability distribution, techniques have been presented to effectively deal with this issue. A common BUGS model syntax to generate a linear model is presented in Figure 3-5. As observed, the model requires defining a likelihood function and prior distributions for the model parameters. The code includes a *for* loop which goes from 1 to N. In this case N represents the number of cases (i.e. data points) in a data file. After running the code the program updates the priors defined for each model parameter (i.e. alpha, beta and tau) using the data and generates posterior values for such parameters which are stored in a new data file. These data represent the results of the

[4] JAGS stands for Just Another Gibbs Sampler
[5] BUGS stands for Bayesian Inference Using Gibbs Sampler

model. In case posterior predictive distributions are required this can be obtained by using the posterior parameters to generate new predictive outcomes (y.pred in Figure 3-5).

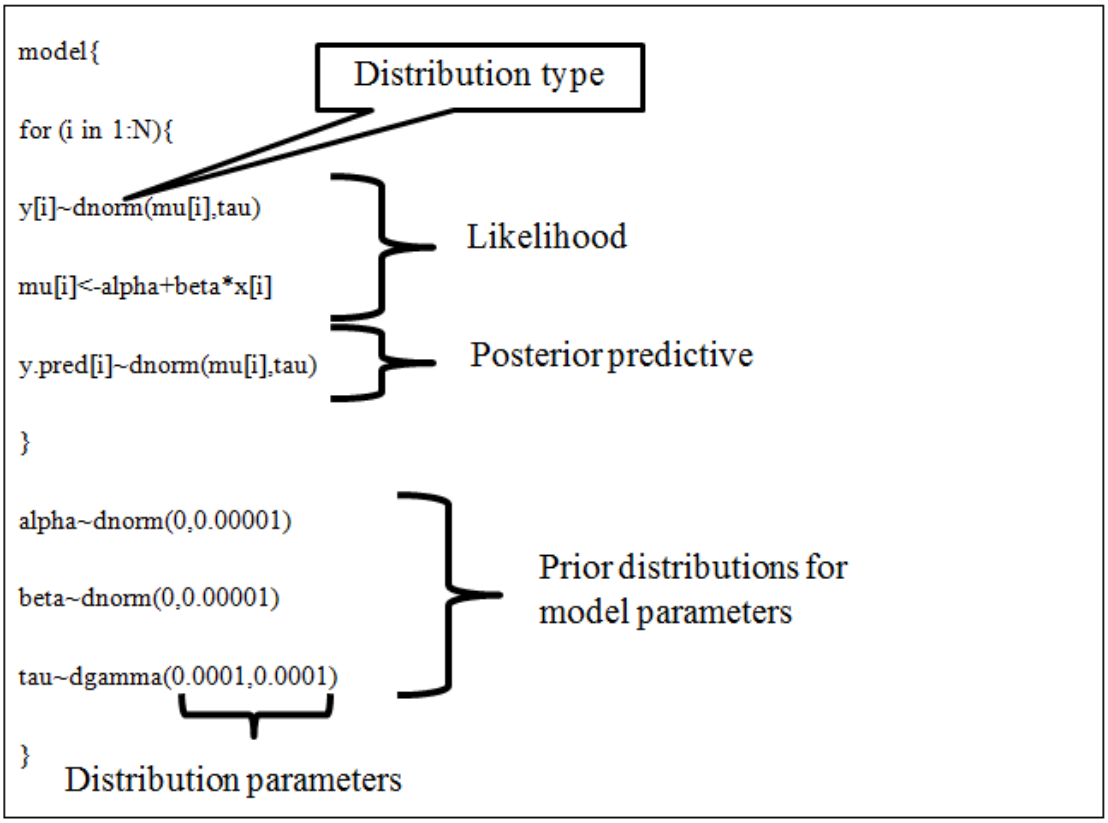

Figure 3-5: Example of model definition in BUGS language

3.2.4.1 Hierarchical models

Bayesian models are particularly suitable for a type of modelling technique known as hierarchical modelling. These models are useful to predict new instances when data are gathered about similar (but not equal) units (e.g. treatment process performance). Hierarchical models are used to represent models with multiple levels of parameters in which prior distributions have been also defined on the prior parameters associated with the likelihood parameters (Ntzoufras, 2011). Therefore, parameters for each particular unit are conditionally independent between them given "population" parameters. Hierarchical models can be considered as a step in between completely independent parameters (i.e. all units different) and identical parameters (i.e. all units equal). More information about hierarchical models can be found in Lunn *et al.* (2012); Gelman *et al.* (2014).

3.2.5 Model evaluation

This section describes the model evaluation tools used for checking the model correctness in terms of its convergence and performance. In the next sub-section, a number of convergence check tests are outlined including some not used in this thesis. Only the convergence tests used in this thesis are described. Performance metrics described next were all used in this thesis.

3.2.5.1 Convergence check

The success of an MCMC calculation needs to be checked to provide evidence that convergence has been achieved and that the results are representative of the posterior distribution with sufficient accuracy and stability. Various metrics and tests have been created to check the convergence of the chains in a MCMC including visual inspection of the chains (i.e. traceplots), level of autocorrelation and convergence tests (e.g. Gelman-Rubin diagnostic, Raftery-Lewis diagnostic, Heidelberger-Welch diagnostic, and Geweke convergence test) (Ntzoufras, 2011). In this thesis, four methods were used to check the convergence of the chains including visual inspection of the chains, level of autocorrelation, the Gelman-Rubin diagnostic, and the Geweke convergence test.

Visual inspection is performed by plotting the posterior parameters over the iteration number and checking that the traceplots overlap when more than one chain is used. If any of the chains varies only within a limited range of values or changes its values very gradually, this behaviour is taken as an indication of failure to convergence (Kruschke, 2014). Informally speaking, the chains have been described as looking like a "fat hairy caterpillar" (Lunn *et al.*, 2012) (Figure 3-6 a). Assessing autocorrelation in the chains is done by measuring the correlation within a chain for different lag values (Figure 3-6 b). The desired outcome for the autocorrelation is to obtain a value close to zero. The effective sample size can also be obtained from the autocorrelation coefficients by using Equation 3-26. The shrink factor or Gelman-Rubin diagnostic (Figure 3-6 c) comprises checking the convergence of two or more parallel chains. It performs the test using ANOVA type diagnostic, providing a shrink factor. Shrink factors close to one indicate convergence. The Geweke convergence test uses a single chain to test the convergence and applies a Z test to check whether the means of two different subsamples of the chain are equal (beginning and end of the chain). If the hypothesis for the two sample-means being equal does not fail to be rejected, convergence cannot be assumed (Ntzoufras, 2011) . More details about its calculation and use of the Gelman-Rubin and Geweke tests can be found in Ntzoufras (2011).

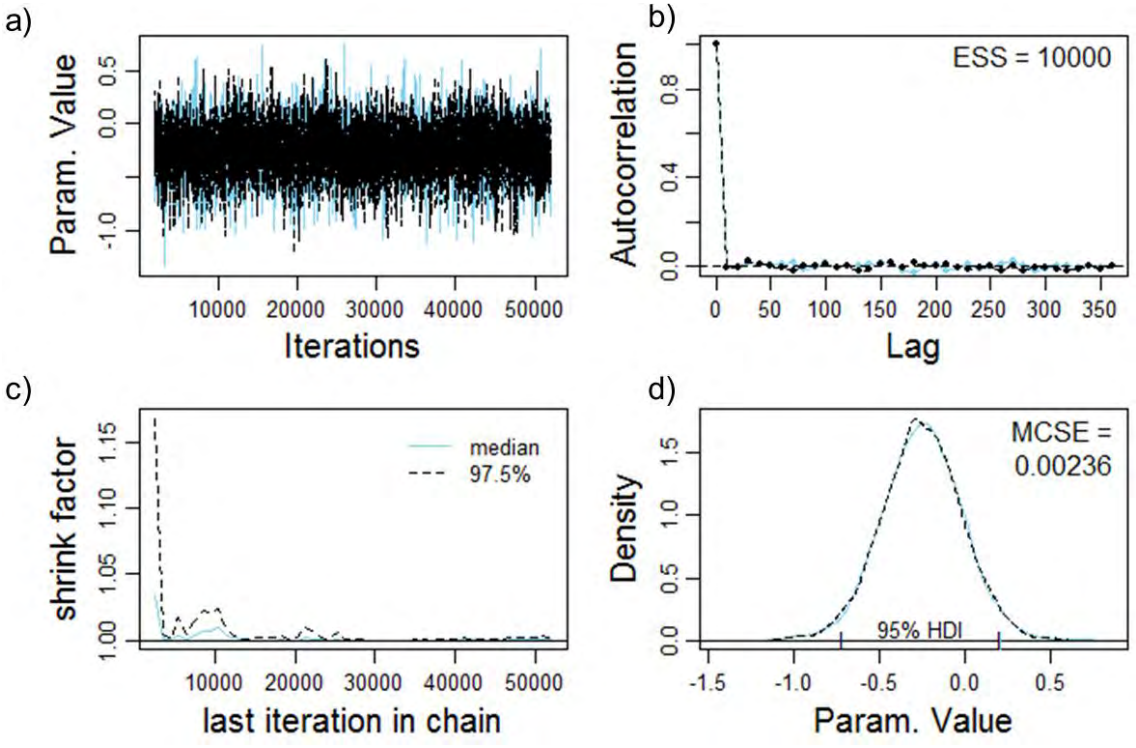

Figure 3-6: Diagnostics used for checking convergence including a) traceplots, b) autocorrelation, c) Gelman-Rubin statistic, d) comparison of densities between chains.

$$ESS = N/\left(1 + 2\sum_{k=1}^{\infty} ACF(k)\right) \qquad \text{Equation 3-26}$$

Where ACF is the autocorrelation factor for a k lag value and N is the number of iterations for the chain.

3.2.5.2 Performance metrics

In the same way as BNs are evaluated through the log-likelihood or AIC, Bayesian models can also be evaluated using similar approaches. A common metric used for Bayesian models is the Deviance Information Criterion (DIC) score which combines a measure of goodness of fit or mean deviance ($\bar{D}$) with a measure of complexity or penalty term, (pD or the effective number of parameters). The number of effective parameters term (pD) is used because the number of parameters in a Bayesian model cannot commonly be defined as in the case of BNs. Lower DIC values indicate a better, more parsimonious fit to the data. The deviance (D) is defined as presented in Equation 3-28.

$$DIC = \bar{D} + pD$$
$$= D(\bar{\theta}) + 2 \cdot pD$$

Equation 3-27

$$D(\theta) = -2log\ p(y|\theta)$$

Equation 3-28

$$pD = \bar{D} - D(\bar{\theta})$$

Equation 3-29

The models developed through Bayesian analysis have continuous outcomes requiring different types of metrics compared to BNs. The metrics used to compare the performance of the Bayesian analysis models in this thesis included adjusted determination coefficient (R^2), and mean squared error (MSE). To evaluate the models under instances not used during model construction, cross-validation (Section 3.1.4.1) was employed. $R^2_{adjusted}$ (Equation 3-30) ranges between 0 and 1 and describes the proportion of variation in the response that is explained by the predictors.

$$R^2_{adjusted} = 1 - (1 - R^2)\left[\frac{n-1}{n-(k-1)}\right]$$

Equation 3-30

Where R^2 is the coefficient of determination, n is the sample size and k is the number of independent variables in the regression equation. Because a stochastic result for $R^2_{adjusted}$ is obtained during Bayesian analysis, the average $R^2_{adjusted}$ under posterior was used to compare the models.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

Equation 3-31

To perform a proper comparison of the models' performances, the evaluated data were normalised (0-1) (Equation 3-32) before computing the metrics so that results could be compared without having the effect from data ranges.

$$x_{norm,i} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

Equation 3-32

Where $x_{norm,i}$ is the normalised observation i, x_i is the observation i, x_{min} is the minimum value for x, and x_{max} is the maximum value for x.

3.3 Developed chapters and applied methodologies

Six case studies were carefully selected to cover a range of validation and reliability topics using the methodologies explained above. Each case study and its corresponding chapter and topic is presented in Table 3-5. The range of processes covered by the case studies comprised activated sludge treatment (CS-1), ultrafiltration membranes (CS-2), ozonation (CS-3), chlorine disinfection (CS-4), UV disinfection (CS-5) and an advanced multi-barrier treatment system (CS-6). The selected topics were aligned with the areas of study covered in the National Validation Framework for Water Recycling (NatVal) project in Australia (Roser *et al.*, 2015; Robillot *et al.*, 2016).

Table 3-5: Cases studies their corresponding topics in this thesis

Case study, chapter	Topic
Case study 1 (CS-1), Chapter 4	Assessment of parameters affecting Log10 reduction values achieved by activated sludge treatment
Case study 2 (CS-2), Chapter 5	Validation of ultrafiltration processes using Bayesian analysis
Case study 3 (CS-3), Chapter 6	Parameters for ozonation performance assessment
Case study 4 (CS-4), Chapter 7	Probabilistic assessment of chlorination performance targets for secondary treated wastewater using Bayesian analysis
Case study 5 (CS-5), Chapter 8	Assessing reliability by using Fault tree analysis and Bayesian networks during UV disinfection
Case study 6 (CS-6), Chapter 9	Improving the quantification multi-barrier system Log Reduction Values using Bayesian networks

BNs were applied in case studies CS-1, CS5 and CS6. Specific methods applicable to BNs used in each case study are detailed in Table 3-6. Bayesian methods were used in case studies CS-2, CS-3 and CS-4. Particular methods relevant to each chapter are specified in Table 3-7.

Table 3-6: Chapters trialling BN models and applied methods

Used methods	CS-1	CS-5	CS-6
Automated structure learning	✓		
Elicited structure		✓	✓
Parameter learning	✓		
Elicited parameters		✓	✓
Validation of model with data	✓		
Sensitivity to evidence	✓	✓	✓
Sensitivity to parameters		✓	

Table 3-7: Chapters trialling Bayesian methods and applied methods

Used methods	CS-2	CS-3	CS-4
Hierarchical model	✓		✓
Non-hierarchical model	✓	✓	
Validation of model with data	✓	✓	✓

3.4 Unsuccessful application case studies

Four other case studies were planned and executed to analyse further applications of BNs and Bayesian methods. However, the results were unsuccessful in responding the proposed research questions. A brief description of the aim of the case studies is presented in Table 3-8. The issues encountered on these case studies and the lessons learned from them are discussed in the Conclusions (Chapter 10).

Table 3-8: Description of unsuccessful case studies

Case study	Description
Case study 7	Analyse the relationship between direct integrity test data and turbidity during a full-scale ultrafiltration using SCADA data.
Case study 8	Investigate the use of alternative predictors for LRV during full-scale ozonation of secondary treated sewage.
Case study 9	Analyse predictors for LRV during full-scale activated sludge treatment.
Case study 10	Model impacts on reverse osmosis treatment performance using elicited information.

Chapter 4: Assessment of parameters affecting Log_{10} reduction values achieved by activated sludge treatment

This chapter has been published in the following journal article:

Carvajal, G., Roser, D. J., Sisson, S. A., Keegan, A., & Khan, S. J. (2015). Modelling pathogen log_{10} reduction values achieved by activated sludge treatment using naïve and semi naïve Bayes network models. Water research, 85, 304-315. DOI: 10.1016/j.watres.2015.08.035

4.1 Introduction

Activated Sludge (AS) is widely employed at municipal wastewater treatment plants to achieve secondary treated effluent quality sufficient for environmental discharge or further treatment. The primary objective of AS is large reductions in biochemical oxygen demand (BOD_5). Concurrent nitrogen removal has also been targeted by AS. Key process control parameters include solids retention time (SRT), mixed liquor suspended solids (MLSS), hydraulic retention time (HRT) and temperature. Performance verification is focused on water quality parameters including BOD_5, chemical oxygen demand (COD) ammonium (NH_4^+), nitrite (NO_2^-) and nitrate (NO_3^-), total Kjeldahl nitrogen (TKN), alkalinity, pH, turbidity and total Suspended Solids (SS) (Metcalf & Eddy Inc. *et al.*, 2014).

Pathogen reduction has not generally been a key aim of AS. However, with increased interest in water reuse, there has been growing interest in understanding and optimising the performance of AS for the improvement of microbial water quality (Wen *et al.*, 2009).

Several studies have now reported qualitative and semi-quantitative relationships between AS operational performance monitoring data and pathogen LRVs (Stadterman *et al.*, 1995; Robertson *et al.*, 2000; Suwa & Suzuki, 2001). However, there is still no consistent methodology for quantitatively relating process performance parameter data to LRVs, such as when assigning and validating LRV credits. A likely reason is the limited degree to which the relationships between AS operational parameters and LRV outcomes have been defined. Developing models to predict LRVs from commonly monitoring parameters would allow assessment of when the system is operating reliably and define safe operational envelopes, which is essential for process validation.

This chapter explored the use of Bayesian networks (BNs) as tools for explaining, quantifying and predicting AS pathogen removal efficiency where a substantial operating and water quality parameter data set is available. In this study "naïve" and "semi-naïve" Bayes models (NB and SNB respectively) were constructed using *C. parvum* and *G. lamblia* LRVs as target variables. Models were evaluated for their predictive capacity using various performance metrics. The best models describing LRV variance in response to reactor operating conditions were then identified. Relevant AS operating and monitoring parameters to predict pathogen LRVs were evaluated. Finally, practical use and model interpretation of BNs were conducted.

4.2 Methods

The stepwise modelling approach developed for this study is summarized in Figure 4-1. This approach is expected to be generally applicable to data collected for other water treatment processes. A central aim was to assess the predictability of the class nodes/variables of interest (pathogen LRVs in the present case) based on water treatment system control parameters, water quality monitoring and derived parameters. This procedure provides a basis for constructing models systematically and selecting the best from those available after implementing step 4). Common terminology, acronyms and abbreviations used in this chapter are presented in Table 4-1. For convenience these terms are also reproduced in the Glossary. Bayesian networks were used for this study because they permit the analysis of multiple variables for prediction without the need of parametric assumptions. BN model construction and evaluation were introduced in Section 3.1.

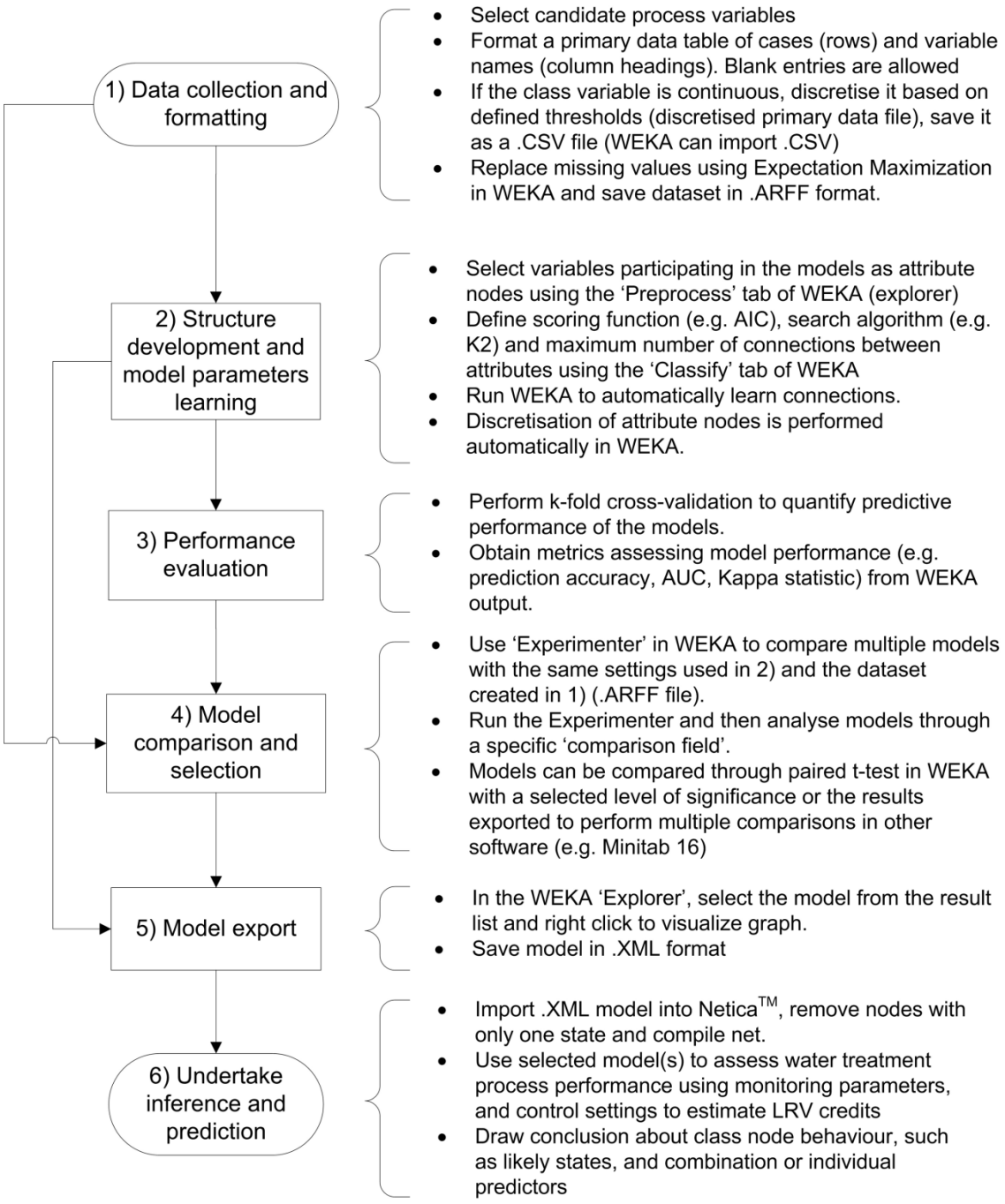

Figure 4-1: Flowchart for model development, evaluation and selection.

Table 4-1: Key Bayesian Network abbreviations and terminology relevant to model validation

Abbreviation /Acronym	Meaning	Explanation / Use /Comments	Reference
PA	Prediction Accuracy	Quantifies the number of correctly predicted values divided by the total number of cases.	(Witten & Frank, 2005)
KS	Kappa statistic	Measures the agreement between model predictions and actual values as a metric in the range [-1,1]. KS = 1 means perfect agreement, KS = 0 means that agreement is equal to chance, and KS = −1 means "perfect" disagreement.	(Marcot, 2012)
AUC	Area Under the Curve for the receiver operating characteristic curve	AUC ranges between 0 and 1, where 1 represents perfect matching, 0.5 reflects totally random models, and <0.5 indicates models generating predominantly inaccurate predictions.	(Korb & Nicholson, 2011)
LL	Log-Likelihood score	Measures how well the data fit each model. Used to compare models with the same variables and dataset but different node/arc structure. Higher scores reflect a better fit.	(Koller & Friedman, 2009)
TPR	True positive rate	Rate of correct positive predictions (high reductions).	
FPR	False positive rate	Failure to detect low reductions when they occurred.	
TNR	True negative rate	Rate of correct negative predictions (low reductions).	
FNR	False negative rate	Failure to detect high reductions when they occurred.	
NB	Naïve Bayesian Network	Bayesian network with a class node as the only parent of the remaining nodes.	

Abbreviation /Acronym	Meaning	Explanation / Use /Comments	Reference
SNB	Semi-Naïve Bayesian Network	Naïve Bayesian network in which attribute nodes are allowed to be connected one another.	
AIC	Akaike Information Criterion score,	Information-theoretic scoring function, which trades off the model's goodness of fit with its complexity.	(Kjræulff & Madsen, 2012)
ZeroR	-	Baseline model, it can be seen as a network without arcs.	
BAN	Bayesian network augmented naïve Bayes	Semi-naïve Bayes model. Two or more arcs between attributes are allowed.	
TAN	Tree Augmented naïve Bayes	Semi-naïve Bayes model. At most one arc between attributes is allowed.	
IDEA	Intermittently Decanted Extended Aeration	Semi-batch activated sludge reactor with cycles of influent feeding, aeration, decanting and effluent withdrawal.	
Attributes	-	Variables hypothesized as related to a class node in a NB and SNB models.	

4.2.1 AS system

The used data were obtained from a study of a 150 litre working volume Intermittently Decanted Extended Aeration (IDEA) AS pilot plant (Figure 4-2) (Flapper *et al.*, 2012). During that study many physicochemical, microbial and operational data were collected.

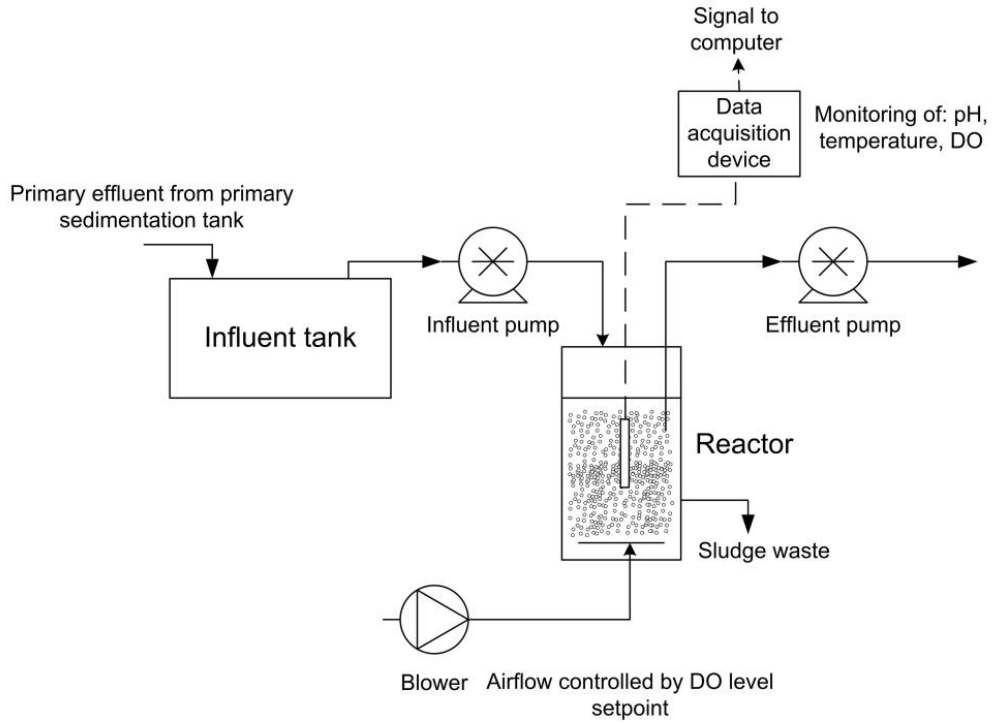

Figure 4-2: Scheme of the activated sludge reactor

Key features of the pilot plant study were as follows: Primary effluent was collected from a full scale wastewater treatment plant in Victoria, Australia and used as influent for the pilot plant. The reactor was operated in a three-stage cycle comprising: i) (top up) influent feed and aeration, ii) settling and iii) (partial) supernatant decanting. The reference operating conditions for the reactor were HRT=24 hours and SRT=15 days (three experimental runs). Additional operating conditions investigated were: HRT=24 hours and SRT= 10 days (three experimental runs); HRT=24 hours and SRT= 20 days (one experimental run); HRT= 7.5 hours and SRT=15 days (two experimental runs) (Flapper *et al.*, 2012). *C. parvum* oocysts were added to the influent tank (3.6 $\log_{10}$ oocysts/L) to ensure effluent concentration data were uncensored and sufficient for estimating LRVs. The mixed liquor dissolved oxygen (DO) concentration was maintained at 1.5 $mg.L^{-1}$ and the reactor was operated at 14.6-27.1 ^{o}C (Flapper *et al.*, 2012). Key operating and water quality parameters, measured or controlled in this study, included: three reactor operating parameters (SRT, MLSS, HRT), seven microbial water quality parameters (F-RNA bacteriophage, *E. coli*, Total coliforms, enterococci, *C. perfringens*, *G. lamblia*, *C. parvum*) and eleven physicochemical parameters (COD, BOD_5, NH_4^+, NO_2^-, NO_3^-, TKN, Alkalinity, pH, Turbidity, SS, Temperature). Pathogen and indicator LRVs were computed from temporally matched concentrations in the reactor inlet and outlet. A total of 98 records were available for the BN analysis. Unlike regression analyses which require complete data sets, BN construction was able to use records with missing values.

4.2.1.1 Modelling software

Four candidate models were constructed and evaluated using a variety of performance measures. The models were designed to quantify the influence of operating and water quality parameters on confirmed *C. parvum* and *G. lamblia* LRVs. Models were designed and evaluated using the Waikato Environment for Knowledge Analysis (WEKA) data mining software v. 3.6.11 (Hall *et al.*, 2009). Final model usage was performed in Netica™ Bayesian modelling software (Norsys, 2015). Models were also constructed without the naïve assumption using bnlearn package in *R* (Scutari, 2010). In this case, hill-climbing algorithm with Akaike Information Criterion (AIC) score and various arc restrictions were used as input before the structure was automatically learned.

The database was formatted to facilitate processing using WEKA and Netica™. The data were first compiled in a single spreadsheet table comprising records (rows) and variables (columns). For each model, initial WEKA processing then involved selection of a class node and its manual discretisation into 2 states. Because WEKA ignores missing values for the class node, records lacking *C. parvum* and *G. lamblia* data were removed when learning NB and SNB structures.

The final LRV datasets consisted of 88 and 75 records for *C. parvum* and *G. lamblia* respectively. The remaining data records still included some missing values for other reactor operational and water quality parameters: MLSS and Temperature (5-16% of records) and SS, pH, NO_3^-, TKN and COD (2-3% of records). The Expectation Maximization (EM) imputation method in WEKA was used to replace the missing values. EM uses a multivariate normal model to impute missing values. The reliability of WEKA was confirmed by also running the EM multiple imputation methods in AMELIA II package in R (Honaker *et al.*, 2011) which offers several options for data pre-processing.

The model designs generated by WEKA were exported as .XML files and imported into Netica™.

4.2.2 Model design

An NB model and three SNB (two TAN and one BAN) models were constructed for each of the two pathogens. WEKA's automated structure learning tool defined the arcs and nodes' states in the networks, using the dataset of Flapper et al. (2012). The two TAN model structures were developed by applying the Chow and Liu (1968) (TAN (1) model) and the K2 Hill Climbing (TAN (2) model) (Cooper & Herskovits, 1992) algorithms. K2 algorithm requires a fixed ordering of the variables in the dataset as input. The variable ordering and the explanation for such selection can be found in Appendix 1. The BAN model was constructed in the same manner as the TAN

(2) model but up to two nodes were allowed as parents in addition to the class node. In this approach, the learning processes are treated as optimisation problems where WEKA's search algorithms maximise a scoring function applicable to BNs, in this case, the AIC (Kjræulff & Madsen, 2012). The AIC scoring function trades off the model's goodness of fit with its complexity. "Structured learning" as employed by WEKA is designed to find the BN structure that best describes the statistical relationship between variables. The rationale for using structured learning was to develop models in a systematic and objective fashion. These models were also compared with their baseline model (ZeroR) equivalents. ZeroR models predict the mode (most repeated state) for a nominal class or the mean for a numeric class (Witten & Frank, 2005). When viewed as a Bayesian network a ZeroR model appears as a set of nodes without connections. So for example, in the case of *C. parvum,* the class node had two states (LRV<1 and LRV≥1) which were distributed 51.1% (LRV<1) and 48.9% (LRV≥1) and ZeroR would always predict LRV<1.

4.2.3 Model parameters, discretisation and learning

The most appropriate number of states for the remaining nodes (i.e. their discretisation) of each model was also determined by WEKA. WEKA optimises thresholds of the attributes based on the class variable (i.e. *C. parvum* and *G. lamblia* LRV node) using the minimum description length principle (Fayyad & Irani, 1993). Where only one state was defined for a variable, the corresponding node was concluded not to contribute to the classification process and was discarded from the NB model. The same nodes and states were also used by WEKA to define the TAN and BAN models.

Although Netica™ has a TAN learning wizard based on the Chow and Liu algorithm, the algorithm was not used to initially learn model structures because Netica™ did not permit comparison to other models and cross-validation. The final nodes and states of the two NB models are presented in Figure 4-3. The best SNB (TAN(2)) models are shown in Figure 4-4. The remaining SNB (TAN (1), BAN) models are presented in Appendix 1.

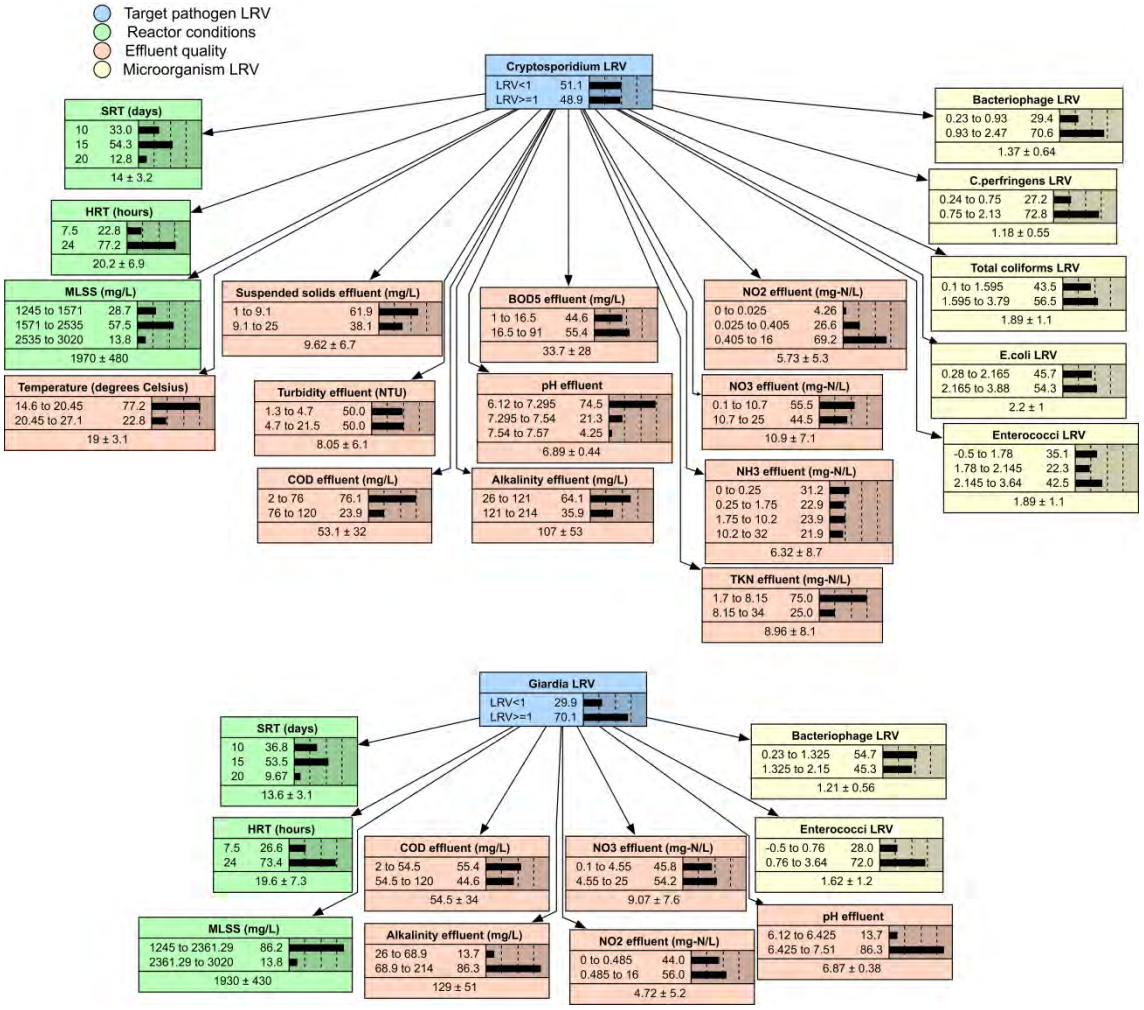

Figure 4-3: Naïve Bayes models for *C. parvum* LRV (top) and *G. lamblia* LRV (bottom) showing discretisation ranges.

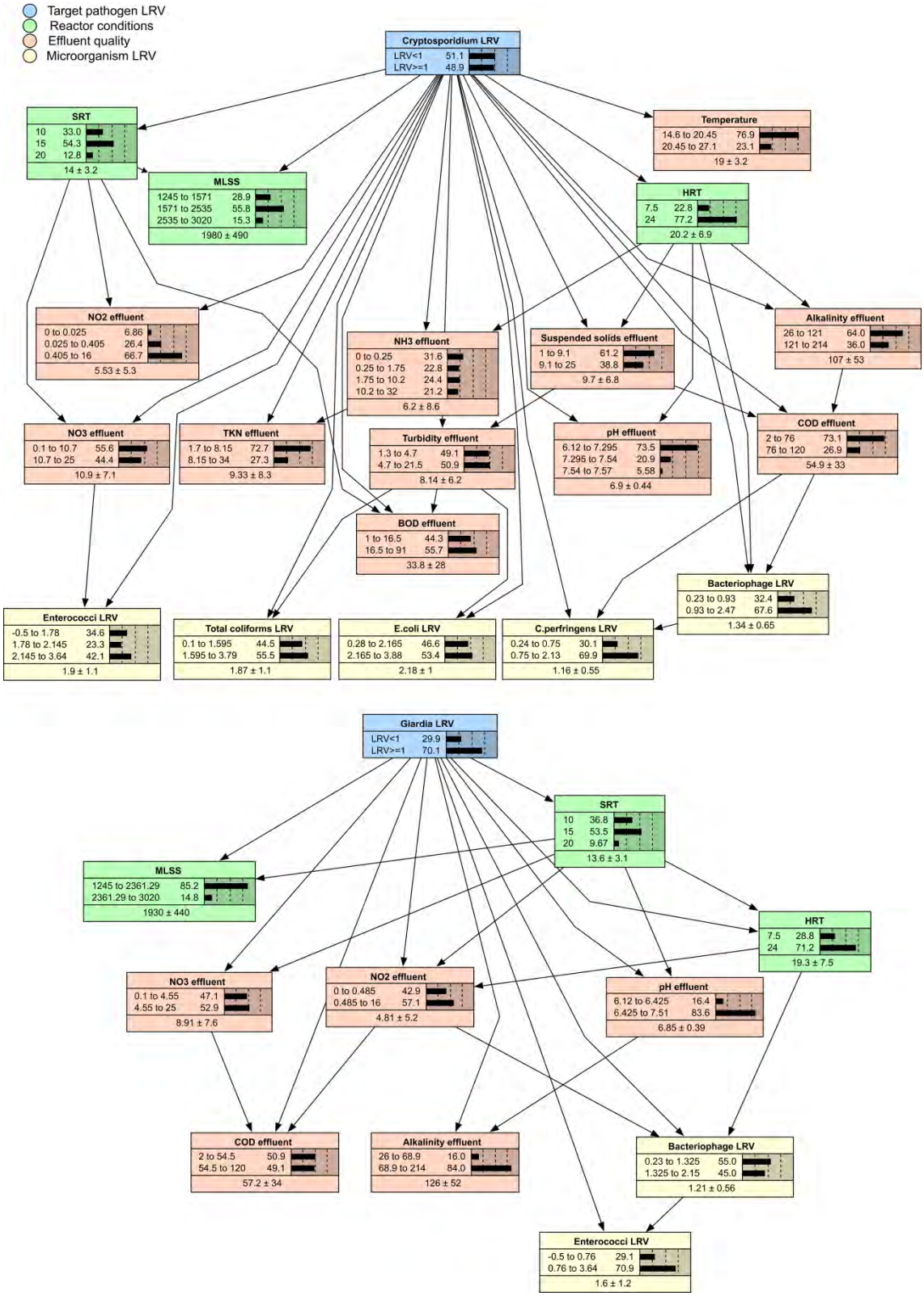

Figure 4-4: Optimum semi-Naïve Bayes net for *C. parvum* LRV (BAN) (top) and *G. lamblia* (BAN) (bottom).

4.2.4 Model evaluation and validation

Due to the low ratio of data records to nodes (*ca* 5:1 and 10:1 for *C. parvum* and *G. lamblia* respectively), stratified 10-fold cross-validation was performed to confirm model stability when undertaking validation using WEKA.

During the cross-validation test, WEKA updated the probabilities of the network with each case, except for the unobserved class nodes (LRV nodes), and then generated state probabilities for those nodes which were then compared against their actual values. The output of this analysis was a comparison of predicted and the real data (metrics in Table 4-2 and Table 4-3). Of the two possible LRV states the higher value was taken as the "positive" result for error calculation purposes. WEKA was also used to estimate prediction accuracy for LRV nodes using single (Table 4-3), and multiple node groups, such as the different coloured operational, control and monitoring groups in Figure 4-3.

Eight different performance gauging measures recommended for BNs were used to compare the NB, SNB and ZeroR models: i) 3 model prediction performance metrics - prediction accuracy (PA), Kappa statistic (KS), area under the curve (AUC) for the receiver operating characteristic curve; ii) one goodness of fit metric - log-likelihood score (LL); and iii) four error matrix metrics - true positive rate (TPR), false positive rate (FPR), true negative rate (TNR) and false negative rate (FNR) (Witten & Frank, 2005; Korb & Nicholson, 2011; Marcot, 2012). Ten-fold cross validation was performed ten times to assess the variation in the metric estimates from different data randomizations. A one-way sensitivity to findings analysis was also performed. This analysis consisted of assessing the effect that each variable had on a target variable and is presented in Appendix 1.

A statistical analysis of the performance metrics' results was performed by nonparametric methods in Minitab 16 (Minitab, 2010). The Kruskal-Wallis test was conducted to determine whether there was a significant difference among the five models' performance metrics medians. When the Kruskal-Wallis test indicated significant differences, Dunn's Test was used for the multiple comparisons (n=10) among the individual groups with a family alpha probability of 0.1 equivalent to an individual pairwise comparison alpha probability of 0.01 (type I error).

4.3 Results and discussion

4.3.1 Identifying the best models

A comparison of nine performance metrics is presented in Table 4-2 (for acronym description see Table 4-1). These were calculated using the 10-fold cross validation procedure for the four candidate models and the baseline ZeroR model. All four refined models performed significantly better (P. < 0.01) than the baseline ZeroR in almost all metrics with the exception of TNR and FPR for *C. parvum*, and TPR and FNR for *G. lamblia*. Because a threshold of 1 LRV splits the data in 51.1% (LRV<1) and 48.9% (LRV≥1) for *C. parvum*, ZeroR always predicted the removal to be LRV<1. This result meant that the testing cases where LRV<1 were always predicted correctly (TNR=1) and no false positives were obtained (FPR=0) (predicting LRV≥1 when testing cases LRV<1) but the true positive rate was 0%. An equivalent analysis was undertaken for *G. lamblia*. Excluding the LL score, no significant differences (P. >0.01) in performance were observed between the metrics for the *C. parvum* NB and three SNB models.

Table 4-2: Arithmetic mean ± standard deviation of performance measures from the 10-fold cross validation for *C. parvum* and *G. lamblia* for the naïve and semi-naïve models.

Pathogen	Performance measure	NB	TAN (1)	TAN (2)	BAN	ZeroR
C. parvum	PA	93.2±7.90[a]	88.6±8.79	91.1±8.26	91.5±8.24	51.1±4.87
	KS	0.86±0.16	0.77±0.18	0.82±0.17	0.83±0.16	0.00±0.00
	AUC[b]	0.96±0.07	0.95±0.08	0.95±0.09	0.95±0.08	0.52±0.15
	LL[c]	-755±42	-564±43	-583±44	-582±43	-1131±52
	TPR[b]	0.98±0.07	0.90±0.13	0.95±0.10	0.96±0.10	0.00±0.00
	FNR[b]	0.02±0.07	0.09±0.13	0.05±0.10	0.04±0.10	1.00±0.00
	TNR[b]	0.89±0.14	0.87±0.15	0.88±0.15	0.88±0.14	1.00±0.00
	FPR[b]	0.11±0.14	0.13±0.15	0.12±0.15	0.12±0.14	0.00±0.00
G. lamblia	PA	81.0±12.6	83.2±13.5	82.6±12.7	84.4±12.9	70.7±4.41
	KS	0.54±0.33	0.59±0.33	0.57±0.32	0.60±0.33	0.00±0.00
	AUC	0.87±0.17	0.86±0.18	0.87±0.17	0.86±0.17	0.35±0.19

Pathogen	Performance measure	NB	TAN (1)	TAN (2)	BAN	ZeroR
	LL	-413±60	-345±37	-350±39	-347±38	-491±76
	TPR	0.85±0.14	0.88±0.16	0.88±0.15	0.91±0.14	1.00±0.00
	FNR	0.15±0.14	0.12±0.16	0.12±0.15	0.09±0.14	0.00±0.00
	TNR	0.72±0.31	0.72±0.32	0.70±0.32	0.70±0.32	0.00±0.00
	FPR	0.28±0.31	0.28±0.32	0.30±0.32	0.30±0.32	1.00±0.00

[a] Standard errors were calculated considering a sample size n=100

[b]AUC and rates were computed considering LRV≥1 as the target range

[c] Log-likelihood score (LL) was computed from the complete dataset, therefore no std. error was estimated

The AUC metric for both pathogens indicated good classification and adequate prediction performances with a high ratio of true positive to false positive results. The fourth metric, log likelihood (LL), indicated the NB networks were significantly inferior to the semi-naive models ($P<0.01$). SNB models for both pathogens showed no significant differences in the LL score ($P<0.01$), indicating similar fit to the data.

Overall, these metrics indicated a significant improvement in the prediction results in the NB and SNB models over the ZeroR model for both pathogens and the SNBs over their NB equivalents based on the LL scores. However, the *G. lamblia* model metrics showed ≈2 times greater variability (standard deviation) compared to the *C. parvum* models.

The final 4 metrics measured True negative rate (TNR), False Positive Rate (FPR) (Type I error), True Positive Rate (TPR) and False Negative Rate (FNR) (Type II error). All *C. parvum* NB and SNB models and the *G. lamblia* BAN model especially predicted reductions very well when they occurred (TPR metric). However, the FPR metric (crediting a plant with a LRV≥1 when the opposite occurred) was nearly 3 times greater in the case of *G. lamblia* (FPR=0.28-0.30).

Another criterion for comparing SNBs was whether the network structures were causally valid and logical. Though the metrics were comparable, WEKA created a *C. parvum* TAN (1) model which included illogical arcs which were absent from the TAN (2) and BAN models, for example ammonia controlling HRT and bacteriophage LRV controlling temperature (see Appendix 1).

Similarly, there was an illogical arc from SRT to HRT in the three *G. lamblia* models (see Appendix 1). Supporting the conclusion, they provided comparable descriptions, the TAN (2) and BAN models had similar structures for both pathogens (see Appendix 1), possibly due to them using the same search algorithm and score function (the AIC).

Increased uncertainty was observed in the attribute nodes' probability distributions when these were connected to each other in the SNBs. This uncertainty was reflected in the attribute nodes conditional probability tables as uniform distributions for combinations of parent node states not found in the data. This behaviour was expected as the dataset was limited in size but was not seen as significant problem sinc these networks were primarily designed to estimate classification LRVs and not estimate other parameters given these LRVs.

WEKA also allowed generating "learning curves" for the two NB models by sequentially adding or removing 10% of the data during model construction and testing. During BN learning, the average KS (prediction agreement) metric for *C. parvum* remained stable at 0.86 once more than 70 percent of the training data were incorporated and was 0.8 even when only 20 percent of the data had been incorporated. Similarly, for *G. lamblia*, the KS statistic plateaued at 0.54 once 80 percent of the data had been incorporated. These stable plateaus indicated the data sets were sufficiently large for predicting the correct LRV range and obtaining models which were as accurate as possible given the data available.

The results of the model development and evaluation, using five different imputed datasets, indicated that the variation in the model performance was negligible (<1% difference) for both pathogens. This result meant that the missing value imputation method did not significantly affect the measured performance of the models. The acceptable proportion of missing values in a dataset will depend on the specific context, including the degree of correlation between the variables. To quantify the influence of different proportions of missing values on model performance, the dataset was split into training (80%) and testing (20%) datasets and then a percentage of values was randomly removed from the training dataset. The effect on performance was assessed using AUC scores. The naïve Bayes model for *C. parvum* returned an AUC score of 0.95 or higher provided less than 30% of values were missing. It was concluded that the number of missing values in the actual data sets was insufficient to substantially influence the final model performance.

Overall, it was concluded that for *C. parvum,* all NB and SNB models performed similarly for most of the metrics in Table 4-2. Consequently, LL and qualitative model assessment (logical structure) was used in this case to discriminate between the models. The *C. parvum* TAN(2) model

(Figure 4-4) achieved similar LL to the other SNBs (Table 4-2), but unlike TAN(1) its structure provided more insights into the system's behaviour (see Appendix 1). On the other hand, it is possible that BAN model may have been overfitted by the number of permitted connections. Thus TAN(2) was selected as best for predicting *C. parvum* LRVs. For *G. lamblia*, the same results as for *C. parvum* were obtained when comparing the models. The TAN(2) (Figure 4-4) model was also selected as the best model for *G. lamblia*.

4.3.2 Operational control and monitoring parameters as predictors of protozoan LRVs

As well as informing overall model performance, the PA and FPR metrics can be used to assess the predictive capacity of individual nodes alone (Table 4-3).

Table 4-3: Individual attributes evaluation through AUC score (mean±standard deviation) for *C. parvum* and *G. lamblia*.

Predictor	***C. parvum***	***G. lamblia***
Baseline[b]	0.52±0.15[a]	0.35±0.19[a]
SRT	**0.89±0.09**	0.76±0.17
HRT	0.71±0.13	0.55±0.15
MLSS	0.75±0.13	0.67±0.17
Temperature	0.61±0.11	-[c]
SS	0.85±0.10	-
Turbidity	**0.90±0.10**	-
COD	0.72±0.13	0.78±0.13
BOD_5	0.81±0.13	-
pH	0.58±0.11	0.71±0.16
Alkalinity	0.84±0.11	0.70±0.16
NO_2	0.70±0.13	0.81±0.14
NO_3	**0.89±0.09**	0.66±0.13
NH_4^+	0.81±0.15	-
TKN	0.71±0.13	-

Predictor	*C. parvum*	*G. lamblia*
Bacteriophage LRV	0.71±0.13	0.66±0.14
C. perfringens LRV	0.70±0.13	-
Total coliforms LRV	**0.91±0.09**	-
E. coli LRV	**0.91±0.10**	-
Enterococci LRV	**0.90±0.10**	0.66±0.10

[a] mean AUC based on 10-fold cross validation repeated 10 times

[b] "No evidence" no attributes are considered in the evaluation (Zero R model)

[C] attribute was not included in the NB or SNBs

Bolded values indicate good predictor nodes.

In the case of *C. parvum* there were several instances where AUC scores were comparable to those obtained using all nodes (Table 4-2). Turbidity, enterococci, *E. coli* LRV and total coliforms LRV returned high AUC scores (≥0.9). SRT, SS and nitrate also generated a high (0.85-0.9) scores. SRT association with LRV indicated manipulating this variable might be used to maximise LRVs.

For *G. lamblia*, none of the attributes obtained an AUC > 0.9. The two highest scores were achieved with nitrite (AUC=0.81), COD (AUC=0.78), and SRT (AUC=0.76). These results were consistent with the poorer overall *G. lamblia* model performance.

The predictive potential of groups of reactor and physicochemical parameters, and microbial indicator LRVs was assessed by selecting only the variables in such groups (coloured groups in Figure 4-3) during the model construction phase. For *C. parvum*, the three reactor settings, SRT, HRT and MLSS together achieved an AUC of 0.93±0.07. Similar predictive power was obtained using a combination of all five microbial indicators (AUC=0.95±0.07) and the eleven physicochemical water quality parameters together (AUC=0.94±0.10).

Disappointingly, for *G. lamblia* the three reactor parameters together were much less accurate in their prediction potential (AUC=0.67±0.17) while the two microbial indicators combined yielded only a slightly better AUC of 0.71±0.13. All five physicochemical parameters (Figure 4-3) provided a prediction comparable to the complete model (AUC=0.84±0.18). This performance evaluation showed that it is possible to predict the removal of *G. lamblia* for a threshold of 1 LRV

under the given set of operating conditions. However, the performance metrics were not as good as the ones obtained for *C. parvum*. Moreover, it was not possible to obtain single operating or water quality parameters with an average AUC higher than 0.80, except from nitrite (AUC=0.81). It was concluded that operating parameters included in this study were not good indicators of the *G. lamblia* removal mechanisms in the activated sludge system. This weak relationship could also be observed in the scatterplots which do not evidence trends or clustering as in the case of *C. parvum* and indigenous microbial indicators (Figure 4-5). A possible reason is high *G. lamblia* input concentrations variability compared to *C. parvum.* This is discussed further below.

Finally, PA and FPR were also calculated for all 171 possible combinations of two predictors for *C. parvum* and 45 for *G. lamblia*. Some high accuracy combinations were identified. However, their usefulness was unclear. Because of the relatively small number of input data records high accuracy could have occurred by chance so the results are shown for information only in Appendix 1.

The absence of useful microbial indicator parameters for predicting of *G. lamblia* LRVs reflected the weak association between *G. lamblia* and indicator LRVs generally (Figure 4-5). The correlation between the variables can be more clearly seen through the locally weighted Kernel smoothers included on each scatterplot (Figure 4-5). The randomness of the scatter in all *G. lamblia* LRV plots is consistent with low model prediction power, for example, between bacteriophage and enterococci LRVs. Conversely, the clear correlation between *C. parvum* and *E. coli* LRVs (as well as enterococci and total coliforms – not shown) is also evident. This outcome is consistent with the high predictive power of the full *C. parvum* models (Table 4-2).

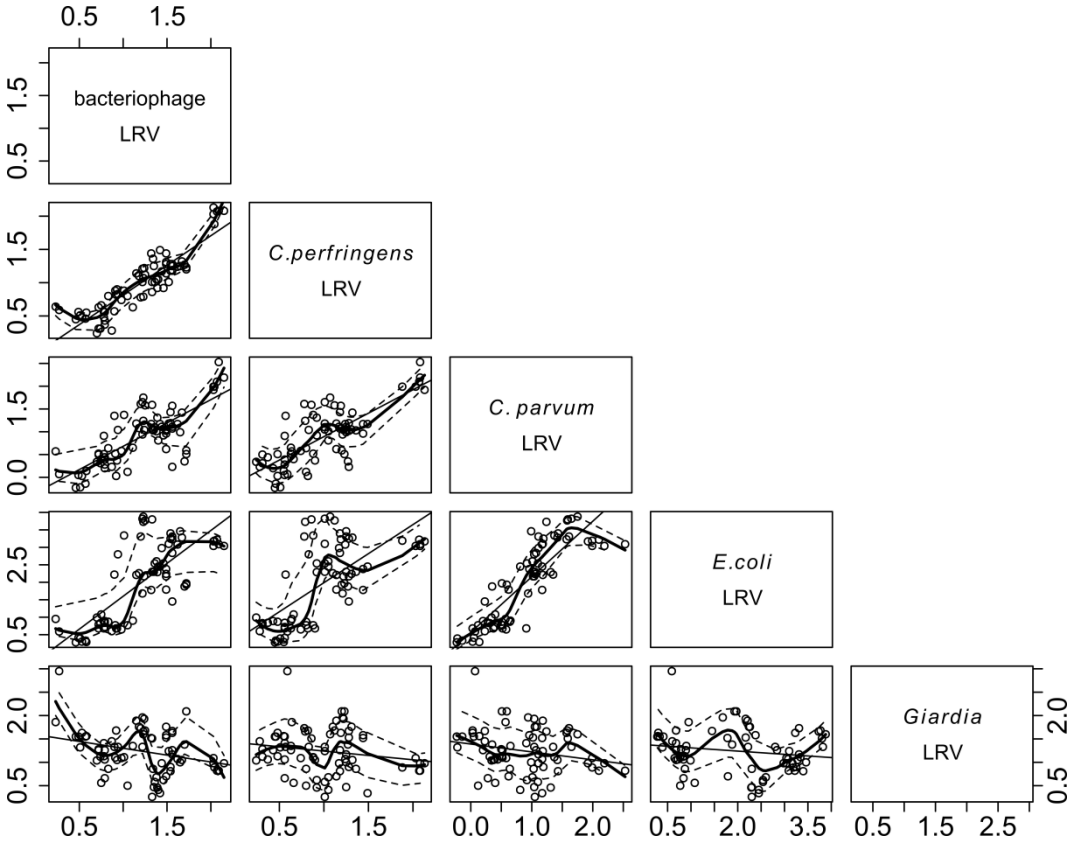

Figure 4-5: Scatterplot matrix for bacteriophage LRV, *C. perfringens* LRV, *E. coli* LRV, *C. parvum* LRV and *G. lamblia* LRV. Graph includes a linear regression fit, smoothers and their 95% confidence intervals.

4.4 Estimating log_{10} reduction credits of protozoan pathogens for activated sludge

4.4.1 Use of semi-Naïve Bayes Net models to estimate log credits

Both the optimized TAN(2) models and the presented modelling approach have potential applications and implications for setting AS operational and monitoring parameters and predicting protozoan pathogen reductions for the selected ranges. The *C. parvum* TAN(2) model quantified i) how operational, physicochemical microbial indicators related to removal and process settings for maintaining a removal range given by the model, and ii) what log_{10} credit might be assigned where AS is optimised for BOD_5 and nitrogen removal instead.

The model can be used to determine when the process is operating reliably for reducing pathogen numbers i.e. when the conditions for which the LRV≥1. These conditions are not necessarily the

maximum pathogen reductions the system could achieve, but the performance that the model is able to reliably predict based on the available data.

By contrast the *G. lamblia* model and metrics indicated that its reduction is less well understood, and operating and monitoring parameters cannot as yet be tuned to optimize *G. lamblia* removal. That being said the average LRV was ≥1 indicating log credit assignment is still possible for AS even though indirect monitoring and removal optimization is not yet possible.

The *C. parvum* AUC score also suggested that a high degree of LRV prediction was possible using only one monitoring parameter including total coliform LRV, *E. coli* LRV, enterococci LRV, and turbidity. Most usefully this list includes a real-time predictor, turbidity.

The achievement of ≈1 $\log_{10}$ removal for *G. lamblia* or *C. parvum* was not as striking as the >3 LRV reductions achieved with purpose designed disinfection agents. However, the result was robust. Not only did the model show the reduction was real but the prediction metrics confirmed the model reflected real trends in monitoring variables and were not the result of overfitting. This outcome indicates that in the future, AS systems may be further optimised for improved pathogen removal, and the techniques described in this study will be suitable for demonstrating any increased pathogen reduction. Separately this study demonstrates how robust treatment targets can be robustly estimated for other novel or unconventional disinfection and contaminant treatment processes.

The optimum discretisation thresholds and sizes of datasets to obtain high levels of model performance (e.g. prediction accuracy > 0.9) will depend on the specific model, process, characteristics of the process, the number of nodes and dataset. Such data set characterization is undertaken as part of the initial data mining by statistically experimenting with the data set and subsets on a case by case basis. The option for estimating minimum data set size adopted in this case was the use of "learning curves" (Frank *et al.*, 2000) where prediction performance metrics are determined for increasing sample sizes until a stable plateau is reached. It was found that > 61 data records were required to obtain stable performance for both the simplest *C. parvum* and *G. lamblia* (NB) models Other machine learning and data mining techniques (e.g. decision trees) can also inform on whether datasets are sufficiently large for models to be robust (Witten & Frank, 2005).

Node state discretisation threshold, though was necessarily defined arbitrarily. It was required to consider the influence of between state boundaries on prediction accuracy, while ensuring there were sufficient records corresponding to each state to permit to estimate the performance credibly. Because of this issue, state thresholds currently need to be defined empirically and selecting a

threshold where almost all observations are allocated to one state must be avoided if possible. In this study the impact of the protozoan thresholds being 1.0, 1.5 and 2.0 LRV was also assessed. The AUC scores indicated that all models maintained good performance. However, PA was not adequate for LRVs of 1.5 and 2.0, as these returned negligible improvements over their equivalent ZeroR models. Put another way, it could be possible to construct models able to predict higher LRVs using the higher thresholds, however, the states to be predicted became too "unbalanced", and the models tended to over-predict the majority class.

Using the "Enter findings" Netica™ function the most likely conditions (posterior probabilities) when LRVs were ≥1 were identified. This analysis was visualised in a tornado chart (see Appendix 1). Consistent with expectations, for *C. parvum* LRVs ≥1 were obtained when turbidity, SS, COD, TKN and alkalinity were in their lower ranges (P > 0.9) and microbial indicator LRVs (*E. coli*, enterococci, FRNA bacteriophage, *C. perfringens* and total coliforms) were in their higher ranges (P. > 0.9). The optimal reactor parameter settings (SRT, MLSS and HRT) when *C. parvum* LRV was ≥1 and BOD_5 and ammonia were - MLSS (1140-1571 mg/L), lowest SRT (10 days), and higher HRT (24 hours), their lowest, lowest and higher ranges respectively.

Recognizing the limitations of the *G. lamblia* model, it also noted that the conditions most associated with *G. lamblia* LRV ≥1 were high range pH (P > 0.95) and alkalinity (P > 0.96) and low range MLSS (P = 0.73). But counterintuitively these conditions were also associated with lower bacteriophage and enterococci LRVs and higher range COD. In light of these puzzling results and the poor model performance it was concluded that *G. lamblia* reduction needs further investigation.

Few studies have investigated the association between protozoa LRVs by AS and operational and microbial indicator variables. However, comparison with a recent literature review indicated that the BN model pathogen and indicator LRVs were consistent with other research and protozoan removal is inversely correlated with SRT, effluent organic carbon and effluent SS (Flapper *et al.*, 2010). A constraint on the *C. parvum* model's value is that LRVs are more difficult to measure than concentrations as both influent and effluent data are required. Accordingly, the use of effluent indicator concentrations in place of LRVs was also evaluated and similar trends were observed.

The reproducibility of the *C. parvum* data was likely enhanced by seeding of standardized *C. parvum* oocysts into the influent of the test reactor at 3.6 $\log_{10}$ oocysts/L. The results of seeding these microorganisms contrast with the highly variable oocyst numbers and biotypes that would normally be encountered in wastewater influent but was essential within the earlier study to ensure

valid LRVs would be established. *C. parvum* numbers in real wastewater influent are much more variable ranging from 10^1-10^4 oocysts/100L (Harwood *et al.*, 2005). Conceivably this increase and stability of oocyst numbers could have improved the model by reducing this source of variance in the LRV estimates and accounted for the contrasting performance of the *C. parvum* and *G. lamblia* models. This said *G. lamblia* cysts are generally present at higher numbers than *C. parvum* (10^4-$10^{5.5}$ cyst per 100L in influent) (Harwood *et al.*, 2005) and vary less between seasons.

4.4.2 Finding relationships without the naïve assumption

A model for *C. parvum* was constructed without forcing the naïve assumption during the structure learning. The final network (Figure 4-7) could be grouped into four groups of variables slightly different from the groups defined in Figure 4-3. The outcomes could be interpreted as the reactor setting and sludge environment as distinct causes of the removal performance for both indicators and pathogens. Variables indicative of these two groups were assumed to determine or reflect whether the system achieved high or low removal of indicators and pathogens.

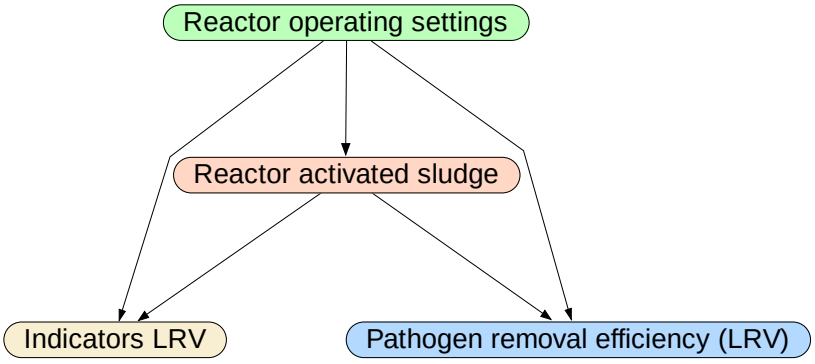

Figure 4-6: Relationships between groups of variables for "causal" model.

The BN in Figure 4-7 shows the specific arcs found by the *hill-climbing* algorithm through bnlearn (see Appendix 1). As observed, the activated sludge conditions and reactor operating settings influence all the microbial LRVs. Because the pathogen LRV is a child node in this case (treated not as a target as in the naïve case), more diverse types of queries can be explored (e.g. influence on other microbial indicators). The arcs determined from the structure learning provided insights about the most important variables for each variable.

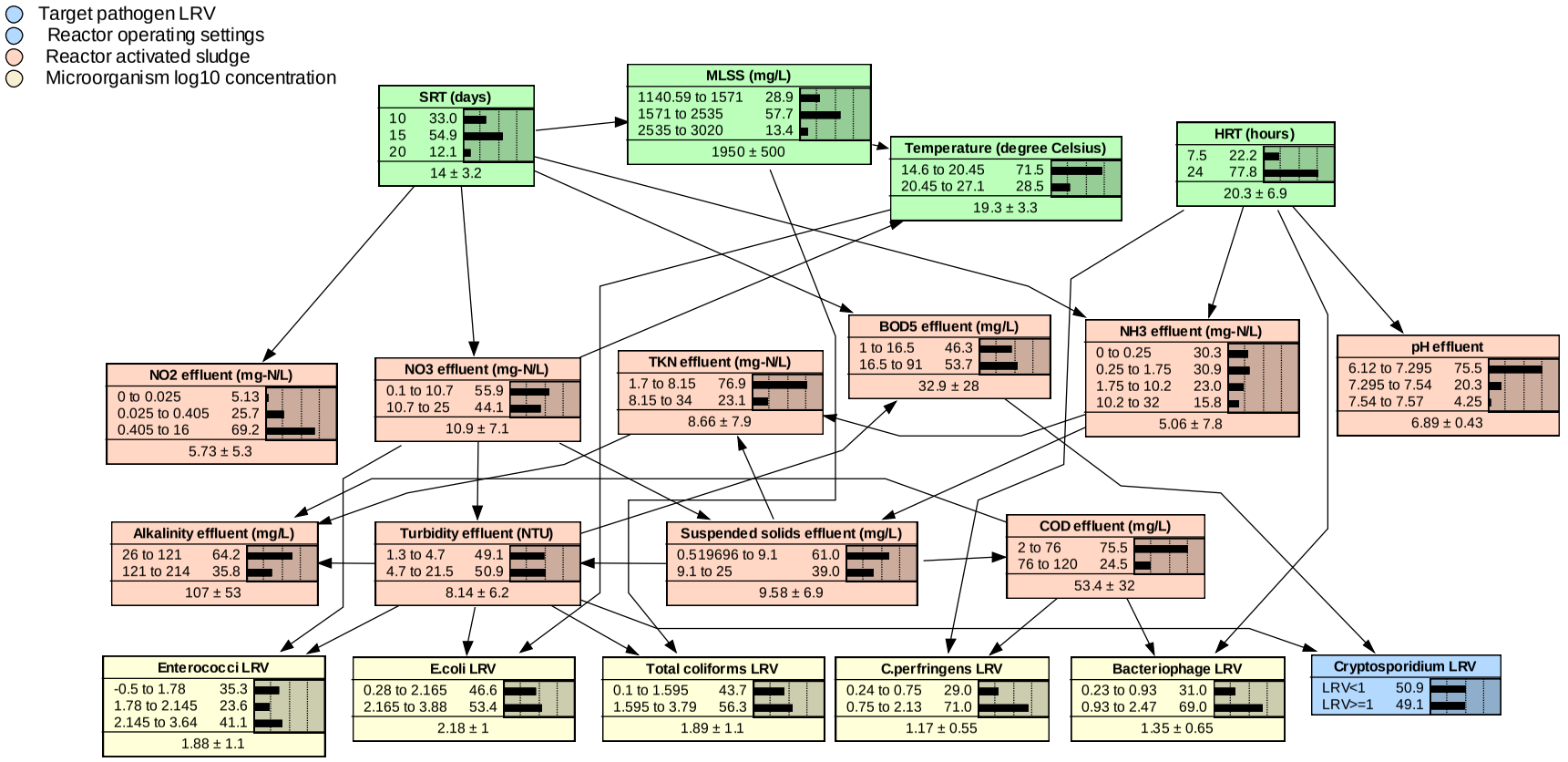

Figure 4-7: BN model for *C. perfringens* using hill climbing algorithm

4.4.3 Sensitivity analysis for model without naïve assumption

One way sensitivity analysis was performed to identify the most influential variables for *C. parvum* LRV through relative mutual information (Figure 4-8). In general, sensitivity analysis is directly influenced by the structure of the network and the estimated parameters. A node which is far away from a target node having many intermediate nodes will have less effect than a node which is directly connected. In this case, for example turbidity has an arc to *C. parvum* LRV, whereas pH, the least influential, is only indirectly connected. In this model, *C. parvum* LRV can be potentially influenced by any node given this particular structure. But the results of the sensitivity analysis clearly identify the most useful monitoring parameters. The most influential variables were also found to be the most relevant for the prediction of *C. parvum* according to the results presented in Table 4-3 for the naïve models.

