Development of a Model to Predict Bulking in Full-scale Wastewater Treatment Plants, and the Impact of Bulking in the Receiving Environment

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy: Health Sciences in the Faculty of Health Sciences at the Durban University of Technology

Nashia Deepnarain

April 2021

Supervisor: Prof. Faizal Bux

Co-supervisor: Prof. Sheena Kumari

Co-supervisor: Prof. Poovendhree Reddy

Co-supervisor: Prof. Thor Axel Stenström

DECLARATION

I hereby declare that this dissertation, submitted for the degree Doctor of Philosophy: Health Sciences at the Durban University of Technology, is composed of my original work and contains no material previously submitted to any other institution for academic qualifications. Where use is made of any author's work, it has been duly acknowledged.

_____ _____

Nashia Deepnarain (Student number: 20515606) *Date*

As the candidate's supervisory team, we hereby approve this dissertation for final submission.

Supervisor: Prof. Faizal Bux

_____ _____

Signature *Date*

Co-supervisor: Prof. Sheena Kumari

_____ _____

Signature *Date*

Co-supervisor: Prof. Poovendhree Reddy

_____ _____

Signature *Date*

Co-supervisor: Prof. Thor Axel Stenström

_____ _____

Signature *Date*

DEDICATION

I would like to dedicate this work to my late father Basanth (Kemme) Deepnarain. "Dad my eyes still search for You, my heart still calls out to You and my soul knows that You are at peace."

&

To my mother, Sandra Deepnarain, thank you mom for your guidance, support, encouragement and eternal love.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to:

- God Almighty, for giving me strength, courage and peace of mind during the most difficult times of my journey.
- My academic supervisor, Prof. Faizal Bux, your role as the father of IWWT is legendary. Thank you for your instrumental support, guidance and encouragement, my heartfelt thank you for giving me many opportunities to help me develop in my career. You have always made the time and I am infinitely grateful for your insights and wisdom.
- My co-supervisors; Prof. Sheena Kumari, Prof. Poovendhree Reddy and Prof. Thor Axel Stenström, I appreciate your patience, expert advice, constructive criticism and your belief in me throughout this study. I am eternally grateful to all for always making the time and for guiding me on this journey.
- Special gratitude to Dr. Mahmoud Nasr for your guidance and instrumental role in statistics and publications.
- Thank you to the Institute for Water and Wastewater Technology and the National Research foundation for support.
- My appreciation and heartfelt thanks to my family from IWWT: Dr. Trisha Mogany, Dr. Ismail Rawat, Dr. Isaac Denis Amoah, Dr. Luveshan Ramanna, Dr. Zinhle Marrengane, Dr. Leanne Pillay, Dr. Oluyemi Awolusi, Dr. Abimbola Enitan, Jashan Gokal, Keith Chetty, Sasha Lee Pillay, Julian Arran and Karen Reddy. Thank you all for the lifelong memories, humor and support throughout my journey.
- To my extended family and friends, thank you for your support and encouragement.
- To my brother, Avinash Deepnarain and sister-in-law Nivasha Deepnarain, thank you both for your love, encouragement and support.
- To my sister Kayshia Deepnarain and brother-in-law Brenton Cohen Subrayen, thank you for helping me see that ray of hope, when I never believed there was one.
- To my baby sister, Sanam Deepnarain, thank you for your troublesome moments which always made me laugh and keep centered.
- To my best friend, Kriveshin Pillay, a special thank you for your remarkable ideas, technical expertise and your guidance. I will forever treasure your love and support, especially through my most challenging life experiences.

ABSTRACT

Sludge bulking has been a continuous operational hurdle affecting the solids-liquid separation in wastewater treatment plants (WWTPs) worldwide. Excessive growth of filamentous bacteria is the primary and common cause of sludge bulking, which can have negative impacts on the wastewater treatment efficiency. Filamentous bacteria serve as the backbone structure of flocs which assist in the sludge settling process, however, their prolific growth result in slow sedimentation due to inadequate settling of flocs.

The main focus of this research was to develop a model to assist in a clearer understanding of the bulking sludge phenomenon in relation to filamentous bacterial growth and to identify predictors of bulking in different biological nutrient removal (BNR) WWTPs. The growth of filamentous bacteria and sludge bulking in different WWTPs and its association with sludge bulking incidents were evaluated using different statistical models [*viz.* artificial neural networks (ANN), principal component analysis (PCA), cluster analysis and Decision Trees]. In addition, the effect of bulking on pathogen discharge and its potential impact on the community was assessed using a microbial risk assessment model.

A total of seven WWTPs were investigated to identify the most common and dominant filamentous bacteria during bulking and non-bulking periods. A total of ten filamentous bacterial species were identified in this study with their dominance varying across the selected WWTPS during the sampling period. Based on the filament index scale ranging from 1 (None filament) to 7 (Excessive filament), the developed ANN model predicted sludge volume index (SVI) in relation to the abundances of ten filamentous species as model inputs. Among the filamentous bacteria identified, Eikelboom Type 0041 attained the highest impact on SVI, followed by *Gordonia* spp., *Nostocoida limicola*, and *Thiothrix* spp.

Developing a model for a WWTP, with proper calibration and validation against plant operational data, can allow for proper evaluation of filamentous bacteria associated to bulking, with effective mitigating strategies. Hence, in this study, a Decision Tree model was further implemented as a novel approach in the form of a case study to evaluate the effect of influent wastewater characteristics and plant operational parameters on the dominant filamentous bacteria and sludge bulking for prediction and control. Various factors such as pH, temperature, dissolved oxygen (DO), sludge retention time (SRT), food-to-microorganisms (F/M) ratio, soluble chemical oxygen demand (sCOD), total COD (tCOD), NH_4^+-N, total Kjeldahl nitrogen

(TKN), phosphorus as phosphate (PO_4^{3-}-P), TP, and total suspended solids (TSS) were considered to have an impact on filamentous dominance.

High bulking incidents were observed during long SRT and nutrient deficient (low F/M) conditions. However, a negative correlation was observed with soluble sCOD and ammonium-nitrogen (NH_4-N). Type 0092 was the dominant species largely responsible for sludge bulking in the selected plants, which prevailed at low F/M (< 0.08 kg COD/kg MLSS d^{-1}) conditions. The secondary filaments *Candidatus Microthrix parvicella* increased in their abundance at low temperature (< 15.5°C), causing an increase in SVI at lower ambient temperatures. In addition, an increase on *Thiothrix* spp. was linked with the unbalanced ratio between readily biodegradable COD and nutrient conditions.

The last objective of this study provided an assessment from an environmental health perspective, by investigating the impacts of bulking on the receiving environment, using a quantitative microbial risk assessment (QMRA) approach. This was done by studying the difference in selected microbial pathogen abundance during bulking and non-bulking conditions using qPCR. *Salmonella* was the most dominant species of the investigated microorganisms, during the study period (2270– 96733 copies ng^{-1} of DNA) followed by *E. coli* (4133 – 76847 copies per ng of DNA); whereas, *Mycobacterium* was the least (542 – 3340 copies ng^{-1} of DNA). During high bulking with SVI >200 mL g^{-1}, positive correlations were found between the selected pathogens in the final effluent.

The QMRA model was applied to investigate the safety of treated effluent for (a) children, women, and men during recreational activities, (b) farmers during irrigation practices, and (c) consumers of edible plants (vegetables). The QMRA values during all bulking events exceeded the tolerable risk of 10^{-4} (*i.e.* less than one case of infection per 10 000 people) per year, as recommended by the world health organization (WHO). In addition various disinfection scenarios such as chlorination, ultraviolet (UV) and ozonation were tested to control the risks associated with pathogenic bacteria, for further information of safe disposal and reuse of the treated effluent. The application of UV provided the most effective treatment to reduce the pathogenic bacteria, except for the case of children that were exposed to *Salmonella* infection. To the best of my knowledge, the probable health risks associated with the discharge or reuse of WWTPs effluents under different sludge bulking events have not yet been systematically evaluated using QMRA.

This research can potentially lead to the development of appropriate model systems for bulking control in full-scale WWTPs, while highlighting some of the significant contributors, environmental impact and mitigation strategies. The outcomes of this research will contribute to the current global body of knowledge in relation to predictive models for filamentous bulking control in full-scale WWTPs.

LIST OF PAPERS

Nashia Deepnarain, Mahmoud Nasr, Sheena Kumari, Thor Axel Stenström, Poovendhree Reddy, Kriveshin Pillay, Faizal Bux. (2020). Artificial intelligence and multivariate statistics for comprehensive assessment of filamentous bacteria in wastewater treatment plants experiencing sludge bulking. *Environmental Journal and Innovation.* 19: 1-14. https://doi.org/10.1016/j.eti.2020.100853. (Presented in Chapter 3 and Appendix 3).

Nashia Deepnarain, Mahmoud Nasr, Sheena Kumari, Thor Axel Stenström, Poovendhree Reddy, Kriveshin Pillay and Faizal Bux. (2019). Decision Tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant. *Process Safety and Environmental Protection*. 126: 25-34. https://doi.org/10.1016/j.psep.2019.02.023. (Presented in Chapter 4 and Appendix 4).

Nashia Deepnarain, Mahmoud Nasr, Isaac Dennis Amoah, Abimbola Enitan, Poovendhree Reddy, Thor Axel Stenström, Sheena Kumari and Faizal Bux. (2020). Impact of sludge bulking on receiving environment using quantitative microbial risk assessment (QMRA)-based management for full-scale wastewater treatment plants. *Journal of Environmental Management*. 267: 1-9. https://doi.org/10.1016/j.jenvman.2020.110660. (Presented in Chapter 5 and Appendix 5).

LIST OF CONFERENCE PRESENTATIONS

Nashia Deepnarain, Sheena Kumari, Poovendhree Reddy, Thor Axel Stenström, Jordache Ramjith, Kriveshin Pillay and Faizal Bux. (2016). A conceptual model for filamentous bulking control in biological nutrient removal systems. Water Institute of Southern Africa (WISA), conference and Exhibition. 16-19 May 2016.

Oluyemi Olatunji Awolusi, Kriveshin Pillay, **Nashia Deepnarain**, Faizal Bux and Sheena Kumari. Population dynamics of key functional microbes associated with nutrient removal in different full-scale BNR systems. Water Institute of Southern Africa (WISA) 2018 Biennial Conference and Exhibition. Cape Town International Conventional Centre, Cape Town, South Africa. 26-28 June 2018.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
ARX	Auto-Regressive Exogenous
ASM	Activated Sludge Model
ASP	Activated Sludge Process
BNR	Biological Nutrient Removal
BOD_5	Biochemical Oxygen Demand
CART	Classification and Regression Trees
CBR	Case-Based Reasoning
CFD	Computational Fluid Dynamic
CLM	Cumulative Logistic Model
CTAB	Cetyltrimethyl Ammonium Bromide
CVI	Cause Variables Identification
DAPI	DNA Stain 4′, 6-Diamidino-2-Phenylindol
DNA	Deoxyribonucleic Acid
DO	Dissolved Oxygen
DWS	Department of Water and Sanitation
EPA	Environmental Protection Agency
EPS	Extracellular Polymeric Substance
F/M	Food to Microorganism Ratio
FI	Filamentous Index
FISH	Fluorescent *in situ* Hybridization
FNN	Fuzzy Neural Network
IAWPRC	International Association on Water Pollution Research and Control
MARS	Multivariate Adaptive Regression Splines
MBR	Membrane Bioreactor
MCRT	Mean Cell Residence Time
MiDAS	Microbial Database for Activated Sludge
MLSS	Mixed Liquor Suspended Solids
MLVSS	Mixed Liquor Volatile Suspended Solids
MSE	Mean Squared Error
NH_4-N	Ammonium-Nitrogen
O_3	Ozone

PBS	Phosphate-buffered Saline
PCA	Principal Component Analysis
PHB	Poly -3-Hydroxybutyrate
PID	Proportional Integral Derivative
PO_4^{3-}-P	Phosphorus
PST	Primary Settling Tank
QMRA	Quantitative Microbial Risk Assessment
qPCR	Quantitative Polymerase Chain Reaction
RAS	Return Activated Sludge
RFM	Random Forest Model
RWQC	Recreational Water Quality Criteria
SCADA	Supervisory Control and Data Acquisition
sCOD	Soluble Chemical Oxygen Demand
SORBFNN	Self –organising Radial Basis Function Neural Network
SRT	Sludge Retention Time
SS	Suspended Solids
SST	Secondary Settling Tank
SVI	Sludge Volume Index
SVR	Support Vector Regression
tCOD	Total Chemical Oxygen Demand
TKN	Total Kjeldahl Nitrogen
TN	Total Nitrogen
TP	Total Phosphorus
TSS	Total Suspended Solids
UCT	University of Cape Town
UV	Ultraviolet
WAS	Waste Activated Sludge
WHO	World Health Organization
WWTP	Wastewater Treatment Plant

CHAPTER ONE: INTRODUCTION

Globally, rapid urbanisation, depletion of available freshwater resources and surface water pollution are the major converging contributors to water scarcity. Wastewater reclamation is one of the main possibilities to overcome such increasing pressures on water resources (Salgot and Folch, 2018). According to the South African wastewater services, approximately 1000 municipal wastewater treatment plants (WWTPs) are in operation, which treat approximately 7500 ML d^{-1} of wastewater (DWA, 2012a). However, currently, more than 80% of marginally treated industrial and municipal wastewater is released into the environment, thus representing one of the largest sources of surface water pollution in South Africa (Ntombela *et al.,* 2016, Boretti and Rosa, 2019). The water quality of many rivers continues to deteriorate at unprecedented rates, primarily due to failing wastewater treatment facilities containing high levels of nutrients and waterborne pathogens (Momba *et al.*, 2006, DWA, 2012b, Edokpayi *et al.*, 2017, Adegoke *et al.*, 2018, Amoah *et al.*, 2018a). Most waterborne pathogens that cause disease in humans and animals (*viz*. cholera, amoebiasis, diarrhea, typhoid fever, and gastroenteritis, shigellosis, salmonellosis, giardiasis and cryptosporidiosis) originate from animal and human fecal wastes that pass through the final effluent of inadequately treated WWTPs into receiving water bodies. This poses severe health risks for water reuse applications (WHO, 2016, Amha *et al.*, 2017, Amoah *et al.*, 2018b).

The activated sludge process (ASP) is still the most widely used wastewater treatment technology globally and in South Africa. The ASP was initially designed to remove carbonaceous material from wastewater and consisted of a simple configuration involving an aeration tank coupled to a secondary settling tank (SST), for nutrient removal and sludge settling, respectively (Diehl and Faras, 2013, Mulas *et al.*, 2015). The initial ASP process design was subjected to several modifications, and have gradually transpired into various existing biological nutrient removal (BNR) processes which further included anoxic and/or anaerobic selectors to facilitate nutrient removal (Jeppsson, 1996, Grady *et al.*, 2011). These modifications however, have also led to severe challenges to the treatment plants including hindered settling due to the excess growth of filamentous bacteria (Martins *et al.*, 2004, Hu *et al.*, 2012).

The implementation of new technologies in South African WWTPs had caused many operational problems in full-scale systems, which include improper sludge settling and the requirement of longer sludge ages to facilitate nitrification in winter seasons. These operational problems often lead to failure in achieving the required wastewater effluent discharge standards (Musvoto *et al.*, 1999, Hu *et al.*, 2003, Ekama, 2010). The country depends largely on surface water for urban, industrial and agricultural requirements. Currently, few municipalities within the country have implemented several demand management strategies such as wastewater reuse which presents promising alternatives to curb growing demands to alleviate declining freshwater resources (Adewumi *et al.*, 2010). This will also add pressure on municipal and industrial WWTP facilities, to ensure optimum performance of the plant to achieve the desired objectives for wastewater reclamation.

Formation of well-balanced, compact flocs with efficient settling properties is imperative for the optimal functioning of BNR processes. The activated sludge flocs consist largely of microorganisms which aggregate together by a sticky "gel-like" substance, known as the extracellular polymeric substance (EPS) (McSwain *et al.*, 2005a). The microbial community within the ASP form ecological interactions within the floc structure and contribute to flocculation, therefore allowing sludge to settle effectively. Filamentous bacteria usually form the backbone of the floc structure. An optimally working ASP generally will have a well-balanced microbial community structure within the floc that jointly contribute to achieving a high-quality effluent. However, under suboptimal conditions, an imbalance in microbial communities could lead to numerous nutrient removal problems including the excessive growth of filamentous bacteria. The excessive growth of filamentous bacteria can extend from the flocs core to the outer surface, thus forming irregular diffuse flocs and inter-floc bridging that can affect the sludge settling process (Smets *et al.*, 2006, Wagner *et al.*, 2015). Based on the presence and abundance of filamentous bacteria, the flocs are generally classified as: 1) ideal flocs (moderate levels of filamentous bacteria conserved within floc), 2) pin point flocs (lack of filamentous bacteria, leading to a very low SVI and a turbid effluent) and 3) diffused flocs (excessive filamentous bacteria causing bulking conditions and resulting in a high SVI) (Martins *et al.*, 2004, Mesquita *et al.*, 2011).

Sludge bulking is a common phenomenon which can be described as the inability to separate solids from the treated effluent (Wanner, 1994, Ramin *et al.*, 2014). The loss of functional biomass as a result of bulking and/or foaming conditions, can also significantly reduce the

nutrient removal efficiency and lead to failure in achieving the required wastewater effluent discharge standards. Bulking conditions can also lead to increased discharge of suspended solids which can results in spike in pathogen concentrations in the final effluent that can contaminate rivers and water catchment areas, thus causing an adverse effect on the receiving environment. More than 50% of wastewater treatment processes in many countries such as China, Australia, Italy, USA, Denmark and South Africa experience sludge bulking often due to the proliferation of filamentous bacteria (Richard *et al.*, 2003, Tandoi *et al.*, 2006, Lou and Zhao, 2012, Mielczarek *et al.*, 2012, Welz *et al.*, 2014, Lever, 2015). Many operational and environmental conditions are known to cause sludge bulking in WWTP *viz*. low dissolved oxygen (DO), low food to microorganism (F/M) ratio, long sludge retention times (SRT), high grease and oil content and low temperatures (Jenkins *et al.*, 2003, Martins *et al.*, 2004, Mielczarek *et al.*, 2012, Milobedzka *et al.*, 2016, Rossetti *et al.*, 2017).

Globally, in municipal WWTPs, filamentous species such as *Microthrix parvicella*, *Haliscomenobacter hydrossis*, Types 0041, 0803 and 0092 belonging to phylum *Chloroflexi* have often been identified. Whereas, in industrial WWTPs, *Thiothrix* spp., Type 021N and Type 1851 are commonly reported (Mielczarek *et al.*, 2012, Nittami *et al.*, 2019). In South Africa, former surveys conducted by Blackbeard *et al.* (1986), reported that out of the 111 AS processes surveyed, 33 were BNR systems, from which 27 WWTPs experienced bulking episodes. In their study, filamentous bacteria that were known to proliferate under low F/M ratios, were commonly identified. Among these, Eikelboom Type 0092, Type 1851 and Type 0041 were dominant in approximately 75% of all South African WWTPs. A recent survey in the Western Cape Province indicated that the filaments are still dominant in these BNR systems (Welz *et al.*, 2014). The dominant filamentous bacteria identified in their study in descending order of frequency were: Type 0092 (68.9%), *M. parvicella* (39.8%), Type 0041 (33.9%), Type 021N (29.4%), Type 1851 (28.7%), Gram positive branching bacilli (21.5%), Type 0803 (5.9%), and *Thiothrix* spp. (Welz *et al.*, 2014).

Even though extensive research has been carried out on filamentous bacteria, there is still no systematic method to control their growth and they are found to occur sporadically causing bulking episodes, worldwide (da Motta *et al.*, 2003, Rossetti *et al.*, 2017). Therefore, knowledge of specific nutrient requirements and operational parameters that can control the activity of filamentous bacterial types may be useful in controlling their proliferation (Martins, 2004, Martins *et al.*, 2004). Various statistical models currently explain the complexities of

biological systems with the function of micro-organisms (Belanche *et al.*, 2000, Lou and Zhao, 2012). Lou and Zhao (2012) reported on sludge bulking incidents in a WWTP in China, and correlated SVI to the plant operational parameters using principal component regression analysis and artificial neural networks (ANN). However, the sludge bulking phenomenon could not be completely explained and the mechanisms of the inner signal processing of the model were unknown. Singh *et al.* (2008) used the support vector regression (SVR) and the ANN model to predict sediment removal efficiencies in the settling basins. The authors reported that the SVR model performed statistically better in comparison to ANN. All three models are data-driven (based on input and output data generated), which analyses extensive data sets, formed by computational programming. However, these models pose many challenges in understanding and interpreting the intermediate and output results, and its application with filamentous bacteria was limited (Singh *et al.*, 2008).

Modeling of WWTPs faces many challenges, as these engineered processes include physical, chemical and biological phases, as well as sensitive instrumentation for collecting large data. The data can be noisy and uncertain, therefore algorithms can be produced to process such data, by reducing its dimensionality, and to determine important variables. Optimization of these biological processes is another major challenge, as these models are generally nonlinear and dynamic. Although many research studies towards modelling of sludge bulking, have been carried out, the available prediction technologies are still not as developed and need further investigation. Proper WWTPs modeling techniques, including information on the microbiology, physicochemical properties, empirical knowledge and process operational parameters are necessary for optimal operation (Banadda *et al.*, 2011, Chun *et al.*, 2017, Fan *et al.*, 2018, Chmielowski *et al.*, 2019).

Therefore, three major research objectives were designed in this study, to contribute to the knowledge of filamentous bulking in full-scale wastewater treatment plants. The first objective included profiling of filamentous bacteria in selected WWTPs using cluster analysis and ANN (chapter 3, see paper I in Appendix 3). The second objective focused on the development of a Decision Tree model to predict bulking incidence (chapter 4, see paper II in Appendix 4). The third objective of this study implemented the use of a quantitative microbial risk assessment model (QMRA) approach to determine the impact of bulking on the receiving environment (chapter 5, see paper III in Appendix 5).

1.1. RATIONALE OF THE PRESENT STUDY

A great demand currently exists for the development of biological process models that can predict and control sludge bulking in WWTPs. The development of novel techniques to control sludge bulking by the use of current modelling approaches such as artificial intelligence is relevant to WWTPs. The identification of filamentous bacteria via their morphological characteristics are often useful, however can be misleading based on imprecise filamentous bacteria identified and human biases. The current development of molecular techniques allows for proper identification and quantification of filamentous bacteria with higher accuracy and reliability. A significant part of this thesis is dedicated to a comprehensive assessment of dominant filamentous bacteria based on a filamentous index (FI) scale that were related to sludge bulking using ANN and Decision Trees. Currently, Decision Trees are useful tools for modelling sludge bulking, which form clear patterns and provide high accuracy. The Decision Tree model had been successfully used in many fields, however limited to wastewater treatment systems. Monitoring the filamentous bacteria populations at different conditions using the ANN and decision tree model can be feasible and simple approaches to alleviate bulking. This thesis provide effective modelling tools to predict bulking using SVI. The use of SVI is the current gold standard to characterize sludge bulking. However, more reliable methods such as online sensors are needed to determine sludge settleability and to encourage a wider application of sludge bulking using artificial intelligent approach.

A QMRA model was further applied to determine the probable health risks associated with the discharge or reuse of WWTPs effluents under different sludge bulking events. Currently, this is the first study which described the exposure routes and the risks of infection that were associated with selected pathogenic bacteria, such as *E. coli* O157:H7, *Salmonella*, and *Mycobacterium*, encountered by different sludge bulking events from full-scale WWTPs. This study would be helpful for the management of human health risks associated with effluent wastewater containing pathogens, *i.e.*, particularly concerning the case of sludge bulking. Substantial work is needed to expand on the framework of QMRA on enhancing the precision outcomes using some environmental aspects pointed out in this study.

1.2. PROJECT AIMS AND OBJECTIVES

The overall aim is to predict the abundance of filamentous bacteria and sludge bulking using a Decision Tree model, and to measure the impact of bulking on the risk of pathogenic infection for exposed communities. This was achieved by the following objectives:

a) To profile filamentous bacteria associated with sludge bulking in selected full-scale biological nutrient removal processes, using artificial neural networks and multivariate regression analysis (see paper I presented in Appendix 3).

b) To predict the abundance of filamentous bacteria and bulking in relation to influent physical-chemical characteristics and operational parameters in full-scale biological nutrient removal processes using a Decision Tree model (see paper II presented in Appendix 4).

c) To evaluate the potential health risks associated with filamentous bulking on the receiving environment using a quantitative risk assessment model (see paper III presented in Appendix 5).

CHAPTER TWO: LITERATURE REVIEW

2.1. WASTEWATER TREATMENT

Globally, the United Nations World Water Development Report, estimated that 80% of wastewater (over 95% in some developing countries) is released into the receiving environment without efficient treatment, thus posing severe environmental and human health risks (UNESCO, 2019, United Nations, 2018). The global water demand expects more than 30% increase above the current water use by 2050 (Boretti and Rosa, 2019). Reclaimed water can be useful to our society, further contributing to water security and sustainable development. Treated wastewater can be beneficial for irrigation and industrial processes. However, the wastewater sector requires new improvements concerning wastewater treatment infrastructure and technology, including proper training and monitoring schemes for efficient treatment (Voulvoulis, 2018).

A wastewater treatment plant (WWTP) can be described as an operational facility that consists of a combination of various processes *viz*. physical, chemical and biological for the treatment of domestic and industrial wastewater (Hreiz *et al.*, 2015). The treatment of wastewater is mainly concerned with the removal of floatable material, settleable and biodegradable organic compounds [(*viz*. carbonaceous oxygen demand (COD) and nutrients such as nitrogen (N) and phosphorus (P)], and pathogens (Van Loosdrecht *et al.*, 2015). In general, a conventional activated sludge wastewater treatment is divided into three core treatment steps *viz.* primary, secondary and tertiary treatment, to ensure the complete removal of pollutants in wastewater (Figure 2.1). The primary treatment step involves the physical removal of oils, greases and floatable organic material together with the sedimentation of large particles by the use of screen bars and sedimentation tanks (Figure 2.1). Primary treatment can achieve up to 50% reduction of pollutants by physical treatment methods. Primary settling tanks (PSTs) also known as primary clarifiers is generally located after the screen bars and often equipped with mechanical scrapers which skim oils and greases from the surface to allow further sedimentation (Spellman, 2003).

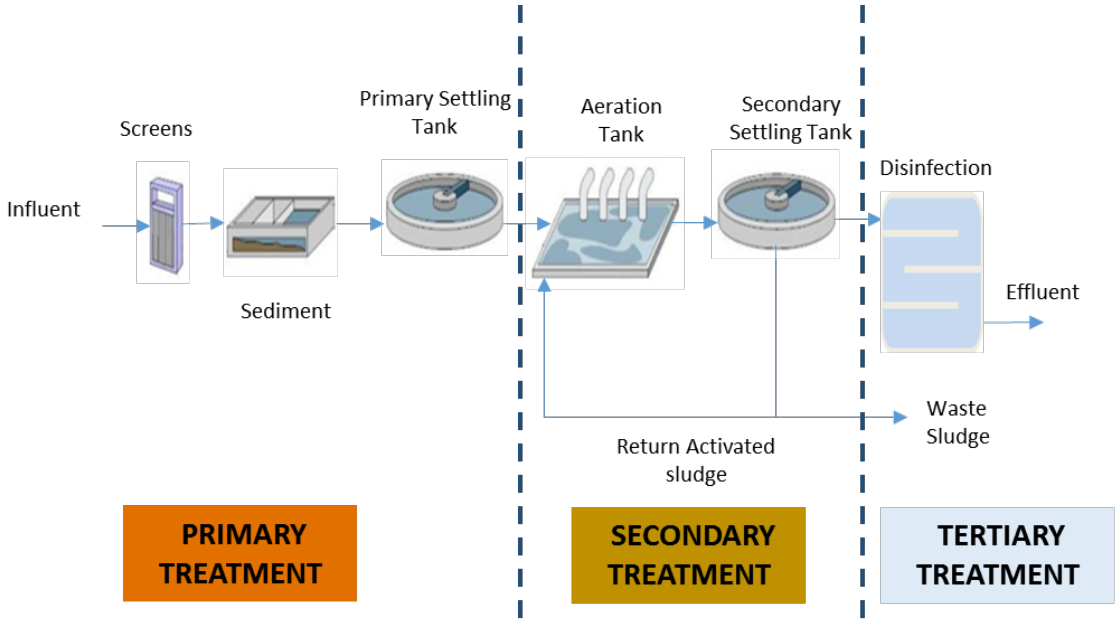


Figure 2.1. Flow diagram of a typical wastewater treatment process (Pei *et al.*, 2019).

The secondary biological treatment of a conventional ASP consists of aerobic tanks, coupled to a secondary settling tank (SST), for COD removal and sludge settling respectively. The successful operation of secondary treatment mainly relies on maintaining a high mixed liquor suspended solids (MLSS), containing concentrated microorganisms, capable of high rate nutrient removal in the aerobic tanks and sludge separation in the SSTs (Gernaey *et al.*, 2004, Ramin *et al.*, 2014, Han *et al.*, 2018a). A biological network of microorganisms, largely in the form of flocs degrades and removes pollutants by the use of oxygen which is supplied either by mechanical agitation or by diffused aeration in the aeration tank. The microorganisms together with oxygen, are capable of degrading up to 95% of the organic matter during secondary treatment (Martins *et al.*, 2004). The SST, also known as a clarifier, is a hydraulically sensitive unit operation, whereby sludge flocs settle via gravity sedimentation. The separation of biomass from the liquid in the SST is necessary to achieve a high-quality effluent, thus producing an end product relatively free of microorganisms (Martins *et al.*, 2004). Whereas, a larger fraction of biomass is returned to the biological aerobic reactors via the return activated sludge (RAS) stream to maintain the MLSS in the system to an optimum level (Takacs, 2008).

Further treatment is essential to improve the water quality and to ensure the removal of pollutants, whereby the final SST effluent channel through tertiary treatment (*viz*. ozone, chlorine, UV contact tanks) and subsequently discharged into receiving water bodies. The WWTPs require improved tertiary treatment to reduce the discharge of residual pollutants to the

receiving environment, from the secondary treated effluent. Moreover, improved monitoring techniques are essential to eliminate peak-pollution loads caused by problems such as bulking or overloading of the biological WWTP. Membrane filtration techniques have shown to be effective in reducing bacteria and viruses substantially decreasing the number of hazardous microbes in the final effluent. Methods such as filtration could be carried out before the disinfection unit to further improve the efficiency of disinfection, in the case when more efficient techniques are essential to reduce the microorganisms (Quach-Cu *et al.*, 2018).

2.2. CONVENTIONAL ACTIVATED SLUDGE PROCESS

Globally, the ASP is the most common technology, for more than 100 years known for its convenience and simplicity, to treat domestic and industrial wastewater (Hreiz *et al.*, 2015, Rossetti *et al.*, 2017). The ASP is an integral component of the biological wastewater treatment process, composed largely of aerobic biological reactors for the nutrient removal coupled to secondary settling tanks for the biomass separation (SSTs) (Figure 2.2) (Van Loosdrecht *et al.*, 2015, Puyol *et al.*, 2017). Many research studies focused largely on nutrient removal which has contributed to efficient optimization and control. However, the solids-liquid separation via gravity separation is a critical stage of the WWTP, due to problems that could severely affect the compaction and settling process, primarily as a result of excessive filamentous bacteria (Eriksson *et al.*, 1992).

The first mathematical model on the ASP was initiated in the early 1980s by the International Association on Water Pollution Research and Control (IAWPRC) (Henze *et al.*, 1987, Chun *et al.*, 2017). Edward Arden and William T. Lockett had significantly contributed to the initial mathematical model of the ASP (Ardern and Lockett, 1914). Their study showed that the addition of sludge into wastewater, together with oxygen, caused the organic matter to be removed rapidly by the microorganisms (Hartmann, 1999, van Haandel and Van der Lubbe, 2012). The microbial consortium primarily consists of bacteria and protozoa which form flocs and their composition depend on the influent wastewater characteristics and the plant operational conditions.

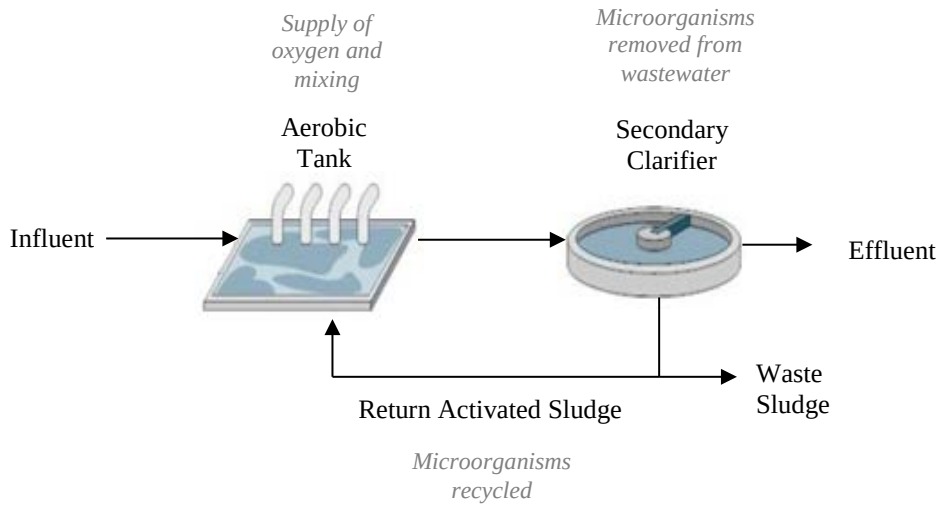


Figure 2.2. Schematic diagram of a conventional activated sludge process.

To maintain the microbial consortium of the ASP, approximately 40-60% of thickened underflow sludge volume in SSTs is recycled *via* the return activated sludge (RAS) stream, which in turn transfers the settled sludge from the SSTs to the aerobic reactors, whereas the remaining part is wasted (Figure 2.2). If the RAS rate is too low, solids remain in the settling tank, resulting in loss of solids (MLSS) and a septic return activated sludge in the aeration tank. Filamentous bacteria can proliferate during such conditions, thus causing a continuous cycle of sludge bulking (Spellman, 2003). Therefore, it is important to maintain a balance between sludge wastage and recycle flows. The waste activated sludge (WAS) is an important operational parameter, since it allows to maintain the desired MLSS concentration, food to microorganisms' ratio (F/M), and sludge age in the system. The microbial populations best suited to treat the raw incoming sewage can be selected through continuous recycling and wastage of sludge (Rustum, 2009).

Bulking has been the scourge of activated sludge ever since the transition from the ASP to completely mixed, continuous flow-through systems. Plug-flow reactors had resulted in more stable treatment plant performance to control bulking, however nutrient removal was challenging in these systems (Ardern and Lockett, 1914). To this end, several modifications of the ASP such as plug flow systems and tapered aeration systems developed in the early 1930s, which were thought to reduce the prevalence of filamentous bacterial and thus increase the settleability of the sludge. These systems were modified to assist in treating influents with high organic loading and encourage nitrification over longer hydraulic retention times, as well as to

further overcome the imbalances of diffused oxygen (Grady *et al.*, 1999, Seviour *et al.*, 2010). The return sludge mixed with the raw influent at the head of the aeration tank, require a higher DO, since the microbial biomass utilise oxygen at a much faster rate for degradation of organic materials. Towards the end of the reaction tank, the concentration of organic material is reduced by microorganisms thereby concomitantly lowering the oxygen demand. Hence, the tapered aeration system was designed to distribute the oxygen across the tank relative to the biomass requirements (Seviour *et al.*, 2010). However, lack of oxygen in some regions (towards the outlet of the primary tank), due to inadequate mixing, often hinder settling of the biomass. Biomass that settled at the bottom turn septic resulting in excessive filamentous bacteria (*viz. Thiothrix* spp., Type 021N, Type 1851 and *N. limicola*) (Seviour and Nielsen, 2010).

The process plant configurations had further evolved to reduce effluent nutrient (*viz*. nitrogen and phosphorus) concentrations to environmentally acceptable levels, as the final effluent discharge license requirements became more stringent. Further selective pressures for biological nutrient removal (BNR) are still imposed to encourage the growth of selective bacterial populations to ensure nitrogen (N) and phosphorus (P) removal (Barnard and Comeau, 2014, Khunjar *et al.*, 2014, Speirs *et al.*, 2019).

2.3. BIOLOGICAL NUTRIENT REMOVAL PLANTS

The initial design of the BNR process, included an anoxic tank to the ASP configuration, to enhance nitrification and denitrification (Chudoba *et al.*, 1973, Grady *et al.*, 1999). However, BNR processes may need a longer mean cell residence time (MCRT) to allow complete nitrification and denitrification to occur, which typically leads to foam or excessive filamentous bacteria problems. Further knowledge of the history and development of the different biological nutrient removal process configurations have been explained in detail by various authors (Albertson, 1991, Jeppsson, 1996, Seviour *et al.*, 2010, Van Loosdrecht *et al.*, 2015).

Barnard (1983), reported that the inclusion of the anaerobic tank followed by aerobic tanks, assisted in the removal of phosphorus such as the development of a 5-stage Phoredox process. However, due to the partial nitrification in these systems, a modified 3-stage process was later configured (Figure 2.3). This design include an internal recycle from the aerobic to the anoxic to for complete nitrification and denitrification. Whereas, the return activated sludge (RAS) from the final effluent to the anaerobic tank at the beginning of the system, is selective for

phosphorus removal. However, the disadvantage of this system lead to excess nitrates in the RAS, which inhibit proper phosphorus release in the anaerobic tanks (Jeppsson, 1996, Seviour and Nielsen, 2010).

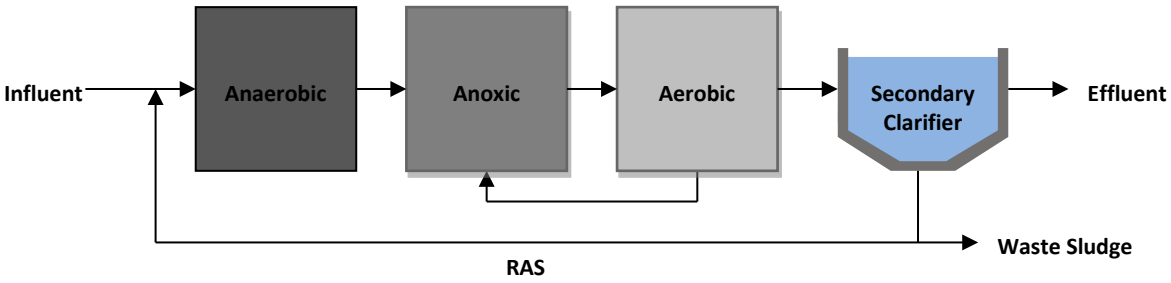


Figure 2.3. BNR plant with anaerobic, anoxic and aerobic stages (3 Stage phoredox process) (Seviour *et al.*, 2010).

To further improve on the nitrification and phosphorus removal process, these systems were later modified into a Johannesburg process and later, a University of Cape Town (UCT) process (Seviour *et al.*, 2010). The Johannesburg process was configured from the Phoredox system, by introducing a denitrification (pre-anoxic) tank before the anaerobic tank, to overcome the negative effect of NO_3^- on enhanced biological phosphorus removal. The high biomass concentration was intended to ensure complete denitrification in the pre-anoxic tank. With regards to the UCT process, RAS enters into the anoxic tank and the mixed liquor enters then enters the anaerobic tank. The UCT process was designed to ensure NO_3^- is completely denitrified before reaching the anaerobic tank (Seviour *et al.*, 2010). Like most BNR systems, these configurations use longer sludge ages (> 20 days) due to internal recycles, which can result in filamentous bulking. Many surveys on BNR WWTPs showed inadequate sludge settling, of these processes due to excessive filamentous bacteria. Among some of the factors contributing to filamentous abundance, the use of alternating selectors (anoxic to aerobic tanks) and low organic loading favor their prolific growth (Chudoba *et al.*, 1973, Musvoto *et al.*, 1999, Seviour *et al.*, 2010, Mielczarek *et al.*, 2012, Rossetti *et al.*, 2017).

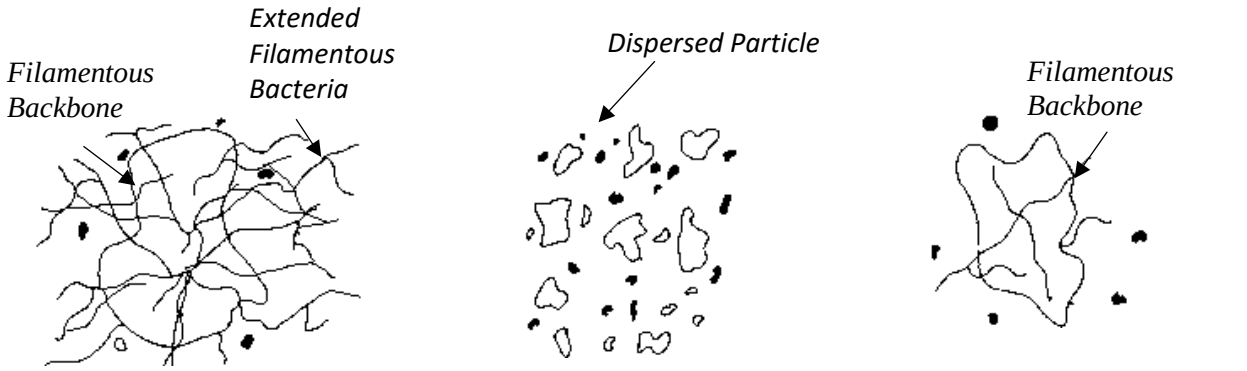
2.4. FLOC STRUCTURE AND FILAMENTOUS BACTERIA

Microorganisms aggregate and cluster together by a "gel-like" substance, known as the extracellular polymeric substance (EPS) layer to form floc (Grady *et al.*, 1999, Govoreanu *et al.*, 2003, McSwain *et al.*, 2005b, Burger *et al.*, 2017). London forces, electrostatic interactions and hydrogen bonds occur between the EPS and microbial clusters, thus, binding the microbes within the EPS layer, which enable flocculation (Tian *et al.*, 2006). Activated sludge flocs can grow typically to a range of 30 to 1800 μm in diameter (Etterer, 2006). Organic matter and nutrients (*viz.* carbon, nitrogen and phosphorous) are effectively removed from the wastewater by the support of firm healthy flocs, hence, providing a high-quality effluent (Martins *et al.*, 2004).

The primary microorganisms within the flocs, which forms the biotic community, for the nutrient removal process includes heterotrophic bacteria, nitrifiers, denitrifiers, phosphate accumulating organisms (PAOs) and glycogen accumulating organisms (GAOs). These functional microbial groups are imperative for nutrient removal (Henze *et al.*, 2008). Filamentous bacteria can be described as "threadlike" microbes that form a firm backbone structure of the floc and are the major microbial group responsible for floc formation in the ASP (Cenens *et al.*, 2000, Guo and Zhang, 2012). The size, density, shape and abundance of filamentous bacteria determine flocs stability and settling (Cydzik-Kwiatkowska and Zielinska, 2016). Thus, the microbes within the activated sludge floc play an integral role in forming ecological interactions within the floc structure which contribute to flocculation, settling and nutrient removal.

The floc structure can be used as an indicator of process performance, whereby, ideal flocs may be as a result of moderate levels of filamentous bacteria that are retracted within floc (Martins *et al.*, 2004, Mesquita *et al.*, 2011). Moderate levels of filamentous bacteria are essential for the formation of stable, compact flocs and therefore efficient settling in the SST (Figure 2.4 c). During ideal operational conditions, the floc formers together with filamentous bacteria (Figure 2.4 c) form well-balanced compact flocs that settle efficiently. The flocs structural characteristics (*viz.* size, density, shape) and content of filamentous bacteria inevitably determine flocs stability and settling (Grady *et al.*, 1999, Govoreanu *et al.*, 2003, Smets *et al.*, 2006).

In WWTPs, quantitative changes between autotrophic, heterotrophic and filamentous bacteria are largely affected by influent wastewater characteristics, operational parameters and geographic location (Cydzik-Kwiatkowska *et al.*, 2012, Guo and Zhang, 2012, Mielczarek *et al.*, 2012, Zhang *et al.*, 2012, Burger *et al.*, 2017). Pin-point flocs occur during long sludge ages and low F/M ratios with very few or no filaments present. Pin-point flocs are usually identified by the formation of small activated sludge flocs which can be easily broken up and lacks a filamentous back-bone, thus forming inadequate settling properties that often result in a turbid effluent from the clarifiers (Figure 2.4 b) (Jenkins *et al.*, 2003). The various WWTP operational parameters and variations of the influent wastewater composition can also cause excessive filamentous bacteria. The profuse growth of filamentous bacteria interferes with flocs compaction and result in the formation of weak, open floc structures, thus hindering the solid-liquid separation process, commonly referred to as filamentous bulking (Figure 2.4 a) (Jenkins *et al.*, 2003, Lee *et al.*, 2003, Martins *et al.*, 2004, Burger *et al.*, 2017).



a) Filamentous bulking Activated sludge floc

- Filamentous bacteria dominant
- Filaments interfere with settling and compaction
- High SVI

b) Pin point flocs

- Low or no filamentous bacteria
- Weak small flocs
- Turbid supernatant
- Very Low SVI

c) Ideal, non-bulking activated sludge floc

- Filamentous and floc formers are in balance
- Strong, large flocs
- No inter-floc bridging
- Clear supernatant
- Moderate SVI

Figure 2.4. Floc structure and filamentous bacteria during ideal, non-ideal bulking conditions (Jenkins *et al*., 2003).

The distribution and prevalence of various filamentous bacterial types can vary from plant to plant and their presence are demonstrative of specific conditions within the system. Further knowledge on microbial functional groups in terms of their identity and physico-chemical properties is necessary to address cost-effective treatment strategies, by modifying their composition in the ASP (Rossetti *et al.,* 2017). For instance suppressing the growth of filamentous bacteria by selective control measures can mitigate bulking conditions. The remediation of bulking currently relies on the identification and monitoring of filamentous bacteria in the activated sludge systems, however, conventional identification methods can be misleading and may lead to contradictory results. Thus far there is still no universal solution to the bulking problem and these methods lack experimental verification (Mielczarek *et al.*, 2012, Zhang *et al.*, 2019).

To date, several filamentous bacteria are not classified and unavailable in pure cultures as they cannot be cultured by the use of current laboratory methods (Martins *et al.*, 2004, Rossetti *et al.*, 2017). Further understanding their ecophysiological traits had most often been ambiguous and unsuccessful, therefore preventing detailed studies of these microorganisms (Richard *et al.*, 2003, Martins *et al.*, 2004, Mielczarek *et al.*, 2012). Hence, the condition of the plant operation under which bulking sludge and excessive filamentous populations occurs is usually only marginally documented. Possible solutions to avoid the occurrence of the problematic filamentous bacteria can be found, by evaluating their ecophysiology, either in pure cultures or monitoring their *in-situ* characteristics in full-scale WWTPs. Knowledge of specific nutrient requirements and operational parameters that can enhance the activity of a single filamentous bacterial type may provide information in controlling their growth (Martins *et al.*, 2004).

In several distributed WWTPs from China, Singapore and USA, Guo and Zhang (2012) found that the percentage of filaments varied from 1.86 to 8.99% with the main filamentous bacteria identified as *N. limicola* I and II, *Mycobacterium fortuitum*, Type 1863 *Acinetobacter* and *M. parvicella* (Guo and Zhang 2012). Other studies found that *M.parvicella, H. hydrossis*-like bacteria and Types 0041, 0803 and 0092 belonging to phylum *Chloroflexi* as the predominating filamentous bacterial populations in municipal and domestic WWTPs (Mielczarek *et al.*, 2012, Cydzik-Kwiatkowska and Zielinska, 2016, Kowalska *et al.*, 2016). *Thiothrix* spp., *Gordonia* spp. and Type 1851 are generally prevalent in industrial WWTPs (Seviour and Nielsen, 2010, Nittami *et al.*, 2019). These filamentous bacteria proliferate in many different process designs and can cause severe bulking conditions. However, owing to their complex behavior and various

influential parameters, a thorough evaluation of these organisms in full-scale WWTPs can be challenging. Identification of filamentous bacteria that are responsible for bulking enables operators to select optimal solutions for bulking control in full-scale WWTPs (Kanagawa *et al.*, 2000, Ramothokang *et al.*, 2006a, Mielczarek *et al.*, 2012, Rossetti *et al.*, 2017).

The most successfully applied bulking control measures are supplementation of nutrients and the use of selectors. However, changes to the reactor design and plants operational process (*viz.* reducing SRT and loading, increasing temperature, increasing DO concentration, low sulphur concentrations, increasing the RAS and the changing reactor design to plug flow systems), were found to be successful measures (Azimi and Zamanzadeh, 2006, Kumari *et al.*, 2009, Yang *et al.*, 2009, Rossetti *et al.*, 2017). Despite several investigations on sludge bulking, the problem still occurs globally. Further knowledge about the WWTP, e.g. morphological and physiological characteristics of bacteria, flocculation, microbial interactions, as well as model systems to improve sludge bulking predictions is imperative to gain better insight about the bulking phenomenon (Martins *et al.*, 2004).

2.5. CURRENT TAXONOMY OF FILAMENTOUS BACTERIA

The taxonomy of filamentous bacteria may include three sub-disciplines *viz.* classification, nomenclature and identification. The classification step describes the arrangement of a group of microorganisms (*taxa*) based on their similarities and differences. Bacterial nomenclature is regulated by an international agreed set of rules, with distinctive names given for each taxon to allow effective communication among bacteriologists. For instance, the prefix "*Candidatus*" is used for uncultured microorganisms, with insufficient characterisation or information on its 16S rRNA sequence (e.g. *Candidatus Microthrix parvicella*). Identification compares an unknown to known isolates and can only be constructed after such isolates have been properly classified. Essentially a pure culture is required, however, to date very few filamentous bacterial isolates were cultured and properly identified. Hence, genotypic and phenotypic characteristics together are the best approach for proper classification of the filamentous bacteria, to reflect and ascertain much information about the organism's biological properties (Kaetzke *et al.*, 2005, Seviour and Nielsen, 2010, Jiang *et al.*, 2015, Rossetti *et al.*, 2017).

Identification of filamentous bacteria can be done by conventional microscopy techniques and are generally confirmed by molecular methods such as FISH with the application of fluorescently tagged oligonucleotide probes which specifically targets selective filamentous

bacteria using 16S rRNA gene sequence. Further studies focus on the rapid quantification of microorganisms, using molecular qPCR approaches, Sanger sequencing and next generation sequencing/metagenomics approaches. These methods are rapid and sensitive for detection and quantification based on using 23S and 16S rRNA gene sequence analyses (Kumari *et al.*, 2009, Seviour and Nielsen, 2010, Mielczarek *et al.*, 2013, Jiang *et al.*, 2015).

2.5.1. Filamentous Bacteria Identification Methods

2.5.1.1. Conventional Microscopy

Conventional microscopy methods are relatively simple and (Seviour and Nielsen, 2010, Rossetti *et al.*, 2017). Conventional microscopic identification is a simple, rapid and inexpensive approach to detect filamentous bacteria based on their morphological characteristics (*viz.* cell shape, presence of sheath, attached growth etc.). These methods are still commonly used for characterisation of filamentous bacteria. Microscopic examination to identify filamentous bacteria, following Eikelboom keys are based on wet mounts and staining methods (*viz*. Grams Stain, Neissers Stain, PHB test, sulphur test). The conventional method is still essential to date, for the elucidation of filamentous bacteria in WWTPs (Eikelboom, 2000, Jenkins *et al.*, 2003, Seviour and Nielsen, 2010, Rossetti *et al.*, 2017). In municipal WWTPs, approximately 30 filamentous morphotypes have been identified which mostly cause bulking and foaming. However, about ten of these morphotypes account for at least 90% bulking incidences (Jenkins *et al.*, 2003, Seviour and Nielsen, 2010).

Morphological identification of filamentous bacteria is imperative since information on most filamentous microbes are incomplete. Several research studies, depend on morphological identification methods of filamentous bacteria, to evaluate bulking control measures, since only a few species can be identified reliably (Adonadaga, 2016, Rossetti *et al.*, 2017, Perez *et al.*, 2018). Detailed studies on the ecophysiology of filamentous bacteria had revealed large variations among them, therefore, suggesting that individual control measures are necessary for each type. The ecophysiology of each individual filamentous type can assist microbiologists and engineers to control the prolific growth, however, thorough knowledge of the WWTP design, operational parameters and influent wastewater characteristics are further pivotal information in selecting the proper control measures, and so detailed monitoring of the bulking problem is recommended (Rossetti *et al.*, 2017).

2.5.1.2. Molecular Methods: FISH, qPCR and Metagenomics

Molecular methods such as FISH, real-time qPCR and metagenomics adds to information about the physiological characteristics of filamentous bacteria in wastewater treatment systems. An in-depth understanding of filamentous bacteria is needed to improve bulking in WWTPs.

Fluorescent *in situ* hybridization (FISH) is a useful molecular method that can be used to identify specific microbial cells, using 16S ribosomal RNA (rRNA)-targeted oligonucleotide probes (Amann *et al.*, 1995). The application depends on short oligonucleotide probes, which are described as synthetic sequences of nucleic acid that are designed to complement the target sequence, to a conserved region of RNA in a particular group of organisms (Coskuner, 2002). The probes can be designed to target different phylogenetic groups ranging from family to species level (Keller and Manak, 1989). The oligonucleotide probes are labelled with a fluorophore molecule also known as a fragment of the fluorescent dye (*viz*., fluorescein, tetramethyl rhodamine), which binds to the complementary rRNA target sequence. Fluorescein is one of the most popular fluorophores, which attach to the target region of the cells and these molecules emit light during excitation. The process requires optimum conditions for the hybridisation and washing steps with high temperatures to allow specificity of the probe binding. This allows for the colorimetric detection and quantification of the hybridized cells by the means of an epifluorescence microscope or by flow cytometry (Keller and Manak, 1989, Nielsen *et al.*, 2009b). Currently, the rapid FISH technique is available for a few filamentous bacterial species (Table 2.1).

Filamentous bacteria can undergo morphological changes and convert to a unicellular form, hence, FISH can be used as a reliable technique for identification (Carr *et al.*, 2005, Ramothokang *et al.*, 2006b). The disadvantage of the FISH technique is that it can only be successful if the fluorescently tagged probe can access and hybridise the rRNA oligonucleotide of the specific organism. Due to the high G+C content of the cell wall of certain bacteria such as *M. parvicella* and *Gordonia* spp., acid or enzyme treatment is usually required before hybridization as the probes have difficulty in perforating through the cell wall (Carr *et al.*, 2005). Moreover, only a few oligonucleotides are currently available to identify a few species of filamentous bacteria, whereas some morphologically identified filamentous bacteria (in both industrial and domestic WWTPs) could not be positively identified by FISH due to the lack of

probes available. The FISH probes target largely the phylotypes and not the individual filamentous bacteria species (Martins, 2004, Nielsen *et al.*, 2009b, Seviour and Nielsen, 2010).

Difficulties associated with FISH methods during image analysis also included photobleaching and interferences by autofluorescence. Photobleaching is caused by the destruction of the fluorophores at a high light intensity, whereas autofluorescence is a result of background noise or natural fluorescence caused by biological structures that can emit light (Nielsen *et al.*, 2009b). Hence, quantification of the targeted filamentous bacteria by image analysis can result in biased results (Moter and Göbel, 2000). Although filamentous bacteria such as *Saccharibacteria*, including those with filamentous morphotype, have been frequently detected in activated sludge by FISH analysis, the relationship between their abundance and bulking issues is still unclear (Takenaka *et al.*, 2018).

Andersen *et al.* (2019) and Nierychlo *et al.* (2018), designed FISH probe CFX64 which hybridized solely to filamentous bacteria Eikelboom Type 0092 morphotype, and is in agreement with previous *in situ* characterization of the Type 0092. The filamentous bacteria Type 0092 contributed to inter-floc bridging that is characteristic of filamentous bulking episodes. Since no pure cultures of Type 0092 are available, the current name *Ca. Amarolinea* for the species is listed in Table 2.1. Bulking problems in WWTPs were often observed, with a high abundance of Type 0092 (*Ca. Amarolinea*).

Table 2.1 lists the most common filament bacteria from municipal WWTPs that were identified via the FISH technique with selective 16S rRNA oligonucleotide probes. The Microbial Database for Activated Sludge (MiDAS) survey, further covers more than 50 Danish WWTPs, and gives a comprehensive overview of filamentous bacteria, indicating their 16S rRNA gene amplicons and current nomenclature (McIlroy *et al.*, 2015b, Nierychlo *et al.*, 2018). This platform contains critical information on their diversity, distribution, functional importance and morphology (McIlroy *et al.*, 2015a, Cydzik-Kwiatkowska and Zielinska, 2016).

Table 2.1. Current global taxonomy of filamentous bacteria isolated from municipal activated sludge, including 16S rRNA oligonucleotide probe sequence (MiDAS)

Phylogeny	Genus/Species name	Morphotype	16S rRNA gene sequence 5'-3'	References
Alpha-proteobacteria	*Neomegalonema*	Type 021N	CTG TCA CCG AGT CCC TTG C + CGG GAT GTC AAA AGG TGG	(Thomsen *et al.*, 2006)
Beta-proteobacteria	*Sphaerotilus natans*	*Sphaerotilus natans*	CAT CCC CCT CTA CCG TAC	(Wagner *et al.*, 1994a)
	Leptothrix discophora	*Leptothrix*	CTC TGC CGC ACT CCA GCT	
Gamma-proteobacteria	*Thiothrix nivea,*	*Thiothrix* spp.	CTC CTC TCC CAC ATT CTA	(Wagner *et al.*, 1994a)
	T. unzii,	*Thiothrix* spp.	*CCT TCC GAT CTC TAT GCA +	(Kanagawa *et al.*, 2000)
	T. fructosivorans,	*Thiothrix* spp.	*CCT TCC GAT CTC TAC GCA	(Kanagawa *et al.*, 2000)
	T. defluvii	*Thiothrix* spp.		
	T. eikelboomii	Type 021N group II	GCA CCA CCG ACC CCT TAG	(Kanagawa *et al.*, 2000)
	T. disciformis,	Type 021N group I	TGT GTT CGA GTT CCT TGC	(Kanagawa *et al.*, 2000)
	T. flexilis	Type 021N group III	CTC AGG GAT TCC TGC CAT	(Kanagawa *et al.*, 2000)
	Acinetobacter spp.	Type 1863	ATC CTC TCC CAT ACT CTA	(Wagner *et al.*, 1994b)
	Leucothrix mucor	*Leucothrix mucor*	CCC CTC TCC CAA ACT CTA	(Wagner *et al.*, 1994a)
	Beggiatoa spp.	*Beggiatoa*	GCC TTC CCA CAT CGT TT	(Manz *et al.*, 1992)

Phylogeny	Genus/Species name	Morphotype	16S rRNA gene sequence 5'-3'	References
Bacteroidetes	*H. hydrossis*	*H. hydrossis*	GCC TAC CTC AAC YTG ATT	(Schauer and Hahn, 2005)
Chloroflexi	*Calidilinea Amarolinea*	Type 0092	TCT ACC TAA GCA GAC CGT TC	(Nierychlo *et al.*, 2018, Andersen *et al.*, 2019)
	Candidatus Promineofilum	Type 0092, B45 group	TCC CGG AGC GCC TGA ACT/	(Speirs *et al.*, 2009)
			TCC CGA AGC GCC TGA ACT/	(Speirs *et al*., 2009)
			GGT GCT GGC TCC TCC CAG	(Speirs *et al*., 2009)
			AAT TCC ACG AAC CTC TGC CA	(Beer *et al.*, 2002)
	Kouleothrix aurantiaca	Type 1851	GGC TCC GTC TCG TAT CCG	(Beer *et al.*, 2002, Nittami *et al.*, 2019)
	Chloroflexi	Type 0041/0675/1851/1701	GGG ATA CCG TCC TTG TCT CT	(Nierychlo *et al*., 2018)
	Candidatus Sarcinathrix	Type 0914	TTG ACT CCG GCA GTC CCA CT	(Nierychlo *et al*., 2018)
	Candidatus Catenibacter	Type 0041	CCG CCA CTT TCA RGG ATA C +	(Speirs *et al.*, 2019)
			AWG TAC CCY CTC ACG TTC GAC	

Phylogeny	**Genus/Species name**	**Morphotype**	**16S rRNA gene sequence 5'-3'**	**References**
Actinobacteria	*Candidatus M. parvicella*	*M. parvicella*	GCC GCG AGA CCC TCC TAG+ CCG GAC TCT AGT CAG AGC	(Erhart *et al.*, 1997)
	Candidatus M. calida	Thin *M. parvicella*	TTC GCA TGA CCT CAC GGTTT	(Levantesi *et al.*, 2006)
	Gordonia amarae	*Gordonia amarae*-like organisms (GALO)	CAT CCC TGA CCG CAA AAG C	(de los Reyes *et al.*, 1997, de los Reyes *et al.*, 1998)
	Candidatus N. limicola	*N. limicola II*	GGC TCC GTC TCG TAT CCG/ CAA GCT CCT CGT CAC CGT T	(Liu and Seviour, 2001)
Firmicutes	*Tricoccus flocculiformis*	*N. limicola I*	CGC CAC TAT CTT CTC AGT	(Liu *et al.*, 2000, Trebesius *et al.*, 2000)
Planctomycetes	*Candidatus Nostocoida*	*N. limicola* type III	CCC AGT GTG CCG GGC CAC/ AGC ATC CAG AAC CTC GCT/ CCA TCGGC GAG CCC CCTA	(Liu and Seviour, 2001)

***probes target *Thiothrix* spp. and Type 021N groups I, II and III**

Real-Time qPCR methods enable the quantification of individual microbes from their 16S rRNA gene-specific copy numbers as compared to the total bacteria 16S rRNA copies present in AS samples. Different filamentous bacterial species were identified using species specific primers (*viz*. *Ca. Amarolinea*; *M. parvicella, Gordonia amarae, Candidatus Saccharibacteria,* (formerly known as candidate division TM7), *Thiothrix* spp. and *S.natans*. *Thiothrix* spp. degenerate primers for amplification of the chaperonin subunit (cpn60) were successfully applied and investigated in industrial pulp and paper WWTPs (Liao *et al.*, 2004, Kaetzke *et al.*, 2005, Vervaeren *et al.*, 2005, Vanysacker *et al.*, 2014, Asvapathanagul *et al.*, 2015, Andersen *et al.*, 2019).

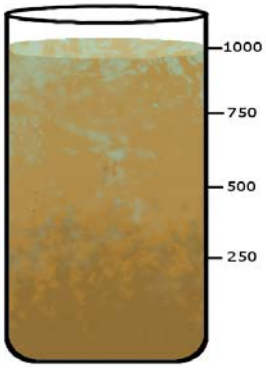
The real-time, *in situ* growth rates of specific filamentous species in complex environments, is essential to know for mathematical modelling and for investigating ecological interactions. If the growth rate of filamentous bacteria in activated sludge can be quantified in real time, researchers can more accurately assess the effect of operational parameters and wastewater characteristics and further improve the mathematical modelling of filamentous bulking (Nguyen *et al.*, 2016). Application of these conventional molecular techniques, however have difficulty in obtaining the complete profile of filaments. These techniques require the precise design of probes and primers, by trial and error approaches which are still lacking (Guo and Zhang, 2012, Cao and Lou, 2015).

Advanced sequencing methods such as for example high throughput sequencing is believed to be a promising method for monitoring filamentous bacteria. Cao *et al.* (2016), reported that the total filamentous bacteria could be identified and quantified through the high-throughput Ion Torrent PGM sequencing. Based on their study, *H. hydrossis*, Type 0411, *S.natans*, and Type 1863 were the dominant species and considered as possible bulking filaments in Macau WWTP (Cao *et al.*, 2016). Compared to qPCR and FISH methods, metagenomics sequencing provided more accurate data on the diversity of the microbial community, however, further studies are required for a better comparison (Guo and Zhang, 2012, Cao *et al.*, 2016).

2.6. SLUDGE BULKING

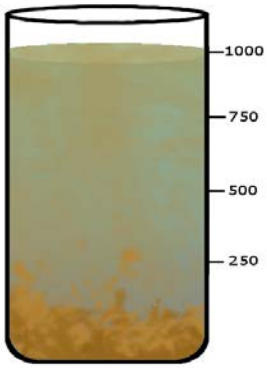
Bulking is a common operational problem in WWTPs, which is primarily due to the excessive growth of filamentous bacteria. The term bulking describes sludge which settles inadequately, as a result, solids overflow through the weirs of secondary clarifiers to the final effluent and a reduced sludge load is recycled into the system. This not only reduces the water quality of the final effluent, it also affects the removal of nutrients such as nitrogen and phosphorus, since biomass which is necessary for proper functioning is washed out from the system (Lakay *et al.*, 1988, Rossetti *et al.*, 2017). The bulking phenomenon is largely influenced by the floc size and sludge concentration together with the filamentous abundance. During ideal operational and environmental conditions, the filamentous bacteria forms a balance with the floc forming bacteria, thus improving the settling efficiency (Sezgin *et al.*, 1978, Sezgin *et al.*, 1982, Gerardi, 2006, Han *et al.*, 2018b). A sudden change in these parameters may favour the increased growth of filamentous bacteria, which may cause filaments to extend into the bulk solution, thus forming irregular and diffuse floc structures, as well as interfloc bridging (Burger *et al.*, 2017).

Globally, the SVI method is one of the most common, rapid measures which is widely used by researchers in science and engineering to determine the activated sludge settleability (Jenkins *et al.*, 2003, Zhang *et al.*, 2019). Sludge volume index (SVI), is an indicator commonly used in the wastewater industry to measure sludge bulking (Burger *et al*., 2017; Zhang *et al*., 2019). It can be defined as, the volume in millilitres, occupied by 1 g of sludge that settles after 30 minutes (Figure 2.5). A sludge volume index (SVI) of more than 150 mL g^{-1} is generally considered as a bulking sludge, while severe bulking sludge has an SVI of more than 250 mL g^{-1}. Generally, effective settling of sludge has SVI of 50-120 mL g^{-1} (Lou and Zhao, 2012, Han *et al.*, 2018a, Zhang *et al.*, 2019). Among the SVI methods, stirred SVI (small increments of stirring) and diluted SVI have also been used for the routine monitoring of settled sludge in WWTPs (Martins *et al*., 2004, Seviour and Nielsen, 2010).



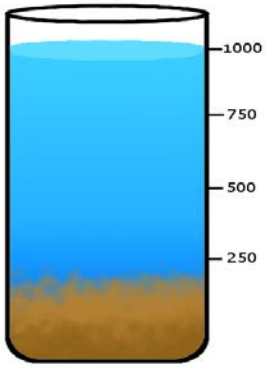
Severe sludge bulking – inadequate sludge settling

- Settling >250mL g^{-1}, may exceed TSS standards
- MLSS appear spongy or elastic in nature
- Turbid effluent



Rapid sludge settling

- Settling < 50 mL g^{-1}
- Sludge settles rapidly
- Turbid effluent



Efficient sludge settling

- Efficient sludge settling between 50-120 mL g^{-1}
- Clear effluent

Figure 2.5. Sludge volume index test describing inadequate, rapid and efficient sludge settling.

A limited number of studies further investigated flow and viscoelastic properties (e.g. sludge suspension) of the sludge settling characteristics via rheology meters (Wagner *et al.*, 2015, Wang *et al.*, 2016a). Wagner *et al.* (2015), investigated the impact of different filamentous bacterial species on the settling velocity and on the viscosity using rheology meters, for a better understanding of the sludge settling. However, results had some limiting factors and required further optimisation and calibration to achieve satisfactory rheological characterisation. In addition, the quantification of filamentous bacteria via total average length was inadequate, thereby requiring further development of new approaches for future studies.

2.7. SLUDGE BULKING CONTROL IN WASTEWATER TREATMENT PLANTS

Non-specific control techniques such as chlorination, hydrogen peroxide and ozonation is commonly used to alleviate sludge bulking in full-scale WWTPs, however, such methods can even harm the non-targeted microbial population. Moreover, they only offer short-term control of bulking problem and are costly for application in full-scale WWTPs due to increased chemical use (van Leeuwen, 1992, Jenkins *et al.*, 2003, Rossetti *et al.*, 2005). Chlorine is one of the first technologies proposed to control filamentous bulking and it is still commonly used as a disinfectant in effluent wastewater in the form of chlorine gas. However, the effects of chlorine on slow-growing bacteria such as nitrifiers take a long time to recover when affected by oxidants, hence disrupting the nitrogen removal process (Caravelli *et al.*, 2006). Additional disadvantages of chlorine treatment in wastewater plants are the formation of carcinogenic compounds and chlorine residuals such as trihalomethanes, which are toxic to aquatic life.

The use of ozone (O_3) treatment to reduce filamentous bulking has proven to be partially successful in full-scale WWTPS (van Leeuwen, 1988). The author further showed that no significant effect occurred on the BNR process *viz.*, nitrification and phosphorous removal. van Leeuwen (1988) also showed that an ozone dose of 2 mg O_3 L^{-1} reduced diluted SVI from 180 to less than 100 mL g^{-1}. Doses up to 30 mg O_3 L^{-1} had no effect on the nutrient removal process. The major advantage of O_3 is that it's relatively pure, which does not form harmful by-products that affect the environment (Wijnbladh, 2007). Studies have shown that O_3 improved sludge settleability and more effectively inhibit filamentous bacteria than the floc formers in the ASP (Saktaywin *et al.*, 2005, Sankaran *et al.*, 2008). However, O_3 is an unstable gas and has to be generated on-site for use (EPA, 1999).

Methods selective for specific filaments are often preferred, if the relationships between the real-time growth-rates of filamentous bacteria in addition to the wastewater conditions are well understood (Martins *et al.*, 2004). Therefore, knowing the *in situ* growth rates of the problematic filamentous bacteria in activated sludge would improve the use of specific control methods for predicting filamentous biomass levels (Rossetti *et al.*, 2017).

2.8. PLANT OPERATIONAL PARAMETERS, WASTEWATER CHARACTERISTICS AND ENVIRONMENTAL CONDITIONS AFFECTING FILAMENTOUS BACTERIA

The use of specific control methods aims to identify selective parameters which can favour the growth of problematic filamentous bacteria. Influent wastewater composition (*viz*. COD, NH_4, NO_3, NO_2, PO_4^{3-}-P), plant operating parameters (*viz*. DO, F/M ratio, pH, sludge age and MLSS) and environmental conditions such as rainfall and temperature could affect the microbial population (Martins *et al.*, 2004, Milobedzka *et al.*, 2016, Nittami *et al.*, 2019). Theoretically, filamentous bacteria predominate under long sludge ages, low DO, low F/M ratio and low nutrient conditions (Table 2.2) (Jenkins *et al.*, 2003, Martins *et al.*, 2004, Vaiopoulou *et al.*, 2007). Only a few studies had investigated the filamentous bacterial community compositions in full-scale wastewater treatment plants (WWTPs), in relation to these factors using predictive models. Mathematical modeling can be useful tools to study such complex ecosystems with several influential factors acting together (Martins *et al.*, 2004). Thus, knowing the conditions that enhance the activity of dominant filamentous bacteria can be useful in controlling their growth in the activated sludge system (Martins *et al.*, 2004, Tandoi *et al.*, 2006, Rossetti *et al.*, 2017). Further studies that can provide larger with appropriate modelling techniques are needed to evaluate factors which affect filamentous bacteria.

Table 2.2. Suggested control methods for problematic filamentous bacteria under specific operational conditions

Filamentous bacteria	Conditions	Suggested Cure Methods	Reference
H. hydrossis Types 0041 Type 0675 Type 0092 Type 1851 Type 021N	Low F/M long sludge age	Reduce MLSS in the aeration basin and increase the F/M. Increase the substrate concentration available to the AS, introduce batch or plug-flow characteristics to the aeration basin.	(Martins *et al*., 2004) (Seviour and Nielsen, 2010)
Type 1701 *S. natans* *H. hydrossis*	Low DO	Increase dissolve oxygen level in the plant (>1.5 mg L^{-1}), increase the sludge retention time (SRT).	(Richard *et al.*, 2003)
Thiothrix I and II. Type 021N	Nutrient deficiency (Nitrogen & Phosphorus) Septic sludge, presence of H_2S	Addition of lacking nutrient, chemical oxidation (*viz*., chlorine, hydrogen peroxide, potassium permanganate), or chemical precipitation (ferric chloride), Addition of lime or other alkaline agent to the aeration basin to raise the pH to 7.5. Control of influent waste septicity (organic acids and H_2S). Influent wastewater septicity can be treated by pre-aeration (which releases odors).	(Richard *et al.*, 2003)

Filamentous bacteria	Conditions	Suggested Cure Methods	Reference
M. parvicella *Gordonia amarae* *(*Formally known as *Nocardia amarae)*	Fatty acids (Grease and Oil) low F/M Low temperature	Still unclear but the most recommended solutions are: • Removal of lipids, grease and oil content by the use of a primary settling tank. • Increase oxygen level (>1.5 mg L^{-1}) (Note that higher aeration causes more foam formation, due to the physical action of more air present. Many operators reduce aeration when foaming occurs to reduce the foam, but this only causes more filament growth in the long term). • Minimise ammonium concentration (< 1 mg L^{-1}). • Reduce sludge age. • Dosage with chemicals (chlorine), however not always effective due to the high dose required. • Use of coagulants- polyaluminium chloride (PAX) dosages ranging from 1.5 to 4.5 g Al^{+3}/kg MLSS.	(Richard *et al.*, 2003) (Pal *et al.*, 2014)

To achieve the desired plant performance, a proper balance is imperative between the amount of organic matter (food) available, microorganisms and DO in the WWTP. The majority of problems result from an imbalance of these three parameters (Spellman, 2003, Rustum, 2009). Filamentous bacteria are largely affected by wastewater composition, nutrient concentrations, and environmental conditions that exist in the biological reactors.

2.8.1. Effects of F/M ratio and Sludge Age on Filamentous Bacteria

The time in which solids are retained within the systems is referred to as sludge age or solids retention time which generally are between 4 -10 days in a conventional AS plant. However, the sludge age in many BNR plants often exceeded 10 days, particularly during lower ambient temperatures to sustain the slow growing nitrifying population and to allow nitrification and denitrification to occur (Hu *et al.*, 2003, Ekama, 2010, Hu *et al.*, 2012). Floc forming microbes tend to utilise the nutrients hence forming relatively high biomass and as a result, a low F/M ratio may be formed. These observations directed to the accumulation of the low F/M filamentous population group (Table 2.2) (Martins *et al.*, 2004). Globally, several surveys showed the dominance of filamentous bacteria types *viz*. Type 0092, Type 0041, Type 0675, Type 1851, *Gordonia* spp. and *M. parvicella* which usually proliferate during low F/M conditions below 0.15 kg COD/kg MLSS d^{-1} in enhanced biological phosphorus removal plants (Hugenholtz *et al.*, 2001, Richard *et al.*, 2003, Kumari *et al.*, 2009, Speirs *et al.*, 2019). Milobedzka *et al.* (2016) reported that longer sludge age and low F/M positively correlated with Chloroflexi. A similar observation was made by Speirs *et al.* (2011), in Australian BNR WWTPs.

Research studies in South Africa, evaluated possible reasons for the proliferation of filamentous bacteria in SA WWTPs (Casey *et al.*, 1994b, Casey *et al.*, 1994a, Casey *et al.*, 1995, Ekama, 2010, Ekama, 2015). Studies have shown that proliferation of low F/M filamentous bacteria was generally improved by the continuous recycle of sludge between anoxic -aerobic conditions. However, these microbes proliferated during reduced intermittent aeration conditions and alternating anoxic-aerobic conditions. The significance of the anoxic selector is to reduce the readily biodegradable COD (RBCOD) following the aeration tank, thus also selecting for floc forming bacteria. However, high RBCOD in the anoxic tanks, had

increased filamentous growth (*viz*., *S. natans*, Type 1701, *H. hydrossis*, *Thiothrix* spp., *N. limicola* spp, Type 021N and Type 1851) (Xin *et al.*, 2008, Amanatidou *et al.*, 2016).

2.8.2. Effects of Nutrients on Filamentous Bacteria

Milobedzka *et al.* (2016) investigated key operational parameters in relation to filamentous bacteria of full-scale WWTPs with bulking problems, which indicated positive correlations of N and P with Type 1851. A similar observation was also found in a study conducted by Nittami *et al.* (2019). Investigations on the competition between filamentous bacteria and floc forming bacteria at different substrate concentrations showed specific growth rates of filamentous bacteria versus soluble RBCOD, whereby, both filamentous and floc formers differed. The growth rate for many filamentous organisms was higher than for floc formers at low COD concentrations, whereas at high COD concentrations, filamentous growth was inhibited (Chudoba *et al.*, 1973, Henze *et al.*, 2008). In WWTPs with low substrate concentration, filamentous bacteria proliferate and outcompete floc formers by growing profusely and extending out of the flocs. These filaments extract nutrients from the bulk liquid rather than within the flocs itself but the growth of floc formers are suppressed due to much lower nutrients diffusing within the flocs (Martins *et al.*, 2004, Henze *et al.*, 2008).

2.8.3. Effects of Dissolved Oxygen on Filamentous Bacteria

To maintain a balance, between organic matter available and diverse microbial populations in the WWTP, the DO concentration is usually maintained at 2 mg L^{-1}, as lower levels can limit the growth of floc formers and encourage the prevalence of filamentous bacteria. Filamentous bacteria generally proliferate at low DO (<0.4 mg L^{-1}) (Jenkins *et al.*, 2003, Davies, 2005). Filamentous bacteria Type 021N and Type 1851 and *Thiothrix* spp., were commonly observed at low DO levels (< 1 mg O_2 L^{-1}) (Martins *et al.*, 2003). However, some filamentous bacteria *viz*., *N. limicola* and *M. parvicella*, proliferate over a wide range of DO concentrations. Some reports on filamentous bulking associations to DO are often inconsistent and contradictory, since the performance of WWTPs can be affected by multiple factors. There are relatively few reports investigating the influence of multiple parameters on sludge settleability during low DO conditions (Rossetti *et al.*, 2005, Fan *et al.*, 2017a, Rossetti *et al.*, 2017).

2.8.4. Temperature Effects on Filamentous Bacteria

A report on the global microbial distribution of activated sludge WWTPs, showed that the *Chloroflexi* group was among the most frequent abundant filamentous bacterial populations especially during the summer and autumn months (Mielczarek *et al.*, 2012). "*M. parvicella*" was mostly found during winter and spring, in response to the lower mixed liquor temperatures (Rossetti *et al.*, 2005). Many research studies found that bulking varied seasonally, with highest SVI often occurring during colder temperatures (Kruit *et al.*, 2002, Parker *et al.*, 2004, Jones and Schuler, 2010). In both pure cultures and full-scale WWTPs, *M. parvicella* can grow at temperatures as low as 7°C, hence they have a competitive advantage during the cold season (Rossetti *et al.*, 2005). In Czech Republic, a system with a compartmentalized pre-denitrification zone had experienced severe bulking which was caused by *M. parvicella*, only when the mixed liquor temperature had dropped below 15°C for an extended period (Wanner, 1994). In Italy, similar seasonal patterns were reported for a WWTP treating domestic and industrial wastewater (Rossetti *et al.*, 2005). The Johannesburg Northern WWTP in South Africa, showed that *M. parvicella* occurred in the winter (<15°C) and Type 0092 dominated in the summer (Kumari *et al.*, 2009).