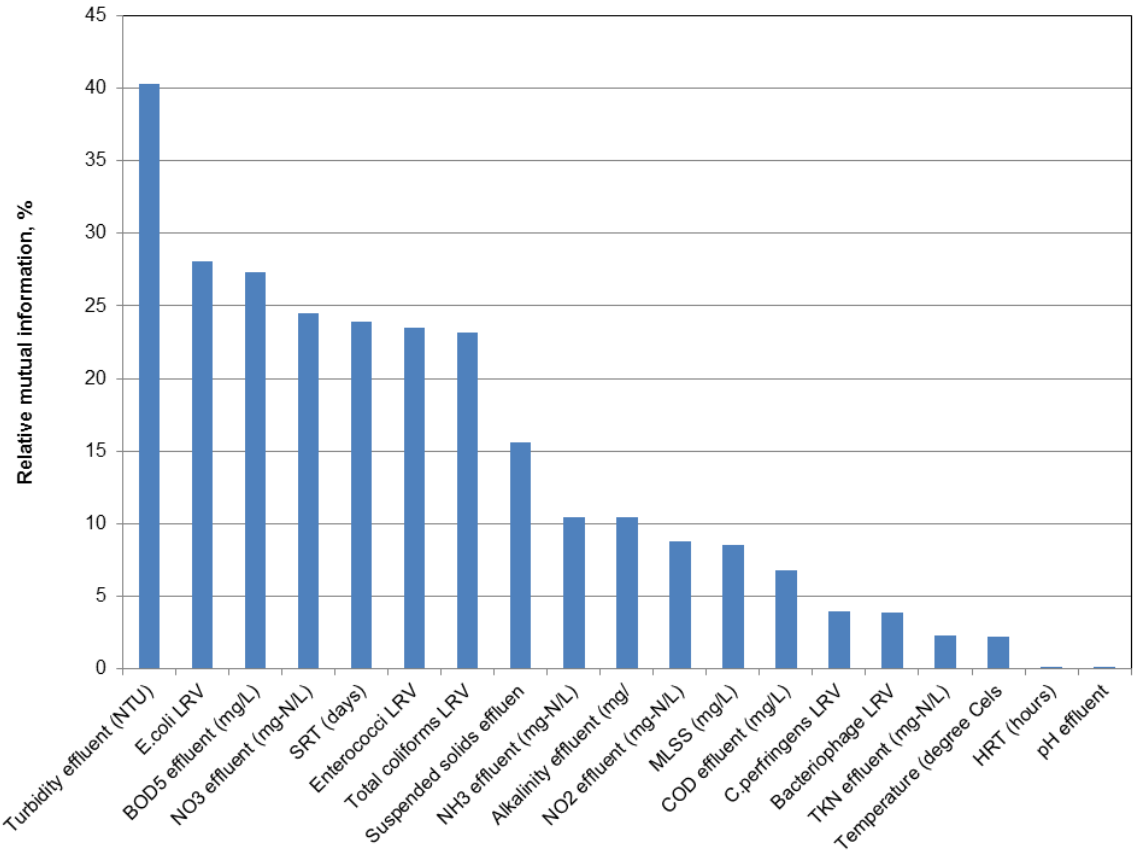

Figure 4-8: Sensitivity to findings for removal efficiencies according to operating parameters.

4.4.4 Semi-naïve v. causal Bayesian modelling of water treatment processes

The models presented in this study were not primarily designed to reflect cause-effect relationships. That is, they are not "causal" BN models. However, they were shown to provide stable credible prediction of *C. parvum* LRVs. Among the reasons the semi-naïve Bayes approach was trialled, rather than a causal network approach, was the limited size of the available dataset compared to the number of possible variables, ease of model construction, avoidance of human bias and limited prior knowledge of the likely associations among the variables.

Causal models have advantages and different potential uses. For example, they would allow the whole system to be represented by a single network and conceptually provide insights to the system behaviour. In causal models, derived numerical probabilities can be considered as representations of the probabilities of occurrence of a particular event. However, a disadvantage of causal models is that where a system is not well understood mechanistically and there are many dependent and independent nodes variables, the number of plausible models multiplies rapidly making parsimony a concern.

Semi-naïve Bayes models, on the other hand, allow the strong assumption of node independency given the target variable to be relaxed. These models are an intermediate step between the naïve Bayes model and a causal model and empirical experience in other fields has shown they can be very reliable (Korb & Nicholson, 2011). This approach also allowed dispassionate model construction using various performance metrics in a stepwise fashion based on rules developed for BNs generally. WEKA allowed assessing whether there were sufficient data records to generate stable model structures and what were credible discretisation thresholds. The performance metrics provided information about which nodes were most likely to influence LRVs. The metrics also allowed comparison of i) different model options, ii) their respective predictive power, and iii) assessment of whether the best models were credible or provided no improvement over the ZeroR model.

Real world activated sludge plants will differ in many respects for the pilot AS including SRT ranges, HRT ranges, temperature ranges, and MLSS concentration ranges and different data sets. So although the LRVs estimated here are valuable *per se* it may be preferable to repeat the model development process in such systems rather than use the models themselves uncritically. Nevertheless, the method used here can be adapted to this task of constructing new candidate models and determining which if any validly describe the system being characterised.

4.5 Conclusions

A conceptual alternative to directly measuring pathogen removal efficiency is to predict LRVs using cost effective microbial and physicochemical monitoring, control parameters system operating conditions. However, conventional parametric statistical analyses have not yielded sufficiently convenient tools which describe AS processes and relate variables. In this investigation, naïve and semi-naive Bayes models to predict and manage pathogen reductions were developed and assessed. A real-world data set was used to evaluate and quantify significant relationships between operating and monitoring parameters and estimate removal of two pathogens. The developed methodology is objective, systematic and applicable to analysing water treatment processes more generally. This study also identified operational parameters potentially useful for the prediction of *C. parvum* removal efficiency. Conversely the lack of success in modelling *G. lamblia* suggested that its removal by AS is not sufficiently understood and cannot yet be quantified based on removal of microbial indicators, even though assignment of average reduction credits of ≥ 1 $\log_{10}$ is still reasonable judging by the raw LRV probability density function.

Key outcomes from this study included:

- Useful predictors for *C. parvum* reduction included turbidity, SS, total coliform bacteria LRV and enterococci LRV.

- SRT, COD and nitrite were potential predictors of *G. lamblia* LRV. However, their AUC score was less than or equal than 0.81 indicating more work is needed before they can be reliably applied to this task.

- No microorganisms alone were reliably correlated with, or good predictors, of *G. lamblia*. This result highlighted the need to better understand the relationship between the removal of *G. lamblia* and other AS microbes.

- Naïve and semi-naive Bayes modelling of a real AS plant could reduce the costs of direct pathogen monitoring and encourage the gathering of informative process data which would permit LRV credits to be linked to the system's operating conditions.

- NB and SNB models can be used to understand whether optimal LRVs can be achieved concurrently with satisfactory BOD_5 and nitrogen removal.

Although causal BNs were not constructed, the non-causal models provide a reference and starting point for such modelling by identifying those variables most likely to be useful when constructing causal models with the minimum of nodes. The SNB models provide an objective way of estimating the maximum accuracy that is possible with a causal Bayes model. The models are relatively easy to understand which should assist uptake by non-experts in Bayesian networks. Finally, the method here can reduce disagreements between model developers about what form BNs should take.

Chapter 5: Validation of ultrafiltration processes using Bayesian analysis

This chapter has been published in the following journal article:

Carvajal, G., Branch, A., Sisson, S.A., Roser, D.J., van den Akker, B., Monis, P., Reeve, P., Keegan, A., Regel, R. and Khan, S.J. (2017). Virus removal by ultrafiltration: Understanding long-term performance change by application of Bayesian analysis. Water research. 122, 269-279. DOI: 10.1016/j.watres.2017.05.057

5.1 Introduction

The work presented in this chapter was undertaken to develop a new application of BNs to assess the reliability and effectiveness of and ultrafiltration (UF) process for advanced water treatment. This assessment was undertaken by examining multiple UF skids operating in parallel and their change over time. Data obtained were used to calculate and compare individual and aggregated virus log reduction values for the multiple skids. The results obtained demonstrate how treatment can be effectively quantified using Bayesian hierarchical modelling analyses.

Through size exclusion, UF membranes provide an effective barrier against waterborne pathogens such as *Cryptosporidium* and viruses (Asano *et al.*, 2007). Although there is no established online monitoring system to directly test integrity issues of relevance to particles as small as viruses (Antony *et al.*, 2012), challenge testing with representative virus surrogates, such as MS2 bacteriophage (MS2), can be conducted to attribute an LRV to a UF process during treatment performance validation (USEPA, 2005). During challenge testing, feed and filtrate samples are taken over the course of a filtration cycle and analysed to determine the virus LRV. When MS2 challenge testing is performed at full scale, common practise is to select UF skids for testing individually to minimise the effects of uncontrolled variables and assess both single skid and between-skid LRV variation. Challenge testing provides an instantaneous measure of LRV, which may confirm that a UF skid can meet required performance targets. In the past, LRVs have been typically derived by water quality analysis of time-matched samples preceding and following a treatment process of interest (Smeets *et al.*, 2008); as in this study, analysis of UF feedwater and filtrate, taken at the same time. When considering inherent LRV variability of the process, parametric distributions may be fitted to these data or bootstrapping may be used to generate descriptive statistics without assuming an underlying distribution (Smeets *et al.*, 2008). However, uncertainties in the parameters of these distributions are often not considered and historical challenge testing information is usually not incorporated. Incorporation of historical test data is particularly relevant when re-validation (i.e. a repeat of challenge testing on the same skid) is mandated as part of a water safety plan.

When validation of membrane systems are conducted, individual and overall skid performances need to be estimated from skid-specific data and combined into an overall LRV by pooling challenge test data as described by USEPA (2005) methods. Determination of LRV for the whole UF system in this manner assumes two somewhat contradictory approaches: 1) considering total independence between skids for the individual performance and 2) completely equal performance for the overall LRV. The first assumption disregards any similarities between the skids which could be employed during the performance estimation, while the latter assumption discounts the

impact of historical events that are confined to sub-units, or in this case, to individual skids. Although the methodology is commonly used and can provide reasonable outcomes, it neglects the incorporation of parameter uncertainty and the information that other skids in the system could provide to improve individual skid LRV estimates. As a consequence, less accurate results may be generated if predicting new outcomes.

A more complete approach would be to consider all skids performing in a similar, but not necessarily in the same fashion, and functioning as inputs to an overall performance model. This desired model characteristic may be achieved when a Hierarchical Bayesian Model (HBM) is employed. A HBM consists of a Bayesian model on which prior distributions have been also defined on the prior parameters associated with the likelihood parameters (Ntzoufras, 2011). The work presented in this chapter applied a HBM to address the above needs. A number of model assumptions were tested using real UF challenge test data, from a full-scale plant, taken across multiple skids and also at five years apart. The optimised model was applied to compare LRVs of individual skids for the same and different years, quantify uncertainty in the distribution parameters and generate skid-specific and skid-independent predictive distributions.

5.2 Materials and methods

5.2.1 Ultrafiltration system

The UF system of a full-scale water reuse treatment plant in South Australia (36 ML/d maximum capacity) was used to collect data for this study. This system was initially validated in 2010 and then re-validated after five years for virus removal, using MS2 as a microbial surrogate (Reeve *et al.*, 2016). To minimise the impact of feedwater variability, and due to the typically low densities of autochthonous F-RNA bacteriophage, challenge testing was conducted using a laboratory-grown stock of MS2 F-RNA bacteriophage. Preparation of MS2 and dosing into the UF system has previously been reported by Reeve *et al.* (2016). Briefly, MS2 stock was dosed from a premixed batching tank upstream (suction side) of the UF skid feedwater pump, at a flowrate sufficient to achieve a challenge concentration of 10^5 plaque forming units per mL of feedwater (PFU mL^{-1}). Dosing commenced 10 – 30 min prior to challenge testing to allow mixing and equilibration throughout the UF skid, and continued for the entire test duration (4 hours).

The current UF system comprised 8 parallel skids, each containing 120 membrane modules (Figure 5-1). These were hollow fibre polyvinylidene difluoride (PVDF) membranes with a nominal pore size of 0.04 μm, an asymmetric structure, and filter flow from outside to inside. The plant has a design flux of 52.1 L m^{-2} h^{-1} and a maximum design flow of about 1500 m^3 h^{-1}. The

UF system processes chlorinated secondary effluent which is further pre-chlorinated to minimise biological fouling. Following UF, pathogens are further reduced by UV and chlorine disinfection. End uses for the product water include unrestricted municipal irrigation, industrial applications, and dual reticulation. The LRV accredited for the UF system when first validated upon installation was 2.5 $\log_{10}$ for viruses (Reeve *et al.*, 2016). Accreditation of this performance was subject to compliance with process control parameters including pressure decay rate (PDR) (<4.8 kPa min^{-1}), resulting from daily pressure decay testing (PDT) and effluent turbidity (daily average <0.15 NTU and not >0.3 NTU for 30 continuous minutes). Compliance for these CCPs was derived in consultation with the regulatory agency, which oversees the operation of this plant (SA Health). These turbidity requirements were based on performance of the membranes during the initial validation exercise when new.

5.2.2 Data collection

Challenge testing was performed in three non-consecutive days (17 June, 8 July and 28 July) for 2010, during commissioning with newly installed membranes (9-months since installation), and in one day (17 June) for 2015. For the purposes of challenge testing the secondary effluent was not pre-chlorinated. The intention of the initial validation study, was to obtain a conservative set of virus LRV, that the UF unit could continuously be expected to achieve, to assure safe water supply for end users (Reeve *et al.*, 2016). Indeed, membranes will foul over time and this accumulated fouling layer will likely enhance observed virus removal (Antony *et al.*, 2012). Therefore, it was expected that challenge testing with a membrane with fouling removed should yield results that represent conservative, worst case removal performance. Challenge testing immediately following chemical cleaning is also recommended in the USEPA membrane filtration guidance manual, for the same reasons (USEPA, 2005). Accordingly, prior to challenge testing, the membranes were backwashed and chemically cleaned to minimise the potential of accumulated fouling leading to increased LRV. As such, the MS2 LRVs presented are considered to be conservative. Onsite chemical cleaning involved a clean in place (CIP) regime that included a five-hour sequential soak, circulate and rinse stages, with sodium hypochlorite (200 mg/L w/v) or sulphuric acid (0.1% w/v) and citric acid (0.5% w/v). A membrane rinse followed using potable water. The chemical cleaning success was corroborated by the increase in skid flow, and decreases in transmembrane pressure and resistance, post clean. Prior to sampling, PDT was conducted to confirm that the membrane PDR was below the limit value of 3.5 kPa min^{-1}, indicating nominal membrane integrity. Grab samples were collected at UF skid feed and filtrate sampling taps. Sampling was performed at four equally time-spaced points in a single filtration cycle. As presented in Figure 5-1, a total of 43 paired samples (i.e. membrane feed and filtrate) were collected from 4 skids (skids 2,4,5 and 6) during the commissioning testing in 2010, while a total

of 19 paired samples were collected for 5 skids (skids 2,4,5,6 and 8) during re-validation in 2015. All data observations were above the limit of detection (LOD) of 1 PFU/100 µL. However, some of the observations for 2010 were close to the limit of detection with few counts per plate (≤10 PFU) and were considered unreliable. Therefore, the limit of quantification of the method was used, which was defined as 10 PFU/100 µL (ISO, 2013b). Counts obtained below the limit of quantification were considered censored in this study. Because LRVs were calculated as the $\log_{10}$(influent concentration)-$\log_{10}$(effluent concentration), some right censored values with multiple censored limits were obtained. During initial challenge testing in 2010, all skids were operated at a flux of 52 L m^{-2} h^{-1} with a filtration run time of 30 minutes. After 5 years of operation, some irreversible fouling had accumulated on the membranes. As a consequence, the membranes would require more frequent backwashing and CIP. For the re-validation testing in 2015 it was necessary to decrease flux to 41-47 L m^{-2} h^{-1} to provide a filtration cycle of 20 minutes. Samples were collected in triplicate and analysed on the same day of collection. MS2 were quantified using the plaque assay double-agar layer method with *Escherichia coli* Famp (ATCC700891) as host (USEPA, 2001).

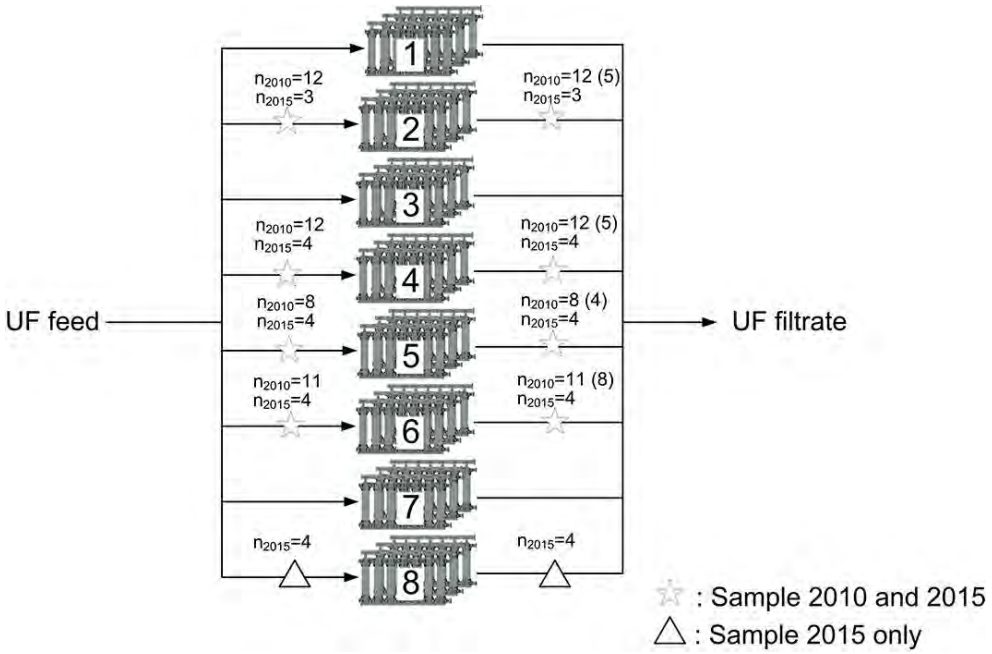

Figure 5-1: Sample points for the validation and re-validation testing. n_y is the number of samples taken at a particular sample point during sample period in year y. Numbers in parentheses indicate the number of uncensored observations.

5.2.3 Software for Bayesian analysis

Parametric Bayesian analysis was used in this study because of its suitability to facilitate model construction and interpretation of hierarchical models. Bayesian analysis also permits to explore uncertainty and variability explicitly. Model construction and evaluation for Bayesian models were explained in Section 3.2.

The Bayesian models for validation were implemented using the *JAGS* software package (Plummer, 2003) using Markov Chain Monte Carlo (MCMC) methods. The process of model construction and inference was conducted in R through the open source *r2jags* package (Su & Yajima, 2015). More information can be found in Chapter 3, Materials and Methods.

5.2.4 Model design for validation for processes in parallel

The validation model for parallel UF processes was constructed using a Bayesian approach to compare process performance between skids and years, and to predict the performance of new skids. For the evaluation of performance of a UF process, which consists of a set of parallel membrane skids, a hierarchical Bayesian model (HBM) was employed, which was named $HBM_{\mu\sigma}$ (Table 5-1) as shown in the directed acyclic graph (DAG) in Figure 5-2. The graphical representation of the model as a DAG provides the basis for calculation by indicating the joint relationship between the variables in the model through local relationships. In a DAG, nodes (ellipses) represent variables and arrows represent direct dependence between variables. Two variables are identified as "parent" and "child" nodes if there is an arc from the former to the latter (Korb & Nicholson, 2011). In this diagram, data ($y_{i,j,k}$) are represented as a small rectangle. The rectangular "plates" indicate repetition. Application of a HBM was considered to be appropriate because all skids are subject to the same process conditions incorporating the same membrane characteristics, and so they can be expected to perform similarly, but there will also be skid-specific variations in performance. In the $HBM_{\mu\sigma}$, global and individual skid performance parameters were used to characterise the system. This $HBM_{\mu\sigma}$ configuration allowed learning about the performance characteristics of the skids, not only from the data originating from a particular skid, but also through incorporation of information observed from the remaining parallel skids. The $HBM_{\mu\sigma}$ assumed that parameters from the different skids are neither identical, nor completely independent. The $HBM_{\mu\sigma}$ provided information about the degree of similarity between skids, which was then used to predict the performance for non-observed skids. The $HBM_{\mu\sigma}$ was appropriate for these data, because it estimated the underlying performance of each membrane skid while simultaneously estimating the overall performance and consistency of the group. The $HBM_{\mu\sigma}$ had three distinct levels (i,j,k), which corresponded to sample number (i), skid

number (j) and year of sampling campaign (k) out of a total number of samples (m), skids (n) and years (o), respectively. Parameters for the distributions were positioned on different levels with global parameters for the whole system located in the uppermost level, indicating that they are unique. Comparison of data for individual units and years against normal probability plots returned correlations higher than 0.92 (Appendix 2) indicating that normality assumption was adequate for these data. Therefore, the data (y_{ijk}) were assumed to follow a normal distribution with parameters mean ($\dot{\mu}_{jk}$) and standard deviation ($\dot{\sigma}_{jk}$) (Equation 5-1).

$$y_{ijk} \sim Normal(\dot{\mu}_{jk}, \dot{\sigma}_{jk}) \quad \text{Equation 5-1}$$

Each skid was assumed to have individual (but not independent) mean and standard deviation parameters. A potential location shift of the parameters between the two campaigns of analysis for each skid *j* was captured through δ_j for the mean (Equation 5-2), and ε_j for the log of the standard deviation (Equation 5-3). Values of the shifts (δ_j and ε_j) would reveal any change in skid performance captured by a shift in the distribution or change in the variability. In this manner, it can be written:

$$\dot{\mu}_{jk} = \mu_j + I(k) \cdot \delta_j \quad \text{Equation 5-2}$$

$$\log(\dot{\sigma}_{jk}) = \log(\sigma_j) + I(k) \cdot \varepsilon_j \quad \text{Equation 5-3}$$

where

$$I(k) = \begin{cases} 0, & k = 1 \\ 1, & k = 2 \end{cases} \quad \text{Equation 5-4}$$

Equation 5-4 is an indicator function, taking the value 0 in the first year of analysis (*k=1*) and 1 in the second year (*k=2*). Normal distributions were also assumed for the mean (μ_j) (Equation 5-5), its shift (δ_j) (Equation 5-6), logarithm of standard deviation ($\log(\sigma_j)$) (Equation 5-7) and its shift (ε_j) (Equation 5-8). The parameters for these distributions are located on the top level providing the connection between the skids' performances. The logarithm of the standard deviation was used as suggested by Lunn *et al.* (2012) because it can take any real value instead of only positive values as in the case of the standard deviation. This consideration is important as the logarithm of the standard deviation can be parameterised using a normal distribution, which

is convenient in this case. A second option is parameterising the standard deviation directly as an inverse Gamma distribution with parameters alfa and beta. Both approaches are generally considered valid and accepted (Lunn *et al.*, 2012).

$$\mu_j \sim Normal\left(\mu_\mu, \sigma_\mu\right) \quad \text{Equation 5-5}$$

$$\delta_j \sim Normal\left(\mu_\delta, \sigma_\delta\right) \quad \text{Equation 5-6}$$

$$\log(\sigma_j) \sim Normal\left(\mu_\sigma, \sigma_\sigma\right) \quad \text{Equation 5-7}$$

$$\varepsilon_j \sim Normal\left(\mu_\varepsilon, \sigma_\varepsilon\right) \quad \text{Equation 5-8}$$

When interpreting results, μ_j and $\mu_j + \delta_j$ represent the distribution mean of the *j*-th skid in the first and second campaigns respectively, and μ_μ and $\mu_\mu + \mu_\delta$ represent the distribution means of all skids in the first and second campaigns respectively. Similar interpretations are available for the other parameters. Diffuse prior distributions for the parameters presented in Equation 5-5 to Equation 5-8 were defined as Normal(0,100) for μ_μ, μ_δ, μ_σ and μ_ε, and Gamma(0.001,0.001) for σ_μ^{-2}, σ_δ^{-2}, σ_σ^{-2} and σ_ε^{-2} (Lunn *et al.*, 2012). Such a normal prior distribution is practically flat within the values of interest for this problem (i.e. 0-5). The Gamma(0.001, 0.001) distribution is commonly employed as a prior for the precision ($\tau = \sigma^{-2}$) of the normal distribution because it functions as an approximation to the Jeffrey's prior $P_J(\tau) \propto \tau^{-1}$, thereby favouring large values of standard deviation in an adequate fashion (Lunn *et al.*, 2012).

The $HBM_{\mu\sigma}$ model represents the most flexible model in which each skid was permitted to have individual (but not unrelated) parameters for the level of performance (mean) and variability (standard deviation). Four nested variations of the $HBM_{\mu\sigma}$ (Appendix 2) were also considered as shown in Table 5-1, to identify the most appropriate model for the observed data. These models corresponded to variations of the fully hierarchical model resulting from different assumptions on the standard deviation of the hyperparameters (σ_μ, σ_δ, σ_σ, σ_ε). Full details are presented in Appendix 2. Summary characteristics of each model are also included in Table 5-1. A comparison between these models is useful to demonstrate the ability of HBMs to adequately analyse LRV data.

While the normal distribution is frequently used for describing LRVs (Haas & Trussell, 1998), other parametrisations are also possible including the use of gamma distributions and Pareto (Haas & Trussell, 1998; Teunis *et al.*, 1999; Teunis *et al.*, 2009).

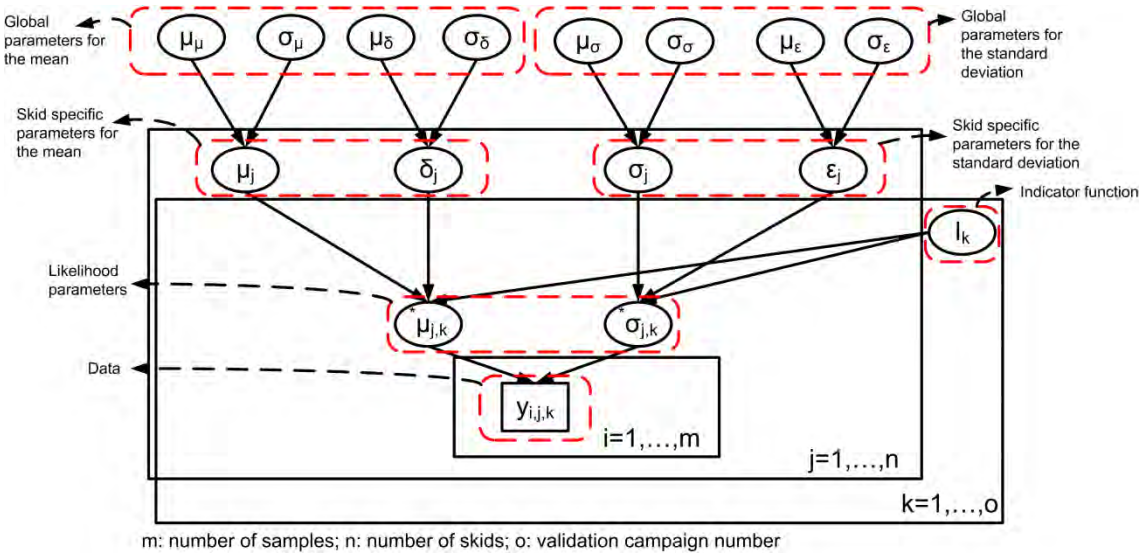

Figure 5-2: Full hierarchical validation model with exchangeable parameters for μ and σ.

Table 5-1: Models evaluated and main assumption

Figure	Model type	Assumptions	Abbreviation
Figure 5-2	Hierarchical model (μ and σ exchangeable)	Population distribution for parameters μ and σ	$HBM_{\mu\sigma}$
Figure A4 [a]	Hierarchical model (only μ exchangeable)	Population distribution for parameter μ and identical σ	HBM_{μ}
Figure A5 [a]	Hierarchical model (only σ exchangeable)	Population distribution for parameter σ and identical μ	HBM_{σ}
Figure A6 [a]	Non-hierarchical (identical μ and σ)	Same μ and σ across skids	$NHBM_{same}$
Figure A7 [a]	Non-hierarchical (independent μ and σ)	Unique μ and σ for each skid	$NHBM_{indep}$

[a] Appendix 2

The first model variation (HBM_{μ}) assumes that the variability in the skids' performance is the same in each year, but the level of performance was specific to each skid. That is, the HBM_{μ} has exchangeable parameters only for the means, and identical standard deviations for the skids sampled during the same year (Appendix 2, Figure 12-4). This configuration is equivalent to letting the standard deviations, σ_{σ} and σ_{ε} approach zero in the fully hierarchical model. The second model variation (HBM_{σ}) assumes skid-specific variability but the same level of performance across all skids for each year. This configuration means that the HBM_{σ} assumes exchangeable parameters only for the standard deviation, and identical means for the skids sampled during the same year (Appendix 2, Figure 12-5). As before, the HBM_{σ} represents a variation of the $HBM_{\mu\sigma}$ in which the standard deviations related to the mean, σ_{μ} and σ_{δ}, tend to zero. The third model option was a non-hierarchical Bayesian model (NHBM). $NHBM_{same}$ assumed that all the skids had the same level of performance and variability, which is equivalent to pooling all skid-specific data within a year and fitting a normal distribution. This model results from letting all hyperparameter standard deviations in $HBM_{\mu\sigma}$ approach zero (Appendix 2, Figure 12-6). The final model $NHBM_{indep}$ assumes individual (and independent) parameters for the level of performance and variability for each skid. This model arises when all hyperparameter standard deviations for the hyperparameter become large and approach infinity (Appendix 2, Figure 12-7).Deviance Information Criterion (DIC) (Spiegelhalter *et al.*, 2002) was used to compare the goodness of fit between the proposed HBM and NHBM models.

Pairwise comparisons were used to assess the differences in performance between skids for the same year and the decrease in performance after five years of operation. The comparisons were performed by using the parameter estimation approach (Kruschke, 2013; Kruschke, 2014). This method uses a concept known as the Region Of Practical Equivalence (ROPE), which indicates a small range of parameter values that are deemed to be equivalent to the null value when used in a particular application. When the 95% credible interval (CI) is completely contained within the ROPE it may be interpreted that the null value cannot be rejected, whereas when the 95% CI interval excludes the ROPE it can be interpreted that there is strong evidence that the null value can be rejected. In this study the ROPE for the null hypothesis was defined as -0.1 to 0.1. Therefore, an increase or decrease in 0.1 LRV was considered to be enough to result in an improvement or deterioration of the performance. The range was considered to be sufficiently large to capture the case when there are no differences and sufficiently small to capture real differences. A sensitivity analysis to the selected ROPE was also conducted to test the importance of this assumption.

5.3 Results and discussion

5.3.1 Model comparison

The five candidate models were compared by their DIC score, with additional consideration of the model goodness of fit and the effective number of parameters for each case (Table 5-2). In all, these results suggest that the model with exchangeable parameters for the mean (HBM_{μ}) is the most appropriate model; this model was used to analyse the system. It has been suggested that a magnitude of more than five DIC points indicates a substantial difference between models (Lunn *et al.*, 2012). The results presented in Table 5-2 indicated that the NHBM with independent parameters ($NHBM_{indep}$) and the HBM with hierarchical parameters for the mean and standard deviation ($HBM_{\mu\sigma}$) had the best fit according to the mean deviance. This outcome is not unexpected, as using more parameters to describe a dataset will usually result in a better fit to the data. However, the large number of effective parameters used in the $NHBM_{indep}$ resulted in a comparatively large DIC score compared to the other candidate models. The hierarchical model with exchangeable parameters for the mean alone (HBM_{μ}) had similar fit to the fully independent NHBM and fully hierarchical HBM according to the mean deviation. However, the larger effective number of parameters for the $HBM_{\mu\sigma}$ produced an increase in the DIC of about one unit compared to the HBM_{μ}. The hierarchical model with exchangeable parameters for the standard deviation (HBM_{σ}) had the second worst fit according to both the mean deviance and DIC score, indicating substantial evidence for mean performance differences between skids. The identical parameters model ($NHBM_{same}$) showed the worst fit according to the mean deviance. At the same time, this model had the lowest number of effective parameters.

Table 5-2: Analysed model variations, their abbreviations and results for mean deviance ($\overline{D}$), effective number of parameters (pD) and deviance information criterion (DIC).

Model	Abbreviation	$\overline{D}$	pD	DIC
Exchangeable parameters for μ and σ	$HBM_{\mu\sigma}$	33	11	44
Exchangeable parameters for μ	HBM_{μ}	35	8	43
Exchangeable parameters for σ	HBM_{σ}	43	7	50
Identical parameters	$NHBM_{same}$	46	4	50
Independent parameters	$NHBM_{indep}$	33	19	52

$\overline{D}$: mean deviance, pD: effective number of parameters, DIC: deviance information criterion.

5.3.2 Ultrafiltration validation results

Data were fitted to normal distributions and summary statistics were computed as shown in Table 5-3. The results indicated that for 2010 the sample mean and standard deviation of skids 5 and 6 were lower than the sample mean and standard deviation of skids 2 and 4. For 2015, the sample mean performance was similar across skids. However, sample standard deviations obtained for skids 6 and 8 were lower than the standard deviations for skids 2, 4 and 5. The results of fitting HBM_{μ} (Appendix 2, Figure 12-4) are presented as box plots and probability density functions. Box plots were modified to depict the 2.5^{th}, 25^{th}, 50^{th}, 75^{th}, and 97.5^{th} percentiles (with the box representing the usual quartiles, and the whisker the extent of the central 95% of the probability distribution). In this study, 2.5^{th} and 97.5^{th} percentiles were used in place of more commonly reported 5^{th} and 95^{th} percentiles, to maintain consistency and avoid confusion with material presented in Section 5.3.3 (pairwise comparisons part) that required the use of the 95% credible interval. The results obtained from the analysis include skid-specific posterior predictive distributions (Figure 5-3), and skid-specific posterior distributions for the parameter means (Figure 5-4) and standard deviations (Figure 5-5). Additional results include distributions for the differences in performance between years for the means and for the logarithm of the standard deviations (see Appendix 2). Skid-independent posterior predictive distributions were also generated for both analysed years (Figure 5-8).

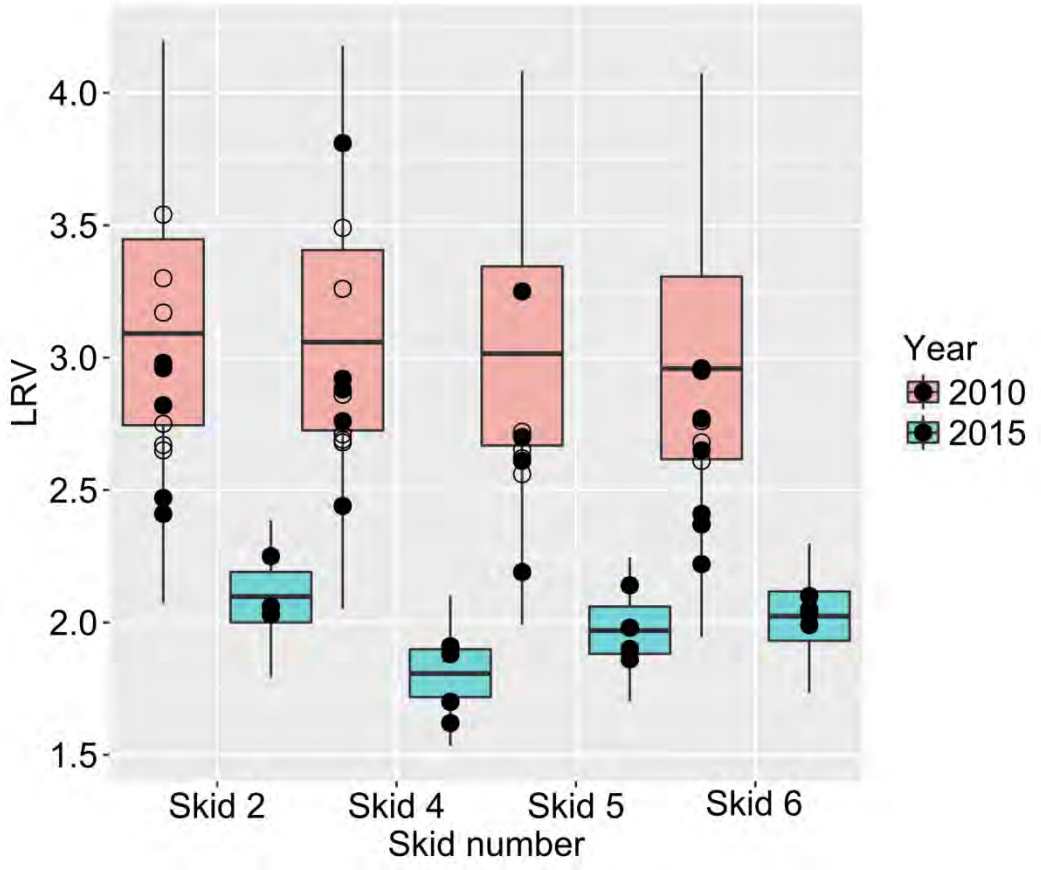

Figure 5-3: Skid-specific posterior predictive distribution with observed data points superimposed. Void circles indicate right censored values.

Table 5-3: Summary statistics for the data assuming normal distribution

Unit	2010		2015	
	Mean(95% CI)	SD [b]	Mean(95% CI)	SD
2	3.09(2.35, 3.83)	0.38	2.11(1.85, 2.37)	0.13
4	3.16(2.34, 3.96)	0.41	1.78(1.46, 2.08)	0.16
5	2.68(2.49, 2.87)	0.10	1.97(1.69, 2.24)	0.14
6	2.78(2.50, 3.06)	0.14	2.04(1.93, 2.14)	0.05
8	NA[a]	NA	1.90(1.74, 2.06)	0.08

[a] Not available

[b] Standard deviation

It is observed from Figure 5-3 that there was a marked decrease in the performance between years in all skids. The superimposed data reveals a credible fit to the posterior distributions for each skid. The difference in performance means for each skid between years (δ_j) was above zero for less than 0.1% of observations (Appendix 2, Figure 12-1), which suggests a decrease in the mean performance. Similarly, the logarithm of the standard deviations between years (ε) was less than zero for less than 0.1% of observations (Appendix 2, Figure 12-2) suggesting a decrease in variability between years. It is possible that the difference in variability between samples for the two years may have been affected by the size of the data. However, the uncertainty in the posterior parameters is usually increased when data are limited. The decrease in uncertainty in this case is more likely due to consistent observed performance. The removal performance differences between years for each skid can be further observed by the large differences in skid specific distribution means (Figure 5-4) and standard deviations (Figure 5-5). The posterior mean of the overall performance for all skids obtained from the hyperparameters (μ_μ and μ_δ) was 3.1 and 2.0 $\log_{10}$ removal units for the years 2010 and 2015, respectively. The posterior mean of the overall difference for all skids (μ_δ) was -1.1 (95% credible interval (CI): (-1.4, -0.7)), while the difference in the logarithm of the standard deviation (ε) had a mean of -1.4 (95% CI: (-1.9, -0.9)). The mean difference between standard deviations (Figure 5-5) was -0.37 (95% CI:(-0.60, -0.23)).

UF membranes are vulnerable to physical damage (i.e. cuts and abrasion of fibres) from foreign bodies in feed water, as well as chemical degradation from regular exposure to cleaning solutions used to reduce fouling or prolonged storage. Depending on the membrane materials, chemical exposure can result in changes in membrane surface chemistry (i.e. surface charge and hydrophobicity), mechanical strength and also pore size (Childress *et al.*, 2005; Arkhangelsky *et al.*, 2007; Puspitasari *et al.*, 2010). Regardless of the failure mode, cumulative integrity failure is expected in UF systems after a prolonged operational period, with one estimate suggesting 1 broken fibre per 100,000 fibres per year to be typical (Gijsbertsen-Abrahamse *et al.*, 2006). Increased levels of integrity failure would be expected to reduce the effectiveness of size exclusion and result in decreased removal capacity of the UF system, which corresponds with the decrease in mean performance observed within HBM_μ.

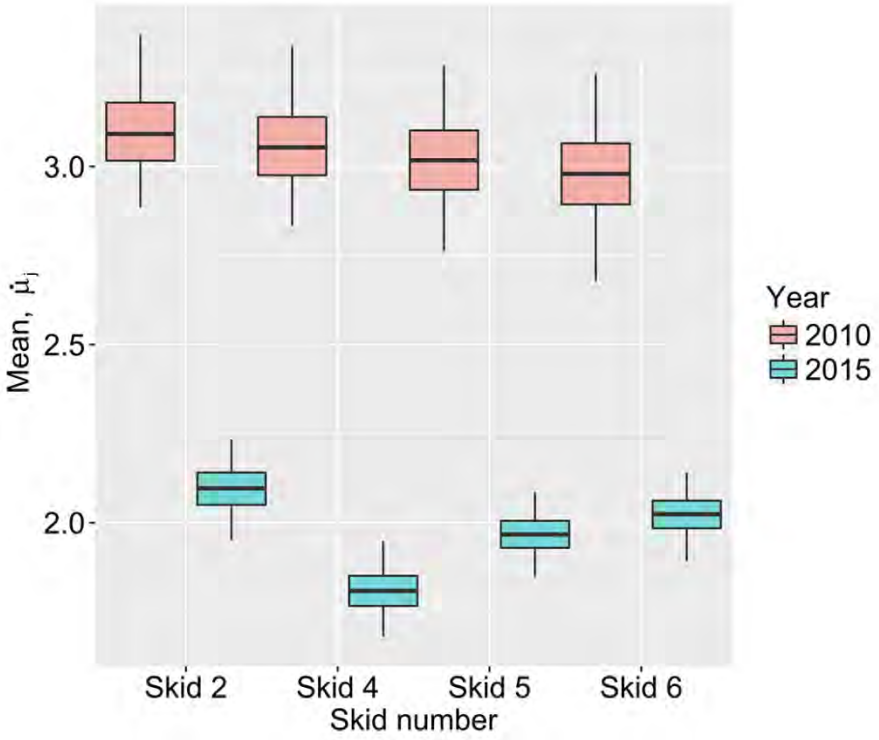

Figure 5-4: Skid specific mean ($\dot{\mu}_j$) distributions.

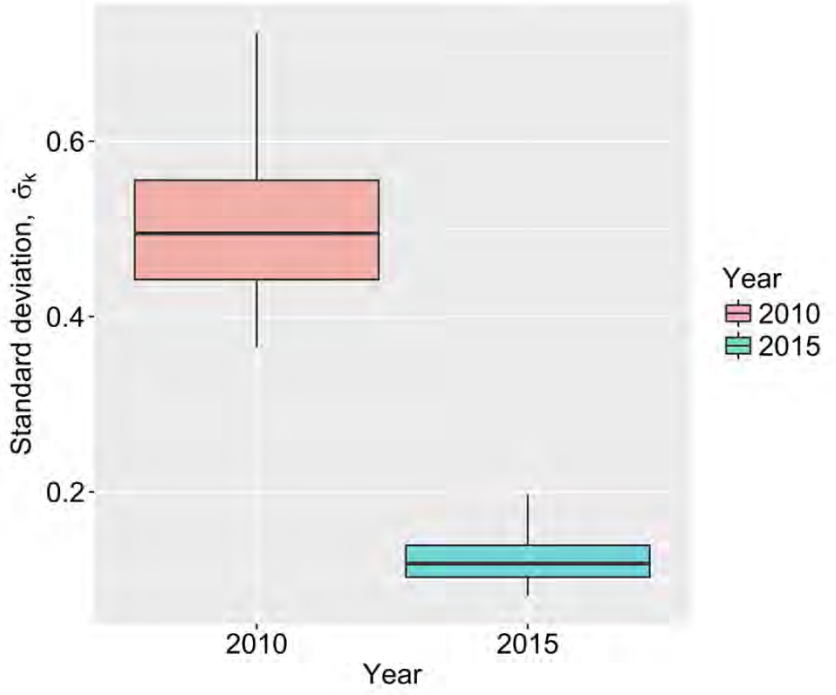

Figure 5-5: Year specific standard deviation ($\dot{\sigma}_k$) posterior distributions.

5.3.3 Pairwise comparisons

Pairwise comparisons to analyse differences in the mean between skid performances during the same year ($\dot{\mu}_{j=a,k} - \dot{\mu}_{j\neq a,k}$) (Table 5-4) indicated that for 2010 none of the skid comparison differences were below zero less than 5% of the time (Table 5-5, "Year 2010", "%<0 column") or above zero less than 5% of the time (Table 5-5, "Year 2010", "%>0 column"). However, the ROPE analysis showed that the credible differences were approximately zero (Table 5-4, "Year 2010" column), but that there was uncertainty in the estimation leading to only between 50% and 65% of the posterior distribution being contained within the ROPE. Measuring what percentage of the distribution is below or above zero provides information about the level of difference in performance between pairs of skids. For 2015, the paired differences between skids 4 and 2 (4-2), 4 and 5 (4-5), and 4 and 6 (4-6) were less than 5% of the time above zero (Table 5-5, "Year 2015" column), providing some evidence that there was lower performance from skid 4. Considering the ROPE, in all cases the ROPE overlapped with the 95% CI (Table 5-4, "Year 2015" column). However, the ROPE of the difference between skids 4 and 2 barely overlapped with the 95% CI. For 2015, the ROPE contained between 3.8% and 68% of the posterior distribution indicating that the results were uncertain about the difference in performance. As described by Reeve *et al.* (2016), skid 4 was subjected to a turbidity spike which was defined as a hazardous event as it exceeded the maximum limit of the turbidity analyser. However, skid 5 was also exposed to this event, without showing any evident detrimental change on its performance, except for the difference in performance compared to skid 2 (2-5) (7.6% of the time below zero). The results from analysis with HBM_{μ} suggest that factors other than the confined turbidity spike may have affected the performance of skid 4. The performance decrease measured as the difference in the means between years of each skid (δ_j) was also used to compare skid performance ($\delta_{j=a} - \delta_{j\neq a}$). The differences between skid 4 and 2 (4-2), and 4 and 6 (4-6) were below zero less than 5% (but higher than 1%) of the time (Table 5-5, "Difference" column) suggesting that skid 4 had a larger decrease in performance than skids 2 and 6. The third largest difference was obtained from the comparison between skids 4 and 5 (4-5) which was almost 10% below zero. In all cases the ROPE partially or completely overlapped with the 95% CI (Table 5-4, "Difference" column) and contained between 14% and 56% of the posterior distribution. The evidence presented in this section suggested that there was a lower performance from skid 4 in 2015. However, in all cases for that year the ROPE only partially excluded the 95% CI, indicating that no strong conclusions regarding the difference in performance can be drawn from these results. Although with some uncertainty, the evidence suggested that all skids performed similarly and that no large differences were observed during the 2010 validation period. Widening the ROPE would have yielded the same conclusions for the year 2010 and produced an increased

uncertainty regarding the decrease in performance for skid 4 in 2015 by a larger overlap between the ROPE and the 95% CI.

Table 5-4: Median (95% credible interval) for the difference in mean performance between skids ("Year 2010" and "Year 2015") and difference between decrease in performance between years ("Difference").

Skids comparison	Year 2010	Year 2015	Difference
2-4	0.03(-0.24,0.31)	0.29(0.07,0.47)	-0.24(-0.60,0.03)
2-5	0.06(-0.14,0.39)	0.13(-0.05,0.31)	-0.05(-0.32,0.27)
2-6	0.09(-0.08,0.49)	0.07(-0.10,0.25)	0.03(-0.20,0.43)
4-5	0.02(-0.19,0.41)	-0.16(-0.33,0.02)	0.18(-0.06,0.60)
4-6	0.05(-0.18,0.52)	-0.22(-0.39,-0.02)	0.27(-0.01,0.78)
5-6	0.03(-0.18,0.34)	-0.06(-0.22,0.11)	0.08(-0.14,0.43)

Table 5-5: Percentage of the distribution lower than zero (%<0) and higher or equal to zero (%≥0) for the difference in mean performance between skids ("Year 2010" and "Year 2015") and difference between decrease in performance between years ("Difference").

Skids comparison	Year 2010		Year 2015		Difference	
	%<0	%≥0	%<0	%≥0	%<0	%≥0
2-4[a]	37.2	62.8	0.05*	99.5	96.9	3.1*
2-5	25.1	74.9	7.6	92.4	81.1	18.9
2-6	17.1	82.9	19.9	80.1	45.7	54.3
4-5	40.4	59.6	96.2	3.8*	10.3	89.7
4-6	32.4	67.6	98.1	1.9*	2.2*	97.8
5-6	38.3	61.7	76.1	23.9	15.6	84.4

[a]: if the order of the subtraction is reversed, the distribution below and above zero changes conversely

*Lower than 5 percent

5.3.4 Comparison between HBMμ and NHBM$_{indep}$

Box plots comparing the skid-specific posterior predictive distributions using the hierarchical model and the completely independent model (no common hyperparameters) are presented in Figure 5-6 for year 2010 and Figure 5-7 for year 2015. Compared to the predictive distributions of the independent model, those of the hierarchical model are pulled towards a common mean value (which is the primary assumption behind the hierarchical model, and the predictive uncertainty of some skids (notably skid 6) has increased as a result. With application of the HBMμ, skids with fewer available LRV data points were supplemented by the information in the hyperparameters, which were estimated using additional LRV data points from all parallel skids. This outcome means that the skids with fewer data points borrow information from the remaining skids, through the hierarchical structure of the model. In general, the skids contributed to the population parameters in proportion to the amount of data available from each. This behaviour is especially advantageous when each skid has few data points or when limited data are collected for some skids, whereas larger numbers of records are available for other skids. In general, for HBMs, group-specific parameters (i.e. skid specific parameters) are pulled together closer to the mode of the population-level parameters (i.e. μ_μ, μ_δ, σ_μ^{-2} and σ_δ^{-2}), which is known as "shrinkage". Shrinkage is a desirable property provided that the parameter estimates are potentially less affected by random sampling noise compared to a NHBM. The shrinkage pulls in the estimates of accidental outliers thereby reducing potential false alarms (Kruschke, 2014). In 2010 it was observed that skid 5 returned a broader 95% credible interval when assuming complete independence compared to the hierarchical model. The hierarchical effect on the model generated shrinkage in the variability of the estimates. The opposite effect was also observed in 2015 for skid 6. In this case, the variability was increased by the connection to the other measured skid performances. Skid 6 presented a very narrow distribution when the parameters were estimated independently from the other skids. The tight distribution indicated a very consistent performance. However, four data points to generate the posteriors might not be sufficient for a reliable prediction and the HBM would play a key role in improving the estimates by widening the 95% credible interval.

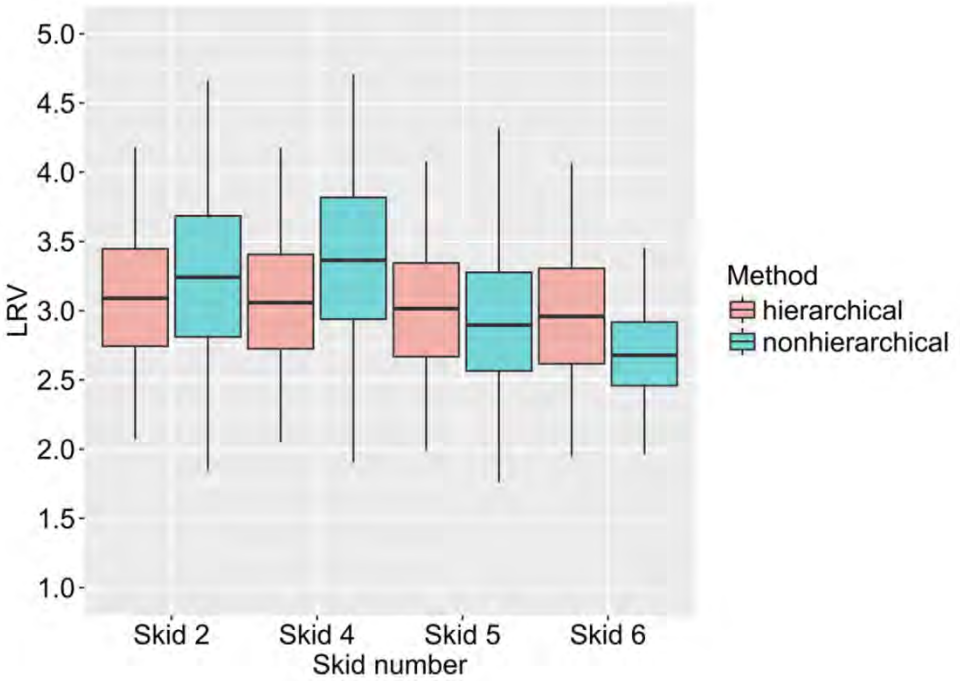

Figure 5-6: Skid-specific posterior predictive distribution comparisons between the hierarchical and completely independent models in 2010.

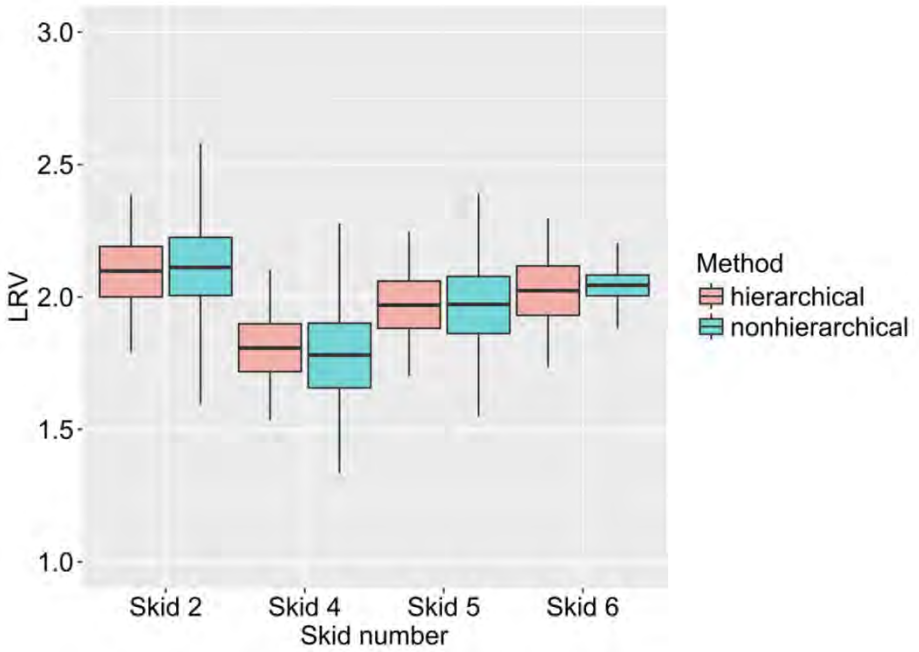

Figure 5-7: Skid-specific posterior predictive distribution comparisons between the hierarchical and completely independent models in 2015.

5.3.5 Skid-independent posterior predictive distributions

HBM_{μ} was used to produce skid-independent posterior predictive probability density functions (Figure 5-8 bottom row) – i.e., a predicted distribution of an as yet unobserved skid. Data from the observed skids have been superimposed (open circles for uncensored values and vertical lines for censored values) to visualise their location with respect to this distribution. A fifth skid, not previously challenge tested in 2010, was measured during the 2015 sampling campaign (Figure 5-1) and included on the comparison as black triangles. All measured data points were located inside the interval located between the 2.5th percentile and 97.5th percentile. It can be observed that the distribution is more concentrated near the median for 2015 with a narrower interquartile range (IQR: 1.8-2.2) compared to the year 2010 (IQR: 2.7-3.4). The 2.5th percentile for both distributions indicates that predicted performance in 2015 was inferior compared to 2010. Note that the observed data points appear to be more narrowly distributed in 2015 than in 2010, as the credible interval extends further than the data range. The underlying reason for this behaviour was linked to the variability in performance between skids for year 2015. In addition, a limited number of data points can lead to higher uncertainty in the estimated posterior parameters as in the case of 2015 data, which can lead to greater predictive uncertainty. The data for 2015 are narrowly distributed compared to the posterior predictive distribution obtained from the selected hierarchical model. In this case the variability in the distribution comes from both the between-skid and between-sample variance. Few observations for each skid can potentially increase the uncertainty in the estimations by affecting the variability at any of these two levels (between-skid and/or between-sample). The posterior predictive distributions for the model with identical parameters were also generated for comparison as shown in Figure 5-8 (top row). For the year 2010 the difference between the probability distributions was minimal, whereas for the year 2015 the probability distribution had a wider credible interval for the HBM, which provided more conservative estimates. It is likely that the full range of uncertainty of future skid performance is incorporated into this prediction. Incorporating additional uncertainty is desirable for QMRA because larger variability allows addressing conditions in which low or no microbial reduction is achieved (Ito *et al.*, 2016).

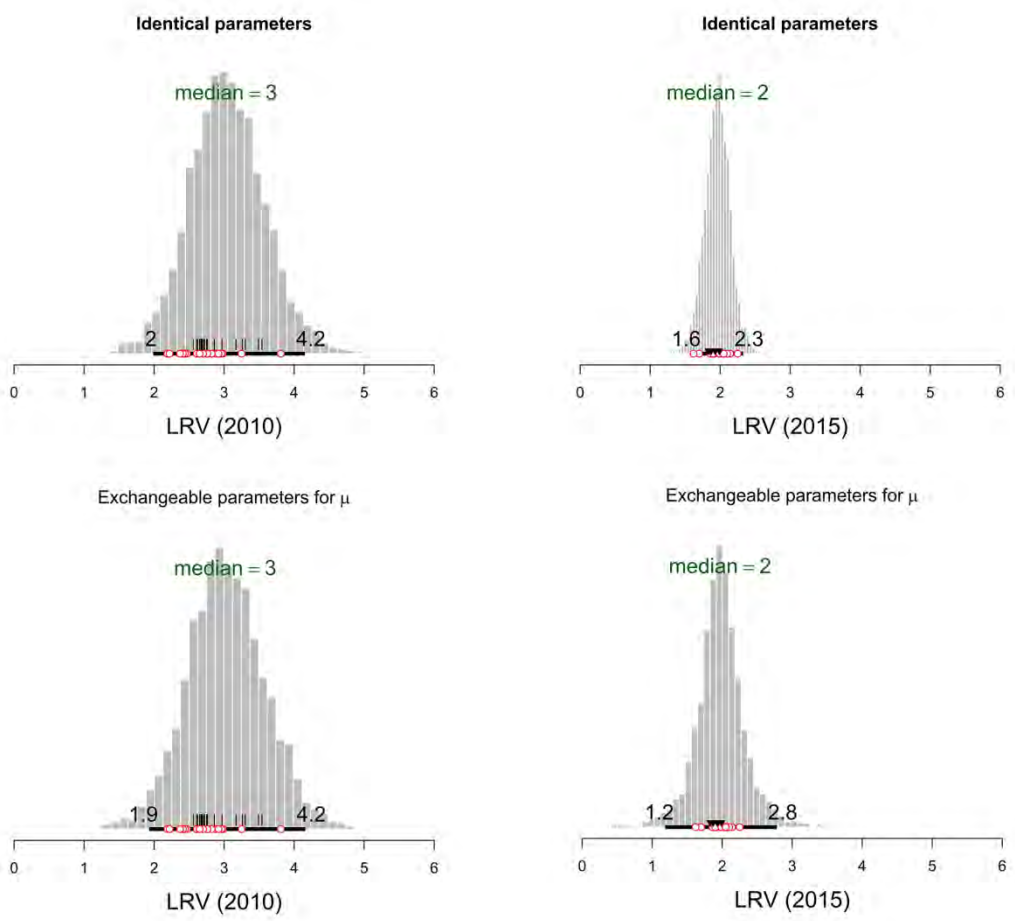

Figure 5-8: Posterior predictive distributions for 2010 and 2015 with superimposed data points using $NHBM_{same}$ (top row) and HBM_{μ} (bottom row). The horizontal axis shows the LRV. Open circles are the data points used in the model, black triangles are data points from skid 8 (only measured in 2015) and vertical black lines are right censored values. The 2.5th and 97.5th percentiles are also indicated on each tail of the probability distribution.

The simulated parameters for the skid-independent distributions were used to generate cumulative distribution functions with 95% credible intervals (Figure 5-9). The narrower credible interval in 2010 indicates less uncertain results compared to 2015. These results are useful to estimate uncertainty for the percentiles of the distribution. In this case, the 95% credible interval for the 2.5th percentile was between ~1.5 and ~2.5 $\log_{10}$ for 2010 and between ~1 and ~2.5 $\log_{10}$ for 2015.

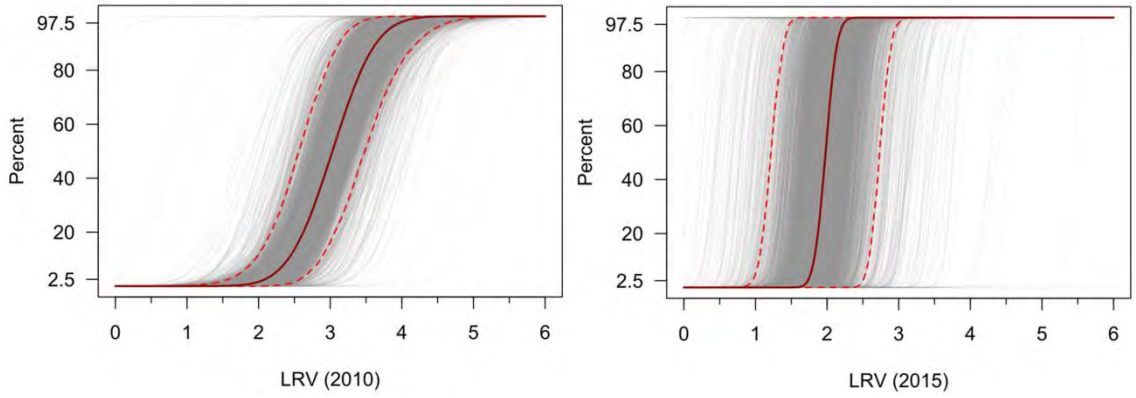

Figure 5-9: Posterior predictive cumulative distribution functions for 2010 and 2015 with 95% credible intervals.

5.3.6 Variance partition coefficient

The Variance Partition Coefficient (VPC) measures the proportion of total variation attributable to higher level units of the model indicating the importance of the hierarchical structure (Goldstein *et al.*, 2002). In this case VPC is defined as follows

$$VPC = \frac{\sigma_{\mu}^2+\sigma_{\delta}^2}{\sigma_{\mu}^2+\sigma_{\delta}^2+\dot{\sigma}^2}$$ Equation 5-9

VPC values close to 1 reveal the importance of the hierarchical model to the quality of the model (Ntzoufras, 2011). The posterior distribution of the VPC (Figure 5-10) indicates that for the year 2010 the variability between skids is small compared to the variability within sets of samples for a particular skid, as the majority of the posterior distribution is located at small proportions (VPC: mean=0.090, 95% CI: (0.002, 0.514)). For 2015, the VPC posterior showed the opposite behaviour compared to 2010, with the variability between the skids strongly contributing towards the total variability (VPC: mean=0.731, 95% CI: (0.215, 0.984)). Accordingly, the overall conclusion from consideration of the VPC, is that the hierarchical structure of the model was less important in 2010 than in 2015. This outcome can also be observed in the comparison between the $NHBM_{same}$ and the HBM_{μ} presented in Figure 5-8 for both years. In the case where the $NHBM_{same}$ is used to construct the skid-independent predictive distributions for 2015 instead of the HBM_{μ}, this distribution resulted in much more optimistic predictions for performances based on 2.5th or 5th percentiles than can realistically be expected to occur. Although the skid-independent distributions generated in both cases ($NHBM_{same}$ and the HBM_{μ}) contain the observations within the 95% credible interval, the distribution generated by HBM_{μ} incorporates variability between skids. This variability is not considered by $NHBM_{same}$ so more precision is

expected in this case. The issue with this aspect is that the prediction of performance of skids not measured can be observed outside of the 95% credible interval for $NHBM_{same}$, if inter-skid variability is not taken into account.

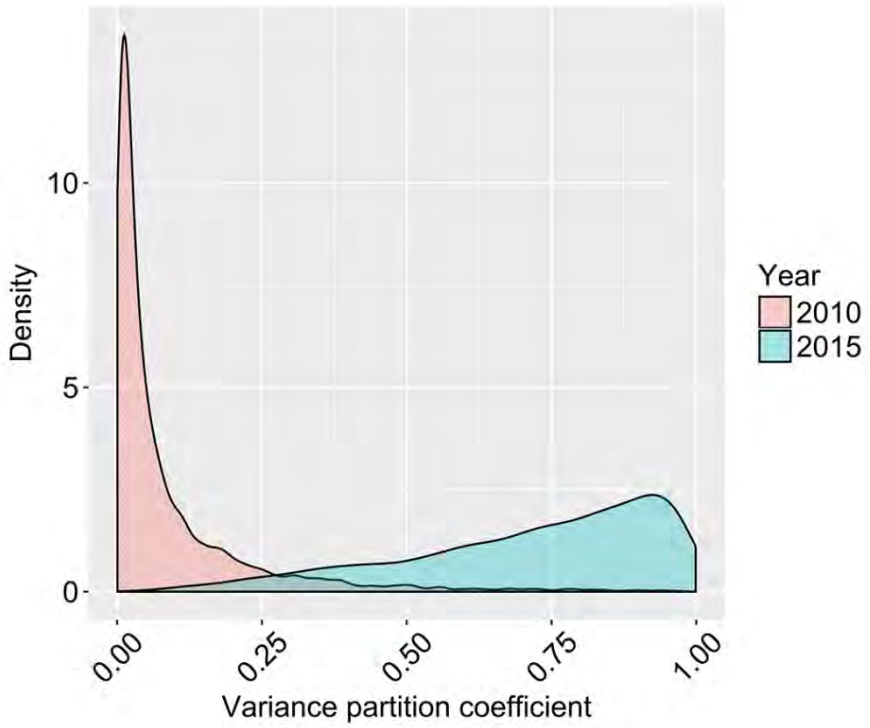

Figure 5-10: Posterior distribution of the variance partition coefficient for the HBM_{μ} in 2010 and 2015.

5.3.7 Censored data

It is important to note that the sample results can be either censored (right, left or interval) or not censored values. It is common when evaluating LRV data to encounter right censored data (i.e. a lower bound on the LRV), because a well operated process will often remove the target microorganism to below the available limit of detection (LOD) in the treated water, particularly when there is a low density in the feed water. Missing and/or censored values typically increase the uncertainty of the estimated parameters and complicate statistical analysis and inter-temporal comparison. Censored data are occasionally omitted or replaced by half the limit of detection (Brown & Mac Berthouex, 2002; Helsel, 2005). However, in this study, censored data are easily managed in this analysis by integrating the respective Normal density term over the censored region, which implies that censored observations are considered in the calculation of distribution parameters. Simply removing or replacing the censored values produces biased estimates of both mean and standard deviation (Brown & Mac Berthouex, 2002). Other methods have been

proposed to deal with this issue including the "trimmed mean", "Winsorized mean", graphical methods and maximum likelihood, among others. Methods including substitution of censored values are only useful when less than 15% of the data are censored in the case of trimmed mean method or less than 25% of the data are censored in the case of Winsorized mean (Brown & Mac Berthouex, 2002). In cases where uncertainty needs to be estimated for the distribution parameters, maximum likelihood and Bayesian methods are required. Bayesian estimation of LRVs from validation data for wastewater treatment processes including left censored values have been studied by Kato *et al.* (2013); Ito *et al.* (2015); Ito *et al.* (2016), and Kato *et al.* (2016). These studies have also analysed the effect of the number of non-censored observations on the accuracy of LRVs and the influence on distribution parameters of pairing of influent to effluent samples.