High temperatures decreased the growth rate of *M. parvicella*, even when the substrate concentrations of lipids and fats were not limiting. *M. parvicella* was eliminated at 29 °C, indicating that high temperatures not only decreased the availability of lipids and fats, but also allowed other bacteria to out-compete its growth (Rossetti *et al.*, 2017). The proliferation of *M. parvicella* at low temperature was suggested to be attributed to its effect in reduced solubility of lipids which concentrated on the surface of the aeration tanks. Therefore, lipids were more easily available for *M. parvicella* present in the scum layers, although this hypothesis has not yet been verified experimentally (Rossetti *et al.*, 2017).

2.9. PREDICTIVE MODELLING APPROACH TO CONTROL BULKING

The mechanistic description of the sludge bulking phenomenon in a general model is a challenge, since the kinetic growth properties of filamentous bacteria are difficult to assess. Sludge bulking control can be challenging, due to multiple variables such as influent wastewater composition and plant operational parameters. Hence, modelling could be used as

a tool to understand the complex behavior of filamentous bacteria under varying operational conditions. This information would assist the operators to preempt bulking and could assist in implementing appropriate measures to control bulking (Rustum, 2009). To circumvent the lack of a mechanistic understanding of sludge bulking, emerging soft computing techniques and their implication for practical application in full-scale WWTP are discussed. The input information of the WWTPs and the internal model is used in these soft computing methods, which returns the output information associated with the multiple factors that can influence sludge bulking (Han *et al.*, 2018a).

The understanding of sludge bulking via several modelling approaches has advanced since the last decade (Banadda *et al.*, 2011, Bagheri *et al.*, 2015, Chun *et al.*, 2017, Szeląg *et al.*, 2017). Predictive models can achieve early warning for filamentous sludge bulking which can be divided into the following categories, *i.e.*, white-box, gray-box and black-box models. The physical knowledge of a process is used in white-box modelling. Response data such as monitoring of a process and a universal model set can be described as black-box modelling. The black-box approach, often involves making assumptions, although physical parameters are not necessary (De Coninck *et al.*, 2015). Grey-box models refer to knowledge of how the system components interact but may not have detailed knowledge about internal program functions and operation. The Grey-box approach includes tools that cater for the situation whereby the prior knowledge of the process is not comprehensive enough for satisfactory white-box modelling and, purely empirical black-box methods do not suffice since the involved physical processes are too complex (De Coninck *et al.*, 2015). Grey and black-box models are also called inverse models which are relatively simple for the design of predictive models and they do not require a detailed understanding of the system, compared to white-box models. In addition, the latter models are cost-effective to perform prediction, which attempts to automatically capture the systematic processes and to link input to output variables, hence perceived as an alternative when mechanistic models are not available.

Prediction models for the operation, control and management of WWTPs can be approached from many classified model systems such as mathematical models (*viz*. ASM1-3) and statistical models (*viz*. knowledge-based systems, case-based reasoning, neural nets and hybrid approaches). This section highlights the evolution of applicable model systems which predicts sludge bulking events, considering the advantages, disadvantages and control strategies. Many studies focus on the predictions of water quality parameters of the influent or effluent

wastewater characteristics (Ranmin *et al.*, 2014, Bagheri *et al.*, 2015, Liu *et al.*, 2017, Han *et al.*, 2018b, Deepnarain *et al.*, 2019) (Table 2.3).

Table 2.3: Models used for prediction of filamentous bulking

Model	Parameters incorporated into model	Significant Findings	Country	References
Kinetic selection theory	Filamentous bacteria, Floc formers	Filamentous bacteria were found to contain a higher growth rate at lower substrate concentrations, hence, outcompeting floc formers under nutrient limiting conditions, however the kinetic growth rate did not apply to all filamentous bacteria.	Czechoslovakia	(Chudoba *et al.*, 1973)
The first mathematical model	soluble organic substrates, DO, floc shapes and sizes	This model predicted the average growth rate of filamentous bacteria (*Sphaerotilus natans*) over floc forming bacteria.	Taiwan	(Lau *et al.*, 1984)
ARX models	SVI, Filamentous length, Floc characteristics	Filament length showed a strong positive correlation with SVI values, however its predictive capacity was limited. Other combination of inputs was suggested.	Germany	(Banadda *et al.*, 2005, Smets *et al.*, 2006)
Risk assessment model	NH_4, NO_3, NO_2, and biodegradable substrates, F/M, DO	Severe risk for sludge settling problems and causes considered within the risk model: filamentous bulking were due to low DO, low organic loading, nutrient deficient conditions, low F/M. The model quantify the	Spain, Denmark, Sweden	(Comas *et al.*, 2008)

Model	Parameters incorporated into model	Significant Findings	Country	References
		risks associated with simulated control strategies and evaluates the severe risks leading the process towards microbiology-related separation problems.		
Three-layered feedforward ANN	Predicts SVI using temperature, COD, NH_4^+-N, TN, MLSS, SS, BOD, pH, and TP.	Determines complicated and nonlinear correlations, Requires long training times due to a large number of weights and neurons between layers.		(Lou and Zhao, 2012)
PCA	Predicts SVI using temperature, COD, NH_4^+-N, TN, MLSS, SS, BOD, pH, and TP.	Extracts useful information from large datasets, however, requires a data normalization step		(Lou and Zhao, 2012)
Cumulative logit model	Determines the effects of pH, COD, temperature, F/M, NH_4-N, DO, and PO_4^3-P on the abundance of filamentous bacteria.	Efficient for predicting discrete functions. It does not require data tuning. Important and independent variables should be used.	South Africa	(Deepnarain *et al.*, 2015)

Model	Parameters incorporated into model	Significant Findings	Country	References
Genetic algorithm based ANN	Predicts SVI using pH, DO, MLVSS, TSS, COD, TN, and temperature	Simulates complex functional relationships using an unrestricted number of inputs and outputs. Requires adjustment of the network parameters and connecting weights.		(Bagheri *et al.*, 2015)
Sensitivity analysis based ANN	Predicts SVI using MLSS, COD, TN, pH, and DO.	Attains long-range predictions even in the occurrence of data noise. Overfitting problem can occur if the error of testing is larger than that of training.		(Han *et al.*, 2018a, Han *et al.*, 2018b)
Decision Tree	Predicts SVI using pH, temperature, DO, SRT, F/M ratio, sCOD, tCOD, NH_4^+-N, TKN, PO_4^3-P, TP, and TSS.	Improves the prediction accuracy by considering the growing and pruning strategies. Requires large amounts of data records.	South Africa	(Deepnarain *et al.*, 2019)
Modified Random Forest	Predicts settleability on the basis of temperature and inflow	The results of analysis indicate that modified random forests demonstrate the best predictive abilities. Best results in predicting sludge settleability were obtained with temperature.	Poland	(Szeląg *et al.*, 2017)

2.9.1. Mechanistic Models

Mechanistic models describe the internal interactions between the hydrodynamic and sludge distribution in the ASP which are necessary for the accurate prediction of the SST performance (Raman, 2014). The activated sludge model (ASM) 1, was developed to describe the removal of carbon (C) and nitrogen (N) with the simultaneous consumption of oxygen and NO_3 as electron acceptors (Van Loosdrecht *et al.*, 2015). Organic matter measured in the form of carbonaceous oxygen demand (COD) was incorporated into this algorithm. Both nitrogenous and organic compounds are further fractioned to consider the biodegradability and solubility of the activated sludge process. However, over the last three decades, the removal of nutrients such as nitrogen and phosphorus from wastewater, has been further emphasised and the ASM model have advanced due to the increase in human population and industrialization (Jeppsson, 1996, Gernaey *et al.*, 2004). Thus, the model further extended to ASM2 and ASM3 which incorporated more fractions of COD removal and new processes for the enhanced biological phosphorus removal.

An elaborative review on the history of ASMs can be found in Jeppsson (1996) and Gernaey *et al.* (2004). Initially, the organic removal and sludge settling were the basic functions of the aeration and secondary settling tanks, respectively. However, not much focus was done on relating the specific identified filamentous bacteria to the sludge bulking problem. The lack of a general solution to bulking sludge led to further investigations on the microbial populations, specifically the dominant filamentous bacteria responsible for bulking (Martins *et al*., 2004). Models based on the first principals of ASM also exist for SSTs, however, due to a lack of a mathematical relationship between floc characteristics and sludge settling, they are still limited to process design and research purposes. Uncertainty in the model inputs and the simplified mathematical framework fundamentally may limit the accuracy of the model (Tchobanoglous *et al.*, 2003, Newhart *et al.*, 2019).

Initial development of models applicable to filamentous bacteria, began with the kinetic selection theory, elucidating their competition with floc-formers, which is perhaps still the most widely used theory in activated sludge. This empirical model used Monad kinetics to model the growth of microorganisms, based on the assumption that floc - formers and filamentous bacteria have specific growth rates for a given substrate (Chudoba *et al.*, 1973). Filamentous bacteria were found to have a higher growth rate at lower substrate concentrations, hence, outcompeting floc formers under nutrient limiting conditions (Chudoba *et al.*, 1973). However,

these theories still lack experimental verification. Following the modelling approaches to understand complex systems such as the ASP, Lau *et al.* (1984), further developed the first mathematical model for activated sludge bulking, which predicted the average growth rate of filamentous bacteria (*Sphaerotilus natans*) over floc forming bacteria. Parameters included in the model were soluble organic substrates, dissolved oxygen (DO), floc shapes and sizes which were used to predict the volume average growth rate selective for *S.natans*. However, the kinetic parameters were only selective for *S.natans* that was characterized in wastewater and lack a general framework for the total microbial selection (Table 2.3). Although many attempts are made by means of mathematical modelling (Monod Kinetics), to explain the proliferation of filamentous bacteria, these models still lack experimental validation (Cenens *et al.*, 2000, Seki *et al.*, 2004, Lou and de los Reyes, 2005).

One –dimensional (1-D) SST mathematical models are widely used for the dynamic simulation of WWTPs which further advanced into second order 1-D SST models. The first order 1-D models describes the solid separation process and predict the sludge blanket dynamics during wet-weather flow. The second order 1-D model includes the dispersion term in the mass balance, to overcome the numerical limitations of the first order, thus simulating the hydraulics of the SSTs in a much wider range of dynamic flow conditions. These models describe the hindered and compression settling, but they assume no variation in the sludge physical properties and thus the effect of filamentous bulking effects on the overall process performance was not considered. The practical application of these models is still a challenge in full-scale WWTPs due to several difficulties, such as lack of efficient solutions and reliable model calibration strategies (Flores-Alsina *et al.*, 2009, Li, 2016).

To further calibrate the above models, a global sensitivity analysis (GSA) model was developed which performed a comparative assessment of the sensitivity of WWTP model outputs (Ranmin *et al.*, 2014, Li, 2016). The identification and calibration of these models were developed using a computational fluid dynamic (CFD) tool. The CFD tool uses a computational software including applied mathematics and physics to analyse the flow of gases or liquids, which have been used to predict internal flow and transport of solids in the SSTs, mainly used for the purposes of designing, optimising or troubleshooting. The second order 1-D model best described the key output settling parameters which depended on influent flow and hydraulic parameters. However, these models have not been implemented to common practice in full-scale WWTPs. Moreover, further research and development are necessary for a more mechanistic based flow dependent hydraulic sub-model in the second order 1-D SST model.

The CFD model was used to identify sludge settling velocity and rheological characteristics which accounted for the impact of filamentous bulking. Parameters such as hindered, transient, and settling parameters of the settling velocity model were estimated in laboratory scale tests. However, these mathematical models lack a general mechanism that can explain the proliferation of selective filamentous bacteria and if each specific type taxonomically described is able to provide definitive strategies for bulking control (Alsina *et al.*, 2009, Ramin *et al.*, 2014).

The computational fluid dynamics (CFD) model further accounted for flocculants, hindered and compression settling characteristics, observed in the settling process of activated sludge, which described the flow behavior of a system (Karpinska and Bridgeman, 2016). Hydrodynamics and the distribution of sludge were predicted under varying filamentous populations. The results suggests that it is fundamental to account for transient and compression settling when modelling filamentous bulking. Wagner *et al.* (2015), reported that the impact of the different filamentous bacterial populations on the sludge settling velocity can vary. Quantification of the total average length of filamentous bacteria can carry limited information in characterising the effects of bulking on the sludge viscosity, thereby requiring new approaches. However, effective prediction of the biomass characteristics was limited and required regular monitoring of the process. Further testing and validation under significantly higher filamentous abundance are currently ongoing (Wagner *et al.*, 2015, Karpinska and Bridgeman, 2016).

2.9.2. Classical Control Method

Classical control methods are simple model systems which are verifiable with available on-line instrumentation. Proportional–integral–derivative (PID) and cascade control, are well known classical control methods for set point tracking of DO concentrations and feedforward controllers of influent ammonia to handle disturbance rejection both in a linear and non-linear manner (Amand, 2010). Classical control methods were successfully investigated and compared to other control strategies, however, these models cannot solve activated sludge separation problems and have many limitations (Vrečko *et al.*, 2003, Yong *et al.*, 2005, Zhang *et al.*, 2008, Amand, 2010). Classical control methods are limited to a single control technique and lack the knowledge about the mechanism involved in the ASP (Martínez *et al.*, 2003). Scientists argue that such a simple control strategy cannot guarantee process performance, as

the ASP is such a complex multidimensional, nonlinear system with interactions between several biological process variables (Amand, 2010).

2.9.3. Knowledge-based Systems

Knowledge based reasoning integrates all factors in an intelligent system for the supervision and control of wastewater systems. Studies of such complex systems is also currently being exploited by neural intelligence, logistic regression and decision support systems to model non-linear systems. Since the inception of the earlier deterministic models, much improvement has been made in the last decade. Significant effort had been emphasized on non-linear regression models, also, such as decision support systems for the prediction and control of sludge settling problems in complex biological systems. However, there are some limitations, due to lack of information from full-scale systems and these models require extensive data input (Okubo *et al.*, 1994, Zhu and Simpson, 1996, Serra *et al.*, 1997).

2.9.4. Case-based Reasoning

A case-based reasoning (CBR) technique for knowledge management in complex operational systems permits past historical experiences to solve new problems that can arise in a process and can be a potential support tool for WWTPs (Martínez *et al.*, 2003, Wiese *et al.*, 2003). Technical operators implement decisions for selective control methods of similar solids separation problems that are based on operational manuals and their past experience. The knowledge and historical experiences related to solids separation are generally collated in an accessible database which can be valuable information for operators and researchers. Figure 2.6 describes a CBR cycle which gives support to the WWTP operator, consisting of nine daily steps and three additional actions for each case study.

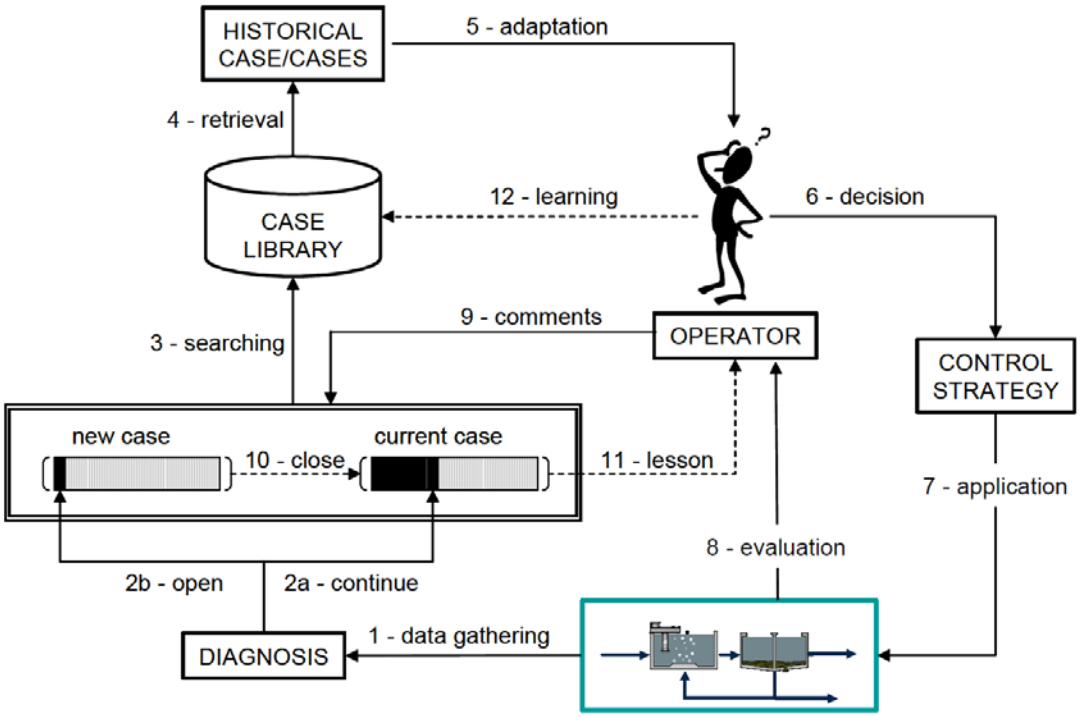


Figure 2.6. Diagram of the reasoning cycle of a CBR system (Martínez *et al.*, 2003).

In the first step, the data is gathered daily from the process to diagnose the process status. The current and new cases (step 2) are then opened. Step 3 explores the case library to search and retrieve (step 4) the historical cases which best match the current data. For this system 21 variables were considered. The operator then receives fundamental information from previous experiences such as process variables and control strategies (step 5), to establish the best control strategy (step 6) that can be further applied to react to the problem (step 7). The daily evaluation (step 8) of results and control strategy in the current case is registered via comments of the operator (step 9) (Figure 2.6). Steps 10 -12 describes the new case registered into the case library (Martínez *et al.*, 2003).

The CBR tool aims to provide critical information based on past experiences to engineers, operators and researchers. However, most investigations were presented as simplified daily case studies, and had serious limitations when faced with complex problems such as inadequate solids separation in the ASP. To this end, artificial intelligence began in the 1970s, to give support to operators when managing such complexities of WWTPs. Fuzzy control methods as

an extension to classical control was later used to handle the nonlinearities of the ASP both in set-point control and in aeration volume control. These methods had evolved to control aeration and plant operation (Amand, 2010).

2.9.5. Neural Nets and Hybrid Approaches

Artificial neural networks are the most commonly used for predictive modelling, adaptive control and applications whereby the dataset can be trained and validated. The ANN is a non-parametric models which is built with a three-layer feed-forward network consisting of an input layer, one or more hidden layers and an output layer, which connects, whereby all inputs are modified by weights. This in turn, can minimize the error between the network outputs and the targets. The number of hidden layers can be determined by trial and error. The nodes in each layer are connected by the weights and adjusted through a training process to obtain an optimised model. These weights are adjusted to minimize the error between the out-put and a given in-put data (Lou and Zhao, 2012). The ANN model utilise the interconnected nodes to form a network that can model complex functional relationships. The ANN model can analyse extensive input data to solve problems such as bulking, which allow non-linear simulations and good fit of the data when multiple factors are included in a given system. The neural network is essentially an algorithm which process information, based on the interconnection of nodes following a weighting system. The ANN model can solve a problem based on past experiences and heuristic knowledge, functioning like a human brain, whereby previously solved scenarios form neurons that makes new decisions, classification and predictions (Zou *et al.*, 2009, Abiodun *et al.*, 2018). A large data set with extensive input data is required to solve these intense problems, which further allows non-linear simulations and good fit of the data when multiple explanatory factors are included in a system (Ching *et al.*, 2018).

The ANN model showed a better performance to principal component analysis and multiple linear regression in terms of model fit statistics, when dealing with noisy and incomplete data (Noori *et al.*, 2010). Belanche *et al.* (2000) have modelled effluent TSS as an indication of plant performance with the ANN, based on both qualitative and quantitative variables, since high bulking could lead to a high TSS in the effluent clarifier (Belanche *et al.*, 2000). The authors showed that qualitative information such as microscopic examination, and process observations such as the presence of foam proved to be invaluable information for plant output performance. However, TSS is an insufficient representative of sludge bulking events,

moreover, the ANN model needed further validation and calibration. The earlier ANN models were difficult to compute with limited understanding or explanation in the hidden layer, due to its restricted black box approach. Proper monitoring of sludge bulking using methods such as SVI or rheology meters and monitoring of filamentous bacterial populations together with plant operational parameters and physico-chemical wastewater characteristics are more profound information that can provide a holistic approach to bulking predictions. Further validation and calibration of these models should be explored on larger data sets with different outcomes or significant explanatory variables (Wagner *et al.*, 2015).

Lou and Zhao (2012), predicted sludge bulking events using SVI as an output, that was applied to the ANN model, which further included influent factors such as temperature, pH, COD, BOD, SS, NH_4, TN, and MLSS. The authors advocated the model as a useful tool for bulking control. The data had fitted better with the ANN model as opposed to a principal linear regression (PCR) analysis, thus providing better prediction results with fitting accuracy. In general, the structures of ANN can be classified as feedforward neural networks (FNNs) and recurrent neural networks (RNNs). Bagheri *et al.* (2015) applied an artificial neural network (ANN) model with 76 data points to predict SVI (*i.e.*, output) using several inputs including DO, pH, TSS, COD, and TN. Their study found that SVI was considerably influenced by the influent TN (as a single input attribute), and TN and COD (as a group of two inputs). In recent years, the ANN model have been further modified and used for monitoring, control and simulation of activated sludge processes (Han *et al.*, 2018b, Han *et al.*, 2018c, Pisa *et al.*, 2019). Yu *et al.* (2015), developed ANN together with a Grey Markov model, and the results demonstrated that the proposed method can facilitate a real-time prediction of bulking sludge. This method was then modified to a self –organising recurrent radial basis function neural network (SORRBFNN) to predict SVI to detect bulking sludge (Figure 2.7) (Han and Qiao, 2013). These ANN models were based on fuzzy neural networks (FNNs) with backpropagation or its other variations to model nonlinear systems. However, one of the main drawbacks of FNNs is that they are essentially static input-to-output maps and their capability for representing nonlinear systems is limited. Floc structural characteristics and filamentous dominance associated with the sludge bulking, should be incorporated into these models for better predictions of sludge bulking events (Han and Qiao, 2013).

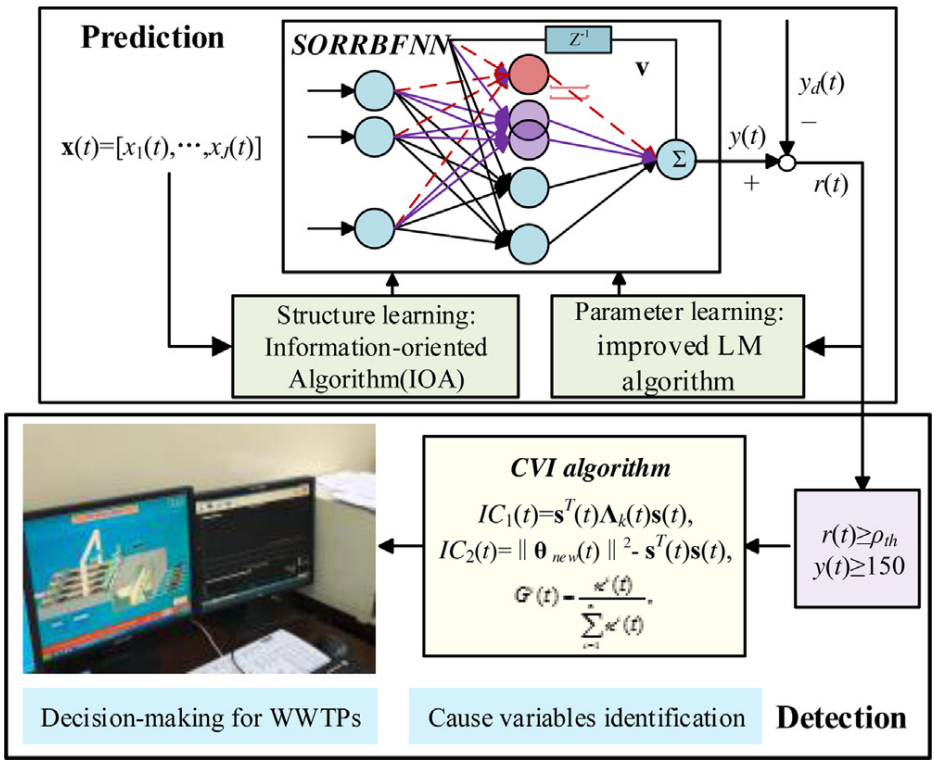


Figure 2.7. The flow chart of an intelligent detection method (Han *et al.*, 2018b).

Han *et al.* (2018b) later developed SORBFNN together with cause variables identification (CVI). The recurrent radial basis function neural networks (RRBFNN) provided long range predictions compared to FNN, even in the presence of measurement noise due to their structures. The information –oriented algorithm and the improved Lavenberg – Marquardt algorithms were used to adjust the structure and parameters of the SORRBFNN to improve the flexibility for predicating values of SVI with suitable accuracy and these models were better for modelling non-linear systems. The CVI algorithm was designed to determine the fault variable with high accuracy, which was capable of detecting the cause of bulking sludge and further successfully applied real-time in a full-scale WWTP. However, these models have an overfitting problem which can occur if the error of testing is larger than that of training (Han *et al.*, 2018b).

2.9.6. The Cumulative Logistic Model

The cumulative logistic model (CLM) had been successfully applied in social and biomedical sciences particularly dealing with categorical data (Park, 2005, Agresti, 2007). In a recent study, the model had been implemented with data from a full-scale biological nutrient removal plant, relating identified filamentous bacteria to plant operation and environmental conditions (Deepnarain *et al.*, 2015). The CLM can be described as an ordinal non-linear regression analysis used for predicting the outcome dependent variable and the probabilities describing the possible outcome are modelled as a function of explanatory variables, using a logistic function. The model was useful in predicting the abundance of filamentous bacteria under specific plant operational parameters (Deepnarain *et al.*, 2015). For instance, low DO (<1 mg O_2 L^{-1}) and F/M (<0.1 kg COD/kg MLSS d^{-1}) ratios predicted the dominance of Eikelboom Type 1851 (FI: ≥4), which also coincided with literature (Beer *et al.*, 2002, Martins *et al.*, 2004). As a result, the distribution of probabilities investigated indicated the role of DO and F/M over the other operational parameters mentioned in controlling Type 1851. Moreover, Eikelboom Type 021N was dominant during low COD (<750 mg COD L^{-1}) and high NH_4-N levels (≥32 mg L^{-1}), regardless of F/M and DO being at low or high levels, hence, indicating a strong influence of ammonia on Type 021N. However, these findings require further validation among multiple WWTPs (Deepnarain *et al.*, 2015).

2.9.7. Risk Assessment Model and Decision Trees

Comas *et al.* (2008) and others designed the risk assessment model, to alert plant operators of the solids separation problem in real-time. However, the usefulness for sludge bulking prediction was limited, since the model only provided a warning when the problem was already at the developed stage. The authors suggested reliable sensors and online data of the sludge settling characteristics. The integrated model evaluated simulations based on several control strategies and scenarios. The sludge settleability characteristics of SSTs were assessed to determine the level of risk, whereby input data included real-time on-line data from the plant, such as, NH_4, NO_3, NO_2 and biodegradable substrates. The results showed that the risk assessment model was able to detect the conditions of sludge bulking events in the SST. However, sludge settling parameters were not adjusted and most variables were measured inconsistently and infrequently to the point of sampling, due to the lack of proper on-line

sensors. The authors intended to configure a risk assessment model for microbiology solids separation problems, using the simulation output. This assessment was expanded from a previous version in conjunction with a panel of wastewater experts, to clearly classify the data in the form of Decision Trees. These models required extensive data and expert knowledge in the field (Comas *et al.*, 2008, Alsina *et al.*, 2009). Figure 2.8 shows the knowledge related to the risk of filamentous bulking proliferation which was synthesized into a Decision Tree that included three main branches (causes): low DO, nutrient deficiency and low F/M ratios and substrate limiting conditions. Additional work was required on the risk results and the thresholds triggering the bulking effect.

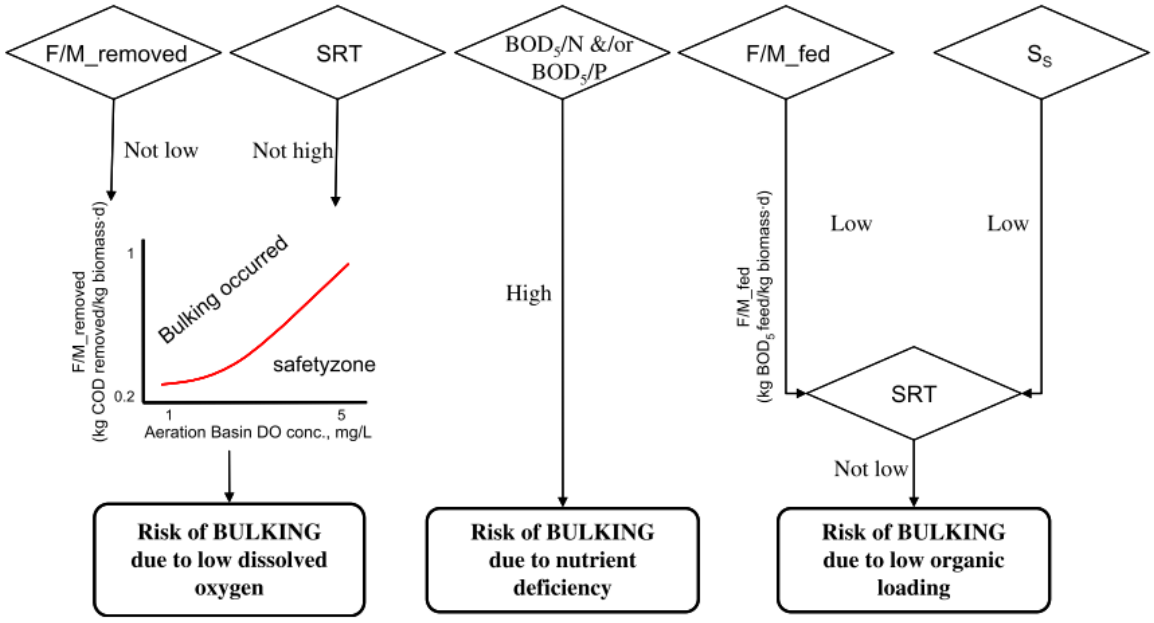


Figure 2.8. Developed Decision Tree, evaluating the risk of filamentous bulking sludge (adapted from Comas *et al.*, 2008).

The usefulness of the new risk assessment model demonstrated for objective comparison of several control strategies. However, this model can also be useful for other purposes, such WWTP design and optimization, scenario testing or operational improvement. Alsina *et al.* (2009), further investigated filamentous bulking sludge and its effects on the WWTP process performance, using the IWA benchmark simulation model No.2 and a risk assessment model which was also integrated into a knowledge based Decision Tree. Their study detected

favorable conditions for the development of filamentous bulking sludge. Results demonstrated that including the effects of filamentous bulking in the SST model, provided a more realistic plant performance.

Decision Trees are reliable and effective decision making tools in data mining which can evaluate large, complex data in the form of tree like structures, to infer useful patterns and provide high classification accuracy. This decision making tool can handle both categorical and numerical data, which have been successfully used in many fields of research (Shwe and Oo, 2016). Decision Trees create a type of flow chart consisting of nodes (referred to as "leafs") and a set of decisions to be made based of node (referred to as "branches"). The leaf and branch structure forms a hierarchical representation that mimics the form of the tree. Decision Tree learning is one of the most widely used and practical methods for inductive inference and is an important tool in machine learning and predictive analysis. A classification tree searches through each independent variable to find a value of single variable that best splits the data into 2 (or more) groups. Decision Trees can be computed very quickly and are simple to understand and interpret. Important insights can be generated, based on experts describing a situation (its alternatives and probabilities) and their preferences for outcomes.

To date, studies have shown data-driven methods that were developed as effective, alternative ways to predict sludge bulking, extracting data directly from the process (Figure 2.9) (Banadda *et al.*, 2011, Bagheri *et al.*, 2015, Chun *et al.*, 2017, Han *et al.*, 2018a). A dynamic auto-regressive exogenous (ARX) model was used to predict the SVI values. This model investigated as a function of organic loading and digital image analysis information. Results showed the online prediction of SVI values. The ARX model followed an input (image analysis) and output (SVI) simulations, also known as the black box approach, which uses the ARX command (Banadda *et al.*, 2005, Smets *et al.*, 2006, Alsina *et al.*, 2009). Small changes in the data of the applied Decision Tree model can drastically effected the structure of a tree. Hence, to improve the performance, many trees are fitted and predictions are aggregated across the trees.

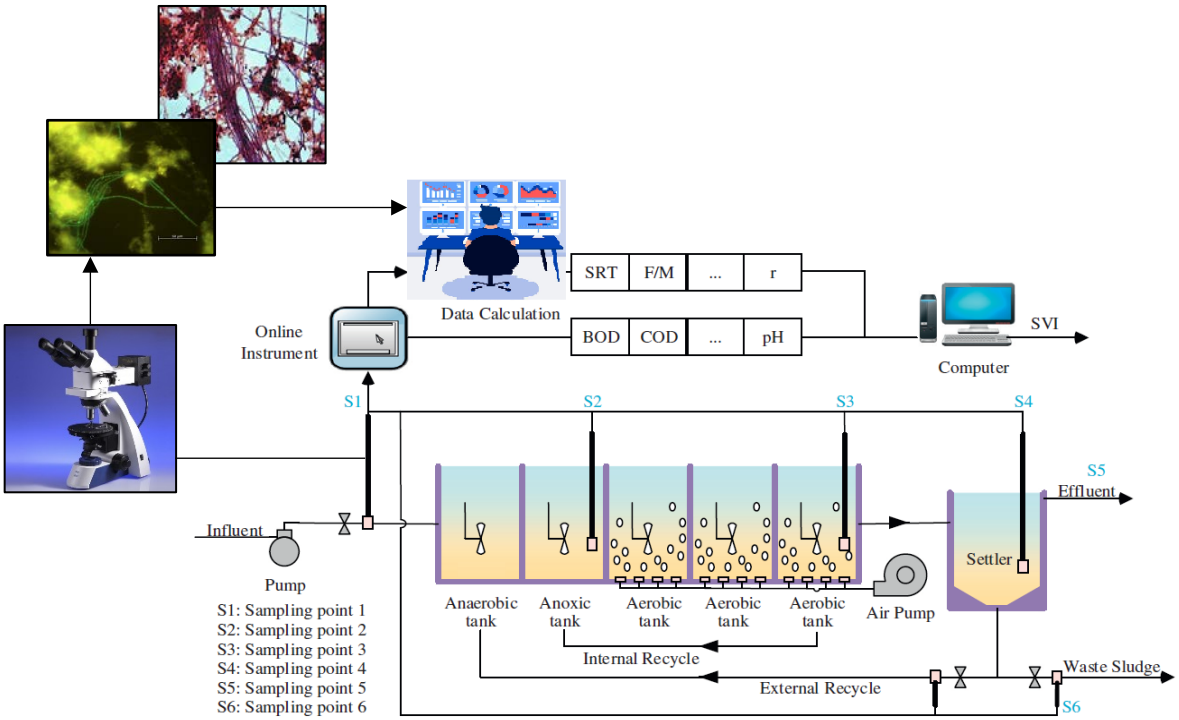


Figure 2.9. Data analysis and prediction of SVI (Han *et al.*, 2018a).

A Random Forest model takes a Decision Tree concept even further by producing multiple Decision Trees (Ali *et al.*, 2012). The approach first takes a random sample of the data and identifies a key set of features to grow each Decision Tree. Random variables are selected to grow a Decision Tree, then evaluate the performance of the tree is then evaluated. A second random set is subsequently selected and evaluated. This is done to observe how the first Decision Tree performed compared to the second tree and it continues to iterate through this process. An entire forest with random features, also known as a Random Forest is generated. The random forest model (RFM) can handle many variables, which does not require cross validation. It estimates which of the variables are important. In the final RFM, the most important variables can be detected via a mean decrease accuracy. Important risk factors that can detect low or high levels of bulking can be identified. For instance low F/M levels may cause a greater risk for bulking. However, operational parameters (such as DO, F/M ratios, sludge age), influent wastewater characteristics and filamentous bacteria concentrations may be the most important variables that can predict and control bulking events (Figure 2.9). The RFMs have a higher level of accuracy compared to other predictive algorithms (Pedrazzani *et*

al., 2016). Pedrazzani *et al.* (2016) applied a RFM to a WWTP located in Northern, Italy. The RFM indicated the most important variable which might provide additional insight to conventional indices such as sludge biotic index and sludge index. Both were modulating the weight of each covariate (namely, each microfauna component) and allowing interactions and nonlinearities. Therefore, it would be of interest to evaluate this method on WWTP data containing samples from different plants. Moreover, the determination of covariates contribution requires a significant monitoring period, were seasonal variability is also considered. Szeląg *et al.* (2017) had applied multivariate adaptive regression splines (MARS), random forests (RF) and modified random forests (RF+ SOM). The MARS model is a non-parametric form of regression method, whereby the relationship is a linear regression spline function of a conditional character. The MARS method is an extension of the classical approach of explanatory variables in the regression model, which automatically model the nonlinearities and the interactions between variables (Szeląg *et al.*, 2017). Random forests are classified within black box modelling techniques, an interesting solution is also increasingly used for modelling both the quantity and quality of upstream and downstream of the WWTP as well as forecasting processes taking place in different places of the plant. The results indicated that modified random forests demonstrated best predictive abilities. Recent developments are aimed at improving the quality of activated sludge settling models. However, the information on the growth kinetics of filamentous bacteria and their properties are still lacking (Szeląg *et al.*, 2017).

Among the prediction models, Decision Tree is a good method which has a high prediction performance. Compared to other data mining approaches, the Decision Tree model provides easy interpretation and understanding of the data, which are relatively quick and simple to build. Decision Trees can handle data from various scales and can be used for real time sludge bulking events (Nefeslioglu *et al.*, 2010, Khosravi *et al.*, 2018). Knowledge on Decision Trees application for filamentous bulking control in full-scale BNR WWTPs is scarce and require further developments.

2.10. QUANTITATIVE MICROBIAL RISK ASSESSMENT

Access to safe, sustainable water to support the livelihoods, human wellbeing, ecosystems and social-economic development (e.g. Water used for industry and agricultural irrigation) is of vital concern to the economy. Microbial pathogens present in surface water can cause serious health issues *viz*. gastroenteritis, amoebiasis, salmonellosis, dysentery, cholera, typhoid fever, hepatitis-A and diarrhea (Genthe *et al.,* 2013). Globally, one of the largest sources of surface water pollution is effluents from WWTPs, primarily from municipal WWTPs, though a few types are responsible for the majority of waterborne illnesses (Seto *et al.*, 2016). Common causes of diarrhoeal diseases and gastrointestinal infections are pathogenic bacteria (*viz*., *Escherichia coli* O157:H7, *Salmonella typhimurium*, *Campylobacter jejuni*, *Shigella dysenteriae*, *Vibrio cholera*, *Cryptosporidium* spp. and *Giardia* spp.) which are found in the final effluents of WWTPs, thereby posing a serious health risk (Teklehaimanot *et al*., 2015).

One of the fundamental importance of WWTPs is to protect the receiving environment by removing the microbiological pollutants. Wastewater systems require improved water management systems and operational strategies, to monitor and regulate the use of treated water in an efficient and safe way (Pereira *et al.*, 2019). The final effluent is generally further treated using disinfection methods such as chlorine, UV or ozonation (Vaezi *et al.,* 2004, Macauley *et al.,* 2006, Adeyemo *et al.,* 2019). However, these methods are not always effective, mainly due to the common operational challenges experienced by wastewater treatment facilities. The lack of appropriate infrastructure can lead to poor management of wastewater systems that may attribute to increasing risks in urban areas. Appropriate management of water resources are of valuable importance to achieve sustainability in social and economic development (Sdiri *et al.*, 2018). The United Nations Food and Agriculture Organization has reported that the agricultural sector is the largest user of water and wastewater globally, accounting for approximately 70% of water use worldwide (FAO, 2011), which illustrates the potential scale of this health risk (Kouame *et al.*, 2017). Farmers rely on natural surface water for irrigation, however such water is not treated or monitored regularly for proper use and are occasionally used untreated. Pathogens originating from the effluent of wastewater treatment facilities can be of detrimental health risks to farmers, consumers of irrigated crops and communities living in close proximity to wastewater resources.

Health risks associated with exposure to pathogens in wastewater or wastewater contaminated surface water may be assessed either directly or indirectly (Dickin *et al.*, 2016). One such indirect approach that has received global attention in assessing risks of microbial infections in water, food and the environment is the quantitative microbial risks assessment (QMRA) technique. QMRA is a tool built on modelling the potential human health risks from exposure to different pathogens (e.g. human pathogenic viruses, protozoa, and bacteria) contained in food and/or water (Haas *et al.*, 2014, Beaudequin *et al.*, 2016). It is a well-structured approach that integrates information and data on pathogen concentration in the hexposure medium (water or food) with mathematical models to examine the exposure and spread of microbial pathogens (Havelaar, 2012). The QMRA approach/tool is comprised of four interrelated steps;

2.10.1. Problem formulation:

The overall context of the risk assessment is presented in this section. It covers the reference pathogens, exposure pathways, hazardous events and health outcomes of interest. The boundary of the assessment is defined at his stage in order to determine the specific risk management question that will be addressed.

2.10.2. Exposure assessment:

This stage addresses the magnitude and frequency of exposure to the pathogen/s of interest via the identified exposure pathways and hazardous events.

2.10.3. Dose-response relationships:

This is the health effect assessment which links exposure dose to probability of infection or illness and the probability of morbidity and mortality.

2.10.4. Risk characterization:

This step of QMRA involves the combination of the three earlier steps to assess the health impact based for the pathogen, exposure route and dose response chosen.

2.10.5. Risk management:

Despite not initially considered a major step in QMRA, recent assessments have added risks management as a critical step in health management. This step presents possible ways of managing the risk estimates established in the QMRA process. It usually entails the presentation of risk reduction interventions.

The QMRA technique is a tool which can be used to evaluate potential human health threats, based on theoretical and practical assessments, incorporated into a risk management framework

(Westrell, 2004). The QMRA approach has been applied extensively by different researches evaluating risks of infection associated with wastewater treatment (Dias *et al.*, 2019), wastewater reuse (Beaudequin *et al.*, 2016, Soller *et al.*, 2018), surface water quality (Abia *et al.*, 2016, Petterson *et al.*, 2016) and drinking water systems (Mohammed and Seidu, 2019, Owens *et al.*, 2020).

CHAPTER THREE: THE ASSESSMENT OF DOMINANT FILAMENTOUS BACTERIA USING ARTIFICIAL NEURAL NETWORKS AND PRINCIPAL COMPONENT ANALYSIS

3.1. INTRODUCTION

The efficiency of the sedimentation process is critical in activated sludge WWTPs that governs the capability and potential of the entire treatment system (Nasr, 2018). Generally, the settling properties of sludge is described in terms of SVI, expressing the amount (in mL) occupied by 1 gram of suspended solids in the mixed liquor (usually 1 L sample) for 30 min (Martins *et al.*, 2004, Han *et al.*, 2016). For instance, activated sludge having SVI over 150 mL g^{-1} could be subjected to sludge bulking, which hinders the operation of the ASP (Deepnarain *et al.*, 2019). Low SVI values indicate that the activated sludge is dense, and thus, the biomass has high settleability in the final clarifier (Lou and Zhao, 2012). One of the main reasons for sludge bulking is mass proliferation of filamentous bacteria. In this context, monitoring and assessment of the filamentous abundance are important domains of research to attain proper management of WWTPs (Fan *et al.*, 2020).

The filamentous bacteria identification is mainly carried out based on the morphological and staining characteristics of microorganisms (Deepnarain *et al.*, 2015). These features include filament shape and length, trichome cell shape, Gram and Neisser staining, and detection of intra-cellular energy storage products such as poly -3-hydroxybutyrate (PHB) (Eikelboom, 2000, Jenkins *et al.*, 2003, Zhang *et al.*, 2019). However, conventional identification techniques are insufficient to distinguish between different filaments having the same morphological and phenotypic characteristics (Dias *et al.*, 2016, Abusam *et al.*, 2017). Alternatively, fluorescence *in situ* hybridization (FISH) has been successfully used as a rapid and highly sensitive technique to identify filamentous organisms in bulking activated sludge and for obtaining semi-quantitative data on filamentous bacteria (Nielsen *et al.*, 2009b, Nittami *et al.*, 2017). The FISH method could also be employed to evaluate the effects of operating conditions and physicochemical factors on the growth of filaments in WWTPs (Speirs *et al.*, 2017, Huber *et al.*, 2018).

Several filamentous types, including *Candidatus M. parvicella*, Type 1701, *Thiothrix* spp., Type 021N, *Gordonia* spp., Type 0092, Type 1701, Type 1851 and Type 0041, have been reported to cause severe bulking events in nutrient removal WWTPs, treating mainly municipal

activated sludge (Thomsen *et al.*, 2002, Lee *et al.*, 2003, Seviour and Nielsen, 2010, Xu *et al.*, 2014, Besha *et al.*, 2017). In South Africa, only three surveys have been reported on the identification and abundance of filamentous bacteria, whereby high populations of *M. parvicella* and Type 0041 were frequently observed in activated sludge systems. These surveys had found no associations between the filamentous bacteria, plants operational parameters and wastewater physico-chemical characteristics, which had shown their ability to adapt to wider parameter ranges (Blackbeard *et al.*, 1988, Bux and Kasan, 1994, Lacko *et al.*, 1999, Welz *et al.*, 2014).

The growth of each filamentous species is generally influenced by various operating conditions, *viz*., dissolved oxygen (DO) concentration, food-to-microorganisms (F/M) ratio, sludge retention time (SRT), and/or sulfur concentration and influent wastewater characteristics (e.g., pH and nitrogen and phosphorus concentrations (Martins *et al.*, 2004, dos Santos *et al.*, 2015, Deepnarain *et al.*, 2019, Fan *et al.*, 2020). Accordingly, the relative abundance of each filament is considered as a "fingerprint" that is influenced by specific environmental conditions in the treatment units (Lee *et al.*, 2003). Moreover, there are no universal strategies that could be used to control the excessive growth of filamentous populations (Xu *et al.*, 2014). Hence, the function and physiological behavior of filamentous bacteria in WWTPs still need comprehensive studies. This can be achieved with the assistance of advanced modelling and statistical methods (Fan *et al.*, 2018, Han *et al.*, 2019, Fan *et al.*, 2020).

Full-scale WWTPs comprise various physical, chemical, and biological processes, which are complex and difficult to describe using conventional linear models (Nguyen *et al.*, 2019). Artificial neural network (ANN) is an effective and powerful tool that has been widely used to handle and solve the non-linear relationships, even under fluctuating environmental conditions (Hamdy *et al.*, 2018, Han *et al.*, 2019). Large historical data obtained from real WWTPs can be imported into an ANN model for the prediction of non-linear behaviors with high accuracy and adequacy (Nasr, 2018). Moreover, the ANN model can be employed for process optimization, giving more robust and accurate results compared to regression-based mathematical models (Fawzy *et al.*, 2018). ANN has a parallel processing structure, in which several neurons or nodes is operated concurrently to mimic the biological nervous systems of humans (Bagheri *et al.*, 2015). This black-box model has an adaptive and self-learning ability to reveal the influence of each independent input on a dependent response (output) (Boztoprak

et al., 2015). Recently, ANN has been broadly applied to model and simulate the biological processes in WWTPs (Han *et al.*, 2018b).

Large data retrieved from WWTPs can also be interpreted and classified using principal component analysis (PCA) to extract valuable information (Abusam *et al.*, 2017). This multivariate approach is suitable for identifying correlations between random variables and describing essential features of WWTPs via data dimensionality reduction (Awolusi *et al.*, 2018). In PCA, the dimensionality of such datasets is reduced based on two axes, known as principal components (PCs). The first principal component (PC1) retains the largest variance direction in the data, whereas the second principal component (PC2) is orthogonal to PC1 and it represents the second maximum source of variation in the data (Cao and Lou, 2015, Nasr, 2018). The PCA technique can also be used to compress a set of data-points into specific numbers of clusters. These methods have recently found several applications for the assessment of bacterial abundance in full-scale WWTPs (Awolusi *et al.*, 2018).

The focus of this chapter is, therefore, to use ANN and PCA as tools to assess the dominance of filamentous bacteria in seven South African wastewater treatment processes. This chapter provides an overview of the prevalent filamentous bacteria that had the most impact on sludge bulking, using the ANN model, which ranks the filament species according to the order of importance. Further, PCA was employed to correlate the main features of each WWTP with the abundance of filamentous bacteria.

3.2. MATERIALS AND METHODS

3.2.1. Sampling from Full-scale Wastewater Treatment Plants

The seven selected wastewater treatment processes in this study are located in three South African provinces, *i.e.*, Gauteng, KwaZulu-Natal, and Western Cape. These plants were designed for the requirements of biological nutrient removal (BNR), *viz.*, UCT process configuration, three-stage Bardenpho process, modified Ludzack-Ettinger (MLE) process, and five-stage Bardenpho system. Table 3.1 lists the specifications and operational features of the selected BNR WWTPs. For the microbiological analysis, grab samples were collected from the end of the aeration tank (1L) and transported on ice to the laboratory and processed within 24 hours. The wastewater samples were collected monthly from the aeration tanks to determine sludge settleability, floc structures, and filamentous abundance.

Table 3.1: Process configuration, operational parameters and filamentous abundance from seven WWTPs in South Africa

	WWTP I	WWTP II	WWTP III	WWTP IV	WWTP V	WWTP VI	WWTP VII
Province	Western Cape	Western Cape	Western Cape	Gauteng	Gauteng	KZN	KZN
Configuration	5-stage Phoredox Process (Modified Bardenpho)	Modified University of Cape Town	Modified University of Cape Town	3-stage Phoredox process	3-stage Phoredox process	Modified Ludzack-Ettinger	Modified University of Cape Town
Wastewater source	Domestic	Domestic	Domestic and industrial (5%)	Domestic and industrial (5%)	Domestic and industrial (5%)	Domestic	Domestic Industrial (5%)
Aeration type	Surface aeration	Surface aeration	Surface aeration	Fine and course bubbles using new aeration diffuser system configured with new turbo blowers	Fine and course bubble using old aeration-single-stage centrifugal blowers, with replaced disc diffusors	Surface aeration	Surface aeration
Primary settling tank	No	No	No	Yes	Yes	Yes	No
ADWF Operational design capacity (ML d^{-1}):	12.5	7.5	14	45	40	25	7.2

	WWTP I	WWTP II	WWTP III	WWTP IV	WWTP V	WWTP VI	WWTP VII
Total Design Capacity							
Balancing tank	No	No	Yes	Yes	Yes	No	No
Anaerobic digester	Yes	Yes	Yes	No	No	Yes	No
Chemical dosing	None	None	None	Ferric chloride	Ferric chloride	None	Aluminum chloride
SVI (mL g^{-1})	155.85±7.20	165.05± 2.27	167.62±12.15	50.13±13.52	141.48±10.52	167.95±7.18	178.33±6.19
MLSS (mg L^{-1})	4837±637.55	5346 ±752.26	5465±657.11	4384±481.66	4514±526.05	5105±750	5071±597
Sludge age (days)	30	25	15	20	20	25	30
F/M (kg COD/ kg MLSS d^{-1})	0.13	0.11	0.14	0.11	0.11	0.14	0.18
floc characteristics	Excessive filamentous bacteria; open, irregular and diffuse floc structures	Excessive filamentous bacteria; open, irregular and diffuse floc structures	Excessive filamentous bacteria; open, irregular and diffuse floc structures	Lower filamentous abundance; compact flocs	Lower filamentous abundance; compact flocs	Excessive filamentous bacteria; open, irregular and diffuse floc structures	Excessive filamentous bacteria; open, irregular and diffuse floc structures

3.2.2. Filamentous Bacteria Identification

3.2.2.1. Conventional Staining and Microscopic Methods

The microbial community associated with sludge bulking was assessed using staining and microscopic methods. Initially, the morphotypes and staining reactions of filamentous bacteria were performed using the classification system of Eikelboom (Eikelboom *et al.*, 1998). Wet mounts, Gram and Neisser staining, cellular inclusions of poly-β-hydroxybutyrate (PHB), and sulphur storage test reactions were conducted to evaluate the physical factors (e.g., floc shape, size, and morphology), filament abundance, and the impact of filaments on floc quality.

In addition, cell shape, presence or absence of a sheath, and bacteria with attached growth were used to distinguish the filamentous organisms (Jenkins *et al.*, 2003). Further, the filament index was rated on a scale that ranged from 1 to 7, where 1 denoted no filamentous organisms, and 7 represented excess growth of filamentous organisms (Deepnarain *et al.*, 2019). For instance, the filamentous abundance was categorically ranked as “none”, “few”, “some”, “common”, “very common”, “abundant”, and “excessive” for filament index records of “1”, “2”, “3”, “4”, “5”, “6”, and “7”, respectively. A total of 15-20 random images were captured for each sample, to assess the abundance of each identified filamentous bacterial spp.

3.2.2.2. Fluorescent in situ Hybridisation

The preliminary identified filamentous bacteria based on morphological characteristics were further confirmed via FISH using 16S rRNA based specific oligonucleotide probes. The mixed-liquor samples were initially pre-treated, fixed, and dehydrated. Subsequently, the samples were treated with lysozyme (5 mg L^{-1}) and incubated at 25°C for 20 min. The pre-treatment step was conducted using 1x Phosphate-buffered saline (PBS) and 8% paraformaldehyde, and then the samples were washed in PBS and stored in PBS: ethanol (1:1, v:v) at -20°C until further use (Nielsen *et al.*, 2009b). The fixed samples were sonicated for 45 seconds at 2 watts using an ultrasonic liquid processor (XL-2000; Tri-Lab Support, USA) to shear and break up the floc structure into smaller pieces.

The selected filamentous bacteria were targeted using 16S rRNA-specific oligonucleotide probes, as listed in Table 3.2. The probes were labelled with a fluorescent dye 5-(and-6)-carboxyfluorescein diacetate N-succinimidyl ester, also known as 5(6)-FAM (CFSE; MWG-Biotech, Ebersberg, Germany; A subsidiary company of Roche Products (Pty) Ltd. South Africa). The EUB 338 probe, which is a general probe for most bacteria, was employed as the positive control, whereas hybridization without probe was applied as the negative control.

After sonication, 10 μL from the sample were plated on wells of Teflon-coated slides, and dried at 48°C for 10 min. The slides were subjected to a dehydration step using an ethanol series in 50-mL polypropylene tubes (3 min each in 50%, 80%, and 100% ethanol). This step is used to eliminate the excess water from the bacterial cells, improving the image resolution during microscopy.

The hybridization step was carried out under the conditions of high stringency (Nielsen *et al.*, 2009b). Hybridization buffer (Milli-Q water; 5 M NaCl; 1 M Tris-HCl; 10% sodium dodecylsulfate; 20 - 45% formamide) was prepared for each percentage of formamide. The prepared hybridization buffer together with the specific gene probe (50 ng μL^{-1} working solution) was added onto each well and incubated overnight at 46 °C. Following incubation, the slides were rinsed with warm distilled water and transferred onto a pre-warmed wash buffer (1M Tris/HCL; 10% SDS; 5 M NaCl; 0.5 M EDTA) in 50 mL polypropylene tubes. The tubes were incubated at 48°C for 45 min. After incubation, the slides were counter-stained with the DNA stain 4′, 6-diamidino-2-phenylindol (DAPI) (0.25 $\mu g\ mL^{-1}$). Further, each well was covered by a volume of 10 μL DAPI and maintained under dark condition for 10 min; *i.e.*, the slides were rinsed with distilled water and left to air-dry overnight. A drop of Vector Shield (mounting agent containing anti-bleaching) was added onto the slides and enclosed by a cover slip.

Table 3.2: 16S rRNA based oligonucleotide probes and target microorganisms (Nielson *et al*., 2009)

Probe	**Sequence (**5'-3')	**FA (%)**	**Target microorganism**	**Reference**
EUB 338	GCTGCCTCCCGTAGGAGT	35	Most bacteria	(Daims *et al.*, 1999)
EUB338-II	GCAGCCACCCGTAGGTGT	35	*Plantomycetales*	(Daims *et al.*, 1999)
EUB338-III	GCTGCCACC CGTAGGTGT	35	*Verrucomicrobiales*	(Daims *et al.*, 1999)
G123T	CCT TCC GAT CTC TAT GCA	40	*Thiothrix* spp.,*T. nivea, T. unzii, T. fructosivorans, T. defluvii,* Type 021N group I, II, III	(Kanagawa *et al.*, 2000)
G1B	TGT GTT CGA GTT CCT TGC	30	Type 021N group I	(Kanagawa *et al.*, 2000)
G2M	GCA CCA CCG ACC CCT TAG	35	Type 021N Grp II	(Kanagawa *et al*., 2000)
G3M	CTCAGGGATTCCTGCCAT	30	Type 021N Grp III	(Kanagawa *et al*., 2000)
21N	TCC CTC TCC CAA ATT CTA	35	Type 021N	(Kanagawa *et al*., 2000)
CFX223	GGT GCT GGC TCC TCC CAG	35	Type 0092	(Speirs *et al.*, 2009)
CFX197	TCC CGG AGC GCC TGA ACT	40	Type 0092	(Speirs *et al*., 2009)
GNS B941	AAA CCA CAC GCT CCG CT	35	Type 0041, Type 0675, Type 0092	(Gich *et al.*, 2001)

Probe	**Sequence (**5'-3')	**FA (%)**	**Target microorganism**	**Reference**
CFX 1223	GGT GCT GGC TCC TCC CAG			(Speirs *et al.*, 2009)
SNA	CAT CCC CCT CTA CCG TAC	45	*Sphaerotilus natans*	(Wagner *et al.*, 1994a)
CHL 1851	AAT TCC ACG AAC CTC TGC CA	20	Type 1851	(Beer *et al.*, 2002)
MPAMIX	GGA TGG CCG CGT TCG ACT GCC	20	*Candidatus M. parvicella*, *Candidatus*	(Erhart *et al.*, 1997)
(MPA60	GCG AGA CCC TCC TAG		*M. calida*	
MPA223	CCG GAC TCT AGT CAG AGC			
MPA645)				
NLIMI 91	CGC CAC TAT CTT CTC AGT	20	*N. limicola I*	(Liu and Seviour, 2001)
NLIMII 175	GGC TCC GTC TCG TAT CCG	40	*N. limicola II*	(Liu and Seviour, 2001)
NLIMIII 301	CCC AGT GTG CCG GGC CAC	20	*N. limicola III* strains	(Liu and Seviour, 2001)
G.am205	CAT CCC TGA CCG CAA AAG C	30	*Gordonia amaerae*	(de los Reyes *et al.*, 1998)
Myc657	AGT CTC CCC TGY AGT A	30	*Mycobacterium* subdivision Mycolata	(Davenport *et al.*, 2000)

*FA – Formamide 20 - 45%, for adequate hybridisation for probe binding, an optimum formamide concentration is used, thus, preventing the loss of fluorescence signal in the target cell (Pernthaler *et al.*, 2001)

The slides were visualized under an AxioLab ApoTome microscope (Carl Zeiss, Germany) containing FLUOS Fluorochrome/Filter Set. Image analyses were performed by AxioVision imaging software (Version 4.6, Carl Zeiss Microimaging GmbH, Germany).

3.2.3. Sludge Settling Test

For the sludge volume index (SVI in mL g^{-1}) test, a mixed liquor sample with a known volume of 1 L was allowed to precipitate in a measuring cylinder for 30 min. Subsequently, the volume of settled sludge per gram of solids was recorded. Sludge volume index of ≥150 mL g^{-1} is usually an operational measure used to determine sludge bulking (Jenkins *et al.*, 2003, Boztoprak *et al.*, 2015, Deepnarain *et al.*, 2019). SVI was calculated using equation 3.1 according to Jenkins *et al.* (2003).

$$SVI = \frac{V \times 1000}{MLSS} \qquad \textbf{Equation 3-1}$$

where, V is the settled volume of sludge after 30 min (mL L^{-1}), 1000 is a conversion factor (mg g^{-1}), and $MLSS$ is the mixed liquor suspended solids (mg L^{-1}).

3.2.4. Artificial Neural Network Modelling

The ANN is built from a set of interconnected elements (*i.e.*, units, neurons, or nodes), which are arranged in multiple layers. In this study, a feed-forward ANN model was used to predict the SVI data using ten inputs of filamentous bacteria, namely Type 0092, *Thiothrix* spp., Type 1851, *M. parvicella*, Type 1701, *N. limicola*, Type 0041, *Sphaerotilus natans*, *Gordonia* spp., and Type 021N. The ANN system is trained in a proper way, whereby it maps the input values to output values, according to causes and effects, which works by a pattern recognition system. The ANN model builds a pattern with all the numbers in the input layer and ANN recognizes the pattern and memorizes the target values.

The ANN structure is composed of (a) an input layer that collects the records from the 10 input attributes, (b) a hidden layer containing 15 neurons, and (c) an output layer designating the model output (*i.e.*, SVI). Hence, the network configuration is defined as 10 – 15 – 1. This

configuration was chosen based on several trial-and-error procedures. The weights and activation functions were used to transform the data, in which the Levenberg-Marquardt optimization was used for the training stage to adjust the weight and bias values. The data points were distributed into 70% for training, 15% for cross-validation, and 15% for testing the ANN model.

The ANN computations were performed using MATLAB (R2015a), according to the following steps:

The input layer comprised ten neurons that connected the input signals to fifteen neurons in the hidden layer. The neurons in the input layer ($x_{10\times1}$) were multiplied by a weight matrix ($w_{15\times10}$) and added to a constant bias ($b_{15\times1}$). The results were summed up and adapted by a tan-sigmoid "Tansig" activation function to limit the product between −1 and 1. The outputs of the hidden layer ($y_{15\times1}$) were estimated by Equation (3.2).

$$y_{15\times1} = \text{Tansig}\left(\sum w_{15\times10} \cdot x_{10\times1} + b_{15\times1}\right) \qquad \textbf{Equation 3-2}$$

Similar to the previous step, each node in the middle layer was interconnected to that of the last layer. The responses of the hidden layer ($y_{15\times1}$) were multiplied by weights ($w_{1\times15}$) and added to a constant bias ($b_{1\times1}$). The product was summed up and transformed into an output ($z_{1\times1}$) via a linear "Purelin" activation function. The output of this stage is represented by Equation (3.3).

$$z_{1\times1} = \text{Purelin}\left(\sum w_{1\times15} \cdot y_{15\times1} + b_{1\times1}\right) \qquad \textbf{Equation 3-3}$$

3.2.5. Principal Component Analysis

Filamentous bacteria were monitored in seven wastewater treatment processes and then described as categorical data (FI: 0-6) using multivariate analysis. The PCA was conducted using the sequential steps of Awolusi *et al.* (2018): (a) data standardization into standard deviation = 1 and mean = 0 (Appendix 2; Figure 1), (b) computation of covariance matrix, (c)

estimation of eigenvectors and eigenvalues (Appendix 2; Figure 2), (d) calculation of principal components (*i.e.*, PC1 corresponds to the largest eigenvalues), and (e) reduction of dataset dimensions to enhance the quality of representation. The results of PCA were then used for clustering via the k-means algorithm. These steps were performed using the function "Princomp", in the MATLAB (version R2015a) software package.

3.3. RESULTS AND DISCUSSION

3.3.1. Subjective Scoring of Filamentous Bacteria Abundance

Figure 3.1 shows the wet mount observations of the floc structural characteristics during different bulking conditions. The filamentous bacteria formed bridging among the flocs and enlarged the surface area, leading to a loose and less compacted structure. The filaments scorings for the occurrence and dominance of filamentous bacteria were based on the FI scale (FI: 1-6) (Jenkins *et al.*, 2003, Lee *et al.*, 2003). Figure 3.1a shows the presence of filamentous bacteria with the classifications of "most dominant" and "excessive". The shape and strength of flocs were irregular, diffuse, and weak, corresponding to subjective scores of 5 and 6. During "excessive" bulking conditions, filamentous microorganisms tend to protrude from activated sludge flocs, forming "filament-to-filament" or "floc-to-filament" aggregates with low density (Jin *et al.*, 2003). Figure 3.1b represents the "dominant" filamentous bacteria, having irregular and open flocs, classified as FI score of 4 with filaments in all flocs at medium density (*i.e.* 5 – 20 filaments per floc). Figure 3.1c illustrate the subjective scores of 2 and 3 (*i.e.* "few" and "common" respectively) with a low density of 1 – 5 filaments per floc. Figure 3.1d illustrates the "rare" filamentous bacteria, having a subjective FI score below <2. In this group, the floc structures are more firm, compact, and spherical in shape (Jenkins *et al.*, 2003, Lee *et al.*, 2003).

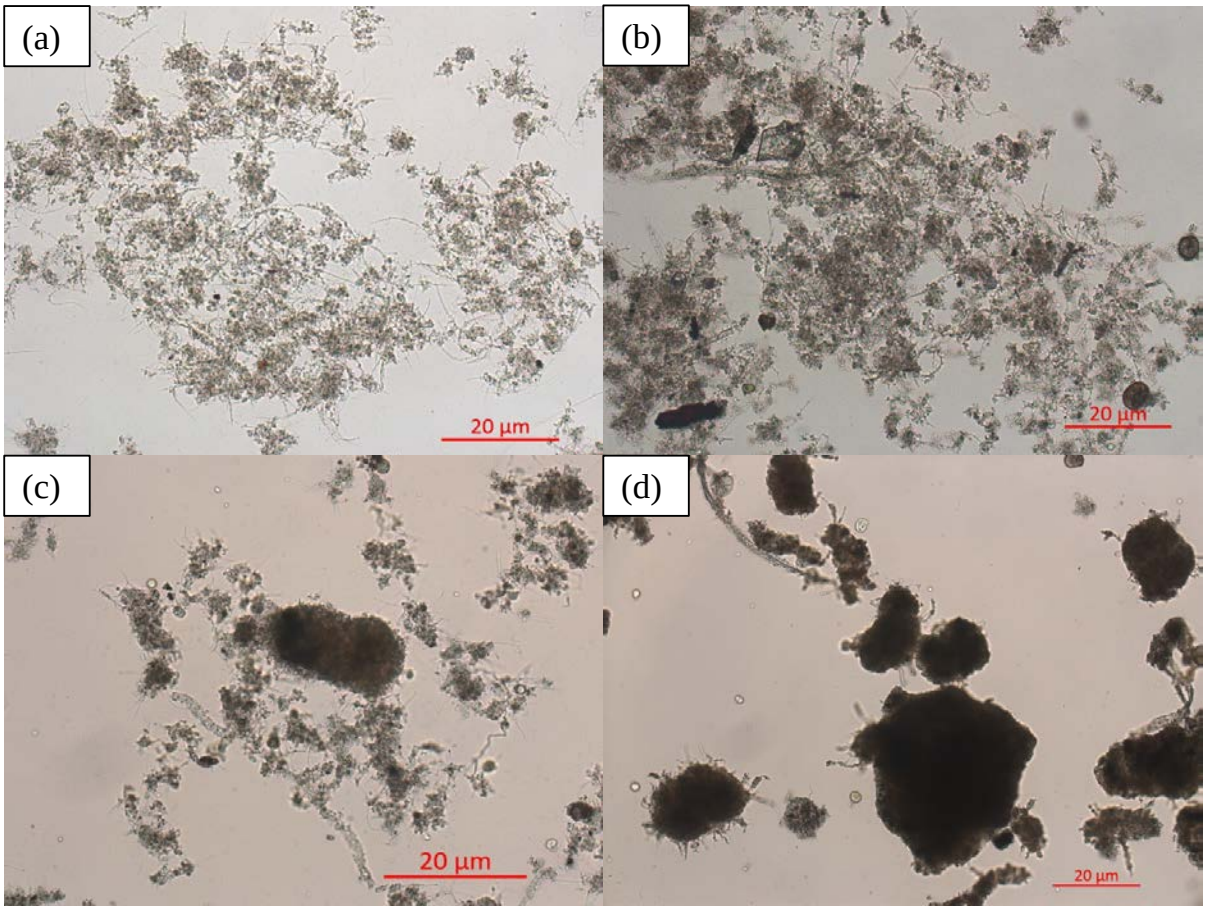


Figure 3.1. Wet mount observation of floc structural characteristics: (a) irregular and diffuse flocs with very common filamentous bacteria (b) “excessive” filamentous bacteria, indicating irregular and open flocs, (c) “few” to “common” filamentous bacteria, within more compact flocs, and (d) “rare” filamentous bacteria within compact, firm flocs.

3.3.2. Standard Staining Techniques (Gram and Neisser) for Filamentous Detection

Figure 3.2 shows Neisser and Gram staining reactions of filamentous bacteria. Figure 3.2a demonstrates Neisser positive of the dominant filamentous bacteria Type 0092 detected mainly within flocs, in which both filament branching and attached growth were absent. Moreover, Figure 3.2b displays Gram positive *Candidatus M. parvicella* with coiled and twisted morphology. Figure 3.2c indicated that Type 0041 and Type 1701 appeared with abundant attached growth (*i.e.* the bacterium Type 0041 is defined as a Gram variable, sheathed, curved, and unbranched filament) (Williams and Unz, 1985, Nittami *et al.*, 2014). Thomsen *et al.* (2002) found that Type 0041 appeared in about 88% of WWTP samples, indicating the “most abundant” filamentous bacterium with attached growth in activated sludge. In Figure 3.2d,

Type 021N appeared as Gram negative and unbranched, and the septa was clearly visible with the absence or very few of attached growth forms. The features of Type 021N have also been reported in a previous study by Guo *et al.* (2012) in a continuous activated sludge system.

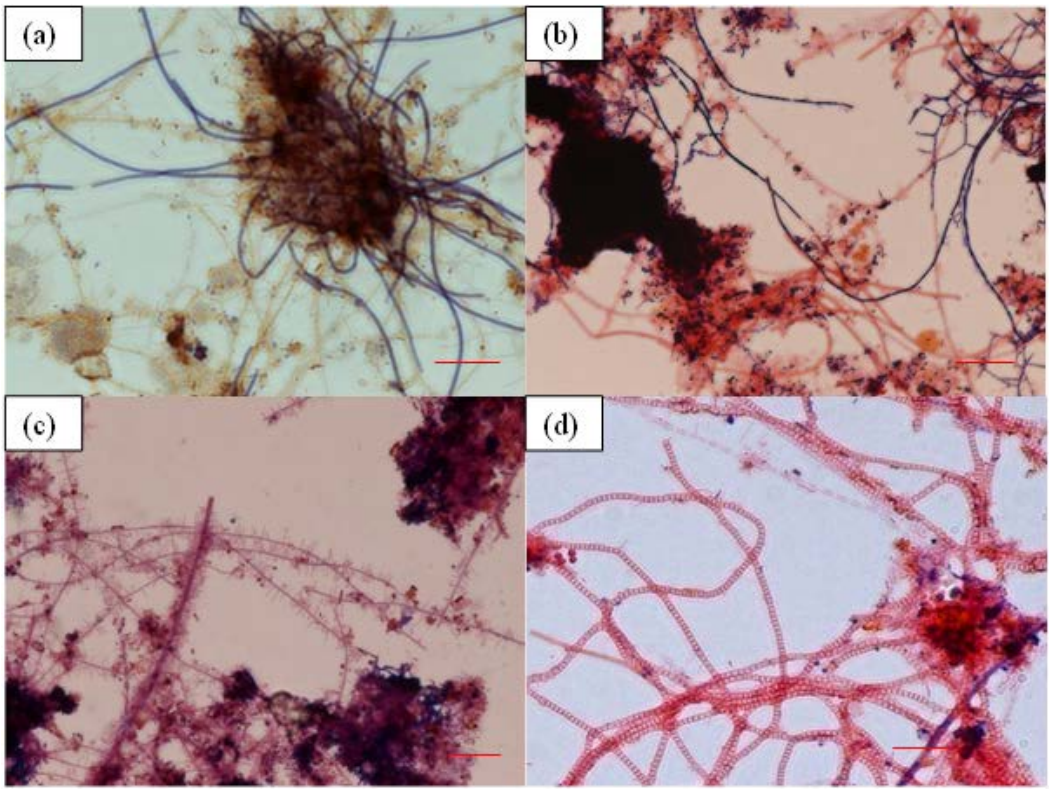


Figure 3.2. Images of filamentous bacteria at 20 μm: (a) Neisser positive Type 0092 found mainly within flocs, (b) Gram positive *M. parvicella* (c) Type 0041 and Type 1701 with abundant attached growth, and (d) Gram negative Type 021N, showing visible septa.

3.3.3. Fluorescence *in situ* Hybridisation Images of Filamentous Bacteria

An optimal sonication power of 2W for 45 seconds was observed to loosen up the floc structure without causing severe damage to the filamentous structure. The sonication step was done to break up the flocs, which also proved to be a successful pre- treatment method for FISH analysis. Ramothokang *et al.* (2003) used sonication at 20W for 10 seconds, however such high sonication treatment was done to break the filaments to the same size as floc forming bacteria for further isolation and cultivation of filamentous bacteria.

The FISH images were used to ensure the identification and abundance of specific filamentous species (Figure 3.3). For instance, the presence of Type 0092 was confirmed using the Phylum-targeted CFX-MIX probes (Figure 3.3 a) and DNA staining with DAPI (Figure 3.3 b). The filament *M. parvicella* was observed using MPA-MIX probe (Figure 3.3 c) and DNA staining with DAPI (Figure 3.3 d). Similarly, Lemmer *et al.* (2005) applied the FISH technique to detect Type 0092 filaments using CF319 probe and *M. parvicella* using MPA-mix (*i.e.*, MPA60, MPA223, and MPA645). In addition, Graveleau *et al.* (2005) detected Type 0092 and *M. parvicella* using both standard staining techniques (Gram, Neisser, and PHB) and FISH probes (Graveleau *et al.*, 2005). Similar observations have also been reported by You and Sue (2009), suggesting that the FISH and DNA sequencing methods could be more appropriate for estimating the filamentous microbial diversity in activated sludge WWTPs.

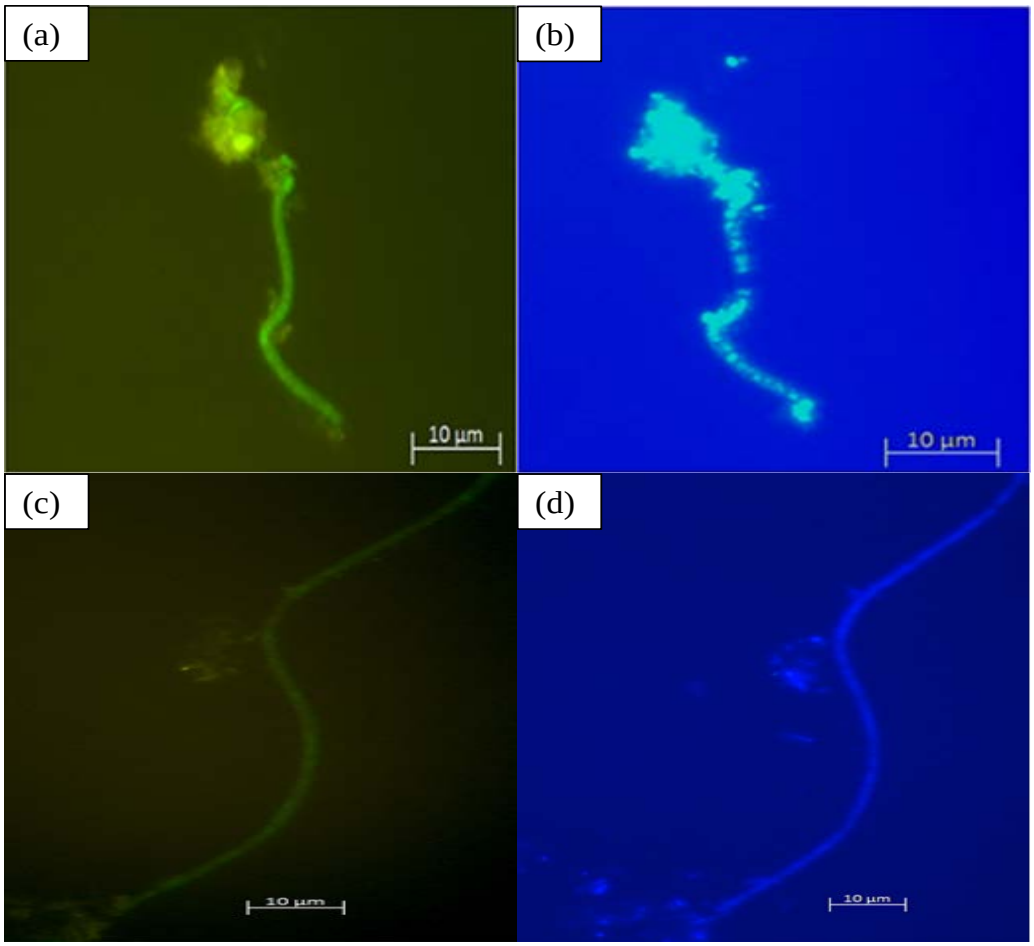


Figure 3.3. FISH images on dominant filamentous bacteria at 10 μm: (a) Phylum-targeted CFX-mix probes for *Chloroflexi* detection of Type 0092, (b) DNA staining with DAPI of Type 0092, (c) MPA-mix probe for detection of *M. parvicella*, and (d) DNA staining with DAPI of *M. parvicella*.

3.3.4. Relative Importance of Each Filament Using Artificial Neural Network Simulation

3.3.4.1.ANN Performance

The basic structure of ANN, consists of 3 layers, *i.e.* input data which considers the incoming signals, some of which can be manipulated, for instance other dominant filamentous types, to predict the outcome SVI (Figure 3.4a). The inputs can be assessed to examine the most critical parameter. The intermediate hidden layer is where the data is processed. The hidden layer controls the weights of the network which are trained and continuously changes until the total error of the training set is below acceptable error.

The mean squared error (MSE) between the experimental data of SVI and the model output (predicted) was used to evaluate the performances of training, validation, and testing (Figure 3.4b):

During the training stage, the ANN parameters (*i.e.*, weights and biases) were altered to obtain a model output congruent to the target data. The training step is stopped based on either of the following criteria, whichever occurs first (Fawzy *et al.*, 2018): (1) the magnitude of the performance gradient reaches a minimum threshold of 10^{-5}, or (2) the number of validation checks becomes 6. As shown in Figure 3.4b, the initial MSE of the training dataset was 542, which tended to decrease with a successive number of trials until reaching the maximum validation checks of 6. Hence, the progress of the training process was terminated according to the maximum validation failures, achieving 6 consecutive epochs (trials) (Appendix 2; Figure 3). This pattern revealed that the learning procedure successfully improved the ANN performance during training.