5.4 Conclusions

The Hierarchical Bayesian model (HBM) presented in this study provided a number of advantages over independent skid analyses, or analyses based on pooling data over skids. These included the use of individual skid data to inform population parameters, mutual sharing of information between individual skids, a simple comparison of performance between skids and years, and skid independent prediction of future performance. In this work, five candidate model forms (three HBM and two NHBM) were proposed and evaluated for prediction of virus LRV originating from UF challenge test data. Of the candidate model forms evaluated, HBM_{μ} (exchangeable parameters for u) appeared to be the most appropriate based on consideration of DIC, alongside goodness of fit and effective number of parameters, indicating the significant potential in application of HBMs for analysis and prediction of challenge test data. The HBM_{μ} was then applied to compare the LRV obtained for the UF system when new, and after 5 years of operation. As expected, UF LRV performance had decreased after 5 years of operation with a posterior mean of the overall difference between years for all skids (μ_{δ}) of -1.1 (95% CI: (-1.4, -0.7)). The difference on log standard deviation between years (ε) had a mean of -1.4 (95% CI: (-1.9, -0.9)). Both of these outcomes are significant, as although membrane ageing may be expected to result in lower removal performance, the results presented here suggest that this loss in performance may remain more stable after prolonged operational periods.

Chapter 6: Parameters for ozonation performance assessment

This chapter has been published in the following journal article:

Carvajal, G., Branch, A., Michel, P., Sisson, S.A., Roser, D.J., Drewes, J. E. and Khan, S.J. (2017). Robust evaluation of performance monitoring options for ozone disinfection in water recycling using Bayesian analysis. Water research. DOI: 10.1016/j.watres.2017.07.079

6.1 Introduction

Ozonation of wastewater provides effective reduction of colour, odour, UV_{254} absorbance (Ni *et al.*, 2002; Paraskeva & Graham, 2005), and trace organic chemicals (Hübner *et al.*, 2015). In addition, ozone can be an effective disinfectant of more resistant pathogens such as *Cryptosporidium* and *Giardia* (Gamage *et al.*, 2013). It complements biological filtration by transforming finer and recalcitrant organic matter into more biodegradable compounds (Reaume *et al.*, 2015). In ozonation of secondary treated sewage, removal efficiencies for microbial indicators in log_{10} reduction value (LRV) units have been reported to range from 2 to 8 LRVs for *E. coli* (Ni *et al.*, 2002; Paraskeva & Graham, 2005; Tripathi & Tripathi, 2011; Gerrity *et al.*, 2012; Gamage *et al.*, 2013), 0-1.7 LRVs for *C. perfringens* (Xu *et al.*, 2002), and 1-8 LRVs for FRNA coliphage (Xu *et al.*, 2002; Gerrity *et al.*, 2012; Gamage *et al.*, 2013; Merlo *et al.*, 2015). Similar indicative LRVs have been reported for pathogenic species of enteric bacteria (2-6 LRV), enteric viruses (3-6 LRV), *Giardia* (2-4 LRV), and *Cryptosporidium* (1-2 LRV) (NRMMC *et al.*, 2008). Despite common and accepted usage, it has been reported that concentration integrated over contact time (CT) may not correlate well with LRV following ozonation due to the rapid depletion of ozone in the presence of oxidisable organic matter (Xu *et al.*, 2002). Alternative water quality parameters have been proposed, including change in UV_{254} absorbance (ΔUVA), change in total fluorescence (ΔTF), and ozone demand (ratio of ozone dose to total organic carbon (O_3:TOC) or dissolved organic carbon (O_3:DOC)), that indirectly indicate the oxidation demand of organic material and reliability of the system for microbial disinfection (Gerrity *et al.*, 2012; Gamage *et al.*, 2013). Although removal of trace organic compounds has been more reliably predicted using these control parameters, predicting the inactivation of microorganisms remains uncertain. In response ozone demand (O_3:TOC or O_3:DOC ratio) has been recommended as a parameter to compare potential ozonation effectiveness on water of different origins (Buffle *et al.*, 2006). Alternatively formation of bromate has been proposed as a potential predictor for microbial disinfection during drinking water ozonation by Von Gunten et al.(2001). However, low concentrations of bromide and interferences during quantification (i.e. via ion chromatography) produced by the presence of high concentrations of chloride in the treated wastewater matrix restrict the use of bromate as a surrogate. Some recent studies have attempted to find a general relationship between non-microbial surrogates and microorganism LRV by combining data obtained from different plants around the world (Gerrity *et al.*, 2012) and within the United States (Gamage *et al.*, 2013). These studies were interrelated with the data used by Gamage *et al.* (2013) being included in the study conducted by Gerrity *et al.* (2012). Although uncertainty was large, statistically significant correlations between microbial and non-microbial surrogates were reported in both studies. Although not measured, there was still unfortunately

appreciable uncertainty in these latter studies (Gerrity *et al.*, 2012; Gamage *et al.*, 2013). The work employed seeded, laboratory-cultivated, standard microorganism rather than autochthonous species. Also seeded microorganism could be less resilient to disinfection chemicals because they may not be as shielded by particles or biofilms. In addition, only one strain of each target microbial group was used in seeded studies in comparison to autochthonous microorganisms which are necessarily more diverse (Shelton *et al.*, 2006; Nenonen *et al.*, 2012). Other potential sources of uncertainty are variations between analytical methods, including detection limits, analysis times, potential interferences and storage and handling requirements. All of these can increase uncertainty when attempting to quantify associations between microbial and non-microbial surrogates and disinfectant effectiveness. To appropriately incorporate the resulting uncertainty, robust methods are needed which assess relationships between key disinfection variables and capture the influence of these uncertainties.

Bayesian data analysis techniques are one potential approach which meets these needs. They can incorporate prior information from multiple known variables, and provide estimates of the magnitudes and likelihoods of candidate variables. The application of Bayesian analysis techniques to water treatment processes is relatively new, having been applied to incorporate uncertainty and interpolate CT values during chlor(am)ination of secondary effluent (Carvajal *et al.*, 2017) and to predict microbial disinfection rates during UV disinfection (Qian *et al.*, 2004) and chlorination (Sivaganesan *et al.*, 2003) for drinking water treatment. This study has assessed the applicability of Bayesian analysis to wastewater ozonation using a bench scale ozonation system. In addition, for comparison, the proposed method and findings have been applied to results obtained by Gamage, Gerrrity and colleagues. In this present study, autochthonous somatic coliphages, *Clostridium perfringens*, *Escherichia coli,* and total coliforms indigenous to secondary treated effluent were analysed. This approach compares with Gamage, Gerrity and colleagues' selection of type strain *E. coli* and MS-2 F-RNA coliphage. In line with previous work, surrogate parameters assessed for monitoring microbial removal efficiency included ozone demand, ΔUVA, ΔTF, and components concentrations obtained from Fluorescence excitation emission matrices (EEMs) through Parallel Factor (PARAFAC) analysis (Harshman, 1970). This study evaluated the sensitivity of the removal efficiency with the predictors and generated regression models to test their reliability. A Bayesian analysis was used to estimate the uncertainty related to regression parameters. Predictive capabilities of the predictors were evaluated through a number of metrics including the Deviance Information Criterion (DIC) score, adjusted R squared ($R^2_{adjusted}$), mean squared error (MSE) using leave-one-out cross-validation, and visual assessments of posterior predictive distributions.

6.2 Materials and methods

6.2.1 Bench-scale ozonation

Experiments were conducted using an ozone batch reactor with a maximum working capacity of 3 L, which was equipped with inline sensors capable of monitoring ozone concentration and flow rate on gas inlet and exhaust streams (Figure 6-1). The reactor included temperature control which was maintained at 20 °C and a constant stirring speed of 700 rpm for all experiments. A simple condenser was used as a dehumidifier on the ozone gas outlet of the reactor (BMT DH 3b, Germany), to remove known measurement interferences due to liquid water prior to entering the ozone analyser. Given the dehumidifier was a simple, low surface area condenser it is unlikely that significant loss of ozone occurred due to its installation. In the unlikely event that ozone loss occurred due to the dehumidifier, the results would estimate a higher than true ozone dosage required to achieve a certain LRV. In this regard, the obtained results are believed to be conservative. The ozone generator (BMT 803 BT, Germany) produced ozone from pure oxygen. Ozone concentrations were measured via UV ozone photometers (BMT 964, Germany) installed on the gas inlet and the exhaust pipework of the reactor. A LabVIEW virtual instrument (National Instruments, USA) was used to process sensor data and calculate the applied (Equation 6-1) and consumed ozone dosages (Equation 6-2) (mg/L).

$$Applied\ ozone\left[\frac{mg}{L}\right] = \int_0^t \frac{[O_3]_{in} \cdot Q_{gas}}{V_L} \cdot dt \qquad \text{Equation 6-1}$$

$$Consumed\ ozone\left[\frac{mg}{L}\right] = \int_0^t \frac{([O_3]_{in} - [O_3]_{out}) \cdot Q_{gas}}{V_L} \cdot dt \qquad \text{Equation 6-2}$$

Where $[O_3]_{in}$ and $[O_3]_{out}$ are the concentration of ozone in the inlet and outlet gas streams, respectively; Q_{gas} is the gas flowrate, and V_L is the volume of liquid in the reactor; and t represents the reaction time.

Prior to each experiment, the reactor was filled with 1.5 L of secondary treated wastewater obtained from a full-scale wastewater treatment plant in Garching, Germany, which had been filtered through a 4-7 µm cellulose filter (Macherey-Nagel MN 1672) to simulate tertiary filtration to minimise the potential interference of variable solids size and loading. Five sets of experiments were conducted, with each one including five different applied ozone dosages (2, 4, 6, 8 and 10

mg/L). The first set of experiments used applied ozone dosages 1.6, 3.2, 4.8, 6.4 and 8 mg/L. Although they were different to the values used in the other experiments, they were also included in the dataset. The inclusion of these data was considered valid as the purpose of this study was to find relationships between LRVs and monitoring parameters using a continuous range of values rather than discrete point conditions. The duration of each experiment was approximately 10 minutes including ozone injection, which lasted between 30 and 120 seconds. To summarise reactor mixing performance and ozone adsorption behaviour, applied and consumed ozone dosages were averaged for each run and included in Appendix 3.

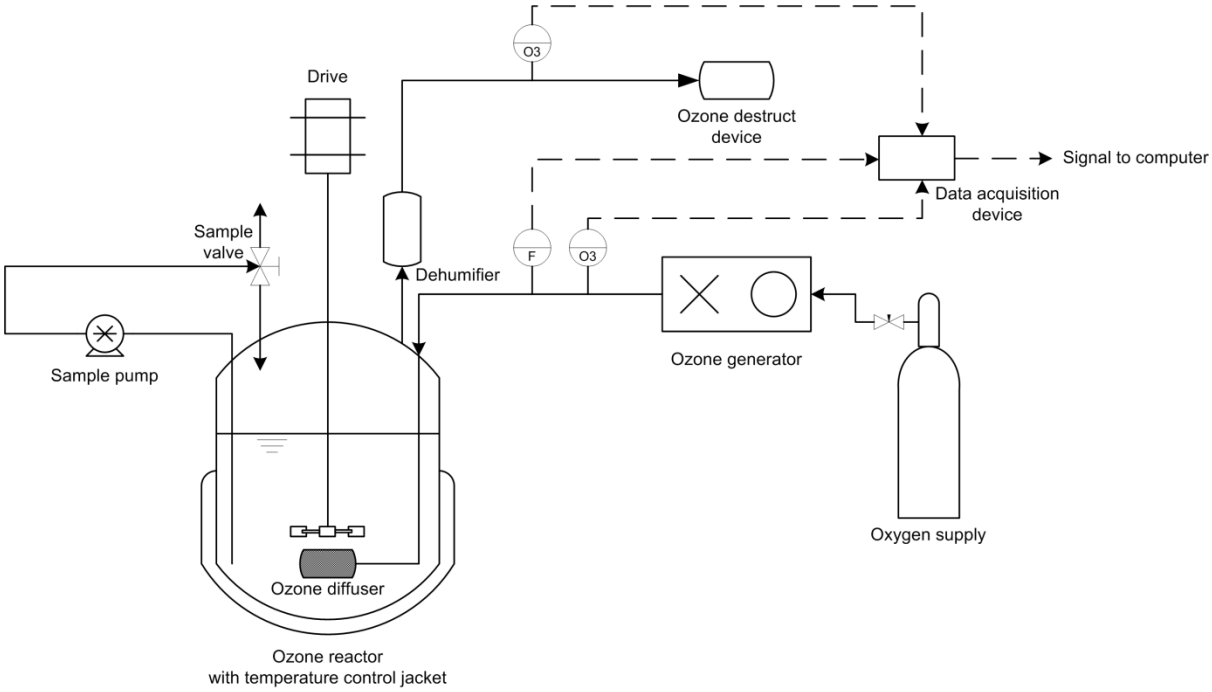

Figure 6-1: Process diagram of the bench-scale ozonation system

6.2.2 Measures of water quality

UV absorbance at 254 nm for each sample was measured using a Shimadzu UV-1601 (Shimadzu Europa GmbH) UV/visible spectrometer. Fluorescence excitation emission matrices (EEMs) were obtained using a Horiba Aqualog Fluorescence spectrophotometer (Horiba Instruments Inc., Edison, NJ, USA). The scanning wavelength ranges were 230-599 nm for both excitation and emission wavelengths. Interfiltering effects were expected to be small as UV absorbance at 254 nm was less than 0.16 absorbance units for all samples. Due to the low expected level of interfiltering, EEM measurements were performed on undiluted filtered samples. EEM intensities were standardised by subtraction of a blank, normalisation to Raman units (R.U.) and corrected for interfiltering effects using Aqualog software and then processed using PARAFAC (Harshman,

1970) with the software package Solo (Eigenvector Research Inc. Manson, USA). PARAFAC is a statistical tool used to decompose 3D-EEM data and allow the identification and quantification of independent underlying signals known as "components" (Murphy *et al.*, 2013). DOC samples were collected in glass vials, filtered through 0.45 μm cellulose nitrate filters (Sartorious 11306-47-N) and acidified to pH < 3 (with hydrochloric acid) prior to analysis in Vario TOC cube (Elementar GmbH). Turbidity was quantified using a Hach 2100 QiS turbidity meter (HACH, USA).

The obtained EEM data were supplemented with comparable data obtained from the full-scale wastewater treatment plant Munich Gut Marienhof for the PARAFAC analysis, to increase the dataset size and improve model resilience and accuracy. The generated model included 90 wastewater samples from different sections of the treatment process, covering water matrices obtained from the plant influent, after biological nutrient removal, denitrifying tertiary granular media filtration, UV disinfection, and an advanced oxidation process (UV/H_2O_2). To assure PARAFAC model accuracy, three diagnostic metrics were used, including Core Consistency, Split Half Analysis, and Total Variance. The use of only one metric can often provide misleading information, therefore more than one metric is usually recommended (Murphy *et al.*, 2013). Core consistency indicates the appropriateness of the model with values ranging between 0% and 100%, where 100% represents the perfect score (Kompany-Zareh *et al.*, 2012). The split half analysis measures the agreement between excitation and emission loading modelled for the calibration and test datasets. Total variance measures the percentage of variability explained by the generated model.

Somatic coliphages were quantified using the double agar layer technique (SM 9224 C) (APHA, 2005), using *E. coli* CN-13 (American Type Culture Collection #700609) as the host and Phi X174 coliphage (ATCC # 13706-B1) as the positive control. *C. perfringens* were enumerated using the tryptose sulphite cycloserine (TSC) agar (Merck 1.11972.0500), and incubated anaerobically at 37 °C for 24 h. *C. perfringens* samples were pre-treated at 60 °C for 15 minutes to inactivate vegetative bacteria (ISO, 2013b). Chromocult (Merck 1.10426.0500) was used to enumerate both *E. coli* and total coliforms, which were incubated at 37 °C for 24 h. Bacterial indicators after ozonation were quantified using the membrane filtration inoculation technique (Method 9215D, (APHA, 2005)). In this case a desired volume of sample (typically 5, 50 and 100 mL) was filtered through a 47 mm diameter, 0.45 μm gridded filter membrane (Millipore, S-Pak, type HA). The filter membrane was then transferred onto the surface of a well dried plate of selective agar. For *C. perfringens*, the filter was overlaid with a layer of molten TSC agar cooled to below 40 °C to prevent colony decolouration by oxygen exposure. *C perfringens, E. coli* and

total coliforms were quantified in colony forming units per 100 mL (CFU/100mL) and somatic coliphages in plaque forming units per 100 mL (PFU/100mL).

6.2.3 Sample collection

After performing the experiment and once the ozone concentration in the off-gas was determined to be negligible (0.01 mg/L), the ozone purging (with oxygen) was stopped and the sample was taken from the sample port. Microbial analyses were performed the same day and the analysis of bulk organic parameters was performed within 48 h.

6.2.4 Parameters of ozone effectiveness

As discussed in the introduction, there are a number of possible primary and derived measurements proposed for assessing ozonation disinfection including ΔUVA, ΔTF, ozone demand and bromate formation. Previous studies have shown that some of them are more promising for predicting the removal of microorganisms during ozonation of wastewater. In this study ΔUVA, ΔTF and ozone demand were selected based on the studies conducted by Gamage *et al.* (2013) and Gerrity *et al.* (2012). Additionally, reductions in "component" concentrations obtained through PARAFAC processing were also estimated.

Somatic coliphages were chosen because they are usually present in secondary treated effluent in higher numbers than the commonly used FRNA coliphage. *C. perfringens* spores have been recommended as a suitable indicator for ozonation due to their high oxidant resistance (Tyrrell *et al.*, 1995; Xu *et al.*, 2002). *E. coli* and total coliforms were selected as they represent common enteric bacterial pathogens and are proven indicators of faecal contamination (Asano *et al.*, 2007). LRVs were calculated by taking the Log_{10} ratio for each different microorganism density assayed before and after ozone exposure. Total fluorescence (TF) was obtained from the summation of all the intensities for each EEM. Delta TF and UVA were obtained from the subtraction of the initial measurement value and the after-treatment measurement, all divided by the initial measurement value.

Data from previous studies (Gerrity *et al.*, 2012; Gamage *et al.*, 2013) were extracted into a .CSV file using the WebPlotDigitizer tool (Rohatgi, 2011) which permits extraction of data from plots.

6.2.5 Data analysis software and Bayesian analysis

Parametric Bayesian analysis was used in this study because generation of an equation to define the association between operational parameters and process performance is necessary. Bayesian

analysis also allows the inclusion of priors and assess uncertainty and variability explicitly. Bayesian analysis for this type of models was introduced in Section 3.2.

Bayesian methods for analysing regression models were implemented using the *JAGS* software package (Plummer, 2003). Model construction and inference were conducted in *R* through the open source *r2jags* package (Su & Yajima, 2015). The data were analysed using linear and non-linear regression models to compare the results between studies and check the fit to the data. All models assumed normal distribution of the errors. Accordingly, the response variable (y) is normally distributed with mean (μ) (dependent on the predictor x) and standard deviation (σ). In mathematical notation this is $y_i \sim Normal(\mu_i,\sigma)$ for i=1,..,n. Diffuse prior distributions for the parameters were specified as Normal(0,100) for the parameters of the linear or non-linear relationships and Gamma(0.001,0.001) for the precision ($\tau = \sigma^{-2}$) (Lunn *et al.*, 2012). Scatterplots showing the association in the data, model fit and 95% posterior predictive intervals were constructed for each combination of microbial indicators and surrogate parameters. The Deviance information criterion (DIC) (Spiegelhalter *et al.*, 2002), adjusted coefficient of determination ($R^2_{adjusted}$), and mean squared error from leave-one-out cross validation (LOOCV) were employed to compare the performance of the predictor variables. DIC is used to compare the goodness of fit between the proposed models. LOOCV was used to test the model with observations not previously used in the construction of the model. In this case the metric used to quantify the model's predictive abilities was the MSE. Mean squared error measures the deviation of the predicted values $\hat{y}_i$ from the real values y_i, with higher MSE values indicating larger error in the prediction. The average MSE under posterior was used to compare the models. As part of this procedure, the data were normalised (0-1) before computing the metrics so that results could be compared across microorganisms and different studies.

6.3 Results

6.3.1 Water quality measures

Common water quality parameters for the secondary effluent used in the ozonation experiments are summarised in Table 6-1. Nitrite was not measured because the water was stored for 3 to 5 hours after the filtration step before being used for the ozonation experiments. Additionally, at the initial phase of each experiment, the water was aerated with oxygen to zero the O_3 sensors. Nitrite concentration was measured during one of the trials to confirm its low value (< 0.02 mg/L).

Table 6-1: Summary of secondary treated wastewater quality measurement

Parameter (units)	**Range[b,c]**
DOC (mg/L)	6.2-14
pH	7.4-8.0
Turbidity (FNU)	0.8-2.4
Conductivity (mS/cm)	1.13-1.22
Alkalinity (mg/L as $CaCO_3$)	216-250
UVA influent	0.13-0.16
UVA reduction, ΔUVA (%)	10-48
TF reduction, Δ TF (%)	32-77
Ozone demand (O_3:DOC) (mg O_3/mg)	0.08-0.67
E. coli (log_{10} (CFU/100 ml))[a]	2.8-5.0
Total coliforms (log_{10} (CFU/100 ml))	4.2-7.0
C. perfringens (log_{10} (CFU/100 ml))	2.5-3.1
Somatic coliphage (log_{10} (PFU/100 ml))	3.2-3.5

[a] CFU: colony forming units

[b] Five batches of wastewater were each analysed.

[c] For each reduction e.g. Δ TF, LRV; 25 measurements were made per analyte.

EEMs were processed and two components were found to be present through PARAFAC analysis. The methods and analytical adjustments were based on the work conducted by Stahlschmidt *et al.* (2016). The three diagnostic metrics indicated that the PARAFAC model was valid and sufficiently resilient to be applied. The core consistency for the two-component model achieved 100% for the calibration dataset and 95% for test dataset (i.e. data used in this study). The split half analysis resulted in 93% indicating a satisfactory agreement. The total variance score was 83.6%, suggesting that the model was able to explain a large proportion of the variance in the data. The received Excitation/Emission peak combinations for each identified component were $\lambda_{ex/em}$=233/442 nm (component 1) and $\lambda_{ex/em}$=230/349 nm (component 2) as shown in the component contour plot in Figure 6-3. Fluorescence intensity was gradually reduced for the five

increasing O_3:DOC ratios as observed in Figure 6-2. The components found through PARAFAC analysis suggested that the organic matter can be associated to humic acid-like compounds (Peak A) (Coble, 1996) and tryptophan-containing compounds (Peak T2) (Hudson *et al.*, 2007).

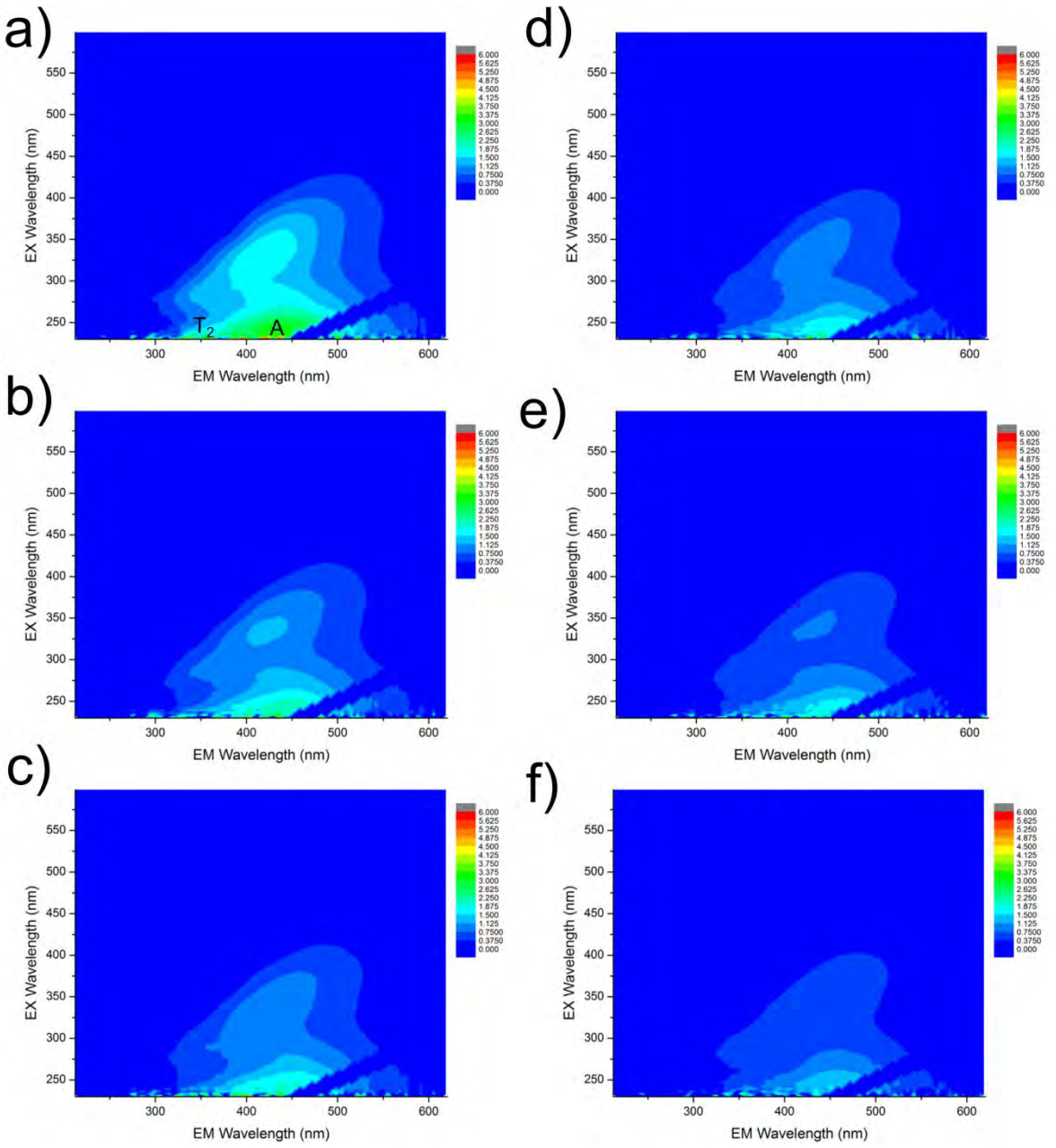

Figure 6-2: Extract of results of initial EEM spectra for the initial secondary effluent water matrix (a) and EEM for the five different ozone dosages , 2 (b), 4 (c), 6 (d), 8 (e) and 10 mg/L (f). Peaks A and T_2 are indicated on (a).

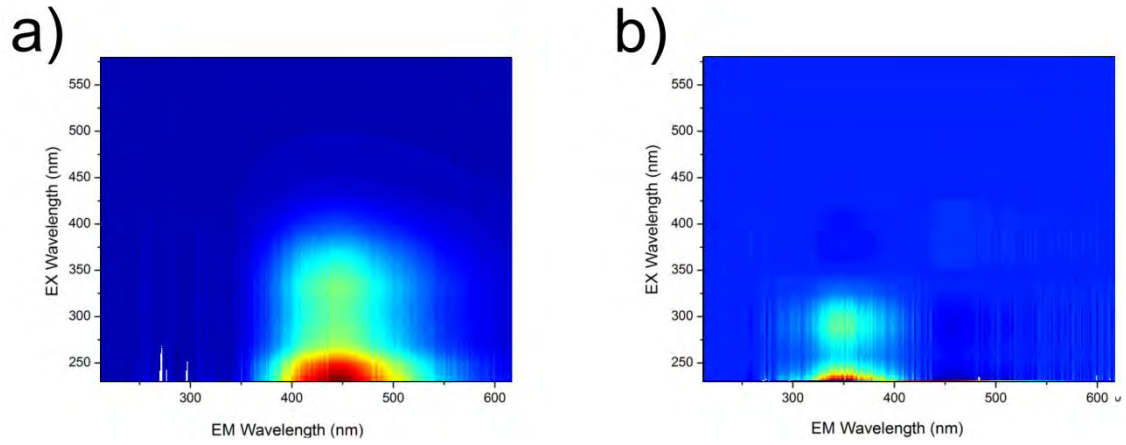

Figure 6-3: Extracted results of the model components (a = component 1, b = component 2)

6.3.2 Quantifying the relationships between ozone parameters

Ozone demand (O_3:DOC or O_3:TOC) was plotted against ΔUVA and ΔTF (Figure 6-4 a-d). The ozonation parameters showed a positive relationship between these variables with low uncertainty in fitted power functions ($y = a.X^b$) as used by Gamage *et al.* (2013) and Gerrity *et al.* (2012). The decreasing slope with increasing ozone demand (Figure 6-4) can be explained by the depletion of reactive moieties by organic matter (Buffle *et al.*, 2006). The coefficients for the power function for this current study and the study conducted by Gamage *et al.* (2013) are presented in Table 6-2. Coefficients include their 95% credible interval (CI) in parentheses. The empirical model fitted the data of both studies well, with an average $R^2_{adjusted}$ for ΔUVA of 0.89 for this current study and 0.90 for the previous study (Gamage *et al.*, 2013), and an average $R^2_{adjusted}$ for ΔTF of 0.86 for this current study and 0.85 for the previous study (Gamage *et al.*, 2013).

Table 6-2: Coefficients of the fitted curve for the relationship between water quality parameters

Variables	**Coefficients for this study a(95% CI), b(95% CI)**	**Coefficients for Gamage *et al.* (2013)[a] a(95% CI), b(95% CI)**
O_3:DOC (TOC) -ΔUVA	0.71(0.63 ,0.80), 0.75(0.63, 0.88)	0.48(0.46, 0.51), 0.53(0.44, 0.63)
O_3:DOC (TOC) -ΔTF	0.98(0.90,1.07), 0.46(0.38, 0.55)	0.87(0.83, 0.91), 0.29(0.24, 0.36)

[a] Gerrity *et al.* (2012) did not report their results for the relationship between ozone parameters.

The relationships between ozone demand, ΔUVA, and ΔTF for this current study (Figure 6-4 a and b) and the study conducted by Gamage *et al.* (2013) (Figure 6-4 c and d) indicate that for the present study, the mean ratio of the consumed ozone to the applied ozone was approximately 50%. For the experimental ranges of ozone demand used in the present study, the maximum achieved removals were 48% for UVA and 77% for TF while Gamage *et al.* (2013) achieved a maximum removal of 63% for UVA and 95% for TF. Although, this current work suggested significant differences to the previous studies (Gerrity *et al.*, 2012; Gamage *et al.*, 2013) in terms of operational settings, the ozonation effects were still of similar magnitudes. Also, the data analyses provided a comparison of the sensitivity of nominally similar microorganisms to ozone and estimates of the uncertainty obtained by fitting the models. These results were achieved despite using a lower ozone demand range.

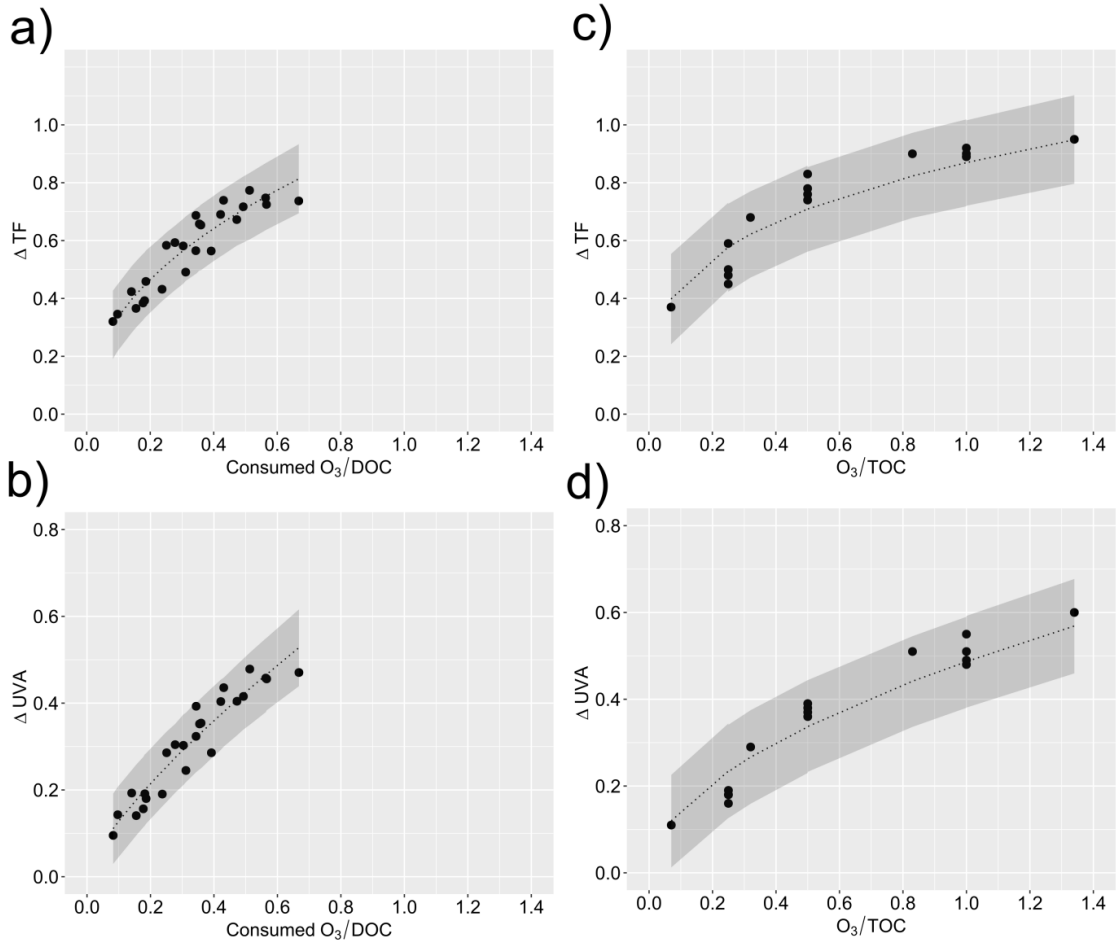

Figure 6-4: Relationship between ozone demand (O_3:DOC) and ΔUVA and ΔTF for this current study (a-b) and O_3:DOC and ΔUVA and ΔTF obtained from Gamage *et al.* (2013) (c-d). The dotted line indicates the power function model calculated from the median of the posterior predictive distributions of the corresponding data set. Shaded region indicates the 95% posterior predictive invervals.

6.3.3 Relationship between ozonation parameters and disinfection performance

6.3.3.1 Analysis of predictors from the current study

Standard linear models with ($y=a+bx+cx^2$) and without a quadratic term ($y=a+bx$) were fitted through Bayesian methods to the data, using the surrogate parameters and LRV for each microbial indicator. The results showed similar fits between all three surrogate parameters and LRVs for all the indicators (Figure 6-5, Figure 6-6). In general, models incorporating a quadratic term provided similar fit to the univariate linear regression models. Therefore, for simplicity, only the detailed results for the linear models are presented here. All 95% credible intervals for the posterior distributions of the slopes of the linear models excluded zero and were positive, indicating strong evidence of a positive association between the surrogates and microbial indicator LRVs. For *E. coli*, ΔTF and ΔUVA showed better predictive capabilities than ozone demand according to the assessment metrics applied (DIC, $R^2_{adjusted}$ and MSE) (Figure 6-5 a-c). The models' slopes revealed that a 23% (95% credible interval (CI): (18, 29)) change in TF corresponded to a one unit change in LRV (Figure 6-5 b), while a 20% (95% CI: (15, 27)) change in UVA corresponded to a unit change in LRV (Figure 6-5 c), which implies that ΔTF may be more sensitive for detecting LRV change than ΔUVA.

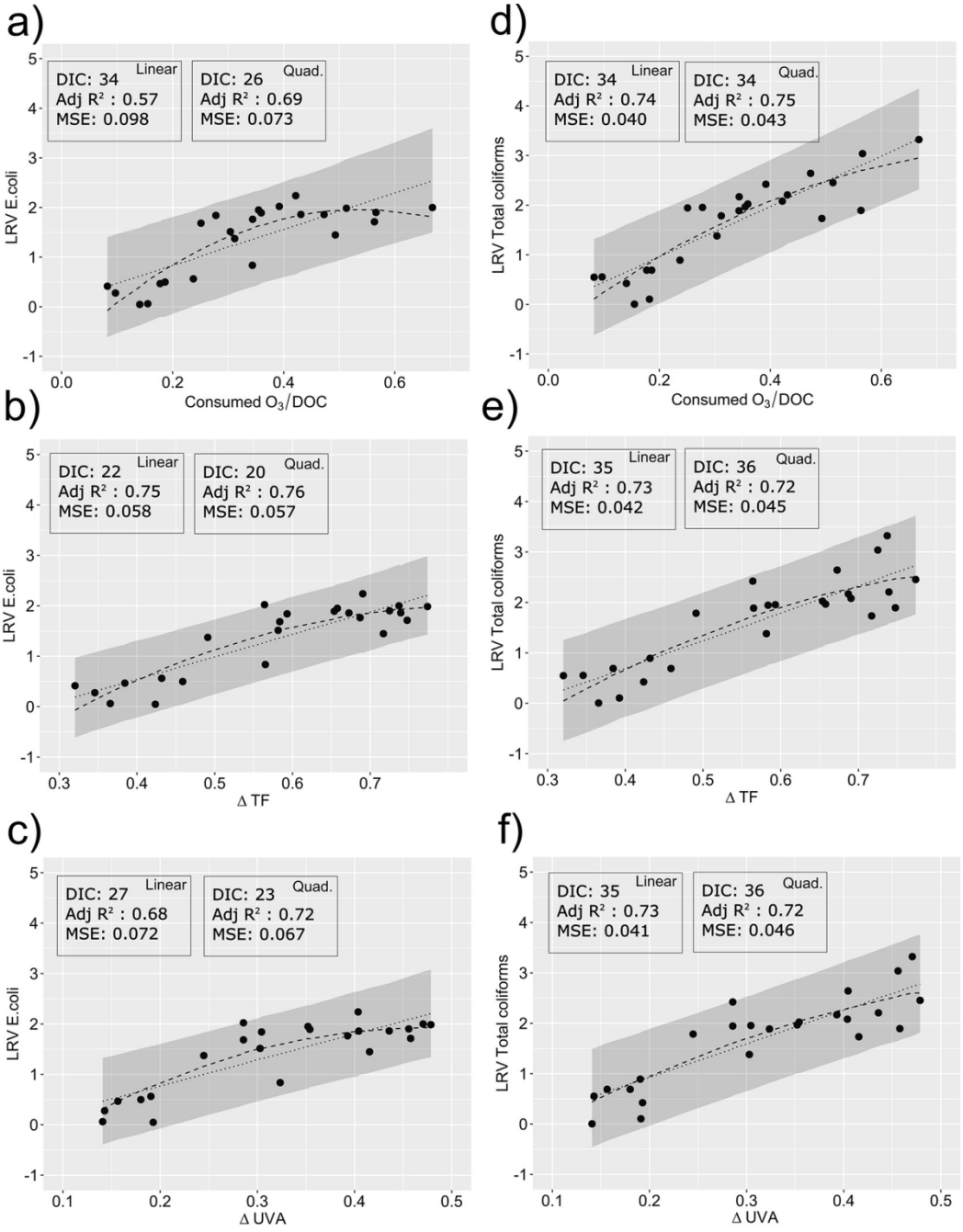

Figure 6-5: Scatterplots with linear fit between LRV for *E. coli* and ozone demand (a), ΔTF (b) and ΔUVA (c); and scatterplots with linear fit between between LRV for total coliforms and ozone demand (d), ΔTF (e) and ΔUVA (f). The dotted line indicates the linear regression model calculated from the median of the posterior predictive distributions of the corresponding data set. The dashed line indicates the quadratic regression model calculated from the median of the posterior predictive distributions of the corresponding data set. Shaded region indicates the 95% posterior predictive invervals for the linear regression model.

For total coliforms, the assessment metrics (Figure 6-5 d-f) presented similar fits with all ozonation parameters. From the model, 18% (95% CI: (15, 24)) of change in TF corresponded to a one unit change in LRV, while a 15% (95% CI: (12, 20)) in UVA corresponded to a unit change in LRV. Total coliforms behaved similarly to *E. coli*. The lowest MSE of all the microorganisms were obtained for total coliforms across all three predictors.

For *C. perfringens* the assessment metrics suggested a better association with ΔTF and ΔUVA than ozone demand (Figure 6-6 a-c). *C. perfringens* was more resistant than the other studied microorganisms with a maximum of approximately 1 LRV. Based on the models' slopes, a 71% (95% CI: (53, >100)) reduction in TF corresponds to a unit change in LRV while for UVA, a 59% (95% CI: (43, 91)) reduction corresponds to a unit change in LRV. Similar to *E. coli* and total coliforms, ΔTF was more sensitive to change in LRV of *C. perfringens* than UVA.

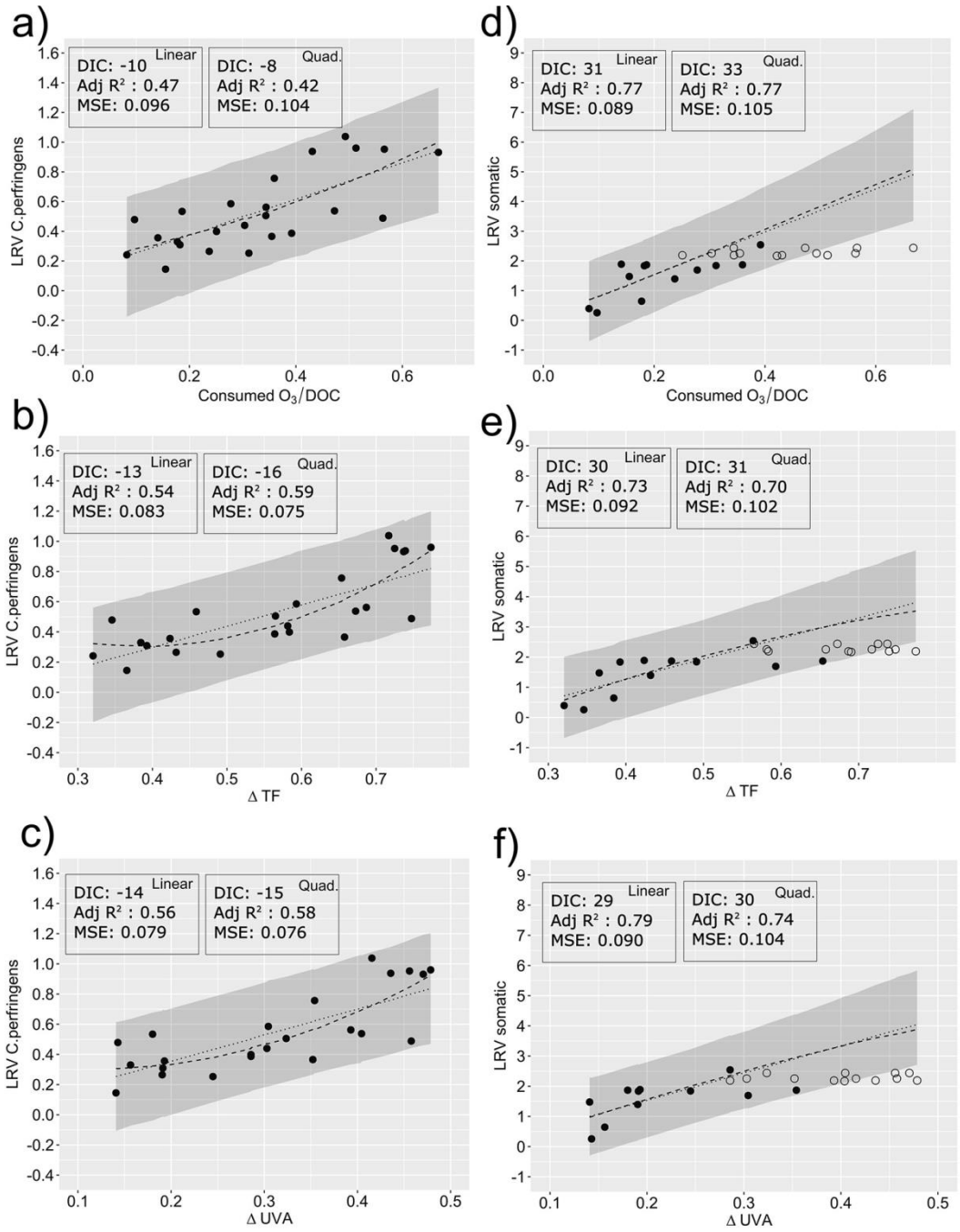

Figure 6-6: Scatterplots with linear fit between LRV for *C. perfringens* and ozone demand (a), ΔTF (b) and ΔUVA (c); and scatterplots with linear fit between between LRV for somatic coliphages and ozone demand (d), ΔTF (e) and ΔUVA (f). Open circles indicate removal is higher than the recorded value due to censored data. The dotted line indicates the linear regression model calculated from the median of the posterior predictive distributions of the corresponding data set. The dashed line indicates the quadratic regression model calculated from the median of the posterior predictive distributions of the corresponding data set. Shaded region indicates the 95% posterior predictive invervals for the linear regression model.

For somatic coliphages the removal efficiency was sufficiently high that 52 percent of observations post-treatment, were below the detection limit of 10 PFU/100mL, resulting in a high proportion of censored data. Nonetheless, using the Bayesian analysis techniques, it was still possible to generate the linear regression coefficients by integrating the respective Normal density term over the non-detected region. In this case, the assessment metrics indicated a similar fit for all the surrogate parameters (Figure 6-6 d-f). It can be observed in Figure 6-6 (d-f) that there was a positive relationship between the evaluated predictors and the somatic LRV. According to the slope results, a 14% (95% CI: (10, 23)) of change in TF corresponds to a unit change in LRV. In the case of UVA, an 11% (95% CI: (7, 17)) change corresponds to a unit change in LRV.

6.3.3.2 Use of components of EEM spectra to predict microbial removal

The two model components found through PARAFAC were analysed separately and together for their predictive capabilities using univariate and multivariate linear regression models, respectively. The results (Table 6-3) indicated that component 1 reduction ($\lambda_{ex/em}$=233/442 nm) provided a better fit than component 2 reduction ($\lambda_{ex/em}$=230/349 nm) for all studied microorganisms except from *E. coli*. The assessment metrics results for component 1 provided similar conclusions to the results obtained for ΔTF (Figure 6-5, Figure 6-6) except from *C. perfringens*. No improvement was observed when using both components together. Overall PARAFAC analysis provided no better indication of LRV magnitude than the other ozonation measures.

Table 6-3: linear model comparison according to DIC score, $R^2_{adjusted}$ and MSE for this study using components

Microorganism	Component(s)	DIC	R-squared	MSE
E. coli	1	22	0.75	0.058
	2	20	0.78	0.053
	1 and 2	21	0.78	0.058
Total coliforms	1	32	0.76	0.038
	2	38	0.70	0.047
	1 and 2	35	0.74	0.042
C. perfringens	1	-11	0.51	0.089
	2	-10	0.48	0.094

Microorganism	Component(s)	DIC	R-squared	MSE
	1 and 2	-9	0.46	0.099
Somatic coliphage	1	31	0.72	0.100
	2	34	0.66	0.112
	1 and 2	33	0.72	0.153

6.3.3.3 Analysis of predictors from previous studies

When Bayesian techniques were applied to the data reported from previous studies, the fitted linear models indicated that in all cases the 95% credible interval for the posterior distributions on the slopes excluded zero and were positive, as in this current study. Unlike the data collected by Gamage *et al.* (2013), the results reported by Gerrity *et al.* (2012) included right-censored values (16% for *E. coli* and 40% for MS2). $R^2_{adjusted}$ and MSE results (Figure 6-7 and Figure 6-8) and indicated higher uncertainty in the models fitted to seeded *E. coli* in both previous studies (Gerrity *et al.*, 2012; Gamage *et al.*, 2013) compared to the microbial indicators from the experiments conducted in this study. In contrast, MS2 models produced more adequate model fits, and provided comparable results to the results obtained in this current study (Figure 6-7 and Figure 6-8). For *Bacillus subtilis* spores, most of the values were close to zero and after reaching a specific change of UVA and TF, a positive relationship was observed, leaving too few data points to fit the model. The resulting metrics (Figure 6-7 and Figure 6-8) suggested that for *E. coli,* all three predictors were similar in their predictive potential, whereas for MS2 both ΔTF and ΔUVA provided better prediction than the O_3:TOC ratio. Models with a quadratic term were also fitted to the data, resulting in a similar fit to the linear models according to $R^2_{adjusted}$ and MSE for *E. coli* and $R^2_{adjusted}$ for MS2.

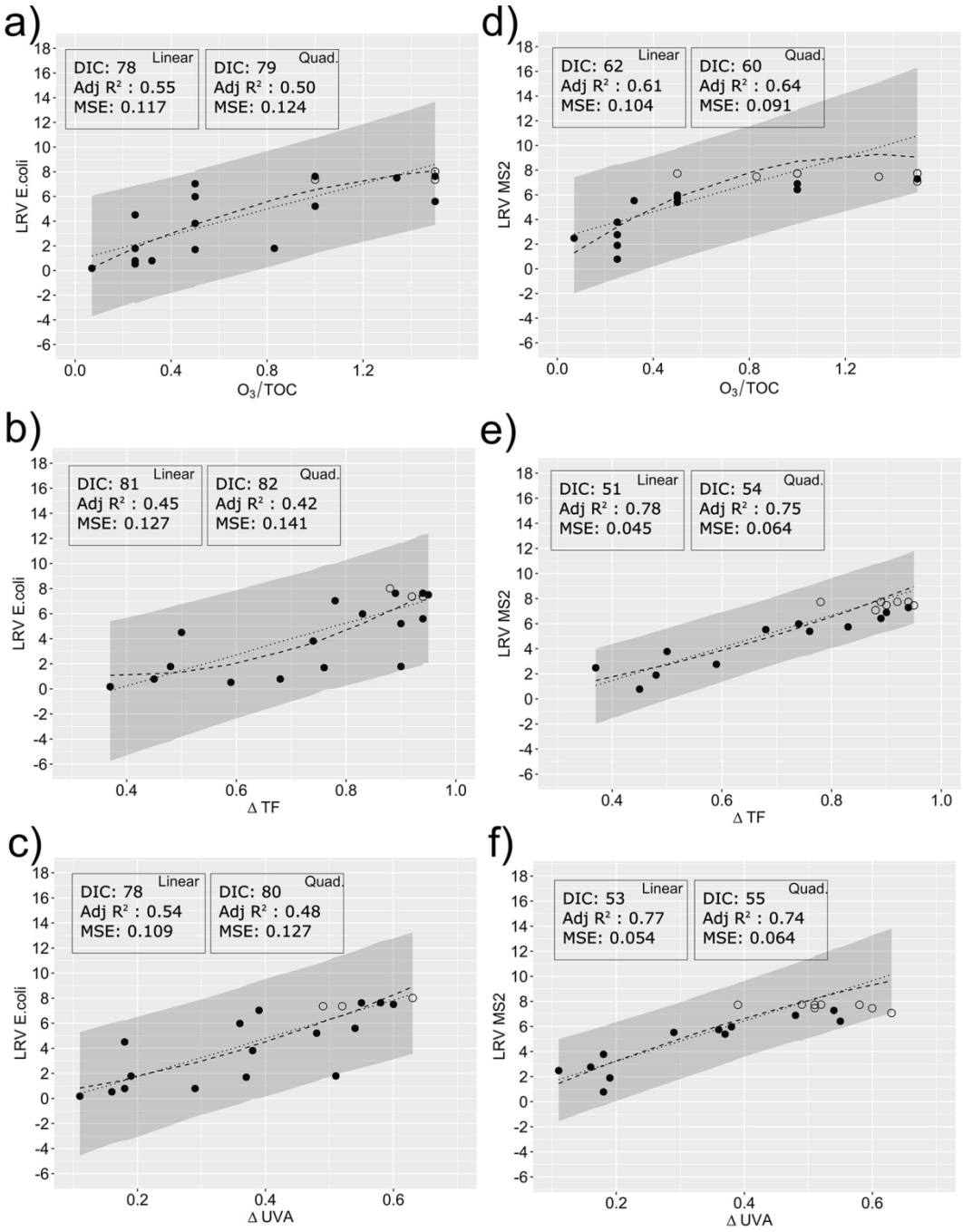

Figure 6-7: Scatterplots with linear fit between surrogate parameters O_3/TOC, ΔTF and ΔUVA and LRV for *E. coli* (a-c) and MS2 (d-f) for data collected by Gamage *et al.* (2013). Open circles indicate removal is higher than the recorded value due to censored data. The dotted line indicates the linear regression model calculated from the median of the posterior predictive distributions of the corresponding data set. The dashed line indicates the quadratic regression model calculated from the median of the posterior predictive distributions of the corresponding data set. Shaded region indicates the 95% posterior predictive invervals for the linear regression model.

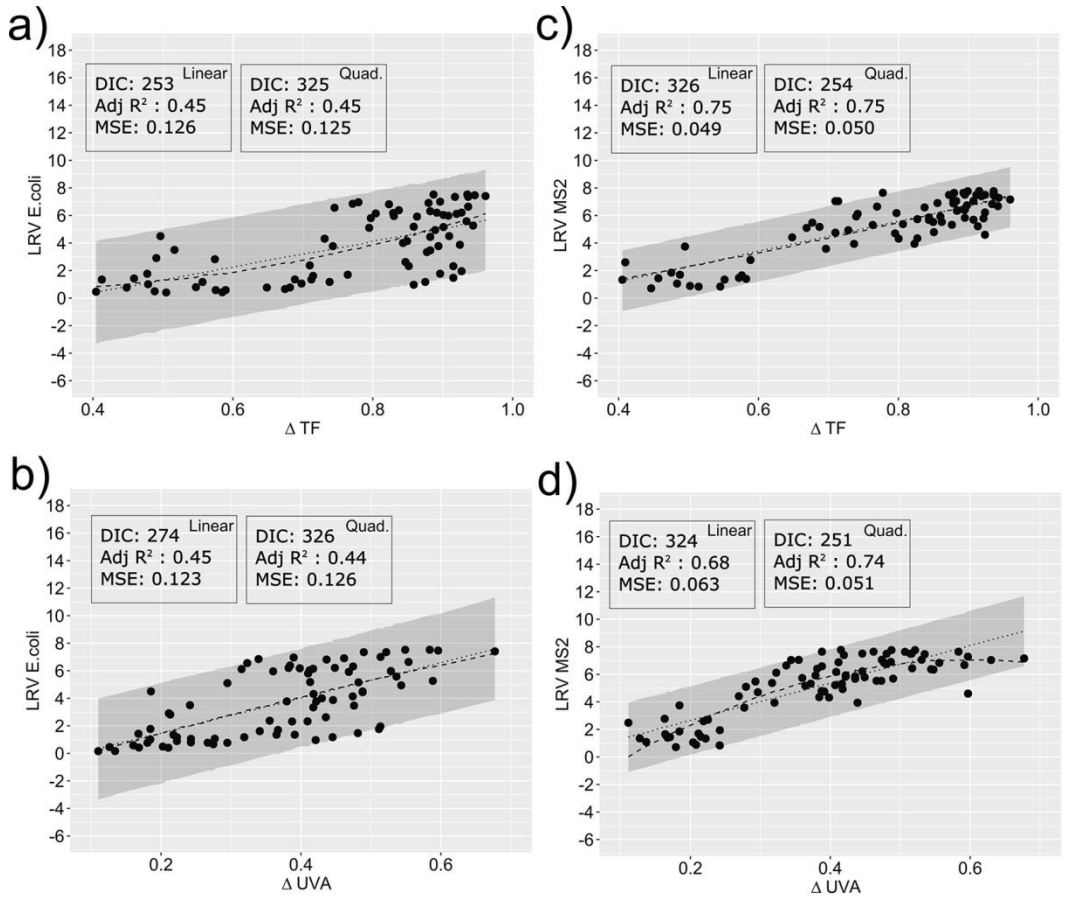

Figure 6-8: Scatterplots with linear fit between LRV for *E. coli* and MS2 with surrogate parameters ΔTF and ΔUVA for *E. coli* LRV (a-b), and MS2 LRV (c-d) for Gerrity *et al.* (2012). The dotted line indicates the linear regression model calculated from the median of the posterior predictive distributions of the corresponding data set. The dashed line indicates the quadratic regression model calculated from the median of the posterior predictive distributions of the corresponding data set. Shaded region indicates the 95% posterior predictive invervals for the linear regression model.

Slope results from the models fitted to the data reported by Gamage *et al.* (2013) showed that to achieve a unit change in LRV for *E. coli*, corresponded to a 8% (95% CI: (5, 15)) TF reduction, while a 7% (95% CI: (5, 11)) UVA reduction corresponded to the same microbial removal performance. For MS2, an 8% (95% CI: (6, 10)) reduction in TF corresponded to a unit change in LRV, while the same removal performance was associated with a 6% (95% CI: (5, 9)) UVA reduction (Figure 6-8). Similar slope results were obtained from the models fit to the data reported by Gerrity *et al.* (2012). The maximum achieved LRVs for the Gamage and Gerrity studies were much higher than for this present study (*E. coli* 5 $\log_{10}$, total coliforms: 3.3 $\log_{10}$, *C. perfringens*: 3.1 $\log_{10}$, and somatic coliphages: 3.5 $\log_{10}$) with values exceeding 7.5 $\log_{10}$ approximately for both seeded MS2 and *E. coli*.

6.4 Discussion

6.4.1 Comparison of model parameters

In this study the effectiveness of ozonation with that reported by two earlier studies were compared using the slope parameter of the regression models (Table 6-4). From the results obtained, it appears that, based on the posterior mean and 95% CI of the slope parameters, the sensitivity to disinfection of the indigenous microbial indicators to ozone treatment is lower than for seeded microorganisms (i.e. autochthonous species were more resistant). This result is consistent with previous studies, which have shown that environmental populations of microorganisms can be more resistant to disinfection than cultured populations (Smeets, 2010). Gerrity *et al.* (2012) also performed a full-scale investigation to validate the results of correlations found between LRVs and water quality parameters. However, they found that achieving high levels of inactivation was increasingly difficult when the influent concentrations were low (compared to a spiked experiment) possibly due to the effect of shielding.

Table 6-4: Posterior slope parameter for the fitted regression models[a,b]

M.I	**This study**			**(Gamage *et al.*, 2013)**			**(Gerrity *et al.*, 2012)**		
	O3:DOC	**ΔTF**	**ΔUVA**	**O3:TOC**	**ΔTF**	**ΔUVA**	**O3:TOC**	**ΔTF**	**ΔUVA**
EC[c]	3.6 (2.4,4.9)	4.4 (3.4,5.5)	5.1 (3.7,6.6)	5.2 (3.1,7.6)	13 (6.7,19)	15 (8.9,22)	-	9.3 (7.1,12)	13 (9.7,16)
TC[c]	5.1 (3.9,6.3)	5.5 (4.2,6.7)	6.7 (5.1,8.2)	-	-	-	-	-	-
CP[c]	1.2 (0.7,1.7)	1.4 (0.9,1.9)	1.7 (1.1,2.3)	-	-	-	-	-	-
MS2[c]	-	-	-	5.6 (3.1,7.6)	13 (9.9,17)	16 (12,22)	-	11 (9.6,12)	14 (12,16)
SC[c]	7.4 (4.7,11)	6.9 (4.4,10)	9.2 (5.9,14)	-	-	-	-	-	-

[a] A larger slope, indicates greater possible demonstration of organism LRV per unit change in monitored parameter.

[b] Values in parentheses are the 95% CIs.

[c] EC: *E. coli*, TC: total coliforms, CP: *C. perfringens spores*, MS2: MS2 coliphages, SC: somatic coliphages.

The posterior distributions of the slopes for the three studies using ΔUVA as a predictor for *E. coli* removal are shown in Figure 6-9. The slope coefficients for both this present work and the previous studies (seeded 1 and 2 in Figure 6-9) showed similar values with greater spread in the case of the data from Gerrity *et al.* (2012) compared to the data from Gamage *et al.* (2013). These results are expected as a larger dataset reduces the standard error of the coefficient. In the current study, it could be observed that the probability distribution for the slope coefficient was narrower with a mean of 5 units. Two reasons the results from the two previous studies were similar are likely to be, common authors and hence methodology, and because one dataset was contained within the other (i.e. Gerrity *et al.* (2012) incorporated data from Gamage *et al.* (2013)). Also, the higher uncertainty observed in these previous studies for *E. coli* may be a consequence of using heterogeneous water samples from different sites. The current study showed that it is possible to reduce the uncertainty of the outcomes by generating data from a single site. To address this source of variance, where data are obtained from different treatment plants, a hierarchical model could be constructed to account for random effects associated with different sites and estimate the uncertainty of population parameters.

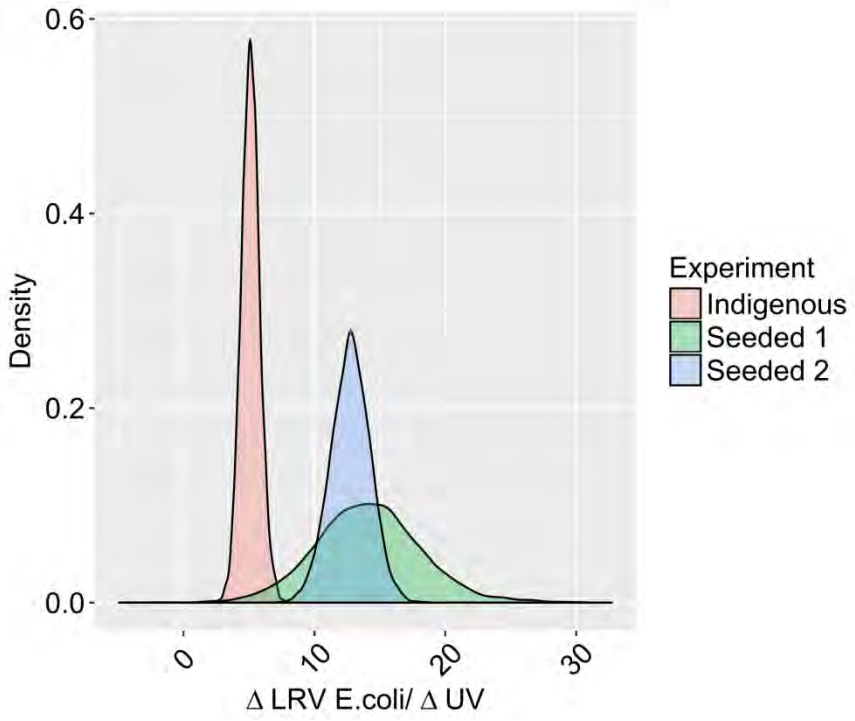

Figure 6-9: Posterior density plots comparing the slope coefficient from the fitted linear models for indigenous microorganisms and seeded from previous studies. Seeded 1 represents the data obtained by Gamage *et al.* (2013) and Seeded 2 represents the data obtained by Gerrity *et al.* (2012).

6.4.2 Evaluation of parameters of ozone effectiveness

Initial water quality tests showed low concentrations of indigenous FRNA coliphage in the secondary treated wastewater (≤10 pfu/10 ml) which made this measure unsuitable for the present study. In its stead, somatic coliphages were used which are usually more abundant in wastewater samples and present similar sensitivity to ozone compared to FRNA coliphages (Tyrrell *et al.*, 1995). Although *C. perfringens* spores are more resistant than vegetative bacteria and coliphages, the results indicated that they possessed appreciable sensitivity to ozone and that O_3:DOC ratio, ΔTF, and ΔUVA can be used to predict their removal. When comparing results from this study with previous studies using seeded microbial indicators (Gerrity *et al.*, 2012; Gamage *et al.*, 2013), it appeared that indigenous microorganisms showed lower sensitivity to ozone which results in a more conservative estimate of removal performance. This finding needs further investigation as it has a bearing on ozone disinfection validation which increasingly use seeded microorganisms.

Assignment of removal credits to treatment processes is subject to the reliable monitoring of operational parameters which should be able to indicate when the system is not performing correctly and corrective actions are required (NRMMC *et al.*, 2008). In general, the data presented a better fit to the models constructed with ΔPeak A (humic acid-like), ΔTF, and ΔUVA than when fitted to ozone demand. For the same target LRV, a higher reduction in ΔTF is necessary when compared to ΔUVA, which means that ΔUVA can potentially demonstrate higher LRVs. Although, ΔPeak A, ΔTF and ΔUVA were all useful for monitoring ozonation performance, fluorescence measurements are potentially less suitable when compared to ΔUVA when considered on the basis of maximum possible demonstrated LRV. In addition, UV transmittance sensors which typically measure at 254 nm are already commercially available for monitoring UV disinfection systems, which means that ΔUVA is a more readily available technique to be incorporated online in real time monitoring of water recycling systems. ΔUVA is already widely applied for monitoring ozone performance (usually for colour removal or DOC oxidation prior to BAC). However, a quantitative relationship with disinfection performance is not typically available or applied in full-scale water treatment applications. The present work suggests such surrogate measures though are quite practical.

6.4.3 Advantage of using Bayesian analysis to evaluate predictors

The data analysed in this study could have been interpreted using non-Bayesian techniques such as conventional linear and non-linear regression analyses. However, Bayesian analysis provided two unique advantages including posterior distributions and incorporation of censored values.

Bayesian posterior distributions can be interpreted as true probability statements about unknown parameters (Hamada *et al.*, 2008). Therefore, information stating the range of values and their probability of occurrence for a particular parameter truly represent the uncertainty in that parameter. Instead, from a classical approach point of view, confidence intervals must be constructed from repeated sampling of the data. In terms of propagating the uncertainty in the parameters of a model to estimate posterior predictive distributions, Bayesian analysis can effectively handle such tasks. This feature is important during monitoring as there will always be uncertainty when inferring a performance from operational or water quality parameters. With Bayesian analysis, the full probability distribution reflecting the uncertainty of an outcome can be obtained.

Censored data are commonly wrongly omitted or replaced by half of the limit of detection (Helsel, 2005). Censored values also have an influence on the fitting of models and affect the uncertainty of the estimations. Through the use of Bayesian analysis, this study was able to incorporate the influence of censored values into the estimation of model parameters. Instead of making assumptions about the censored values, Bayesian analysis uses all the information available for them which is defined by the assumed probability distribution (i.e. normal distribution in this case).

6.5 Conclusions

Conventional use of the monitoring parameter CT during wastewater ozonation becomes problematic when ozone rapidly reacts with organic matter and its dissolved residual concentration cannot be accurately or continuously measured. Alternative surrogate parameters have been proposed including ozone demand (O_3:DOC (TOC) ratio), ΔUVA, and ΔTF. Previous studies using seeded microorganisms have found positive associations between the removal of microorganisms and these alternative parameters. However, high uncertainty and potentially optimistic results may have been obtained from these experiments. This study focused on the monitoring of ozonation of secondary treated wastewater using the previously proposed surrogates and indigenous microorganisms. Results of this work were compared to previous findings in terms of uncertainty and disinfection performance. Key outcomes from this study included:

- Ozone demand, ΔUVA, and ΔTF were all correlated with removal of indigenous microorganisms

- In general, the data presented better fit to the models constructed with ΔTF and ΔUVA than those constructed with ozone demand with ΔUVA the most cost effective and simple, and hence promising on line measure, of LRV.
- *Clostridium perfringens* spores presented higher resistance to ozone treatment compared to coliform bacteria and somatic coliphages.

Experiments performed on a single wastewater source in this study possibly resulted in a lower uncertainty than the models developed over multiple sources from previous studies. Although global surrogate correlations are possible, it is recommended that the generated models should be site-specific to decrease excessive uncertainty. Further investigation into the correlation of non-microbial surrogates with pathogens is also needed to assess whether microbial indicators are appropriately conservative indicators of the process performance is also advisable. Bayesian methods are recommended because they facilitate the uncertainty measurement in the model parameters, incorporate censored values and generate posterior predictive distributions.

Chapter 7: Probabilistic assessment of chlorination performance targets for secondary treated wastewater using Bayesian analysis

This chapter has been published in the following journal article:

Carvajal, G., Roser, D.J., Sisson, S.A., Keegan, A. and Khan, S.J. (2017). Bayesian belief network modelling of chlorine disinfection for human pathogenic viruses in municipal wastewater. Water research 109, 144-154. DOI: 10.1016/j.watres.2016.11.008

7.1. Introduction

Disinfection dose requirements for water and wastewater treatment are conventionally expressed as the product of disinfectant concentration and contact time (CT), required to achieve a predetermined reduction in microbial numbers. Achievement of target CT values is dependent upon meeting various factors for each pathogen type, including pH, temperature, disinfectant concentration, ionic strength, and suspended particles (Jensen *et al.*, 1980; LeChevallier & Au, 2004). Consequently, assigning log reduction value (LRV) credits depends on optimising such variables as well as ensuring the primary CT product.

Previous work has sought to establish CT values necessary to achieve LRVs for various pathogens in biologically treated (activated sludge) municipal wastewater (Keegan *et al.*, 2012).The products of that research included a number of linear models and tables which related target CTs to various pH and turbidity combinations (VDoH, 2013). This approach to deriving and representing CT-LRV relationships is in line with international best practice for defining the disinfection CT requirements for drinking water (USEPA, 2003). However, as the number of disinfection controlling factors increases, interpolating values between experimental data points and communicating the information in tabular form or as linear models becomes increasingly problematic. Further, the increasing popularity of Quantitative Microbial Risk Assessment (US-EPA & USDA/FSIS, 2012) means that in future, LRV point estimates may not be sufficient, and measures of model uncertainty and variability will be needed.

Bayesian networks (BNs) offer an alternative approach for relating chlorination LRVs to CT and wastewater quality parameters and incorporate parameter uncertainty and variability. They also offer a convenient means for performing scenario exploration and inference, and hence prediction of disinfection performance under diverse conditions. Continuous BNs are models which involve the use of continuous variables without the need for discretisation.

To facilitate interpolation and CT estimation and prediction, the use of Bayesian multilayer perceptron (BMLP) models were investigated to derive continuous relationships between virus LRV, pH, turbidity and CT, and perform interpolation considering uncertainty in the model parameters. The Bayesian integration within the BMLP model transforms this model into a BN, introducing features such as stochastic representation of the parameters and predictions, as well as computation of queries on target variables given a set of observations. A multilayer perceptron (MLP) model is a type of neural network composed of layers of neurons (elements that generate a transformation of the inputs) with an input layer, at least one hidden layer, and an output layer (Priddy & Keller, 2005). In MLP models the inputs of the neurons in one layer come from the

outputs of neurons in a previous layer. Neurons in one layer are connected to the previous layer through weighted connections. These models can solve non-linear problems and perform prediction with high accuracy in multivariate settings.

This study assessed the application of a continuous BN for estimating chlorination and chloramination LRVs for human virus removal during wastewater treatment, while accounting for the influence of pH, turbidity and CT variance. The broader aim of the study was to investigate the utility of BNs for quantifying the effectiveness of chlorination of treated wastewater in a multivariate context and to present a simple and practical tool for interpolation of CT values and incorporation of uncertainty and variability.

7.2. Materials and methods

7.2.1. Data extraction and model construction

Chlorine disinfection data were obtained from a previously published project report of bench-scale batch experiments using secondary treated wastewater from the Bolivar Wastewater Treatment Plant in Adelaide, South Australia (Keegan *et al.*, 2012). This wastewater was seeded with two viruses, Coxsackievirus B5 for estimating free chlorine LRVs, and Adenovirus 2 for estimating monochloramine LRVs. These viruses were selected because they are known to exhibit high resistance to chlorine and monochloramine inactivation respectively (Liu *et al.*, 1971; Payment *et al.*, 1985). Two parameters, pH and turbidity, were varied to determine CT values for virus inactivation under a range of conditions.

Inactivation experiments were conducted at pH 7.0, 8.0 and 9.0, and at three turbidity values (2, 5, 20 NTU), at a constant low (conservative) temperature of 10°C. Both viruses were seeded at concentrations of *ca* 10^5/mL to allow measurement of up to at least 4 logs inactivation. The datasets produced by Keegan *et al.* (2012) consisted of 226 records for Adenovirus and 154 records for Coxsackievirus. CT values were calculated from the integrals (areas under the curves) of residual free and combined chlorine concentrations vs. sample contact time. Unlike the original study, in the present analysis the disinfection conditions and corresponding LRV data were not obtained from fitted linear models, but by using the raw replicated LRV measurements obtained in that study. The aim for Keegan *et al.* (2012) was to construct CT tables for specific whole-number LRV values (1, 2, 3 and 4) in line with conventional past practice (USEPA, 2003). This work presents the construction of a BN, which can produce a target CT value for any combination of input variables.

Prior to analysis, the raw data (Appendix 4) were extracted and arranged in a table (.CSV file) with columns representing variables and rows presenting the experimental cases. This arrangement was used to facilitate importation to *R* programming (R-project, 2014) environment and BMLP model construction. The process of model construction and parameter definition was conducted in *R* using *JAGS* through the freely available *R2jags* package (Plummer, 2013).

The two overall structurally identical but parametrically different models, one for each virus, represented the conventional procedure followed by the US Environment Protection Agency (USEPA, 2003). The models were constructed considering three continuous variables, target LRV, pH and turbidity as predictors of target CT. The BN model is represented in Figure 7-1, which shows the choice of distributions for each identified node using *JAGS* nomenclature. Explanation of each variable in the model is provided in Table 7-1.

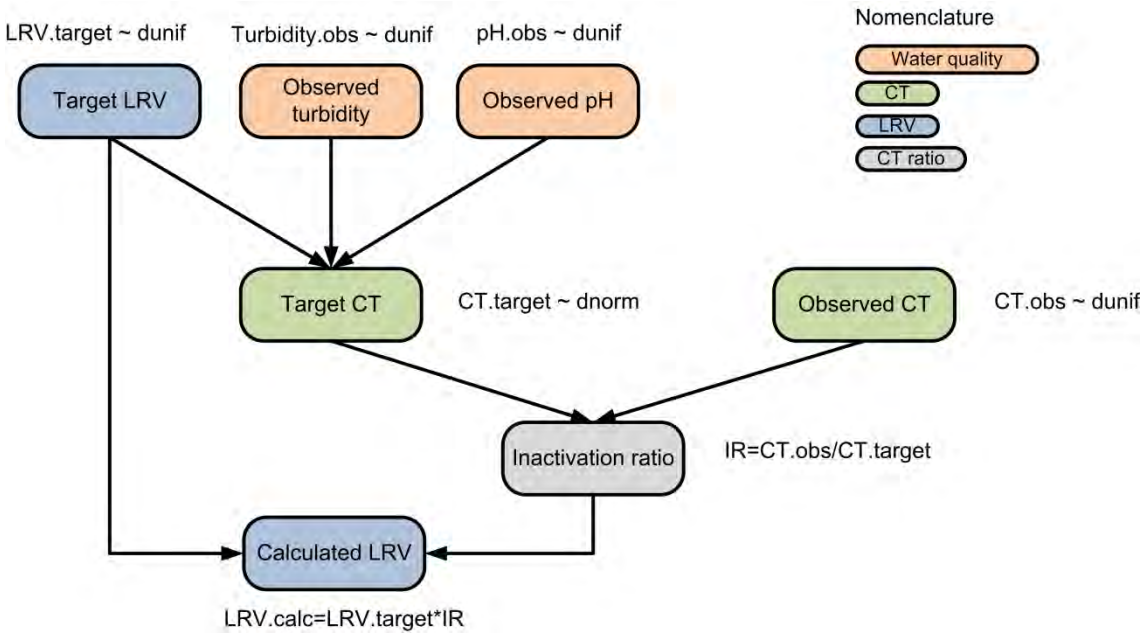

Figure 7-1: Bayesian network for chlorination model. For model nodes and consistency with *JAGS* notation, *dunif* indicates a uniform distribution, and *dnorm* indicates a normal distribution.

Table 7-1: Variables/Nodes considered in the study and their description

Variables	**Description**	**Adenovirus** **Monochloramine** **Range**	**Coxsackievirus** **Free chlorine** **Range**
Observed pH	Measured pH during disinfection operation	7 – 9	7 - 9
Observed turbidity	Measured turbidity during disinfection operation	2 – 20	2 - 20
Target LRV	Desired target LRV	0 – 4.7	0 – 5.3
Observed CT	Measured CT during disinfection operation	≥0	≥0
Target CT	Required CT to achieve a given target LRV	71 - 11488	2.3 – 53
Calculated LRV	Computed from the multiplication of inactivation ratio and Target LRV	≥0	≥0
Inactivation ratio	Ratio of observed CT and target CT (predicted)	≥0	≥0

Other water quality parameters may also affect chlorination LRV. However, such variables were controlled throughout these experiments by the use of the same biologically treated wastewater matrix in all experiments. The model thus assumes that target LRV, pH and turbidity are conditionally dependent given target CT. This means that without knowing the value of the target CT, a modification to any of these three nodes will not produce any change to the other two.

A BMLP model was employed to capture the relationships between the variables. BMLP model use enabled consideration of pH and turbidity as continuous variables and therefore facilitated interpolation of values intermediate with the selected experimental settings. In this model, the error estimates in the target CT were assumed normally distributed. Normal distribution was also preferred due to its simplicity. This assumption has been previously used to derive confidence intervals in the predicted outcomes of multilayer perceptron (Chryssolouris *et al.*, 1996; He & Li, 2011). That is

$$CT \sim N(\mu; \sigma^2)$$ Equation 7-1

where μ is the predicted dependant value (target CT) as a function of the target LRV, turbidity and pH, and σ is the standard deviation of the model errors. Unlike the conventional neural network approach, the BMLP model considers uncertainty in the model parameters which introduces an additional source of uncertainty for the CT value estimation. The resulting model is presented in Section 7.2.3. Inactivation ratio (IR) and actual LRV (LRV.act) are variables which were computed from the outcomes of stochastic variables using Equation 7-2 and Equation 7-3 respectively (USEPA, 2003).

$$\mathrm{IR} = \frac{\mathrm{CT.obs}}{\mathrm{CT.target}}$$ Equation 7-2

$$\mathrm{LRV.calc} = \mathrm{LRV.target} \cdot \mathrm{IR}$$ Equation 7-3

where *CT.obs* is the CT measured from the chlorine disinfection process, *CT.target* represents the target CT from the model, *LRV.target* is the target LRV from the model and *LRV.calc* is the calculated LRV.

The model (Figure 7-1) can be used to query information for LRV, CT, pH or turbidity given one or more variables observed. Depending on the case, these variables can be seen as target or observed variables as explained in Table 7-1. The target CT is used along with the CT of the process (observed CT) to estimate a LRV using Equation 7-2 and Equation 7-3. Explanation of model queries and observations is provided in Appendix 4.

7.2.2. Use of *JAGS* for BN construction and inference

Inference in continuous BNs was performed through Markov Chain Monte Carlo (MCMC) which is used to obtain joint and marginal posterior distributions (see Chapter 3, Materials and methods). In this study, *JAGS* was employed because of its simplicity in model definition and capability of dealing with different types of data formats (Plummer, 2013). The model definition and model inference steps in *JAGS*, used to obtain posterior distributions were as follows, while the code is presented in detail in Appendix 4. The key steps were:

1. Model definition: Arcs between nodes and distributions are defined in Figure 7-1. Uninformative priors are defined.

2. Incorporation of model and data into *JAGS*: see Appendix 4.

3. Model updating with 2E5 samples and generation of 2E6 samples from the model.

4. Convergence checking to verify that the output was valid: Visual inspection of trace plots and Geweke convergence test.

7.2.3. Bayesian Multilayer Perceptron model structure

The BMLP model was constructed using R2jags in R and consisted of an "n" number of neurons in a single hidden layer as shown in Figure 7-2. A single hidden layer was used to avoid overfitting and because a single hidden layer is capable of approximating any nonlinear function (Hornik *et al.*, 1989). Each neuron in this layer was defined by the softplus function (Nair & Hinton, 2010) shown in Equation 7-4. The output layer was defined as a linear combination of a specific number of hidden neurons as presented in Equation 7-5. All input variables were normalised in the range [0,1] using Equation 7-6 to minimise the bias in the model for one variable over another (Priddy & Keller, 2005). Posterior distributions for the parameters (i.e. weights and biases) in the BMLP model were found through Bayesian inference using MCMC Gibbs sampling (Geman & Geman, 1984) and assuming non-informative priors (see Appendix 4). The model also included a hierarchical structure for the weights and biases to regularise the estimates as defined by Lampinen and Vehtari (2001). Due to lack of identifiability in the parameters, order constraints for the parameters were included (Müller & Insua, 1998). A preliminary comparison between models including two, three, four and five neurons indicated that three neurons provided better fit to the data for both the Adenovirus and Coxsackievirus models according to the deviation information criterion score (Lunn *et al.*, 2012). Initial values were defined as recommended by Vehtari *et al.* (2000) and ten random seeds were tested through the DIC score.

$$neuron_j = \ln(1 + e^{\sum_i w_i \cdot x_i + w_{0j}}) \quad for\ i = 1,2,3\ and\ j = 1, \ldots,\mathrm{n} \qquad \text{Equation 7-4}$$

$$neuron_0 = \sum_{j=1}^{4} (w_0 + w_j \cdot neuron_j) \qquad \text{Equation 7-5}$$

$$normalized\ variable = \frac{unnormalized\ variable - min}{max - \min} \qquad \text{Equation 7-6}$$

where w_i and w_j represents weights, w_0 and w_{0j} are bias terms, and x_i represents a normalised variable

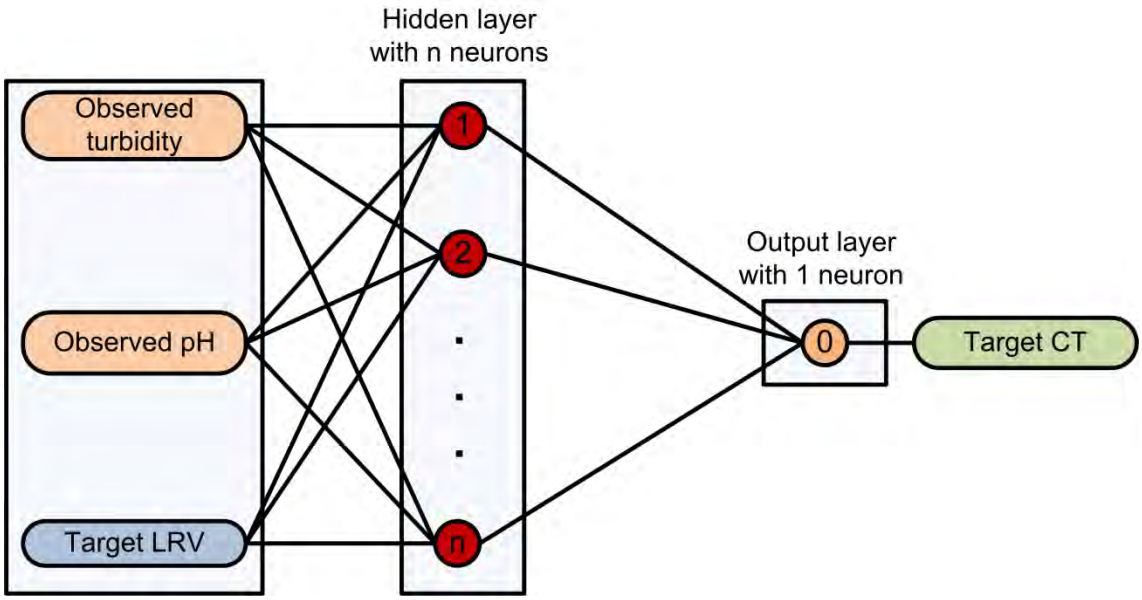

Figure 7-2: Multilayer perceptron diagram with n neurons in hidden layer

An advantage of neural networks is their flexibility to model variable relationships without the need for an underlying mechanism or complete understanding of the process. However, this flexibility also entails the possibility of more than one model providing suitable results for the analysed problem. In this study, the selection process permitted a unique model for each microorganism. An alternative solution is the averaging of several models' outputs. The approach used in this study did not follow this option to keep the modelling process simple.

7.2.4. Sensitivity analysis

The Adenovirus and Coxsackievirus models were assessed for their sensitivity to turbidity, pH and residual concentration times time. The methodology employed for this analysis consisted of selecting specific values for LRV and varying the turbidity while keeping pH constant and varying pH while keeping turbidity constant. A tridimensional plot was also constructed by changing pH and turbidity at the same time for a defined LRV value of 3 logs.

7.3. Results

7.3.1. Bayesian MLP model development and performance

The Bayesian multilayer perceptron model parameters (i.e. the weights and biases) were obtained for Adenovirus (monochloramine) and Coxsackievirus (free chlorine). High correlation coefficient (r>0.96) was found with the models of both viruses between the actual LRV and the posterior predictive LRV mean (Figure 7-3). The correlation and visual check of the predictions against actual LRVs indicated that the model correctly fitted the data and that the fitted values (Figure 7-3) closely match the ideal outcome (line= 1 to 1 relationship). The proportion of standardised residuals lying outside the (-2,2) range was 6.1% for the Adenovirus model and 4% for the Coxsackievirus model, while the proportion of the residuals lying outside the (-3,3) was 0.5% for the Adenovirus model and 0.7% for the Coxsackievirus model. These results supported the normality assumption used in the models (Ntzoufras, 2011).

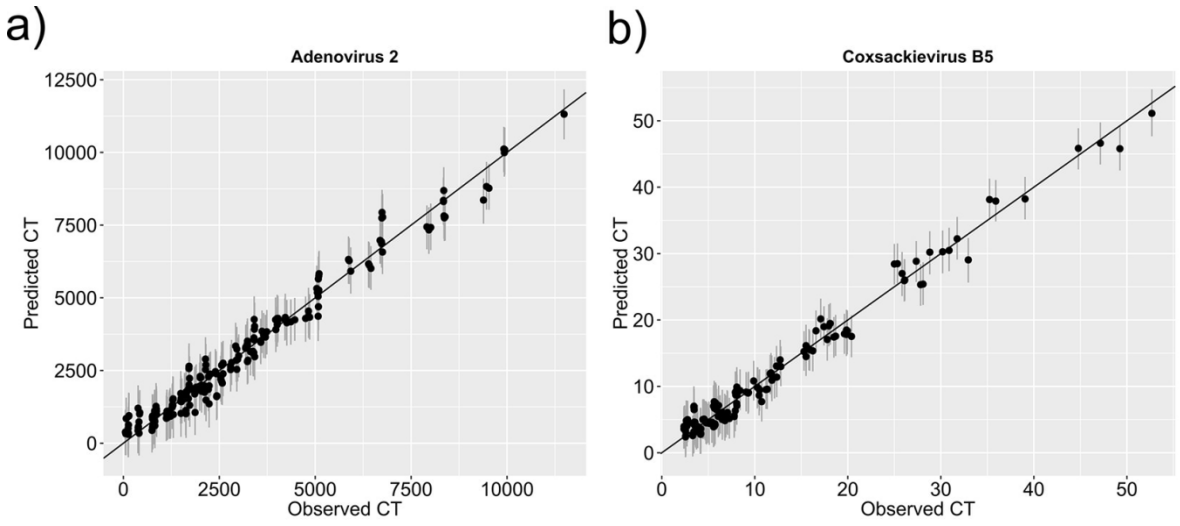

Figure 7-3: Scatter plot of observed LRV versus predicted LRV using BMLP model for a) Adenovirus and b) Coxsackievirus. Vertical bars indicate the 2.5th and 97.5th percentiles from the predicted values.

7.3.2. Calculation of LRV

The use of CT tables during monitoring is based on the comparison of observed CT values versus the target CT. The ratio of these two estimates, the inactivation ratio IR, indicates whether the process is achieving the target LRV (IR≥1) or not (IR<1). Further, IR itself can be probabilistic allowing the confidence that treatment is satisfactory to be estimated. After computing the IR,

LRV estimation can be performed by multiplying the IR by the target LRV. To model this scenario in the BN model, the variables “target CT”, “inactivation ratio” and “calculated LRV” were incorporated. To illustrate, the Coxsackievirus model was used to derive information about the calculated LRV from a target LRV of 4 log for a turbidity of 5 NTU, with CT and pH defined by distributions from real operational chlorination clear-well data (Walker *et al.*, 2016). The distribution for pH was normal (mean=6.99, standard deviation=0.139, Figure 7-4 a), while the distribution for CT was lognormal (location=4.51, scale=0.148, Figure 7-4 b). Under this scenario, the inactivation ratio was always higher than 1, with a mean of 16.6 (Figure 7-4 d). The associated LRV mean was 66 (5^{th} percentile=34), which reflected an operation in which the chlorine addition far exceeded the requirements for this target microorganism. When specific conditions need to be evaluated, the water quality parameters pH and turbidity can also be used as deterministic values. The model could be used to determine the CT in the process required to achieve a specific exceedance probability for the desired LRV performance. The CT could be obtained by selecting the desired percentile value from the target CT distribution. The observed CT has then to be aimed for that particular value. For example, if the 5^{th} percentile of a target CT is 20 $mg.L^{-1}.min$ for a LRV of 3 logs, then if the observed CT is set to this same value of 20 $mg.L^{-1}.min$, the 5^{th} percentile obtained from the calculated LRV will correspond to the desired 3 logs target.

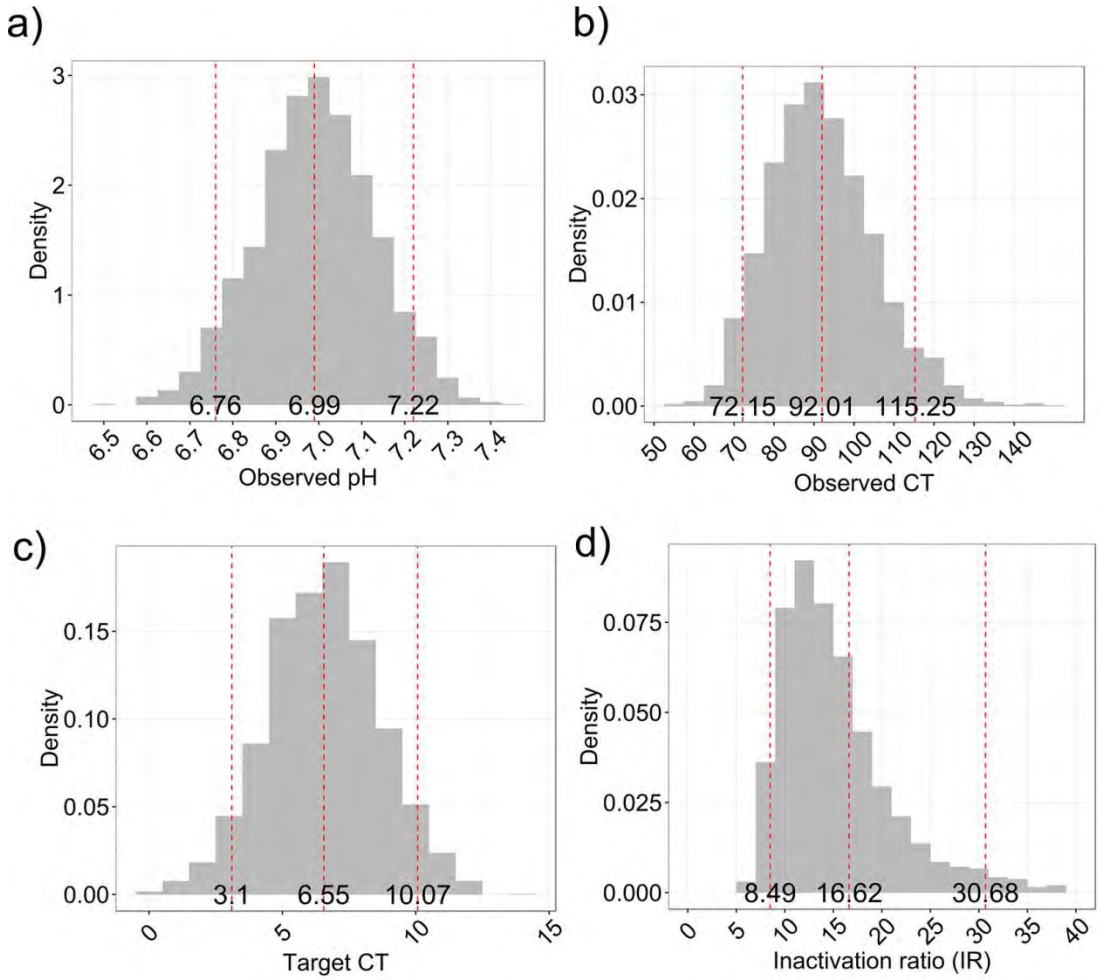

Figure 7-4: LRV estimation using, a) observed pH and, b) CT distributions during a real chlorination process (Walker *et al.*, 2016). The target CT c) is obtained from the BMLP model and the inactivation ratio d) is the ratio between the observed CT and target CT.

7.3.3. Inferring pH and turbidity given target operating conditions.

To investigate the capabilities of the model for inferring water quality conditions (i.e. pH and turbidity) required to achieve a target LRV given a target CT, two queries were assessed. The first query consisted of inferring the turbidity required to achieve a LRV of 4 with a CT of 20 $mg.L^{-1}.min$ and pH 9 for the Coxsackievirus model (Figure 7-5 a). The prior distribution for turbidity was assumed to be uniform between 2 and 20 NTU reflecting the choice of uniform distribution or full uncertainty (if the turbidity probability distribution for the water is not available). The results for the posterior distribution are presented in Figure 7-5 (a). The simulation results indicated that to achieve these targets, the turbidity should be low with a 95% probability of being below 2.8 NTU. Additionally, using the training dataset, it was observed that turbidity less than or equal to 5 NTU was associated with these operating conditions of pH, CT and LRV. The second

query involved obtaining the posterior distribution for pH given a CT of 20 mg.L^{-1}.min, LRV of 4 and turbidity of 20 NTU for a uniform prior pH distribution between 7 and 9. As observed in Figure 7-5 (b), this query indicates that pH is required to be close to 7.0 to achieve this LRV for this high turbidity value.

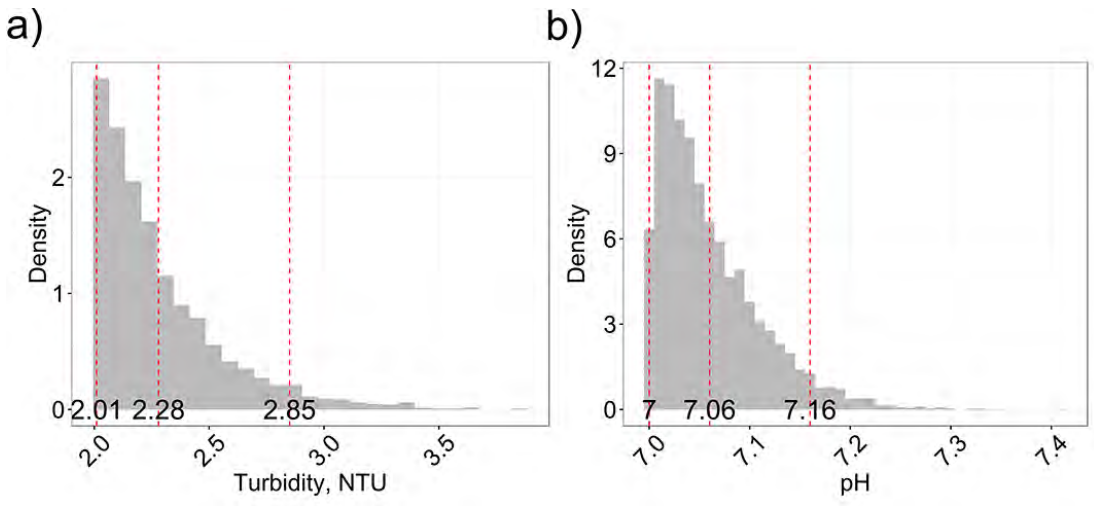

Figure 7-5: Posterior distribution for turbidity and pH after observing CT=20 mg.L^{-1}.min and LRV 4 with pH=9 (a) and turbidity 20 NTU (b) for free chlorine.

7.3.4. Sensitivity analysis

A sensitivity analysis was performed to study the impacts of pH and turbidity variation individually on target CT. To perform the sensitivity analysis, the posterior median target CT values were estimated while varying turbidity and holding pH fixed at 7.0, and while varying pH and fixing turbidity at 2 NTU. Three levels of target LRV were considered for both monochloramine and free chlorine (LRV=2, 3, 4). The results are presented in Figure 7-6. As is visually apparent, pH has a greater effect on mean CT than turbidity, given the conditioned values of each variable, for both Adenovirus and Coxsackievirus. As can be observed, the relationships between pH and CT, and turbidity and CT are clearly nonlinear, which is an important consideration during interpolation when using kinetic constants as in Chick-Watson's model (Haas & Heller, 1989) or CT tables. The impact of pH on Adenovirus target CT was considerably greater than the impact of turbidity. For Adenovirus, at higher values of target LRV the effect of increasing pH intensified, whereas the effect of turbidity remained stable. For Coxsackievirus, the impact of turbidity was minor for LRV of 2, whereas for LRVs of 3 and 4 there was a steeper increase in the target CT values. For this same virus, the impact of pH showed a large increase on the target CT values for the three LRVs.

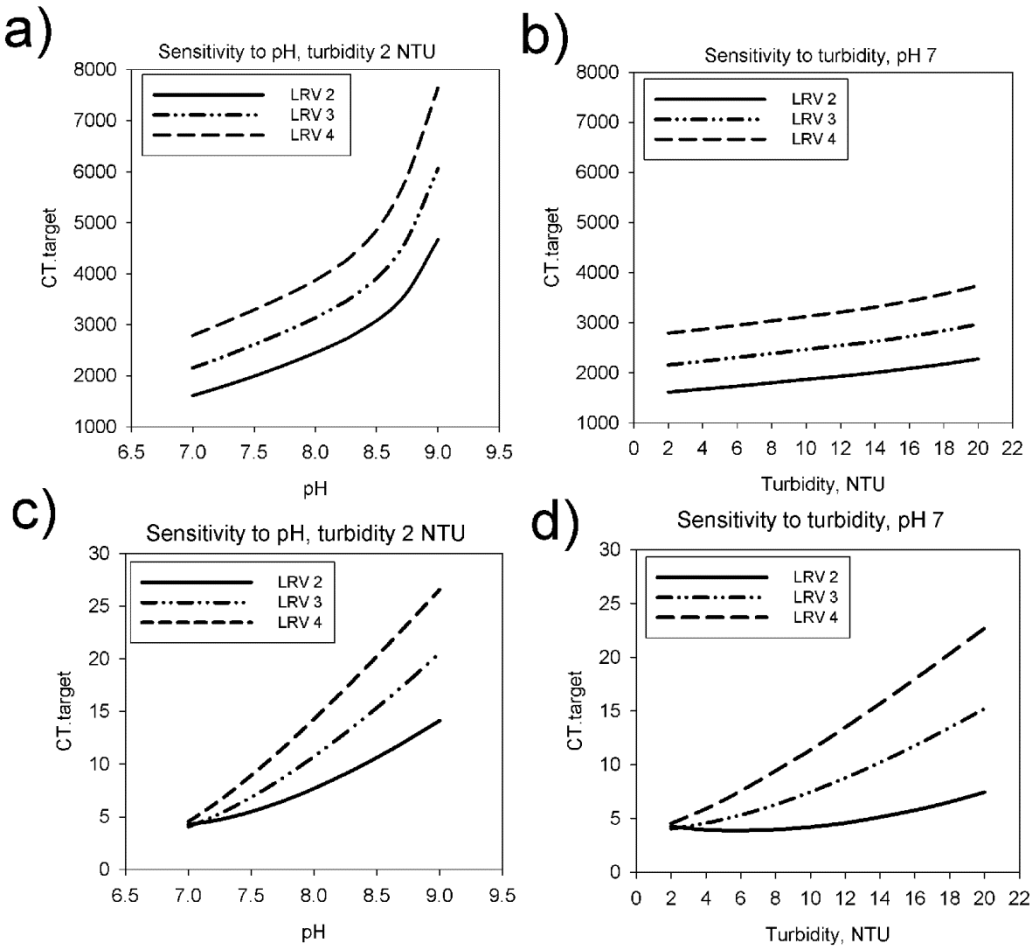

Figure 7-6: Sensitivity of target CT to turbidity and pH for LRV 2, 3 and 4. Sensitivity to, a) pH for Adenovirus and chloramination, b) turbidity for Adenovirus, c) pH for Coxsackievirus and chlorination and d) turbidity for Coxsackievirus.

The combined effect of turbidity and pH on CT at LRV=3 for both the Coxsackievirus and Adenovirus models can be seen using a 3-D plot (Figure 7-7). For Adenovirus (Figure 7-7 a) the effect of pH on CT was clearly influenced by the turbidity value. At turbidity 20 NTU, there was a marked increase in CT from pH 7 to 9. By contrast, as the turbidity was reduced, the steepness of the surface declined. For Coxsackievirus (Figure 7-7 b) the pH effect on CT had a lesser impact at lower turbidity values. Three dimensional plots were also constructed by interpolating the target CT values at specific conditions of turbidity and pH defined by the experiments using simple interpolation techniques (i.e. linear or quadratic models). These results were used to compare the result obtained by the BMLP model with the actual data. The model usefully provided reasonable interpolations for the range of turbidity values at which no data were available (between 5 NTU and 20 NTU). In the present instance it was possible to illustrate the interrelationship graphically.

But where more parameters were incorporated a model would be necessary to efficiently identify optimum water quality conditions.

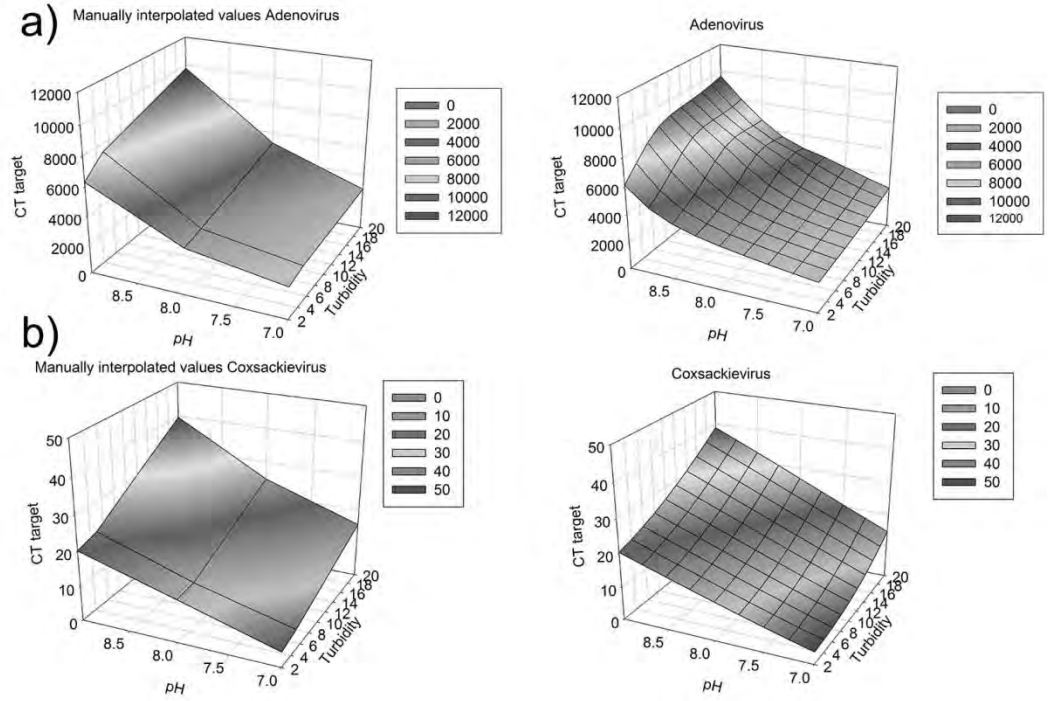

Figure 7-7: Combined effect of pH and turbidity on a) Adenovirus target CT (for LRV 3) and b) Coxsackievirus target CT (for LRV 3).

7.3.5. Impact of pH and turbidity on target CT and comparison to guidelines

Similar to the approach used for CT tables, in the proposed model the target CT value can be obtained from water quality conditions (e.g. pH and temperature) and the target LRV. The influence of observed turbidity and pH on target CT given a fixed target LRV was analysed through two scenarios assuming best and worst pH and turbidity conditions for both disinfection models. The scenarios assessed were pH=7 and turbidity=2 NTU for best conditions, and pH=9 and turbidity=20 NTU for worst conditions. Both scenarios were evaluated for a target LRV of 4. The difference between the optimum and worst operational conditions was approximately 10-fold for free chlorine and 5-fold for monochloramine (Figure 7-8).

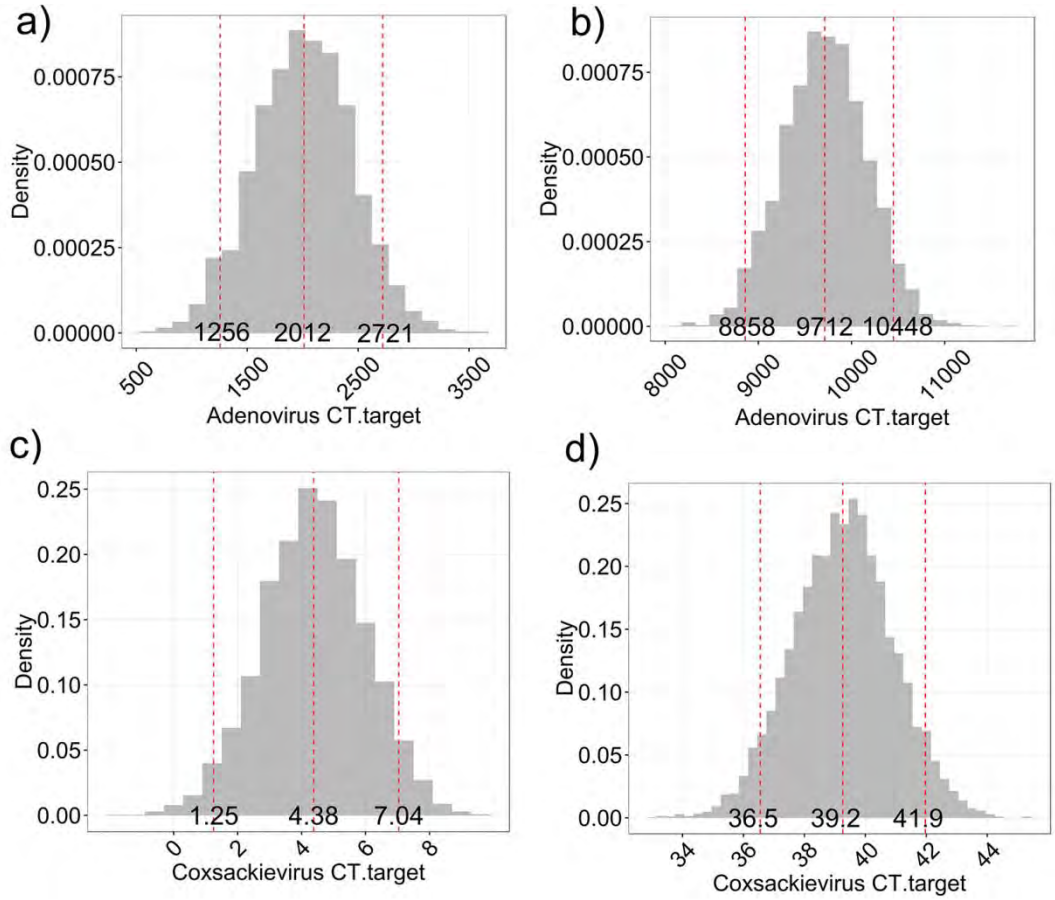

Figure 7-8: CT posterior probability distribution for optimum and worst conditions of pH and turbidity with a LRV of 3. Panels indicate Adenovirus under a) optimum (pH=7, turbidity=2), and b) worst (pH=9, turbidity=20) conditions, and Coxsackievirus under c) optimum (pH=7, turbidity=2), and d) worst (pH=9, turbidity=20) conditions.

The target CT results for the defined levels of turbidity, pH and LRV adopted in the Victorian state recycled water guidelines (Keegan *et al.*, 2012; VDoH, 2013) were compared to the results of the models as shown Table 7-2. In only five cases (one for Adenovirus and four for Coxsackievirus) was the recommended CT not inside the estimated 2.5^{th}-97.5^{th} percentile range in this study. Minor discrepancies between these estimates were expected as the input for the model estimation was the raw triplicate data instead of their mean values as in the case of CT table estimates. Determination coefficients between the table values and model median estimates were 0.98 for Adenovirus and 0.94 for Coxsackievirus indicating good agreement between the CTs obtained from the models and tables. Model estimates were considered more informative than CT tables due to their capability for interpolation and estimating uncertainty in the model outputs.

Table 7-2: Comparison between CT values from Victorian guidelines (VDoH, 2013) **and this study for monochloramine and free chlorine disinfection.**

			Monochloramine		Free chlorine	
NTU	pH	LRV	CT mg.min/L Table	CT mg.min/L This study Mean ($P_{2.5th}$,$P_{97.5th}$)	CT mg.min/L Table	CT mg.min/L This study Mean ($P_{2.5th}$,$P_{97.5th}$)
2	7	1	977	1091 (468,1742)	3	3.5 (0.3,6.5)
		2	1681	1609 (985,2262)	4	4.2 (1.1,7.4)
		3	2386	2010 (1153,2850)	5	4.4 (1.2,7.5)
		4	3090	2728 (2098,3350)	6	5.1 (2.1,8.2)
	8	1	1494	1727 (1107,2363)	7	4.8 (1.7,7.9)
		2	2318	2466 (1853,3128)	10	7.5 (4.4,10.7)
		3	3141	3153 (2526,3785)	13	11.2 (8.1,14.3)
		4	3965	3903 (3291,4531)	16	14.4 (11.4,17.5)
	9	1	3154	3360 (2721,4010)	10	8.1 (4.9,11.3)
		2	4393	4726 (4102,5385)	16	14.2 (11.2,17.5)
		3	5631	6048 (5422,6704)	21	21.2 (18.1,24.5)
		4	6870	7556 (6911,8226)*	27	26.2 (23,29.5)
5	7	1	1201	1153 (530,1808)	3	2.9 (0,5.9)
		2	1914	1692 (1070,2343)	4	4.3 (1.2,7.4)
		3	2628	2232 (1621,2869)	5	5.8 (2.8,9)
		4	3341	2844 (2229,3485)	7	7.6 (4.6,10.7)
	8	1	1596	1848 (1225,2493)	9	4.7 (1.6,7.8)*
		2	2541	2633 (2017,3286)	13	8.5 (5.3,11.6)*
		3	3486	3359 (2756,4019)	18	13.6 (10.5,16.6)*
		4	4431	4167 (3565,4816)	23	18.1 (15.1,21.2)*

NTU	pH	LRV	Monochloramine		Free chlorine	
			CT mg.min/L Table	CT mg.min/L This study Mean ($P_{2.5th}$,$P_{97.5th}$)	CT mg.min/L Table	CT mg.min/L This study Mean ($P_{2.5th}$,$P_{97.5th}$)
	9	1	4400	4253 (3635,4902)	10	8.6 (5.4,11.8)
		2	5967	5889 (5284,6547)	16	16 (13,19.2)
		3	7535	7442 (6838,8112)	23	24.3 (21.1,27.6)
		4	9102	9127 (8452,9852)	29	30.2 (27.1,33.4)

*CT from guidelines outside obtained interval

7.4. Discussion

7.4.1. Regression models versus multilayer perceptron

The methodology employed on this study relied on neural networks to describe the nonlinear relationships between the input variables and LRV. The major advantage of neural networks is their capability to model those relationships without imposing any structure on the relationship between the interacting factors. Simpler models including multiple linear regressions with and without non-linear terms were also tested for their fit. However, preliminary results using conventional least squares for the regression and backpropagation for the multilayer perceptron indicated that the MLP fitted the data better with a higher determination coefficient ($R^2_{regression}$: CB5=0.72, AD2=0.81; R^2_{MLP}: CB5=0.98, AD2=0.98) and smaller root mean squared error ($RMSE_{regression}$: CB5=5.2, AD2=990; $RMSE_{MLP}$: CB5=2.1, AD2=446). Another alternative was the model developed by Smith *et al.* (1995). However, while the model provided adequate fit for the Coxsackievirus data (R^2_{Smith}: 0.96, $RMSE_{Smith}$: 2.1), the Adenovirus data were not very well fitted (R^2_{Smith}: 0.82, $RMSE_{Smith}$: 1007). The alternative models were checked by Bayesian inference and plots with the model fits are presented for comparison in Appendix 4.

7.4.2. Developing models from experimental results

Development of models to describe the relationship between LRVs and independent variables is central to probabilistic use of water quality monitoring data in the prediction of process performance e.g. for QA/QC and QMRA purposes. Building a model with the variables presented

in this study enabled the interpolation of experimental trial data and the capacity to investigate a broader range of disinfection scenarios for a continuous target variable compared to the fixed experimental conditions. The BMLP model enabled to obtain estimate probability distributions for both their input and output variables considering both uncertainty in the parameters and variability of the measurements. Further, the use of the Bayesian inference capabilities allowed backwards (e.g. prediction of pH, turbidity provided CT and LRV are observed) as well as forward reasoning (e.g. prediction of CT provided pH, turbidity and/or LRV are observed) to be performed. The BMLP model used in this study permitted defensible interpolation of CT, turbidity, pH and LRV values, while treating all variables as continuous. Moreover, it provided tools for predicting the variability of target CTs during chlor(am)ination. The Bayesian aspect of the model allowed querying target variables including CT, pH or turbidity given a set of observed evidence. In the scenarios presented in this study, the prior distributions for the query variables were uniform reflecting full uncertainty in their estimates. As the choice of prior affects the results of the posterior distributions, the particular measured water quality probability density functions for pH and turbidity during chlorination should be included as priors. Incorporating these priors would allow obtaining a posterior distribution specific to the analysed process.

While a number of data mining tools for process monitoring during wastewater treatment have been reported (Toifl *et al.*, 2010; Dürrenmatt & Gujer, 2012; Haimi *et al.*, 2013), the use of BNs in this area is currently limited. However, the potential application of this BN based learning offers a number of important advantages, which conventional statistical tools do not. Another possible BN application involves development of predictors of pathogen removal by wastewater treatment processes (Carvajal *et al.*, 2015).

The relationship between LRVs and CTs is usually defined for fixed water quality conditions especially temperature, pH and turbidity. Disinfection models for the analysis of these data commonly use kinetic constants and parameters which are defined separately for each condition (Haas, 2004). Previously, multivariate regression models have been generated to perform interpolation for potable water chlorination (Clark *et al.*, 1989; Smith *et al.*, 1995). However, this approach has been limited to predicting a single LRV corresponding to a single CT, which must then be linearly extrapolated to an LRV corresponding to an observed CT. The approach does not usually include uncertainty in the model parameters. Haas (2004) highlighted the uncertainties introduced by interpolation and investigated the use of neural networks to predict inactivation of *Giardia lamblia* by free chlorine in drinking water. Neural networks have also been developed to predict the inactivation of *Cryptosporidium parvum* using ozone and chlorine dioxide (Janes & Musilek, 2007b) as well as bacteria using free chlorine and chloramine (Janes & Musilek, 2007a). These models were all designed for the purpose of predicting microbial inactivation considering

variation in pH, temperature, residual concentration and contact time. Assessment of uncertainty and variability in the model prediction was not assessed on these studies.

7.4.3. Need for stochastic prediction

The modelling illustrated here offers a way for maintaining the traditional use of CTs in managing disinfection while making the outputs more probabilistic in line with the movement toward more quantitative methodology (e.g. US-EPA and USDA/FSIS (2012)). Conventional Monte Carlo style probabilistic models are one possible approach. However, Bayesian inference for model definition outperforms conventional Monte Carlo simulation in terms of naturally defining uncertainty in the model parameters, allowing interaction between data and parameters and producing estimates from backwards inference (Smid *et al.*, 2010). The alternative analysed in this study exploits the power of Bayesian inference and neural networks without sacrificing the benefits of continuous distribution. BMLP models can efficiently represent large amounts of information, such as might be collected in a water quality monitoring database table, in a compact and efficient manner. Though the model structure is strictly "acyclic" the flow of information in BMLP models is not limited to the direction of the arcs (Korb & Nicholson, 2011).

Current guidelines for chlorine disinfection promote a conservative estimation of LRVs reflecting peak flowrates, minimum operating volumes and free chlorine residual concentration at peak flows at the effluent of the reactor. At the same time, the tables which are central to predicting target CTs cannot easily incorporate interpolation or the uncertainty associated with CT and LRV point estimates. Similarly, tabulated CT values provide a conservative estimate for the required removal by rounding or using a point estimate from the confidence interval for CT after fitting a model (Clark & Regli, 1993; Sivaganesan & Rice, 2003). In combination, these effects lead to a highly conservative calculation for the achieved LRVs. This conservative dosage of disinfectant may lead to excessive amounts of by-products formation.

Treating CT and LRV estimates as probability density functions has a number of advantages. It provides an effective means for evaluating quantifying inherent reliability of disinfection processes while explicitly acknowledging the variability and uncertainty of the process data and models and the vulnerability of systems operating outside specifications. In this manner, the probability that a treatment target has or has not been achieved can be estimated closer to reality and the assumptions and uncertainties underlying these estimates become clear. That said, the models developed in this study still reflect the limits of the available data and there were two important constraints which need recognition. Firstly, other well documented influences on chemical disinfection were not included such as temperature. Secondly, the input data are

necessarily restricted in the validation experiment ranges they cover. Although temperature was not analysed, the CT measurements were undertaken at a low (conservative) temperature. The model used in this study was constructed from experimental data considering specific ranges for the input and output variables. Therefore, the data define an envelope in which the model is known to behave correctly. Using the model outside this envelope does not necessarily mean that the estimates will be incorrect, but that they are not possible to be verified. Extending these ranges would also extend the envelope for the model to reliably predict. Incorporating new data records or variables would be a simple task by using the existing algorithms for BMLP model construction used in this study.

Experimental bench-scale batch reactor analysis to obtain CT values has been the common way of defining the chlorination requirements as this is equivalent to an ideal plug flow reactor. The reasons are of practical implication, especially the lower cost compared to full-scale experiments. Bench-scale reactor experiments have proven to be more reliable and more controllable than full-scale settings (Gyürék & Finch, 1998). Due to the specific hydraulic characteristics of the full-scale reactors, obtaining data would be very specific for that particular configuration, making it difficult to generalise their results. The non-ideal reactor flow conditions are usually considered by assuming the contact time as the T_{10} or the time it takes for 10 % of the water pass through the reactor.

A related additional benefit of the BMLP modelling was that the probabilistic form of the input and predicted CT values captures the uncertainty more effectively than conventional algorithm based models. Even when a conventional model has been fitted to a widely accepted algorithm format such as Chick-Watson (LeChevallier & Au, 2004), the kinetic parameters incorporate a level of uncertainty and/or confidence intervals. Reflecting this uncertainty, a minimum target CT and accompanying safety factor, instead of a single point estimate, has been recommended by Sivaganesan *et al.* (2003). The approach to modelling presented here eliminates the need for such conservatism and provides a method for making implementation of this recommendation routine.

The proposed approach does not attempt to change the methodology currently employed in regulations and guidelines. This study seeks to solve the issues of interpolation and incorporation of probabilistic estimation during chlorine disinfection. Interpolation of water quality characteristics is useful when log credits change over time because of changing operational conditions. In this case, there is no need for worst case scenario assumptions when the CT values are obtained from the tables at conditions not previously tested. Moreover, a continuous estimation of the process performance can be obtained with the BMLP model. Stochastic estimates permit the calculation of exceedance probabilities instead of compliance/not

compliance cases. All models have a degree or error; in this case it is observed that the model does not match the estimates from the tables published by DOH (2013) in some instances (turbidity 5 NTU and pH 8). A simple model was fitted with the data for the latter conditions at which the models presented discrepancies. The model was found to fit the data closely with a determination coefficient of 0.97. Using this model to find the CT values at specific LRVs, it was observed that the BMLP model produced the same or very similar values. It is possible that these differences are due to very conservative assumptions used during the CT table construction.

7.5. Conclusions

The influence of water quality parameters including pH and turbidity in treated wastewater for recycling must be quantified and considered in CT calculations to ensure efficient disinfection of target pathogens. This study has demonstrated that BN models can be used to jointly model the effect of three variables (target LRV, pH and turbidity) during chlor(am)ination on the target CT value estimation.

Key outcomes from this study included:

- The combined/overall model provided a tool for obtaining LRV distributions for multiple scenarios and estimating CTs for target LRV given certain or uncertain pH and turbidity.
- The sensitivity analyses showed that pH had higher impact on LRV than turbidity within the ranges covered by the CT experiments.
- The use of the Bayesian multilayer perceptron model permitted interpolation between the defined levels of pH and turbidity avoiding the use of manual interpolation and conservative assumptions.

Incorporating the uncertainty in the model parameters plus the variability of the operational process parameters in this way resulted in a more realistic characterisation of chlor(am)ination disinfection process.

Further improvements to the conceptual model could incorporate the effect of other water matrix characteristics including temperature, dissolved organic matter and ionic strength. Conventional tools for communicating and representing this information and data rely primarily on tables which do not incorporate probabilities, or achieve full graphical representation and parameter interaction and analyse multiple scenarios and queries. By contrast BNs address this and offer further advantages in knowledge representation including explicit association between the variables and probabilistic outcomes.

Chapter 8: Assessing reliability by using Fault tree analysis and Bayesian networks during UV disinfection

8.1 Introduction

Reliability for water reuse treatment systems must be quantitatively understood to ensure that an appropriate level is achieved to match the consequences of failure. Quantitative reliability assessment requires appropriate tools such as Bayesian networks (BNs) to undertake structured evaluation of failure modes and likelihoods. However, few studies have explored this subject with regard to water treatment processes, revealing the need for further investigation.

The aim of this case study was to evaluate the use of BNs for analysing the reliability of a UV reactor treatment for water reuse. UV systems are commonly found during tertiary treatment and represent an important barrier for pathogens and trace chemicals when used for advanced oxidation. Models were constructed based on information found in published literature and by communication with practitioners. The analysis included the model construction, reliability estimation and factors influencing the reliability of the system.

Fault tree analysis (FTA) was used to define the causal structure of the BN model. Fault tree analysis was previously outlined in Section 2.5.3. In a FTA causal factors are deductively identified and organised in a logical manner using a tree diagram (IEC/ISO, 2009). Logical gates are used to represent the relationships between the causal factors and the top event. Mapping FTA into BNs has previously been applied in several studies (Bobbio *et al.*, 2001; Montani *et al.*, 2005; Khakzad *et al.*, 2011) as described in Chapter 2. However, these applications were not related to water treatment systems.

UV systems comprise components that can be repaired. In this case, to analyse the reliability of the UV system, repair has to be considered. In repairable systems, availability is used instead of reliability. Availability is defined as the fraction of time a component or system is able to perform its required function (O'Connor & Kleyner, 2012). Long term availability is computed by considering the mean time between failures (MTBF) and mean time to repair (MTTR). In this study, the probability that sensors were in a failed condition was estimated using unavailability (i.e. the complement of availability).

Due to data limitations constraints, this study focused on a theoretical analysis. Even though the study was conceptual, the focus in this chapter was to demonstrate the usefulness of BNs for water reuse reliability analysis. Therefore, it is important to consider that there were a number of occasions where illustrative data are used.

8.2 Materials and methods

8.2.1 Bayesian network, software and evaluation

Bayesian networks were used in this study because the interactions between the various components in the UV system need to be evaluated. BNs permit explicit representation and assessment of probabilistic associations between variables. BNs methodology was introduced in Section 3.1.

Models were constructed and evaluated in Bayes Server 7.18 (Bayes Server, 2017). This software package has the special feature of permitting various types of analyses (see Section 3.1.5 for explanation). The model was evaluated using the mutual information, *lift*, *difference* and sensitivity to parameters described in Sections 3.1.4.5 and 3.1.4.6 (Kjræulff & Madsen, 2012). Evidence was entered on different nodes to assess their impact to the target variable. These analyses were undertaken using the methodology previously explained in Section 3.1.4.

8.2.2 System explanation and model construction

This study was undertaken based on the assessment of a theoretical UV system for the purpose of investigating a conceptual reliability analysis. The configuration of the theoretical UV system was based on a real system (operating at Glenelg water reuse plant in South Australia), but with some minor modifications. In the real system, six trains are used, whereas in the analysed model, only two trains were considered. The theoretical system was assumed to consist of two trains with two reactors per train and forty lamps per reactor. Each reactor has a UV intensity (UVI) sensor, while each train has a flowrate sensor. UV transmissivity (UVT) is measured before the water enters the process. The sensors influence the control of parameters including flowrate and UVI, except for UVT which depends on the performance of upstream processes. As observed in Figure 8-3 and Figure 8-4, the sensors affect the reliability of the system (system dose) through a low UVI, low UVT or high flowrate. Included model variables and their descriptions are presented in Table 8-1. All variables have been selected to correspond to one of two states, *Yes* or *No*.

Table 8-1: Description and states of variables in the model

Variable	Description
Low dose not detected?	A low UV system dose has not been detected.
Train low dose not detected?	A low UV train dose has not been detected.
Reactor low dose not detected?	A low UV reactor dose has not been detected.
Low UVI not detected?	Low UVI was not detected causing a hazardous event.
UVI sensor giving too high values?	UVI sensor fails giving higher values than reality.
Too high flowrate not detected?	High flowrate is not detected by the monitoring system causing a hazardous event.
Flowrate sensor fails giving too low values?	The flowrate sensor fails giving lower values than reality.
Low UVT upstream affecting dose?	Low UVT is not detected and affected UV dose producing a hazardous event.
Low UVT	Low UVT from upstream processes.
UVT sensor giving too high values?	UVT sensor fails giving higher values than expected.

A flowchart diagram was constructed to explain the steps of model development (Figure 8-1). As observed, model construction was an iterative process at which each step was checked before continuing to the next. The first step ("Define model objective") consisted of defining the purpose or aim of the model. In this case the objective was not detecting a low UV dose. The second step consisted of defining the structure of the model. In this case the structure was based on a FTA approach using expert elicited information and technical literature as explained in Section 8.2.4. The third step was used to define the parameters of the model from expert elicited data (Section 8.3.2). The fourth step consisted of evaluating the model through sensitivity and scenario analysis as explained in Sections 0 to 8.3.5. Finally, the model was checked to decide whether it was accepted or not. If the model was not accepted, modifications to the parameters or structure would need to be made. It is important to note that by incorporating changes to the structure, modification of model parameters would be necessary.

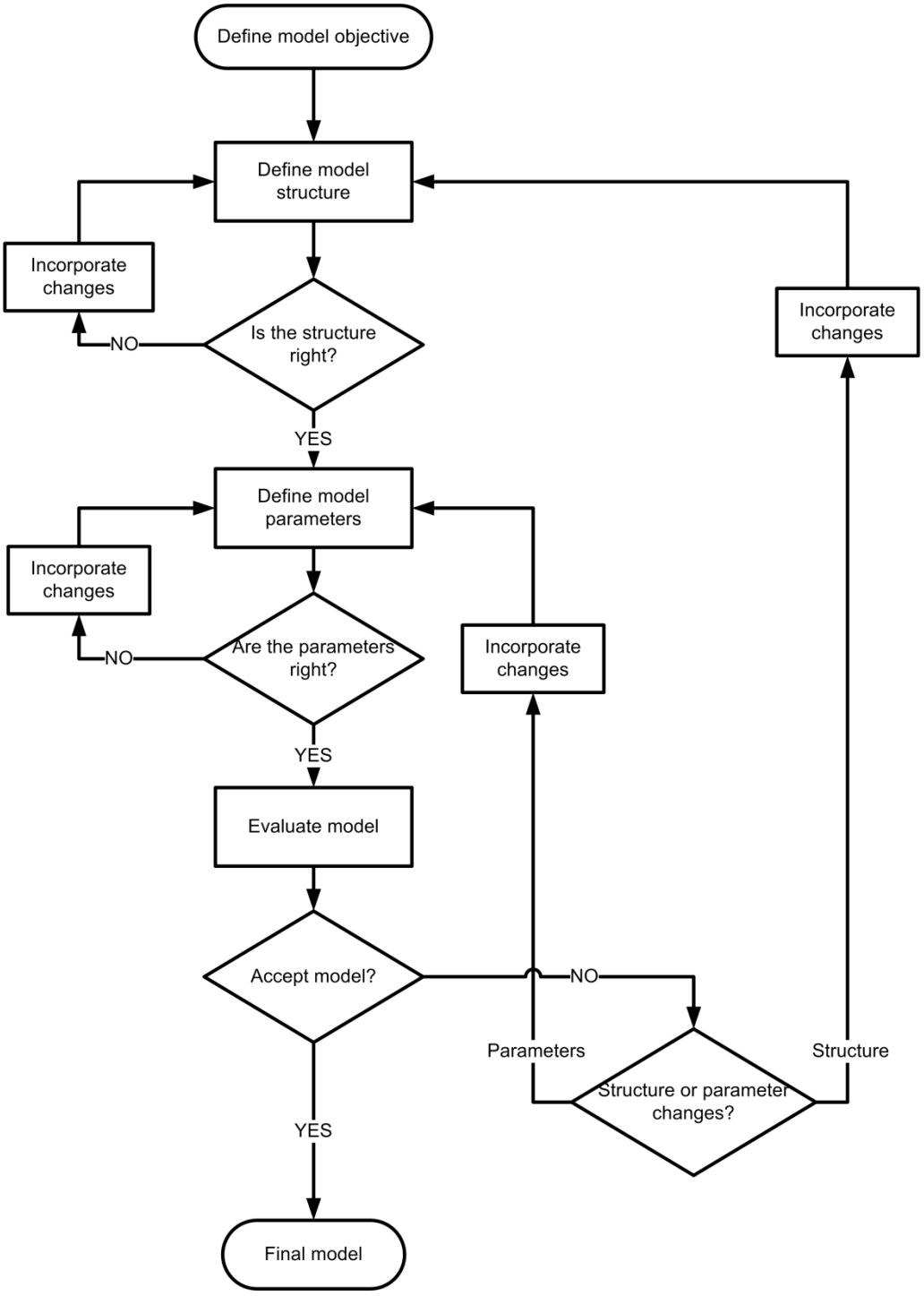

Figure 8-1: Flowchart diagram for BN model development

8.2.3 Available data and estimation of parameters

Reliability data for advanced water treatment processes are scarce with only a few reports gathering information regarding failure and repair rates from full-scale systems (Forss & Ander, 2011; Tng *et al.*, 2015). The datasets from these reports contain data in the form of failure rates

and repair times for a number of components on advanced water treatment systems. However, the available identified data for UV systems were not sufficiently detailed to generate a model to analyse reliability. To develop a reliability model, the required information should include data about subcomponents, such as sensors and lamps. To address this limitation, one possible appropriate approach is to collect these data through an expert knowledge elicitation process (Lindhe *et al.*, 2009; Lindhe *et al.*, 2012). The consulted expert in this study is the Product and Application Manager for UV & Ozone of one of the leading water technology companies worldwide (Xylem Water Solutions Australia). He possesses more than eleven years of experience on UV system design and validation for water treatment. The elicitation approach consisted of asking the qualified expert to provide his "best guess" regarding specific parameters, in this case, both failure rates and repair times. A common approach for eliciting values consists of requesting the expert(s) to specify two or three values, commonly the 5th percentile, median and 95th percentile (Lindhe *et al.*, 2009; Lindhe *et al.*, 2012; Morris *et al.*, 2014). Then, these values are used to fit a probability distribution through a maximisation method. In this study, 5th percentile, median and 95th percentiles were requested from the expert. The expert thus had the freedom between providing two or three values. To keep the model simulation simple, a point estimate (mean) of the fitted distribution was used.

Availability (Equation 8-1) was used to compute the probability of having a component working in the system.

$$A = \frac{MTBF}{MTBF+MTTR} = \frac{\mu}{\mu+\lambda} \qquad \text{Equation 8-1}$$

$$\mu = \frac{1}{MTTR} \qquad \text{Equation 8-2}$$

$$\lambda = \frac{1}{MTBF} \qquad \text{Equation 8-3}$$

Where A is the availability, μ is the repair rate, λ is the failure rate. In these equations the rates are considered constant (exponential failure and repair times). Unavailability ($\bar{A}$) is defined as 1-availability (1-A).

An OR gate in a FTA (Figure 8-2 a) is used to represent an event which occurs if and only if at least one the n basic events (E_1,…, E_n) occurs. For the case of system unavailability this can be represented by Equation 8-4:

$$\bar{A} = 1 - \prod_{i=1}^{n}(1 - \bar{A}_i) = \Pr(E_1 U \ldots U E_n) = 1 - \prod_{i=1}^{n}(1 - \Pr(E_i)) \quad \text{Equation 8-4}$$

An AND gate (Figure 8-2 b) is used to represent an event which occurs if and only if the n events (E_1,…,E_n) occur simultaneously. The system unavailability in this case is represented by Equation 8-5:

$$\bar{A} = \prod_{i=1}^{n}\bar{A}_i = \Pr(E_1 \cap \ldots \cap E_n) = \prod_{i=1}^{n}\Pr(E_i) \quad \text{Equation 8-5}$$

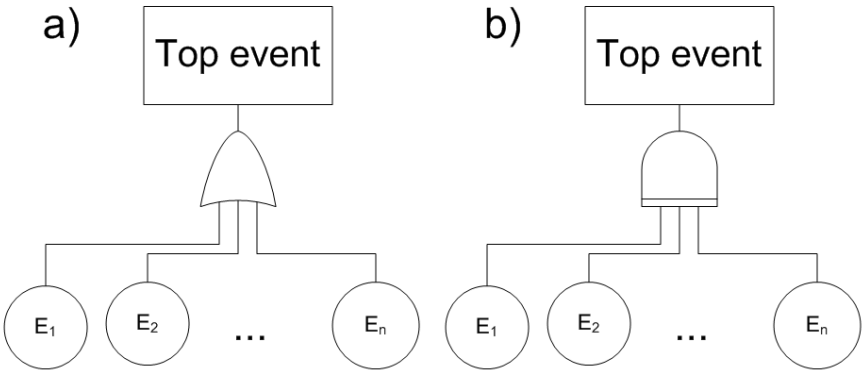

Figure 8-2: Representation of gates of FTA used in this study. a) OR gate, b) AND gate.

8.2.4 Fault tree analysis representation

The UV reliability model was constructed based on technical guidance for the operation and assessment of UV disinfection systems (USEPA, 2006) and expert elicitation with a UV disinfection expert (as described above). In this analysis, a hazardous event was defined to occur when any of the sensors related to the dose calculation indicate better performance than reality, so both the system and operator would be unable to detect an underperformance. The aim of this model was to assess the likelihood of obtaining a non-detected low system dose. The model assumed that providing that the sensors work properly, events impacting the dose, such as fouling of the lamp sleeves, ageing and failure of lamps can be diagnosed and corrected by the operator in a timely manner. In the case of UVT falling below the threshold value, the corrective action

after the sensor captures the event, is to cease water production. At the plant used as the basis for this model, if UVT falls to below 50% for more than 30 min, an automatic plant shutdown is triggered (GPA Engineering, SA Water, AllWater, 2011).

Fault tree analysis (Figure 8-3) was used to generate the base model for the BN reliability model. FTA provides the causal structure and the type of associations between the variables through the logic gates. In this case, two types of gates including OR and AND gates (see Section 8.2.3) were used. The FTA model aids in targeting specific causal factors for a particular undesired event.

Considering the objective of the model, the top event in the FTA (Figure 8-3) was defined as [*Low system dose not detected?*]. This event occurs when both trains have low dose (event: [*Train low dose not detected?*]). Train low dose is produced if any of the reactors presents a low dose ([*Low reactor dose not detected?*]). A low reactor dose occurs if the flowrate is too high ([*Too high flowrate not detected?*]), if the UV intensity is too low ([*Low UVI not detected?*]) or if there is low UVT ([*Low UVT affecting dose?*]). Low UVI takes place if the [*UVI meter fails giving too high values*] (Figure 8-4). Flowrate is too high if the [*flowrate sensor fails giving too low values*] (Figure 8-4). Low UVT causes issues to the trains UV dose when there is a [*low UVT*] and the [*UVT sensor fails giving too high values*]. As observed in Figure 8-3 there are common cause events in the system (represented by the same colour). These events in the FTA are repeated on various parts of the model. The FTA model in Figure 8-3 was encoded into a BN (Figure 8-4).