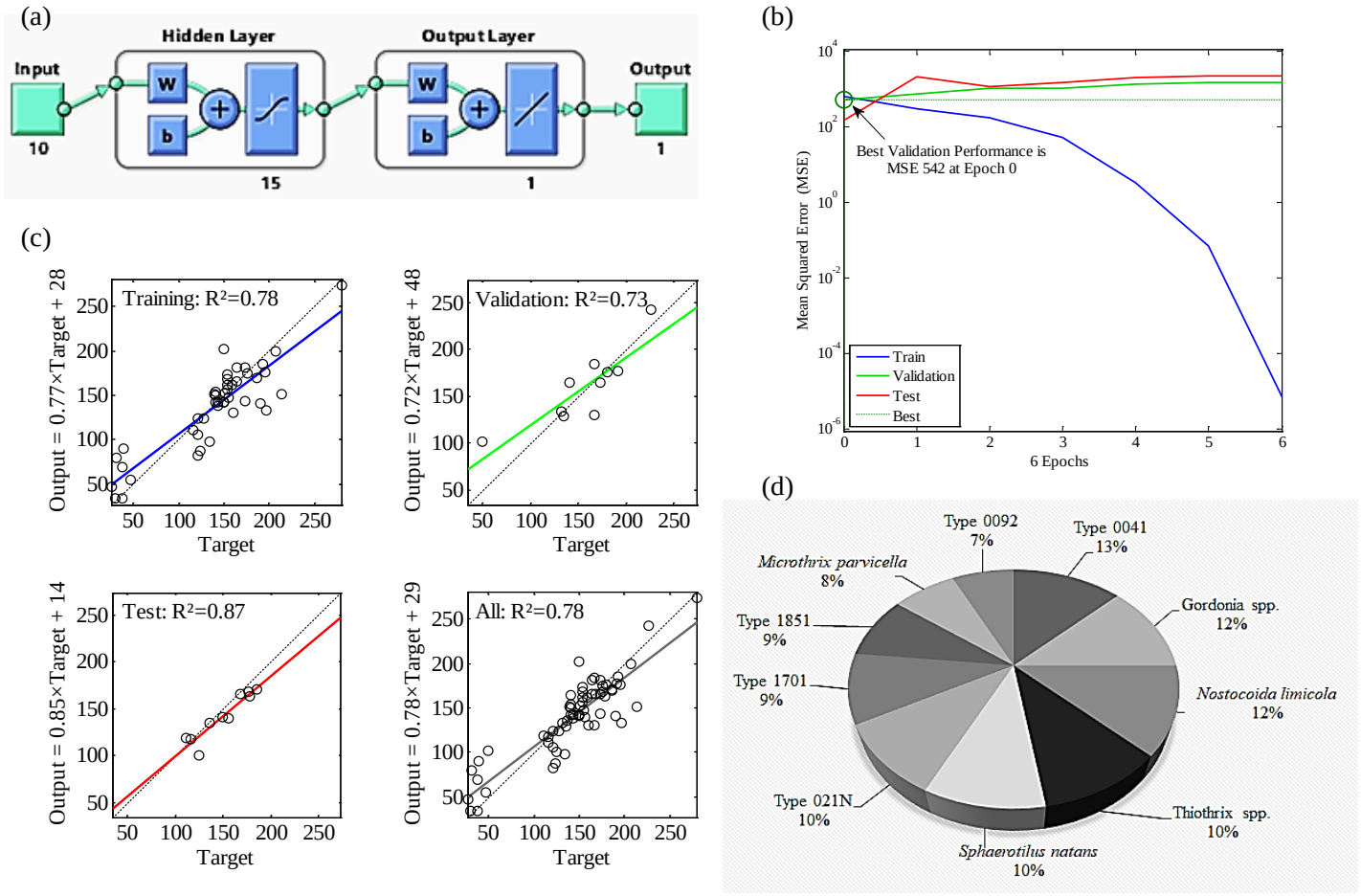


Figure 3.4. ANN simulation for prediction of SVI (an output) using abundance of filamentous bacteria (Ten inputs) Type 0041, *Gordonia* spp., *N. limicola*, *Thiothrix* spp., *Sphaerotilus natans*, Type 021N, Type 1701, Type 1851, *M. parvicella*, and Type 0092: (a) ANN model configuration 10 – 15 – 1, (b) ANN performance, (c) ANN prediction accuracy, and (d) ANN relative importance of each filament on SVI.

The performance function of the validation dataset was also examined along with the training phase to eliminate any inaccuracies from the network. The validation step was employed to evaluate the generalisation ability of the model as a predictive tool within the range of the data records used for training (Bagheri *et al.*, 2015). The validation process stopped earlier at an epoch number 0, *i.e.* once the MSE began to increase (Figure 3.4 b), signifying that the network initiated to overfit the data. At the point of overfitting, the model generates errors and/or random noises while importing new input data. Hence, this early stopping criterion is used to avoid overfitting problems and improve generalisation capacity.

Finally, the testing phase was undertaken using new independent records that were not considered for either training or validation. The error (in terms of MSE) of the test set enlarged after the initial iteration, indicating that the validation and testing curves were comparable. This observation implied that the proposed ANN model had a satisfactory robustness and generalisation capabilities (Nasr, 2018).

Based on the above findings, the ANN model displayed the highest performance at epoch 0 with a MSE (the performance function) of 542 (Figure 3.4b). At this point, the adjustable weights and biases were estimated, which were used to determine the relative importance of each input (See section 3.3.4.3. Sensitivity analysis).

3.3.4.2.Regression Plot

Figure 3.4c shows a correlation graph of the training, validation and test curves, where the solid line represents the model. This figure also indicate the coefficient of determination (R^2 values) between the measured SVI and the predicted outputs for training, validation, and test. In each graph, the dashed and solid lines represent the actual and perfect fits, respectively (Figure 3.4c). The overall R^2 value of the entire dataset was 0.78, revealing that the designed network could describe 78.0% of the variability of SVI episodes. As a result, good prediction power, for accuracy and generalisation performance, thus indicating that ANN can handle well the non-linear relationships between SVI and filamentous bacteria. In this context, the ANN model could be employed to offer acceptable predictive accuracy for the SVI data due to filamentous bacteria in WWTPs.

3.3.4.3.Sensitivity Analysis

Garson (1991) reported that the connection weights estimated from the ANN model at the optimum performance could describe the relative predictive importance of each independent input. Hence, the partitioning step of the input-to-hidden and hidden-to-output weights was applied using equation 3.4 to arrange the input attributes by order of importance.

$$I_j = \frac{\sum_{m=1}^{m=N_h}\left(\left(\frac{\left|W_{jm}^{ih}\right|}{\sum_{k=1}^{N_i}\left|W_{km}^{ih}\right|}\right)\times\left|W_{mn}^{ho}\right|\right)}{\sum_{k=1}^{k=N_i}\left\{\sum_{m=1}^{m=N_h}\left(\frac{\left|W_{km}^{ih}\right|}{\sum_{k=1}^{N_i}\left|W_{km}^{ih}\right|}\right)\times\left|W_{mn}^{ho}\right|\right\}}$$

Equation 3-4

where, I_j is the relative importance of jth input factor, N_i is the inputs number (N_i = 10), N_h is the number of hidden neurons (N_h = 15), and W is the connection weight, in which the subscripts "*k*", "*m*" and "*n*" refer to input, hidden and output neurons, respectively, and the superscripts "*i*", "*h*" and "*o*" denote input, hidden and output layers, respectively.

The relative importance of every input regarding its contribution to the model output was also confirmed using the "Weights method", as reported by Gar Alalm and Nasr (2018). Figure 3.4d shows that Type 0041 had the highest impact (12.91%) on SVI, followed by *Gordonia* spp. (12.09%). This observation was followed by *N. limicola* (11.98%) and *Thiothrix* spp. (10.47%). The importance values of *Sphaerotilus natans*, Type 021N, and Type 1701 were almost similar. Moreover, all the percentages for the contribution of filaments to SVI were relevant, and thus, no input could be excluded during multivariate statistical procedures.

The ANN algorithm described the relative magnitude of importance of filamentous bacteria as a significant predictor variable based on the FI scale: 0 -6, in its connection with outcome SVI variable by dissecting the model weights.

3.3.5. Spatial Variation of Filaments Using Principal Component Analysis

Figure 3.5 shows a detailed investigation of the spatial distribution of filamentous bacteria with the investigated WWTPs. As per comparison of both graphs, filamentous bacteria present, for instance identified as *Gordonia* spp., Type 1701, *N. limicola* and *M. parvicella* were common in WWTPs I, II, III, VI, while Type 0092 and 0041 were common in IV and V.

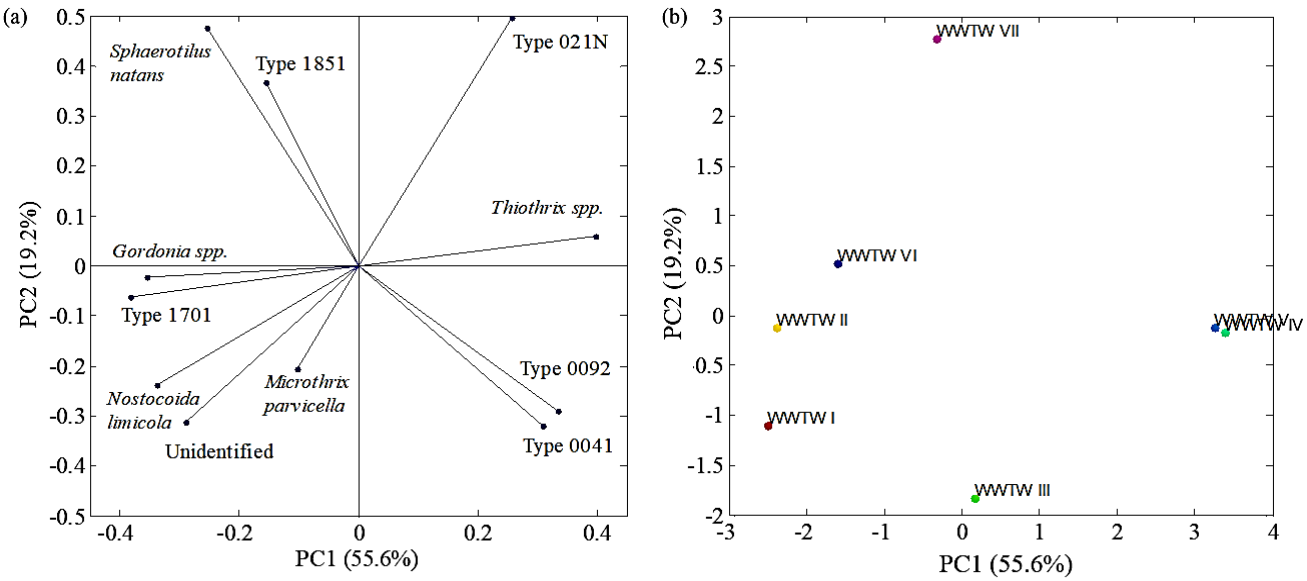


Figure 3.5. PCA results for distribution of filamentous bacteria in studied WWTPs: (a) Loading plot, and (b) Cluster analysis.

3.3.5.1.Dominance of Nostocoida limicola

As shown in Figure 3.4d, *N. limicola* had the third highest impact on SVI. The filamentous bacterial morphotype *N. limicola* appeared on both -PC1 and -PC2 directions (Figure 3.5a). Hence, *N. limicola* filaments showed a high relative abundance (FI: >3) in WWTPs I, II, III, and VI (Figure 3.5b), and they possibly increased due to the high SRT, low F/M conditions (<0.14 kg COD/ kg MLSS d^{-1}) (Table 3.1) and low nutrient conditions. Similar to previous WWTPs, the *N. limicola* morphotype caused incidents of bulking and foaming in activated sludge systems, worldwide (Liu *et al.*, 2000, Petrovski *et al.*, 2012). The coiled filament

N. limicola contains discoid cells, splitting into three morphotypes with different cell dimensions, *viz*., I (small cells), II (medium cells), and III (large cells) (McKenzie *et al.*, 2006, Seviour *et al.*, 2006). *N. limicola* is one of the few filaments reported to grow in anaerobic conditions that can utilise different organic substrates such as carbon, proteins and lipids. Nowak and Brown (1990), reported that *N. limicola* are prevalent in conditions of low sludge ages and low organic loading and their growth can be controlled by means of anoxic selectors or reduced sludge age.

3.3.5.2. Dominance of Gordonia spp.

Figure 3.4d indicated that *Gordonia* spp. were the second contributor to SVI, with a relative importance of 12.09%. These bulking-related species tend to reduce the specific gravity of the flocs and increase the SVI (Insel *et al.*, 2014). *Gordonia* spp. contain the presence of mycolic acids on their cell walls, hence these microbes often attach to gas bubbles and flocs, which floats to the surface of biological reactor tanks (Kougias *et al.*, 2014). As shown in the loading plot of PCA (Figure 3.5a), *Gordonia* spp. located on the extreme direction of -PC1; hence, they were detected largely in WWTPs I and II (Figure 3.5b). This result could be linked to the longer sludge ages (>25 days) compared to other treatment plants (Table 3.1), as previously reported by Richard *et al.* (2003). In a similar study, Dunkel *et al.* (2017) found that Nocardiaform-like organisms (*Gordonia* spp.) were the dominant bulking and foaming bacteria in industrial activated sludge plants. The “common” (FI: 3) *Gordonia* spp. in WWTPs I and II could also be attributed to slowly biodegradable compounds such as lipids (Marrengane *et al.*, 2011). In addition, *Gordonia amarae* were found abundant in membrane bioreactor (MBR) systems operated for nutrient removal (Insel *et al.*, 2014). *Gordonia* spp. could also physiologically thrive under anaerobic, anoxic, and aerobic environments at low DO levels (Insel, et al., 2014). However, Zhang *et al.* (2019), found that the filamentous bacteria *Gordonia* spp. were less dominant in oxidation ditch bulking sludge. In another study, Liu *et al.* (2015) revealed that phages could be applied to control foaming and bulking events caused by *Gordonia* spp., leading to improved sludge settling capabilities.

3.3.5.3. Dominance of Type 0041

The bacterium Type 0041 was the first contributor to SVI with a relative importance of 12.91% (Figure 3.4d). In addition, this bacterium appeared in all WWTPs with various dominance levels, suggesting its ability to sustain under different environmental conditions. This finding complied with a previous study by (Seviour *et al.*, 1994), showing that the filament morphotype Type 0041 could be defined as "all zone growers". Similarly, Thomsen *et al.* (2002) suggested that Type 0041 was associated with the bulking and foaming events in nutrient-removing WWTPs. Moreover, the filamentous bacteria Type 0041 proliferate due to favourable conditions of WWTP III, IV and V (Figure 3.5b). These microbes play a vital role in the formation of flocs, since they are integrated within the floc matrix. However, low F/M and long sludge ages were possible contributing factors to their proliferation in the selected systems (Richard *et al.*, 2003, Nielsen *et al.*, 2009a).

3.3.5.4. Dominance of Type 1701

As shown in the PCA findings (Figure 3.5a), Eikelboom Type 1701 distributed on the -PC1 axis. Hence, this filament was dominant in WWTPs I, II, and VI with possible conditions of low DO by surface aerators and/or low F/M. The filament Type 1701 is Neisser negative and Gram negative (Thomsen *et al.*, 2002), which has been detected with attached growth in activated sludge systems (see Figure 3.3c). Type 1701 is a bulking filament, similar to Type 0041, these microbes assist in flocculation, with heavy attached growth. The proliferation of this type can affect the sludge settleability.

3.3.5.5. Dominance of Thiothrix spp.

Thiothrix spp. was detected in all WWTPs; however, dominant in WWTPs IV and V, which could be due to the diffused aeration systems that supplied efficient DO and increased sludge age (20 days) (Figure 3.5b). Moreover, according to the ANN results, *Thiothrix* spp. was recorded as the fourth contributor to SVI with a relative importance of 10.47% (Figure 3.4d). These microbes have been associated with high sulphide concentrations, high organic loading rates and nutrient deficient conditions in full-scale WWTPs (Nielsen *et al.*, 2009a, Henriet *et*

al., 2017). Rosette-like structures were commonly observed with no attached growth (Appendix 1). *Thiothrix* spp. have been linked to high SVI > 200 mL g^{-1} in activated sludge WWTPs treating domestic and industrial wastewater sources (Williams and Unz, 1985, Nielsen *et al.*, 2009a, Wang *et al.*, 2015). In another work, Henriet *et al.* (2017) reported that the growth of *Thiothrix* spp. was highly associated with DO levels of 1.4 – 4.0 mg L^{-1}. The application of some strategies such as (a) mixing tanks (anoxic or aerobic stages), (b) contact units with no aeration, or (c) controlled aerobic selectors could limit the excessive growth of *Thiothrix* spp. (Martins *et al.*, 2004).

3.3.5.6. Dominance of Sphaerotilus natans

Sphaerotilus natans have Neisser negative, Gram negative, and unbranched characteristics. In this study, *S. natans* were related to a subjective FI scoring <2 in WWTPs, suggesting "few abundance" levels. This result could be because the WWTPs were operated at low F/M ratios of 0.11 – 0.18 kg COD/kg MLSS d^{-1}; however, an F/M level of about 1 kg COD/kgMLSS d^{-1} would favour the growth of *S. natans* (Figueroa *et al.*, 2015).

3.3.5.7. Dominance of Type 021N

Type 021N existed in WWTPs IV, V, and VII, which could be because this filamentous bacterium tends to grow under a nutrient deficit condition (*i.e.*, Low TN) (Han *et al.*, 2019). Type 021N was most excessive in WWTP VII system, as also seen in Figure 3.5. WWTP VII is a modified UCT process, operating as a simultaneous nitrification-denitrification system, which also operate at a very low DO and long sludge age to stabilise nitrifiers. Type 021N has the ability to consume NO_3 as an electron acceptor (Martins *et al.*, 2004). Moreover, the dominance of Type 021N in these plants could be linked to the direct addition of septic sludge to the preliminary treated wastewater. Guo *et al.* (2012) found in a complete mixed reactor treating domestic wastewater, that cetyltrimethyl ammonium bromide (CTAB) could be used as a biocidal agent to inactivate the chlorine-resistant Type 021N bacteria.

3.3.5.8. Dominance of Type 1851

Type 1851 bacterium is recognised as a Neisser negative and Gram positive strain (Thomsen *et al.*, 2002). In this study, Type 1851 appeared as "few" filamentous in most WWTPs during high SVI data. Similarly, Thomsen *et al*. (2002) reported that the Eikelboom morphotype Type 1851 existed in only 4 – 5% of WWTP samples. However, Type 1851 was "very common" at WWTP VI (FI: >4). The proliferation of filamentous bacterium Eikelboom type 1851 could be largely associated with low DO levels (*i.e.*, below 1.1 mg L^{-1}) (Nittami *et al.*, 2019).

3.3.5.9. Dominance of Type 0092

Type 0092 (Gram-negative) exhibited "abundant" and "excessive" scores in most WWTPs (see images of Type 0092 in Figures 3.2a and 3.3a). The excessive growth of Type 0092 has been linked to severe issues for solids separation in biological nutrient removal WWTPs (Madoni *et al.*, 2000, Liu *et al.*, 2015). Moreover, the Eikelboom morphotype Type 0092 has been defined as "all zone grower", due to its existence in anaerobic, anoxic, and aerobic zones (Speirs *et al.*, 2009, Insel *et al.*, 2014). Insel *et al.* (2014) suggested that limited aeration capacity (DO 0.2 – 0.3 mg L^{-1}) could allow the prolific growth of Type 0092 in aerobic units. However, the "some" and "common" abundance of Type 0092 WWTPs I and II could be because these plants were treating COD using activated sludge configurations.

3.3.5.10. Dominance of Candidatus Microthrix parvicella

In this study, *M. parvicella* appeared in most WWTP samples, however excessive in WWTP III during high SVI events, especially during the winter season (Figure 3.5a, b). *M. parvicella* is a Gram-positive and unsheathed filamentous bacterium (Figure 3.2b), which has been detected in several WWTPs in South Africa (Kumari *et al.*, 2009, Deepnarain *et al.*, 2019). The growth of *M. parvicella* could be due to the low F/M (< 0.2 kg COD/kg MLSS d^{-1}) in nutrient removal WWTPs treating domestic wastewater (Eikelboom, 2000).

3.3.6. Future Considerations for Mitigating Filamentous Bulking Sludge

In this study, an ANN model was employed for the prediction of SVI, which could be used as a reliable indicator to control the filamentous bulking problems. Additional input variables, e.g. operating conditions (temperature, sludge age, and hydraulic retention time), sludge morphological properties, and wastewater characteristics (BOD, N and P), would be imported into the ANN model to improve its predictive usefulness. Further, the input variables of the modified ANN model would be adjusted to control the sludge bulking problem occurrence by limiting the SVI output to be lower than 150 mL g^{-1}. This objective can be efficiently attained by the ANN approach without having a deep understanding of the complete mechanisms and the relationships among the input attributes (Nasr, 2018).

Furthermore, in this study, different species of problematic filamentous bacteria were assessed in seven BNR processes treating primarily domestic wastewater. For this purpose, the PCA approach was employed to give comprehensive information on the causes of filamentous growth in WWTPs. The clustering statistics depicted that floc-formers attempted to outcompete filamentous bacteria of the WWTPs IV and V, which could be associated with the use of efficient aeration system (*i.e.*, diffusion) in the 3-stage Phoredox process. In a similar study, Abusam *et al.* (2017) applied the PCA and clustering techniques to identify environmental conditions associated with the abundance of filaments.

Interestingly, the application of multivariate statistical analysis along with FISH images in this study would provide a scientific basis to control the proliferation of filamentous bacteria in other WWTPs situated in similar climates to South Africa. Although the FISH method was able to visualise the filamentous bacteria, it is considered an expensive and time-consuming technique that employs targeted probes and requires a fluorescent microscope (Bokulich *et al.*, 2012). Other molecular techniques, such as polymerase chain reaction (PCR), PCR followed by denaturing gel electrophoresis (DGGE), qPCR, Illumine-Miseq sequencing, and next generation sequencing (NGS), could be used for obtaining accurate and complete characterisation of the microbial community. These techniques would be employed in future studies to improve the predictive ability of the mathematical models, leading to a better understanding and control of sludge bulking at full-scale WWTPs.

In a comparable study, Bagheri *et al.* (2015) applied ANN to predict SVI using 7 operating parameters and they found that artificial intelligence techniques could be used to optimise the weights and thresholds of the model. Similarly, Boztoprak *et al.* (2015) revealed that ANN had a high prediction accuracy of a SVI output using 33 input parameters obtained from automated image acquisition. Han *et al.* (2016) assessed inputs of MLSS, COD, TN, pH, and DO that were used to predict SVI values efficiently. Han *et al.* (2018a), used a fuzzy neural network model for the prediction of sludge bulking, which would be a valuable option for the cases of poor sensor sensitivity and disturbance environment. Other modelling techniques such as a logistic model (Deepnarain *et al*., 2015), and a regression Tree model (Deepnarain *et al.*, 2019), have also been employed to predict the proliferation of filamentous microorganisms at full-scale WWTPs.

Based on these studies, the advantage of the ANN model is the determination of the most influential filament using the "Weights method" (see Figure 3.4. d), which could be used to improve the knowledge of sludge characteristic for further control of filamentous bulking. Moreover, the proposed models would provide an efficient tool to predict, simulate, and control sludge bulking in WWTPs having approximately similar influent wastewater properties and operational conditions. Moreover, the presented models were developed based on monitoring the sludge bulking events in seven WWTPs, giving a broad range of applicability in other wastewater treatment systems.

3.4. CONCLUSIONS

This study analysed and assessed the microbial community associated with sludge bulking in seven WWTPs operated under different operational conditions using PCA and ANN. The main conclusions are:

- Staining and microscopic methods, along with FISH images, revealed that filaments formed bridging among the flocs and enlarged the surface area, leading to a loose and poor compacted structure.
- Subjective scores based on microbial investigation were an effective way to determine the relative quantities and dominance of filamentous bacteria.

- ANN with a structure of 10 – 15 – 1 was able to predict the SVI data using ten attributes of filaments bacteria. The ANN results depicted that Eikelboom Type 0041 contributed to the highest impact on SVI, followed by *Gordonia* spp., *N. limicola*, and *Thiothrix* spp.
- Based on multivariate analysis, the distribution of filamentous species among seven WWTPs was adequately reported, signifying that the implementation of efficient aeration system (*i.e.*, diffusion) in the 3-stage Phoredox process enhanced the settling properties of flocs.
- Principal component analysis and clustering showed that most WWTPs experienced inadequate sludge settling properties due to the prevalent filamentous bacteria identified; however, the application of efficient aeration system (*i.e.*, fine bubble, diffused aeration) in the three-stage Phoredox process improved the settling characteristics of flocs. Operational conditions that caused filament overgrowth in each WWTP were also determined. PCA was employed to correlate the main features of each WWTP with the abundance of filamentous bacteria.
- The outputs of this study are of growing importance to potentially produce feasible solutions and ease of control for the operation of WWTPs experiencing frequent bulking problems.

CHAPTER FOUR: DECISION TREE FOR IDENTIFICATION AND PREDICTION OF FILAMENTOUS BULKING AT FULL-SCALE ACTIVATED SLUDGE WASTEWATER TREATMENT PLANTS

4.1. INTRODUCTION

The effectiveness of sludge settleability in SSTs is considered one of the important features that identify the functionality of activated sludge systems (Lou and Zhao, 2012). However, the overgrowth of undesirable microorganisms resulting from inadequate system operation and/or unbalanced carbon and nutrient ratios can reduce the solid precipitation efficiency (Mielczarek *et al.*, 2012). Poor sludge settleability can block the piping system, reduce the oxygen transfer efficiency in mixed liquor, deteriorate the performance of the biological processes, and cause a carryover of solid particles with the final effluent (Martins *et al.*, 2004). Hence, proper statistical models should be employed to assess and evaluate the settling performance in activated sludge systems.

The development of well-settling sludge (dense and strong flocs) is influenced by a balanced growth between floc-forming and filamentous bacteria (Wagner *et al.*, 2015). Sludge bulking occurs when a large number of filamentous microorganisms generates inter-floc bridging, reduces biomass density, and diffuses bio-flocs having poor precipitation properties (Dunkel *et al.*, 2016). The condition of poor sludge settleability is characterised by a high sludge volume index (SVI), which can exceed 120 mL g^{-1} (Bagheri *et al.*, 2015).

Recently, various efforts have been exerted to determine the appropriate statistical and computational modelling techniques that can evaluate and predict sludge bulking scenarios (Comas *et al.*, 2008, Lou and Zhao, 2012, Bagheri *et al.*, 2015, Deepnarain *et al.*, 2015, Han *et al.*, 2018a). Multivariate statistical tools such as principal component analysis (PCA) can be employed to describe the correlations between wastewater characteristics and SVI in activated sludge systems. For example, Lou and Zhao (2012) employed PCA to predict SVI using several inputs such as temperature, chemical oxygen demand (COD), ammonium-nitrogen (NH_4^+-N), total nitrogen (TN), and total phosphorus (TP). PCA can extract useful information from large datasets and describe complex relationships between input attributes and target variables (Nasr and Zahran, 2016). PCA is used to transform and reduce a complex multi-dimensional system into orthogonal variables, known as principal components (PCs) (Jolliffe, 2002). PCs having

eigenvalues greater than one can be used to represent a high percentage of total variation in the investigated variables (Lou and Zhao, 2012).

Decision Tree technique can also be applied to predict sludge bulking events, and to select the most relevant input variables. Decision Tree is considered a data mining method that employs a flowchart-like tree structure for partitioning a set of data into discrete sub-groups or nodes (Yu *et al.*, 2010). The flowchart structure of Decision Tree consists of root, intermediate, and leaf nodes (Kim *et al.*, 2014). The root node, which is the most important predictor with respect to the target variable, partitions all records into two or multiple subgroups (Breiman *et al.*, 1984). The intermediate node expresses a condition or a binary split test for each independent variable; *i.e.*, the top edge is linked to a parent node, whereas the bottom edge is associated with child or leaf nodes (Kim *et al.*, 2001). The leaf node represents the result of a combination of decisions, being categorical outcomes for classification trees and numerical results for regression trees. Decision Trees also contain branches (*i.e.*, chance outcomes), which are emanated from both root and internal nodes (Loh and Shih, 1997). Decision Tree provides adequate categorisation, generalisation, and prediction of a given population using a set of if-then rules. Decision Tree can deal with undefined observations in the dataset by classifying the records into independent branches, and hence, the entire dataset is used in the investigation (Yeo and Grant, 2018). Decision Tree can also display both linear and non-linear relationships of large datasets in simple-to-interpret visualisations. These benefits tend to improve the predictive accuracy of both continuous and categorical target variables.

Previous studies have investigated the ecophysiological traits and characteristics of filamentous microorganisms using laboratory strategies (Wagner *et al.*, 2015, Dunkel *et al.*, 2016). However, laboratory scale might be unreliable to cultivate various filamentous species (Mielczarek *et al.*, 2012). Hence, in this study, a full-scale WWTP with an activated sludge system situated in Gauteng, South Africa was monitored for two years. Several variables that described the seasonal variation in influent wastewater characteristics, effluent quality, operational conditions, and filamentous bacteria in the aeration tanks were recorded. PCA and Decision Tree techniques were employed to evaluate and model the filamentous sludge bulking and to determine the factors that controlled the proliferation of filaments. The model input factors contained pH, temperature, dissolved oxygen (DO), sludge retention time (SRT), food-to-microorganisms (F/M) ratio, soluble COD (sCOD), total COD (tCOD), NH_4^+-N, total Kjeldahl nitrogen (TKN), phosphorus as phosphate (PO_4^3-P), TP, and total suspended solids (TSS). These parameters were selected because they significantly varied over time ($p < 0.05$).

Fluorescent *in situ* hybridisation (FISH) analysis was used to determine the abundance levels of various filamentous species such as *M. parvicella*, *Thiothrix* I & II, and Eikelboom Types 0041, 0092, and 021N. The environmental factors that affected the proliferation of these species were also determined.

4.2. MATERIALS AND METHODS

4.2.1. Wastewater Treatment Facilities

A WWTP located in Gauteng, South Africa was monitored monthly for 2 years (*i.e.*, 2015 and 2016). The WWTP was employed for the treatment of domestic wastewater in addition to 5% industrial effluents. The feed wastewater was screened, de-gritted, and collected in balancing tanks. The biological treatment system was operated as a 3-stage Phoredox process containing anaerobic (29%), anoxic (29%), and aerobic (42%) zones. The mixed liquor discharged from the bioreactors was subjected to quiescent conditions in the SSTs. The WWTP was operated using return activated sludge (RAS) from the underflow of the SSTs to the head of the biological basins, and internal recycle (IR) of mixed liquor. The waste activated sludge (WAS) was removed from the treatment units and pumped to the dissolved air flotation (DAF) units for thickening.

The system with diffused aeration was selected as the system had online sensors to measure DO and wastewater characteristics, this data was easily accessible from the plant operators. The selected WWTP was relatively stable compared to other systems surveyed in this study.

4.2.2. Sampling

Through the monitoring process, wastewater samples were collected at appropriate points (inlet and outlet) to determine the plant performance. The effluent quality was compared with the allowable discharge standards reported by the Department of Water Affairs of South Africa (DWA, 1998). The samples were preserved at a temperature below 4°C in a portable icebox containing ice packs and transported to the laboratory within 2 h. Composite samples of the mixed liquor were collected from the beginning, middle, and end of aeration tanks using an aluminum dipper (1 L) connected to a wooden handle. About 1/3 of the sampling bottle was kept empty to ensure the survival of microorganisms under the aerobic conditions. To avoid the growth/decay of microbial cells after harvesting, samples were subjected to fixation, cell wall

permeabilisation, and hybridisation following the protocols of Nielsen *et al.* (2009b). The fixed samples were kept in the freezer (-20 °C) for further analyses.

4.2.3. Bacterial Analysis

4.2.3.1. Identification of Filamentous Bacteria

The morphological classification and microscopic examination of filamentous bacteria were determined using staining characteristics, as specified by Eikelboom (2000) and Jenkins *et al.* (2003). The filamentous abundance were subsequently, categorically ranked as follows: “none”, “few”, “some”, “common”, “very common”, “abundant”, and “excessive” (Jenkins *et al*., 2003). This FI scale is used to provide relevant information about the tendency of bulking in activated sludge systems.

4.2.3.2. Fluorescent in situ Hybridisation Analysis

The FISH analysis was employed on a monthly basis to confirm the identification of filamentous bacteria. The FISH procedures were performed based on the guidelines described by Nielsen *et al.* (2009). Table 3.2 lists the specific 16S rRNA probes used for targeting the filamentous organisms during hybridisation (Amann *et al.*, 1995, Daims *et al.*, 1999, Kanagawa *et al.*, 2000, Gich *et al.*, 2001, Speirs *et al.*, 2009).

4.2.4. Analytical Analysis

The influent flow rate, pH, temperature, and DO were recorded by the respective WWTPs personnel using a supervisory control and data acquisition (SCADA) system. Nitrogen and phosphorus were analyzed by Aquakem Gallery Photometric Auto-analyser (Thermo Scientific, Germany). The analyses were conducted according to the protocols outlined by the US Environmental Protection Agency (EPA) (USEPA, 1983). Other parameters such as COD and MLSS were measured according to the standard methods of the American Public Health Association (APHA, 1998).

4.2.5. Sludge Volume Index Calculation

The SVI was considered in this chapter as a typical indicator for the assessment of sludge settling performance, as also described in chapter 3. The SVI was calculated using equation 3.1 according to Jenkins *et al.* (2003)

4.2.6. Statistical Analysis

4.2.6.1. Principal Component Analysis

The PCA is designed to transform the original variables into new and orthogonal axes, called principal components (PCs) (Nasr and Zahran, 2016). The PCs represent the directions of the maximum variance, and thus, provide information on the most meaningful parameters.

The eigenvalues define the associated variance of the PCs, whereas the loadings are the weights of the original variables in each PC. The PCA involves the following steps (Jolliffe, 2002): (a) all parameters are z-scale standardised with zero mean and unit variance (Figure 4.1a), that is, to ensure that all the measurements have equal weight in the analysis, (b) the correlation matrix is calculated from equation 4.1, (c) the eigenvectors and eigenvalues are estimated, and the eigenvalues are sorted in a descending order, (d) the eigenvector with the highest eigenvalue represents the most dominant principal component of the data set, namely PC1, and (e) the second component (PC2) is computed under the constraint of being orthogonal to PC1 and having the second largest variance.

The PCA calculations were performed using the functions “pca” and “pcacov” in MATLAB software (version R2015a).

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y)}{(n-1)\sigma_x\sigma_y}$$

Equation 4-1

where, μ_x and μ_y are the sample means of X and Y, respectively, and σ_x and σ_y are the sample standard deviations of X and Y, respectively.

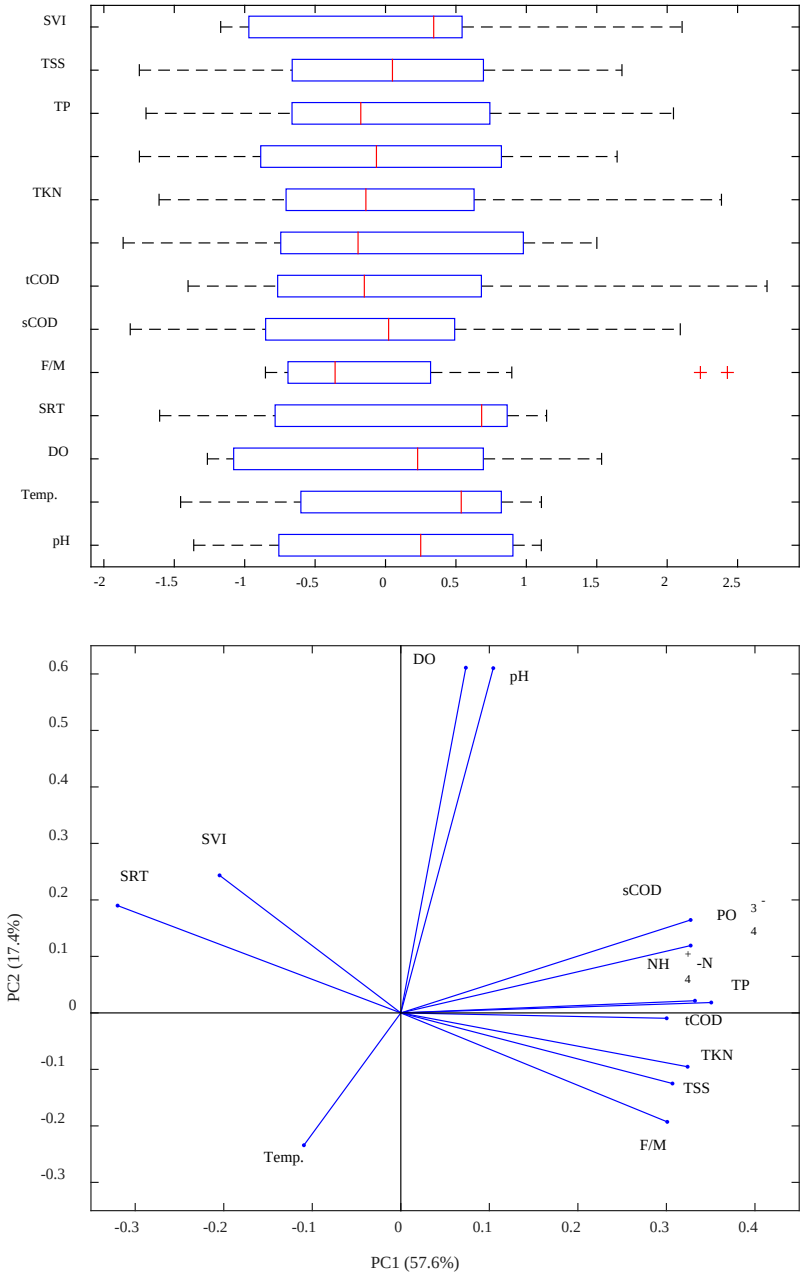


Figure 4.1. (a) Normalised variables in Box-and-Whisker plot, and (b) loading plot of PCA.

4.2.6.2. Decision Tree Method

Decision Trees classify instances or examples into branch-like segments by taking paths from the root node through internal nodes to leaf nodes. In this method, the root node implies the predictor that gives the best split of the target class values (Yu *et al.*, 2010). Further, splitting is

applied iteratively to the subgroups (internal nodes or child partitions) until leaf nodes are obtained. Each internal node contains splits and holds two or more child nodes (Loh and Shih, 1997). Pruning is used to avoid overfitting problems by removing unnecessary nodes and optimising the tree size. The pruning process can either be pre-pruning by preventing the generation of unimportant branches or post-pruning by removing branches from the established tree (Breiman *et al.*, 1984). The Decision Tree process is completed when (a) the class label of the leaf node has the same target class value, (b) every predictor has already been used to split a partition, and (c) no more records for a particular value of a predictor variable.

The learning procedure of Decision Trees is achieved via the training and validation steps. The training process is used to build and evaluate the Decision Tree model by minimising the difference between the measured and predicted outputs. The validation step is employed to determine the suitable tree size that avoids data overfitting and the loss of generalisation ability. These procedures are applied to estimate the accuracy of the Decision Tree by comparing predicted outputs with actual data. The optimal final model is achieved when the accuracy becomes acceptable, and thus, the Decision Tree can be employed for classification and prediction purposes using a new dataset (Yu *et al.*, 2010).