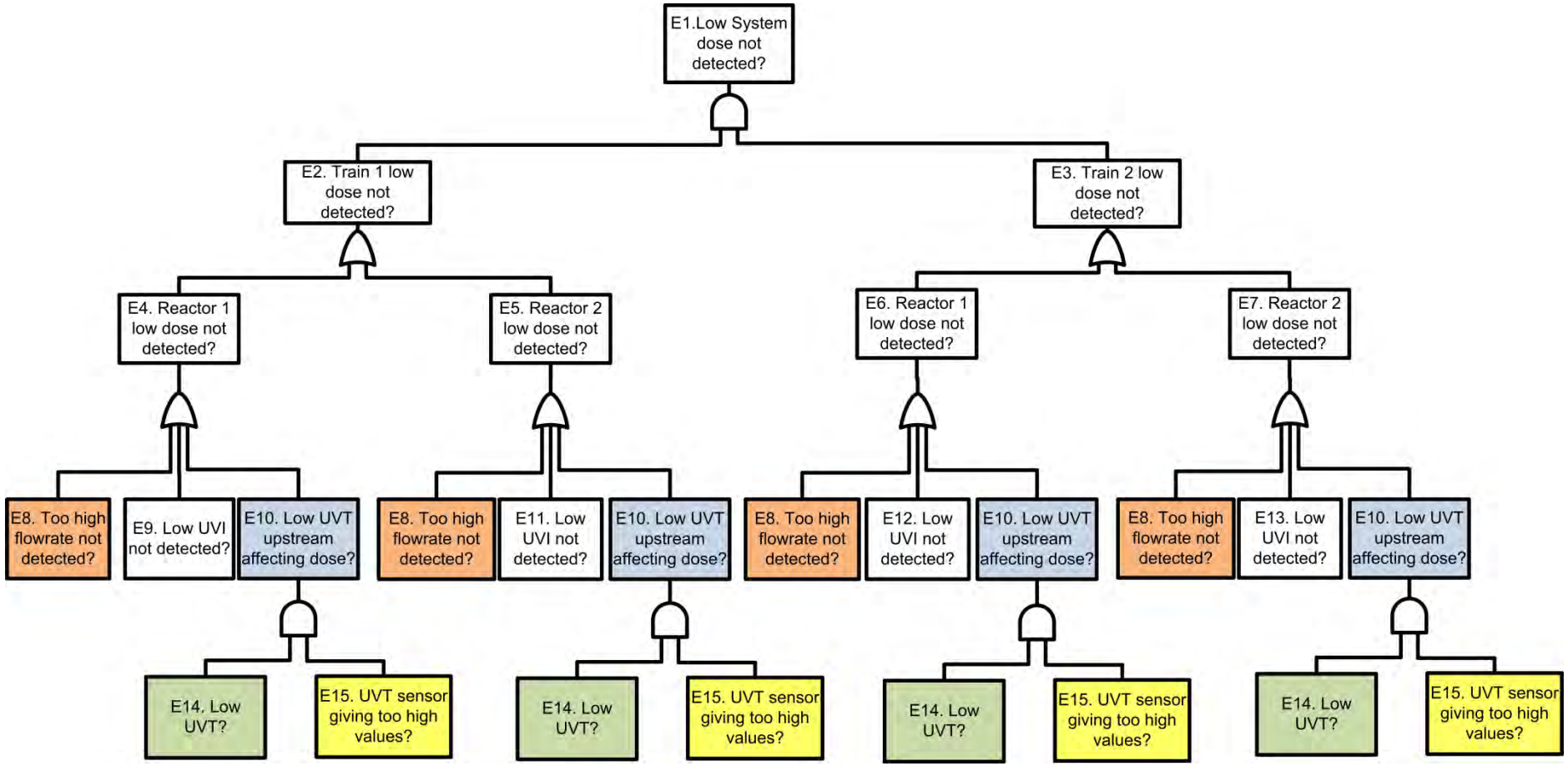

Figure 8-3: Fault tree representation of the UV system. Blocks in colours indicate the same events.

8.3 Results and discussion

8.3.1 Bayesian network model

The BN derived in this study was represented as a directed acyclic graph (DAG) (Figure 8-4), which shows the causal relationships between the different variables in the system. As observed, the target variable is [*Low system dose not detected?*]. However, any variable in the model can be assessed. Each variable in the model had states *Yes* and *No.* The BN model is presented in Figure 8-5.

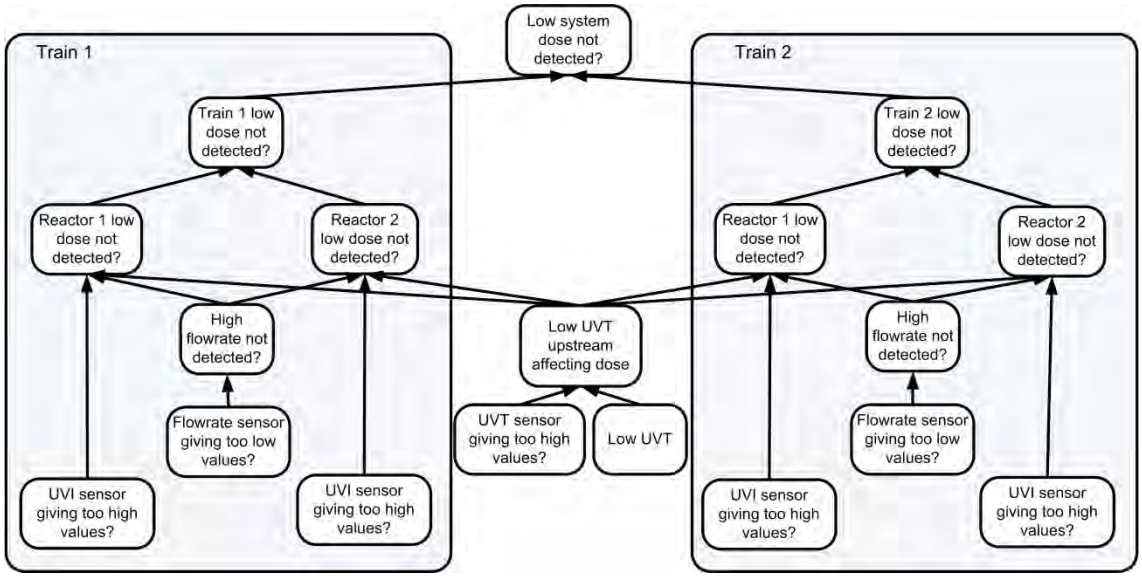

Figure 8-4: DAG for UV system reliability assessment

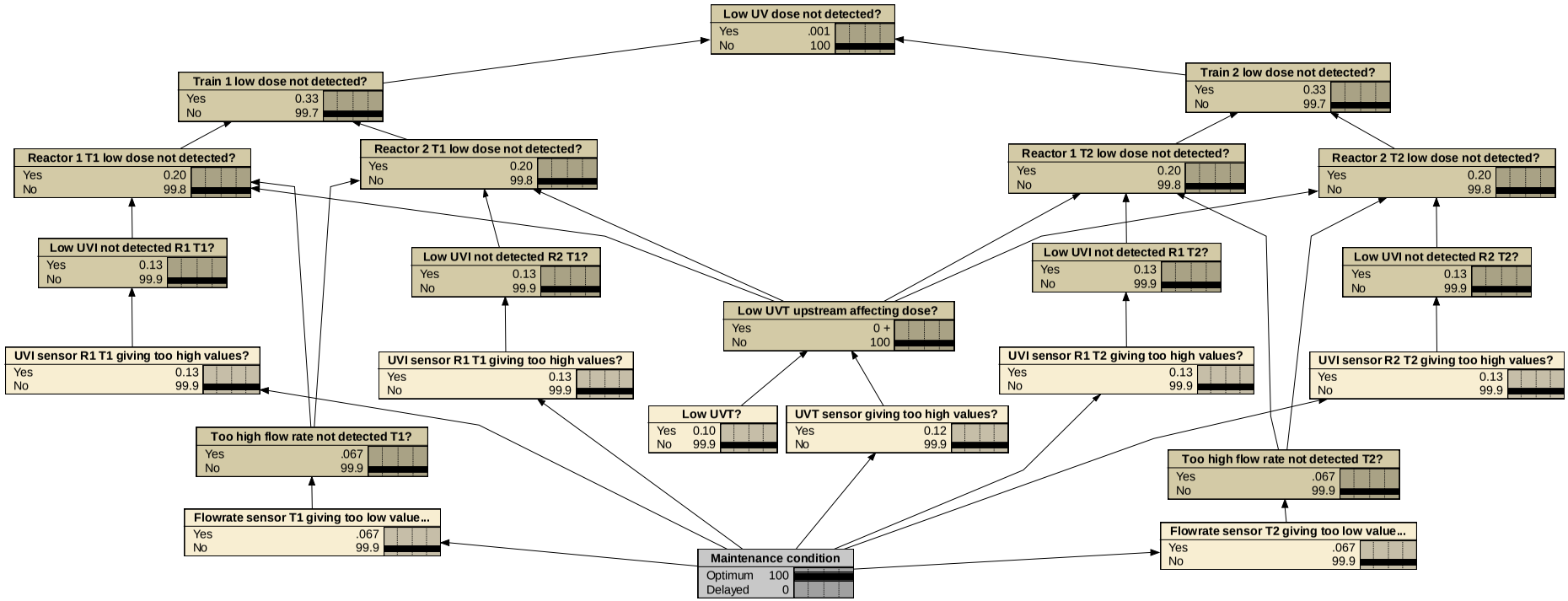

Figure 8-5: BN for the UV system reliability assessment.

8.3.2 Elicited values as parameters distributions

The elicited values for the time between failures (Table 8-2) were fitted to Gamma distributions (Table 8-3). Gamma distributions were selected because of their flexibility and because they can only take positive values. Furthermore, Gamma distribution can be used as a conjugate prior to the exponential distribution (Gelman *et al.*, 2014). Once real data (number of failures and the specific time period) are collected, they can be used to update the prior Gamma distribution. Repair times were considered constant and equal to 24 hours when maintenance logistics are optimum. Repair times are often heavily dependent on the logistics of the company. Therefore, a worst case-scenario was considered in which maintenance delays affected the repair times. Repair times under delayed conditions were obtained by expert elicitation (Table 8-3). The BN model can incorporate this information and assess its impact over the target variable. A variable with name [*Maintenance*] and two states (optimum and delayed) was included in the model as a parent variable of [*UVI meter fails giving too high values*], [*flowrate sensor fails giving too low values*] and [*UVT sensor fails giving too high values*] to capture its impact. The failure rates (λ) were estimated as the inverse of the mean time between failures (1/MTBF) (Table 8-4). In the same manner, the repair rates were calculated as the inverse of the constant repair time (1/MTTR). For [*Low UVT*], the threshold was defined as 50% (GPA Engineering, SA Water, AllWater, 2011). UVT was assumed to be 99.9% of the time above the threshold considering a consistent performance of filtration before UV disinfection. This value was based on real UVT data from a water reuse scheme in Alice Springs using sand media filtration (Figure 8-6). These unpublished data were provided by the expert consulted in this study. It is important to note that capturing out of specification circumstances is rare in well operated systems like the one producing these data. Because the objective is studying reliability, the interest is in potentially very rare occurrences which are unlikely to have been captured in a normal data acquisition or process monitoring period. Using long term data can assist in capturing events affecting the reliability of the system.

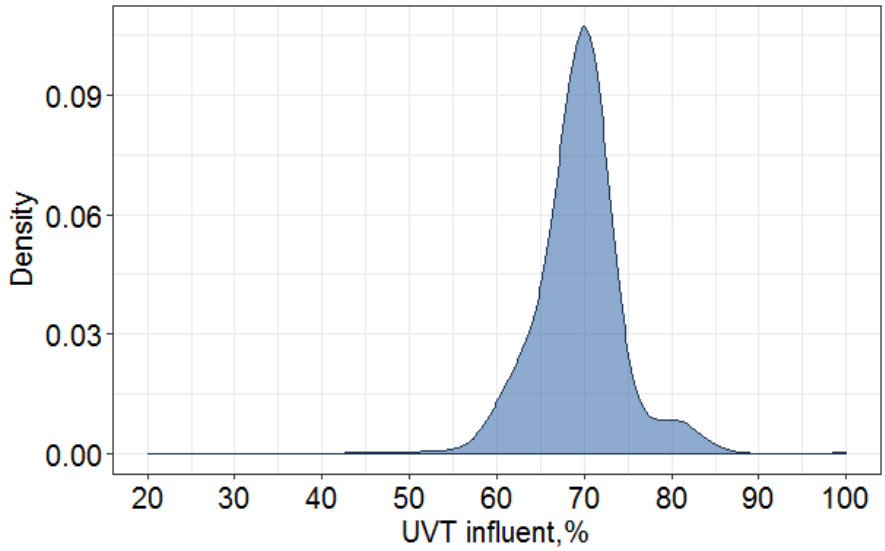

Figure 8-6: Smoothed histogram for the influent UVT

Table 8-2: Raw elicited values for sensors failure times (from expert elicitation)

Component	Elicited 50th percentile	Elicited 95th percentile
UV intensity sensor	2 years	3 years
Flowrate sensor	4 years	6 years
UVT sensor	2 years	5 years

Table 8-3: Failure times and repair times for the components in the UV system

Component	Elicited Failure time	Repair time optimum	Repair time delayed
UV intensity sensor	Gamma(14.6,7.13) years	24 hours	60 days
Flowrate sensor	Gamma(14.6,3.56) years	24 hours	15 days
UVT sensor	Gamma(2.60,1.14) years	24 hours	30 days

Table 8-4: Mean failure rates, repair rates used in the BN

Component	Failure rate	Repair rate optimum	Repair rate delayed
UV intensity sensor	6E-05 $hour^{-1}$	4 E-02 $hour^{-1}$	7E-04 $hour^{-1}$
Flowrate sensor	3E-05 $hour^{-1}$	4E-02 $hour^{-1}$	3E-03 $hour^{-1}$
UVT sensor	5E-05 $hour^{-1}$	4E-02 $hour^{-1}$	1E-02 $hour^{-1}$

Two types of impacts to [*Reactor low dose not detected?*] were tested including an OR gate and the noisy-OR gate. A noisy-OR gate in BNs makes the assumption of independence between the causal factors to decrease the number of parameters required to define the conditional probability table. For this target event, any of the three causal factors can result in the reactor producing a low dose. Therefore, a noisy-OR gate would be adequate in this case. The effect of each causal variable over the event has to be determined individually with one probability. The noisy-OR assumption relaxes the assumption taken by the OR gate. Therefore, the occurrence of an event does not necessarily provoke the target event to occur with 100% probability. The required parameters for the noisy-OR gate are presented in Table 8-5. These parameters were assumed and not based on real data due to data limitations constraints. Noisy-OR can also incorporate an inhibition factor, which can be understood in this case, as the probability of having an out-of-specification dose given that none of the causal factors are present. In this case, the leak was as assumed to be zero.

The BN model using OR gates indicated that the probability of having a [*Low system dose not detected=yes*] was 1E-05 (0.001%), while the probability for this same outcome obtained for the model using noisy-OR gates was 9E-06 (0.0009%). These results indicate a very reliable system with a yearly failure time of around 5 to 6 minutes.

Table 8-5: Parameters used in the noisy-OR gate for the event [*Reactor low dose not detected?*]

Conditional probability	Value
P(Reactor low dose not detected? =yes|Too high flowrate not detected=yes)	0.5
P(Reactor low dose not detected? =yes|Low UVI not detected=yes)	0.9
P(Reactor low dose not detected? =yes|Low UVT upstream affecting dose=yes)	0.8

8.3.3 Sensitivity to evidence

Mutual information and relative mutual information were used to evaluate the sensitivity of the node [*Reactor low dose not detected?*] to the other model variables. Higher values of mutual information and relative mutual information indicate higher reduction of uncertainty. The results for the model using OR gates and noisy-OR gates are presented in Table 8-6. As observed, the sensitivity results were similar for both types of models. From the three sensors considered in this model, the UVI sensor had the highest impact on the system UV dose with relative mutual information of 14% (14% for the noisy-OR model). The other two sensors produced relative mutual information of 7% (5% for the noisy-OR model) for the flowrate sensor and 3% (3% for the noisy-OR model) for the UVT sensor. Sensitivity to evidence is a function of both the parameters and structure of the model. Even though there are two UVI sensors per train (one for each reactor), their relative mutual information was double the one obtained for the flowrate sensors. Because any of the two reactors failing can cause the train to fail, they do not provide real redundancy to the system. The impact shown by the UVT sensor can be explained by the fact that its impact largely depends on the probability of having low UVT. Modifying the probability of [*Low UVT?=no*] from 0.001 to 0.01 produced an increase in the relative mutual information of [*UVT sensor giving too high values?*] of 24%. This sensitivity is relevant to consider because UVT is not a controlled parameter and depends on the performance of upstream processes. Therefore, the UVT sensor reliability becomes more important as the performance of the previous processes is reduced.

Table 8-6: Mutual information (MI) and relative mutual information (RMI) for model using OR and noisy-OR gates

	OR gate		Noisy-OR gate	
Variable	MI[a]	RMI[b]	MI	RMI
Train low dose not detected?	7.0E-05	46%	5.6E-05	47%
Reactor low dose not detected?	4.1E-05	27%	2.8E-05	24%
Low UVI not detected?	2.1E-05	14%	1.7E-05	14%
UVI sensor giving too high values?	2.1E-05	14%	1.7E-05	14%
Low UVI not detected?	2.1E-05	14%	1.7E-05	14%
Low UVT upstream affecting dose?	1.4E-05	9%	1.3E-05	10%
Too high flowrate not detected?	1.0E-05	7%	6.6E-06	5%
Flowrate sensor giving too low values?	1.0E-05	7%	6.6E-06	5%
Low UVT?	4.4E-06	3%	4.3E-06	4%
UVT sensor giving too high values?	4.2E-06	3%	4.1E-06	3%

[a]: Mutual information

[b]: Relative mutual information

Under delayed maintenance conditions, the probability of having [*Low system dose not detected=yes*] increased to 0.023 (2.3%) using the OR gate model. The mutual information under delayed maintenance increased for all nodes except for [*UVT sensor giving too high values?*] compared to the optimum maintenance conditions (Table 8-7). Despite the higher mutual information for the delayed conditions, the relative mutual information reduced for [*Flowrate sensor giving too low values?*] and [*UVT sensor giving too high values?*]. This result implies that although the variables provide information about the target event, this amount of information is low compared to the uncertainty of the target variable.

Table 8-7: Mutual information (MI) and relative mutual information (RMI) when considering optimum maintenance and delayed maintenance

	Optimum		Delayed	
Variable	MI[a]	RMI[b]	Variable	MI[a]
Train low dose not detected?	7.0E-05	46%	4.5E-02	41%
Reactor low dose not detected?	4.1E-05	27%	1.7E-02	16%
Low UVI not detected?	2.1E-05	14%	1.5E-02	14%
UVI sensor giving too high values?	2.1E-05	14%	1.5E-02	14%
Low UVI not detected?	2.1E-05	14%	1.5E-02	14%
Low UVT upstream affecting dose?	1.4E-05	9%	1.3E-04	0%
Too high flowrate not detected?	1.0E-05	7%	1.7E-03	2%
Flowrate sensor giving too low values?	1.0E-05	7%	1.7E-03	2%
Low UVT?	4.4E-06	3%	1.8E-05	0%
UVT sensor giving too high values?	4.2E-06	3%	7.0E-07	0%

[a]: Mutual information

[b]: Relative mutual information

8.3.4 Sensitivity to parameters

Sensitivity to parameters was used to test the influence of individual (one-way sensitivity analysis) and pairs of variables parameters (two-way sensitivity analysis) over the target variable [*Low system dose not detected?*]. The results can be presented as a curve for the case of individual variables or as a surface for the case of a pair of variables. The results for individual parameters for the sensors and [*Low UVT*] (Figure 8-7) on one train indicated that the target variable ([*Low system dose not detected?*]) was slightly impacted by the change in parameters of the tested variables. Increasing the marginal probabilities for the state *Yes* for the tested variables had a positive increase in the probability of having a [*Low system dose not detected=yes*]. The largest influence was observed for the [*UVI sensor giving too high values*] and [*flowrate sensor giving too low values*] (Figure 8-7) with a maximum value of 0.0035 approximately (0.35%). [*Low UVT*] and [*UVT sensor giving too high values?*] presented the lowest influence over the target variable.

Sensors failing individually in one train had little influence on the reliability of the whole system. An explanation of this behaviour is that the system provides redundancy by requiring the two trains to fail for the system to fail. Therefore, having only one train failing because of failure in flowrate or UVI sensor does not cause a system failure. Similarly, low UVT and UVT sensor failure must occur simultaneously to generate a low dose in the system.

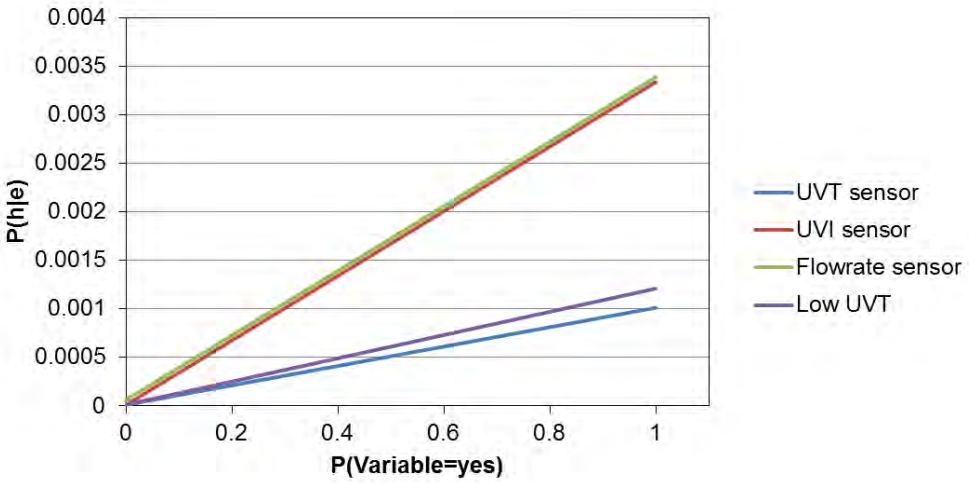

Figure 8-7: Sensitivity of *Low system dose not detected?* to parameters for sensors and *Low UVT*. No evidence was incorporated.

Two pairs of variable parameters were analysed for their influence on the target variable ([*Low system dose not detected=yes*]). The two pairs of parameters included [*UVT low=yes*] and [*UVT giving too high values=yes*], and [*flowrate sensor T1 giving too low values=yes*] and [*flowrate sensor T2 giving too low values=yes*] (Figure 8-8). The results reveal that any particular variable by its own does not generate a high impact on the target variable. However, a simultaneous change on the parameters of both variables produces a substantial impact on the probability of the target variable. Similar behaviour was observed for both pairs of variables. This latter outcome is important to consider, for example when using the same types of sensors on each train sharing common manufacturing or installation defects or limitations.

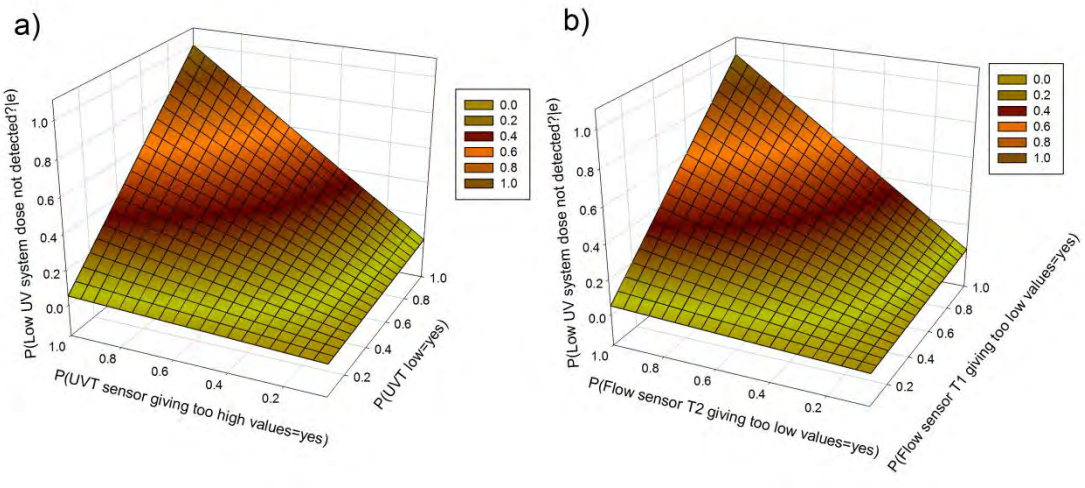

Figure 8-8: Sensitivity of *Low system dose not detected?* to pair of parameters a) UVT low and UVT sensor giving too high values and b) Flow sensors T1 and T2 giving too low values. No evidence was incorporated.

8.3.5 Evaluating scenarios

The BN model was used to generate and evaluate scenarios. An interesting scenario consists of a [*Low UV system dose not detected=yes*] and a [*Low UVT upstream affecting dose=no*] under optimum conditions of maintenance (assuming there is prior probability of having optimum or delayed maintenance of 0.5 for each one). According to the model, given the evidence, the probability of this scenario is 6E-06 (0.006%) for the model with OR gates and 8E-06 (0.0008%) for the model with noisy-OR gates. The probabilities indicate that this is a very unlikely scenario.

The posterior probabilities, *lift* and *difference* for the query variables in the previous scenario are presented in Table 8-8 for the model using the OR gate. As observed, the outcomes indicate that the probability of the state=yes had an increase between 10.6 and 12.9 approximately. Bayes factor (or ratio of the lift between two query variables) for [*Flowrate sensor giving too low values?*] and [*UVI sensor giving too high values?*] was 3.6 (i.e. 37.8/10.6), suggesting that the observations provide more support for [*Flowrate sensor giving too low values?*] as the most likely cause.

Table 8-8: Probability, *lift* and *difference* of nodes given evidence [*Low system UV dose not detected=yes*] and [*Low UVT upstream affecting dose=no*]

Variable	Probability (state=yes)	*Lift* (state=yes)	*Difference* (state=yes)
Low system UV dose not detected?	1	86.8	9.8E-01
Low UVT upstream affecting dose?	0	0	-1.8E-05
UVT sensor giving too high values?	0.0012	0.067	-1.7E-02
Low UVT?	0.001	0.999	-1.2E-06
Train low dose not detected?	1	12.9	9.2E-01
Reactor low dose not detected?	0.6	12.9	5.6E-01
UVI sensor giving too high values?	0.4	10.6	3.6E-01
Low UVI not detected?	0.4	10.6	3.6E-01
Flowrate sensor giving too low values?	0.2	37.8	1.9E-01
Too high flowrate not detected?	0.2	37.8	1.9E-01

The number of UV trains in parallel defines the level of reliability by providing redundancy to the system. By increasing the number of parallel trains, the probability of [*Low UV dose not detected*] decreased as depicted in Figure 8-9. The results revealed that three or more trains provided equivalent reliability values for the system, without a significant improvement when incorporating 4 or more trains. These outcomes can assist in designing optimum systems to balance reliability with cost from over-engineering.

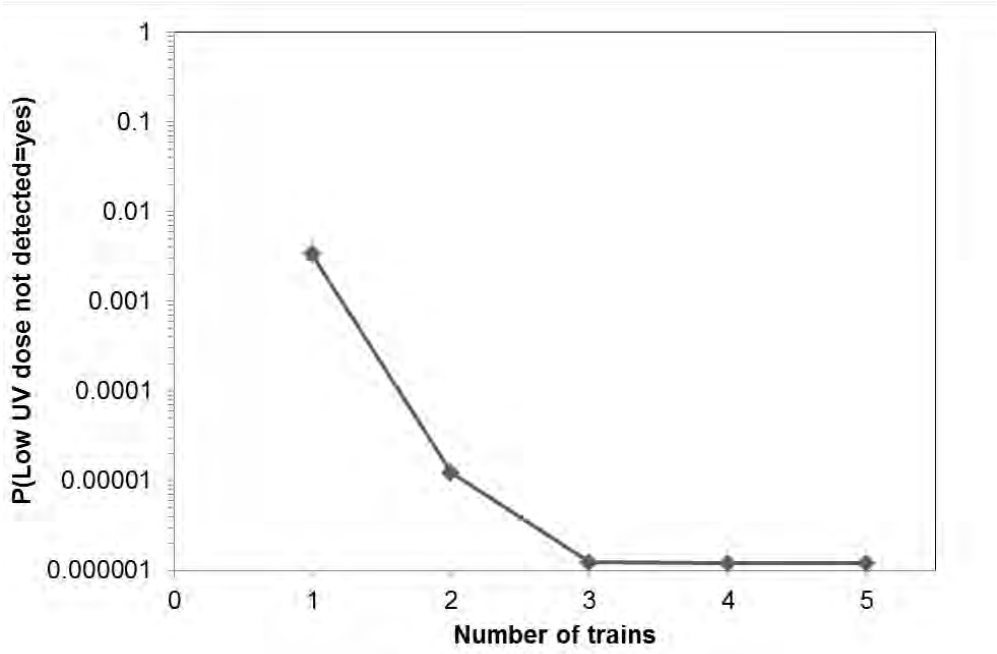

Figure 8-9: Probability of low UV dose not detected=yes vs the number of trains

A limitation of this scheme is that formal validation for this model would require long term failure data to test the accuracy of the model predictions. In fact, any meaningful failure statistics would require data collected from several plants. Due to the limited real data available for this system, the sensitivity and scenario analyses were deemed as partial validation procedures for testing the correct behaviour of the model. The model presented in this chapter can be used as a base to evaluate the reliability of similar UV systems. Collection of site-specific data would be required to update the parameters of the network.

A more sophisticated method could incorporate the whole probability distribution and simulate the results many times through Monte Carlo simulation. This methodology would require programming, for example such as which could be undertaken through the Application Programming Interface (API) available for Bayes Server 7.18 (Bayes Server, 2017).

8.4 Conclusions

Bayesian networks were used to model the event of having an undetected Low UV system dose. The model was based on a fault tree analysis and then mapped into a BN. The model structure was based on literature and the parameters obtained by expert elicitation. The BN targeted UVI, UVT and flowrate sensors as the crucial components of the UV system, determining the reliability of the system by informing the system or operators about malfunctions. From this illustrative analysis, the modelled system reliability was high with approximately 5-6 min of

undetected low dose per year. From the sensitivity analysis, UVI sensor reliabilities were the identified as the most important variables, followed by flowrate sensor reliabilities and UVT sensor reliability and UVT variability. When using two trains, special account should be taken for failures occurring to sensors close in time on both trains as a consequence of common manufacturing or installation defects or limitations. Three trains in parallel were found to provide effectively equivalent reliability to four or more trains in parallel. Such insight provides valuable information for system design optimisation. BNs facilitated scenario analysis and the assessment of potential system variations to improve reliability. Through the use of noisy-OR gates it was possible to relax the assumptions of the OR logic gate used in FTA, potentially incorporating more realistic representation of system behaviour.

Chapter 9: Improving the quantification multi-barrier system Log Reduction Values using Bayesian networks

9.1 Introduction

Advanced water treatment plants, used to control water quality risks in water reuse systems, are composed of multiple treatment barriers. Each of these treatment barriers can be regarded as having a distinct level of performance in terms of the contribution it makes to chemical or pathogen removal or inactivation. Meaningful estimation of the magnitude and variability of pathogen removal efficiencies is required to assess exposure of these hazards to people, by for example quantitative microbial risk assessment (QMRA).

Some authors have dealt with QMRA calculations through the use of Monte Carlo simulations (Chaudhry *et al.*, 2017; Pecson *et al.*, 2017; Soller *et al.*, 2017) or Bayesian networks (BNs) (Beaudequin *et al.*, 2015b). When applied effectively, the main advantages of BNs over Monte Carlo simulations are the interactive graphical representations and inference capabilities (Smid *et al.*, 2010). However, previous researchers have shown that the resolution achieved by BNs may not be sufficiently high to represent full probability density functions, resulting in a loss of information (Beaudequin *et al.*, 2015a; Beaudequin *et al.*, 2017). High resolution is particularly relevant when very low concentrations and PDF tail probabilities are required to be calculated. However, these problems can potentially be solved through the use of non-parametric continuous/discrete BNs (Cooke *et al.*, 2007).

The main aim of this study was to stochastically model a multiple barrier advanced water treatment system for the reduction (i.e. LRV) of three common index pathogens (i.e. Norovirus, *Salmonella spp.*, and *Cryptosporidium spp)* using non-parametric BNs. The model incorporated the reduction performances and the change of the microbial concentrations across the system. Dose-response and ingested volumes were not included in the model for simplicity. Consequently, the model provides a probabilistic assessment of LRV performance variability across the multiple barrier system, which may then serve and an input to QMRA calculations. The BN model was constructed based on previously published LRV performance data (Chaudhry *et al.*, 2017). A number of scenarios were tested to evaluate the impact of common factors affecting all the barriers simultaneously. This common factor could simulate the effect of high level organisational aspects affecting system performance at all levels. The BN was used to evaluate the influence and importance of each barrier LRV on the final effluent concentration. The model permitted to use the results from validation studies to estimate LRV and estimate the reliability of the system by computing the final effluent pathogen concentration.

A secondary aim of this study was to evaluate the use of non-parametric BNs for modelling pathogen reduction based on empirical models relating LRVs to operational parameters. The

model evaluated the use of chlorine disinfection model based on pH, temperature and free chlorine residual on *Giardia* LRV.

9.2 Methods

9.2.1 Model development

Models were constructed and evaluated using the commercially available BN software, Uninet (version 2.95.47) (Cooke *et al.*, 2007). This software package has the useful feature of supporting continuous, discrete and function nodes using non-parametric conditional correlations (i.e. conditional rank correlations) to capture dependencies between the variables. The joint probability distribution is constructed through a normal copula to define the dependence relations between the nodes. A copula is a mathematical structure used to generate multivariate distributions using any parametric marginal probability distribution and their dependence through a joint distribution (i.e. the copula) (Joe, 2014). A more detailed explanation of the functioning of Uninet can found in Chapter 3.

Continuous variables were used to model concentrations and performance of the barriers. The multiple barrier system for the water treatment processes consisted of six barriers including biological treatment, microfiltration, reverse osmosis, advanced oxidation and chlorination (Figure 9-1). Concentration reduction along the system is a function of the performance of each barrier and the concentration from the previous treatment step. The model, was designed so there was also a common factor connecting the barrier performances. This common cause node can be thought of as impacts associated with high level organisational aspects, for example, maintenance management, power supply, procedures, and training (Reason, 2016) that can affect the system performance at all levels. This common factor node also makes each barrier performance conditionally independent. Therefore, even if nothing is known about the common factor, having information from any one barrier's performance provides relevant information about the others. This behaviour occurs provided the common factor has a correlation factor greater than zero.

The use of BNs for modelling Giardia chlorination inactivation performance from operational parameters was also investigated. The BN incorporated an empirical model with dependent variable CT (i.e. residual free chlorine concentration times retention time) and independent variables pH, temperature and residual free chlorine concentration (Smith *et al.*, 1995). LRV calculation was performed through the inactivation ratio method (USEPA, 2003). This method computes the inactivation through the ratio of the measured CT and the required CT. Measured parameters (i.e. pH, temperature and residual free chlorine concentration and hydraulic retention

time), obtained from real operational data in a drinking water treatment plant in the USA, were fit to parametric probability distributions (Walker *et al.*, 2016).

9.2.2 Definition of barrier performance

Performance data for the barriers (Table 9-1) (secondary treatment, microfiltration, reverse osmosis, UV disinfection and chlorination) and influent concentrations for three pathogens (Norovirus, *Salmonella spp.* and *Cryptosporidium spp.*) were collected from a number of previously published articles (Chaudhry *et al.*, 2017). The three pathogen removals were modelled using the approach presented herein. The data used for the removal performance are presented as $\log_{10}$ reduction values (LRVs) for Norovirus (Table 9-1), *Salmonella spp.* (Table 9-2), and *Cryptosporidium spp.* (Table 9-3), while the influent concentrations are presented in Table 9-4. Influent concentrations were modelled through lognormal distributions.

Table 9-1: LRV distributions used in the multi-barrier model for Norovirus

Process	Norovirus	Source reference
Secondary treatment	normal(2.1,0.78)	Lodder and de Roda Husman (2005)
Microfiltration	normal(0.6,0.1)	Matsushita *et al.* (2013)
Reverse osmosis	normal(4.3,0.34)	Governal and Gerba (1999)
UV disinfection	normal(4.96,0.85)	Sherchan *et al.* (2014)
Chlorination	normal(1.68,0.24)	Francy *et al.* (2012)

Table 9-2: LRV distributions used in the multi-barrier model for *Salmonella spp.*

Process	*Salmonella spp.*	Source Reference
Secondary treatment	normal(3.32,0.76)	Ottoson *et al.* (2006)
Microfiltration	normal(5.96,1.47)	Hong *et al.* (2001)
Reverse osmosis	normal(6,0.6)	Gerba *et al.* (1997)
UV disinfection	normal(3.82,0.34)	Francy *et al.* (2012)
Chlorination	normal(2.57,0.35)	Francy *et al.* (2012)

Table 9-3: LRV distributions used in the multi-barrier model for *Cryptosporidium spp.*

Process	*Cryptosporidium spp.*	Source Reference
Secondary treatment	normal(1.58,1.3)	Ottoson *et al.* (2006)
Microfiltration	normal(4.6,0.96)	Hong *et al.* (2001)
Reverse osmosis	normal(4.5,0.73)	Adham *et al.* (1998)
UV disinfection	normal(2.2,1.17)	Craik *et al.* (2001)
Chlorination	normal(0.41,0.4)	Rose *et al.* (1996)

Table 9-4: Influent concentrations used in the multi-barrier model for the three pathogens

Pathogen	Units	Parameters of lognormal	Reference
Norovirus	gene copies/L	μ=9.095, σ=1.413E-3	Eftim *et al.* (2017)
Salmonella spp.	number/L	μ=7.171, σ=2.985	Koivunen *et al.* (2003); Lemarchand and Lebaron (2003)
Cryptosporidium spp.	number/L	μ=2.262, σ=0.944	Rose *et al.* (1996)

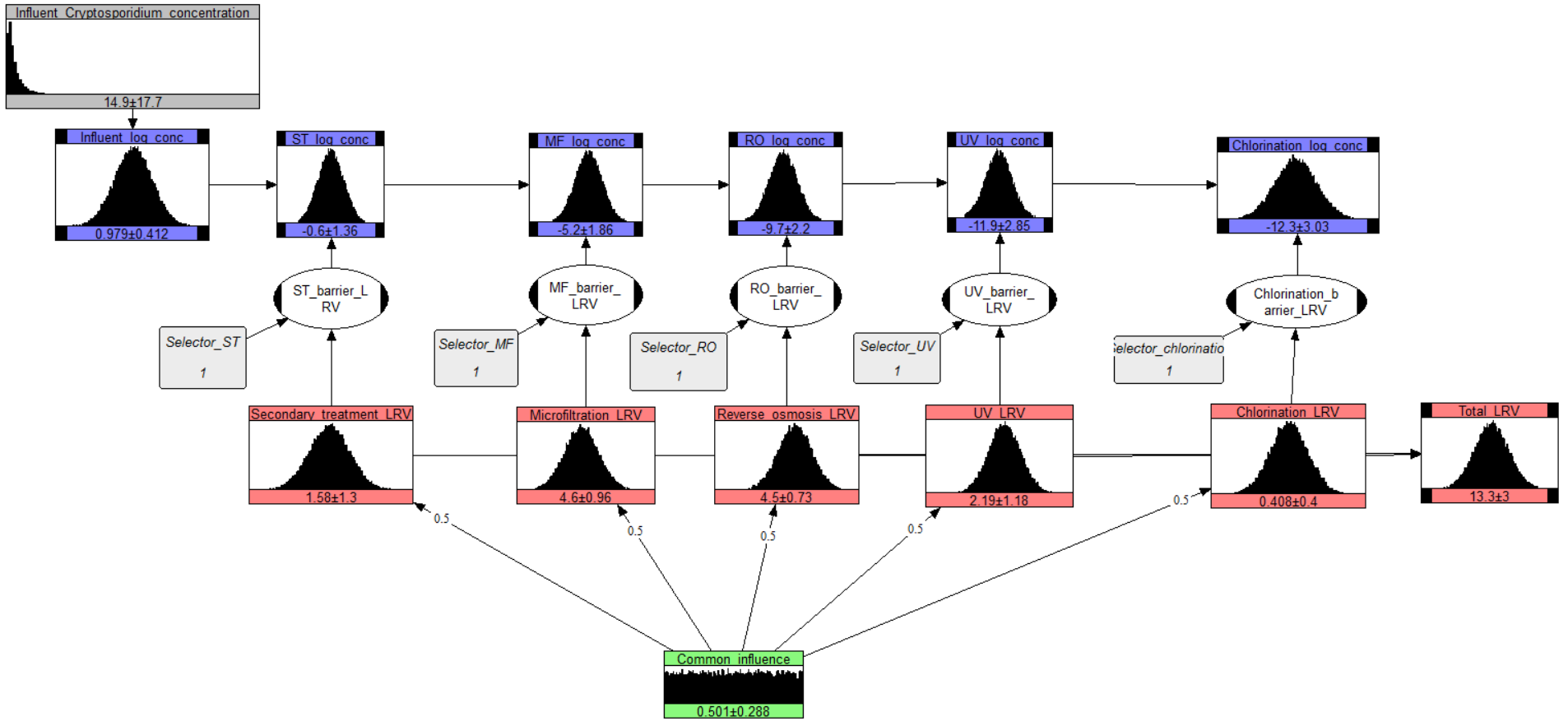

Figure 9-1: BN for the water reuse system using *Cryptosporidium spp.* to illustrate model structure. Blue nodes represent $\log_{10}$ pathogen concentrations, red nodes are the barriers pathogen removal performance, and the green node represent the common factor node. Equivalent models were generated for Norovirus and *Salmonella spp.*.

9.2.3 Model evaluation

The influence from the variability of each barrier LRV was measured by its rank correlation with the final $\log_{10}$ effluent concentration. The impact of each barrier LRV magnitude was measured by removing a specific barrier LRV at a time and calculating the difference of the final $\log_{10}$ effluent concentration with and without the barrier LRV. To simulate the absence of a barrier LRV, the model incorporated a selector. The selector is a discrete variable used to "switch off" a specified barrier in the case that credits for that barrier are assumed to have been lost. The selector could also be used to simulate failures of individual barriers.

9.3 Results and discussion

9.3.1 Impact of barrier variability

The sensitivity of the effluent pathogen concentrations to each barrier performance and the common influence node was measured through rank correlations (Table 9-5). Three different correlations between the common influence and the barrier performances were tested (i.e. 0, 0.5, 1) to analyse their effect on the sensitivity. Considering uncorrelated barriers, (i.e. correlation zero row) the barriers presented dissimilar importance depending on the pathogen. In general, the importance of each node depends on its assumed variability. Furthermore, the importance is relative so it also depends on the importance of the other variables. Therefore, a node with the highest variability will have the largest importance. For Norovirus, the $\log_{10}$ influent concentration and microfiltration presented lower variability than the other variables, returning substantially lower correlation coefficients. For *Salmonella spp.*, the effect of the $\log_{10}$ influent concentration was greater than the effect of other barriers except microfiltration. The two least influencing variables were UV disinfection and chlorination. For *Cryptosporidium spp.*, the greatest impact was observed for secondary treatment, while the least impact was observed for chlorination and $\log_{10}$ influent concentration.

Table 9-5: Rank correlation between variables in the model and pathogen $\log_{10}$ concentrations

Pathogen	RC	*CI*	ILC	ST	MF	RO	UV	Cl
Norovirus	0	0.00	-0.01	-0.61	-0.09	-0.26	-0.67	-0.18
	0.5	-0.74	-0.01	-0.73	-0.42	-0.52	-0.76	-0.48
	1	-0.99	-0.01	-0.99	-0.99	-0.99	-0.99	-0.99
Salmonella spp.	0	0.00	0.56	-0.32	-0.64	-0.25	-0.14	-0.15
	0.5	-0.65	0.45	-0.53	-0.73	-0.49	-0.42	-0.42
	1	-0.93	0.33	-0.93	-0.93	-0.93	-0.93	-0.93
Cryptosporidium spp.	0	0.00	0.17	-0.57	-0.43	-0.31	-0.51	-0.17
	0.5	-0.77	0.12	-0.70	-0.62	-0.56	-0.67	-0.48
	1	-0.99	0.08	-0.99	-0.99	-0.99	-0.99	-0.99

RC: rank correlation between common influence and barriers; *CI*: common influence; ILC: influent $\log_{10}$ concentration; ST: secondary treatment effluent $\log_{10}$ concentration; MF: microfiltration effluent $\log_{10}$ concentration; RO: reverse osmosis effluent $\log_{10}$ concentration; UV: UV disinfection effluent $\log_{10}$ concentration; Cl: chlorine disinfection effluent $\log_{10}$ concentration.

The increase in the rank correlation coefficient between the common influence and the barriers produced a rise in the impact of all variables, and conversely a reduction in the impact from the influent $\log_{10}$ concentration. These results indicated that, when the barriers were assumed independent, each one presented specific importance proportional to its variability. This effect gradually dissipated as higher correlations were assumed, levelling up completely at a correlation of 1. Haas *et al.* (2014) have indicated that correlations should be considered when modelling risks. They stated that the existence of correlations between variables will have an impact on the outcomes derived from such variables. Usually the impact becomes more important at the extremes of the distribution. They present copulas as a useful methodology to deal with this issue.

The effect of correlated barrier performances was evaluated using three different rank correlation values (i.e. 0, 0.5 and 1) and the results are presented in Table 9-6. The results revealed that the variability in the effluent $\log_{10}$ concentration was higher as the assumed correlation between the performances increased. These scenarios are relevant as they could increase the calculated risk during QMRA. An actual degree of correlation between barriers can be investigated from

empirical data using non-parametric BNs through a data-driven approach. In such case, it is important to collect data during the same time intervals at which the system could be affected by common cause conditions. This approach could also be used to generate the performance distributions without the necessity to assume a parametric distribution.

Table 9-6: Summary statistics for the effluent $\log_{10}$ concentration for variable correlations between the common influence and barriers

Pathogen	RC	5th percentile	Median	95th percentile	Standard deviation
Norovirus	0	-11.7	-9.68	-7.65	1.23
	0.5	-12.3	-9.68	-7.07	1.60
	1	-13.4	-9.71	-5.86	2.31
Salmonella spp.	0	-22.2	-18.6	-14.8	2.24
	0.5	-23.1	-18.5	-14.1	2.73
	1	-24.7	-18.6	-12.4	3.74
Cryptosporidium spp.	0	-15.9	-12.3	-8.65	2.21
	0.5	-17.3	-12.3	-7.30	3.03
	1	-19.8	-12.3	-4.69	4.57

9.3.2 Common influence effect on barrier performances

The hypothetical common influence was assumed with a uniform distribution over a range of 0 to 1. This variable can represent a measured factor affecting the performance of the barriers as a common cause. As presented in Table 9-6, various correlations were assumed to analyse the effect of this variable on the model performance. In this case, the variable common influence was not observed so the barriers were conditionally dependant. If this variable is observed, it affects the barriers simultaneously depending on the defined correlation. Two cases were modelled both assuming a correlation of 0.5 between the common influence and the barriers. The first case resulted from the assumption of a low value in the common influence, 0.1. This configuration can be conceptualised as a low level of performance in "operational management" of the plant. As observed in Figure 9-2, all the barrier performances and the total LRV decrease (grey shadows represent the previous distributions). As a consequence, all the $\log_{10}$ concentrations across the

multi-barrier system were increased. The opposite effect was achieved when the common influence was assumed to be 0.9. This outcome can be thought of as a high level of performance in "operational management" of the plant, conceptually reflecting such factors as regular maintenance, personnel training, proactive response to alarms, etc. In this case, the LRVs were increased and, as a result, the $\log_{10}$ concentrations were decreased.

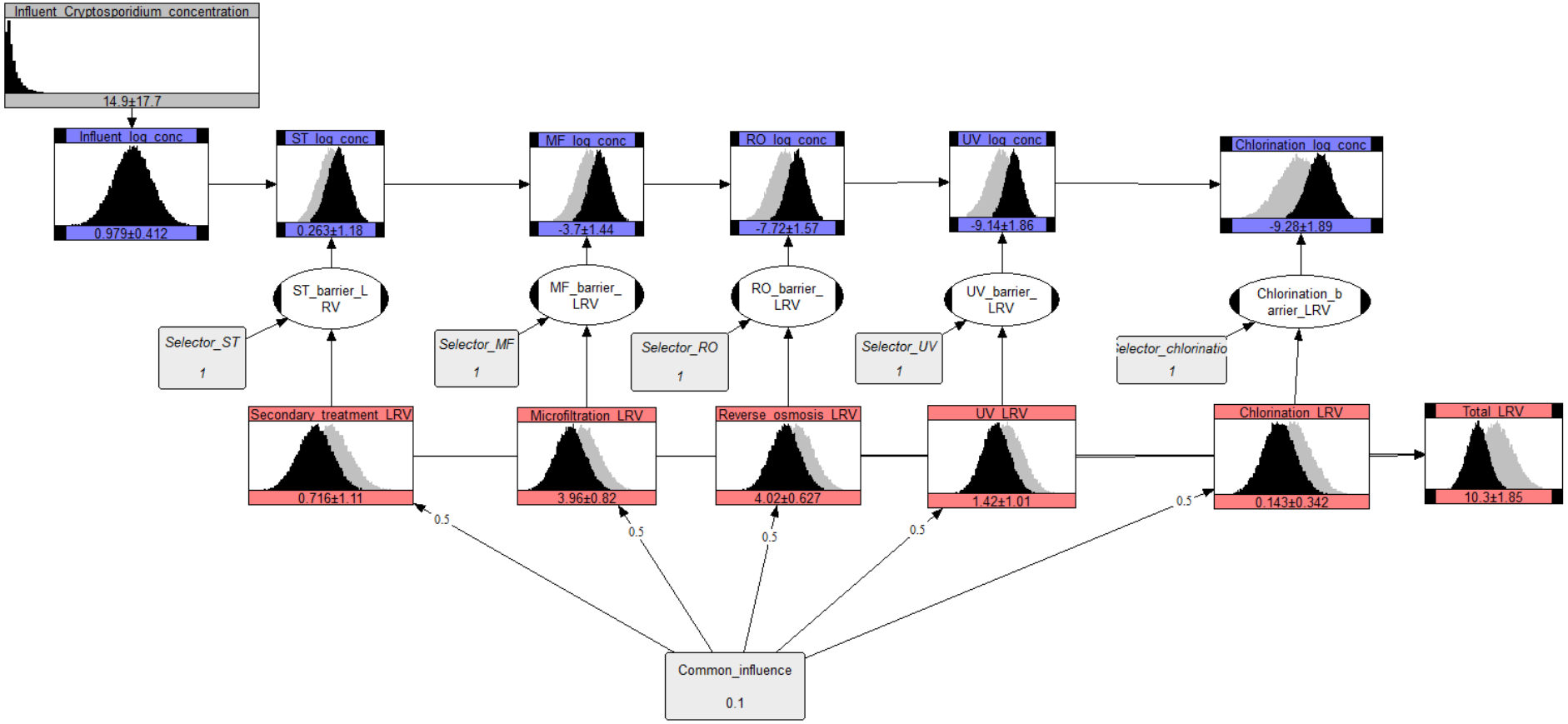

Figure 9-2: Effect of common influence over barriers, negative effect from assumption CI=0.1.

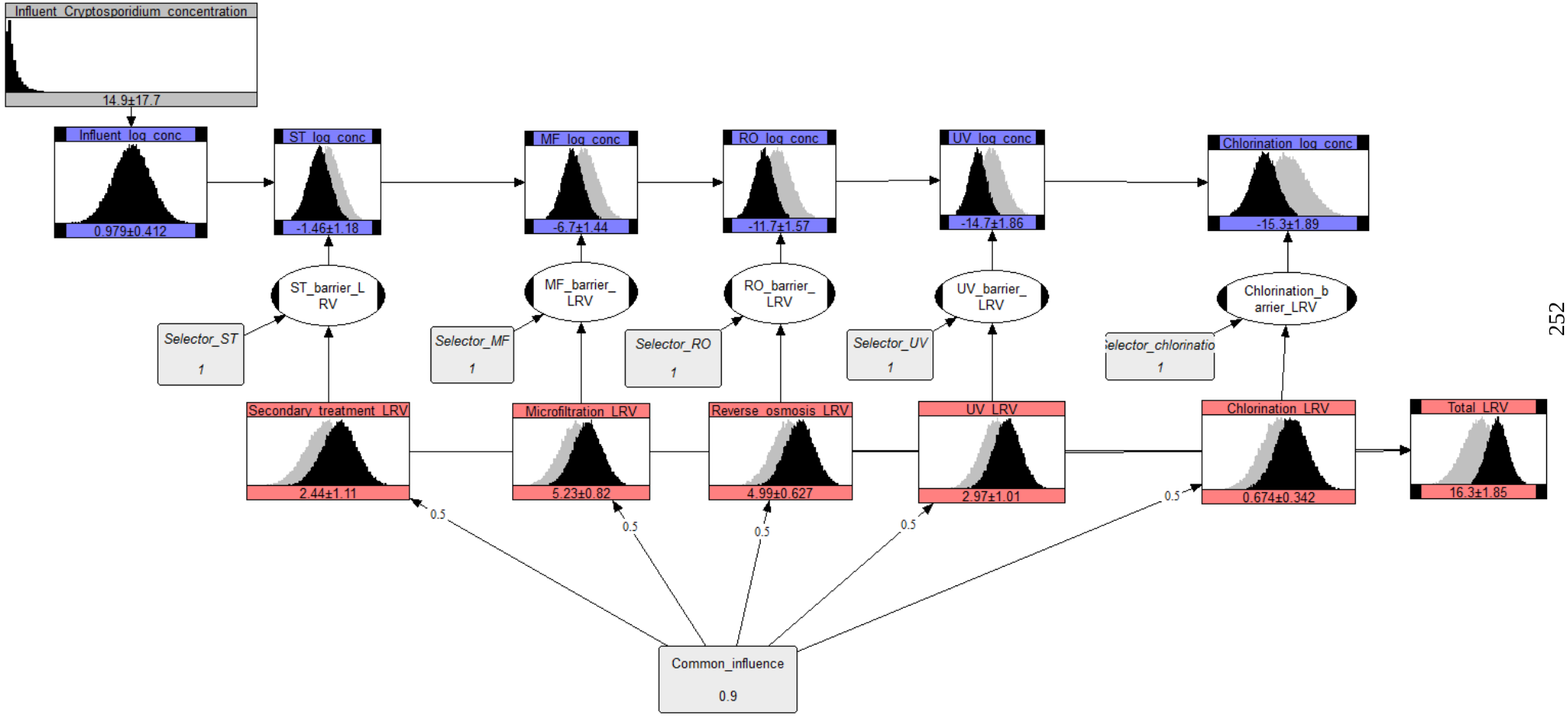

Figure 9-3: Effect of common influence over barriers, positive effect from assumption CI=0.9.

9.3.3 Impact of individual barriers on effluent concentration

The impact of each barrier on the process performance was assessed by "switching off" each barrier, one at a time, using the "selector" node and obtaining the mean and standard deviation of the effluent $\log_{10}$ concentration for Norovirus (Figure 9-4), *Salmonella* spp. (Figure 9-5), and *Cryptosporidium spp.* (Figure 9-6). The assessment employed the model without associations between the barrier performances (i.e. correlation coefficient of zero). The influence of each barrier performance was dependent upon both the magnitude and variability of the LRV assigned to it. The impact was measured as the probability of having an increase in the final plant effluent $\log_{10}$ concentration. For Norovirus, the probability of having an increase in the effluent $\log_{10}$ concentration resulting from not considering the LRVs credited to any one of secondary treatment, reverse osmosis or UV disinfection was greater than 0.95 (Table 9-7). The lowest influence was obtained for microfiltration (probability: 0.82). The variability in the log effluent concentration was not substantially impacted with standard deviations ranging from 0.9 to 1.2 in all cases.

Table 9-7: Percentage of the distribution lower than zero for the difference in log effluent concentrations from "switching off" a particular treatment step. Probabilities higher than 0.95 are presented in bold.

Removed process step	Norovirus	*Salmonella spp.*	*Cryptosporidium spp.*
Secondary treatment (ST)	**0.98**	0.86	0.71
Microfiltration (MF)	0.82	**0.98**	0.94
Reverse osmosis (RO)	**>0.99**	**0.97**	0.93
UV disinfection (UV)	**>0.99**	0.89	0.77
Chlorination (CL)	0.94	0.79	0.55

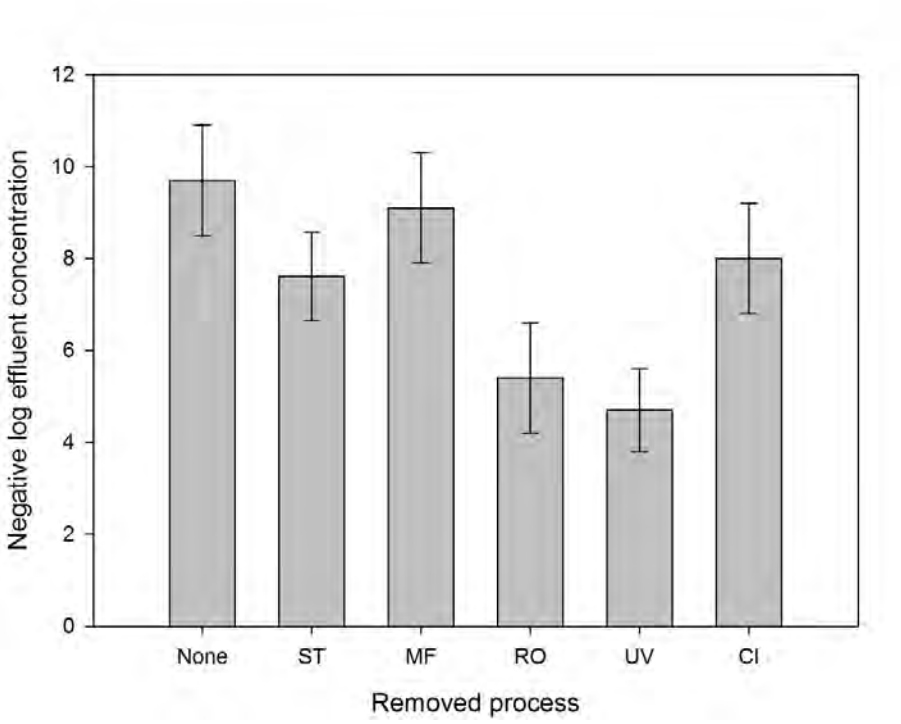

Figure 9-4: Impact of individual process removal for Norovirus. ST is secondary treatment; MF is microfiltration; RO is reverse osmosis; UV is UV disinfection, and Cl is chlorination. Error bars indicate ±standard deviation.

For *Salmonella spp.*, the probability of having an increase in the effluent $\log_{10}$ concentration was higher than 0.95 only for microfiltration and reverse osmosis (Table 9-7). The third greatest impact was obtained for UV disinfection. The least impact was observed when chlorination was not included in the treatment (probability: 0.79). The variability in the log effluent concentration ranged between 1.7 and 2.2. For *Cryptosporidium spp.*, none of the barriers produced an increase in the log effluent concentration with a probability higher than 0.95. However, as observed in Table 9-7, the greatest effects were observed for microfiltration (probability: 0.94) and reverse osmosis (probability: 0.93). By not considering chlorination, the probability of increasing the log effluent concentration was only 0.55. This result was expected as chlorine disinfection is not very effective for *Cryptosporidium spp.* inactivation. The variability in the effluent log concentration was not substantially impacted with standard deviations ranging from 1.8 to 2.2.

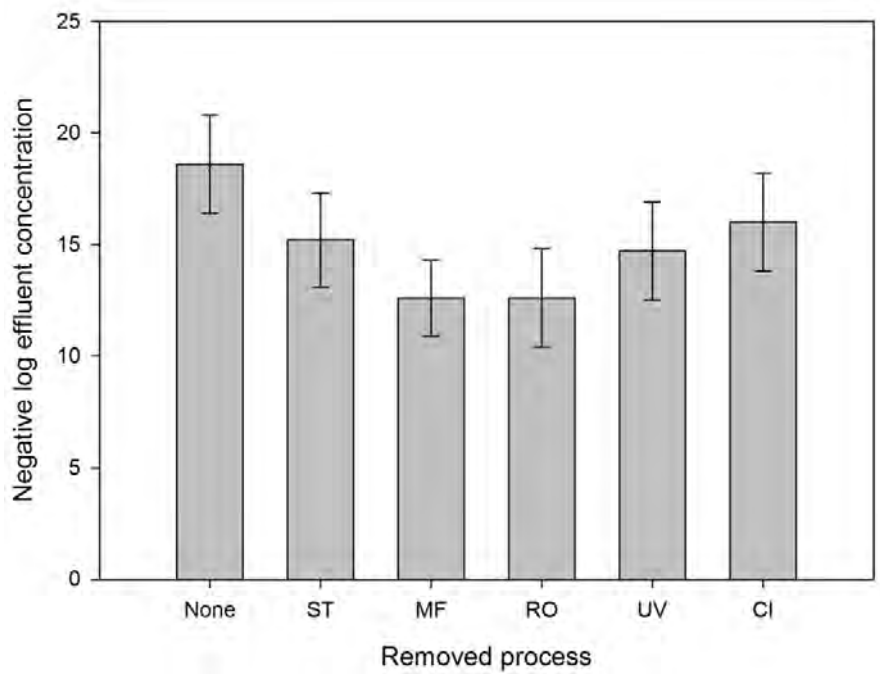

Figure 9-5: Impact of individual process removal for *Salmonella spp.*. ST is secondary treatment; MF is microfiltration; RO is reverse osmosis; UV is UV disinfection, and CL is chlorination. Error bars indicate ±standard deviation.

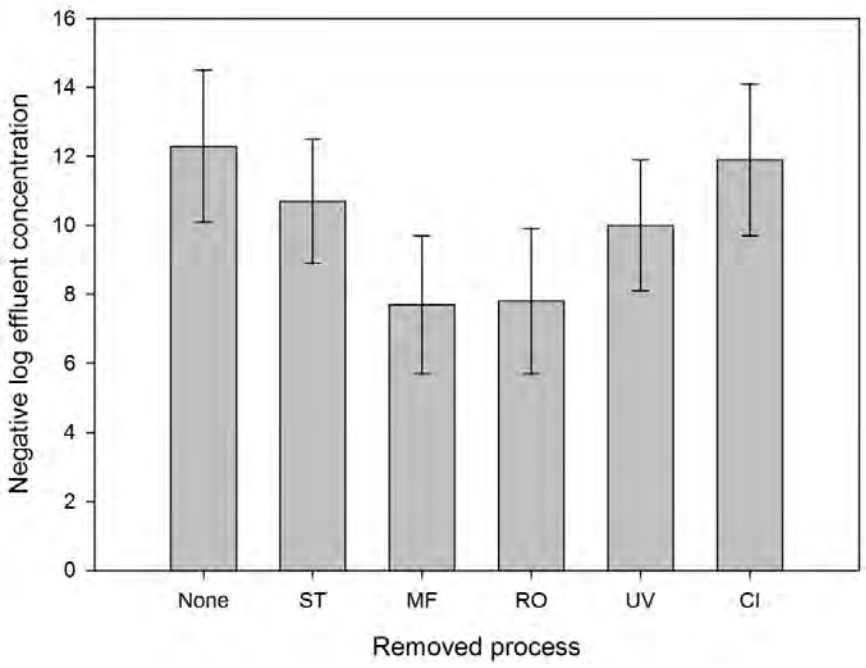

Figure 9-6: Impact of individual process removal for *Cryptosporidium spp.*. ST is secondary treatment; MF is microfiltration; RO is reverse osmosis; UV is UV disinfection, and CL is chlorination. Error bars indicate ±standard deviation.

9.3.4 Modelling pathogen reduction using operational parameters from the literature

When operational data can be related to pathogen LRVs through pre-established empirical or mechanistic relationships such as in chlorination, these relationships can be easily modelled through BNs. A model was constructed to simulate the Giardia LRVs during chlorination based on the inactivation ratio method in a drinking water treatment plant in the USA (Walker *et al.*, 2016) (Figure 9-7). Daily records between 2008 and 2009 were used to fit distributions for temperature, pH, free chlorine residual (*C*) and contact time (T_{10}) for the chlorination clearwell. From the values of *C* and T_{10}, CT values (i.e. concentration times time) were calculated (*CT_actual*). The previous study found that *C* and *CT_actual* could be best fit to lognormal distributions, while *pH* could be best fit to a normal distribution. *Temperature* and contact time (T_{10}) were fit to generalised Beta distributions (Walker *et al.*, 2016). Based on the value of *C*, *Temperature* and *pH,* a required CT to achieve 3 LRVs (*CT_required*) can be computed using an empirical model (Equation 9-1 and Equation 9-2) (Smith *et al.*, 1995). The *CT_actual* was used to calculate the LRV through the inactivation ratio (i.e. LRV=3* $CT_{actual}/CT_{required}$). Rank correlations were determined from the data between the variables T_{10} and *CT_actual* (RC: 0.9),

C and *CT_actual* (RC: 0.61), and T_{10} and *C* (-0.64). These correlation values indicate that *CT_actual* presents a higher influence from T_{10} than *C*. Also, T_{10} and *C* are inversely associated, so when there is short contact time, the residual concentration needs to be high to achieve a specific CT value.

$For\ Temp < 12.5$:

$$CT_{3-log,Giardia} = (0.353 \cdot I)(12.006 + e^{2.46-0.073 \cdot temp+0.125 \cdot C+0.389 \cdot pH}) \quad \text{Equation 9-1}$$

$For\ Temp \geq 12.5$:

$$CT_{3-log,Giardia} = (0.353 \cdot I)(-2.261 + e^{2.69-0.065 \cdot temp+0.111 \cdot C+0.361 \cdot pH}) \quad \text{Equation 9-2}$$

Where I=3 represents the required log removal of *Giardia*, C is the free chlorine residual concentration (mg/L) and temp is the temperature in °C.

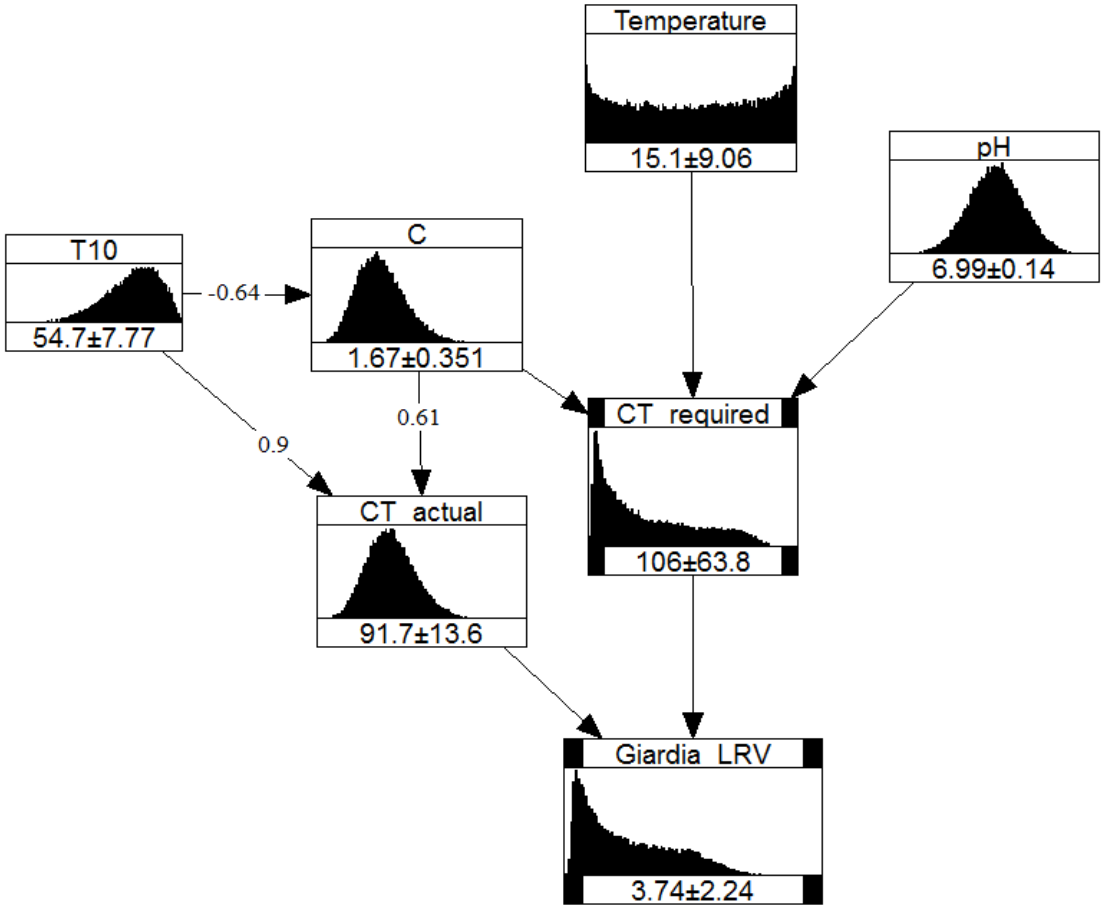

Figure 9-7: Chlorination LRV model for *Giardia*

Based on this model, various inference queries can be performed as shown in Figure 9-9 to Figure 9-12 and Table 9-9 to Table 9-12. The results indicate the effects of conditionalization on the marginal distributions of the remaining variables. Marginal distributions for the nodes without any evidence incorporated into the network can be observed in Figure 9-8. Four scenarios were studied by observing four different nodes, one for each scenario. The base case (i.e. no observations) is shown in Table 9-8 and Figure 9-8 which presents the same marginal distributions as in Figure 9-7.