In this study, the Decision Tree model was used to (a) elucidate the relationships between influent wastewater characteristics and operational conditions (as inputs) and SVI (as an output), and (b) define the environmental factors that mainly influenced the dominance of filamentous species. The function “Fitctree” in MATLAB (version R2015a) was employed to estimate the Decision Trees based on the input variables (*i.e.*, predictors). The function employs the Classification and Regression Trees (CART) algorithm for building the Decision Tree by the binary split at each node (Breiman *et al.*, 1984). The algorithm was used due to its flexibility and applicability to various sorts of data. The input attributes that obtain a high degree of purity of the child nodes are selected using the Gini index method for classification trees and mean squared error for regression trees (Loh and Shih, 1997). This function determines the optimal sequence of pruned subtrees, namely pre-pruning.

4.3. RESULTS AND DISCUSSION

4.3.1. Wastewater Treatment Plant Performance

The investigated activated sludge system was designed for the treatment of a dry weather flowrate that varied from $20.8 \times 10^3 – 35.2 \times 10^3$ m^3 d^{-1} and operated at a SRT of 9.8 – 21.6 days. Table 4.1 lists the measured ranges of wastewater characteristics and the corresponding removal efficiencies obtained during the sampling period. These variables were selected from the database as they covered the organic and nutrient properties of the influent wastewater. The high variation in wastewater characteristics could be because the samples were collected throughout the year (Dunkel *et al.*, 2016). In addition, the study period of two years was satisfactory as it represented all the probable seasonal variations in the wastewater parameters.

Increasing the SRT generally increases energy consumption of WWTPs and reduces waste biomass production. In this study, the low SRTs was due to the system being re-engineered to save cost on aeration. Zvimba and Musvoto (2020) found that this WWTP was under loaded, with issues arising from mechanical failures of dissolved air floatation units and poor wasted sludge removal from the works, thus forcing longer SRTs. To optimize energy efficiencies of the works, (Zvimba and Musvoto, 2020) shut down an entire lane and larger amounts of sludge was wasted. This was evident in this study, since during September, the MLSS concentrations were much lower than usual (<2000 mg L^{-1}), which possibly caused upsets in the system including that of the sludge settling. Low SRTs is generally not recommended due to dispersed growth of microorganisms and pin-point flocs thus causing inadequate clarification and a turbid final effluent quality.

Table 4.1 Operational conditions, influent wastewater characteristics, and corresponding removal efficiencies for WWTP located in Gauteng in South Africa.

	Minimum	Maximum	Average	Std.
pH	7.3	7.6	7.5	0.1
Temp. (°C)	14.0	23.0	19.1	3.5
DO (mg L^{-1})	2.5	4.0	3.1	0.5
SRT (d)	9.8	21.6	16.7	4.3
F/M (kg COD/kg MLSS d^{-1})	0.1	0.4	0.2	0.1
sCOD (mg L^{-1})	27.3	105.6	63.7	20.0
tCOD (mg L^{-1})	71.3	301.8	149.9	56.1
NH_4^+-N (mg L^{-1})	6.0	18.0	12.7	3.6
TKN (mg L^{-1})	9.4	30.7	18.0	5.3
PO_4^{3-}-P (mg L^{-1})	0.4	1.2	0.8	0.3
TP (mg L^{-1})	0.8	3.2	1.9	0.6
TSS (mg L^{-1})	20.5	112.1	67.2	26.7
tCOD removal (%)	60.7	89.2	76.7	8.5
NH_4^+-N removal (%)	93.4	99.2	97.5	1.5
TKN removal (%)	48.3	95.0	85.7	11.5
PO_4^{3-}-P removal (%)	6.2	94.1	75.9	20.5
TP removal (%)	14.6	90.8	73.8	18.8
TSS removal (%)	69.8	97.5	88.6	7.2

The influent wastewater contained tCOD of 149.9 ± 56.1 mg L^{-1}, resulting in an organic load of (1.7 – 10.0) × 10^3 kg/d. Under this condition, the treatment facilities achieved a tCOD removal efficiency of 60.7 – 89.2%. The effluent tCOD fluctuated between 24.5 and 39.5 mg L^{-1}, which complied with the National Water Act (Act No 36 of 1998) of 75 mg COD L^{-1} as general standards and 30 mg COD L^{-1} as specific standards (National Water Act, 1999). The sCOD/tCOD ratios ranged between 0.31 and 0.65, suggesting that about 85% of data had a readily biodegradable substrate = 54%. The nitrogen loads recorded 364 ± 158 kg-NH_4^+/d and 517 ± 233 kg-TKN/d and the removal efficiencies were 93.4 – 99.2% and 48.3 – 95.0%, respectively. The effluent NH_4^+-N was acceptable according to the allowable regulations of 3.0 and 2.0 mg L^{-1} as general and special standards, respectively (National Water Act, 1999).

NH_4-N is the main source of total-nitrogen in the influent wastewater. A large number of independent variables, investigated by Luo and Zhao (2012), included variables such as influent NH_4-N, BOD, COD, TN, TP, TSS in the PCA and the ANN model. These variables had provided more complete data inputs to explain the bulking mechanisms. In this study, influent NH_4-N and TKN were computed into the model, which are also factors known to influence filamentous bacteria. Whereas, common factors affecting nitrification are high COD loads, toxic shock, pH change, low DO, SRT and low temperatures during winter (Seruga *et al.*, 2019). BNR systems are significantly impacted by cold temperatures. In this study, the TKN removal was reduced in the winter months, with reduced SRTs which can also lead to less simultaneous nitrification and denitrification occurring in the structure of the aggregated sludge flocs (Shahzad *et al.*, 2015).

The biological units were also subjected to PO_4^{3-}-P and TP loads of 23 ± 11 and 55 ± 26 kg/d, respectively. The removal efficiencies were 75.9 ± 20.5% and 73.8 ± 18.8% for PO_4^{3-}-P and TP, respectively. The effluent general limits of phosphorus that are applicable for discharge into a water resource in South Africa are 10 mg L^{-1} as general standards and 1 mg L^{-1} as special standards (DWA, 1998). The influent TSS concentration was 67.2 ± 26.7 mg L^{-1}, providing a TSS load of (1.9 ± 1.0) × 10^3 kg/d. The effluent TSS ranged between 2.8 and 163.6 mg L^{-1}, in which several observations exceeded the allowable limit of 10 mg TSS L^{-1} (National Water Act, 1999). About 38% of the total number of SVI observations exceeded the threshold level for sludge bulking (>120 mL g^{-1}). Moreover, the seasonal variation of SVIs ($p < 0.05$) implied that the bulking episodes occurred during the low-temperature seasons (Figure 4.2a). This finding might be due to the presence of filamentous bacteria such as *M. parvicella*, and

Eikelboom Type 0041 and 0092. In a similar study, Madoni *et al.* (2000) found that the main filamentous species involved in bulking problems were *M. parvicella* (53.2%), Type 0041 (11.3%), and Type 021N (9.7%).

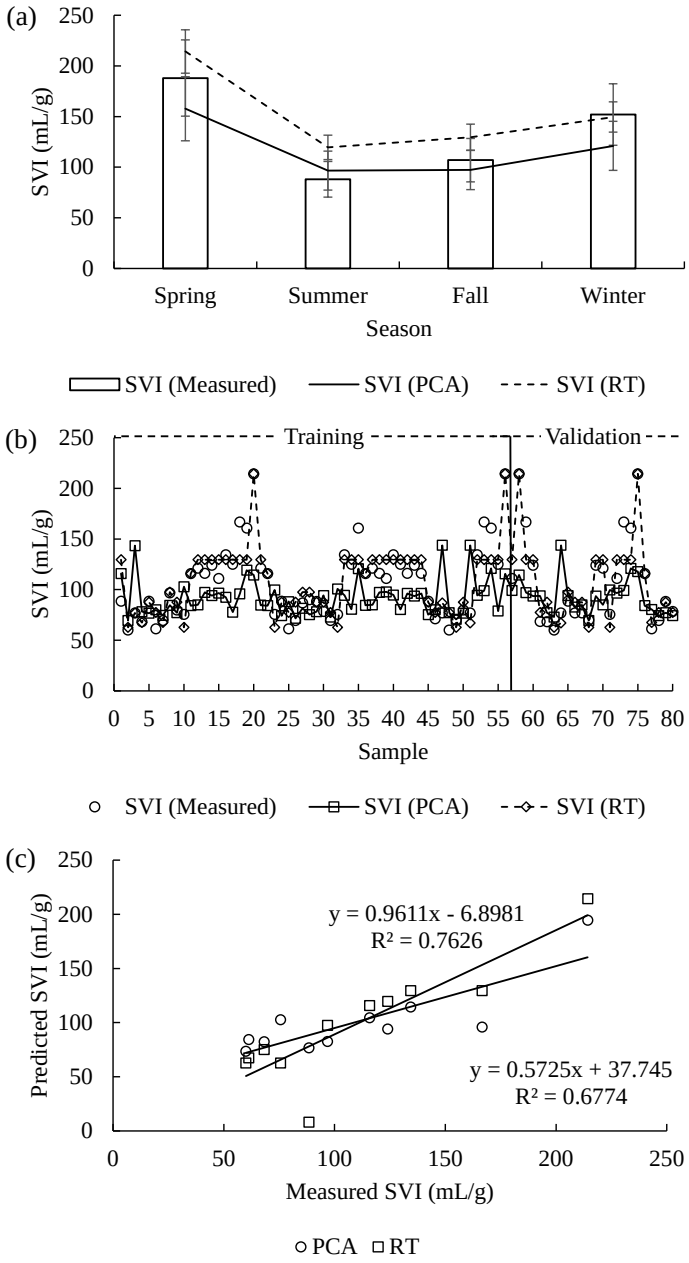


Figure 4.2. (a) Seasonal variation of SVI fitted to PCA and regression tree (RT) models, (b) training and validation procures for the prediction of SVI via PCA (R2 = 0.71 for training and 0.68 for validation) and RT (R2 = 0.84 for training and 0.81 for validation).

The aforementioned results revealed efficient capabilities for the removal of organic and nutrient pollutants, in which the treated wastewater complied with the department of water and sanitation (DWS) regulations during the study period. However, the decrease in settling properties of solids could be associated with the sludge bulking events and the overgrowth of filamentous organisms (Seviour *et al.*, 1994). Under this condition, the tCOD removal efficiency reduced to 63.4%, indicating the detrimental effect of bulking on the effluent quality. The growth of filamentous microorganisms could retard the biodegradation of organic matter due to slower kinetics involved (Ovez *et al.*, 2006).

4.3.2. Principal Component Analysis Application for Prediction of Sludge Volume Index

PCA was applied to group the wastewater characteristics and operational conditions that influenced sludge bulking according to common features. Since the input attributes had different units, variances, and wide ranges of measurements, they were subjected to a standardisation step. Figure 4.1a shows the normalised variables displayed by Box-and-Whisker plot. In figure 4.1a, each variable extended from the 25th to 75th percentiles and had a mean of zero and unit standard deviation. This procedure prevents certain factors to dominate the analysis because of their large numerical values (Nasr and Zahran, 2016).

The first principal component (PC1) had an eigenvalue > 1.0 and captured the highest variance of the dataset with a value of 57.6%. The second and third principal components (PC2 and PC3) explained only 17.4% and 11.1% of the total variance and provided low loadings for organic and nutrient concentrations. Accordingly, PC1 was used in this study to describe the correlations among the investigated parameters and to identify the most important information in the dataset.

Figure 4.1b shows the PCA results identified by PC1 (57.6%) and PC2 (17.4%). PC1 had high positive loadings for F/M (0.30), sCOD (0.33), tCOD (0.30), NH_4^+-N (0.35), TKN (0.32), PO_4^3-P (0.33), TP (0.33), and TSS (0.31). This result showed that PC1 was positively influenced by the studied wastewater characteristics. Moreover, both SVI and SRT had negative loadings on the PC1 coordinate, suggesting that an increase in SVI could be attributed to the increment in SRT. In a similar study, Comas *et al.* (2008) reported that the risk of foaming in activated sludge systems was obtained at high SRT. Flores-Alsina *et al.* (2009) also demonstrated that the high SRT values increased the occasion of sludge bulking. Moreover,

SVI was adversely correlated with NH_4^+-N and sCOD, indicating that the risk of bulking could be due to the deficiency of organic components and nutrient species. Similarly, Comas *et al*. (2008) reported that filamentous bulking could occur due to substrate limiting conditions or nutrient deficiency. However, the optimum nutrient concentration required for bacterial growth varies between each activated sludge system (Wang *et al.*, 2016b).

The second principal component (PC2) depicted that SVI positioned on the opposite side of temperature, suggesting that SVI increased at low-temperature environments. This observation was also explored in the seasonal variation of SVI, as shown in Figure 4.2a. This result could be due to the decrease in lipid solubility at low temperatures, causing specific hydrophobic microorganisms such as *M. parvicella* to outcompete other lipid-utilising microorganisms for oleic acid (Wang *et al.*, 2016b). Moreover, SVI was observed in the opposite direction of F/M, implying that the risk of filamentous bulking could be linked to low F/M ratios. Similarly, studies have indicated that low F/M ratios (< 0.1 kg COD/kg MLSS d^{-1}) encouraged the occurrence of filamentous bulking (Kumari *et al.*, 2009, Deepnarain *et al.*, 2015, Rossetti *et al.*, 2017).

Based on the PCA results, SVI could be predicted using equation 4.2 with R^2 values of 0.71 and 0.68 for the training (n = 56) and validation (n = 24) datasets, respectively (Figure 4.2b). In this model, SRT, COD, and NH_4^+-N were selected as predictors to avoid overfitting. Due to the complexity of the full-scale activated sludge system, the prediction accuracy of PCA with R^2 over 0.7 is acceptable (Wang *et al.*, 2017).

$$SVI = 91.05 + 2.56\,SRT - 0.19\,sCOD - 2.53\,NH_{4^+} - N$$ **Equation 4-2**

In a similar study, Lou and Zhao (2012) employed a PCA technique to predict SVI using several inputs including carbon and nutrient species. Their study depicted that SVI was directly associated with COD, biological oxygen demand (BOD), TP, and TN, whereas it had negative correlations with temperature, and pH.

Sensitivity analysis is always a fundamental element of decision making and in decision trees, which most often focus on the probabilities (Kaminski *et al.*, 2018). The decision makers are often uncertain about the exact parameters of such decision trees and the sensitivity analysis assist in assessing the risk of strategies formed, by identifying how dependent the output such

as SVI is on the selected input variables (i.e. influent wastewater characteristics and plant operational parameters). Sensitivity analysis was done, to learn how the output of the decision making process changes when the input is varied, especially when multiple factors are applied in the Decision Tree. Thus the sensitivity analysis strategy can determine the robustness of the assessment by examining the extent to which the output SVI are affected by changes in the variables or assumptions with aim of identifying the most dependent variable (Kaminski *et al.*, 2018).

4.3.3. Modelling of Sludge Volume Index Using Regression Tree

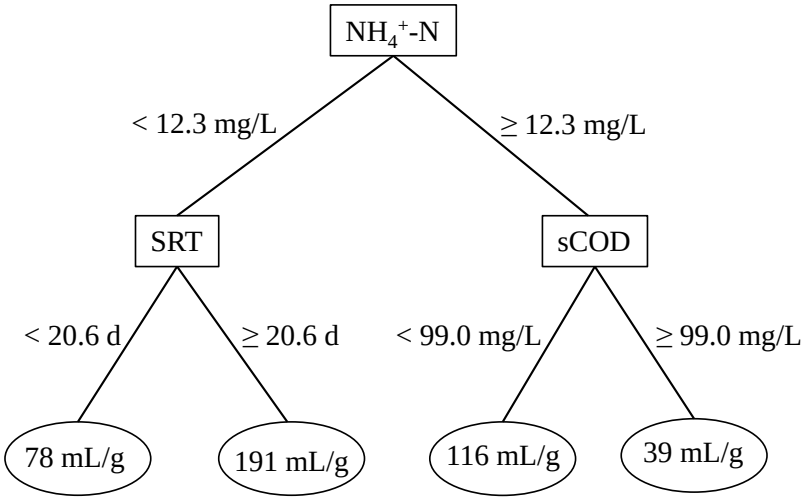
The Decision Tree model was developed using wastewater characteristics and operational conditions (as inputs) and SVI (as an output). It was found that the R^2 value of the training observations was 0.84 (n = 56; Figure 4.2b), and hence, the regression tree fitted the training data well. In addition, the R^2 of the validation dataset was 0.81 (n = 24; Figure 4.2b), suggesting that the developed model would be able to predict SVI values based on new records of the independent variables.

Figure 4.2c shows the generalisation procedure of the developed models using test data that were not employed during the training and validation steps. The R^2 values of PCA and regression tree models were 0.68 and 0.76, respectively. In addition, the average absolute error between the real records and the predicted outputs was 23.3 mL g^{-1} for PCA compared with 14.2 mL g^{-1} for the regression tree model. Hence, regression tree could verify a higher predictive ability of SVI than PCA across a wide range of inputs and applications.

The regression tree in Figure 4.3 demonstrated that the decision node was NH_4^+-N, which provided the main impact on SVI. Moreover, the intermediate nodes were SRT and sCOD. This observation complied with the PCA results, indicating that NH_4^+-N, SRT, and sCOD were amongst the investigated factors that highly influenced SVI. The decrease in NH_4^+-N below 12.3 mg L^{-1} resulted in an increase in SVI to 130 mL g^{-1} (using the regression tree at a pruning level 3 of 4). This observation corresponded to a ratio of C (18)/N (<1), which was lower than the optimum value of C/N = 20/1 for a balanced bacterial growth system (Wukasch, 1993). Hence, insufficient nitrogen concentrations in the influent wastewater could favour the overgrowth of filamentous species and cause possible risks of bulking. Similar to the PCA results, an increase in SVI was noticed at either an increase in SRT (> 20.6 d) or a decrease in sCOD (< 99.0 mg L^{-1}); *i.e.*, see equation 4.2. However, in comparison to the PCA model, the

Decision Tree method provided higher performances for both training and validation procedures (Figure 4.2b). In addition, the Decision Tree model adequately interpreted and displayed the nonlinear correlations between wastewater characteristic and SVI in a simple graphical interface. Bagheri *et al.* (2015) applied an artificial neural network (ANN) model with 76 data points to predict SVI (*i.e.*, output) using several inputs such as DO, TSS, COD, and TN. Their study also found that SVI was considerably influenced by the influent TN (as a single input attribute), and TN and COD (as a group of two inputs). The usefulness of the Decision Tree model derived in this application for the prediction of bulking events, with respect to other existing models, is listed in Table 2.3.

Figure 4.3. Regression tree for prediction of SVI using wastewater characteristics and operational conditions.



4.3.4. Effect of Wastewater Characteristics on Filamentous Bacteria Using Classification Tree

In this investigation, six filamentous bacteria were identified and confirmed using FISH analysis. These microorganisms were *M. parvicella*, *Thiothrix* I and II, and Types 0041, 0092, and 021N. Table 4.2 lists the dominance of filamentous species throughout the study period. The Decision Tree method was employed to classify the proliferation of filamentous microorganisms based on the influent wastewater characteristics and operational conditions.

Each identified filamentous spp. were monitored and categorically ranked using a filamentous index (FI) scale i.e. from none to excessive, as described by Jenkins *et al*., 2004. In this study, sudden changes (eg. variation in SRT or F/M ratio) in the system had led to larger variations of the FI in each of the individual filamentous bacterial spp. The filamentous bacteria were categorised as dominant and in high levels when ranked higher than level 4 on the scale (i.e. 5-20 filaments per floc). The Decision Tree was a suitable method that was used in this study, as it also allowed to handle both continuous and categorical data simultaneously. However, Decision Trees are prone to errors of such categorical data, owing to their differences in perceptions. Though Decision Trees are also able to perform such classification without the use of extensive computation and can provide a clear indication of variables that had most impact for prediction and classification. Knowing the *in situ* growth rates of the dominant filamentous bacteria in bulking activated sludge could improve the application of the model and the use of selective control remedial measures.

Table 4.2 Percentages of filamentous bacteria proliferating in all samples

	M. parvicella	***Thiothrix* I**	***Thiothrix* II**	**Type 0041**	**Type 0092**	**Type 021N**
None	11	0	0	0	0	0
Few	67	50	22	0	0	33
Some	22	28	44	66	17	45
Common	0	22	34	28	28	22
Very common	0	0	0	6	6	0
Abundant	0	0	0	0	33	0
Excessive	0	0	0	0	16	0

The training dataset (n = 56) was used to build the classification tree model, whereas the validation dataset (n = 24) was employed to determine the suitable tree size required to attain the optimal classifications. About 86% of the training observations were correctly classified, and hence, the classification tree fitted the training data well. In addition, 83% of the validating set records were appropriately classified, and thus, the decisions were acceptable. The classification tree models were able to handle categorical attributes and to assign filamentous bacteria to specific classes, such as, “few”, “common”, and “abundant”. The determination of filamentous classification can be illustrated as follows (Figure 4.4):

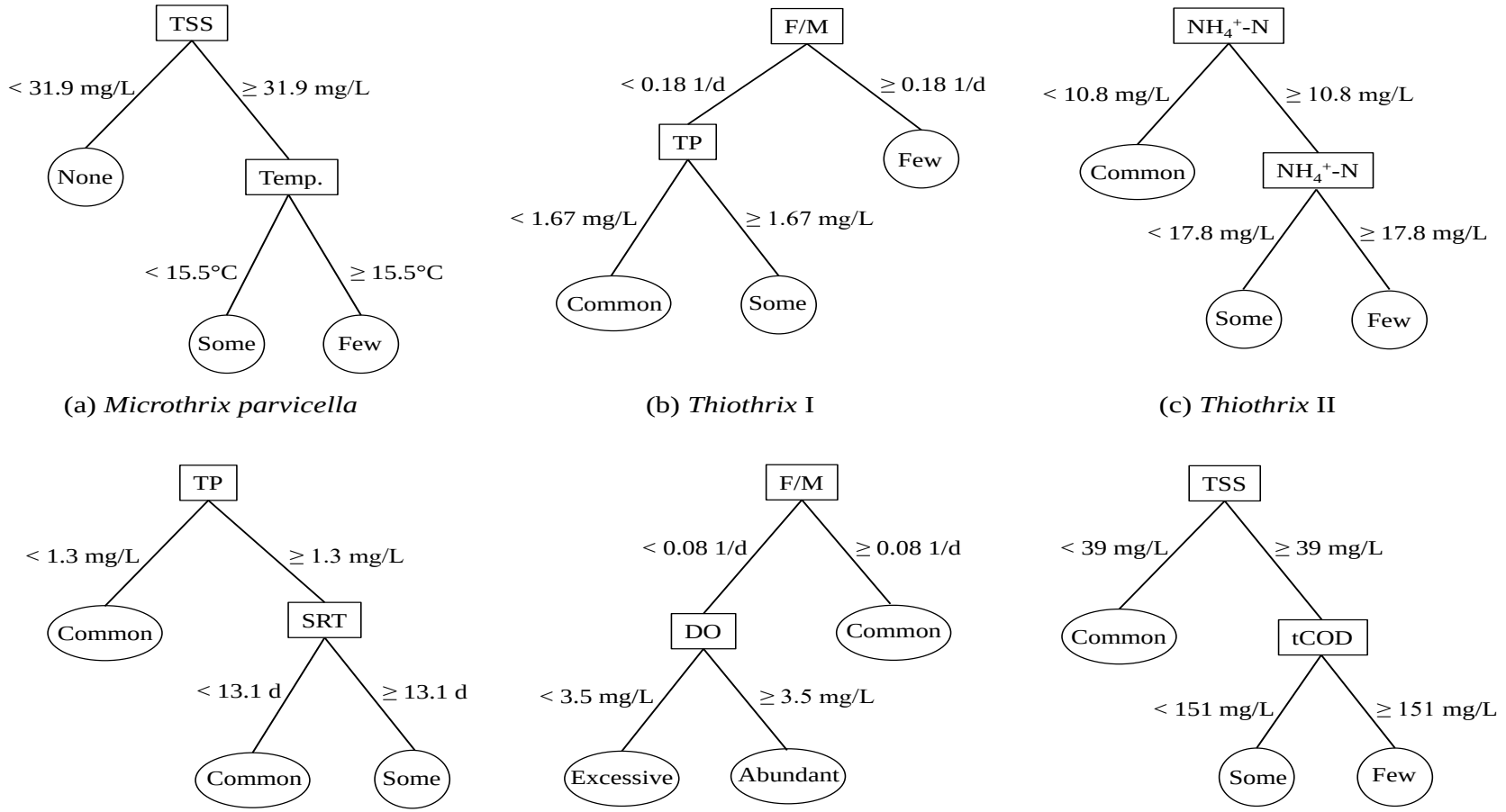


Figure 4.4. Decision Tree for classifications of (a) M. parvicella, (b) Thiothrix I, (c) Thiothrix II, (d) Type 0041, (e) Type 0092, and (f) Type 021N.

The use of machine learning tools such as Decision Trees and ANN and its application to sludge bulking that can enable evaluation of both filamentous bacteria and process control strategies, can yield additional benefits that would not be realized only through previous ASM. These models can determine solutions during dynamic upsets to form improved solutions in real time.

4.3.4.1.Effect of Wastewater Characteristics and Environmental Conditions on Microthrix parvicella

M. parvicella is a Gram-positive bacterium, which has been frequently detected in several WWTPs in South Africa (Lacko *et al.*, 1999, Kumari *et al.*, 2009, Welz *et al.*, 2014). It belongs to the class *Actinobacteria* and has a length of 200 – 400 μm with a diameter of 0.6 – 0.8 μm (Blackall *et al.*, 1996). Domestic nutrient removal plants with low F/M (< 0.2 kg COD/kg MLSS d^{-1}) are favorable for the growth of *M. parvicella* (Eikelboom, 2000). In this study, *M. parvicella* appeared in the collected samples as "none" (11%), "few" (67%), and "some" (22%). The limited proliferation of *M. parvicella* could be linked to the deficiency of long chain fatty acid and lipids content in wastewater influents (Ovez *et al.*, 2006). *M. parvicella* causes serious bulking, and its hydrophobic cell surface can adsorb fats and fatty acids and stimulate the formation of stable scum (Hug *et al.*, 2006). Research indicated that *M. parvicella* was the most common filamentous microorganism involved in bulking and foaming events at low temperatures (Madoni *et al.*, 2000, Kumari *et al.*, 2009, Rossetti *et al.*, 2017).

Figure 4.4a shows the Decision Tree for the classification of *M. parvicella*, in which the generated tree included three terminal nodes, *i.e.*, "none", "few", and "some". *M. parvicella* was mainly affected by TSS (the root node), and its dominance became "None" at lower TSS values (TSS < 31.9 mg L^{-1}). Under this condition, the filament attained a minor influence on the settling velocity of the sludge.

The influent TSS was simulated into the model and was generated in order to understand the effects it has on the SVI and filamentous bacteria. Ebrahimi *et al.* (2016), included influent TSS, MLVSS and dolamite into an ANN model to predict SVI. These parameters were selected due to their relations, high variations, and the availability of continuous experimental results. Direct influence was found with influent TSS concentrations, and results from Ebrahimi et al. (2016), verified that the variables included using the ANN approach increased the reliability of sludge operation and retrofitted its performance in terms of sludge bulking.

Foaming induced by *M. parvicella* can be controlled by reducing the suspended solids concentration in the aeration tank (Madoni *et al.*, 2000). Temperature, *i.e.*, the decision node of sub-tree, had the second largest impact on the dominance of *M. parvicella*. A decrease in temperature below 15.5°C (corresponding to winter) observed the appearance of *M. parvicella* as "some". Under this environment, the settling properties of the sludge were possibly deteriorated. However, the occurrence of *M. parvicella* as "few" was noticed in summer with temperature > 15.5°C. Hence, the development of *M. parvicella* increased at the low-temperature conditions (e.g., winter), equivalent to high SVI events during winter and spring (Figure 4.2a).

Similarly, Eikelboom (2000) revealed that *M. parvicella* grew largely at the end of the winter season, resulting in the occurrence of serious operational problems. Fan *et al.* (2017b) demonstrated that cold climate in winter and spring favored the dominant growth of *M. parvicella*, suggesting that at temperatures over 20°C, some enzymes are produced and destroy the cell wall of *M. parvicella*. In another study, dos Santos *et al.* (2015) found that *M. parvicella* was directly correlated with pH in the aeration tank and it contributed to the biomass washout along with the decline in the treatment performance.

4.3.4.2. Effect of Wastewater Characteristics and Operational Parameters on Thiothrix I & II

Thiothrix spp. strains are Gram-negative filamentous bacteria that form sulphur granules (Eikelboom, 2000). This type is characterised by filament length < 500 μm, round ended rod and cylindrical cell shape, and colony size of 1 – 2 mm (Williams and Unz, 1985). Henriet *et al.* (2017) reported that filamentous bulking caused by *Thiothrix* spp. in dairy WWTPs caused sludge washout and a decrease in biological performance. Moreover, their study demonstrated that the growth of *Thiothrix* spp. is mainly influenced by the presence of volatile fatty acids (VFA) and a relatively low DO concentrations (1.4 to 4.0 mg L^{-1}).

In this study, the classifications of *Thiothrix* I during the investigation were "few" (50%), "some" (28%), and "common" (22%). As can be seen in Figure 4.4b, the root of the classification tree was F/M, whereas the internal node was TP. Moreover, the end nodes were "few", "some", and "common", and the tree height was 2. It was demonstrated that *Thiothrix* I was "Few" at F/M ≥ 0.18 kg COD/kg MLSS d^{-1}. However, when the F/M ratio decreased

below 0.18 kg COD/kg MLSS d^{-1}, the existence of *Thiothrix* I was either "common" or "some", based on the TP concentrations.

Thiothrix II appeared during the study period as "few" (22%), "some" (44%), and "common" (34%). As can be inferred in Figure 4.4c, the influent NH_4^+-N (6.0 – 18.0 mg L^{-1}) was the tree root, and it also appeared as an internal node, indicating that the main factor affecting *Thiothrix* II was NH_4^+-N. The tree terminals were "few", "some", and "common". It was demonstrated that *Thiothrix* II was "common" at NH_4^+-N < 10.8 mg L^{-1}, equivalent to an insufficient nitrogen content of COD (20) to N (<1). The unbalanced ratio between readily biodegradable COD and nutrients favors the overgrowth of *Thiothrix* spp. (Henriet *et al.*, 2017). Similarly, Eikelboom (2000) reported that the occurrence of *Thiothrix* spp. in activated sludge could be due to the lack of either nitrogen or phosphorus species. Further, *Thiothrix* II became either "some" or "few" at NH_4^+-N of 10.8 – 17.8 mg L^{-1} or ≥ 17.8 mg L^{-1}, respectively.

Thiothrix spp. have been observed to cause high SVI > 200 mL g^{-1} in domestic activated sludge processes (Williams and Unz, 1985). Sludge bulking resulting from the excessive growth of *Thiothrix* spp. can be controlled by the implementation of aerobic selectors, small mixing units (anoxic or aerobic), or contact zone (without aeration) (Martins *et al.*, 2004). They can also be limited by using anaerobic and anoxic stages, such as bio-P and denitrifying activated sludge systems (Wanner and Grau, 1989).

4.3.4.1. Effect of Wastewater Characteristics and Operational Parameters on Eikelboom Types 0041, 0092, and 021N

The bacterium Type 0041 is a Gram-variable microorganism having a filament length of 200 – 500 μm and cylindrical, cuboidal, or oval cell shape (Williams and Unz, 1985). Type 0041 is a filament morphotype responsible for the bulking and foaming episodes in biological nutrient removal activated sludge systems, *viz.*, anaerobic–anoxic–aerobic processes (Seviour *et al.*, 1994). The occurrence of bacteria Type 0041 in this study was "some" (66%), "common" (28%), and "very common" (6%). The dominance of Type 0041 could be due to its ability to adapt and thrive in a wide range of environmental conditions in activated sludge systems (Lacko *et al.*, 1999). Similarly, Blackbeard *et al.* (1988) reported that Type 0041 was the main filamentous microorganisms found in South Africa. TP that varied from 0.8 to 3.2 mg L^{-1} was the root node of the tree; *viz.*, the most explanatory variable (Figure 4.4d). The leaves of the classification tree were "some", and "common". Type 0041 became "common" with a decrease

in TP below 1.3 mg L^{-1}. This result corresponded to an unbalanced nutrient and/or substrate-limiting conditions with COD/N/P of 115/9/1. The nutrient requirements for balanced growth of microorganisms and deactivation of filamentous growth during aerobic treatment is C/N/P ratio of 100/5/1 (Wukasch, 1993). Furthermore, an increase in TP over 1.3 mg L^{-1} provided a suitable condition for either "some" at SRT ≥ 13.1 d or "common" at SRT < 13.1 d.

Type 0092 (Gram-negative) is a strict aerobic bacterium (Horan *et al.*, 1988, Speirs *et al.*, 2009), which has been detected in several nutrient removal activated sludge systems (Madoni *et al.*, 2000). In this study, Type 0092 was the prevalent filamentous bacterium observed during high SVI conditions (SVI > 120 mL g^{-1}). dos Santos *et al.* (2015) confirmed the strong relationship between Type 0092 and SVI based on the correlation between filamentous bacteria populations and environmental conditions. Type 0092 was detected throughout the investigated period as "some" (17%), "common" (28%), "very common" (6%), "abundant" (33%), and "excessive" (16%). The development of Type 0092 could be because this filament is able to grow under various aerobic, anoxic, and anaerobic conditions; *i.e.*, recognised as "All zone grower" (dos Santos *et al.*, 2015). The F/M ratio (0.05 – 0.45 kg COD/kg MLSS d^{-1}) was the most informative input variable, in which Type 0092 became "common" at F/M ≥ 0.08 kg COD/kg MLSS d^{-1} (Figure 4.4e). With a decrease in F/M below 0.08 kg COD/kg MLSS d^{-1}, Type 0092 became either "excessive" at DO < 3.5 mg L^{-1} or "abundant" at DO ≥ 3.5 mg L^{-1}, suggesting a poor sludge quality. Madoni *et al.* (2000) also reported that Type 0092 was dominant in plants with a low F/M ratio (< 0.1 kg COD/kg MLSS d^{-1}). These results revealed the major impacts of F/M and DO on the occurrence of Type 0092. Hence, F/M and DO should be optimised to limit the growth of the filament Type 0092.

Type 021N (Gram-negative) has been recognised as multicellular, sheathless, and curled bacterium that simulates blue-green algae (Aruga *et al.*, 2002). Type 021N is a filamentous bacterium able to utilise NO_3 as an electron acceptor (Martins *et al.*, 2004). The proliferation of Type 021N was classified as "few" (33%), "some" (45%), and "common" (22%). The influent TSS was the root node, whereas tCOD was the internal node of the classification tree. The terminal nodes of the classification tree were "few", "some", and "common", and the tree height was 2. Type 021N was "common" with the decrease in TSS below 39.0 mg L^{-1}; however, it recorded either "few" or "some" at TSS ≥ 39.0 mg L^{-1} (Figure 4.4f). The existence of Type 021N as "some" at COD < 151 mg L^{-1} could be related to the favorability of this filaments to dominate at low F/M activated sludge systems, as described by Plante (1990). Aerobic selectors and nutrient addition have been found to limit sludge bulking resulted from the growth of Type

021N (Martins *et al*., 2004). Similar to *Thiothrix* spp., the dominance of Type 021N is positively correlated with the soluble readily biodegradable substrates (Jenkins *et al.*, 2003).

4.3.5. Identified Filamentous Bacteria Type 0092

Figure 4.5 shows the microscopic observations of filamentous bacterium Type 0092. This bacterium was selected because it covered most of the classification categories. During non-bulking conditions (SVI < 120 mL g^{-1}), filamentous microorganisms were less common, and the sludge granules were dense and compact. This observation was noticed in Figure 4.5a, which showed medium-sized flocs with "some" filaments, as identified by the Gram staining method. Moreover, Neisser-stained biomass displaying Eikelboom Type 0092 (violet filaments) indicated "common" (Figure 4.5b) and "very common" (Figure 4.5c) classes. In addition, Figure 4.5d shows "abundant" Type 0092 captured during sludge bulking events. Under this condition, filamentous microorganisms reside inside the microbial flocs forming a sponge-like structure that can alter the hindered settling velocity (Wagner *et al.*, 2015). Neisser staining that indicated "excessive" Type 0092, obtained at SVI >120 mL g^{-1}, is shown in Figure 4.5e. This bacterium forms bridges among the floc structures and prevents them from aggregation and compaction (Eikelboom, 2000); hence, it reduces solid-liquid separation achievements. In addition, the extended filamentous network can capture gas bubbles, which carry the organisms to the surface and stabilize a foam layer. Figure 4.5f demonstrates a FISH image of Type 0092. This microorganism transferred from the sludge flocs into the bulk solution and reduced the removal efficiency of organics and solids.

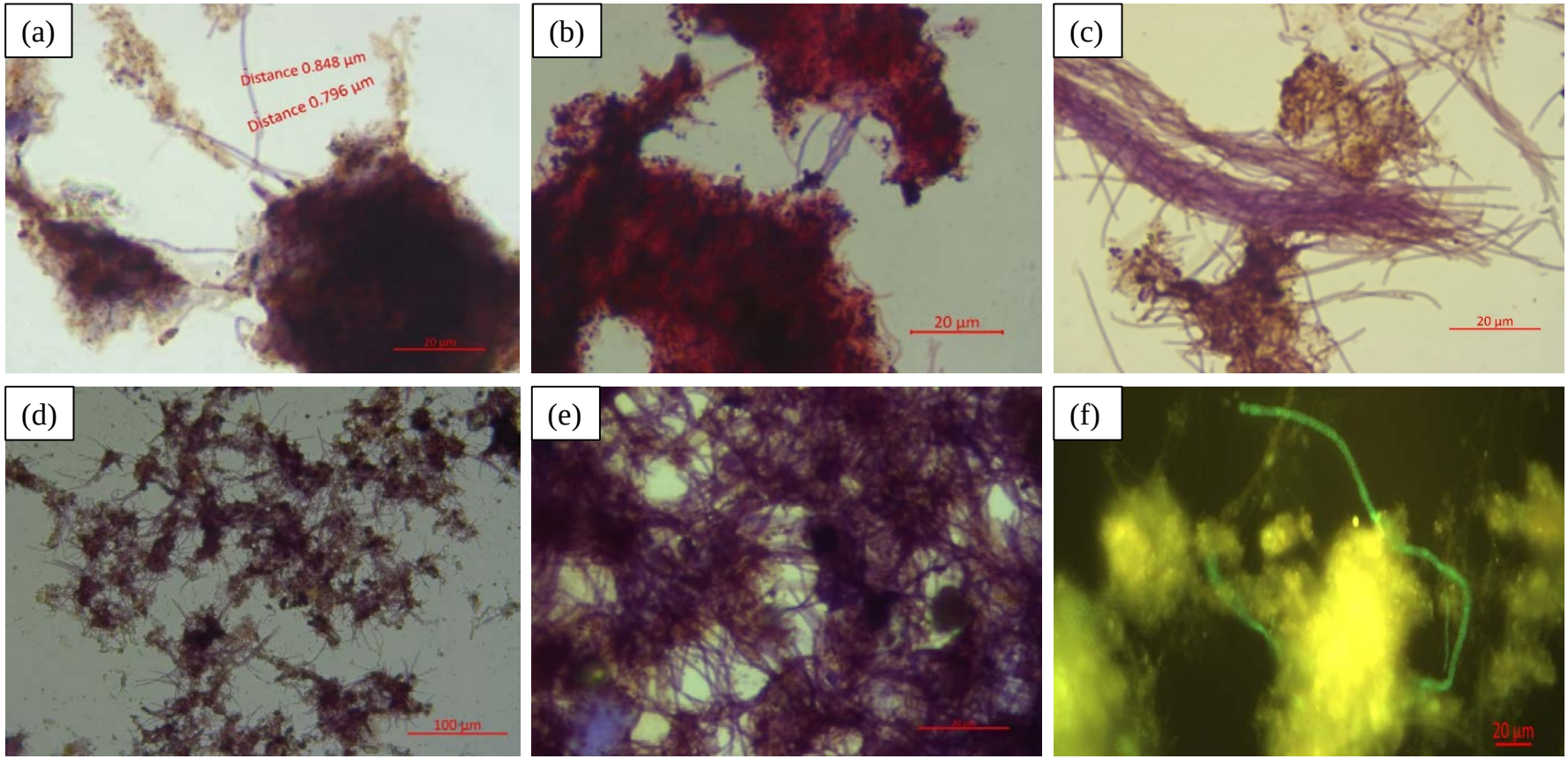


Figure 4.5. Microscopic images of filamentous bacterium Type 0092: (a) “some”, (b) “common”, (c) “very common”, (d) “abundant”, (e) “excessive”, and (f) FISH image.

4.4. CONCLUSIONS

This study aimed at modelling the sludge bulking events in a full-scale activated sludge system using PCA and Decision Tree methods. The relationships between wastewater characteristics, operational conditions, and SVI were interpreted. The environmental factors that influenced the proliferation of filamentous microorganisms were classified. The investigated WWTP was capable of producing treated wastewater that complied with the National Water Act (No 36 of 1998) of South Africa; however, some observations of TSS exceeded the permissible limit of 10 mg L^{-1}. The main conclusions are as follows:

- About 38% of the total number of SVI records were higher than the threshold value of sludge bulking.
- The results obtained from PCA and regression tree models indicated that the increase in SVI was attributed to the increment in SRT, deficiency of organic components and nutrient species, or a decrease in temperature.
- The detected filamentous bacteria were *M. parvicella*, *Thiothrix* I & II, and Eikelboom Types 0041, 0092, and 021N.
- The classification tree results revealed that *M. parvicella* could dominate with a decrease in temperature below 15.5°C, implying the bulking occurrence during the winter and spring seasons.
- Eikelboom Type 0092 was the dominant bacterium responsible for the high SVI events, due to low FM and low DO concentrations. These microbes were classified as “some” (17%), “common” (28%), “very common” (6%), “abundant” (33%), and “excessive” (16%).
- Type 021N was mainly influenced by TSS, whereas the F/M ratio provided a major impact on both *Thiothrix* I and Type 0092. *Thiothrix* II and Type 0041 were primarily impacted by the unbalanced nutrient condition in the feed wastewater.
- The developed models attained acceptable training and validating steps, and thus, they can be used to predict the sludge bulking episodes. However, future studies will be conducted to (a) investigate the applicability of the developed models for the prediction of sludge bulking in other WWTPs operated at different environmental conditions, and (b) develop mathematical models that can predict bulking events based on daily measurements.

CHAPTER FIVE: ASSESSING THE IMPACT OF SLUDGE BULKING ON RECEIVING ENVIRONMENT USING A QUANTITATIVE MICROBIAL RISK ASSESSMENT -BASED FRAMEWORK

5.1. INTRODUCTION

An essential function of WWTPs is to avoid environmental pollution through the reduction of microbiological pollutants and nutrients (Cai and Zhang, 2013). The effluents released from marginally treated WWTPs may contain a wide range of waterborne pathogens, including bacteria, viruses, and parasites (Dickin *et al.*, 2016). During sludge bulking conditions, it is expected that pathogens tend to escape to the effluent due to increased SS, as a result of inadequate settling (Martins *et al*., 2004). This would cause detrimental health risks (e.g., cholera, amoebiasis, diarrhea, typhoid fever, and gastroenteritis) for the reuse application of treated wastewater (Amha *et al.*, 2017, Amoah *et al.*, 2018c). According to the World Health Organisation (WHO) guidelines, the treated wastewater can be provided for consumers at an identified risk level of $\leq 10^{-4}$ infections per year, *viz*., one case or less per 10,000 people per year (WHO, 2006, WHO, 2011, WHO, 2016). Hence, more efforts are required to assess the microbial contamination associated with the application of WWTPs effluents for irrigation and other practices, concerning human health risk.

Wastewater treatment facilities contain a diversity of microbial pathogens including *Campylobacter jejuni*, *Cryptosporidium* spp., *E. coli* O157:H7, *Giardia intestinalis*, *Salmonella typhimurium*, *Shigella flexneri*, *Vibrio cholerae*, and *Mycobacterium* (Haas *et al.*, 2014, Ekwanzala *et al.*, 2018). Among these contaminants, *E. coli* especially *E. coli* O157:H7, has emerged as a pathogen of substantial public health concern (Barak *et al.*, 2005). This microbe can exist in the effluents of WWTPs and resist the treatment and disinfection processes (Ayaz *et al.*, 2014). Moreover, the risk of *Salmonella* infections can be used as a guide of hazard in edible crops irrigated with reclaimed wastewater (Amha *et al.*, 2015). Both *E. coli* O157:H7 and *Salmonella* are the leading causes of crop-related outbreaks, accounting for about 20 and 30% respectively. Both *E. coli* O157:H7 and *Salmonella* are considered the most serious severe food borne pathogens, found even at a low infective dose, in raw vegetables and fruits irrigated with contaminated wastewater (Ayaz *et al.*, 2014). The transmission is largely through eating of undercooked contaminated meat and consumption of raw milk and crops contaminated by water (Kiranmayi and Mallika, 2010).

Mycobacterium has been identified as one of the opportunistic environmental pathogens, tending to enrich among foaming bacteria in WWTPs (Amha *et al.*, 2017). Globally, many species of *Mycobacterium* can result in morbidity and mortality, and exhibit an elevated resistance against disinfectants and antibiotics (Falkinham, 2009). Accordingly, the elimination of these microbial contaminants from WWTPs effluents is a crucial objective to ensure that the treated water is safe for reuse applications.

Quantitative microbial risk assessment (QMRA) is a probabilistic model that integrates data on pathogen abundance, human exposure, and infection to evaluate the potential health impacts related to a polluted environment (Sampson *et al.*, 2017). The components of QMRA include hazard description, exposure estimation, dose-response valuation, and risk classification (Schijven *et al.*, 2019). The QMRA model can provide a clear comparison and assessment of various systems regarding health-based measures (Zhou *et al.*, 2017). This framework can also be used to identify and examine the accessibility of WWTPs effluent for various reuse scenarios (Amha *et al*., 2015). For instance, farmers in some developing countries having inaccessibility to freshwater resources may use treated wastewater for crop irrigation and cultivation (Van Vu *et al.*, 2018). As a result, these irrigated fruits and vegetables may contain unacceptable pathogenic levels (Dong *et al.*, 2017). Furthermore, communities and households living close to WWTPs have to protect themselves against a variety of waterborne pathogens and related diseases (Hamilton *et al.*, 2006, Sunger *et al.*, 2018). The United Nations set sustainable development goals (SDGs), which aim to substantially increase safe water reuse, and also reduce water pollution of receiving water bodies and to further improve the water quality by 2030 (Qadir *et al.*, 2020).

In this study, the potential health risks associated with the discharge or reuse of effluents from WWTPs experiencing sludge bulking were evaluated. For this objective, a QMRA model was applied using the relative abundance of three common waterborne pathogens, *i.e. E. coli* 0157: H7, *Salmonella*, and *Mycobacterium*. The potential health risks associated with low, moderate, and high sludge bulking conditions in WWTPs were incorporated into the model. The QMRA model target was to investigate the safety of treated effluent for (a) children, women, and men during recreational activities, (b) farmers during irrigation practices, and (c) consumers of edible plants (vegetables). The QMRA results were compared to the tolerable risks levels of 10^{-4}; and various disinfection treatment scenarios via chlorination, ultraviolet (UV), and ozonation were

included in the calculations to control the excess risk. To the best of our knowledge, this is the first study on applying the QMRA approach for the management of risks associated with pathogens in wastewater effluents in bulking WWTPs, *i.e.*, particularly for farming activities and hygiene improvements.

5.2. MATERIAL AND METHODS

5.2.1. Sample Collection, Pathogen Enrichment

Four full-scale biological WWTPs, situated in Durban, KwaZulu-Natal Province of South Africa, were selected for this study. The treatment facilities are composed of physical unit operations such as grit chambers, followed by biological reactors (aerators and secondary settling tanks) and chlorination. Wastewater samples (2 L) were collected over five consecutive months (April – August 2019) from the aeration tank and pre-chlorinated effluent. Composite samples of the mixed liquor were harvested, fixed, and transported to the laboratory within 2 hours. The Pre-chlorinated effluent samples were filtered and stored in PBS for DNA extractions. The SVI was estimated as described in chapter 3, section 3.2.3.

5.2.2. Microbial Analysis

5.2.2.1. DNA Extraction

The total genomic DNA was isolated and extracted from sludge samples using a phenol extraction method (Sekiguchi *et al.*, 1998, Enitan *et al.*, 2014). The WWTP effluent containing biomass (2 L sample) was centrifuged at 2600 ×g and 4°C for 20 min, and the supernatant was discarded and pelleted. The concentrated wet biomass (2 mL) was washed with 1X PBS buffer followed by centrifugation. Genomic DNA was recovered by lysing the cells with 700 μL of lysis buffer (50 mM Tris, 5 mM EDTA, 150 mM NaCl, 1% Nonidet P-40, and 1 mM phenylmethylsulfonyl fluoride at pH 8.0). The mixture was then frozen-thawed for 5 minutes in ice-ethanol and subsequently for 3 minutes in a water-bath at 65°C, which was repeated 5 times. Further, samples were pre-treated with 20 μL of Proteinase K (10 mg mL^{-1}) and incubated in a water bath at 37°C for 30 min. The RNA and proteins were separated from the aqueous solution containing DNA with an equal volume of phenol, chloroform, and isoamyl alcohol (25:24:1) followed by chloroform–isoamyl alcohol (24:1). Further, 1 × volume of isopropanol was added

to the genomic DNA precipitate and stored at -20°C overnight. The DNA precipitate was centrifuged at 12000 rpm at 4°C, for 20 minutes. The DNA was then washed with 70% ethanol, and air dried at room temperature for 20 minutes. Subsequently the pellet was dissolved in 100 μL of TE buffer and the purified DNA was stored at -20°C for further use. The quantity and purity of the genomic DNA were assessed via Spectrophotometer (NanoDrop Technologies, ND-1000; USA) at an absorbance value of 260 nm, and the A260/A280 ratio of about 1.8. The purified DNA was stored at -20°C for further processing.

5.2.2.2. Quantitative Polymerase Chain Reaction (qPCR)

The primer sets developed for the quantitative polymerase chain reaction (qPCR) analysis were selected based on their efficiency and specificity in amplifying *invA*, *eae* and *IS6110 genes* (Table 5.1). The total number of the selected bacterial gene copies in the samples were determined based on real-time PCR analysis (C-1000 Touch, CFX 96, Bio-Rad Laboratories Pty., Ltd.; USA). The qPCR amplification and the thermal cycling parameters are summarized in Table 5.2. For all reactions, a final extension step of 72ºC for 2 min was carried out followed by a high resolution melting curve analysis to ensure product specificity. The individual standard curves for the targeted genes were generated from the purified DNA extracted from pure cultures of *Mycobacterium* sp., *Salmonella* spp. and *E. coli*.

For the different pathogens, individual standard curves were prepared using the targeted DNA. To create the copy number standard for the respective target genes, 2 μL of template genomic DNA was amplified using Sso fast green Master Mix (SsoFast™ EvaGreen® Supermix, Bio- Rad Laboratories Pty., Ltd.; USA). For each reaction, 0.5 μL of dNTPs (10 mM), 2 μL of template genomic DNA (10 ng μL^{-1}), 1 μL of each primer (5 μM), and 4 μL of Sso fast green master mix were prepared, and then nuclease-free water was added to a final volume of 10 μL. For each primer set tested against wastewater genomic DNA, a negative control (without template DNA) was examined. The negative controls were included in the aforementioned procedures to ensure that the recorded fluorescence signals were indicative of PCR amplification of the DNA template.

Table 5.1: Primers used for PCR amplification and their target specificity

Target microorganism	Gene	Primer sequence (5' - 3')	Amplicon size	Reference
Salmonella spp.	*InvA* - F	F-GTGAAATTATCGCCACGTTCGGGCAA	284	(Jyoti *et al.*, 2010)
	InvA - R	R-TCATCGCACCGTCAAAGGAACC		
E. coli O157:H7	*Eae* - F	F-GTAAGTTACACTATAAAAGCACCG TCG	106	(Barak *et al.*, 2005)
	Eae - R	R-TCTGTGTGGATGGTAATAAATTTTTG		
Mycobacteria	IS6110 - F	F-CCTGCGAGCGTAGGCGTCGG	123	(Brosch *et al.*, 2002)
	IS6110 - R	R-CTCGTCCAGCGCCGCTTCGG		

Table 5.2. Primers used for PCR amplification of genetic markers for pathogenic bacteria

Target pathogenic bacteria	PCR primer	Initial denaturation		Cycles	Denaturation		Annealing		Extension		Reference
		°C	min		°C	S	°C	S	°C	S	
E. coli O157:H7	*eae* 0157:H7	95	10	40	94	20	55	30	72	45	(Barak *et al.*, 2005)
Salmonella	*invA* gene	94	7	40	95	30	55	60	72	60	(Jyoti *et al.*, 2010)
Mycobacterium	IS6110	95	5	40	95	30	66	30	72	60	(Brosch *et al.*, 2002)

The following calculation was used to estimate the copy number as shown below (equation 5.1):

$$Number\ of\ copies = \frac{Amount\ in\ ng\ x\ Avogadro's\ const.}{Length\ in\ bp\ x\ 1\ x\ 10^{9}\ x\ 650} \qquad \textbf{Equation 5-1}$$

Where, the average weight of a base pair (bp) is 650 Daltons and the Avogadro's constant is 6.022×10^{23}. The DNA of the reference strains were serially diluted to the final concentrations ranging from 10^{1} to 10^{8} genomic DNA copies per PCR reactions.

5.2.3. Quantitative Microbial Risk Assessment Framework

The QMRA framework used in this study is shown in Figure 5.1, whereby the model had factored two variables *i.e.* SVI and pathogen load, that where evaluated in selected WWTPs. Data of SVIs at full-scale WWTPs were collected as reported in a previous study (Deepnarain *et al.*, 2019). Sludge bulking conditions were classified into low (SVI < 100 mL g^{-1}), medium (SVI = 100 – 200 mL g^{-1}), and high (SVI > 200 mL g^{-1}) levels. The relative abundances of three waterborne microorganisms, namely *E. coli* O157:H7, *Salmonella*, and *Mycobacterium*, were quantified using the qPCR assay (see section 5.2.2. Microbial Analysis), and incorporated into the QMRA model. For the analysis of potential risks using the QMRA approach, it was assumed that the copy numbers were of viable bacterial cells. The risks of infection of these pathogens and their impacts on population and downstream watercourse were modelled using the QMRA method. The model was simulated for 10,000 iterations using Monte Carlo techniques for the probability of infections. The @Risk 7.5 software (Palisade Corporation; USA) add-on to Excel was used for the simulation step (Amoah *et al.*, 2018c).

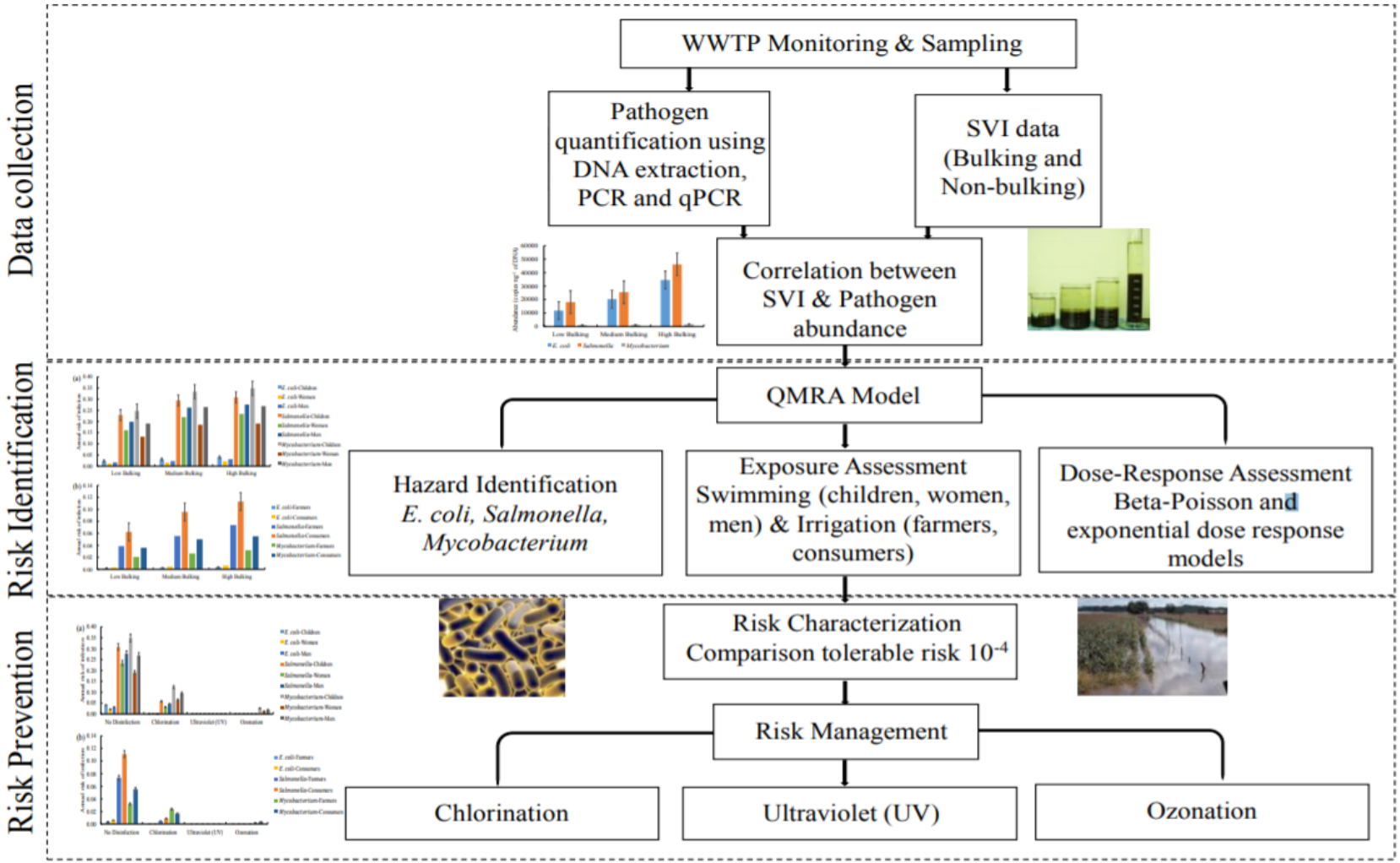


Figure 5.1. QMRA framework for evaluation of health risk infection associated with reuse of WWTPs effluents under different sludge bulking events.

The estimated risks were compared to the tolerable risk thresholds recommended by (a) WHO for infection risk associated with drinking water, 10^{-4} case per person per year (WHO, 2001, WHO, 2006), and (b) Recreational Water Quality Criteria (RWQC) in U.S. Environmental Protection Agency (EPA) for infection risks associated with recreational contacts, 32 cases per 1000 events (USEPA, 2012). Based on the outcome, the risk levels were categorised as; low, moderate, and high risks using probable risks ranges of 0.05 – 0.10, 0.10 – 0.30, and 0.30 – 0.50, respectively. These categories were based on a qualitative assessment of variation in the human health risk associated with the pathogenicity of *E. coli*, *Salmonella*, and *Mycobacterium*, as previously reported by Soller *et al.* (2010). The QMRA model was also used to define the treatment targets in relation to the required level of safety and health-based objective for the local context. This step was done by considering three disinfection scenarios and estimating the corresponding reduction of risks. Similar procedures have been reported by Petterson and Stenström (2015), applying a QMRA model to investigate the inactivation efficiency of several pathogens including *E. coli* O157 and *Campylobacter* in drinking water using free chlorine disinfection.

The QMRA framework involves four steps using experimental, practical, and theoretical assessments, which can be explicitly described as follows (Haas *et al.*, 2014, Amoah *et al.*, 2018c, Amoah *et al.*, 2018b):

5.2.3.1. Hazard Identification

Risks of infections from *E. coli* O157, *Salmonella*, and *Mycobacterium* were assessed via the QMRA model as a potential hazard for this study. The targeted pathogens were identified and quantified via qPCR (Cui *et al.*, 2017).

5.2.3.2. Exposure Assessment

The exposure assessment determines the point of exposure for the populations and the amount of pathogen consumed per scenario. The frequency of exposure per year was also considered in this study. The risks associated with three exposure scenarios were considered as the baseline concentration of pathogens in the effluents and the concentration in the effluent after a bulking

event. Risks assessed the use of water for recreational activities, including swimming (men, women and children), farming and consumers of vegetables.

The scenarios of exposure considered in this study were:

(a) Accidental ingestion of contaminated water during recreational activities such as bathing and swimming. Data for the exposure of swimmers (children, women, and men) to recreational waters were retrieved from a study by Schets *et al.* (2011), conducting a questionnaire approach to approximately 19000 participants. Exposure data covered the amount of water ingested during bathing and the frequency and duration of swimming events (Table 5.3). Further, categorical information was transformed into numerical data to measure mouthfuls of water during head submersion.

(b) Exposure to surface water during irrigation. This component considered the accidental ingestion of water contaminated with WWTP effluent by farmers or other farmworkers during the agricultural practices (Table 5.3) (WHO, 2006).

(c) The dietary intake of vegetables irrigated with surface water by local consumers. This scenario was described by a log-normal function (μ = 0.108, σ = 0.019 mL g^{-1}) (Hamilton *et al.*, 2006). The daily per capita intakes of vegetables for minimum, most likely, and maximum serving portions were assumed as 25, 50, and 75 g, respectively (Table 5.4) (Sant'Ana *et al.*, 2014).

Table 5.3: Amount of water ingested by exposed populations during different activities

Exposed population	Exposure event	Amount ingested	Reference
Men	Swimming	27-34 mL	(Schets *et al.*, 2011)
Women		18-23 mL	
Children		31-51 mL	
Farmers	Irrigation	1-5 mL	(WHO, 2006)

Table 5.4: Amount of surface water ingested by consumers of vegetables

Exposure scenario	Volume of water ingested	Reference
Volume of water on vegetable	Uniform distribution (0.108, 0.019)	(Hamilton *et al.*, 2006)
Daily per capita intake of vegetables	Pert distribution (25,50,75)	(Sant'Ana *et al.*, 2014)

The abundance of total pathogenic bacteria during sludge bulking would be more than that of the non-bulking condition. Hence, the risks associated with these three exposure scenarios were estimated at low, moderate, and high sludge bulking events in WWTPs.

The QMRA model was further used for the disinfection possibility of treated wastewater to eliminate and inactivate the pathogenic bacteria. The doses of Cl_2 residual, UV, and O_3 residual for the inactivation of *E. coli* O157, *Salmonella*, and *Mycobacterium*, as well as the corresponding reduction efficiencies, were adapted from (Sobsey, 1989); (Table 5.5). Results of the treatment cases were incorporated into the QMRA-based disinfection model by re-estimating the concentration of each pathogen for all exposure scenarios. Further, the new pathogenic abundances were used to evaluate possible reduced risks due to the implementation of these disinfection treatments.

Table 5.5. Inactivation of health-related bacteria in effluent wastewater by chlorination, ultraviolet radiation, and ozonation, adapted from Sobsey (1989).

Pathogen	Chlorination	UV	Ozone
E. coli O157:H7	$C1_2$ residual 0.1 – 0.2 mg L^{-1} Reduction 90 – 99%	UV dose 8.2 mWs cm^{-2} Reduction 99.9%	O_3 residual 0.29 – 0.36 mg L^{-1} Reduction 99.9%
Salmonella	$C1_2$ residual 0.1 – 0.2 mg L^{-1} Reduction 90 – 99%	UV dose 8.2 mW s cm^{-2} Reduction 99.9%	O_3 residual 0.29 – 0.36 mg L^{-1} Reduction 99.9%
Mycobacterium	Cl_2 residual 0.3 – 1.0 mg L^{-1} Reduction 40 – 99%	UV dose 60 mWs cm^{-2} Reduction 99.9%	O_3 residual 0.29 – 1.08 mg L^{-1} Reduction 89 – 99%

5.2.3.3.Dose-Response Assessment

The dose-response model was employed to assess the risk of infection for the exposed population based on the dose (concentration) of each pathogen ingested (Table 5.6). This step involved the use of models developed either during feeding trials or disease outbreaks. The models differ between different pathogens. Table 5.6 below presents the different dose-response models used in this study for the risk assessments. The QMRA simulations were implemented with an @ Risk programme (v 7.5, Palisade).

Table 5.6: Dose-response models and parameters chosen for the risk assessment

Pathogen	Model	Parameters	
Mycobacterium	Exponential	K= 6.93E-04	
Salmonella	Beta-poisson	α=1.75E-01	N_{50}=1.11E+06
E. coli	Beta-poisson	α=1.55E-01	N_{50}=2.11E+06

The Beta- Poisson Dose- Response model used for both *E. coli* and *Salmonella* are represented by equation 5.2 (Xie *et al.*, 2016). The exponential model used for *Mycobacterium* is given by equation 5.3 (Haas *et al.*, 2014).

$$P_I(d) = 1 - \left[1 + \left(\frac{d}{N_{50}}\right)\left(2^{\frac{1}{\alpha}} - 1\right)\right]^{-\alpha} \qquad \textbf{Equation 5-2}$$

$$P_I(d) = 1 - \exp(-k \cdot d) \qquad \textbf{Equation 5-3}$$

where, $P_I(d)$ is the risk of infection due to a particular pathogen, *"d"* is the dose (concentration) of pathogen ingested in an identified volume of surface water or crops, N_{50} is the median infection dose equivalent to the number of organisms that will infect 50% of the exposed population ($N_{50} = 2.11\times10^6$ for *E. coli* and 1.11×10^6 for *Salmonella*), and *"α"* is a dimensionless infectivity constant ($\alpha = 1.55\times10^{-1}$ for *E. coli* and 1.75×10^{-1} for *Salmonella*), and *"k"* is the infectivity constant ($k = 6.93\times10^{-4}$ for *Mycobacterium*).

5.2.3.4.Risk Characterisation

In the risk characterisation step, the probability of infection for exposed populations is characterised using all the outputs of the previous actions, *i.e.*, hazard identification, exposure assessment, and dose-response assessment. The risk of infection was determined using the formula in equation 5.4 (Sakaji and Funamizu, 1998):

$$P_I(A) = 1 - [1 - P_I(d)]^n \qquad \textbf{Equation 5-4}$$

where, $P_I(A)$ represents the risk of infection associated with multiple exposures or annual risk, $P_I(d)$ is the risk of infection from a single exposure to a dose *"d"* of the pathogen, and *"n"* denotes the number of days of exposure to the single-dose *"d"*.

5.3. RESULTS AND DISCUSSION

5.3.1. Pathogenic Bacteria Detection and Quantification

Table. 5.7 lists the results of molecular real-time PCR tests generated from 10-fold dilution series of each target gene (*eae*-, *invA*-, and *IS6110*- gene). The constructed standard curves of the qPCR assays revealed adequate amplification efficiencies for the selected primers. Both *eae* 0157:H7 and *invA* assays had almost 100% PCR efficiency for *E. coli* O157:H7 and *Salmonella* primers, respectively. The *IS6110* assay had 92% efficiency for *Mycobacterium*. Estimates of the qPCR efficiency may be influenced by many factors *viz*., probe design, sample contamination, the primer and template concentrations, interfering with the normal reaction (Cui *et al.*, 2017). For qPCR assays, amplification of the targeted DNA associated with the selected primers, efficiencies between 90-110%, is generally acceptable for environmental samples (Svec *et al.*, 2015). The obtained efficiencies (Table 5.7) revealed acceptable accuracy due to the low possibility for the presence of inhibitors during the PCR analysis, as well as the stability of the DNA extraction process (Barak *et al.*, 2005).

Table 5.7. Estimation of parameters obtained from qPCR assays

Parameter	***E. coli* 0157:H7**	***Salmonella***	***Mycobacterium***
Efficiency (%)	103	109.1	92.3
Slope	-3.251	-3.121	-3.522
R^2 of slope	0.994	0.983	0.998

As shown in Figure 5.2, *Salmonella* was the most abundant species during the study period (2270– 96733 copies ng^{-1} of DNA) followed by *E. coli* O157:H7 (4133 – 76847 copies per ng of DNA); whereas, *Mycobacterium* was the least (542 – 3340 copies ng^{-1} of DNA). These pathogens have been known to cause various potential hazards to human health and receiving environments in South Africa (Ekwanzala *et al.*, 2018).

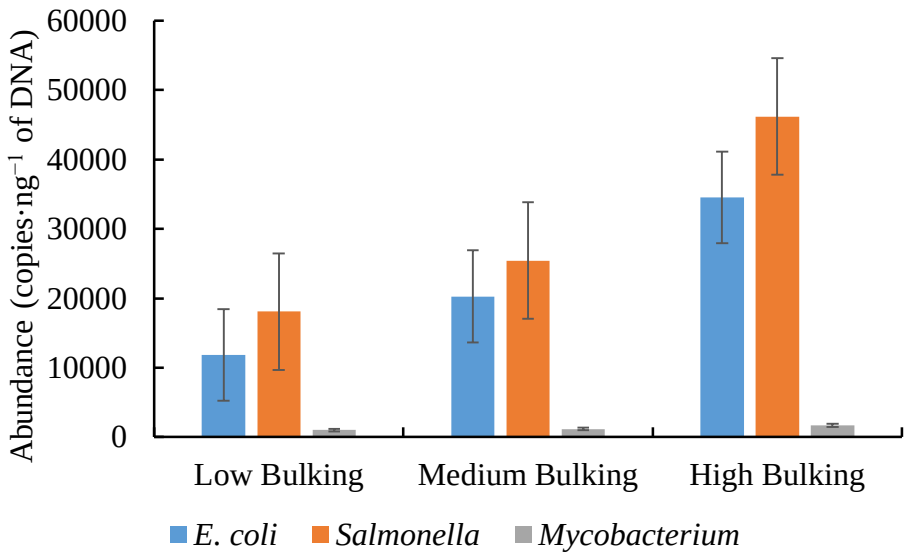


Figure 5.2. Sludge bulking events and abundances of *E. coli* O157:H7, *Salmonella*, and *Mycobacterium* in WWTP effluent.

Figure 5.2 also depicted that an increase in copy number for the selected pathogens occurred mostly during the sludge bulking events in the WWTPs. High bulking incidents (SVI > 150 mL g^{-1}) interfere with the compaction and settling properties of the sludge bioflocs (Martins *et al*., 2004). Bulking could lead to high concentrations of TSS in the final effluent that can exceed the discharge permit limitation, which was also evident in this study. Hence, bulking can reduce the efficiency of the activated sludge, which can further result in the deterioration of WWTP and receiving water bodies (Jenkins *et al.*, 2003, Lou and Zhao, 2012). Chapter four demonstrated that under some unfavourable conditions (e.g., low concentrations of DO and nutrients), filamentous bacteria extend outside the floc structures, and form inter-floc bridging.

Large quantities of *E. coli* O157:H7 were found in the wastewater effluent during high bulking episodes (Figure 5.2). *E. coli* is a common fecal indicator organism of which certain strains can cause diarrheal disease, acute gastrointestinal illness, hemorrhagic colitis, and kidney failure in humans (Ayaz *et al.*, 2014). The detection of *E. coli* in high abundance could be because this pathogenic bacterium is efficiently eliminated via a biological filtration step (Radomski *et al.*, 2011), which was not included in the monitored WWTPs.

A positive correlation ($p < 0.05$) was also found between *Salmonella* in the final effluent and SVI (Table 5.7; Figure 5.2). During the high bulking events, a large amount of *Salmonella*

could pass downstream (*i.e.*, to the receiving environment) through the effluent weir system. Espigares *et al.* (2006) reported that surface water, contaminated irrigation water, and wastewater pathways could release different serotypes of *Salmonella* to the environment. This Gram negative enteric pathogen could be associated with various human diseases such as typhoid fever, gastrointestinal infection, bacteremia, and salmonellosis (Amha *et al.*, 2017). However, the disinfection options such as ozonation and UV has been reported to reduce the negative impacts caused by *Salmonella* (Espigares *et al.*, 2006, Jyoti *et al.*, 2010, Collivignarelli *et al.*, 2017).

Figure 5.2 depicted that the copy number of *Mycobacterium* increased significantly ($p < 0.05$) during high bulking with SVI >200 mL g^{-1}. The relative abundance of *Mycobacterium* increased with sludge bulking, recording average values of 1002, 1174, and 1670 copies ng^{-1} of DNA under the low, medium, and high sludge bulking conditions, respectively (Fig. 5.2). Similarly, Olson and Asvapathanagul (2017) reported that *Mycobacterium* could be associated to bulking in activated sludge WWTPs. Moreover, Cai and Zhang (2013) and Amha *et al.* (2017) revealed that *Mycobacteria* were identified in treated wastewater and urban watershed, causing infections to human receptors (e.g., Cervical lymphadenitis in children). Radomski *et al.* (2011) detected about 5.5×10^5 copies L^{-1} of *Mycobacterium* in the domestic, influent wastewater, which was mostly eliminated via physical-chemical decantation followed by biofiltration. The incidence of *Mycobacterium* in wastewater samples suggested that the WWTP effluent should be treated before disposal or reuse. Hence, in the following sections, the microbial risks of these pathogenic bacteria were assessed using the QMRA technique.

5.3.2. Quantitative Microbial Risk Assessment and Outputs

The data shown in Figure 5.3 (also clearly represented in Tables 5.8 and 5.9) depict the QMRA outcomes after using the wastewater effluent containing bacterial pathogens in (a) recreational activities (swimming) for children, women, and men (Figure 5.3a), (b) farming practice (Figure 5.3b), and (c) irrigation purposes (Figure 5.3b).

The ANOVA analysis showed no statistically significant difference in the risks when comparing the different assessed exposure scenarios (p value ≥ 0.05). However, there was statistically significant difference in the risks of infection with these pathogens under low

bulking and high bulking events (p values ≤ 0.05). Therefore, showing the impact of high bulking on the level of risks posed to the various populations for each reuse scenario studied.

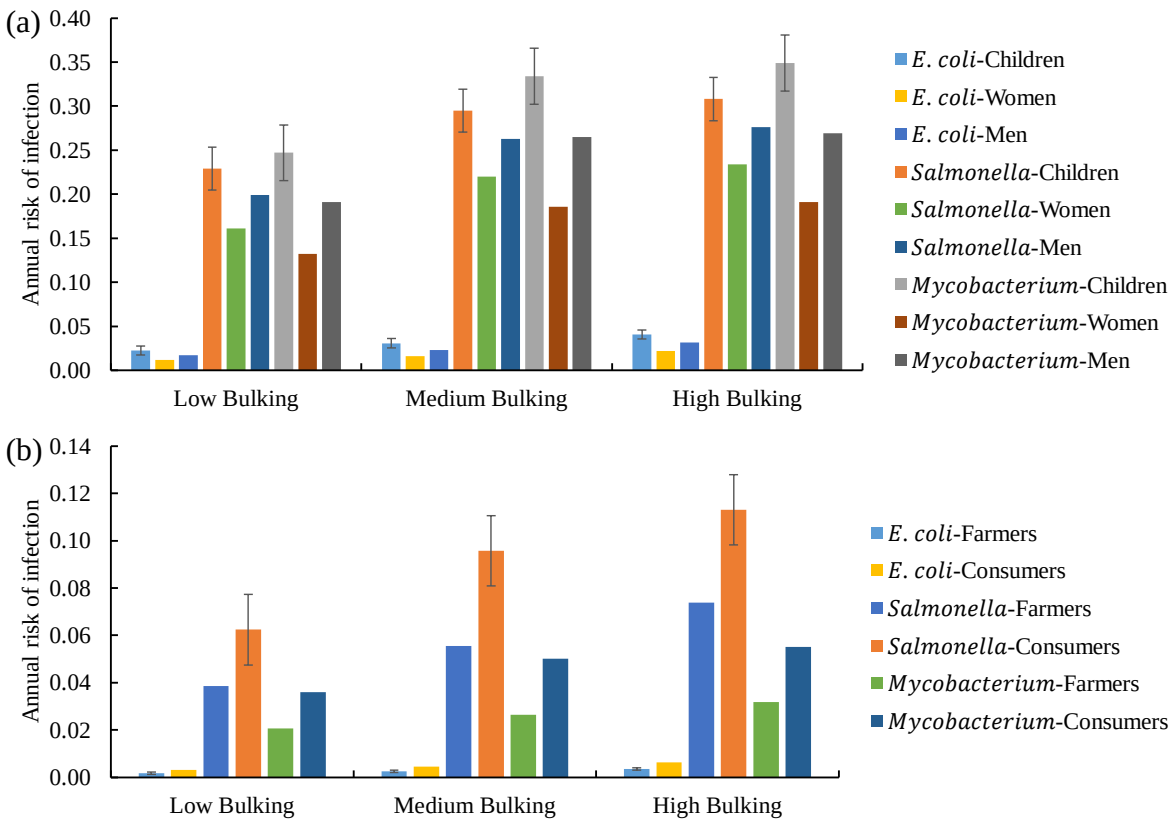


Figure 5.3. Annual risk of infection using the quantitative microbial risk assessment framework for wastewater re-use: (a) recreational (swimming and bathing) activities associated with children, women, and men, and (b) irrigational practices associated with farmers and vegetable consumers.