Table 9-8: Median, 5th and 95th percentiles for the marginal distributions without observations

Percentile	C	T_{10}	CT_actual	Temperature	pH	CT_required	Giardia_LRV
5th	1.2	40	71	1	6.8	35	1.2
50th	1.6	56	91	15	7	86	3.1

Percentile	C	T_{10}	CT_actual	Temperature	pH	CT_required	Giardia_LRV
95th	2.3	65	115	28	7.2	229	7.9

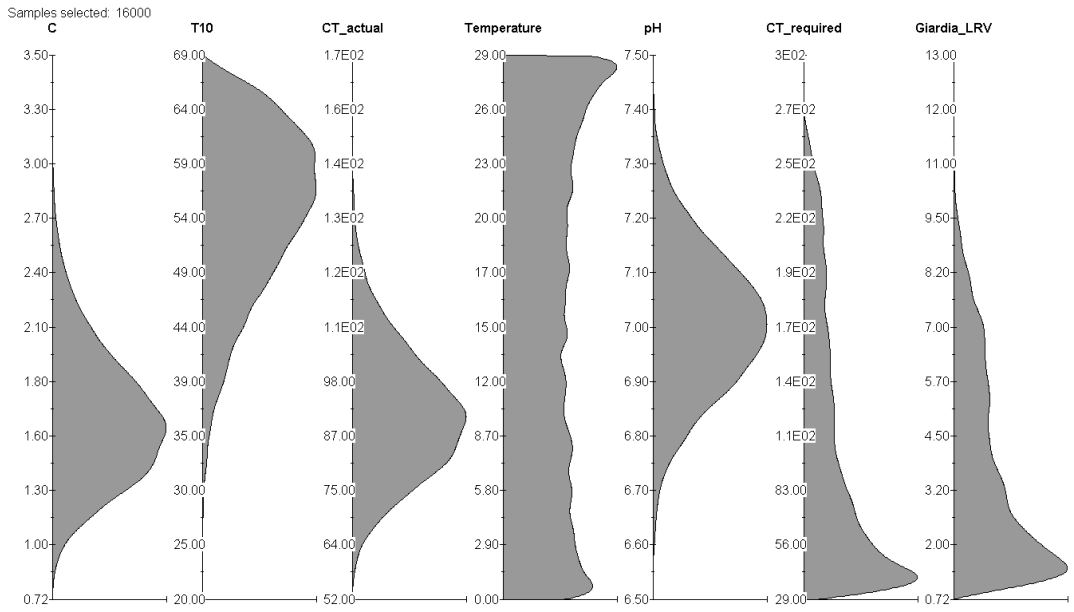

Figure 9-8: Marginal distributions without input evidence

The first scenario presents a case where the conditions leading to a LRV<2 need to be determined (Table 9-9 and Figure 9-9). The results indicated that the most impacted variables with this observation were *Temperature* and *CT_required*. Free residual chlorine (*C*), *CT_actual* and *pH* were not largely affected revealing that these variables are not very important for the LRV under these conditions.

Table 9-9: Median, 5th and 95th percentiles for the marginal distributions after observing Giardia_LRV<2

Percentile	C	T_{10}	CT_actual	Temperature	pH	CT_required	Giardia_LRV
5th	1.1	40	70	0.2	6.8	137	1
50th	1.6	55	88	3.3	7	194	1.4
95th	2.3	65	111	8.4	7.2	250	1.8

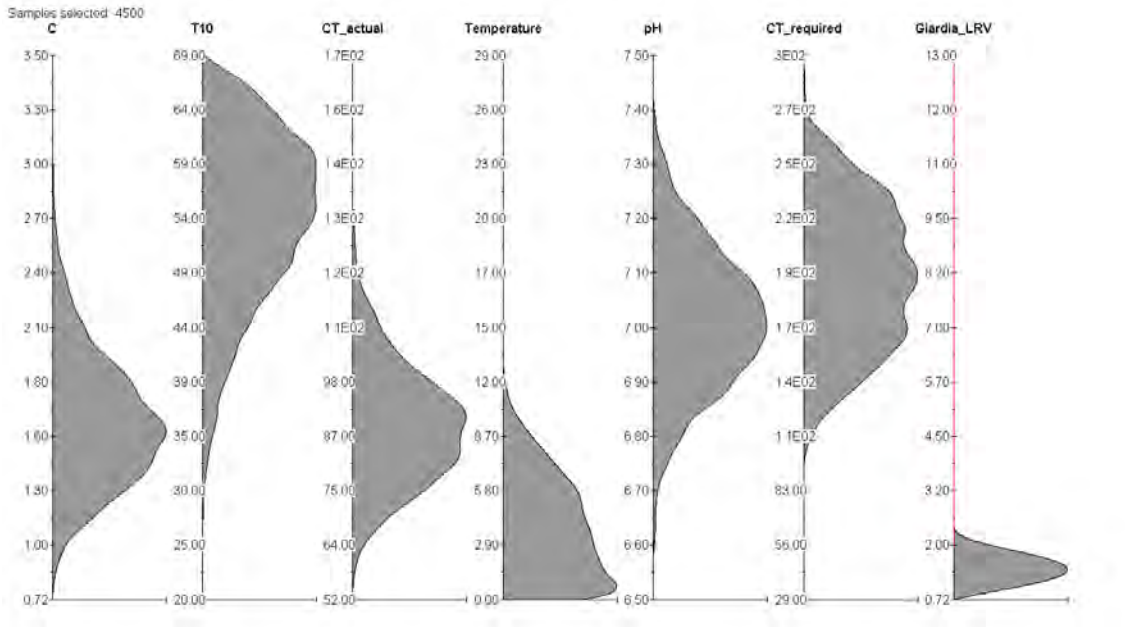

Figure 9-9: Marginal distributions with evidence LRV<2

When conditioning pH to lower than 7, only minor changes were observed for *CT_required* and *Giardia_LRV* confirming the little influence of this variable over chlorination performance in this case (Table 9-10 and Figure 9-10). The slight effect of *pH* could be a consequence of the narrow range of conditions in which the chlorination process worked in this case. It is expected that under *pH* values higher than 7.5, the effectivity of chlorination would present a higher decrease (AWWA & ASCE, 2005).

Table 9-10: Median, 5th and 95th percentiles for the marginal distributions after observing pH<7

Percentile	C	T_{10}	CT_actual	Temperature	pH	CT_required	Giardia_LRV
5th	1.2	40	71	1	6.7	34	1.2
50th	1.6	55	91	15	6.9	83	3.2
95th	2.3	65	115	28	7	216	8.3

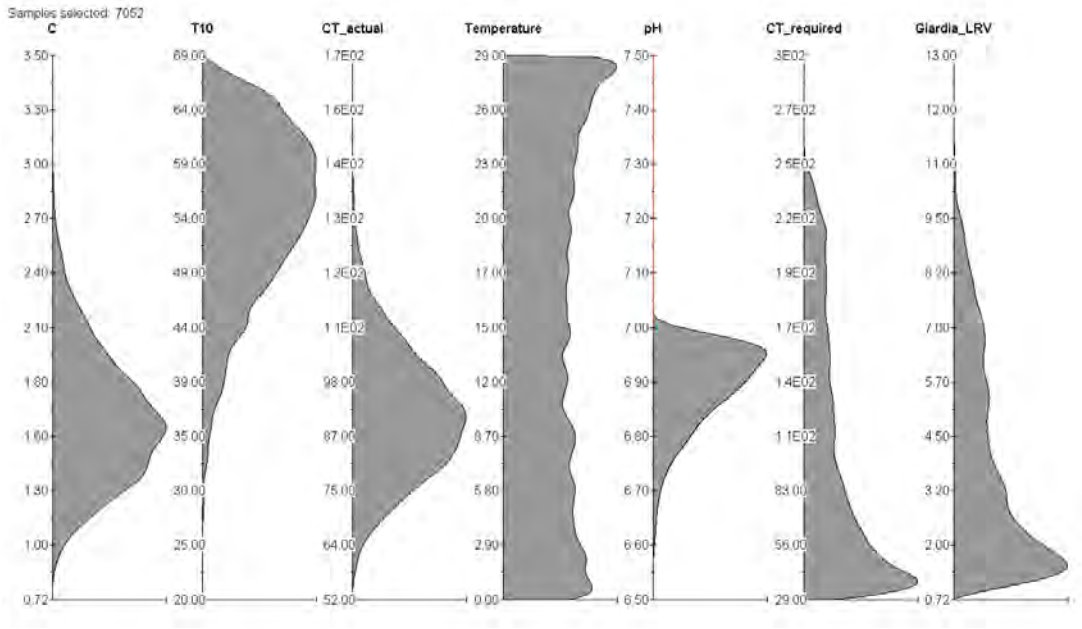

Figure 9-10: Marginal distributions with evidence pH<7

By observing Temperature higher than 15 °C, the impact on *CT_required* and *Giardia_LRV* was substantially higher than in the case of *pH* with an increase of 43% in the median value for *Giardia_LRV* (Table 9-11 and Figure 9-11). By the observed effect of pH and temperature on the acid/base dissociation of hypochlorous acid, it is generally expected that the effectiveness would be more largely affected by pH than temperature (AWWA & ASCE, 2005). However, in this case also the range of conditions under which the chlorine disinfection works would have an effect on the results of the queries. Given the possible range of temperatures on which this system can work, it becomes increasingly important to correctly control the residual chlorine concentration and contact time during cold seasons.

Table 9-11: Median, 5th and 95th percentiles for the marginal distributions after observing Temperature>15 °C

Percentile	C	T_{10}	CT_actual	Temperature	pH	CT_required	Giardia_LRV
5th	1.1	40	70	16	6.8	34	3.1
50th	1.6	55	91	23	7	51	5.4
95th	2.3	65	116	29	7.2	84	8.4

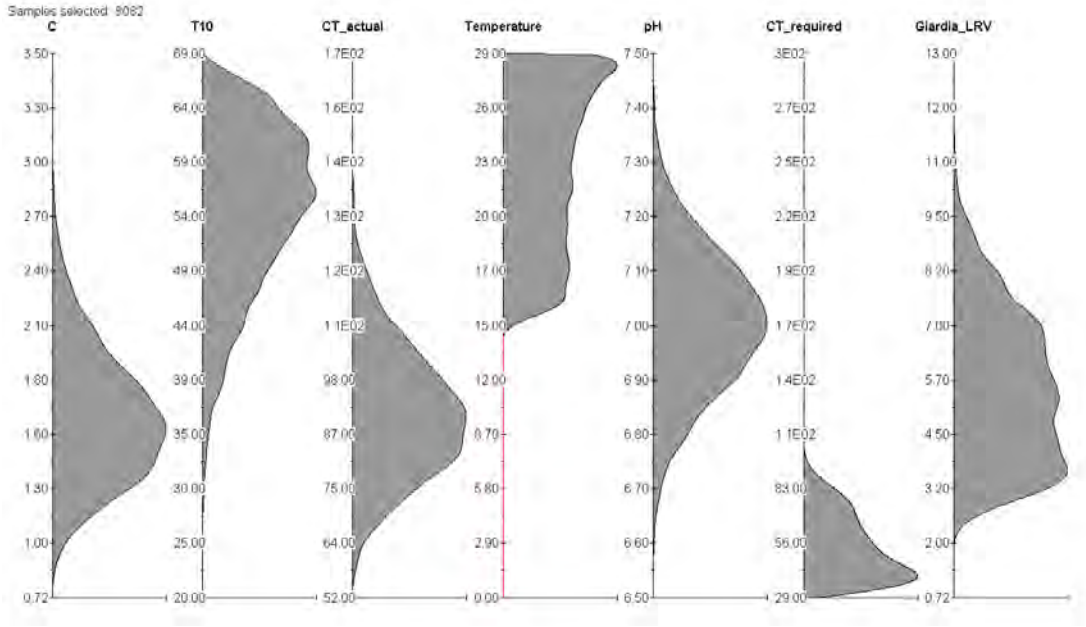

Figure 9-11: Marginal distributions with temperature>15 °C

The last scenario shows the case when the free residual chlorine (*C*) is lower than 1 (mg/L) (Table 9-12 and Figure 9-12). This observation had a large impact on *CT_actual* as a consequence of their observed rank correlations. However, the impact over *Giardia_LRV* was only minor with a decrease in the LRV of 13% approximately. As observed in Figure 9-12, the shapes of the distributions look more variable than in previous cases. The reason of this behaviour is because the model conditioning is performed through sample-based conditioning. Therefore, the shape would depend on the number of samples imposed by the condition. Fewer samples would produce the behaviour observed in Figure 9-12.

Table 9-12: Median, 5th and 95th percentiles for the marginal distributions after observing C<1

Percentile	C	T_{10}	CT_actual	Temperature	pH	CT_required	Giardia_LRV
5th	0.8	59	58	1.2	6.8	32	1.0
50th	0.9	66	71	18	7.0	79	2.7
95th	1.0	68	86	29	7.2	209	6.7

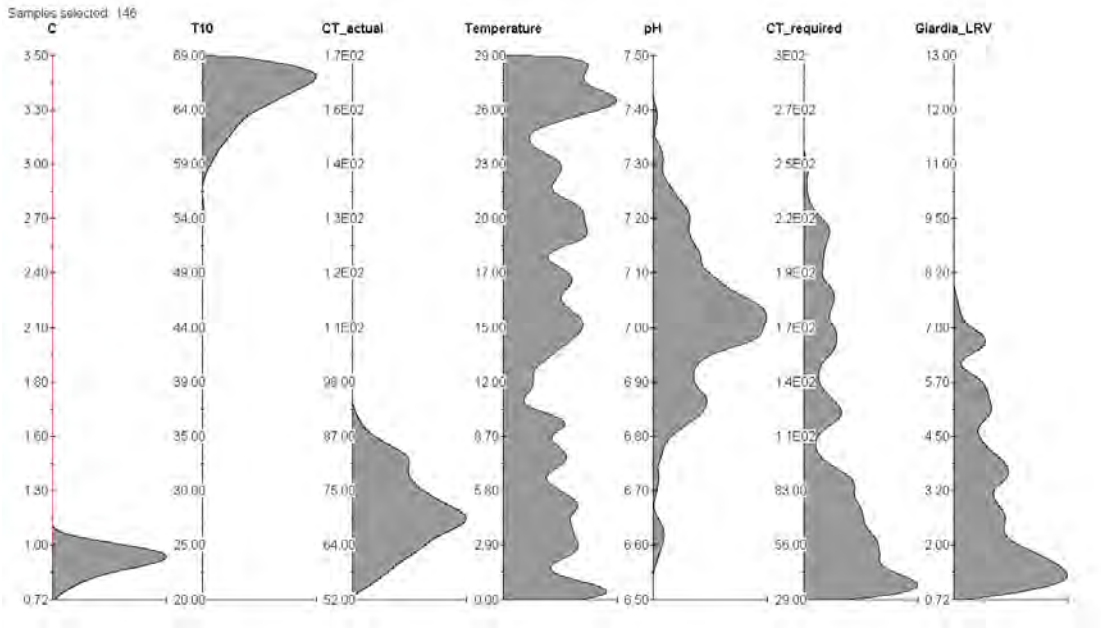

Figure 9-12: Marginal distributions with C<1 mg/L

9.3.5 Comparison with other BN approaches

Various types of BNs can be used to model pathogen reductions a multi-barrier system. The major limitation of these models is often the low resolution obtained because of the discretisation commonly used in these models. In discrete BNs, this issue has been successfully managed by using dynamic discretisation (Fenton & Neil, 2012), which optimises the thresholds every time an inference is performed. Although high resolution is obtained through this method, learning of parameters and representation of correlations is not possible. Furthermore, a large number of parameters would be required to represent the dependencies from impact factors on barriers. A continuous Bayesian network (Scutari & Denis, 2014) can also be constructed through *JAGS* (Plummer, 2003) or Bugs (Lunn *et al.*, 2000). In this case, there is no necessity of discretisation thresholds to be defined. However, the probability distributions need to be previously selected and the models do not employ correlations directly. The approach presented in this chapter has the advantage of permitting the definition of the associations between the variables through correlation coefficients. The model presented in this chapter can also be analysed through a conventional BN after selecting the discretisation thresholds while keeping the same structure as the original non-parametric BN. To capture the associations between the variables, simulated data would need to be obtained from the non-parametric BN. Subsequently these samples would be used to learn the parameters of the model through parameter learning algorithms.

9.3.6 Importance of BNs to validation and reliability

Models presented in this chapter showed the case when the whole multiple barrier system needs to be modelled to measure its reliability. Such model implies determining that specific targets are met even when one or more barriers' performances are compromised. Inputs for the model included LRVs for index pathogens and influent concentrations. These data could be provided from validation studies or estimated through empirical or mechanistic models using operational data. The developed models were capable of representing the whole system performance and measure the importance of each barrier for the final effluent quality.

9.4 Conclusions

Microbial risks have previously been modelled through a multiple barrier approach using Monte-Carlo simulations and BNs. Although both approaches have been shown to be useful in a range of circumstances, there remain limitations, such as lack of inference capabilities in the case of Monte Carlo simulations and low resolution in the case of BNs. This chapter presents new ways of using BN software to model a multi-barrier system using non-parametric continuous/discrete BNs through Uninet.

The approach was demonstrated to be suitable for performing inference and calculating probability distributions with high resolution. The BN could model multiple barriers with interactions and incorporate correlations explicitly. A sensitivity analysis to the modelled effluent $\log_{10}$ concentration with respect to each barrier was measured using correlations. The outcomes indicated that by considering correlated performances, the variability in the log effluent concentration increased. Because the results of this approach serve as inputs for risk analysis, they should be evaluated during QMRA calculations. Incorporating correlated barriers generated a higher sensitivity across all the barriers. The impact of each barrier was assessed probabilistically through this model. In this case the parameter used was the probability of obtaining an increase in the $\log_{10}$ effluent concentration when a barrier was not considered in the model. This approach was also suitable to model probabilistic LRV outcomes calculated from empirical equations using stochastic inputs. From this model a number of queries can be performed to assess scenarios and most important variables. As observed in this chapter, the graphical and stochastic representation of the system through BNs allows effective communication of the LRV performance and the different factors affecting pathogen effluent concentration used during exposure analysis.

Chapter 10: Conclusions and future work

10.1 Conclusions

The aim of the work presented in this thesis was to review and assess appropriate tools for validation and reliability assessment of water treatment processes especially those likely to be used for direct and indirect potable water reuse. The key findings of this research project in relation to the aims in Section 1.2 are presented in Table 10-1. Validation and reliability were studied from the point of view of pathogen removal because pathogens usually represent the most important hazards requiring control for safe water reuse. The literature review (Chapter 2) revealed that commonly used techniques for reliability assessment generally involved causal representation of events and probabilistic assessment of outcomes. It was found that for validation, Monte Carlo has commonly been used as this allows performing calculations with full probability distributions. Various other advanced statistical or machine learning tools were also found to have been applied for treatment process prediction during system control and monitoring. However, these methods were limited in their ability to represent causal associations between variables and in their capacity to facilitate user interaction. The outcomes of the literature review indicated that BNs and Bayesian methods were tools which could be applied to the diverse validation and reliability assessment tasks identified. Useful features of Bayesian tools included explicit representation of uncertainties, probabilistic assessment and inferential reasoning. Through their use, other useful methodologies for reliability assessment such as fault tree analysis, event tree analysis, reliability block diagrams and bow-tie analysis could be either encoded or enhanced.

Table 10-1: Aims and related key findings of this research study

Aim	Key findings
Identify, evaluate and compare different risk and reliability techniques for water reuse processes.	• Relevant tools were identified from standards and technical reliability assessment literature. • Twelve tools were compared and evaluated for their use in water treatment processes. • Machine learning tools were also discussed.

Aim	Key findings
	• BNs and Bayesian analysis were selected due to their wide range of useful features.
Identify and assess operational parameters affecting the reliability of removal of microorganisms and evaluate potential surrogate parameters.	• BNs were used to identify and evaluate relevant parameters for pathogen removal during activated sludge treatment. • Effect of operating conditions on the reliability of system was evaluated. • Bayesian analysis was used for comparing the reliability of operational parameters to predict pathogen reduction during ozonation.
Analyse the validation data of systems to improve prediction of performance reliability and develop a method applicable to treatment systems with parallel subunits.	• A full-scale ultrafiltration system with units in parallel was studied for its virus removal reliability. • Hierarchical Bayesian models showed to improve the predictions of pathogen removal for the system.
Explore improving incorporation of uncertainty into pathogen removal efficiency estimation and promote reliability assessment.	• LRVs were modelled incorporating uncertainty in models parameters during chlorination and ultrafiltration. • Outcomes allowed direct calculation of probabilistic LRVs.
Quantify the reliability of a water reuse system using expert knowledge.	• Expert knowledge was used to develop a model for reliability estimation for a UV system.

Aim	Key findings
	• The model focused on the importance of sensors for the reliability of the system.
Develop and assess a method to analyse reliability in multi-barrier water reuse systems.	• Non-parametric BNs were used to model an advanced treatment system train. • Model captured probabilistic associations between the system components. • The BN allowed measuring the importance of each treatment barrier in relation to the whole reliability of the system.

The review also showed that traditional reliability analysis has been used to assess component failures and system failures. Such failures are usually represented by a yes/no type of event. For water and wastewater treatment processes, though, failures are generally observed as located on a continuum. In these systems, degraded states would be generally perceived by their effect on system performance. Inherent reliability (or performance) and mechanical reliability appear closely related. Validation to investigate system performance is therefore an important task to determine its reliability.

My research then studied the application of BNs to reliability and validation assessment in water and wastewater treatment processes. BNs on their own were found to be somewhat limited for the full range of applications considered. Although BNs have special characteristics that make them unique, they share multiple features with Bayesian statistical methods more generally. These features include inferential reasoning through the Bayes rule, explicit representation of uncertainties, and capability of handling censored and missing values. Consequently, it was decided appropriate to broaden the analysis to include Bayesian models in general, comprising both BNs and Bayesian statistical methods.

Through the development and assessment of six case studies, BNs and Bayesian methods were evaluated and interrogated using a combination of experimentally-derived data, full scale operational data, and expert elicited data. These six cases covered a range of applications including evaluation of monitoring parameters for LRVs, improving LRV calculations, assessment of system reliability and multi-barrier system evaluation. Because these were selected and designed to comprise common tasks encountered during risk analysis in water reuse systems, they may now serve as examples to facilitate application to full-scale real systems. The outcomes of this work will provide valuable alternatives for improving and informing risk assessment in a robust and transparent manner. The work undertaken produced conclusions, described below, relating to validation and reliability assessment.

10.1.1 Bayesian models for validation and inherent system reliability assessment

Reliability can be analysed from various perspectives depending on the specific purpose. From the point of view of LRV performance, variability or inherent reliability is important because it has an impact on the mitigation of risks. Validation is a relevant step that checks that the system is performing appropriately and provides information about how the system could be monitored. Among the most important aspects to consider during treatment process performance validation is the exploration and definition of adequate parameters to monitor LRV performance. BNs were first used to identify and validate potential predictors of pathogen LRV during activated sludge treatment. Next, alternative predictors of LRVs during ozonation were assessed, but in this case using Bayesian analysis. The activated sludge investigation showed that BNs were adequate to predict LRV, assess scenarios and potential predictive parameters. The activated sludge study (Chapter 4) reflects the case where large datasets with many variables are obtained or available, but there is uncertainty about the relative value of the variables monitoring with different objectives (e.g. prediction of LRVs). This case is a common situation when large operational datasets are collected over several years and it is subsequently desired to find associations between the variables to improve system oversight. An example of this situation is where operators wish to fine tune a system based on long term experience and have collected extensive supervisory control and data acquisition (SCADA) data. For the Chapter 4 study BN models were constructed and validated using data collected during experiments on a well-controlled pilot scale reactor. The BN approach proved suitable for a dataset containing missing and censored values which are commonly encountered in real applications. It was found that suspended solids, turbidity and SRT were important operational predictors for *C. parvum* LRVs. *G. lamblia* removal was also modelled through BNs. However, the analysis showed its reduction could not be predicted using the parameters monitored with high reliability in contrast to *C. parvum*. Low predictive potential predictors for this pathogen included SRT and COD, but interestingly not

microbial indicators or *Cryptosporidium spp.* and not optimal reactor operating configuration was identified.

In the third case study, ozonation (Chapter 6), the potential predictors considered included ozone to dissolved organic carbon ratio, difference in UV_{254nm}, difference in total fluorescence and difference in components concentrations from PARAFAC analysis. These were selected based on previous literature. The study used measurements of indigenous indicator microorganisms which have been previously shown to be more resistant than seeded ones during disinfection. The study sought to validate previous findings and also assess and compare the uncertainty of previous and current results. Uncertainty was estimated using linear and non-linear models through Bayesian analysis. Although all the analysed predictors were shown to be useful for prediction of LRVs during ozonation, the difference in UV_{254nm} was most promising because this is a commonly measured parameter and more affordable to collect compared to most of the others. The analysis took into account that all predictors showed equivalent predictive potential. Censored values were also encountered in this study. Bayesian analysis was again able to deal with this type of data and incorporate this information into the model parameter estimates.

Incorporation of uncertainty in model outcomes is important to consider during risk analysis because the full probability distribution could be used to calculate exceedance from specific thresholds. Bayesian analysis proved adaptable for this task because it can estimate uncertainties on model parameters. This model application was demonstrated by the fourth case study on chlorination and chloramination of secondary treated sewage (Chapter 7). In this case study, experimentally derived data were used to create empirical models using Bayesian analysis. The model outputs were a distribution of values instead of a point estimate as is commonly encountered in other modelling approaches (e.g. mechanistic modelling using inactivation rates). In this case, the model was used to estimate required CT values to achieve a specified level of pathogen inactivation under specified ambient conditions (turbidity and pH). These required CT values were then used to calculate performance LRVs from the measured CTs. Further improvements to the conceptual model could incorporate the effect of other water matrix characteristics including temperature, dissolved organic matter and ionic strength.

Case study 2 (Chapter 5) looked at ultrafiltration. UF units are commonly challenge tested through seeded bacteriophages to estimate LRVs. Because of their modular design and the fact that planning and conducting full-scale experiments usually implies large financial costs, it is important to make efficient use of the data for performance estimation (i.e. reliability). Bayesian analysis was used in this case to estimate predictive LRVs from a number of UF units in parallel to obtain system LRVs which combined all data. A number of model variations were evaluated

with a real UF validation dataset. The evaluation indicated that of the four variants assessed, hierarchical Bayesian models were the most appropriate for LRV estimation of the whole system. The models were also useful for improving the LRV distribution estimates of each individual unit where the numbers of measurements were necessarily limited as well as the overall LRV of the system for prediction purposes. For future applications, the type of hierarchical Bayesian model selected will depend on the particular system and conditions, therefore the analysis for a system was seen as a model for how treatment systems with multiple parallel modules could be validated and revalidated. This study did not attempt to be prescriptive, but to show how it can assist in the data analysis during challenge testing. The data evaluated in this study were obtained from real full-scale validation and revalidation work and had to deal with the limitations typically encountered in such operational monitoring campaigns, including censored values, limited observations and differences in performance between units in parallel.

Quantitative microbial risk assessment requires the estimation of multiple barrier performance and concentrations of reference pathogens to estimate exposure assessment. This task is increasingly conducted in a probabilistic format through Monte Carlo simulation. Although Monte Carlo simulation has high resolution and flexibility, it lacks capabilities for post-simulation interaction and inference investigations. BNs have also been used to model multiple barrier systems. However, the resolution found on these models is low, which is problematic when very low probabilities need to be estimated. Non-parametric BNs were used to model a multiple barrier system using data obtained from published literature. This type of BN uses normal copulas and conditional rank correlations to model the dependencies between the variables. In case study 6 (Chapter 9), a BN model was developed to estimate potential scenarios, measure the importance of each barrier and represent correlated performances. The results indicated that when correlations are found between barrier performances, it can increase the variability of the effluent pathogen concentration. It was found that the higher the correlation the higher the contribution of the barriers to the variability of the effluent concentration. By "switching off" each barrier sequentially it was possible to estimate the impact of each barrier to the whole process performance. The impact was measurable as the probability of obtaining an increase in the final effluent concentration. The influence of each barrier depended on its magnitude, variability and the reference pathogen for which it provides reduction estimates. It is expected that this approach could be incorporated in common risk evaluation as it provides more capabilities and valuable characteristics than Monte Carlo simulation. Usefully, the BN methodology is able to model distributions without the need for assuming a parametric distribution and capture the correlations between the variables automatically.

It is important to note that although BNs and Bayesian models in general can deal with missing values, the data still needs to be of good quality. This characteristic conversely implies that large data gaps would probably have an influence on the model outcomes or generate misleading results. Also, data pre-processing is required to remove erroneous information that can affect outcomes. BNs are generally adequate for observational studies, so they are particularly useful for systems subject to uncontrolled factors where uncertainty is present. Water treatment systems represent a typical example of an observational study where factors such as water quality and environmental conditions are random. Long term data analysis through BNs can assist in determining indicators for system performance and evaluate the most important variables for system optimisation.

10.1.2 Bayesian models for traditional hardware reliability assessment

Hardware reliability also needs to be adequately assessed on water reuse systems. Studying how the failure of system components could affect the removal performance is crucial to understand the mechanism of the precipitating hazardous events. Such information is also important to consider during validation of treatment systems. To conduct an appropriate analysis, understanding how the system components interact is fundamental. For such tasks, expert elicitation becomes relevant because the systems are designed to interact with people. Fault tree analysis is one of the commonly used tools to identify causal factors affecting a particular undesired outcome. A UV disinfection system was studied by first targeting the crucial components and then modelling their interactions. A fault tree was constructed based on expert information and literature and model parameters were elicited. The model was then encoded as a BN. The BN facilitated probability calculations because of the presence of common cause events. During the model development stage, it was determined that sensors were fundamental for the correct functioning of the UV system and that under specific failures, hazardous events can be encountered. The hazardous event targeted in this case was a non-detected low UV dose. The results showed that a UVI sensor was the most informative of the system's reliability and that UVT sensor impact depended on the influent UVT from upstream processes. Bayesian networks offered flexibility and a clear representation of the system with explicit causal associations between the system components. One of the useful features trialled was the use of noisy gates which can relax the assumptions made with the conventional logic gates (e.g. AND and OR gates). The type of information obtained in this case-study, once the model has been tested and validated on different plants, should be beneficial to organisations such as a water quality regulator to adapt their requirement on disinfection and dual protection barriers. It can also be part of a risk analysis to determine the needs regarding investing in extra water treatment structures or stand by equipment. Similar analyses can be conducted for other treatment processes where sensors are

similarly crucial but whose relative importance is unclear. Elicitation of data for reliability assessment in this case was necessary because the available data were not adequate or not existing. Lack of reliability data is a general problem within the water treatment sector when data are not stored in an appropriate fashion and therefore become difficult to gather and analyse. Furthermore, data specific for water treatment systems are not commonly found in databases or published material which makes expert elicitation essential. Presenting instances where reliability data can be used to assess a system and improve its performance should help the water authorities to re-think what the important parameters are for them to register for their future analysis of plant and cost optimisation.

10.1.3 Opportunities for future application by water utilities and their regulators

The ongoing evolution of better management practices in the water treatment sector requires the investigation of more versatile tools that can improve the understanding and management of risks and reliability. Modifying the application of conventional methodologies could be a difficult task given the limited number of tools being generally recommended in guidelines and reports. One way of encouraging the interest and implementation of promising approaches is by illustrating their use and utility for addressing current relevant problems and analysing existing systems and data sets. This thesis presents instances of model applications for BNs and Bayesian analysis to real problems and datasets and also shows limitations encountered during analysis of water treatment processes.

The application of new tools by water utilities could also be driven by factors such as how much these methods have been implemented in the wider industry. Generating the interest of water utilities will be crucial in facilitating the more widespread application of new tools. The successful transfer of knowledge across the industry will also be linked to willingness and incentives to share information and communicate models. Transfer of knowledge is one way the industry could benefit and improve in terms of more rigorous risk analysis which it may have not fully exploited so far. The structured representation of BNs and Bayesian methods address such issue by facilitating validation and reliability communication between water utilities and their regulators. The creation of DAGs to understand the systems and how the variables interact would be beneficial for managing specific problems common to the whole sector. Implementing new approaches such as BNs and Bayesian methods has limitations including the need for incorporating new analytical capabilities. In order for these proposed methods to be effectively deployed, it will be necessary to promote water engineers and scientists having a stronger understanding of both system functioning and construction of probabilistic models. They could then work as facilitators and in turn change the way stakeholders generally think about or visualise

treatment systems. Such facilitation would also assist in the elicitation process which would be required particularly in the first stages of model development prior to implementation especially the ability to think causally and probabilistically which is the way risk analysis generally works. BNs and Bayesian analysis are tools that particularly deal with these types of challenges, providing explicit representation of associations and uncertainty.

10.1.4 Lessons learned from unsuccessful case studies

Four case studies based on full-scale processes (three based on operational data and one in elicited information) were also planned and implemented in part. However, the analyses proved unsuccessful in acquiring sufficient information for addressing the proposed research questions and objective. For this reason, these have not been detailed in the case studies above. Nevertheless, the likely reasons behind the negative results were themselves instructive and identified to a degree. This section briefly identifies lessons learnt and opportunities for improvement when similar problems are assessed in the future.

10.1.4.1 Studies with using collected process operation data

The first unsuccessful case study investigated the association between turbidity and direct integrity testing data in an ultrafiltration process in Western Australia. In this case study a large SCADA dataset was obtained including records every ten minutes during one year. Additionally, a more limited microbial validation data set was also obtained. The SCADA dataset was able to be pre-processed to remove any invalid information. However, the timeseries results also suggested that there were other influences in the performance of the turbidity sensors as well as in the results of the pressure decay tests which would be hard to incorporate into a data analysis without much greater knowledge of the system which had been a pilot one and had ceased operation prior to data receipt. For example, there was evidence of sensor drifting and increasing noise. Unfortunately, the information about possible causative malfunctions and maintenance was not available in a form which could be retrieved from the databases. This issue represents a clear example of the necessity of having more accessible systems for data management. Collecting such information would require large resource allocation.

The second unsuccessful case study was similar to the study in Chapter 6 in which alternative predictors for LRVs were investigated. A full-scale ozonation process in Melbourne was intensively sampled during a 3-day campaign. Selected predictors included turbidity, UV_{254nm}, ozone dose to TOC ratio, suspended solids and bromate. The results in this case suggested that despite expectations, none of the predictors provided strong associations to LRVs. However, the bench scale experimental results revealed that associations exist between UV_{254nm} and LRV, and

ozone dose to TOC ratio and LRV. The negative results could be explained by the effect of particulate matter which had a wide size distribution and variable characteristics. Suspended solids are known to affect the resistance of microorganisms by shielding and also contribute to the ozone demand. Characterisation and analysis of the effect of suspended solids on removal performance is therefore recommended. Both these first and second unsuccessful studies illustrated the limitations of mining large sets of data without full details of treatment process set ups.

A third case study endeavoured to assess predictors for LRVs during full scale activated sludge treatment. Data were collected from four treatment plants within Australia including microbial concentrations for *E. coli*, adenovirus, polyomavirus and microviridae. Considered predictors were MLSS, turbidity, pH, temperature, dissolved oxygen, conductivity, ammonia, COD, BOD and suspended solids. The data were analysed through BNs and linear models. However, no clear associations were found for any of the variables with the microbial concentrations or LRVs. Some of the potential causes for the negative results were the large data gaps for the operating parameters. Also, other important predictors such as solids retention time and hydraulic retention time were not collected. Although there are techniques for data imputation, when the values are not randomly missing the mechanisms should be studied beforehand increasing the complexity of the study. It is recommended when studying potential predictors for microbial removal to minimise the occurrence of missing values. Furthermore, to decrease the sources of uncertainty, it is preferable to study a single plant with a larger dataset than a number of plants with small datasets each.

10.1.4.2 Elicitation

Eliciting expert judgment may appear simple as it involves discussions with experts and requesting estimated values for a particular problem. However, elicitation is a complex topic which requires preparation from both the interviewer and the domain expert. There are a number of techniques which seek to avoid common expert biases. Elicitation is usually a lengthy process which requires several hours or even days communicating to the domain experts. Elicitation was successfully employed in Chapter 8 to study the hazardous events associated to sensor failures during UV disinfection. Another case study was prepared to develop the BN structure to analyse the impacts associated to reverse osmosis performance. The general model was visualised in a structured fashion using the diagram in Figure 10-2. The elicitation process consisted of asking a series of questions, such as:

- How does the system remove pathogens?

- What operational settings, inputs can affect pathogen removal performance of the system?
- What impacts can affect pathogen removal? (5 most important)
- What measurements and how they are used in this system to (indirectly) check pathogen removal?
- What impacts can affect these measurements?
- How do operators or the SCADA react to deviations on the measurements related to pathogen removal?
- What impacts can affect the interpretation of results?
- What is the understood relationship between monitoring and performance?

The interviews included an operator and a plant manager each one working on different water reuse plants in New South Wales (one for industrial reuse and the other for discharge into the river and subsequent downstream potable use). One of the main encountered problems was that even though pathogen removal was important during the process operation, it was not directly considered or related to changes in operational conditions. Also, the analysed systems had almost zero non-compliance, so there were limited examples of potential impacts affecting its performance. As a consequence, it was difficult to obtain information about hypothetical hazardous scenarios and their effect on the system. It was believed that other reasons for the difficulty in developing the model was the purpose of the plants' operation. None of these two plants had a direct potable use objective so the health risks associated to pathogens were not of primary importance. In general, each plant also works in different ways depending on the required water quality, type of configuration and logistics. Considering the problems encountered in this case study, an alternative strategy to obtain hazardous events information from experts should be used. An optional approach would be to develop a thorough list of the potential impacts obtained from technical literature and membrane systems suppliers. Then, such list can be used to ask experts working on treatment plants, whether those events apply or not on a particular process operation. The experts can also provide information about how the impacts relate to each other and operational variables. Membrane systems are particularly challenging for this type of modelling approach because it is difficult to diagnose and detect failures. Furthermore, there are no general established tools to efficiently detect pathogen breaches. Expert elicitation is an important field within BNs and should also be considered as a valid data input for the models. Further investigation of the applicability of elicited expert knowledge on BNs for water treatment risk analysis is recommended.

10.2 Recommendations for future work

The work presented in this thesis has described investigations of the use of BNs and Bayesian methods for validation and reliability assessment on water and wastewater treatment processes. In doing so, numerous additional avenues for future research have been identified considering other aspects of reliability and knowledge gaps derived from this investigation. These encompass a broad range of areas including analysis of the various types of risks (i.e. not only microbial risks), use of Bayesian networks for data mining, impact of degraded states on performance, investigation of sensor performances, application of soft systems in multi-barrier systems, incorporation of human reliability assessment and anomaly detection.

10.2.1 Other types of risks

In this thesis, the focus was on reduction of pathogenic risks through various water treatment processes that are increasingly used for water reuse applications. The potential applications of Bayesian analysis tools however are much more extensive, including other types of risks and processes which operate indirectly. For example, for the water reuse sector, relevant risks that could be similarly investigated include organisational risks, trace chemical risks, and loss of reputation risks.

10.2.2 Bayesian networks for data mining

Understanding how the variables interact in a system is crucial to identifying the most important factors and optimising processes. In the water industry sector, large amount of data is generated and archived every day, sometimes without giving additional value. BNs are practical tools for discovering and characterising associations between variables in a system or process. Dedicated algorithms have been developed to deal with structure learning and parameter learning. In this thesis, structure learning algorithms were not fully explored, except from Chapter 4 in which some structure learning algorithms were tested. However, there are greater opportunities within the water industry sector to exploit all the features of BNs. These models could be used to improve risk management practices which could be based on data-driven approaches.

10.2.3 Impact of degraded states on performance

Performance of treatment systems is usually studied under normal operating conditions without considering potential degraded states or hazardous conditions. The work presented in this thesis studied the impact of operational variables (Chapter 4, 6 and 7) and ageing (Chapter 5) on performance. However, there are still knowledge gaps about the effect of software failure and

degraded hardware conditions on removal efficiency. In this case it is important to investigate whether the impacts have an effect over the magnitude and/or variability of the performance. Collecting data from multiple plants having different equipment qualities would provide valuable information in the development of empirical models to understand the effect of degradation.

10.2.4 Sensor performance

Sensors are fundamental for correctly monitoring the performance of a treatment system. In Chapter 8, sensor performances were used to investigate a hazardous event. In this case, expert elicited data were used to derive the parameters of the BN. However, real data should be used if a particular system requires evaluation. Because only specific sensor failures can potentially affect specific hazardous events, the frequency of the occurrence of each failure mode needs to be recorded. Additional information to evaluate sensor performance includes false positive rates and false negative rates. For example, during porous membrane filtration, when a direct integrity test (e.g. pressure decay test) result is higher than a certain threshold, it triggers further inspection such as sonic testing. Any of these tests would not have a hundred percent success rate of detection. Therefore, there may be instances where failures in membranes are inadvertent. Further investigations in this area are recommended.

10.2.5 Application of soft systems to assess multi-barrier reliability impacts

A number of cases were analysed to identify novel applications of Bayesian methods to validation and reliability analysis. From the point of view of reliability of treatment systems, an important future work would be the development of models to study the impacts affecting a multi-barrier system (Figure 10-1). Work presented in Chapter 9 showed how a hypothetical common cause impact can affect the performance of the system. Investigating what types of impacts can affect barriers simultaneously is important for hazardous events analysis. In such models, impacts on the treatment performance and indicators should be included. However, considering the complexity of the problem and the limited data available, soft systems could be a solution. Soft systems can be used to capture high level interactions without the need of a detailed causal analysis involving all components (Fenton & Neil, 2012). They may incorporate organisation culture, resources, processes, and procedures among others. These high level aspects have been considered as important causal latent factors that have produced accidents in the past. These factors can inactivate barriers and have an impact on multiple barriers simultaneously. Finding correct indicators for system performance is also crucial to reliably estimating the state of a system at any point of time. These system performance indicators can be operational indicator, but also be higher level indicators, for example similar to key performance indicators. Such indicators

could reflect the maturity of the organisation in terms of management of risk, operational control, training of operators, quality of crucial equipment, quality of maintenance. It is important to incorporate as well any potential associations between the barrier performances. These are important as they can affect the magnitude or variability of the risks.

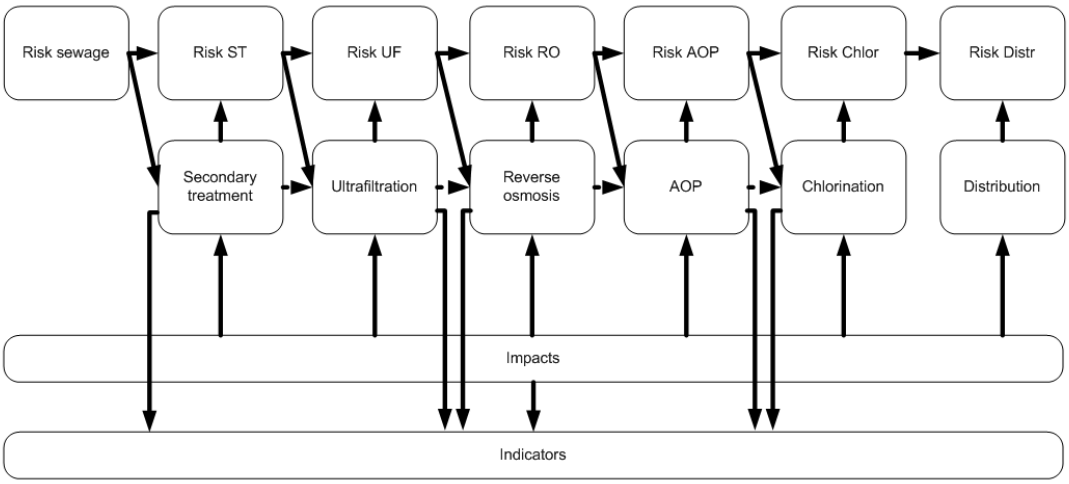

Figure 10-1: Proposed multi-barrier reliability model

10.2.6 Incorporation of human reliability assessment

Given the importance of people to the reliability of water reuse systems during design, operation and the interpretation of monitoring data, human reliability assessment should be of primary concern. Although human reliability analysis has progressed and evolved over the years, using some of the older types of analyses would be a large step ahead and an important start that can facilitate more advanced human reliability analysis in the future. Because human reliability analysis is an area of rapid progress it seems desirable to develop improved human reliability models for water treatment systems. In this manner the gap between what is currently being done in the industry and what is required to be done does not continue widening.

A suggested framework is presented in Figure 10-2 to represent the importance of human reliability in the flow of information in a treatment system. The model incorporates input and output information, impacts and feedbacks. The three blocks represent the main components during measurement of a parameter and the subsequent resulting action. The system provides a barrier against specific hazards. Measurement refers to the sensors that use the system real outputs to provide a measured value which could be precise or biased. Interpretation refers to machine or human interpretation after the sensors have provided the measurements. The interpretation block

generates a reaction which can then be used as feedback for the system or measurement blocks. Human reliability is crucial in the interpretation block because it can produce an erroneous reaction when hazardous events are occurring. It is important to study what factors can impact the human interpretation of measurements. In this regard, BNs have also been found to potentially provide valuable support for human reliability analysis due to its ability to systematise problem scoping and quantification.

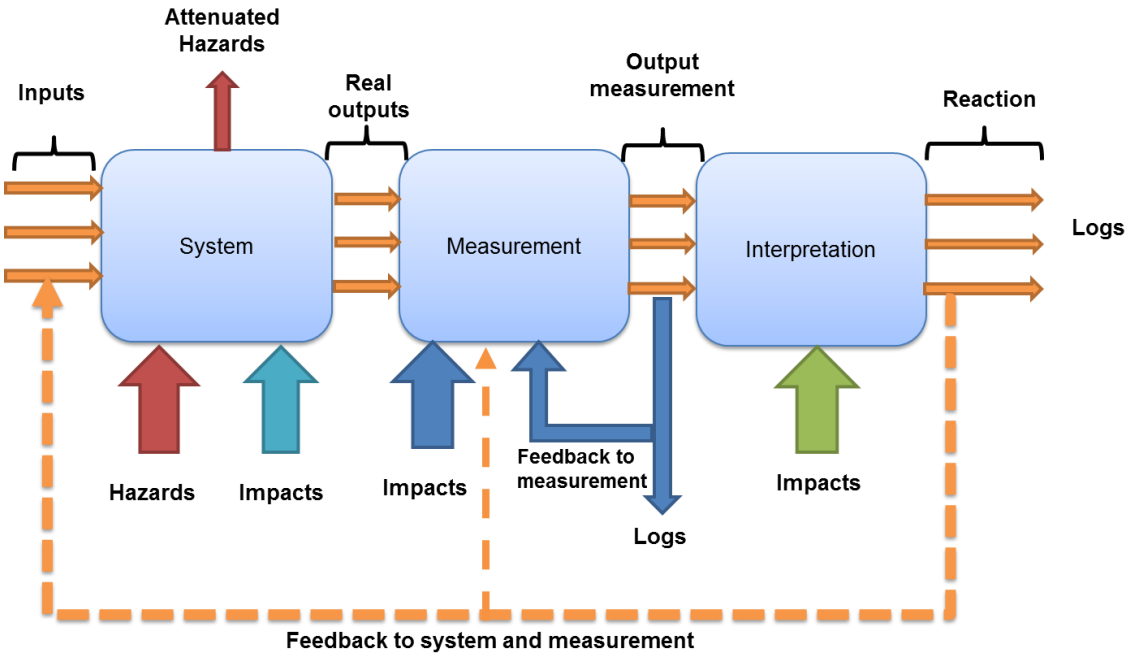

Figure 10-2: Proposed flow of information model for human reliability analysis

10.2.7 Anomaly detection

Plant managers and operators commonly face the problem of how to interpret data to make decisions within process operation. This process involves processing a large amount of data and analysing trends. However, incorrect behaviour is not always evident and only statistical tools can assist diagnosing some problems. Bayesian networks have been used in the past to detect when there is an incorrect behaviour in a system. This approach involves detecting when a very unlikely or conflicting outcome is occurring. The network in this case could be trained with data corresponding to a system working under normal conditions and then inference is performed with online data. The likelihood of the outcomes can then be calculated. Such methodology can be used in any setting within the water reuse system to identify poor performance and minimise hazardous events.

10.2.8 Opportunities and barriers for the implementation of BNs in regulatory practice

Scoping how to implement BNs and Bayesian methods as regular tools to calculate and characterise risks should be considered as the next step for a more robust risk management. Understanding the opportunities and barriers for their implementation feasibility should be investigated. Although some opportunities are clear including defensible decision making, better communication and transparency, there are still a number of barriers and questions which need to be addressed. These barriers include perceived lack of incentives, need for expertise and collection of required data. Finding the right incentives in the industry is fundamental to encourage the use of more sophisticated tools. These incentives can involve the allocation of higher LRV credits when more robust tools are used, inclusion of tools in national guidelines and technical advice from regulators. Expertise in system functioning and model construction is required to convince interested parties that BNs and Bayesian methods are a valuable approach. Collection and interpretation of data are important for effective risk management. However, some data are not easily accessible and are only available if time and resources are allocated for their collection. Considering this issue, it is highly relevant to investigate the benefits of more appropriate data management and associated costs.

It is to be expected that variable local conditions and regulatory paradigms will lead to variable implementation outcomes. As such, assessment of the opportunities and barriers for implementation of BNs in regulatory practice can be expected to be undertaken in multiple jurisdictions worldwide. Some aspect of implementation may prove to be more appropriate and effective in some parts of the world, while other aspects are better suited to different geographic, social and regulatory environments.

Chapter 11: References

Adham, S., Gagliardo, P., Smith, D., Ross, D., Gramith, K. and Trussell, R. (1998) Monitoring the integrity of reverse osmosis membranes. *Desalination,* **119**(1-3)*,* 143-150.

Aguilera, P., Fernández, A., Fernández, R., Rumí, R. and Salmerón, A. (2011) Bayesian networks in environmental modelling. *Environmental Modelling & Software,* **26**(12)*,* 1376-1388.

Antony, A., Blackbeard, J. and Leslie, G. (2012) Removal efficiency and integrity monitoring techniques for virus removal by membrane processes. *Critical Reviews in Environmental Science and Technology,* **42**(9)*,* 891-933.

APHA (2005) Standard Methods for the Examination of Water and Wastewater. Washington, DC, USA.

Arkhangelsky, E., Kuzmenko, D. and Gitis, V. (2007) Impact of chemical cleaning on properties and functioning of polyethersulfone membranes. *Journal of Membrane Science,* **305**(1–2)*,* 176-184.

Asano, T., Burton, F., Leverenz, H., Tsuchihashi, R. and Tchobanoglous, G. (2007) Water Reuse: Issues, Technologies, and Applications. McGraw-Hill, New York.

Assezat, C. (1989) Probabilistic reliability analysis for biological wastewater treatment plants. *Water Science and Technology,* **21**(12)*,* 1813-1816.

Aven, T. (2012) Foundations of risk analysis. John Wiley & Sons, Hoboken, USA.

AWWA and ASCE (2005) Water Treatment Plant Design. 4th Ed.*,* McGraw-Hill, New York.

Baig, A. A., Ruzli, R. and Buang, A. B. (2013) Reliability analysis using fault tree analysis: a review. *International Journal of Chemical Engineering and Applications,* **4**(3)*,* 169-173.

Barbeau, B., Payment, P., Cle, B. and Pre, M. (2000) Evaluating the risk of infection from the presence of Giardia and Cryptosporidium in drinking water. *Quantitative Microbiology,* **2**(1)*,* 37-54.

Barca, E., Del Moro, G., De Sanctis, M., Mascolo, G., Passarella, G. and Di Iaconi, C. (2016) Managing the touristic pressure: performances prediction of an advanced biological system by means of regression trees. *Biochemical Engineering Journal,* **111***,* 43-53.

Barlow, R. E. (1998) Engineering Reliability. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pennsylvania.

Bayes Server. (2017),Website: https://www.bayesserver.com/ Bayes Server Ltd. Bayesian network software, Accessed: *May 2017*.

Bearfield, G. and Marsh, W. (2005) Generalising event trees using Bayesian networks with a case study of train derailment. In: 24th International Conference on Computer Safety, Reliability, and Security (SAFECOMP) (Eds, Winther, R., Gran, B. A. and Dahll, G.) Springer, Fredrikstad, Norway, pp. 52-66.

Beauchamp, N. (2008) Methods for estimating reliability of water treatment processes: an application to conventional and membrane technologies*,* A thesis presented to University of British Columbia, Vancouver. Degree: Master of Applied Science.

Beauchamp, N., Lence, B. J. and Bouchard, C. (2010) Technical hazard identification in water treatment using fault tree analysis. *Canadian Journal of Civil Engineering,* **37**(6)*,* 897-906.

Beaudequin, D., Harden, F., Roiko, A. and Mengersen, K. (2015a) Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. *The Science of the Total Environment,* **541***,* 1393-1409.

Beaudequin, D., Harden, F., Roiko, A. and Mengersen, K. (2017) Potential of Bayesian networks for adaptive management in water recycling. *Environmental Modelling & Software,* **91***,* 251-270.

Beaudequin, D., Harden, F., Roiko, A., Stratton, H., Lemckert, C. and Mengersen, K. (2015b) Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks. *Environment International,* **80***,* 8-18.

Beaudequin, D., Harden, F., Roiko, A., Stratton, H., Lemckert, C. and Mengersen, K. (2015c) Modelling microbial health risk of wastewater reuse: A systems perspective. *Environment International,* **84***,* 131-141.

Bello, G. and Colombari, V. (1980) The human factors in risk analyses of process plants: The control room operator model 'TESEO'. *Reliability Engineering,* **1**(1)*,* 3-14.

Bieder, C., Le-Bot, P., Desmares, E., Bonnet, J. and Cara, F. (1998) MERMOS: Electricite de France (EDF) New Advanced HRA Method. Springer-Verlag, New York.

Blackmore, J., Wang, X., Wang, C.-H., Yum, K.-K., Diaper, C., Zhou, M. and McGregor, G. (2008) Risk assessment and management: a guide for integrated urban water systems, eWater Technical Report. eWater Cooperative Research Centre, Canberra.

Bloom, N. (2006) Reliability Centered Maintenance: Implementation Made Simple. McGraw-Hill, New York.

Bobbio, A., Portinale, L., Minichino, M. and Ciancamerla, E. (2001) Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *Reliability Engineering & System Safety,* **71**(3)*,* 249-260.

Book, G. (2012) Lessons learned from real world application of the bow-tie method. In: SPE Middle East Health, Safety, Security, and Environment Conference and Exhibition Abu Dhabi.

Bourouni, K. (2013) Availability assessment of a reverse osmosis plant: Comparison between reliability block diagram and fault tree analysis methods. *Desalination,* **313***,* 66-76.

Bradshaw, R. A., Illaszewicz, G. L. and Avrahamson, Y. G. (2013) Introducing layer of protection analysis for water safety risk assessments. *Water Quality Research Journal of Canada,* **48**(1)*,* 76-84.

Branch, A. (2016) Validation of membrane bioreactors for water recycling. A thesis in fulfilment of the requirements for the degree of Doctor of Philosophy*,* A thesis presented to UNSW, The University of New South Wales. Degree: PhD.

Branch, A., Trinh, T., Carvajal, G., Leslie, G., Coleman, H. M., Stuetz, R. M., Drewes, J. E., Khan, S. J. and Le-Clech, P. (2016) Hazardous events in membrane bioreactors–Part 3:

Impacts on microorganism log removal efficiencies. *Journal of Membrane Science,* **497**, 514-523.

Brooks, L. E. and Field, K. G. (2016) Bayesian meta-analysis to synthesize decay rate constant estimates for common fecal indicator bacteria. *Water research,* **104**, 262-271.

Brown, L. C. and Mac Berthouex, P. (2002) Statistics for Environmental Engineers. CRC press, Boca Raton, Florida.

Buffle, M.-O., Schumacher, J., Meylan, S., Jekel, M. and von Gunten, U. (2006) Ozonation and advanced oxidation of wastewater: Effect of O3 dose, pH, DOM and HO•-scavengers on ozone decomposition and HO• generation. *Ozone: Science and Engineering,* **28**(4), 247-259.

Burgman, M. (2005) Risks and Decisions for Conservation and Environmental Management. Cambridge University Press, Cambridge.

Burlingame, G. A. and Chalker, R. T. (2017) Risk Management Is Vital to Providing Safe Water. *Opflow,* **43**(8), 20-23.

Cai, B., Liu, Y., Zhang, Y., Fan, Q. and Yu, S. (2013) Dynamic Bayesian networks based performance evaluation of subsea blowout preventers in presence of imperfect repair. *Expert Systems with Applications,* **40**(18), 7544-7554.

California Office of Administrative Law. (2017),Website: https://govt.westlaw.com/calregs/Index?transitionType=Default&contextData=%28sc.Default%29 California Code of Regulations, Title 22: Social Security, Division 4: Environmental Health, Chapter 3: Water Recycling Criteria., Accessed: *May 2017*.

Carvajal, G., Roser, D. J., Sisson, S. A., Keegan, A. and Khan, S. J. (2015) Modelling pathogen log10 reduction values achieved by activated sludge treatment using naïve and semi naïve Bayes network models. *Water Research,* **85**, 304-315.

Carvajal, G., Roser, D. J., Sisson, S. A., Keegan, A. and Khan, S. J. (2017) Bayesian belief network modelling of chlorine disinfection for human pathogenic viruses in municipal wastewater. *Water Research,* **109**, 144-154.

Chaudhry, R. M., Hamilton, K. A., Haas, C. N. and Nelson, K. L. (2017) Drivers of microbial risk for direct potable reuse and de facto reuse treatment schemes: the impacts of source water quality and blending. *International Journal of Environmental Research and Public Health,* **14**(6), 635.

Cheon, S.-P., Kim, S., Kim, J. and Kim, C. (2008) Learning Bayesian networks based diagnosis system for wastewater treatment process with sensor data. *Water Science and Technology,* **58**(12), 2381-2393.

Childress, A. E., Le-Clech, P., Daugherty, J. L., Chen, C. and Leslie, G. L. (2005) Mechanical analysis of hollow fiber membrane integrity in water reuse applications. *Desalination,* **180**(1–3), 5-14.

Chong, H. and Walley, W. J. (1996) Rule-based versus probabilistic approaches to the diagnosis of faults in wastewater treatment processes. *Artificial Intelligence in Engineering,* **10**(3), 265-273.

Chow, C. and Liu, C. (1968) Approximating discrete probability distributions with dependence trees. *Information Theory, IEEE Transactions on Information Theory,* **14**(3), 462-467.

Chryssolouris, G., Lee, M. and Ramsey, A. (1996) Confidence interval prediction for neural network models. *Neural Networks, IEEE Transactions on Neural Networks,* **7**(1), 229-232.

Clark, R. M., Read, E. J. and Hoff, J. C. (1989) Analysis of inactivation of Giardia lamblia by chlorine. *Journal of Environmental Engineering,* **115**(1), 80-90.

Clark, R. M. and Regli, S. (1993) Development of Giardia C· t values for the surface water treatment rule. *Journal of Environmental Science & Health Part A,* **28**(5), 1081-1097.

Cloete, S., Horberry, T. and Head, B. (2013) Urban water system safety: A human factors investigation. *Water: Journal of the Australian Water Association,* **40**(8), 56.

Coble, P. G. (1996) Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. *Marine Chemistry,* **51**(4), 325-346.

Contos, A. (2014) Narrandera Shire Council Drinking Water Management System.Website: http://www.narrandera.nsw.gov.au/cp_content/resources/06.1%20-%20Drinking%20Water%20Management%20System%20-%20ATTACHMENT.PDF Accessed: *June 2017*.

Cooke, R., Kurowicka, D., Hanea, A., Morales, O., Ababei, D., Ale, B. and Roelen, A. (2007) Continuous/discrete non parametric bayesian belief nets with unicorn and uninet. In: Proceedings of Mathematical Methods in Reliability, Glasgow.

Cooper, G. F. and Herskovits, E. (1992) A Bayesian method for the induction of probabilistic networks from data. *Machine Learning,* **9**(4), 309-347.

Cooper, S. E., Wreathall, J., Thompson, C., Drouin, M. and Bley, D. (1996) Knowledge-base for the new human reliability analysis method, A Technique for Human Error Analysis (ATHEANA), Brookhaven National Lab., Upton, NY (United States).

Craik, S. A., Weldon, D., Finch, G. R., Bolton, J. R. and Belosevic, M. (2001) Inactivation of Cryptosporidium parvum oocysts using medium-and low-pressure ultraviolet radiation. *Water Research,* **35**(6), 1387-1398.

Crowl, D. A. (2001) Layer of Protection Analysis: Simplified Process Risk Assessment. Center for Chemical Process Safety of the American Institute of Chemical Engineers, New York.

Cyna, B. (1997) Reliability analyses of water treatment plants. *Water Supply,* **15**(2), 65-74.

Damikouka, I., Katsiri, A. and Tzia, C. (2007) Application of HACCP principles in drinking water treatment. *Desalination,* **210**(1-3), 138-145.

De Dianous, V. and Fiévez, C. (2006) ARAMIS project: A more explicit demonstration of risk control through the use of bow–tie diagrams and the evaluation of safety barrier performance. *Journal of Hazardous Materials,* **130**(3), 220-233.

Démotier, S., Denœux, T. and Schön, W. (2003) Risk assessment in drinking water production using belief functions. In: Symbolic and Quantitative Approaches to Reasoning with Uncertainty Springer, 319-331.

Démotier, S., Odeh, K., Schön, W., Charles, P., Footohi, F. and Allioux, J. (2002) Risk assessment for drinking water production process. In: Proceedings of European Conference on System Dependability and Safety, Lyon, France, pp. 544-550.

Dhillon, B. and Liu, Y. (2006) Human error in maintenance: a review. *Journal of Quality in Maintenance Engineering,* **12**(1)*,* 21-36.

Dhillon, B. S. (2004) Reliability, Quality, and Safety for Engineers. CRC Press, Boca Raton.

Dhillon, B. S. and Balbir, S. D. (1999) Design Reliability: Fundamentals and Applications. CRC Press, Boca Raton.

Dimmock, N. (2000) The Application of HACCP to the Water Industry. In: Proceedings of the 3rd Queensland Environmental Conference: Sustainable Environmental Solutions for Industry and Government; a Focus of Sound, Practical and Economically Viable Solutions for Industry and Government Environmental Engineering Society (Queensland Chapter), pp. 311.

Dominguez-Chicas, A. and Scrimshaw, M. D. (2010) Hazard and risk assessment for indirect potable reuse schemes: An approach for use in developing Water Safety Plans. *Water Research,* **44**(20)*,* 6115-6123.

Donald, M. R. and Mengersen, K. L. (2014) Methods for constructing uncertainty intervals for queries of Bayesian nets. *Australian & New Zealand Journal of Statistics,* **56**(4)*,* 407-427.

Dotan, P., Tal, A. and Arnon, S. (2017) A simple model for estimating the concentrations of natural estrogens in raw wastewater. *Science of The Total Environment,* **575***,* 588-594.

Dowell Iii, A. M. (1998) Layer of protection analysis for determining safety integrity level. *Isa Transactions,* **37**(3)*,* 155-165.

Drewes, J. and Khan, S. (2010) Water reuse for drinking water augmentation, Chapter 16. In: Water Quality & Treatment: A Handbook on Drinking Water. (Ed. Edzwald, J. K.) McGraw-Hill, AWWA, New York.

Duan, R., Tu, J. and Dong, D. (2010) A novel hybrid approach of fault tree and Bayesian networks for fault diagnosis. *Journal of Computational Information Systems,* **6**(11)*,* 3605-3612.

Durga Rao, K., Gopika, V., Sanyasi Rao, V., Kushwaha, H., Verma, A. K. and Srividya, A. (2009) Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. *Reliability Engineering & System Safety,* **94**(4)*,* 872-883.

Dürrenmatt, D. J. and Gujer, W. (2012) Data-driven modeling approaches to support wastewater treatment plant operation. *Environmental Modelling & Software,* **30***,* 47-56.

Eftim, S. E., Hong, T., Soller, J., Boehm, A., Warren, I., Ichida, A. and Nappier, S. P. (2017) Occurrence of norovirus in raw sewage–A systematic literature review and meta-analysis. *Water Research,* **111***,* 366-374.

Eisenberg, D., Soller, J., Sakaji, R. and Olivieri, A. (2001) A methodology to evaluate water and wastewater treatment plant reliability. *Water Science & Technology,* **43**(10)*,* 91-99.

Embrey, D. (1984) Development of slim-maud: a multi-attribute utility approach to human reliability evaluation (NUREG/CP--0048-Vol6). In: Proceedings of the eleventh water reactor safety research information meeting. Washington, DC.

enHealth (2012) Environmental Health Risk Assessment: Guidelines for assessing human health risks from environmental hazards, Environmental Health Standing Committee (enHealth) Australia.

Ericson, C. A. (2005) Hazard Analysis Techniques for System Safety. John Wiley & Sons, Inc., Hoboken, New Jersey.

Ericson, C. A. (2011) Fault Tree Analysis Primer. CreateSpace Inc., Charleston, NC.

Fang, F., Ni, B., Li, W., Sheng, G. and Yu, H. (2011) A simulation-based integrated approach to optimize the biological nutrient removal process in a full-scale wastewater treatment plant. *Chemical Engineering Journal,* **174**(2)*,* 635-643.

Fayyad, U. M. and Irani, K. B. (1993) Multi-interval discretization of continuous-valued attributes for classification learning. In: Proc. IJCAI93, Chambery, France, pp. 1022-1027.

Fenton, N. and Neil, M. (2012) Risk Assessment and Decision Analysis with Bayesian Networks. CRC Press, Boca Raton, Florida.

Ferdous, R., Khan, F., Sadiq, R., Amyotte, P. and Veitch, B. (2013) Analyzing system safety and risks under uncertainty using a bow-tie diagram: An innovative approach. *Process Safety and Environmental Protection,* **91**(1)*,* 1-18.

Fewtrell, L. and Bartram, J. (2001) Water Quality: Guidelines, Standards, and Health: Assessment of Risk and Risk Management for Water-Related Infectious Disease. IWA Publishing, London, UK.

Fitzgerald, S., Owens, C., Angles, M., Hockaday, D., Blackmore, M. and Ferguson, M. (2017) Reframing risk: a risk pathway method for identifying improvement through control and threat analysis. *Water Science and Technology: Water Supply,* ws2017098.

Flapper, T., Campbell, B., Deere, D., Blackbeard, J. and Halliwell, D. (2010) Quantifying pathogen log reduction in Australian activated sludge plants. *Water (Australian Water Association),* **February***,* 56-62.

Flapper, T., Campbell, B., O'Connor, N. and Keegan, A. (2012) Quantification of pathogen removal in Australian Activated sludge plants (Phase 1 and 2).Website: http://clearwater.asn.au//user-data/research-projects/swf-files/72m---7104-quantification-of-pathogen-removal-in-activated-sludge-plant_final_report_.pdf Accessed: *January 2015*.

Flyborg, L., Björlenius, B., Ullner, M. and Persson, K. M. (2017) A PLS model for predicting rejection of trace organic compounds by nanofiltration using treated wastewater as feed. *Separation and Purification Technology,* **174***,* 212-221.

Forss, M. and Ander, H. (2011) Microbiological Risk Assessment of the Water Reclamation Plant in Windhoek*,* A thesis presented to Department of Civil and Environmental Engineering, Division of Water Environment Technology, Chalmers University of Technology Göteborg, Sweden. Degree: Master of Science.

Francy, D. S., Stelzer, E. A., Bushon, R. N., Brady, A. M., Williston, A. G., Riddell, K. R., Borchardt, M. A., Spencer, S. K. and Gellner, T. M. (2012) Comparative effectiveness of membrane bioreactors, conventional secondary treatment, and chlorine and UV disinfection to remove microorganisms from municipal wastewaters. *Water Research,* **46**(13), 4164-4178.

Frank, E., Trigg, L., Holmes, G. and Witten, I. H. (2000) Technical note: Naive Bayes for regression. *Machine Learning,* **41**(1), 5-25.

Fynn, C., Basson, M., Sinkoff, S., Moubray, A. and Nadeau, R. (2007) Applicability of Reliability-Centered Maintenance in the Water Industry. American Water Works Association, Denver, Colorado.

Gale, P. (2002) Using risk assessment to identify future research requirements. *Journal of the American Water Works Association,* **94**(9), 30-2, 34, 36.

Gale, P. (2003) Using event trees to quantify pathogen levels on root crops from land application of treated sewage sludge. *Journal of Applied Microbiology,* **94**(1), 35-47.

Gamage, S., Gerrity, D., Pisarenko, A. N., Wert, E. C. and Snyder, S. A. (2013) Evaluation of process control alternatives for the inactivation of Escherichia coli, MS2 bacteriophage, and Bacillus subtilis spores during wastewater ozonation. *Ozone: Science & Engineering,* **35**(6), 501-513.

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2014) Bayesian Data Analysis. Chapman & Hall/CRC Boca Raton, FL, USA.

Geman, S. and Geman, D. (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* (6), 721-741.

Gerba, C. P., Naranjo, J. E. and Hansan, M. N. (1997) Evaluation of a combined portable reverse osmosis and iodine resin drinking water treatment system for control of enteric waterborne pathogens. *Journal of Environmental Science & Health Part A: Environmental Science and Engineering and Toxicology* **32**(8), 2337-2354.

Gerrity, D., Gamage, S., Jones, D., Korshin, G. V., Lee, Y., Pisarenko, A., Trenholm, R. A., Von Gunten, U., Wert, E. C. and Snyder, S. A. (2012) Development of surrogate correlation models to predict trace organic contaminant oxidation and microbial inactivation during ozonation. *Water Research,* **46**(19), 6257-6272.

Gijsbertsen-Abrahamse, A. J., Cornelissen, E. R. and Hofman, J. A. M. H. (2006) Fiber failure frequency and causes of hollow fiber integrity loss. *Desalination,* **194**(1–3), 251-258.

Goldstein, H., Browne, W. and Rasbash, J. (2002) Partitioning variation in multilevel models. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences,* **1**(4), 223-231.

Governal, R. and Gerba, C. (1999) Removal of MS-2 and PRD-1 bacteriophages from an ultrapure water system. *Journal of Industrial Microbiology & Biotechnology,* **23**(3), 166-172.

Gowland, R. (2006) The accidental risk assessment methodology for industries (ARAMIS)/layer of protection analysis (LOPA) methodology: A step forward towards convergent practices in risk assessment? *Journal of Hazardous Materials,* **130**(3), 307-310.

GPA Engineering, SA Water, AllWater (2011), Glenelg to Adelaide Parklands Recycled Water System (GAP). Operations Manual, Adelaide.

Guldenmund, F., Hale, A., Goossens, L., Betten, J. and Duijm, N. J. (2006) The development of an audit technique to assess the quality of safety barrier management. *Journal of Hazardous Materials,* **130**(3), 234-241.

Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J.-p., Kim, J. H. and Cho, K. H. (2015a) Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences,* **32**, 90-101.

Guo, L., Zhao, Y. and Cui, F.-y. (2015b) A new fault diagnosis method based on Bayesian network model in a wastewater treatment plant of northern China. *Desalination and Water Treatment,* 1-10.

Gyürék, L. L. and Finch, G. R. (1998) Modeling water treatment chemical disinfection kinetics. *Journal of Environmental Engineering,* **124**(9), 783-793.

Haas, C. N. (2004) Neural networks provide superior description of Giardia lamblia inactivation by free chlorine. *Water Research,* **38**(14), 3449-3457.

Haas, C. N. and Heller, B. (1989) Statistics of microbial disinfection. *Water Science and Technology,* **21**(3), 197-201.

Haas, C. N., Rose, J. B. and Gerba, C. P. (2014) Quantitative Microbial Risk Assessment. 2nd Ed., Wiley, New Jersey.

Haas, C. N. and Trussell, R. R. (1998) Frameworks for assessing reliability of multiple, independent barriers in potable water reuse. *Water Science and Technology,* **38**(6), 1-8.

Hadjimichael, A., Comas, J. and Corominas, L. (2016) Do machine learning methods used in data mining enhance the potential of decision support systems? A review for the urban water sector. *AI Communications,* (Preprint), 1-10.

Haimi, H., Mulas, M., Corona, F. and Vahala, R. (2013) Data-derived soft-sensors for biological wastewater treatment plants: An overview. *Environmental Modelling & Software,* **47**, 88-107.

Hajeeh, M. and Chaudhuri, D. (2000) Reliability and availability assessment of reverse osmosis. *Desalination,* **130**(2), 185-192.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. H. (2009) The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter,* **11**(1), 10-18.

Hamada, M. S., Wilson, A. G., Reese, C. S. and Martz, H. F. (2008) Bayesian Reliability. Springer, New York.

Hamilton, P. D., Gale, P. and Pollard, S. J. (2006) A commentary on recent water safety initiatives in the context of water utility risk management. *Environment International,* **32**(8)*,* 958-966.

Hanea, A., Napoles, O. M. and Ababei, D. (2015) Non-parametric Bayesian networks: Improving theory and reviewing applications. *Reliability Engineering & System Safety,* **144***,* 265-284.

Hannaman, G., Spurgin, A. and Lukic, Y. (1984) Human Cognitive Reliability Model for PRA Analysis, EPRI-RP-2170-3. Electronic Power Research Institute, Palo Alto, CA.

Harshman, R. A. (1970) Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multi-modal factor analysis. *UCLA Working Papers in Phonetics,* **16***,* 1-84.

Harwood, V. J., Levine, A. D., Scott, T. M., Chivukula, V., Lukasik, J., Farrah, S. R. and Rose, J. B. (2005) Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Applied and Environmental Microbiology,* **71**(6)*,* 3163-3170.

Havelaar, A. (1994) Application of HACCP to drinking water supply. *Food Control,* **5**(3)*,* 145-152.

He, S. and Li, J. (2011) Confidence Intervals for Neural Networks and Applications to Modeling Engineering Materials. INTECH Open Access Publisher, Rijeka, Croatia.

Health Canada (2007), Canadian Guidelines for Household Reclaimed Water for Use in Toilet and Urinal Flushing, Working Group on Domestic Reclaimed Water of the Federal-Provincial-Territorial Committee on Health and the Environment, Ottawa, Ontario.

Hellier, K. (2003) HACCP at Melbourne Water–Implementation from catchment to tap. *Water Safety,* 29.

Helsel, D. R. (2005) Nondetects and Data Analysis. John Wiley and Sons, New York.

Hokstad, P., Røstum, J., Sklet, S., Rosén, L., Pettersson, T., Lindhe, A., Sturm, S., Beuken, R., Kirchner, D. and Niewersch, C. (2009) Methods for risk analysis of drinking water systems from source to tap–Guidance report on Risk Analysis, Techneau Report, SINTEF, Chalmers University of Technology, TZW, KWR, RWTH Aachen, EU.

Hollnagel, E. (1998) Cognitive Reliability and Error Analysis Method (CREAM). Elsevier, Amsterdam.

Hollnagel, E. (2000) Human reliability analysis. In: International Encyclopaedia of Ergonomics and Human Factors (Ed. Karwowski, W.) CRC Press., Boca Raton.

Honaker, J., King, G. and Blackwell, M. (2011) Amelia II: A program for missing data. *Journal of Statistical Software,* **45**(7)*,* 1-47.

Hong, S., Miller, F. and Taylor, J. (2001) Assessing pathogen removal efficiency of microfiltration by monitoring membrane integrity. *Water Science and Technology: Water Supply,* **1**(4)*,* 43-48.

Hornik, K., Stinchcombe, M. and White, H. (1989) Multilayer feedforward networks are universal approximators. *Neural Networks,* **2**(5)*,* 359-366.

Howard, G. (2003) Water safety plans for small systems: a model for applying HACCP concepts for cost-effective monitoring in developing countries. *Water Science and Technology,* **47**(3)*,* 215-220.

Hrudey, S. E. and Hrudey, E. J. (2005) Safe Drinking Water: Lessons from Recent Outbreaks in Affluent Nations. Wiley Online Library, New York.

Huang, Z., Luo, J., Li, X. and Zhou, Y. (2009) Prediction of effluent parameters of wastewater treatment plant based on improved least square support vector machine with PSO. In: Information Science and Engineering (ICISE), 2009 1st International Conference on IEEE, pp. 4058-4061.

Hübner, U., von Gunten, U. and Jekel, M. (2015) Evaluation of the persistence of transformation products from ozonation of trace organic compounds–A critical review. *Water Research,* **68***,* 150-170.

Hudson, N., Baker, A. and Reynolds, D. (2007) Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters—a review. *River Research and Applications,* **23**(6)*,* 631-649.

IEC (2008) IEC 61078 Analysis techniques for system reliability—Reliability block diagram and Boolean methods. Standards Australia GPO, Australia.

IEC/ISO (2009) Risk management–Risk assessment techniques. IEC/ISO 31010: 2009-11 IEC/ISO, Switzerland.

Ireson, W. G., Coombs, C. F. and Moss, R. Y. (1996) Handbook of Reliability Engineering and Management. Perennial (HarperCollins), New York.

ISO (2005) ISO 9000 Quality management systems — Fundamentals and vocabulary (Third edition). ISO, Switzerland.

ISO (2009) Guide, ISO 73: 2009. Risk management—Vocabulary. ISO, Switzerland.

ISO (2013a) ISO 12489 Petroleum, petrochemical and natural gas industries — Reliability modelling and calculation of safety systems. ISO Switzerland.

ISO (2013b) ISO 14189 Water quality — Enumeration of Clostridium perfringens — Method using membrane filtration. ISO Switzerland.

ISO/TR (2013) Risk management–Guidance for the implementation of ISO 31000, technical report. ISO/TR, Switzerland.

Ito, T., Kato, T., Hasegawa, M., Katayama, H., Ishii, S., Okabe, S. and Sano, D. (2016) Evaluation of virus reduction efficiency in wastewater treatment unit processes as a credit value in the multiple-barrier system for wastewater reclamation and reuse. *Journal of Water and Health,* **14**(6)*,* 879-889.

Ito, T., Kato, T., Takagishi, K., Okabe, S. and Sano, D. (2015) Bayesian modeling of virus removal efficiency in wastewater treatment processes. *Water Science and Technology,* **72**(10)*,* 1789-1795.

Jackson, C. (2017) Reliability Engineering and Management. Amazon Digital Services LLC.

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

Janes, K. R. and Musilek, P. (2007a) Modeling the disinfection of waterborne bacteria using neural networks. *Environmental Engineering Science,* **24**(4), 471-482.

Janes, K. R. and Musilek, P. (2007b) Neural network models of Cryptosporidium parvum inactivation by chlorine dioxide and ozone. *Journal of Environmental Engineering and Science,* **6**(5), 477-482.

Jensen, F., Kjærulff, U. B., Lang, M. and Madsen, A. L. (2002) Hugin-the tool for bayesian networks and influence diagrams. *Probabilistic Graphical Models,* 212-221.

Jensen, F. V. and Nielsen, T. D. (2007) Bayesian Networks and Decision Graphs. Springer, Berlin.

Jensen, H., Thomas, K. and Sharp, D. (1980) Inactivation of coxsackieviruses B3 and B5 in water by chlorine. *Applied and Environmental Microbiology,* **40**(3), 633-640.

Joe, H. (2014) Dependence modeling with copulas. CRC Press, Boca Raton, Florida.

Kaplan, S. and Garrick, B. J. (1981) On the quantitative definition of risk. *Risk Analysis,* **1**(1), 11-27.

Kato, T., Kobayashi, A., Ito, T., Miura, T., Ishii, S., Okabe, S. and Sano, D. (2016) Estimation of concentration ratio of indicator to pathogen-related gene in environmental water based on left-censored data. *Journal of Water and Health,* **14**(1), 14-25.

Kato, T., Miura, T., Okabe, S. and Sano, D. (2013) Bayesian modeling of enteric virus density in wastewater using left-censored data. *Food and Environmental Virology,* **5**(4), 185-193.

Keegan, A., Wati, S. and Robinson, B. (2012) Chlor(am)ine disinfection of human pathogenic viruses in recycled waters. Smart Water Funded Report, SWF 62M-2114.,Website: http://clearwater.asn.au/user-data/research-projects/swf-files/62m---2114-chlorine-disinfection-of-human-pathogenic-viruses-_final_report.pdf Accessed: *May 2015*.

Keller, W. and Modarres, M. (2005) A historical overview of probabilistic risk assessment development and its use in the nuclear power industry: a tribute to the late Professor Norman Carl Rasmussen. *Reliability Engineering & System Safety,* **89**(3), 271-285.

Kelley, D. L. and Allison, R. C. (1979) Fault tree analysis for wastewater treatment. *Journal of the Environmental Engineering Division,* **105**(6), 1105-1112.

Kelley, D. L. and Allison, R. C. (1981) Fault tree analysis and treatment plant instrumentation. *Journal Water Pollution Control Federation,* 43-47.

Khakzad, N. (2015) Application of dynamic Bayesian network to risk analysis of domino effects in chemical infrastructures. *Reliability Engineering & System Safety,* **138**, 263-272.

Khakzad, N., Khan, F. and Amyotte, P. (2011) Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering & System Safety,* **96**(8), 925-932.

Khakzad, N., Khan, F. and Amyotte, P. (2013a) Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection,* **91**(1)*,* 46-53.

Khakzad, N., Khan, F. and Amyotte, P. (2013b) Quantitative risk analysis of offshore drilling operations: A Bayesian approach. *Safety Science,* **57***,* 108-117.

Khan, F. I. and Abbasi, S. (1998) Techniques and methodologies for risk analysis in chemical process industries. *Journal of Loss Prevention in the Process Industries,* **11**(4)*,* 261-277.

Khan, S. and McDonald, J. (2010) Quantifying human exposure to contaminants for multiple-barrier water reuse systems. *Water Science and Technology, ,* **61**(1)*,* 77-83.

Khan, S. J. (2010) Quantitative chemical exposure assessment for water recycling schemes, Waterlines Report Series 27, National Water Commission, Canberra.

Khan, S. J. (2013) Drinking Water Through Recycling. Australian Academy of Technological Sciences and Engineering, Melbourne, Australia.

Khataee, A. R. and Kasiri, M. B. (2011) Modeling of biological water and wastewater treatment processes using artificial neural networks. *CLEAN–Soil, Air, Water,* **39**(8)*,* 742-749.

Kirwan, B. (1994) A guide to practical human reliability assessment. CRC Press, Boca Raton, Florida.

Kjræulff, U. B. and Madsen, A. L. (2012) Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis. Springer Science & Business Media, New York.

Kletz, T. (2001) Learning from Accidents. 3rd Ed.*,* Gulf Profesional Publishing, Oxford.

Koivunen, J., Siitonen, A. and Heinonen-Tanski, H. (2003) Elimination of enteric bacteria in biological–chemical wastewater treatment and tertiary filtration units. *Water Research,* **37**(3)*,* 690-698.

Koller, D. and Friedman, N. (2009) Probabilistic Graphical Models: Principles and Techniques. MIT press, Massachusetts.

Kompany-Zareh, M., Akhlaghi, Y. and Bro, R. (2012) Tucker core consistency for validation of restricted Tucker3 models. *Analytica Chimica Acta,* **723***,* 18-26.

Korb, K. B. and Nicholson, A. E. (2011) Bayesian Artificial Intelligence. Second Ed.*,* Chapman & Hall/CRC Press, London.

Kruschke, J. (2014) Doing Bayesian Data Analysis: a Tutorial with R, Jags, and Stan. Academic Press, Amsterdam.

Kruschke, J. K. (2013) Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General,* **142**(2)*,* 573.

Kutbi, I. I., Sabri, Z. A. and Husseiny, A. A. (1982) Reliability analysis of reverse osmosis plant. *Desalination,* **42**(3)*,* 291-313.

Lampinen, J. and Vehtari, A. (2001) Bayesian approach for neural networks—review and case studies. *Neural Networks,* **14**(3)*,* 257-274.

Lampis, M. and Andrews, J. (2009) Bayesian belief networks for system fault diagnostics. *Quality and Reliability Engineering International,* **25**(4)*,* 409-426.

LeChevallier, M. W. and Au, K.-K. (2004) Water Treatment and Pathogen Control: Process Efficiency in Achieving Safe Drinking-Water. IWA Publishing, London, UK.

Lee, D. S., Lee, M. W., Woo, S. H., Kim, Y.-J. and Park, J. M. (2006) Nonlinear dynamic partial least squares modeling of a full-scale biological wastewater treatment plant. *Process Biochemistry,* **41**(9)*,* 2050-2057.

Lemarchand, K. and Lebaron, P. (2003) Occurrence of Salmonella spp. and Cryptosporidium spp. in a French coastal watershed: relationship with fecal indicators. *FEMS Microbiology Letters,* **218**(1)*,* 203-209.

Lester, R. R., Green, L. C. and Linkov, I. (2007) Site-Specific Applications of Probabilistic Health Risk Assessment: Review of the Literature Since 2000. *Risk Analysis,* **27**(3)*,* 635-658.

Leu, S.-S. and Chang, C.-M. (2013) Bayesian-network-based safety risk assessment for steel construction projects. *Accident Analysis & Prevention,* **54***,* 122-133.

Li, D., Yang, H. Z. and Liang, X. F. (2013) Prediction analysis of a wastewater treatment system using a Bayesian network. *Environmental Modelling & Software,* **40***,* 140-150.

Lifton, G. and Smeaton, P. (2003) Asset risk management in Scottish Water. In: Proc. of the ICE/CIWEM Conference: Risk and reward in asset management delivery–who is best prepared for the challenges ahead? London, UK.

Lindhe, A., Norberg, T. and Rosen, L. (2012) Approximate dynamic fault tree calculations for modelling water supply risks. *Reliability Engineering & System Safety,* **106***,* 61-71.

Lindhe, A., Rosén, L., Norberg, T. and Bergstedt, O. (2009) Fault tree analysis for integrated and probabilistic risk analysis of drinking water systems. *Water Research,* **43**(6)*,* 1641-1653.

Liu, H.-C., Liu, L. and Liu, N. (2012a) Risk evaluation approaches in failure mode and effects analysis: A literature review. *Expert Systems with Applications,* **40**(2)*,* 828-838.

Liu, K. F.-R., Lu, C.-F., Chen, C.-W. and Shen, Y.-S. (2012b) Applying Bayesian belief networks to health risk assessment. *Stochastic Environmental Research and Risk Assessment,* **26**(3)*,* 451-465.

Liu, O., Seraichekas, H., Akin, E., Brashear, D. and Katz, E. (1971) Relative resistance of twenty human enteric viruses to free chlorine in Potomac water. Virus and water quality: occurrence and control. In: 13th Water Quality Conference (Ed, Snoeyink, V.) Department of Civil Engineering, University of Illinois, Urbana-Champaign, pp. 171-195.

Liu, X., Li, H. and Li, L. (2008) Building method of diagnostic model of Bayesian networks based on fault tree. In: Seventh International Symposium on Instrumentation and Control Technology International Society for Optics and Photonics, pp. 71272C-71272C-6.

Lodder, W. J. and de Roda Husman, A. M. (2005) Presence of noroviruses and other enteric viruses in sewage and surface waters in The Netherlands. *Applied and Environmental Microbiology,* **71**(3)*,* 1453-1461.

Lunn, D., Jackson, C., Best, N., Thomas, A. and Spiegelhalter, D. (2012) The BUGS Book: a Practical Introduction to Bayesian Analysis. CRC press, Boca Raton, Florida.

Lunn, D. J., Thomas, A., Best, N. and Spiegelhalter, D. (2000) WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing,* **10**(4), 325-337.

Lyons, M., Adams, S., Woloshynowych, M. and Vincent, C. (2004) Human reliability analysis in healthcare: A review of techniques. *The International Journal of Risk and Safety in Medicine,* **16**(4), 223-237.

Mahadevan, S., Zhang, R. and Smith, N. (2001) Bayesian networks for system reliability reassessment. *Structural Safety,* **23**(3), 231-251.

Marcot, B. G. (2012) Metrics for evaluating performance and uncertainty of Bayesian network models. *Ecological Modelling,* **230**, 50-62.

Markert, F., Engebø, A. and Nielsen, S. (2009) A barrier analysis of a generic hydrogen refuelling station. In: 3rd International Conference on Hydrogen Safety.

Marsh, D. and Bearfield, G. (2008) Generalizing event trees using Bayesian networks. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability,* **222**(2), 105-114.

Martz, H. and Waller, R. (1982) Bayesian Reliability Analysis. Wiley, New York.

Matsushita, T., Shirasaki, N., Tatsuki, Y. and Matsui, Y. (2013) Investigating norovirus removal by microfiltration, ultrafiltration, and precoagulation–microfiltration processes using recombinant norovirus virus-like particles and real-time immuno-PCR. *Water Research,* **47**(15), 5819-5827.

Maxwell, S. and Franssen, I. (2012) Risk Management Framework for Water Planning and Management. Website: https://www.waterconnect.sa.gov.au/Content/Publications/DEWNR/Risk%20Management%20Framework%20for%20Water%20Planning%20and%20Management.pdf Accessed: *November 2013*.

Medema, G., Hoogenboezem, W., van der Veer, A., Ketelaars, H., Hijnen, W. and Nobel, P. (2003) Quantitative risk assessment of Cryptosporidium in surface water treatment. *Water Science & Technology,* **47**(3), 241-247.

Merlo, R., De Las Casas, C., Henneman, S., Witzgall, R., Yu, W., Ramberg, S., Parker, D. and Ohlinger, K. (2015) Process performance of secondary effluent granular media filtration with and without preozonation. *Water Environment Research,* **87**(7), 595-606.

Messner, M. J., Chappell, C. L. and Okhuysen, P. C. (2001) Risk assessment for Cryptosporidium: a hierarchical Bayesian analysis of human dose response data. *Water Research,* **35**(16), 3934-3940.

Metcalf & Eddy Inc., Tchobanoglous, G., Stensel, D., Tsuchihashi, R. and Burton, F. (2014) Wastewater Engineering: Treatment and Resource Recovery. Fifth Ed., McGraw-Hill, New York.

Minitab (2010) Minitab 16 statistical software. *Minitab Inc., State College, Pennsylvania, USA.*

Mkrtchyan, L., Podofillini, L. and Dang, V. N. (2015) Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability Engineering & System Safety,* **139***,* 1-16.

Modarres, M., Kaminskiy, M. and Krivtsov, V. (1999) Reliability Engineering and Risk Analysis: a Practical Guide. CRC press, Boca Raton, Florida.

Mohammadpour, R., Shaharuddin, S., Chang, C. K., Zakaria, N. A., Ab Ghani, A. and Chan, N. W. (2015) Prediction of water quality index in constructed wetlands using support vector machine. *Environmental Science and Pollution Research,* **22**(8)*,* 6208-6219.

Montani, S., Portinale, L. and Bobbio, A. (2005) Dynamic Bayesian networks for modeling advanced fault tree features in dependability analysis. In: Proceedings of the 16th European Conference on Safety and Reliability, Leiden, The Netherlands, AA Balkema, pp. 1415-1422.

Montani, S., Portinale, L., Bobbio, A. and Codetta-Raiteri, D. (2008) Radyban: A tool for reliability analysis of dynamic fault trees through conversion into dynamic Bayesian networks. *Reliability Engineering & System Safety,* **93**(7)*,* 922-932.

Morris, D. E., Oakley, J. E. and Crowe, J. A. (2014) A web-based tool for eliciting probability distributions from experts. *Environmental Modelling & Software,* **52***,* 1-4.

Mullenger, J., Ryan, G. and Hearn, J. (2002) A water authority's experience with HACCP. *Water Science and Technology: Water Supply,* **2**(5-6)*,* 149-155.

Müller, P. and Insua, D. R. (1998) Issues in Bayesian analysis of neural network models. *Neural Computation,* **10**(3)*,* 749-770.

Murphy, K. R., Stedmon, C. A., Graeber, D. and Bro, R. (2013) Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Analytical Methods,* **5**(23)*,* 6557-6566.

Nair, V. and Hinton, G. E. (2010) Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 807-814.

Neil, M., Fenton, N. and Nielson, L. (2000) Building large-scale Bayesian networks. *The Knowledge Engineering Review,* **15**(3)*,* 257-284.

Nenonen, N. P., Hannoun, C., Larsson, C. U. and Bergström, T. (2012) Marked genomic diversity of norovirus genogroup I strains in a waterborne outbreak. *Applied and Environmental Microbiology,* AEM. 07350-11.

Neumann, M. B., von Gunten, U. and Gujer, W. (2007) Uncertainty in prediction of disinfection performance. *Water Research,* **41**(11)*,* 2371-2378.

NHMRC, EPHC and AHMC (2006) Australian Guidelines for Water Recycling: Managing Health and Environmental Risks (Phase 1), Natural Resource Management Ministerial Council, Environment Protection and Heritage Council, Australian Health Ministers' Conference, Canberra.

NHMRC and NRMMC (2011) Australian drinking water guidelines paper 6 national water quality management strategy, National Health and Medical Research Council, National Resource Management Ministerial Council, Commonwealth of Australia, Canberra.

Ni, C., Chen, J., Tsai, Y., Chen, W. and Chen, C. (2002) Ozonation of domestic secondary effluent for recycling and reuse-a pilot plant study. *Water Science and Technology,* **46**(4-5), 361-366.

Niku, S., Schroeder, E. D. and Haugh, R. S. (1982) Reliability and stability of trickling filter processes. *Journal (Water Pollution Control Federation),* 129-134.

Niku, S., Schroeder, E. D. and Samaniego, F. J. (1979) Performance of activated sludge processes and reliability-based design. *Journal (Water Pollution Control Federation),* 2841-2857.

Nokes, C., Fenton, E. and Randall, C. (1999) Modelling the formation of brominated trihalomethanes in chlorinated drinking waters. *Water Research,* **33**(17), 3557-3568.

Norsys (2015) Norsys Software Corp. - Bayes Net Software. *Available at: https://www.norsys.com/*.

NRC (1983) Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington, DC, USA.

NRMMC, EPHC and NHMRC (2008) Australian Guidelines for Water Recycling: Managing Health and Environmental Risks (Phase 2), Augmentation of Drinking Water Supplies, Environment Protection and Heritage Council, the National Health and Medical Research Council and the Natural Resource Management Ministerial Council, Canberra.

Ntzoufras, I. (2011) Bayesian Modeling Using Winbugs. John Wiley & Sons, Hoboken, New Jersey.

O'Connor, P. D. and Kleyner, A. (2012) Practical Reliability Engineering. John Wiley & Sons, Chichester, UK.

Oliveira, S. and von Sperling, M. (2007) Reliability analysis of stabilisation pond systems. *Water Science & Technology,* **55**(11), 127-134.

Oliveira, S. C. and Von Sperling, M. (2008) Reliability analysis of wastewater treatment plants. *Water research,* **42**(4), 1182-1194.

Olivieri, A., Eisenberg, D., Soller, J., Eisenberg, J., Cooper, R., Tchobanoglous, G., Trussell, R. and Gagliardo, P. (1999) Estimation of pathogen removal in an advanced water treatment facility using Monto Carlo simulation. *Water Science and Technology,* **40**(4), 223-233.

Ottoson, J., Hansen, A., Björlenius, B., Norder, H. and Stenström, T. (2006) Removal of viruses, parasitic protozoa and microbial indicators in conventional and membrane processes in a wastewater pilot plant. *Water Research,* **40**(7), 1449-1457.

Paraskeva, P. and Graham, N. (2005) Treatment of a secondary municipal effluent by ozone, UV and microfiltration: microbial reduction and effect on effluent quality. *Desalination,* **186**(1-3), 47-56.

Pasman, H. and Rogers, W. (2012) Risk assessment by means of Bayesian networks: A comparative study of compressed and liquefied H 2 transportation and tank station risks. *International Journal of Hydrogen Energy,* **37**(22), 17415-17425.

Pasman, H. and Rogers, W. (2013) Bayesian networks make LOPA more effective, QRA more transparent and flexible, and thus safety more definable! *Journal of Loss Prevention in the Process Industries,* **26**(3), 434-442.

Passantino, L. and Owen, D. (2005) Integrating UV Disinfection into Existing Water Treatment Plants. AWWA Research Foundation, American Water Works Association, Washington DC.

Payment, P., Tremblay, M. and Trudel, M. (1985) Relative resistance to chlorine of poliovirus and coxsackievirus isolates from environmental sources and drinking water. *Applied and Environmental Microbiology,* **49**(4), 981-983.

Pecson, B. M., Triolo, S. C., Olivieri, S., Chen, E., Pisarenko, A. N., Yang, C.-C., Olivieri, A., Haas, C. N., Trussell, R. S. and Trussell, R. R. (2017) Reliability of pathogen control in direct potable reuse: Performance evaluation and QMRA of a full-scale 1 MGD advanced treatment train. *Water Research*.

Pecson, B. M., Trussell, R. S., Pisarenko, A. N. and Trussell, R. R. (2015) Achieving reliability in potable reuse: the four Rs. *Journal-American Water Works Association,* **107**(3), 48-58.

Petterson, S. R., Signor, R. S. and Ashbolt, N. J. (2007) Incorporating method recovery uncertainties in stochastic estimates of raw water protozoan concentrations for QMRA. *Journal of Water and Health,* **5**(S1), 51-65.

Phan, T. D., Smart, J. C., Capon, S. J., Hadwen, W. L. and Sahin, O. (2016) Applications of Bayesian belief networks in water resource management: A systematic review. *Environmental Modelling & Software,* **85**, 98-111.

Plummer, M. (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd international workshop on distributed statistical computing, Vol. 124 Vienna, pp. 125.

Plummer, M. (2013) rjags: Bayesian Graphical Models using MCMC/2013.Website: https://cran.r-project.org/web/packages/rjags/index.html Accessed: *May 2015*.

Plummer, M. (2015) JAGS Version 4.0. 0 user manual.Website: https://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x Accessed: *May 2016*.

Pollard, S. (2008) Risk Management for Water and Wastewater Utilities. IWA publishing, London, UK.

Pouillot, R., Van Doren, J. M., Woods, J., Plante, D., Smith, M., Goblick, G., Roberts, C., Locas, A., Hajen, W. and Stobo, J. (2015) Meta-Analysis of the Reduction of Norovirus and Male-Specific Coliphage Concentrations in Wastewater Treatment Plants. *Applied and Environmental Microbiology,* **81**(14), 4669-4681.

Priddy, K. L. and Keller, P. E. (2005) Artificial Neural Networks: an Introduction. SPIE Press, Bellingham.

Puspitasari, V., Granville, A., Le-Clech, P. and Chen, V. (2010) Cleaning and ageing effect of sodium hypochlorite on polyvinylidene fluoride (PVDF) membrane. *Separation and Purification Technology,* **72**(3), 301-308.

Qian, S. S., Donnelly, M., Schmelling, D. C., Messner, M., Linden, K. G. and Cotton, C. (2004) Ultraviolet light inactivation of protozoa in drinking water: a Bayesian meta-analysis. *Water Research,* **38**(2)*,* 317-326.

Qin, X., Gao, F. and Chen, G. (2012) Wastewater quality monitoring system using sensor fusion and machine learning techniques. *Water Research,* **46**(4)*,* 1133-1144.

Quinn, B. and Marriott, N. (2002) HACCP plan development and assessment: a review. *Journal of Muscle Foods,* **13**(4)*,* 313-330.

R-project (2014) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012.Website: https://cran.r-project.org/ Accessed: *May 2015*.

Rausand, M. (1998) Reliability centered maintenance. *Reliability Engineering & System Safety,* **60**(2)*,* 121-132.

Rausand, M. and Høyland, A. (2004) System Reliability Theory: Models, Statistical Methods, and Applications. 2nd Ed.*,* John Wiley & Sons, Hoboken, New Jersey.

Ravikrishna, P. and Mhaisalkar, V. (1996) Fault tree analysis for conventional water and wastewater treatment systems. *Journal of the Institution of Engineers (India): Environmental Engineering Division,* 7-13.

Reason, J. (1990) Human Error. Cambridge University Press, Cambridge.

Reason, J. (2016) Managing the Risks of Organizational Accidents. Routledge, Abingdon-on-Thames.

Reaume, M. J., Seth, R., McPhedran, K. N., da Silva, E. F. and Porter, L. A. (2015) Effect of media on biofilter performance following ozonation of secondary treated municipal wastewater effluent: sand vs. GAC. *Ozone: Science & Engineering,* **37**(2)*,* 143-153.

Reeve, P. J., Regel, R., Dreyfus, J., Monis, P., Lau, M., King, B. and Van den Akker, B. (2016) Virus removal of new and aged UF membranes at full-scale in a wastewater reclamation plant. *Environmental Science: Water Research & Technology*.

Risebro, H. L., Doria, M. F., Andersson, Y., Medema, G., Osborn, K., Schlosser, O. and Hunter, P. R. (2007) Fault tree analysis of the causes of waterborne outbreaks. *Journal of Water and Health,* **5***,* 1.

Robertson, L., Paton, C., Campbell, A., Smith, P., Jackson, M., Gilmour, R., Black, S., Stevenson, D. and Smith, H. (2000) Giardia cysts and Cryptosporidium oocysts at sewage treatment works in Scotland, UK. *Water Research,* **34**(8)*,* 2310-2322.

Robillot, C., Le-Clech, P., Pype, M.-L., Sidhu, J., Khan, S. and Monis, P. (2016) National Validation Framework for Water Recycling: Overview of Priority Research, 192220272X, Australian Water Recycling Centre of Excellence, Brisbrane, Australia.

Rohatgi, A. (2011) WebPlotDigitizer.Website: http://arohatgi.info/WebPlotDigitizer/ Accessed: *April 2017*.

Rose, J. B., Dickson, L. J., Farrah, S. R. and Carnahan, R. P. (1996) Removal of pathogenic and indicator microorganisms by a full-scale water reclamation facility. *Water Research,* **30**(11), 2785-2797.

Rosén, C., Rieger, L., Jeppsson, U. and Vanrolleghem, P. (2008) Adding realism to simulated sensors and actuators. *Water Science and Technology,* **57**(3), 337-344.

Rosén, L., Hokstad, P., Lindhe, A., Sklet, S. and Røstum, J. (2007) Generic framework and methods for integrated risk management in water safety plans, Deliverable number D, KIWA and LNEC EU.

Roser, D., Carvajal, G., van den Akker, B., Keegan, A., Regel, R. and Khan, S. (2015) National Validation Guidelines for Water Recycling: Comprehensive Bayesian Recycled Water Validation, 1922202703, Australian Water Recycling Centre of Excellence, Brisbane, Australia.

Sahely, B. S. and Bagley, D. M. (2001) Diagnosing upsets in anaerobic wastewater treatment using Bayesian belief networks. *Journal of Environmental Engineering,* **127**(4), 302-310.

Scutari, M. (2010) R package, bnlearn: Bayesian network structure learning.Website: http://www.bnlearn.com/ Accessed: *June 2015*.

Scutari, M. and Denis, J.-B. (2014) Bayesian Networks: with Examples in R. CRC Press, Boca Raton, Florida.

Shelton, D. R., Karns, J. S., Higgins, J. A., Van Kessel, J. A. S., Perdue, M. L., Belt, K. T., Russell-Anelli, J. and DebRoy, C. (2006) Impact of microbial diversity on rapid detection of enterohemorrhagic Escherichia coli in surface waters. *FEMS Microbiology Letters,* **261**(1), 95-101.

Sherchan, S. P., Snyder, S. A., Gerba, C. P. and Pepper, I. L. (2014) Inactivation of MS2 coliphage by UV and hydrogen peroxide: Comparison by cultural and molecular methodologies. *Journal of Environmental Science and Health, Part A,* **49**(4), 397-403.

Singh, K. P., Basant, N., Malik, A. and Jain, G. (2010) Modeling the performance of "up-flow anaerobic sludge blanket" reactor based wastewater treatment plant using linear and nonlinear approaches—a case study. *Analytica Chimica Acta,* **658**(1), 1-11.

Sivaganesan, M. and Rice, E. W. (2003) Estimation of Giardia Ct values at high pH for the surface water treatment rule. *Journal of Environmental Science and Health, Part A,* **38**(9), 1959-1970.

Sivaganesan, M., Rice, E. W. and Mariñas, B. J. (2003) A Bayesian method of estimating kinetic parameters for the inactivation of Cryptosporidium parvum oocysts with chlorine dioxide and ozone. *Water Research,* **37**(18), 4533-4543.

Smeets, P. and Medema, G. (2006) Combined use of microbiological and non-microbiological data to assess treatment efficacy. *Water Science & Technology,* **54**(3), 35-40.

Smeets, P., Medema, G., Dullemont, Y., Van Gelder, P. and Van Dijk, J. (2008) Improved methods for modelling drinking water treatment in quantitative microbial risk assessment; a case study of *Campylobacter* reduction by filtration and ozonation. *Journal of Water and Health,* **6**(3), 301-314.

Smeets, P. W. (2010) Stochastic Modelling of Drinking Water Treatment in Quantitative Microbial Risk Assessment. IWA Publishing, London, UK.

Smid, J., Verloo, D., Barker, G. and Havelaar, A. (2010) Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment. *International Journal of Food Microbiology,* **139***,* S57-S63.

Smith, D. B., Clark, R. M., Pierce, B. K. and Regli, S. (1995) An Empirical Model for Interpolating C*T Values for Chlorine Inactivation of Giardia lamblia. *J. Water SRTAqua,* **44**(5)*,* 203-211.

Smith, D. J. (2011) Reliability, Maintainability and Risk: Practical Safety-Related Systems Engineering Methods. Butterworth-Heinemann, Oxford, UK.

Solano-Soto, J. and Sucar, L. E. (2001) A methodology for reliable systems design. In: Engineering of Intelligent Systems, 14th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE 2001 (Eds, Monostori, L., Váncza, J. and Ali, M.) Springer, Budapest, Hungary, pp. 734-745.

Soller, J., Eisenberg, J. and Olivieri, A. (1999) Evaluation of pathogen risk assessment framework, International Life Sciences Institute, Risk Science Institute, Washington, DC.

Soller, J., Olivieri, A., Eisenberg, J., DeGeorge, J., Cooper, R. and Tchobanoglous, G. (2006) A public health evaluation of recreational water impairment. *Journal of Water and Health,* **4***,* 1-19.

Soller, J. A., Eftim, S. E., Warren, I. and Nappier, S. P. (2017) Evaluation of microbiological risks associated with direct potable reuse. *Microbial Risk Analysis,* **5***,* 3-14.

Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2007) OpenBUGS user manual, version 3.0. 2. *MRC Biostatistics Unit, Cambridge*.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **64**(4)*,* 583-639.

Spurgin, A. J. (2010) Human Reliability Assessment Theory and Practice. CRC Press, Boca Raton, Florida.

Stadterman, K., Sninsky, A., Sykom, J. and Jakubowski, W. (1995) Removal and inactivation of Cryptosporidium oocysts by activated sludge treatment and anaerobic digestion. *Water Science and Technology,* **31**(5)*,* 97-104.

Stahlschmidt, M., Regnery, J., Campbell, A. and Drewes, J. E. (2016) Application of 3D-fluorescence/PARAFAC to monitor the performance of managed aquifer recharge facilities. *Journal of Water Reuse and Desalination,* **6**(2)*,* 249-263.

Stamatis, D. H. (2003) Failure Mode and Effect Analysis: FMEA from Theory to Execution. American Society for Quality Press, Milwaukee, Wisconsin.

Stan-Development-Team (2016) Stan Modeling Language Users Guide and Reference Manual, Version 2.9.0.

Stapelberg, R. F. (2008) Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design. Springer, Berlin.

Su, Y. and Yajima, M. (2015) R2jags: Using R to run 'JAGS'. *R package version 0.5-7*.

Summers, A. E. (2003) Introduction to layers of protection analysis. *Journal of Hazardous Materials,* **104**(1)*,* 163-168.

Suwa, M. and Suzuki, Y. (2001) Occurrence of Cryptosporidium in Japan and counter-measures in wastewater treatment plants. *Water Science & Technology,* **43**(12)*,* 183-186.

Swain, A. D. (1987) Accident sequence evaluation program: human reliability analysis procedure, Sandia National Labs., Albuquerque, NM (USA); Nuclear Regulatory Commission, Washington, DC (USA). Office of Nuclear Regulatory Research, United States.

Swain, A. D. and Guttmann, H. E. (1983) Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report, Sandia National Labs., Albuquerque, United States.

SydneyWater (2010a) Failure Mode Effects and Criticality Analysis (FMECA) procedure. Website: http://www.sydneywater.com.au/web/groups/publicwebcontent/documents/document/zgrf/mdq2/~edisp/dd_046414.pdf Accessed: *October 2013*.

SydneyWater (2010b) Reliability Block Diagram (RBD) procedure. Website: http://www.sydneywater.com.au/web/groups/publicwebcontent/documents/document/zgrf/mdq2/~edisp/dd_046415.pdf Accessed: *January 2014*.

SydneyWater (2011) Maintenance related clauses for capital and operational projects, version 5. Website: http://www.sydneywater.com.au/web/groups/publicwebcontent/documents/webasset/zgrf/mdq3/~edisp/iconn_ud_dd_047294.doc Accessed: *January 2014*.

SydneyWater (2014) Standards and Specifications. Website: http://www.sydneywater.com.au/SW/plumbing-building-developing/building/providers/standards-and-specifications/index.htm Accessed: *January 2014*.

Taheriyoun, M. and Moradinejad, S. (2015) Reliability analysis of a wastewater treatment plant using fault tree analysis and Monte Carlo simulation. *Environmental Monitoring and Assessment,* **187**(1)*,* 1-13.

Tanaka, H., Asano, T., Schroeder, E. D. and Tchobanoglous, G. (1998) Estimating the safety of wastewater reclamation and reuse using enteric virus monitoring data. *Water Environment Research,* **70**(1)*,* 39-51.

Tang, Y., Wu, S., Miao, X., Pollard, S. J. and Hrudey, S. E. (2013) Resilience to evolving drinking water contamination risks: a human error prevention perspective. *Journal of Cleaner Production,* **57***,* 228-237.

Tao, E., Shen, W., Liu, T. and Chen, X. (2013) Fault diagnosis based on PCA for sensors of laboratorial wastewater treatment process. *Chemometrics and Intelligent Laboratory Systems,* **128***,* 49-55.

Tchórzewska-Cieślak, B., Boryczko, K. and Eid, M. (2011) Failure scenarios in water supply system by means of fault tree analysis. In: Advances in Safety, Reliability and Risk Management (Ed. Soares, G.) CRC Press, Boca Raton, Florida, 2492-2499.

Teunis, P., Evers, E. and Slob, W. (1999) Analysis of variable fractions resulting from microbial counts. *Quantitative Microbiology,* **1**(1)*,* 63-88.

Teunis, P. and Havelaar, A. (1999) Cryptosporidium in drinking water: Evaluation of the ILSI quantitative risk assessment framework, National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands.

Teunis, P., Rutjes, S., Westrell, T. and de Roda Husman, A. (2009) Characterization of drinking water treatment for virus risk assessment. *Water Research,* **43**(2)*,* 395-404.

Tng, K., Currie, J., Roberts, C., Koh, S., Audley, M. and Leslie, G. L. (2015) Resilience of Advanced Water Treatment Plants for Potable Reuse, 1922202347, Australian Water Recycling Centre of Excellence, Brisbane, Australia.

Toifl, M., Diaper, C., O'Halloran, R. and Conscience, W. (2010) Review of process and performance monitoring techniques applicable to large and small scale wastewater recycling systems, CSIRO: Water for a Healthy Country National Research Flagship.

Torres-Toledano, J. G. and Sucar, L. E. (1998) Bayesian networks for reliability analysis of complex systems. In: Progress in Artificial Intelligence—IBERAMIA 98 (Ed. Coelho, H.) Springer, Lisbon, Portugal, 195-206.

Tripathi, S. and Tripathi, B. (2011) Efficiency of combined process of ozone and bio-filtration in the treatment of secondary effluent. *Bioresource Technology,* **102**(13)*,* 6850-6856.

Tyrrell, S. A., Rippey, S. R. and Watkins, W. D. (1995) Inactivation of bacterial and viral indicators in secondary sewage effluents, using chlorine and ozone. *Water Research,* **29**(11)*,* 2483-2490.

US-EPA and USDA/FSIS (2012) Microbial Risk Assessment Guideline: Pathogenic Microorganisms with Focus on Food and Water. US Environmental Protection Agency (EPA); US Department of Agriculture/Food Safety and Inspection Service (USDA/FSIS) Washington, DC, USA.

USEPA (2001) Method 1601: Malespecific (F+) and somatic coliphage in water by two step enrichment procedure, EPA number 821-R-01-030,. Washington, DC.

USEPA (2003) LT1ESWTR Disinfection Profiling and Benchmarking Technical Guidance Manual, Office of Water, United States Environmental Protection Agency, USA.

USEPA (2005) Membrane filtration guidance manual, United States Environmental Protection Agency, USA.

USEPA (2006) Ultraviolet disinfection guidance manual, United States Environmental Protection Agency Washington, USA.

Uyak, V., Toroz, I. and Meric, S. (2005) Monitoring and modeling of trihalomethanes (THMs) for a water treatment plant in Istanbul. *Desalination,* **176**(1-3)*,* 91-101.

van den Akker, B., Trinh, T., Coleman, H. M., Stuetz, R. M., Le-Clech, P. and Khan, S. J. (2014) Validation of a full-scale membrane bioreactor and the impact of membrane cleaning on the removal of microbial indicators. *Bioresource Technology,* **155***,* 432-437.

Van den Akker, B., Whiffin, V., Cox, P., Beatson, P., Ashbolt, N. and Roser, D. (2011) Estimating the risk from sewage treatment plant effluent in the Sydney catchment area. *Water Science and Technology,* **63**(8)*,* 1707.

Vatn, J. (1996) Maintenance optimization: models and methods*,* A thesis presented to Department of Production and Quality Engineering, Norwegian University of Science and Technology, Trondheim, Norway. Degree: PhD.

VDoH (2013) Guidelines for validating treatment processes for pathogen reduction, Supporting Class A recycled water schemes in Victoria, Department of Health, Victoria, Australia, Victoria, Australia.

Vehtari, A., Sarkka, S. and Lampinen, J. (2000) On MCMC sampling in Bayesian MLP neural networks. In: Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on, Vol. 1 IEEE, pp. 317-322.

Verma, A., Wei, X. and Kusiak, A. (2013) Predicting the total suspended solids in wastewater: a data-mining approach. *Engineering Applications of Artificial Intelligence,* **26**(4)*,* 1366-1372.

Von Gunten, U., Driedger, A., Gallard, H. and Salhi, E. (2001) By-products formation during drinking water disinfection: a tool to assess disinfection efficiency? *Water Research,* **35**(8)*,* 2095-2099.

Walker, T., Stanford, B., Alexander, K., Snyder, S., Khan, S. and Robillot, C. (2016) Critical Control Point Assessment to Quantify Robustness and Reliability of Multiple Treatment Barriers of a DPR Scheme, Water Environment & Reuse Foundation, USA.

Wang, X., Ratnaweera, H., Holm, J. A. and Olsbu, V. (2017) Statistical monitoring and dynamic simulation of a wastewater treatment plant: A combined approach to achieve model predictive control. *Journal of Environmental Management,* **193***,* 1-7.

Weber, P. and Jouffe, L. (2003) Reliability modelling with dynamic bayesian networks. In: 5th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SAFEPROCESS'03), IFAC, Washington, DC, USA, pp. 57-62.

Weber, P., Jouffe, L. and Munteanu, P. (2004) Dynamic Bayesian Networks modelling the dependability of systems with degradations and exogenous constraints. In: 11th IFAC Symposium on Information Control Problems in Manufacturing, INCOM'04 IFAC, Salvador-Bahia, Brazil.

Weber, P., Medina-Oliva, G., Simon, C. and Iung, B. (2012) Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence,* **25**(4)*,* 671-682.

Weir, M. H., Pepe Razzolini, M. T., Rose, J. B. and Masago, Y. (2011) Water reclamation redesign for reducing *Cryptosporidium* risks at a recreational spray park using stochastic models. *Water Research,* **45**(19)*,* 6505-6514.

Wen, Q., Tutuka, C., Keegan, A. and Jin, B. (2009) Fate of pathogenic microorganisms and indicators in secondary activated sludge wastewater treatment plants. *Journal of Environmental Management,* **90**(3)*,* 1442-1447.

WHO (2004) WHO Guidelines for Drinking-Water Quality. Third Ed.*,* World Health Organization, Geneva, Switzerland.

WHO (2005) Water Safety Plans: Managing Drinking-Water Quality from Catchment to Consumer. Water, Sanitation and Health, Protection and the Human Environment, World Health Organization, Geneva, Switzerland.

WHO (2006) Guidelines for the Safe Use of Wastewater, Excreta and Greywater: Policy and Regulatory Aspects. World Health Organization, United Nations Environment Programme, Geneva, Switzerland.

WHO (2011) Guidelines for Drinking-Water Quality. Geneva: World Health Organization.

Williams, J. (1986) HEART– a proposed method for assessing and reducing human error. In: 9th Advances in Reliability Technology Symposium University of Bradford, Bradford.

Witten, I. H. and Frank, E. (2005) Data Mining: Practical Machine Learning Tools and Techniques. Third Ed.*,* Morgan Kaufmann, Massachusetts.

WSAA (2015) Drinking Water Source Assessment and Treatment Requirements. Manual for the Application of Health-Based Treatment Targets. WSA 202—2015-1.2 *Water Services Association of Australia*.

Wu, S., Hrudey, S., French, S., Bedford, T., Soane, E. and Pollard, S. (2009) A role for human reliability analysis (HRA) in preventing drinking water incidents and securing safe drinking water. *Water Research,* **43**(13)*,* 3227-3238.

Xu, P., Janex, M.-L., Savoye, P., Cockx, A. and Lazarova, V. (2002) Wastewater disinfection by ozone: main parameters for process design. *Water Research,* **36**(4)*,* 1043-1055.

Zhu, Z. J. and McBean, E. A. (2007) Selection of water treatment processes using Bayesian decision network analyses. *Journal of Environmental Engineering and Science,* **6**(1)*,* 95-102.

Appendices

A.1 Appendix 1

A.1.1 Naïve and Causal BNs

The following example illustrates the difference between naïve and causal BN models in terms of their structure and information required to populate their contingency probability tables (Figure A-1). Note that these models were constructed only as examples and may not reflect the real system. The same variables and states were defined for both models. For simplicity each node is assumed to have two states. The first model is a naïve Bayes model with *E. coli* log reduction value (LRV) as the variable (or class node) to be predicted or classified through three attribute nodes. These attribute nodes have conditional probabilities associated with each combination of states of the class node and attribute nodes. *E. coli* LRV has a marginal probability, not requiring conditional probabilities as it is a root node. In the second example, the *E. coli* LRV node is a child of three parent nodes. In this case, the model represents a causal network in which causal connections follow the direction of the arcs. None of the parent nodes are assumed to be directly associated with one another, but are viewed as conditionally independent from one another. As it can be seen, the combination of parent nodes states defined a large conditional probability distribution for the *E. coli* LRV node.

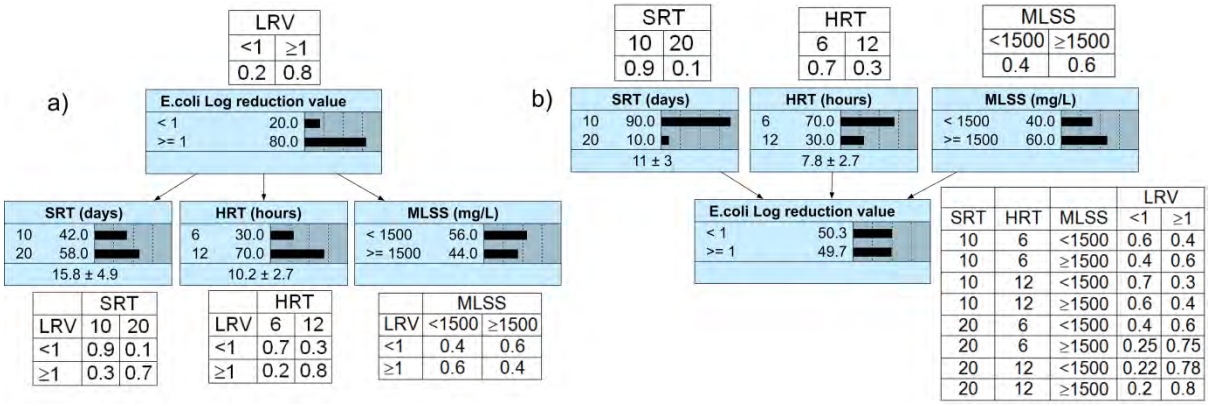

Figure A-1: Non-causal a) and causal Bayesian networks b). Nodes indicate variables and bars indicate the frequencies of each state. Conditional and marginal probabilities are shown in tables.

A.1.2 A.2 Semi-Naïve Bayesian Network Models

The three different SNB models for *C. parvum* (Figure A-2) and *G. lamblia* (Figure A-3) are presented here for comparison. The model nodes are represented as labelled boxes and therefore in this format do not show the probabilities associated to each variable. The TAN(1) and TAN(2) models were obtained with two different search algorithms, whereas the TAN(2) and BAN models were obtained with the same search algorithm but allowing a maximum of two connections between attributes in the case of the TAN (2) models and three in the case of the BAN models.

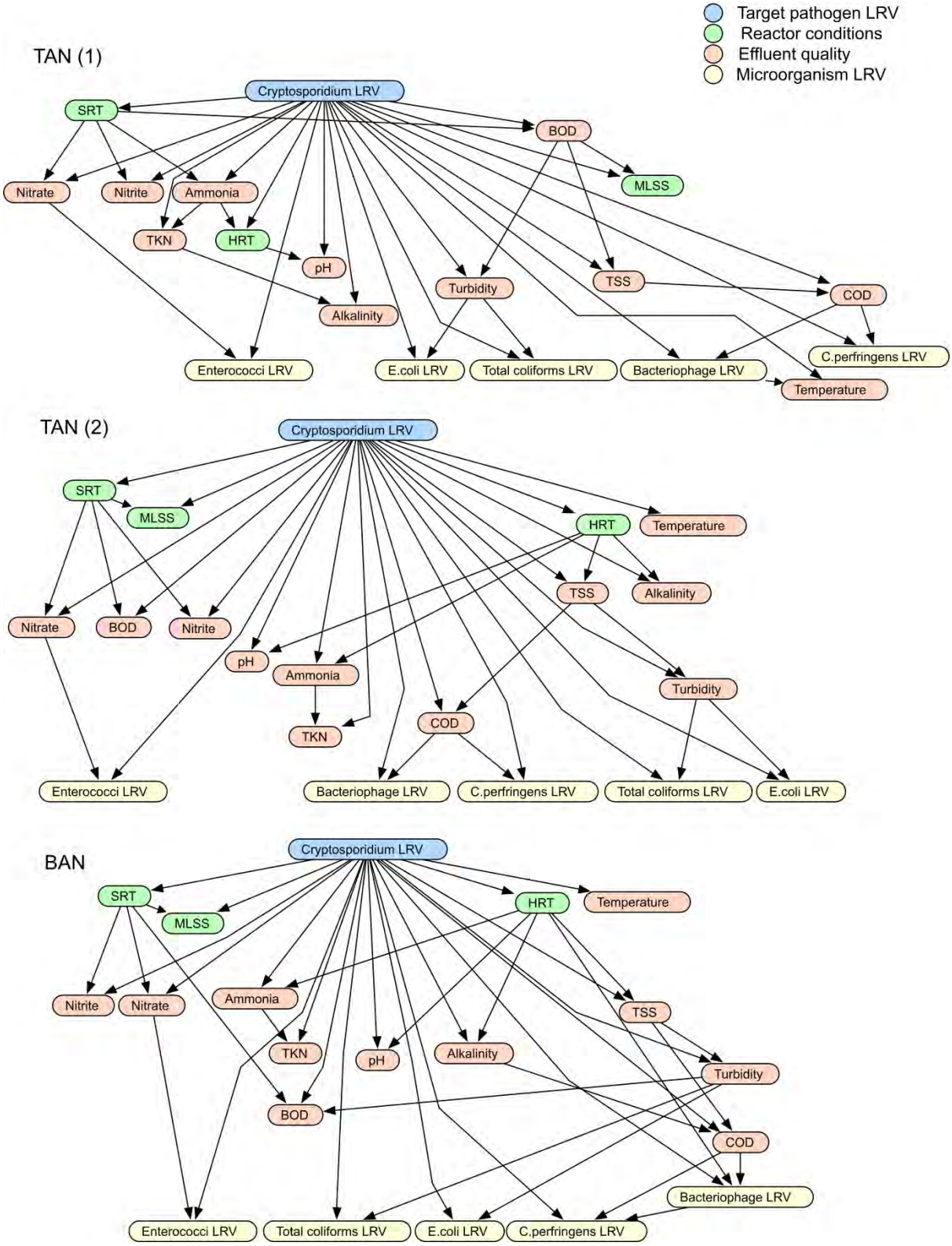

Figure A-2: SNB models for *C. parvum.* Variable categories are presented with distinct colour. Categories include: Reactor parameters in green, physicochemical parameters in pink, microbial indicators in yellow and *C. parvum* LRV in light blue.

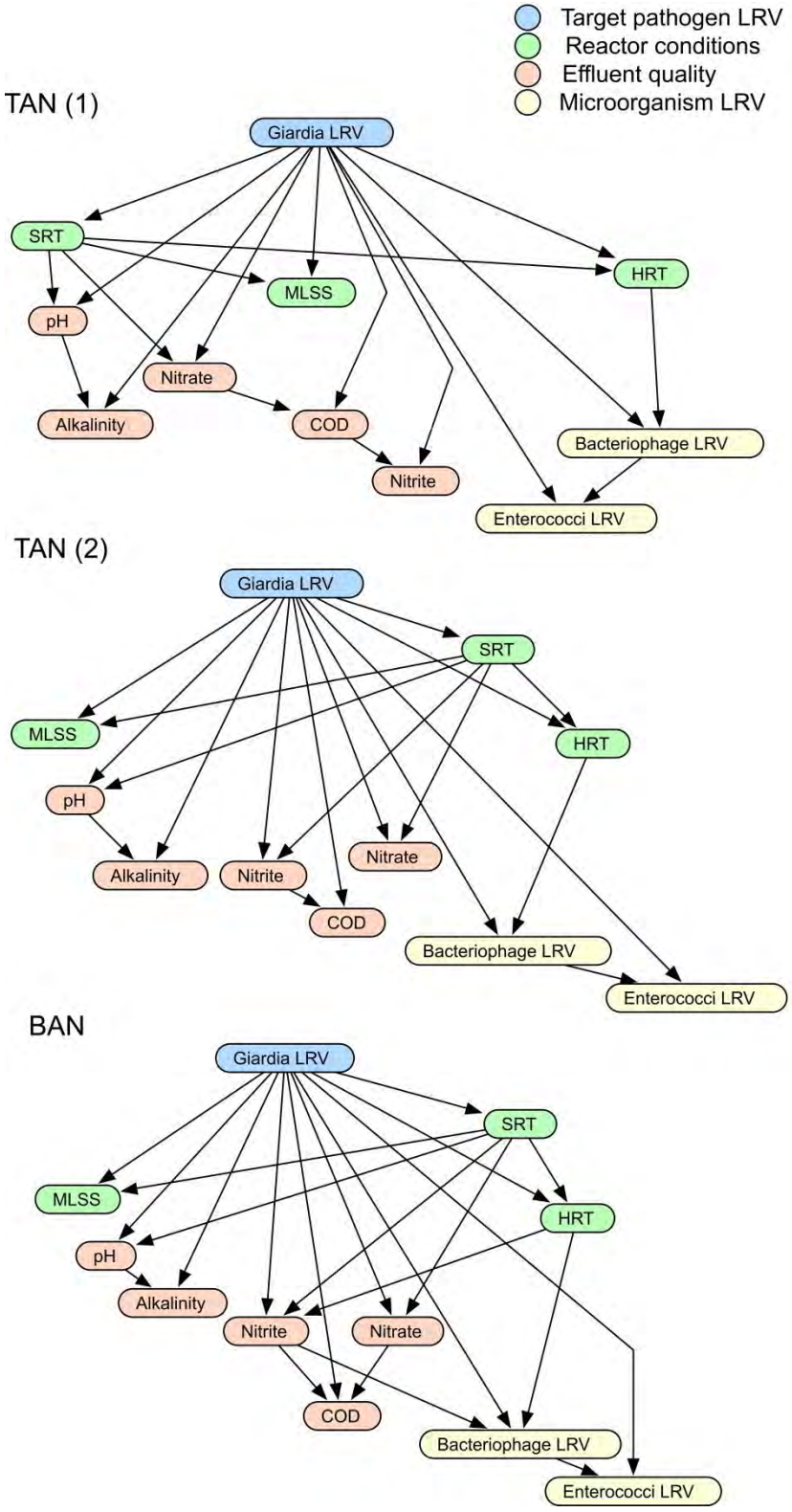

Figure A-3: SNB for *G. lamblia.* Variable categories are presented with distinct colour labels. Categories include: Reactor parameters in green, physicochemical parameters in pink, microbial indicators in yellow and *G.lamblia* LRV in light blue.

A.1.3 Sensitivity Analysis

A one-way sensitivity analysis was performed to investigate the effect of the most important predictors on the variance of the predicted pathogen LRVs as shown in Figure A-4. Consistent with the PA analysis for *C. parvum* (Figure A-4 a), the highest influence was observed for enterococci LRV, followed by total coliforms LRV and *E. coli LRV*. For *G. lamblia* (Figure A-4 b), alkalinity and SRT showed the highest influence.

The sensitivity analysis is illustrated using tornado charts where the changes in the target node for LRV ≥ 1 are obtained by altering the probability of each attribute over the probability space. In the cases shown the BAN *C. parvum* and *G. lamblia* models were analysed as they were the selected optimum models. The line that divides the chart represents the prior probability for the class node which was 0.49 for *C. parvum* and 0.70 for *G. lamblia* approximately. The figures also present the state of the attribute corresponding to the minimum or maximum probabilities. For instance, for the *C. parvum* case, when enterococci LRV is below 1.78 the probability of a LRV higher than 1 would be approximately 0.06. In contrast, when enterococci LRV is above 2.13, the probability of a LRV higher than 1 would be approximately 0.9. It is important to note that effluent water quality parameters included here were in their highest state when the probability of *C. parvum* LRV ≥ 1 was minimal. On the other hand, LRV for enterococci, total coliforms and *E. coli* were on their lowest range state when the probability of LRV ≥ 1 was minimal. A SRT of 15 and an alkalinity in the highest range concentration were associated to a minimal LRV≥ 1 probability for *C. parvum*, whereas these produced the opposite effect on the *G. lamblia* LRV.

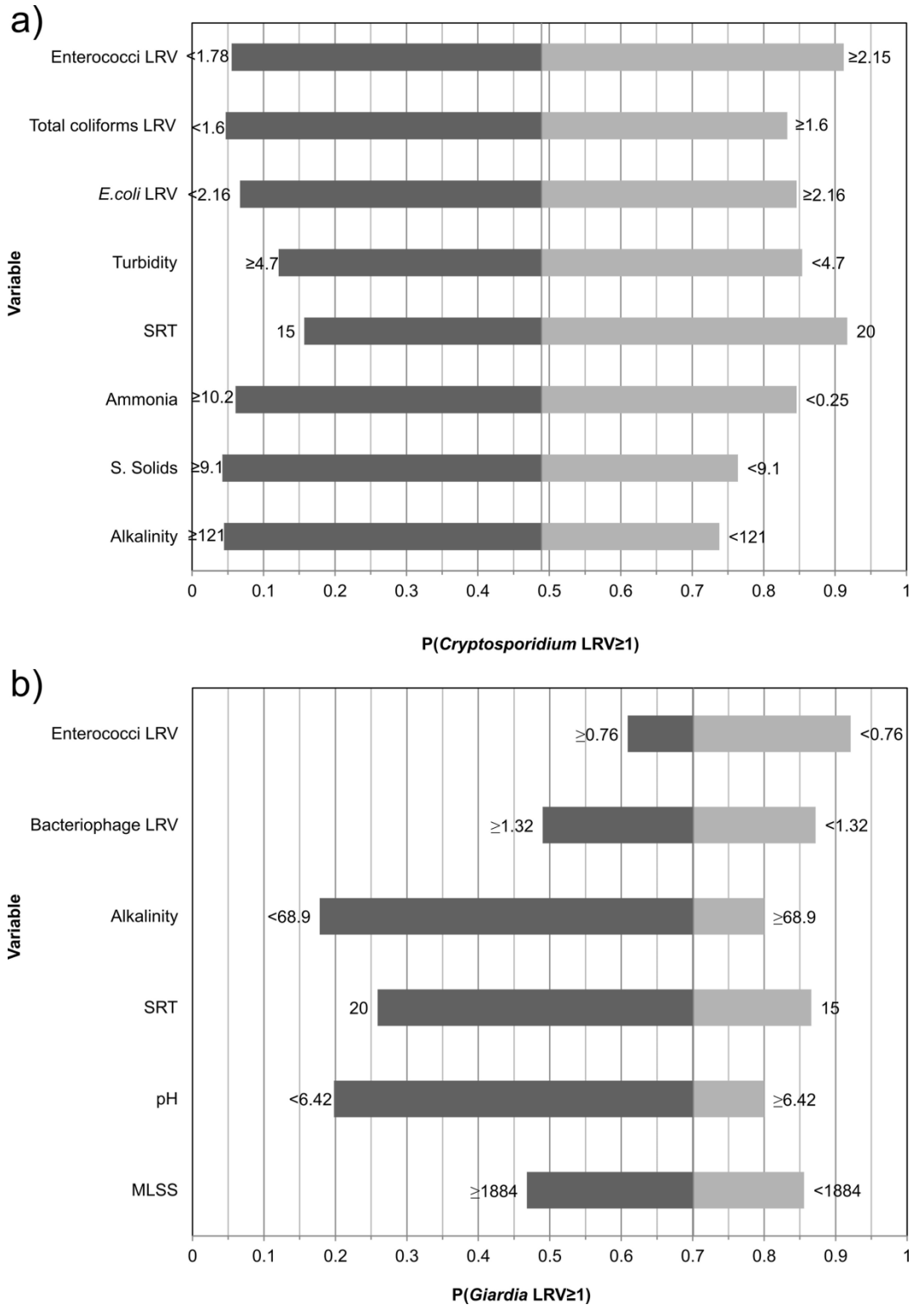

Figure A-4: Sensitivity to findings for removal efficiencies according to most important operating parameters for *C. parvum* BAN model (a) and *G. lamblia* BAN model (b). Bar lengths represent the change in the probability of the state LRV≥1.

Notes:

The vertical line dividing the bars in two in both graphs indicates the prior probability of the class node. Values next to the bars indicate the state of the attribute for the minimum and maximum LRV probabilities.

A.1.4 Order of variables in dataset

K2 algorithm for structure learning depends on a fixed ordering of the variables in the dataset. The ordering used in the present analysis is shown in Table A-1. This arrangement allows the algorithm to search for the best structure by testing the addition of arcs from operating and water quality parameters to microbial indicators and pathogens.

Table A-1: Ordering of variables in datasets for *C. parvum* and *G. lamblia*

Column number	***C. parvum***	***G. lamblia***
1	SRT	SRT
2	HRT	HRT
3	MLSS	MLSS
4	SS	SS
5	pH	pH
6	Temperature	Temperature
7	Turbidity	Turbidity
8	Alkalinity	Alkalinity
9	NH_4^+	NH_4^+
10	NO_2	NO_2
11	NO_3	NO_3

Column number	*C. parvum*	*G. lamblia*
12	TKN	TKN
13	COD	COD
14	BOD_5	BOD_5
15	Bacteriophage	Bacteriophage
16	*C.perfringens* LRV	*C.perfringens* LRV
17	Total coliforms LRV	Total coliforms LRV
18	*E. coli* LRV	*E. coli* LRV
19	Enterococci LRV	Enterococci LRV
20	*G. lamblia* LRV	*C. parvum* LRV
21	*C. parvum* LRV	*G. lamblia* LRV

A.1.5 Individual attributes evaluation

Individual attributes evaluation was analysed through PA and FPR as presented in Table A-2. The results indicate that some of the attributes have good PA but also high FPR, which is not appropriate for this prediction task. From these results there is no clear conclusion about the best predictors. A better outcome can be obtained through the AUC score.

Table A-2: Individual attributes evaluation through prediction accuracy (PA) and false positive rate (FPR) for *C. parvum* and *G. lamblia*. Results are presented as arithmetic mean ± standard deviation

Predictor	***C. parvum***		***G. lamblia***	
	PA	**FPR**	**PA**	**FPR**
No Evidence[b]	51.1±4.87[a]	0.00±0.00[a]	70.7±4.41[a]	1.00±0.00[a]
SRT	89.7±9.51	0.07±0.12	76.0±9.61	0.77±0.28
HRT	70.5±13.2	0.58±0.25	70.7±4.41	1.00±0.00
MLSS	76.1±12.8	0.10±0.15	70.0±13.9	0.44±0.39
Temperature	61.5±12.0	0.12±0.25	-[c]	-
SS	85.2±10.6	0.27±0.19	-	-
Turbidity	90.3±9.52	0.15±0.16	-	-
COD	71.3±13.0	0.56±0.25	63.7±10.1	0.66±0.45
BOD_5	79.8±13.2	0.18±0.20	-	-

Predictor	*C. parvum*		*G. lamblia*	
	PA	**FPR**	**PA**	**FPR**
pH	57.3±10.6	0.51±0.33	82.7±9.66	0.59±0.32
Alkalinity	83.6±11.8	0.32±0.23	81.5±9.89	0.59±0.32
NO_2	70.6±12.3	0.12±0.16	78.4±13.7	0.15±0.25
NO_3	88.7±9.51	0.07±0.12	70.1±11.6	0.78±0.32
NH_4^+	80.0±15.6	0.35±0.29	-	-
TKN	69.1±13.9	0.53±0.26	-	-
Bacteriophage LRV	70.1±13.3	0.53±0.25	68.9±11.3	0.46±0.38
C. perfringens LRV	69.7±13.2	0.54±0.26	-	-
Total coliforms LRV	91.2±8.57	0.15±0.16	-	-
E. coli LRV	90.5±9.79	0.16±0.18	-	-
Enterococci LRV	86.8±9.94	0.08±0.13	69.7±6.96	0.93±0.18

[a] mean PA and FPR based on 10-fold cross validation repeated 10 times

[b] "No evidence" no attributes are considered in the evaluation

[C] attribute was not included in the NB or SNBs

High PA and FPR can be explained with a simple example to show under which conditions these results can be obtained. Suppose 25 cases need to be tested of which 15 correspond to TRUE and 10 to FALSE. The model predicts the cases as shown in Table A-3.

Table A-3: results of prediction for example of high PA and FPR

	Predicted TRUE	**Predicted FALSE**
Actual TRUE	15 (TP)	0 (FN)
Actual FALSE	5 (FP)	5 (TN)

This model would return a false positive rate of 0.5 (FP/(FP+TN)) and a prediction accuracy of 80% ((TP+TN)/(TP+TN+FP+FN)).

A.1.6 Construction of the model without the naïve assumption

The BN model without the naïve assumption was constructed using the bnlearn package in *R*. The following code was used to generate the model using the same automated discretisation algorithm and data used in NB and SNB models. A blacklist (i.e. restricted arcs) (Table A-4) was also incorporated to account for arcs that should not be present in the network. An automated structure-learning algorithm (hill-climbing) was used with AIC score to find the BN structure.

```
#Load bnlearn and RWeka packages
library(bnlearn)
library(RWeka)

#Import data and blacklist

datalearning = read.csv("data.csv")
blacklist = read.csv("blacklist.csv")

#Discretise data automatically using WEKA algorithm
#Microbiallrv is the target microorganisms already discretised

discretedata= Discretize(microbiallrv~.,data=datalearning)

#Remove columns with one state
n.col=ncol(discretedata)
for (i in 1:n.col){

  if (toString(discretedata[1,i])=="'All'"){

    discretedata = discretedata[,-i]

  }else{

  }
  }

#Create network using the discrete data and blacklist as inputs of the
hill-climbing algorithm

dag1 = hc(discretedata, score="aic",blacklist=blacklist)

#Learn parameters of the network
bn1 = bn.fit(dag1,discretedata)

#Export network in Netica's compatible format
write.dsc(bn1,file="BN.dsc")
```

Table A-4: Blacklist used for the hill-climbing structure learning algorithm

from	to	from	to
clostrilrv	SRT	enterococcilrv	NH3out
clostrilrv	HRT	enterococcilrv	NO2out
clostrilrv	SSML	enterococcilrv	NO3out
clostrilrv	SSout	enterococcilrv	TKNout
clostrilrv	pHout	enterococcilrv	CODout
clostrilrv	temperatureout	enterococcilrv	BODout
clostrilrv	turbidityout	cryptoconlrv	SRT
clostrilrv	alkalinityout	cryptoconlrv	HRT
clostrilrv	NH3out	cryptoconlrv	SSML
clostrilrv	NO2out	cryptoconlrv	SSout
clostrilrv	NO3out	cryptoconlrv	pHout
clostrilrv	TKNout	cryptoconlrv	temperatureout

from	to	from	to
clostrilrv	CODout	cryptoconlrv	turbidityout
clostrilrv	BODout	cryptoconlrv	alkalinityout
coliformlrv	SRT	cryptoconlrv	NH3out
coliformlrv	HRT	cryptoconlrv	NO2out
coliformlrv	SSML	cryptoconlrv	NO3out
coliformlrv	SSout	cryptoconlrv	TKNout
coliformlrv	pHout	cryptoconlrv	CODout
coliformlrv	temperatureout	cryptoconlrv	BODout
coliformlrv	turbidityout	coliformlrv	clostrilrv
coliformlrv	alkalinityout	ecolilrv	clostrilrv
coliformlrv	NH3out	enterococcilrv	clostrilrv
coliformlrv	NO2out	cryptoconlrv	clostrilrv
coliformlrv	NO3out	phagelrv	clostrilrv
coliformlrv	TKNout	clostrilrv	coliformlrv
coliformlrv	CODout	ecolilrv	coliformlrv
coliformlrv	BODout	enterococcilrv	coliformlrv
phagelrv	SRT	cryptoconlrv	coliformlrv
phagelrv	HRT	phagelrv	coliformlrv
phagelrv	SSML	clostrilrv	ecolilrv
phagelrv	SSout	coliformlrv	ecolilrv
phagelrv	pHout	enterococcilrv	ecolilrv
phagelrv	temperatureout	cryptoconlrv	ecolilrv
phagelrv	turbidityout	phagelrv	ecolilrv
phagelrv	alkalinityout	clostrilrv	enterococcilrv
phagelrv	NH3out	coliformlrv	enterococcilrv
phagelrv	NO2out	ecolilrv	enterococcilrv
phagelrv	NO3out	cryptoconlrv	enterococcilrv
phagelrv	TKNout	phagelrv	enterococcilrv
phagelrv	CODout	clostrilrv	phagelrv
phagelrv	BODout	coliformlrv	phagelrv
ecolilrv	SRT	ecolilrv	phagelrv
ecolilrv	HRT	cryptoconlrv	phagelrv
ecolilrv	SSML	enterococcilrv	phagelrv
ecolilrv	SSout	clostrilrv	cryptoconlrv
ecolilrv	pHout	coliformlrv	cryptoconlrv
ecolilrv	temperatureout	ecolilrv	cryptoconlrv
ecolilrv	turbidityout	enterococcilrv	cryptoconlrv
ecolilrv	alkalinityout	phagelrv	cryptoconlrv
ecolilrv	NH3out	HRT	SRT
ecolilrv	NO2out	SSML	SRT
ecolilrv	NO3out	SSout	SRT
ecolilrv	TKNout	pHout	SRT
ecolilrv	CODout	temperatureout	SRT
ecolilrv	BODout	turbidityout	SRT
enterococcilrv	SRT	alkalinityout	SRT

from	to	from	to
enterococcilrv	HRT	NH3out	SRT
enterococcilrv	SSML	NO2out	SRT
enterococcilrv	SSout	NO3out	SRT
enterococcilrv	pHout	TKNout	SRT
enterococcilrv	temperatureout	CODout	SRT
enterococcilrv	turbidityout	BODout	SRT
enterococcilrv	alkalinityout	SRT	HRT
enterococcilrv	NH3out	SSML	HRT
enterococcilrv	NO2out	SSout	HRT
enterococcilrv	NO3out	pHout	HRT
enterococcilrv	TKNout	temperatureout	HRT
enterococcilrv	CODout	turbidityout	HRT
enterococcilrv	BODout	alkalinityout	HRT
enterococcilrv	SRT	NH3out	HRT
enterococcilrv	HRT	NO2out	HRT
enterococcilrv	SSML	NO3out	HRT
enterococcilrv	SSout	TKNout	HRT
enterococcilrv	pHout	CODout	HRT
enterococcilrv	temperatureout	BODout	HRT
enterococcilrv	turbidityout		
enterococcilrv	alkalinityout		

A.2 Appendix 2

A.2.1 *R* code for model construction

The various models tested in this study were constructed in Jags using the following codes. The first code shows the model specification for the fully hierarchical model, arrangement of data and model initialisation. The code for the other model alternatives only indicates the structure for the model module specified by "model{…}". The model used the offset method (Lunn et al. 2012) to account for the different size of data in 2010 and 2015.

A.2.1.1 Fully hierarchical validation model with exchangeable parameters for μ and σ

```
#Call the libraries

library(reshape2)

library(R2jags)
```

```
modelstring="

model{


for (k in 1:2){ #k is the year analysed

for (i in 1:offset[k]){ #i is the observation analysed


y[i,k]~dnorm(mu.dot[skid[i,k],k],tau.dot[skid[i,k],k])  #Likelihood

ycen[i,k] ~ dinterval(y[i,k],ycenlim[i,k])  #Incorporation of censored
values into the likelihood




}

}


for (j in skids.num){ #Skid identification 2,4,5 and 6


for (k in 1:2){


#Conditional structure for the mean depending on the year

mu.dot[j,k]<-ifelse(k==1,mu[j],mu[j]+delta[j])


#Conditional structure for the standard deviation depending on the
year

log.sigma.dot[j,k]<-ifelse(k==1,log.sigma[j],log.sigma[j]+epsilon[j])


#Transforming standard deviation to precision

tau.dot[j,k]<-1/(sigma.dot[j,k])^2


log(sigma.dot[j,k]) <-log.sigma.dot[j,k]
```

```
}
#Skid specific means
mu[j]~dnorm(mu.mu,tau.mu)

#Difference between skids' means
delta[j]~dnorm(mu.delta,tau.delta)

#Skid specific standard deviation
log.sigma[j]~dnorm(mu.sigma,tau.sigma)
log(sigma[j])<-log.sigma[j]

#Difference between skids' standard deviations
epsilon[j]~dnorm(mu.epsilon,tau.epsilon)









}


  #Transforming precisions to standard deviation for inspection
  sigma.mu<-1/sqrt(tau.mu)
  sigma.delta<-1/sqrt(tau.delta)
  sigma.sigma<-1/sqrt(tau.sigma)
  sigma.epsilon<-1/sqrt(tau.epsilon)

  #Definition of prior distributions
  mu.mu~dnorm(0,0.0001)
  tau.mu~dgamma(0.001,0.001)
  mu.delta~dnorm(0,0.0001)
```

```
  tau.delta~dgamma(0.001,0.001)
  mu.sigma~dnorm(0,0.0001)
  tau.sigma~dgamma(0.001,0.001)
  mu.epsilon~dnorm(0,0.0001)
  tau.epsilon~dgamma(0.001,0.001)


}
"
writeLines( modelstring , con="modellrv" )



#Data processing

data<- read.csv("data.csv")

# Preparing data for model with offsets method
data2010 <-subset(data,year==2010 & unit<=6)
offset.1<-dim(data2010)[1]
data2010.sub <-data.matrix(data2010[,"lrv"])
data2010.iscen<-data.matrix(data2010[,"isCensored"])
data2010.cenlim<-data.matrix(data2010[,"censorLimitVec"])
data2010.unit<-data.matrix(data2010[,"unit"])

data2015 <-subset(data,year==2015 & unit<=6)
offset.2<-dim(data2015)[1]
data2015.sub <-data.matrix(data2015[,"lrv"])
data2015.iscen<-data.matrix(data2015[,"isCensored"])
data2015.cenlim<-data.matrix(data2015[,"censorLimitVec"])
```

```
data2015.unit<-data.matrix(data2015[,"unit"])



add.col<-function(df, new.col) {n.row<-dim(df)[1]
length(new.col)<-n.row
cbind(df, new.col)
}


y<-data.frame(add.col(data2010.sub,data2015.sub))
names(y)<-c("year1","year2")


ycen<-data.frame(add.col(data2010.iscen,data2015.iscen))
names(ycen)<-c("year1","year2")


ycenlim<-data.frame(add.col(data2010.cenlim,data2015.cenlim))
names(ycenlim)<-c("year1","year2")


skid<-data.frame(add.col(data2010.unit,data2015.unit))
names(skid)<-c("year1","year2")
skids.num<-unique(unique(subset(data, unit<=6))$unit)



offset<-c(offset.1,offset.2)


datalist                                                                =
list(y=y,ycen=ycen,ycenlim=ycenlim,skid=skid,skids.num=skids.num,offse
t=offset) #Data to be used in the model


#Running the model
```

```
parameters                                                              =
c("mu.dot","mu","mu.mu","sigma.mu","delta","mu.delta","sigma.delta","s
igma.dot",


"sigma","epsilon","log.sigma","log.sigma.dot","tau.mu","tau.delta","ta
u.dot", "mu.sigma","sigma.sigma",

               "mu.epsilon","sigma.epsilon")  #  The  parameters  to  be
monitored







burnin = 1000            # Number of steps to burn-in the chains

n.chains = 2               # Number of chains

n.savedsteps=1000        #Number of steps to be saved

thinning=10              #Thinning rate

nIter = ceiling( ( n.savedsteps * thinning) / n.chains )   #Number of
iterations to run



set.seed(123)  #Set seed



#Running the chains

jagsfit.p <-jags(data=datalist, parameters.to.save=parameters,

                 n.iter=nIter,
model.file="modellrv",n.chains=n.chains,n.burnin                        =
burnin,n.thin=thinning)
```

A.2.1.2 Partially hierarchical validation model with exchangeable parameters for μ

```
model{



for (k in 1:2){ #k is the year analysed

for (i in 1:offset[k]){ #i is the observation analysed
```

```
y[i,k]~dnorm(mu.dot[skid[i,k],k],tau.dot[k])  #Likelihood

ycen[i,k] ~ dinterval(y[i,k],ycenlim[i,k])  #Incorporation of censored
values into the likelihood




}

}


for (j in skids.num){ #Skid identification 2,4,5 and 6


for (k in 1:2){


#Conditional structure for the mean depending on the year

mu.dot[j,k]<-ifelse(k==1,mu[j],mu[j]+delta[j])


}

#Skid specific means

mu[j]~dnorm(mu.mu,tau.mu)


#Difference between skids' means

delta[j]~dnorm(mu.delta,tau.delta)


}


  for (k in 1:2){


#Conditional structure for the standard deviation depending on the
year

log.sigma.dot[k]<-ifelse(k==1,log.sigma,log.sigma+epsilon)


#Transforming standard deviation to precision
```

```
tau.dot[k]<-1/(sigma.dot[k])^2

log(sigma.dot[k]) <-log.sigma.dot[k]


  }


  log(sigma)<-log.sigma

  #Transforming precisions to standard deviation for inspection
  sigma.mu<-1/sqrt(tau.mu)
  sigma.delta<-1/sqrt(tau.delta)

  #Definition of prior distributions
  mu.mu~dnorm(0,0.0001)
  tau.mu~dgamma(0.001,0.001)
  mu.delta~dnorm(0,0.0001)
  tau.delta~dgamma(0.001,0.001)
  log.sigma~dnorm(0,0.0001)
  epsilon~dnorm(0,0.0001)


}
```

A.2.1.3 Partially hierarchical validation model with exchangeable parameters for σ

```
model{


for (k in 1:2){ #k is the year analysed

for (i in 1:offset[k]){ #i is the observation analysed


y[i,k]~dnorm(mu.dot[k],tau.dot[skid[i,k],k])  #Likelihood

ycen[i,k] ~ dinterval(y[i,k],ycenlim[i,k])  #Incorporation of censored
values into the likelihood




}

}


for (j in skids.num){ #Skid identification 2,4,5 and 6


for (k in 1:2){


#Conditional structure for the standard deviation depending on the
year


log.sigma.dot[j,k]<-ifelse(k==1,log.sigma[j],log.sigma[j]+epsilon[j])


#Transforming standard deviation to precision

tau.dot[j,k]<-1/(sigma.dot[j,k])^2


log(sigma.dot[j,k]) <-log.sigma.dot[j,k]

}


log.sigma[j]~dnorm(mu.sigma,tau.sigma)
```

```
log(sigma[j])<-log.sigma[j]

#Difference between skids' standard deviations
epsilon[j]~dnorm(mu.epsilon,tau.epsilon)

}

  for (k in 1:2){

#Conditional structure for the mean depending on the year
mu.dot[k]<-ifelse(k==1,mu,mu+delta)

  }

  #Definition of prior distributions
  mu.sigma~dnorm(0,0.0001)
  tau.sigma~dgamma(0.001,0.001)
  mu.epsilon~dnorm(0,0.0001)
  tau.epsilon~dgamma(0.001,0.001)
  mu~dnorm(0,0.0001)
  delta~dnorm(0,0.0001)

}
```

A.2.1.4 Non-hierarchical model with identical parameters μ and σ for all skids

```
model{


for (k in 1:2){ #k is the year analysed

for (i in 1:offset[k]){ #i is the observation analysed


y[i,k]~dnorm(mu.dot[k],tau.dot[k])  #Likelihood

ycen[i,k] ~ dinterval(y[i,k],ycenlim[i,k])  #Incorporation of censored
values into the likelihood




}

}

for (k in 1:2){


#Conditional structure for the mean depending on the year

mu.dot[k]<-ifelse(k==1,mu,mu+delta)


#Conditional structure for the standard deviation depending on the
year

log.sigma.dot[k]<-ifelse(k==1,log.sigma,log.sigma+epsilon)




tau.dot[k]<-1/(sigma.dot[k])^2


log(sigma.dot[k]) <-log.sigma.dot[k]


y.sim[k]~dnorm(mu.dot[k],tau.dot[k])


}
```

```
log(sigma)<-log.sigma

#Priors

mu~dnorm(0,0.0001)

delta~dnorm(0,0.0001)

log.sigma~dnorm(0,0.0001)

epsilon~dnorm(0,0.0001)




}
```

A.2.1.5 Independent validation model with independent parameters μ and σ for each skid

```
model{


for (k in 1:2){ #k is the year analysed

for (i in 1:offset[k]){ #i is the observation analysed


y[i,k]~dnorm(mu.dot[skid[i,k],k],tau.dot[skid[i,k],k])  #Likelihood

ycen[i,k] ~ dinterval(y[i,k],ycenlim[i,k])  #Incorporation of censored
values into the likelihood




}

}


for (j in skids.num){ #Skid identification 2,4,5 and 6


for (k in 1:2){


#Conditional structure for the mean depending on the year

mu.dot[j,k]<-ifelse(k==1,mu[j],mu[j]+delta[j])


#Conditional structure for the standard deviation depending on the
year

log.sigma.dot[j,k]<-ifelse(k==1,log.sigma[j],log.sigma[j]+epsilon[j])




tau.dot[j,k]<-1/(sigma.dot[j,k])^2


log(sigma.dot[j,k]) <-log.sigma.dot[j,k]
```

```
}
#Skid specific means
mu[j]~dnorm(0,0.0001)

#Difference between skids' means
delta[j]~dnorm(0,0.0001)

log.sigma[j]~dnorm(0,0.0001)

epsilon[j]~dnorm(0,0.0001)

log(sigma[j])<-log.sigma[j]

}

}
```

A.2.2 Distributions for the differences in performance between years for the means (δ_j)and for the logarithm of the standard deviations (ε)

The model provided information for the change in performance means between years for each skid (Figure A-5). The results indicated that there was a decrease in performance across all the skids with a larger change observed on skid 4.

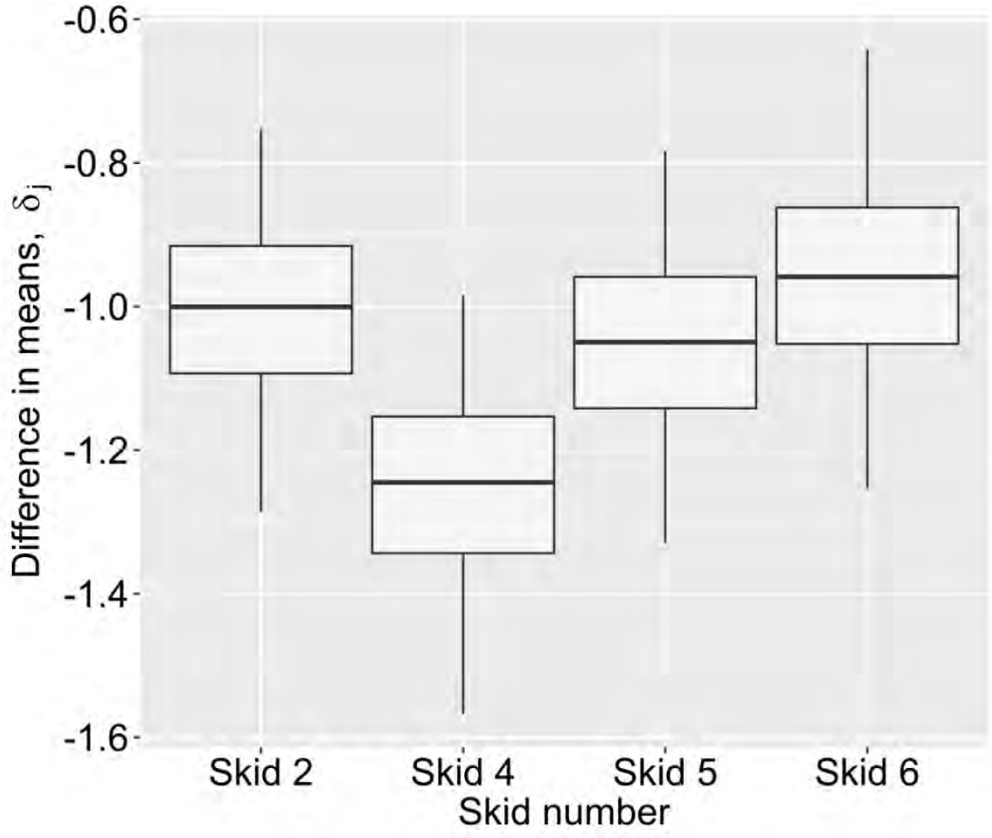

Figure A-5: Box plot for the difference in means between years for each skid (δ_j)

The model also provided information about the change in the log standard deviations between the years (ε) (Figure A-6). The results show that there was a decrease in performance variability after 5 years of operation.

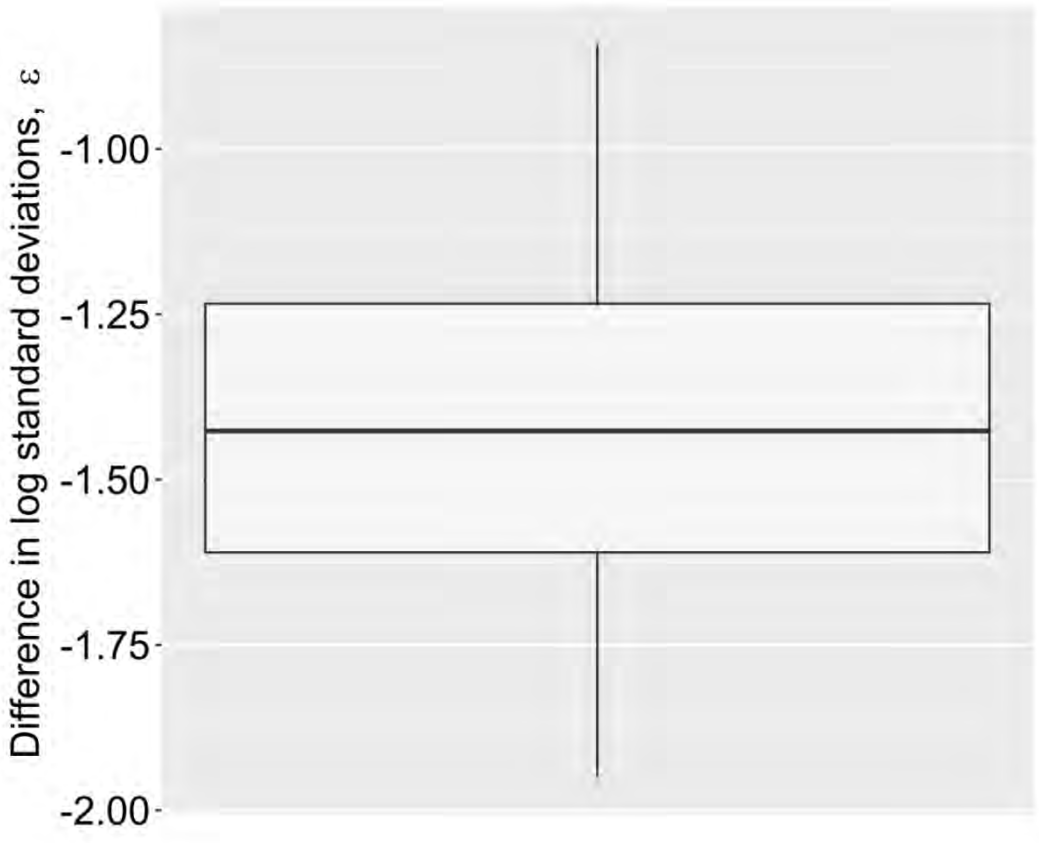

Figure A-6: Box plot for the difference in log standard deviation between years (ε)

A.2.3 Bayesian model development

The Bayesian model analysis can be represented by five distinct steps summarised in Figure A-7. In step 1 the model structure and distributions were defined. In Step 2, the priors were defined for the unknown stochastic model parameters for which posterior distributions would be obtained. In step 3, data were collected and arranged in a suitable format to be incorporated into the model. In Step 4, the MCMC technique was employed to approximate the posterior distributions using the information from steps 1-3. Finally, in step 5 simulated values of the updated parameters were used to generate distributions of future data points, which could then be used for prediction purposes.

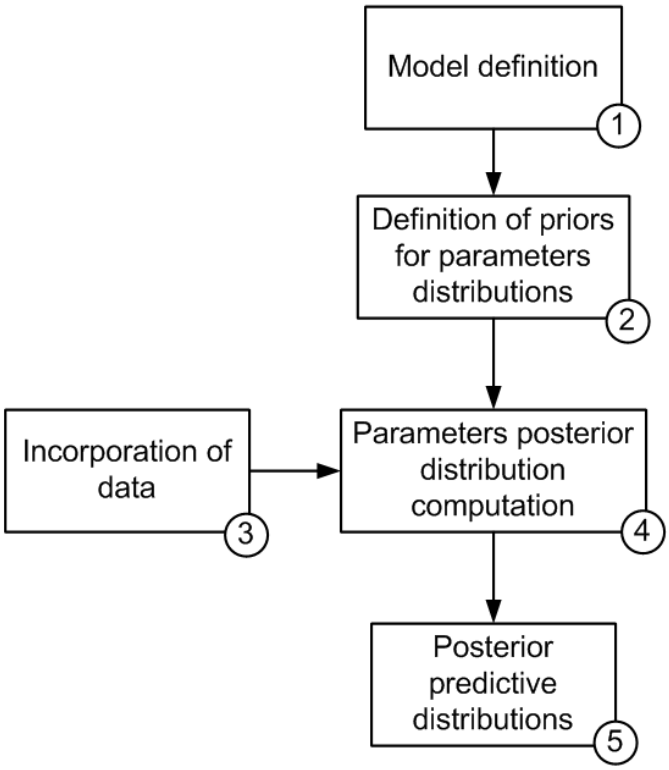

Figure A-7: Construction of posterior and predictive distributions under the Bayesian framework.

A.2.4 Alternative hierarchical and non-hierarchical models

Directed acyclic graphs with the different model options analysed in this study are presented in Figure A-8, Figure A-9, Figure A-10 and Figure A-11.

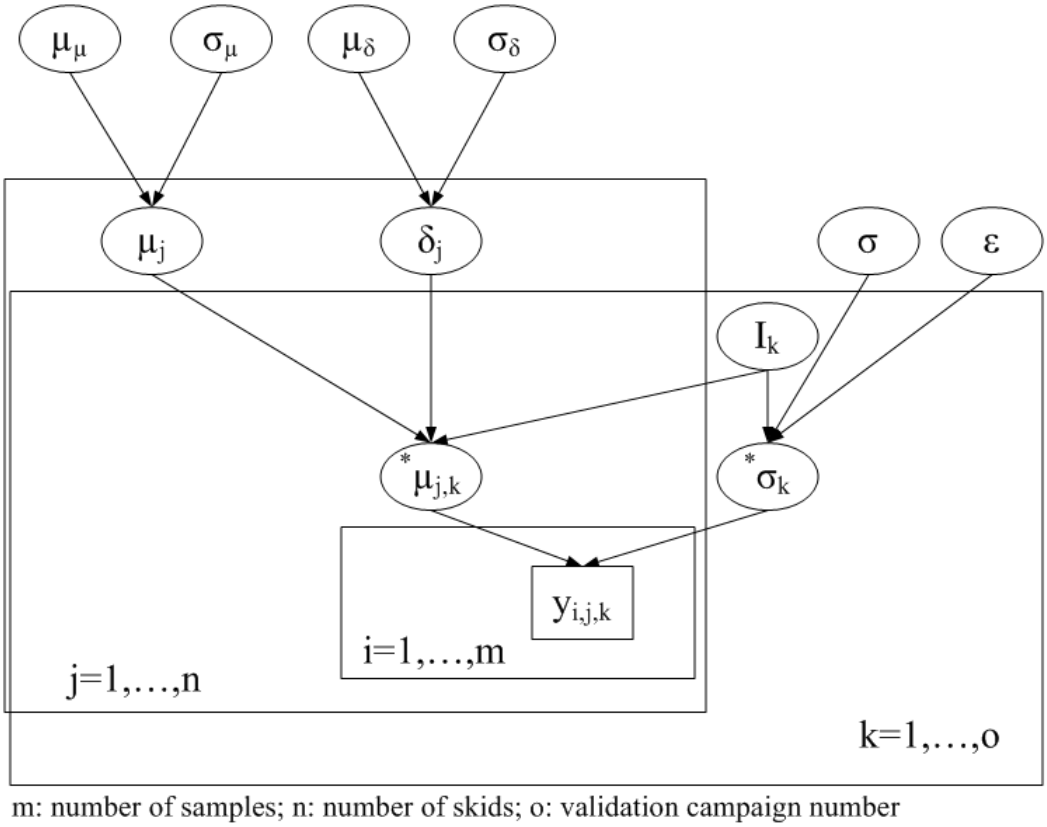

Figure A-8: Partially hierarchical validation model with exchangeable parameters for μ.

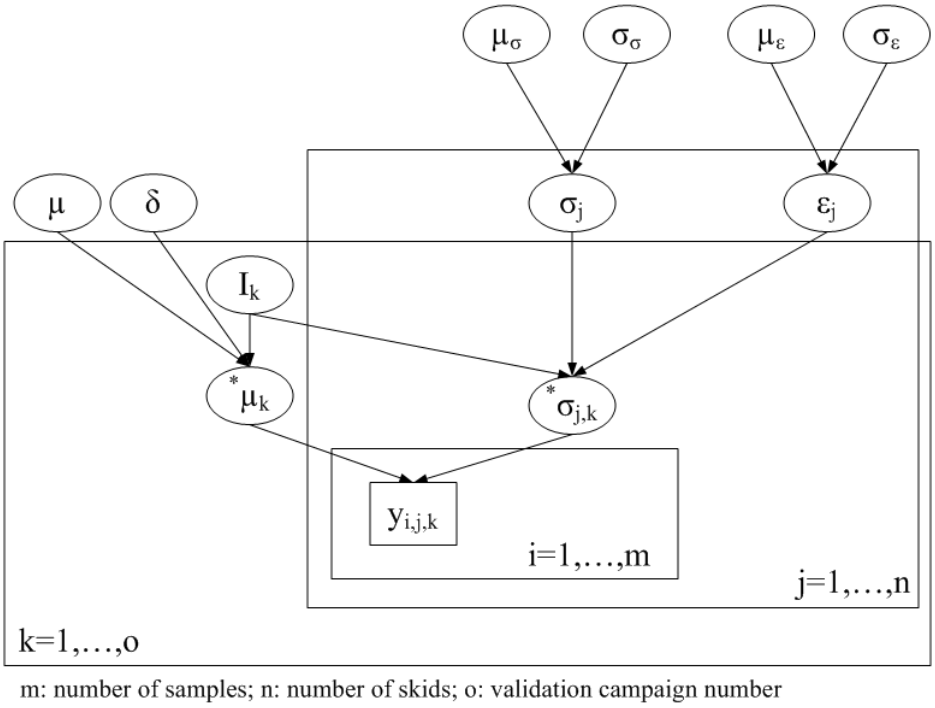

Figure A-9: Partially hierarchical validation model with exchangeable parameters for σ.

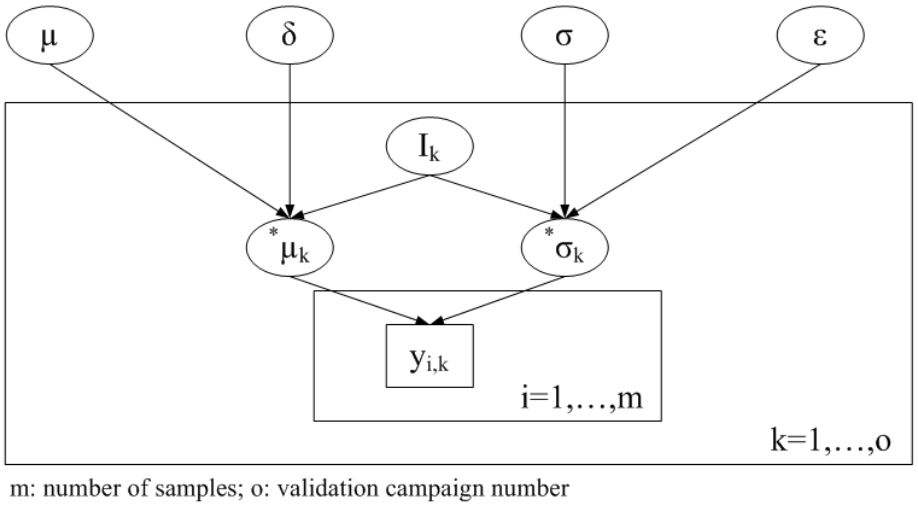

Figure A-10: Non-hierarchical validation model with identical parameters μ and σ for all skids.

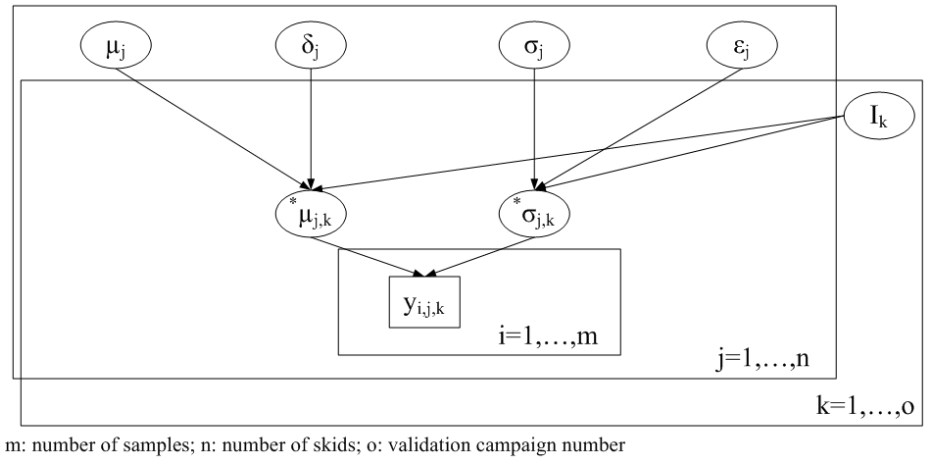

Figure A-11: Independent validation model with independent parameters μ and σ for each skid.

A.2.5 Normality assumption check

Normality was tested through Minitab 16 (Minitab 2010) using probability plots generated from the Reliability/survival option for right censored data. Probability plots show the observations values vs. the percentage of values in the sample that are less than or equal to it. Correlation ("Corr") on Figure A-12 (2010) and Figure A-13 (2015) indicates the Pearson's correlation between the fitted normal distribution and the actual data. The results indicated that the normal distribution is supported by the data with correlation higher than 0.92 for both years.

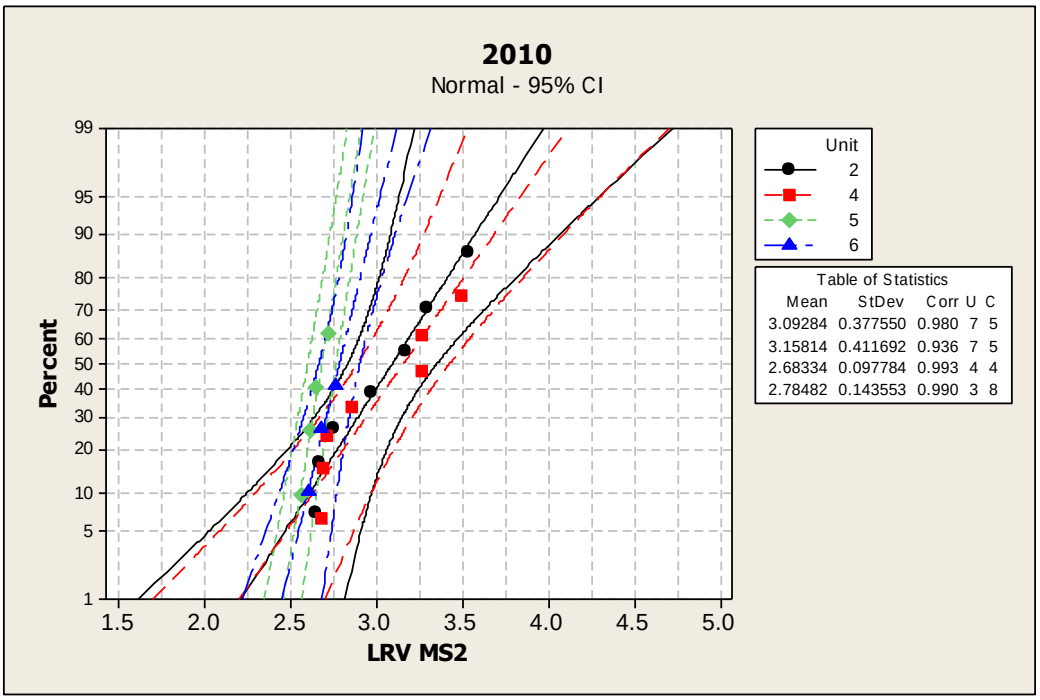

Figure A-12: Probability plot for 2010 data considering censored values. Mean is the average of the observations, StDev is the standard deviation of the observations, "Corr" is the correlation, "U" is the number of uncensored observations and "C" is the number of censored observations.

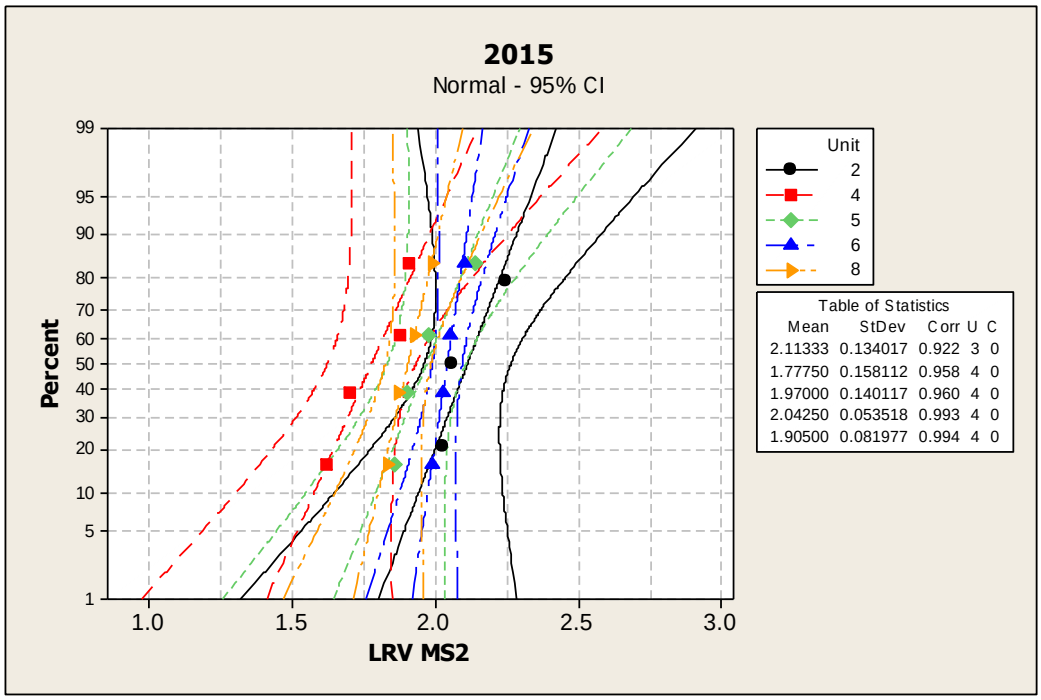

Figure A-13: Probability plot for 2015 data considering censored values. Mean is the average of the observations, StDev is the standard deviation of the observations, “Corr” is the correlation, “U” is the number of uncensored observations and “C” is the number of censored observations.

A.3 Appendix 3

A.3.1 Ozonation experimental results

The following plots show information regarding the ozonation experiments in terms of the inlet ozone in the gas stream (Figure A-14), outlet ozone in the gas stream (Figure A-15), applied ozone dose (Figure A-16), and consumed ozone dose (Figure A-16). The first run of experiments was also considered for the calculation of the mean values. Although the set-points of applied ozone for this first run were slightly lower than the rest of the experiments, the information was considered appropriate as the plots were used to present and compare the relative behaviour of the reactor conditions during each set of experiments. As observed in Figure A-14, the inlet ozone concentration experiences a sharp drop after the system reaches the applied ozone set-point which is produced by the reactor control system automatically stopping the ozone production.

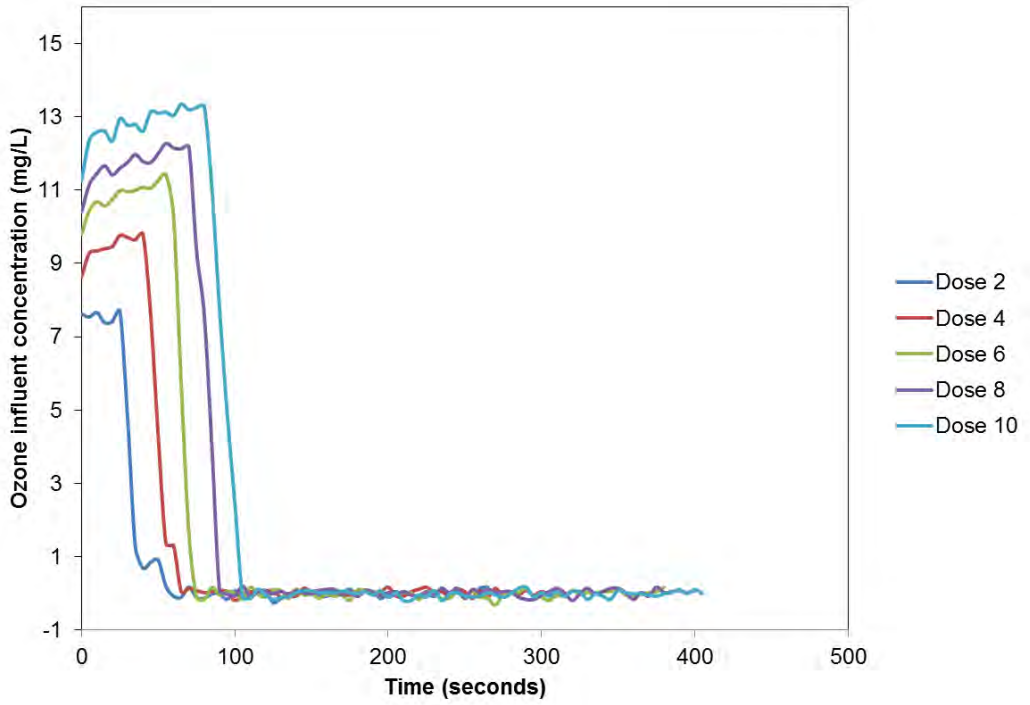

Figure A-14: Average inlet ozone concentration in the gas stream over time

Ozone in the outlet gas stream (Figure A-15) showed a gradual increase and decrease with a single peak. This behaviour was experience because the non-reacting ozone took some time to be stripped from the reactor. Each experiment run was stopped once the outlet ozone concentration was 0.01 mg/L as indicated by the sensor.

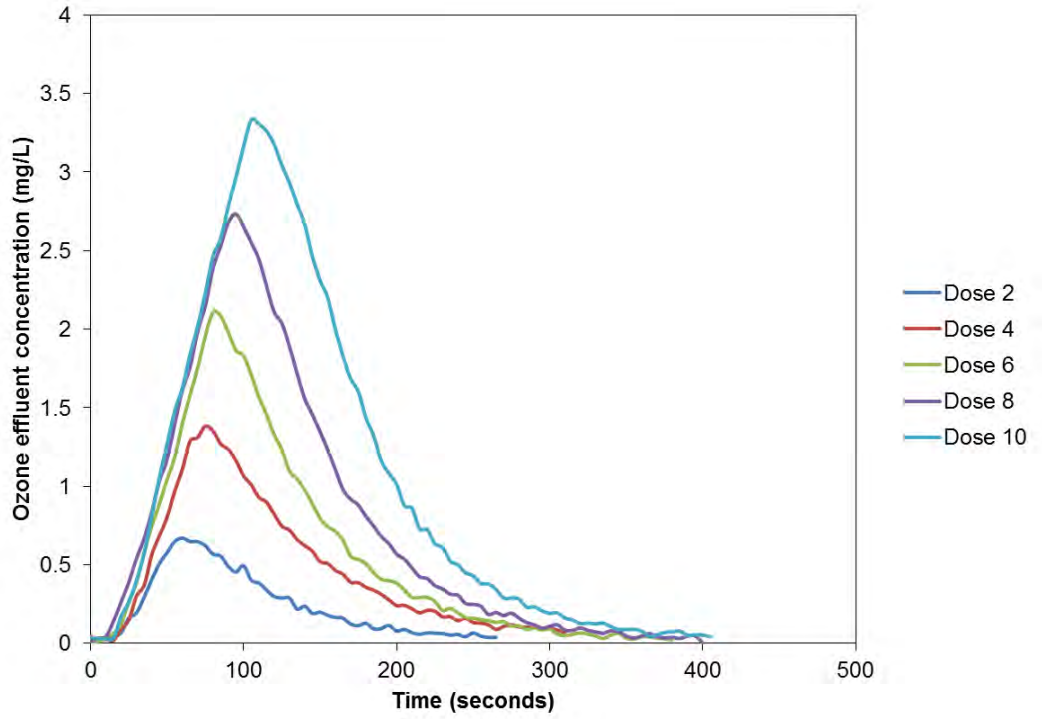

Figure A-15: Average outlet ozone concentration in the gas stream over time

The applied ozone obtained from Equation 6-1 is presented in Figure A-16. As observed, the applied ozone increased steadily until reaching the set-point dose. Once the set-point was reached it did not increase further because the ozone generation was ceased.

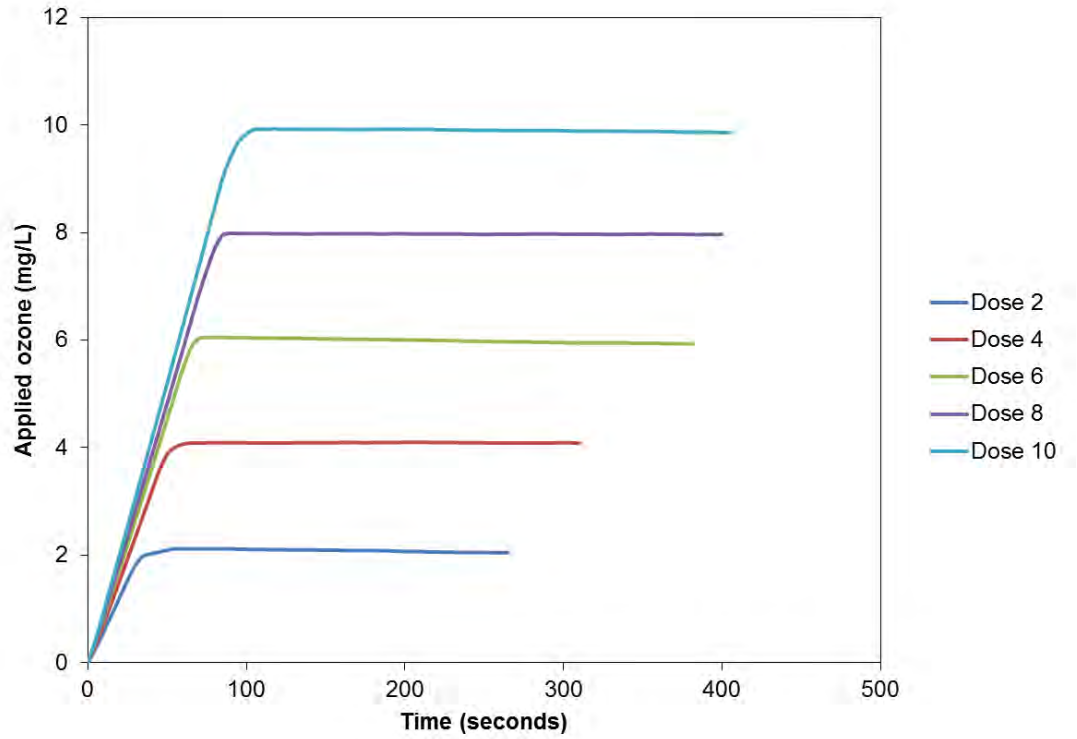

Figure A-16: Average applied ozone over time

The consumed ozone calculated from Equation 6-2 is shown in Figure A-17. The consumed ozone increased linearly until reaching a peak to then decreasing gradually. The final value was used to calculate the ratios of ozone to DOC. As observed, the consumed ozone changed over time. This was expected as the non-reacting ozone was removed from the system gradually.

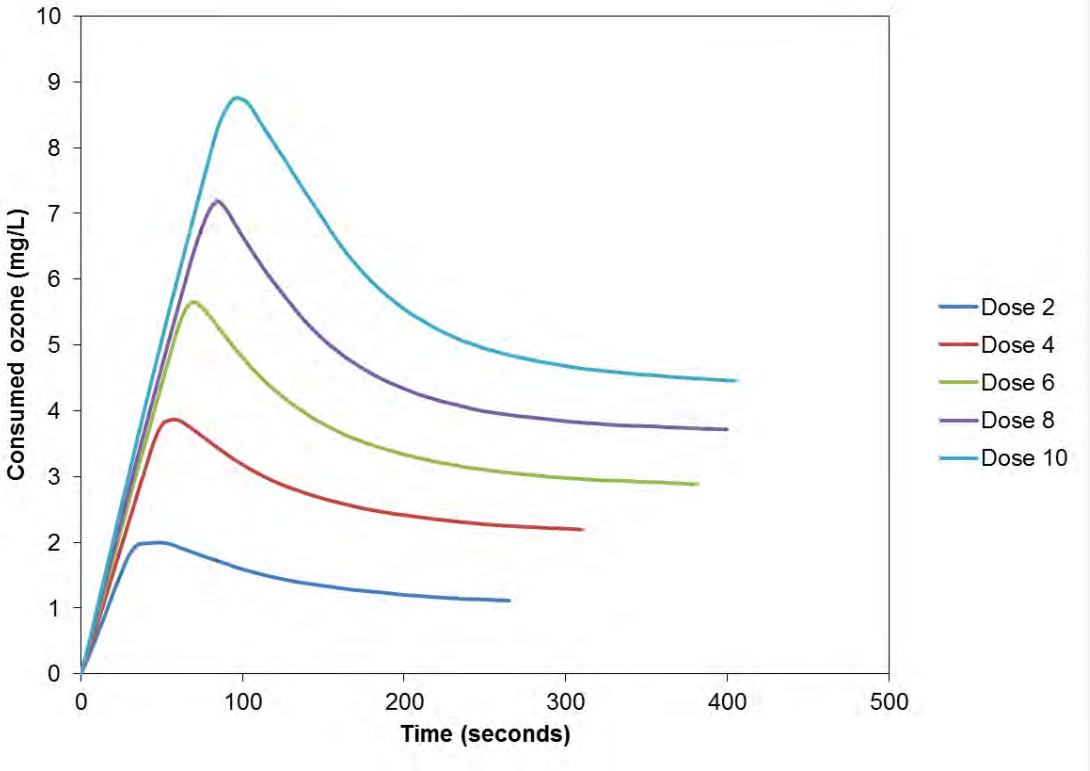

Figure A-17: Average consumed ozone over time

For the calculation of the consumed ozone to DOC ratios, the consumed ozone and DOC concentrations in Table A-5 were employed. As observed the DOC values ranged between 6 and 14 mg/L approximately indicating certain variability in the organic concentrations. The consumed ozone was much lower than the applied ozone. This indicated that only some proportion of the injected ozone reacted. The ratio of the consumed ozone to applied ozone was approximately 0.5.

Table A-5: DOC influent concentrations, applied and consumed ozone for the five samples used during the experiments

Sample	DOC (mg/L)	Applied ozone (mg/L)	Consumed ozone (mg/L)
Sample 1	9.2	1.5; 3.3; 5.0; 6.6; 8.2	0.9; 1.7; 2.5; 3.3; 3.9
Sample 2	14	2.1; 4.2; 6.2; 8.3; 10	1.1; 2.4; 3.2; 4.3; 5.4

Sample	DOC (mg/L)	Applied ozone (mg/L)	Consumed ozone (mg/L)
Sample 3	6.2	2.1; 4.3; 6.2; 8.3; 10	1.1; 2.1; 2.9; 3.5; 4.1
Sample 4	7.7	2.2; 4.4; 6.0; 8.3; 10	1.2; 2.3; 2.7; 3.8; 4.3
Sample 5	8.2	2.2; 4.2; 6.2; 8.3; 10	1.1; 2.0; 2.8; 3.5; 4.2

A.4 Appendix 4

A.4.1 Equations for model construction

The proposed model in this study used the structure of a neural network with one hidden layer and k neurons, three input variables (p=3) and a single dependent variable (Equation A-2).

$$y = f(x; \theta_w) + e \qquad \text{Equation A-1}$$

$$f(x, \theta_w) = w_{k0} + \sum_{j=1}^{k} w_{kj} \cdot \phi_j \left(w_{j0} + \sum_{i=1}^{p} w_{ij} \cdot x_i \right) \qquad \text{Equation A-2}$$

$$e \sim N(0, \sigma^2) \qquad \text{Equation A-3}$$

The θ_w denotes all the parameters w_{ij}, w_{j0}, w_{kj}, w_{k0} which are the hidden layer weights and biases, and the output layer weights and biases, respectively (Equation A-2). The independent variables are represented as x and the dependent variable by y. The random variable e is the model residual (Equation A-1 and Equation A-3). Where $N(0,\sigma^2)$ denotes a normal distribution with mean 0 and variance σ^2 (Equation A-3)

The input variables were normalised in the range [0,1] by using Equation A-4.

$$x_i = \frac{x_i^u - min_i}{max_i - min_i} \qquad \text{Equation A-4}$$

Where x_i^u denotes the unnormalised variable i, min_i is the minimum value for the variable i, and max_i is the maximum value for the variable i.

The activation function ϕ is the softplus function as shown in Equation A-5.

$$\phi_j(x) = \ln(1 + e^x)$$ Equation A-5

Weights and biases were defined with normal distributions with Gaussian distributions as shown in Equation A-6, Equation A-7, Equation A-8 and Equation A-9.

$$w_{k0} \sim N(0, \lambda_\alpha^{-1})$$ Equation A-6

$$w_{j0} \sim N(0, \lambda_\alpha^{-1})$$ Equation A-7

$$w_{kj} \sim N(0, \lambda_\beta^{-1})$$ Equation A-8

$$w_{ij} \sim N(0, \lambda_\gamma^{-1})$$ Equation A-9

Where λ^{-1}'s are the variance hyperparameters.

The fixed values for the highest level hyperparameters for the variances were defined by Equation A-10, Equation A-11, Equation A-12 and Equation A-13.

$$\lambda_\alpha \sim Gamma(r_\alpha, \lambda_\alpha)$$ Equation A-10

$$\lambda_\beta \sim Gamma(r_\beta, \lambda_\beta)$$ Equation A-11

$$\lambda_\lambda \sim Gamma(r_\lambda, \lambda_\lambda)$$ Equation A-12

$$\sigma^{-2} \sim Gamma(r_\sigma, \lambda_\sigma)$$ Equation A-13

Where r's are the r and λ are the parameters of the Gamma distribution.

A.4.2 Comparison of BMLP with other alternative models

Alternative models were tested for their fit to the data including multiple linear models and models with nonlinear terms. The tested equations are presented in Equation A-14, Equation A-15 and Equation A-16.A Bayesian approach was used to fit the data in order to find the posterior parameters. The comparison between observed and predicted CT values is presented in Figure A-18.As it can be observed, the Smith equation provided an adequate fit to the Coxsackievirus data (Figure A-18 c), but poor fit to the Adenovirus data with large residual errors (Figure A-18 d).

$$CT = \beta_0 + \beta_1 \cdot pH + \beta_2 \cdot Turbidity + \beta_3 \cdot LRV \quad \text{Equation A-14}$$

$$CT = \beta_0 \cdot pH^{\beta_1} \cdot Turbidity^{\beta_2} \cdot LRV^{\beta_3} \quad \text{Equation A-15}$$

$$CT = \beta_0 \cdot LRV \cdot (\beta_1 + \exp(\beta_2 + \beta_3 \cdot pH + \beta_4 \cdot Turbidity)) \quad \text{Equation A-16}$$

Where β_0, β_1, β_2, β_3, β_4 denote the parameters of the models. CT is the dependent variable, whereas pH, Turbidity and LRV are the independent variables.

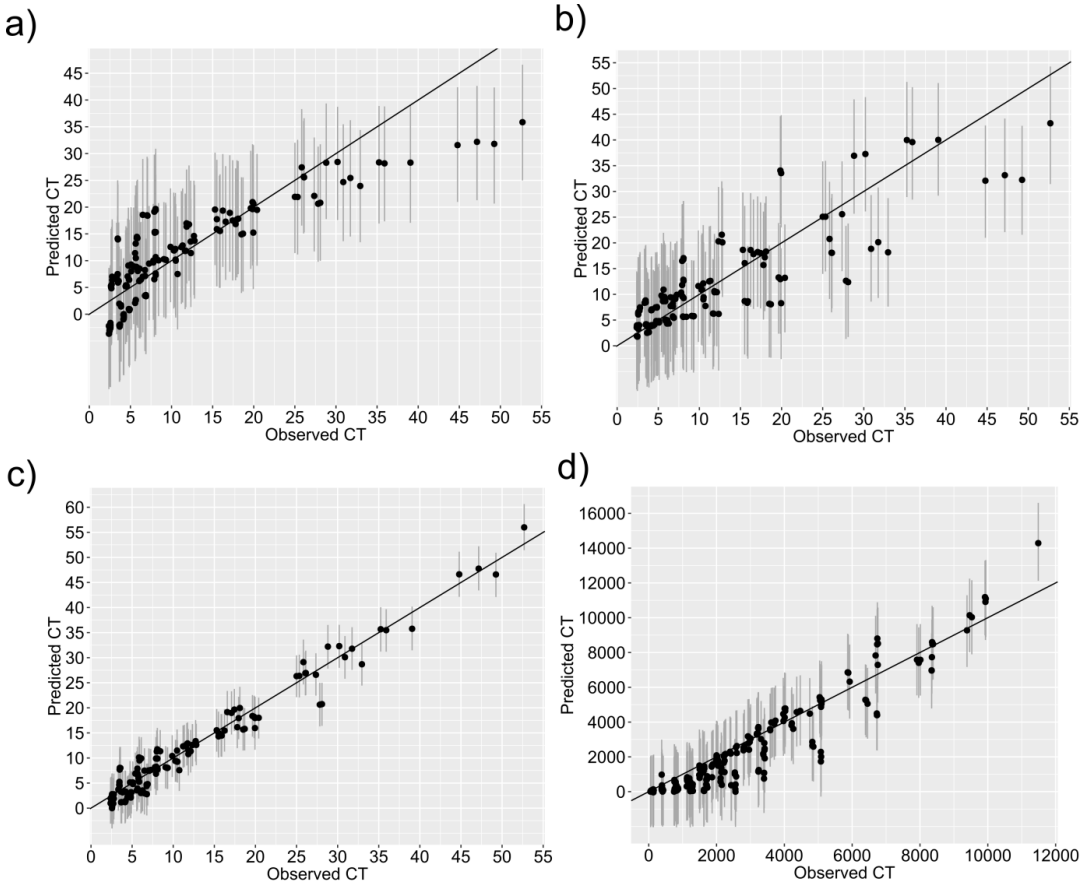

Figure A-18: Comparison of fits for alternative models. Multiple linear regression models with a) non-transformed variables, b) log-transformed variables. Smith equation fitted to c) Coxsackievirus data and d) Adenovirus data.

A.4.3 Jags code for model development

```
model{

for (r in 1:p){

gamma[1,r] ~ dnorm(0,taugamma)


for (l in 2:k){

gamma[l,r]<-gamma[(l-1),r]+delta.gamma[(l-1),r]

}
```

```
for (m in 1:(k-1)){
delta.gamma[m,r]~dnorm(0,0.01)T(0, )
}
}

for (j in 1:k){
bias[j]~dnorm(0,taubias)
beta[j] ~ dnorm(0,taubeta)
}

for (i in 1 : n){
y[i] ~ dnorm(ct.mean[i],tau.y)
y.pred[i]~dnorm(ct.mean[i],tau.y)

for (j in 1 : k){
p.eq[i,j]       <-       beta[j]*log(1+exp(bias[j]+gamma[j,1]*(x[i,1]-
min.tu)/(max.tu+min.tu)+gamma[j,2]*(x[i,2]-min.ph)/(max.ph-
min.ph)+gamma[j,3]*(x[i,3]-min.lrv)/(max.lrv-min.lrv)))
}
ct.mean[i] <- beta0+sum(p.eq[i,])
}

taubeta ~ dgamma(0.001,0.001)
taugamma ~ dgamma(0.001,0.001)
tau.y ~ dgamma(0.001,0.001)
taubias ~ dgamma(0.001,0.001)
beta0 ~ dnorm(0,taubias)
}

#Initial values
```

```
list("taubeta"=0.01,"taugamma"=0.01,"tau.lrv"=0.01,"taubias"=0.01,"bet
a0"=0)
```

A.4.4 R2Jags implementation in *R*

```
#Load required packages

library(R2jags)

library(coda)



#Import data

data<- read.csv("data.csv")



#Starting values

inits.mlp=function(){


list("mu.mubeta"=0,"taubeta"=0.01,"taugamma"=0.01,"tau.lrv"=0.01,"mubi
as"=0,"taubias"=0.01,"beta0"=0)

}



#Input data

x=as.matrix(data[,c(1,2,4)])

y=data$ct




#Number of neurons in hidden layer

k=3



#Number of input variables

p=3



#Input parameters for the model

datalist <- list(
```

```
  k=k,

  p=p,

  n=dim(data3)[1],

  x=x,

  y=y,

  min.tu=min(data$turbidity),

  max.tu=max(data$turbidity),

  min.ph=min(data$ph),

  max.ph=max(data$ph),

  min.lrv=min(data$lrv),

  max.lrv=max(data$lrv)


)



jags.params                                                                =
c("gamma","beta","beta0","tau.y","taubeta","taugamma","taubias")  #  The
parameters to be monitored



#Run model, "model.txt" is the model which is saved as a separate file

jagsfit.p                                          <-jags(data=datalist,
parameters.to.save=jags.params,inits=inits.mlp,

                  n.iter=2.2E6,               model.file                =
"model.txt",n.chains=1,n.burnin = 2E5,n.thin=1000)



#Visualise traceplots

traceplot(jagsfit.p)



#Show summary of the results

print(jagsfit.p)




#Geweke diagnostics
```

```
jagsfit <- as.mcmc(jagsfit.p)

autocorr.diag(jagsfit)

gewe=geweke.diag(jagsfit)

gewe2=sapply(gewe, "[[", "z")


pnorm(abs(gewe2),lower.tail=FALSE)*2


#Plot autocorrelation plots

autocorr.plot(jagsfit)
```

A.4.5 Types of queries performed in the model

A number of queries were analysed in this study depending on the observations provided and the target variable to be assessed. The first example illustrates when it is needed to query a required CT (Target CT) given the values of observed pH, observed turbidity and desired target LRV (Figure A-19).

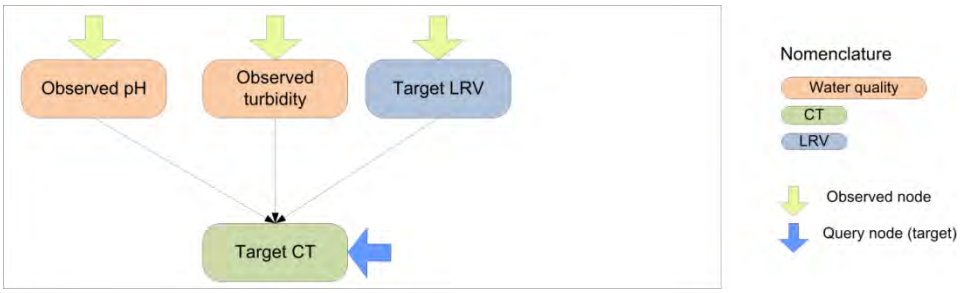

Figure A-19: Scheme of query for target CT given observations of pH, turbidity and LRV.

The second example displays the case when it is required querying the turbidity required to achieve a target LRV given the values of observed pH and CT (Figure A-20). This same type of query is used for pH variable.

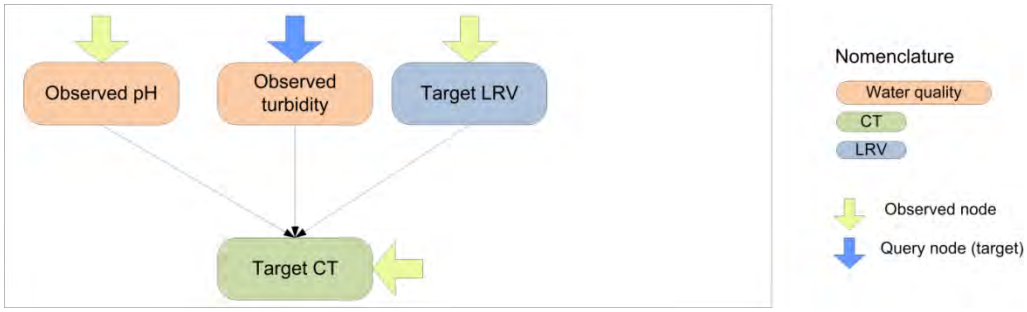

Figure A-20: Scheme of query for required turbidity given observations of pH, required LRV and observed CT.

The final instance shows the case when it is required an estimation of a LRV for an observed CT in the system (Figure A-21). For the estimation, a desired target LRV is selected and pH and turbidity are observe to obtain the target CT. The observed CT from the process is then used to calculate the ratio between the required CT and observed CT (inactivation ratio). Finally, the calculated LRV is computed from the multiplication of the selected LRV (e.g. 4 logs) and the inactivation ratio.

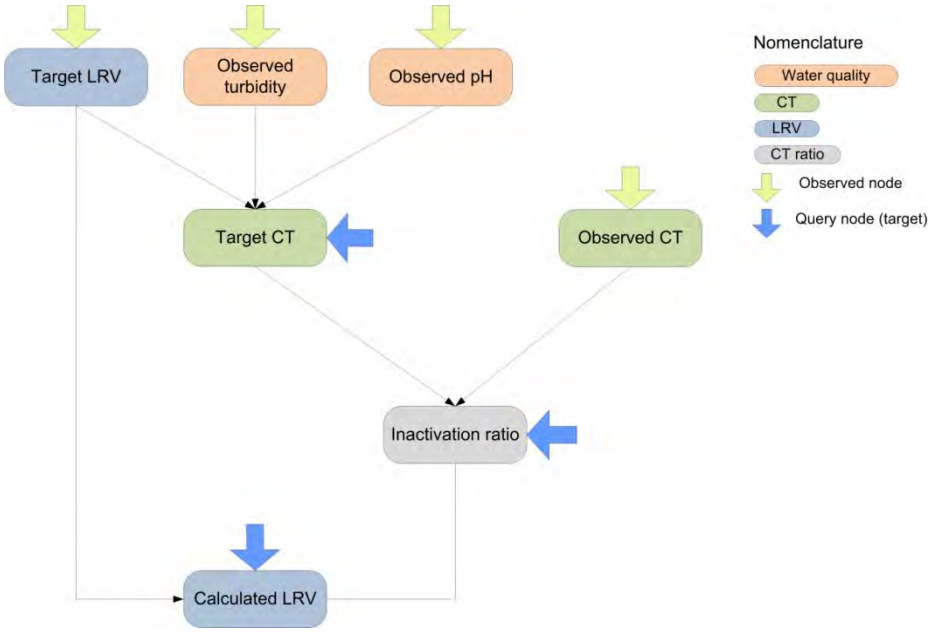

Figure A-21: Scheme of query for inactivation ratio and calculated LRV when the CT is exceeding the range of the study. The required observations are pH, turbidity, target LRV and observed CT.

A.5 Appendices references

Lunn, D., Jackson, C., Best, N., Thomas, A. and Spiegelhalter, D. (2012) The BUGS Book: a Practical Introduction to Bayesian Analysis. CRC press, Boca Raton, Florida.

Minitab (2010) Minitab 16 statistical software. Minitab Inc., State College, Pennsylvania, USA.