Table 5.8. Estimated risks of infection for recreational users (±90% CI)

	Low sludge bulking			Moderate sludge bulking			High sludge bulking		
Pathogen	**Children**	**Women**	**Men**	**Children**	**Women**	**Men**	**Children**	**Women**	**Men**
E. coli O157:H7	2.26×10^{-2} ± 4.82×10^{-5}	1.17×10^{-2} ± 1.31×10^{-5}	1.70×10^{-2} ± 1.76×10^{-5}	3.08×10^{-2} ± 6.28×10^{-4}	1.62×10^{-2} ± 4.57×10^{-4}	2.33×10^{-2} ± 5.54×10^{-4}	4.07×10^{-2} ± 8.73×10^{-4}	2.18×10^{-2} ± 6.77×10^{-4}	3.15×10^{-2} ± 7.87×10^{-4}
Salmonella	2.29×10^{-1} ± 2.45×10^{-4}	1.61×10^{-1} ± 1.08×10^{-4}	1.99×10^{-1} ± 1.10×10^{-4}	2.95×10^{-1} ± 2.74×10^{-3}	2.20×10^{-1} ± 1.81×10^{-3}	2.63×10^{-1} ± 2.46×10^{-3}	3.08×10^{-1} ± 9.02×10^{-4}	2.34×10^{-1} ± 9.03×10^{-4}	2.76×10^{-1} ± 8.92×10^{-4}
Mycobacterium	2.47×10^{-1} ± 4.96×10^{-4}	1.32×10^{-1} ± 1.43×10^{-4}	1.91×10^{-1} ± 1.86×10^{-4}	3.34×10^{-1} ± 1.75×10^{-3}	1.86×10^{-1} ± 1.02×10^{-3}	2.65×10^{-1} ± 1.37×10^{-3}	3.49×10^{-1} ± 1.71×10^{-3}	1.91×10^{-1} ± 1.17×10^{-3}	2.69×10^{-1} ± 1.44×10^{-3}

0.05-0.10	Low risk
0.10-0.30	Moderate risk
0.30-0.50	High risk

Table 5.9. Estimated risks of infection for farmers and consumers associated with reuse of effluents for irrigation (±90% CI)

Pathogen	Low sludge bulking		Moderate sludge bulking		High sludge bulking	
	Farmers	**Consumers**	**Farmers**	**Consumers**	**Farmers**	**Consumers**
E. coli O157	1.79×10^{-3} $\pm 1.13\times10^{-5}$	3.15×10^{-3} $\pm 1.35\times10^{-5}$	2.58×10^{-3} $\pm 1.63\times10^{-4}$	4.48×10^{-3} $\pm 2.25\times10^{-4}$	3.50×10^{-3} $\pm 2.73\times10^{-4}$	6.22×10^{-3} $\pm 3.74\times10^{-4}$
Salmonella	3.86×10^{-2} $\pm 2.16\times10^{-4}$	6.24×10^{-2} $\pm 2.17\times10^{-4}$	5.56×10^{-2} $\pm 6.91\times10^{-4}$	9.58×10^{-2} $\pm 9.38\times10^{-4}$	7.39×10^{-2} $\pm 7.52\times10^{-4}$	1.13×10^{-1} $\pm 8.36\times10^{-4}$
Mycobacterium	2.06×10^{-2} $\pm 1.29\times10^{-4}$	3.60×10^{-2} $\pm 1.54\times10^{-4}$	2.64×10^{-2} $\pm 2.61\times10^{-4}$	5.01×10^{-2} $\pm 3.81\times10^{-4}$	3.19×10^{-2} $\pm 3.30\times10^{-4}$	5.52×10^{-2} $\pm 5.02\times10^{-4}$

0.05-0.10	Low risk
0.10-0.30	Moderate risk
0.30-0.50	High risk

Table 5.10: Probabilistic risk of infection with the incorporation of further treatment options for recreational use (swimming) use of surface water contaminated with WWTP effluents at high bulking condition

		Disinfection Treatment Scenario			
Pathogen	**Exposure**	**No further treatment**	**Chlorination (free chlorine)**	**UV**	**Ozonation**
	Children	$4.07\times10^{-2} \pm 8.73\times10^{-4}$	$2.62\times10^{-3} \pm 2.47\times10^{-4}$	$4.74\times10^{-5} \pm 1.92\times10^{-5}$	$2.84\times10^{-5} \pm 1.27\times10^{-5}$
	Women	$2.19\times10^{-2} \pm 6.77\times10^{-4}$	$1.32\times10^{-3} \pm 1.61\times10^{-4}$	$2.37\times10^{-5} \pm 1.16\times10^{-5}$	$1.42\times10^{-5} \pm 8.14\times10^{-6}$
E. coli O157:H7	Men	$3.14\times10^{-2} \pm 7.87\times10^{-4}$	$1.97\times10^{-3} \pm 1.91\times10^{-4}$	$3.54\times10^{-5} \pm 1.52\times10^{-5}$	$2.12\times10^{-5} \pm 9.82\times10^{-6}$
	Children	$3.08\times10^{-1} \pm 9.02\times10^{-4}$	$5.82\times10^{-2} \pm 7.13\times10^{-4}$	$1.26\times10^{-3} \pm 4.01\times10^{-5}$	$7.58\times10^{-4} \pm 2.57\times10^{-5}$
	Women	$2.34\times10^{-1} \pm 9.03\times10^{-4}$	$3.26\times10^{-2} \pm 4.97\times10^{-4}$	$6.25\times10^{-4} \pm 2.14\times10^{-5}$	$3.75\times10^{-4} \pm 1.34\times10^{-5}$
Salmonella	Men	$2.76\times10^{-1} \pm 8.92\times10^{-4}$	$4.60\times10^{-2} \pm 6.15\times10^{-4}$	$9.28\times10^{-4} \pm 2.99\times10^{-5}$	$5.58\times10^{-4} \pm 1.89\times10^{-5}$
	Children	$3.49\times10^{-1} \pm 1.70\times10^{-3}$	$1.24\times10^{-1} \pm 1.44\times10^{-3}$	$4.30\times10^{-4} \pm 4.88\times10^{-6}$	$2.59\times10^{-2} \pm 3.47\times10^{-4}$
	Women	$1.90\times10^{-1} \pm 1.17\times10^{-3}$	$6.51\times10^{-2} \pm 8.10\times10^{-4}$	$2.12\times10^{-4} \pm 2.27\times10^{-6}$	$1.32\times10^{-2} \pm 1.83\times10^{-4}$
Mycobacterium	Men	$2.69\times10^{-1} \pm 1.43\times10^{-3}$	$9.53\times10^{-2} \pm 1.13\times10^{-3}$	$3.15\times10^{-4} \pm 3.59\times10^{-6}$	$1.96\times10^{-2} \pm 2.65\times10^{-4}$

N.B: Chlorination: 0.2 mg L^{-1} (Cl_2 residual) for both *E. coli* and *Salmonella*, and 0.3 – 1.0 mg L^{-1} (Cl_2 residual) for *Mycobacterium*. UV: 8.2 mW s cm^{-2} for both *E. coli* and *Salmonella*, and 60 mW s cm^{-2} for *Mycobacterium*. Ozonation: 0.29 – 0.36 mg L^{-1} (O_3 residual) for both *E. coli* and *Salmonella*, and 0.29 – 1.08 mg L^{-1} (O_3 residual) for *Mycobacterium*.

0.05-0.10	Low risk
0.10-0.30	Moderate risk
0.30-0.50	High risk

Table 5.11: Probabilistic risk of infection with the incorporation of further treatment options for irrigational use of surface water contaminated with WWTP effluents at high bulking condition

		Disinfection treatment scenario			
Pathogen	**Exposure**	**No treatment**	**Chlorination (Free chlorine)**	**UV**	**Ozonation**
E. coli	Farmers	$3.48\times10^{-3} \pm 2.77\times10^{-4}$	$1.83\times10^{-4} \pm 7.97\times10^{-5}$	$3.51\times10^{-6} \pm 2.31\times10^{-6}$	$2.11\times10^{-6} \pm 1.41\times10^{-6}$
	Consumers	$6.18\times10^{-3} \pm 3.69\times10^{-4}$	$3.48\times10^{-4} \pm 6.67\times10^{-5}$	$6.28\times10^{-6} \pm 2.12\times10^{-6}$	$3.77\times10^{-6} \pm 1.29\times10^{-6}$
Salmonella	Farmers	$7.35\times10^{-2} \pm 7.54\times10^{-4}$	$4.72\times10^{-3} \pm 1.51\times10^{-4}$	$9.66\times10^{-5} \pm 3.90\times10^{-6}$	$5.79\times10^{-5} \pm 2.37\times10^{-6}$
	Consumers	$1.11\times10^{-1} \pm 8.33\times10^{-4}$	$8.84\times10^{-3} \pm 2.17\times10^{-4}$	$1.68\times10^{-4} \pm 6.25\times10^{-6}$	$1.01\times10^{-4} \pm 3.78\times10^{-6}$
Mycobacterium	Farmers	$3.19\times10^{-2} \pm 3.56\times10^{-4}$	$2.37\times10^{-2} \pm 1.61\times10^{-4}$	$3.24\times10^{-5} \pm 3.78\times10^{-7}$	$1.71\times10^{-3} \pm 2.99\times10^{-5}$
	Consumers	$5.52\times10^{-2} \pm 5.07\times10^{-4}$	$1.67\times10^{-2} \pm 2.52\times10^{-4}$	$5.69\times10^{-5} \pm 5.68\times10^{-7}$	$3.29\times10^{-3} \pm 4.73\times10^{-5}$

N.B: Chlorination: 0.2 mg L^{-1} (Cl_2 residual) for both *E. coli* and *Salmonella*, and 0.3 – 1.0 mg L^{-1} (Cl_2 residual) for *Mycobacterium*. UV: 8.2 mW s cm^{-2} for both *E. coli* and *Salmonella*, and 60 mW s cm^{-2} for *Mycobacterium*. Ozonation: 0.29 – 0.36 mg L^{-1} (O_3 residual) for both *E. coli* and *Salmonella*, and 0.29 – 1.08 mg L^{-1} (O_3 residual) for *Mycobacterium*.

0.05-0.10	Low risk
0.10-0.30	Moderate risk
0.30-0.50	High risk

5.3.3. Risk of Infection during Recreational Exposure

The QMRA values for *E. coli* O157:H7 were in the range of 0.01 – 0.04, considering the estimated risks associated with swimming-related activities, for three population sub-groups *i.e.* men, women and children (Figure 5.3a and Table 5.8). The probability of annual infection for children and adults exposed to pathogenic *E. coli* O157:H7 was low during swimming, *i.e.*, compared to the risks estimated from *Salmonella* (0.16 –0.31) and *Mycobacterium* (0.13 – 0.35). Accordingly, medium to high risk values were found during population (men, women, and children) exposure to either *Salmonella* or *Mycobacterium* (Table 5.8). In addition, the QMRA values generally increased with the rise in SVIs, suggesting that the high bulking conditions contributed to the human health risks during swimming (Figure 5.3a). Moreover, in all bulking scenarios, the 10^{-4} risk level (WHO, 2006) was exceeded.

Children were almost seven-fold more likely to be infected with *Salmonella* and *Mycobacterium* as compared to *E. coli* O157:H7. Moreover, children were more susceptible to bacterial infections, representing a risk level of approximately 1.5 to 2.0-fold higher than that for adults. The elevated risks driven by children could be attributed to the accidental water ingestion during swimming (Sunger *et al.*, 2018). The values are based on Schets *et al.* (2011) reporting intake of 31 – 51 mL of water while swimming for children as compared to 18 – 23 mL for women and 27 – 34 mL for men. Similarly, Suppes *et al.* (2016) reported that the annual infection risk for adults (>18) and children (≤18) while swimming were 2.2×10^{-2} and 2.9×10^{-2}, respectively. Their study hypothesized that the greater level of risk for children compared to adults could be attributed to underdeveloped immune systems, and participation in more swimming activities such as splashing, playing, and diving. In addition, Vergara *et al.* (2016) stated that the illness risk for children swimming was 46% higher than for adults. Cui *et al.* (2017) found that the incidences of hand-to-mouth/eye contact of children, caused the greatest infection risk, regarding three pathogens *Mycobacterium avium*, *Salmonella*, and *Pseudomonas aeruginosa*, using the QMRA framework.

5.3.4. Risk of Infection during Agricultural Practice

Based on the health risks linked to the application of treated wastewater in farming activities (Figure 5.3b and Table 5.9), the highest risks of infection were 0.04 – 0.07 for *Salmonella*, which dropped to 0.02 – 0.03 for *Mycobacterium*, and below 0.01 for *E. coli* O157:H7. These risks estimates exceeded the WHO recommended limit of the 10^{-4} risk to develop wastewater reuse, indicating possible adverse health impacts (WHO, 2006). Handling, transportation, and storage of contaminated water could be important infection pathways during farming (Amenu *et al.*, 2016). Van Vu *et al.* (2018) reported that some farmers' actions such as touching the mouth with hand, wiping mouth using arm, and other unskillful and untrained practices during work would increase the possible risks/diseases. Moreover, farmers could accidentally ingest water (about 1–5 mL (WHO, 2006)) while working in the wastewater-irrigated fields (Moazeni *et al.*, 2017). Furthermore, nearby populations could be exposed to microbial agents when employing spray irrigation of wastewater. The relative abundance of pathogenic bacteria (*E. coli* O157:H7, *Salmonella*, and *Mycobacterium*) implied that necessary actions should be considered for wastewater reuse in irrigation.

The effect of bulking scenarios on the concentrations of organisms as further calculated in the QMRA levels was also significant (p-value < 0.05). For instance, under the low sludge bulking condition, at least 3 farmers out of 100 (90% Confidence Interval: CI) were at risk of being infected due to exposure to *Salmonella*. This value increased to 7 out of 100 (90% CI) at a high bulking condition (Table 5.9). As a consequence, it can be concluded that farmers suffered from higher risks due to irrigation with water contaminated by effluents of WWTPs during moderate- to high- bulking scenarios. Confidence interval can be expressed as samples or repeated samples, whereby the procedure can be repeated and the fraction of the calculated CI that encompass the true population parameter would tend towards 90% (Cox and D.V., 1974). Sampson *et al.* (2017) found that the annual risk of infection for farmers in frequent contact with wastewater did not meet the stringent risk benchmark of 10^{-6}. They concluded that some proactive mitigations and behavioral strategies such as the application of drip irrigation, wearing face mask, gloves, and boots, and avoiding unhygienic practices would protect the farmers against adverse health risks (Sampson *et al.*, 2017). Moreover, awareness campaigns directed towards the farmer communities would be useful to reveal the potential risks related to the wastewater irrigation strategies.

5.3.5. Risk of Infection for Vegetable Consumers

Treated wastewater can provide nutrients for plant growth; however, the consumption of wastewater-irrigated crops might cause various infections to end-users (Amha *et al.*, 2015). At least 6 consumers out of 100 may be infected due to the intake of vegetables irrigated with wastewater containing *Salmonella* (Table 5.9). This risk increased during high bulking events, in which the risk estimates were 1.13×10^{-1} ($\pm8.36\times10^{-4}$) (*i.e.*, more than 10 incidences per 100; 90% CI) at SVI > 200 mL g^{-1}. Comparable patterns were noticed with *E. coli* O157:H7 and *Mycobacterium* pathogens during all bulking conditions. Hence, the application of wastewater in agriculture would be an important transmission route of infection, tending to increase the pathogen load and microbiological health hazard. Beaudequin *et al.* (2016) applied the QMRA method to characterise the human health risks associated with wastewater-irrigated lettuce. Their work depicted that some scenarios such as lettuce washing prior to sending to market and withholding irrigation before harvesting could reduce the pathogenic effects (Beaudequin *et al.*, 2016). In Saudi Arabia, Balkhair (2016) found that the use of treated wastewater for crop production could have detrimental environmental and health impacts. Other health risks and routes of exposure associated with the wastewater irrigation practices have been reviewed in a previous paper by Dickin *et al.* (2016). Amha *et al.* (2017) demonstrated the need for disinfection and continuous monitoring of treated wastewater prior to reuse. These results suggested the necessity of water disinfection before land spreading to avoid potential human infections and food crop contamination.

5.3.6. Disinfection Scenarios Using QMRA Outcomes

Disinfection would be a viable option for complying treated wastewater with the health-based target (Kollu and Ormeci, 2012, Schijven *et al.*, 2019). Three disinfection scenarios, *viz.*, chlorination, UV, and ozonation, were therefore theoretically tested under simulated high bulking condition. The reduction efficiency used in this study for all three treatment scenarios was based on published data according to Sobsey (1989) (Table 5.5).

The QMRA levels associated with the application of chlorine, as an oxidant, to inactivate the pathogenic bacteria within the effluent discharge are illustrated in Figure 5.4 (a and b) and (Tables 5.10 and 5.11). As per risk mitigation framework, Cl_2 residual at 0.2 mg L^{-1} for both

E. coli O157:H7 and *Salmonella* were factored, whereas 0.3 – 1.0 mg L^{-1} (Cl_2 residual) was used for *Mycobacterium*. Although chlorination decreased the risks compared to the initial estimates for the direct use of wastewater (Figure 5.4), the reduction of health risks was statistically insignificant ($p > 0.05$). Moreover, the probable risk of infection for both *Salmonella*, and *Mycobacterium* exceeded the tolerable risk of 10^{-4} (WHO, 2006). Previous studies have demonstrated that chlorine might generate harmful disinfection by-products, causing a chemical threat to human health (Dong *et al.*, 2017). In this context, finding alternative disinfectants to chlorine would be a suitable solution. Petterson and Stenström (2015) used QMRA to study the infectious risks related to drinking water systems, and showed that chlorine disinfection inactivated the reference pathogens, e.g., *E. coli* O157:H7, *Giardia*, Rotavirus, *Campylobacter*, and Norovirus. In their study, Log_{10} reductions of the five pathogenic strains were calculated with assumed initial chorine residuals of 0.4 - 1.5 mg L^{-1}. Sanawar *et al.* (2017) assessed the disinfection process to eliminate antibiotic-resistant bacteria (ARB) from marine aquaculture effluents via the QMRA method. Their study revealed that monochloramination could attain an acceptable inactivation efficiency for ARB with the release of a minimum amount of trihalomethanes compared to the case of chlorination.

In Figure 5.4 (a and b), UV was employed to enhance the disinfection efficacy, according to the theoretical values retrieved from Sobsey (1989) (Table 5.5). The UV treatment reduced the levels of risks to below the tolerable threshold, except for the case of children exposed to *Salmonella* infection during swimming (Table 5.10). Kollu and Ormeci (2012) reported that some of the self-aggregated *E. coli* O157:H7 could resist high UV doses such as 90 mJ cm^{-2}. It has been reported that the UV disinfectant could effectively inactivate protozoan parasites within a relatively short contact time (Sobsey, 1989). Furthermore, UV processes are rarely associated with the release of disinfection by-products. Schijven *et al.* (2019) applied QMRA to assess the microbial safety of drinking water at Beerenplaat regarding several pathogens, including *Giardia*, *Campylobacter*, and *Cryptosporidium*. Their study depicted that chlorine dioxide was the essential treatment step; however, UV disinfection was effective with the addition of chlorine dioxide.

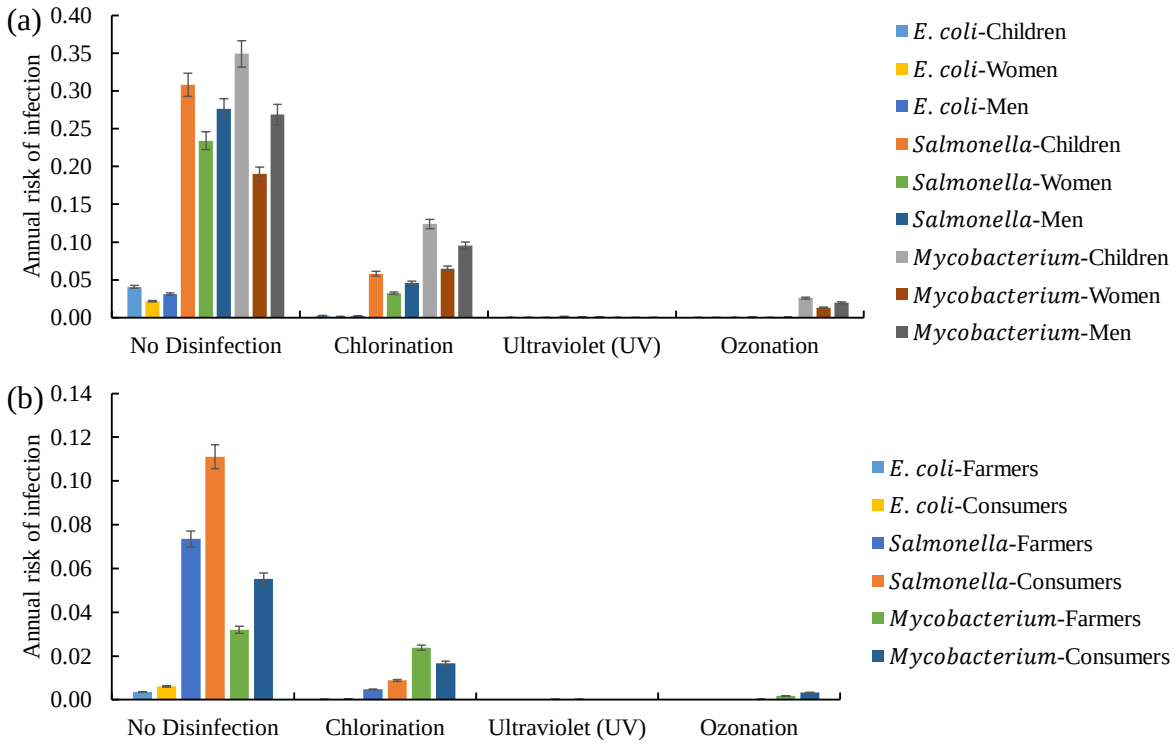


Figure 5.4. Risk mitigation framework at high bulking condition using disinfection scenarios (*viz*., chlorination, UV and Ozonation), during: (a) recreational (swimming and bathing) activities for children, women, and men, and (b) irrigational practices with farmers and vegetable consumers.

Ozone is a strong oxidizing agent that can disinfect a number of microorganisms including *E. coli*, *Giardia* and chlorine resistant bacteria such as *Cryptosporidium parvum* (Dong *et al.*, 2017). The mode of action for ozone is similar to chlorine, which targets the cell membrane, thus resulting in leakage of the cells constituents. Ozone may also diffuse directly through the cell membrane and destructs the chromosomal DNA (Caravelli *et al.*, 2006). In this study, ozonation showed an effective reduction of the estimated risks for all pathogens under the exposure scenarios except for *Mycobacterium* (Figure 5.4; Tables 5.10 and 5.11). The risk of *Mycobacterium* infection during ozonation was greater than the tolerable criterion of 10^{-4} recommended by WHO (2006). In a similar QMRA research study, Zhou *et al.* (2017) found that an O_3/UV contactor attained efficient inactivation of *E. coli* for drinking water systems. Dong *et al.* (2017) compared between ozonation and chlorination for the inactivation of three waterborne pathogens, namely *Giardia*, *Legionella pneumophila*, and *Cryptosporidium*

parvum. Their QMRA work depicted that the treated wastewater could be applied for irrigational landscape reuse. Moreover, Dong *et al.* (2017) indicated that the microplasma ozonation technique was an adequate wastewater disinfection treatment method, showing reduced human health impacts compared to the chlorination scenario.

Based on the aforementioned risk estimates, the UV and ozonation treatments could provide adequate options for attaining the tolerable risks of infection for the exposure scenarios and populations modelled. However, the case of *Mycobacterium* infection risks after ozonation should be reassessed. This finding could be because *Mycobacteria* tend to naturally aggregate in water, making a shield that protects them against disinfection (Loret and Dumoutier, 2019). Moreover, Amha *et al.* (2017) reported that various strains of *Mycobacteria* would be over 100-folds resistant to chlorine than *E. coli*.

Although the disinfection option provided essential results for reducing the human health risk associated with sludge bulking, further studies should focus on the economic feasibility of the chlorination, UV, and ozonation systems. For example, chlorination has been recognised as a cheaper option compared to other disinfection processes; however, it tends to generate toxic by-products, comprising trihalomethanes and haloacetic acids. Moreover, a dechlorination step would be required to avoid the negative impact of high chlorine levels on the aquatic environment, resulting in high capital and operating costs. The operational costs of chemical based disinfection (*i.e.*, Peracid) varied from 0.0114 to 0.0261 € (*i.e.*, about \$0.012 – 0.028 USD using an exchange rate of 1 € = \$1.09 USD) for 1 m^3 of treated water (Luukkonen *et al.*, 2015). The UV requires a pre-treatment step for the removal of turbidity and suspended solids, increasing the initial capital cost of the entire disinfection process. Moreover, the UV reactors would require high-energy consumption, making it not always the best disinfection option, especially in developing countries. However, for long-term operation, UV can be a cost effective solution compared to the chlorination-based systems (Tak and Kumar, 2017). Zhuang *et al.* (2015) found that the disinfection of 1 m^3 of wastewater using chlorination, UV irradiation, and UV/chlorination would cost 0.041, 0.046, and 0.034 Yuan, respectively (*i.e.*, about \$0.0058, \$0.0065, and \$0.0048 USD, respectively using 1 Yuan = \$0.14 USD). Ozone is a strong oxidant that diffuses directly through the cell membrane, and deactivates microorganisms; however, its long-run cost is negatively influenced by the presence of organic constituents such as humic substances and fatty acids in the aqueous phase. Accordingly, each disinfection method has its constraints not only in terms of efficiency, but also in terms of

feasibility, practicability, and cost. Further comparative studies on a life cycle cost analysis of the disinfection systems, including chemical utilisation, energy consumption, UV lamp replacement, land footprint, manpower, plant configuration and size, and miscellaneous equipment repair, are recommended.

5.3.7. Environmental Aspects for Improving the Precision of QMRA Outcomes

This is the first study describing the exposure routes and the risks of infection associated with pathogenic bacteria, *i.e.*, *E. coli* O157:H7, *Salmonella*, and *Mycobacterium spp.*, encountered by different sludge bulking conditions from full-scale WWTPs. These microorganisms exhibit various negative impacts on the environment (Ramirez-Castillo *et al.*, 2015, Loret and Dumoutier, 2019). The QMRA framework (Figure 5.1) was used to examine the utilisation of wastewater effluent for recreational and farming activities. Based on laboratory experiments and additional data obtained from literature, either UV or ozonation treatments would be used to avoid the microbiological risks associated with high bulking episodes (Figure 5.4). The study objectives are apparent; however some environmental aspects should be considered for future researches to enhance the QMRA outputs. For instance, either ethidium monoazide or propidium monoazide could be used to avoid the drawbacks of DNA-based methods, including the inability to discriminate between live and dead cells (Taylor *et al.*, 2014). Furthermore, the thresholds of risk (low, moderate, and high) associated with the environmental exposures to pathogens would be identified by quantitative tools using tracer testing and/or computational fluid dynamics (Petterson and Stenström, 2015). Moreover, despite the advantages of UV, the design and scale-up stages would hinder the widespread and consistent practice. The presence of particulate matters would reduce the efficiency of UV disinfection by shielding the pathogenic microorganisms from UV irradiation (Kollu and Ormeci, 2012). Hence, more researches on the measurements of the actual pathogenic reductions are required. Although ozone was suitable during the investigation, it is unstable in water and may react with natural organic matters to form biodegradable oxygenated by-products (Dong *et al.*, 2017). The optimum disinfectant dose for microbial inactivation should be identified to maintain the hygienic environment and economic benefits. Due to sludge bulking conditions in WWTPs, system operators, farmers, and consumers have to be sensitized about the potential occupational health threat during wastewaters reuse. Moreover, the population, especially in

developing countries, needs to be educated about hygiene regulations before consuming vegetables. The findings of this work would assist private and public sectors, as well as decision-maker, for the management of risks associated with effluent wastewater pathogens in the case of bulking, *i.e.*, particularly concerning farming activities and hygiene improvements.

5.4. CONCLUSIONS

In this study, a QMRA framework was applied, for the first time, to identify the applicability of treated wastewater for reuse in recreational practices for children, women, and men, and irrigation for farmers and vegetable consumers during sludge bulking in WWTPs. It was concluded that:

- During the sludge bulking events, the QMRA outputs showed unacceptable risk levels due to the high abundance of the pathogens, *viz.*, *E. coli* O157:H7, *Salmonella*, and *Mycobacterium*, in the WWTP effluent
- In all bulking scenarios (before disinfection options), the acceptable risk level of one case or less per 10,000 people per year was always exceeded.
- Children were more susceptible to bacterial infections approximately 2 to 3-fold higher than that for adults during swimming-related activities.
- Due to the risk estimates, the bacterial infection among children was 1.5 to 2.0-fold higher than the case of adults during swimming-related activities, and farmers would suffer from higher microbial risks due to irrigation with effluents of WWTPs experiencing moderate to high bulking scenarios.
- For the disinfection scenarios, the estimated risks of infection after chlorination were still exceeding the tolerable risk level for both *Salmonella* and *Mycobacterium*; UV treatment reduced the levels of risks below the tolerable threshold, except for the case of children exposed to *Salmonella* infection during swimming; for ozonation, the *Mycobacterium* infection resulted in higher risks than the allowable criterion of 10^{-4}.
- The outputs of the QMRA model could be employed to enhance the performance of the WWTPs under investigation and similar facilities in other countries; *i.e.*, and that will be the focus of our future work.
- Our future work will focus on enhancing the precision of QMRA outcomes using some environmental aspects pointed out in this study.

CHAPTER SIX: GENERAL CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

In this study, filamentous bulking incidence from seven WWTPs located in three provinces of South Africa were assessed. The filamentous bacteria identified in sludge bulking was determined using conventional microscopic methods, and confirmed using FISH. The dominant filamentous bacteria in descending order of frequency detected included: Type 0092, Type 1851, Type 1701, Type 021N, *M. parvicella*, Type 0041, *Thiothrix* spp., *N.limicola*, *Gordonia* spp. and *S. natans*. The ANN model was developed to predict the SVI by identifying influential predictors of the bulking conditions. Based on a filament index scale from 1 (None filament) to 7 (Excessive filament), the developed ANN predicted SVI data using the abundances of ten inputs of filamentous species. Eikelboom Type 0041 attained the highest impact on SVI, followed by *Gordonia* spp., *N. limicola*, and *Thiothrix* spp. The use of PCA was employed to correlate the main features of the WWTP with filamentous bacteria, which showed that efficient aeration system (*i.e.*, fine bubble, diffused aeration) in the three-stage Phoredox process improved the settling properties of flocs.

This study further attempted to model sludge bulking which further determined control strategies of the dominant filamentous bacteria. Two full-scale wastewater treatment plants which operated as a 3-stage Phoredox process were evaluated for this approach. Principal component analysis and a Decision Tree model were developed to describe the correlations and control strategies between influent wastewater characteristics and operational conditions (as inputs) and SVI as an output. The model input factors were pH, temperature, DO, SRT, F/M ratio, sCOD, tCOD, NH_4^+-N, TKN, PO_4^3-P, TP, and TSS. The classification tree model was used to determine factors that affected the proliferation of the prevalent filamentous microorganisms such as *M. parvicella*, *Thiothrix* I & II, and Eikelboom Types 0041, 0092, and 021N. It was found that SVI increased with increasing sludge retention time, but negatively correlated with sCOD and ammonium-nitrogen. *M. parvicella* had increased with a decline in temperature below 15.5°C, thus causing an increase in SVI. Abundant *Thiothrix* spp. was found to be linked with unbalanced ratio between readily biodegradable COD and nutrients. The filament Type 0092 had contributed to high SVI and prevailed with a decrease in F/M ratio below 0.08 1/d. Based on the acceptable training, validation, and generalisation procedures, the proposed models could be applied for the prediction of sludge bulking episodes. However, accurate quantification methods of filamentous

bacteria would provide better inputs for Decision Trees of the different identified bacteria. The outputs from this study would provide the scientific basis to control problematic filamentous bacteria in other full-scale WWTPs located in similar environmental conditions to South Africa. The use of predictive models has the potential to provide effective controls for the operation of WWTPs experiencing frequent bulking problems.

For the last objective, the potential health risks associated with filamentous bulking on the receiving environment, were evaluated using a QMRA model. Three common waterborne pathogens, *viz*., *E. coli* O157:H7, *Salmonella*, and *Mycobacterium*, were identified and quantified using the qPCR method. The pathogens where assessed from four full-scale WWTPs which had experienced sludge bulking events. The detected pathogens in this study, were incorporated into a QMRA model, to investigate the safety and applicability of treated WWTP effluent for recreational activities and agricultural activities. The QMRA model indicated the risk exposures for children, women, and men during recreational activities and for farmers and vegetable consumers during irrigation practices. Bacterial abundance in the treated WWTPs increased with the response to SVIs, and the QMRA risk values during all bulking events which exceeded the tolerable risk. The application of QMRA can be useful in private and public sectors, assisting decision-makers, for the management of risks that are associated with effluent wastewater pathogens; *i.e.*, particularly concerning water reuse for irrigation practices. Further investigations of the QMRA model can be used to improve the performance of the WWTPs by selecting the optimum disinfection scenario such as UV, Ozonation and chlorine treatment. The estimated risks for *Mycobacterium* and *Salmonella*, found to be less than tolerable risk levels (10^{-4}), provided by UV and ozonation disinfection processes.

The outputs of this study are of growing importance to produce feasible solutions and ease of control for the operation of WWTPs experiencing frequent bulking problems.

6.2. RECOMMENDATIONS

- Molecular based techniques such as qPCR and genome sequencing for the identification of filamentous bacterial populations are further recommended.

- Monitoring bulking using rheology meters can be invaluable information for sludge settling properties, also in this way all significant variables are monitored in real-time, thus, creating a knowledge-based hybrid supervisory systems. For better prediction and control simulations strategies, future variables most influential to filamentous bacterial growth and sludge settling should be included in integrated modeled systems.

- Other factors apart from F/M, sludge age CODs and seasonal effects investigated in this study can predict bulking. Therefore a further investigation is needed for improved prediction and control strategies, evaluating other factors such as C/N ratios that can influence filamentous bulking.

- Significant effort had been emphasized on non-linear regression models, also, leading to decision support systems and modified artificial neural networks for the prediction and control of sludge settling problems in complex biological systems. Few research studies focus on the modelling approach in full-scale WWTPs. Improved monitoring strategies and developed models are needed to evaluate the sludge bulking phenomenon.

- Future prediction and control strategies on operational problems in wastewater systems, are directed towards empirical models in addition to mechanistic models and heuristic knowledge. The out-put data are then integrated into simple decision support systems, whereby risk assessment models can be monitored, signifying valuable information on prediction and control of operational incidences

REFERENCES

National Water Act, 1999. Discharge limits and conditions set out in the National Water Act, Government Gazette No. 20526. *In:* AFFAIRS, D. O. W. (ed.). South Africa.

United Nations, 2018. The Sustainable Development Goals Report 2018. *In:* NATIONS, U. (ed.). New York, US.

ABIA, A. L. K., UBOMBA-JASWA, E., GENTHE, B. & MOMBA, M. N. 2016. Quantitative microbial risk assessment (QMRA) shows increased public health risk associated with exposure to river water under conditions of riverbed sediment resuspension. *Science of The Total Environment,* **566**(1)**:** 1143-51.

ABIODUN, O. I., JANTAN, A., OMOLARA, A. E., DADA, K. V., MOHAMED, N. A. & ARSHAD, H. 2018. State-of-the-art in artificial neural network applications: A survey. *Heliyon,* **4**(11)**:** 1-41.

ABUSAM, A., MYDLARCZYK, A., AL-SALAMEEN, F. & AHMED, M. I. 2017. Identification of the most probable causes for filamentous bacteria over-proliferation in Riqqa wastewater treatment plant, Kuwait. *Desalination and Water Treatment,* **72**(1)**:** 78-84.

ADEGOKE, A. A., AMOAH, I. D., STENSTRÖM, T. A., VERBYLA, M. E. & MIHELCIC, J. R. 2018. Epidemiological evidence and health risks associated with agricultural reuse of partially treated and untreated wastewater: A review. *Frontiers in Public Health,* **6**(337)**:** 1-20.

ADEWUMI, J. R., ILEMOBADE, A. A. & VAN ZYL, J. E. 2010. Treated wastewater reuse in South Africa: Overview, potential and challenges. *Resources, Conservation and Recycling,* **55**(2)**:** 221-231.

ADEYEMO, F. E., SINGH, G., REDDY, P., BUX, F. & STENSTROM, T. A. 2019. Efficiency of chlorine and UV in the inactivation of *Cryptosporidium* and *Giardia* in wastewater. *PLoS One,* **14**(5)**:** 216-40.

ADONADAGA, M. 2016. Reliability of morphological approach for bulking filamentous bacteria identification in activated sludge plants. *International Journal of Applied Science and Technology,* **6**(2)**:** 30-7.

AGRESTI, A. 2007 An Introduction to analysis. *In:* AGRESTI, A. (ed.) *Multicategory logit models.* 2nd ed. New Jersey: John Wiley and Sons, Inc.

ALBERTSON, O. E. 1991. Bulking sludge control–progress, practice and problems. *Water Science and Technology,* **23**(4)**:** 835-46.

ALI, J., KHAN, R., AHMAD, N. & MAQSOOD, I. 2012. Random forests and Decision Trees. *International Journal of Computer Science Issues,* **9**(3)**:** 272-78.

ALSINA, X. F., COMAS, J., RODA, I. R., GERNAEY, K. V. & ROSEN, C. 2009. Including the effects of filamentous bulking sludge during the simulation of wastewater treatment plants using a risk assessment model. *Water Research,* **43**(1)**:** 4527–38.

AMANATIDOU, E., SAMIOTIS, G., TRIKOILIDOU, E., TZELIOS, D. & MICHAILIDIS, A. 2016. Influence of wastewater treatment plants' operational conditions on activated sludge microbiological and morphological characteristics. *Environmental Technology,* **37**(2)**:** 265-78.

AMAND, L. 2010. Control of aeration systems in activated sludge processes – a review. Uppsala University.: IVL Swedish Environmental Research Institute.

AMANN, R. I., LUDWIG, W. & K.H., S. 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews,* **59**(1)**:** 143–69.

AMENU, K., SHITU, D. & ABERA, M. 2016. Microbial contamination of water intended for milk container washing in smallholder dairy farming and milk retailing houses in Southern Ethiopia. *SpringerPlus,* **5**(1195)**:** 1-6.

AMHA, Y. M., ANWAR, M. Z., KUMARASWAMY, R., HENSCHEL, A. & AHMAD, F. 2017. *Mycobacteria* in municipal wastewater treatment and reuse: microbial diversity for screening the occurrence of clinically and environmentally relevant species in arid regions. *Environmental Science and Technology,* **51**(5)**:** 3048-56.

AMHA, Y. M., KUMARASWAMY, R. & AHMAD, F. 2015. A probabilistic QMRA of *Salmonella* in direct agricultural reuse of treated municipal wastewater. *Water Science and Technology,* **71**(8)**:** 1203-11.

AMOAH, I. D., ADEGOKE, A. A. & STENSTRÖM, T. A. 2018a. Soil-transmitted helminth infections associated with wastewater and sludge reuse: a review of current evidence. *Tropical Medicine and International Health,* **23**(7)**:** 692-703.

AMOAH, I. D., REDDY, P., SEIDU, R. & STENSTRÖM, T. A. 2018b. Concentration of soil-transmitted helminth eggs in sludge from South Africa and Senegal: A probabilistic estimation of infection risks associated with agricultural application. *Journal of Environmental Management,* **206**(1)**:** 1020-27.

AMOAH, I. D., REDDY, P., SEIDU, R. & STENSTRÖM, T. A. 2018c. Removal of helminth eggs by centralized and decentralized wastewater treatment plants in South Africa and Lesotho: Health implications for direct and indirect exposure to the effluents. *Environmental Science and Pollution Research,* **25**(13)**:** 12883-95.

ANDERSEN, M. H., MCILROY, S. J., NIERYCHLO, M., NIELSEN, P. H. & ALBERTSEN, M. 2019. Genomic insights into *Candidatus Amarolinea aalborgensis* gen. nov., sp. nov., associated with settleability problems in wastewater treatment plants. *Systematic and Applied Microbiology,* **42**(1)**:** 77-84.

APHA 1998. *Standard methods for the examination of water and wastewater,* Washington, D.C.

ARDERN, E. & LOCKETT, W. T. 1914. Experiments on the oxidation of sewage without the aid of filters. *Journal of the Society of Chemical Industry,* **33**(10)**:** 523-39.

ARUGA, S., KAMAGATA, Y., KOHNO, T., HANADA, S., NAKAMURA, K. & KANAGAWA, T. 2002. Characterization of filamentous Eikelboom type 021N bacteria and description of *Thiothrix disciformis* sp. nov. and *Thiothrix flexilis* sp. nov. *International journal of systematic and evolutionary microbiology,* **52**(1)**:** 1309-16.

ASVAPATHANAGUL, P., OLSON, B. H., GEDALANGA, P. B., HASHEMI, A., HUANG, Z. & LA, J. 2015. Identification and quantification of *Thiothrix eikelboomii* using qPCR for early detection of bulking incidents in a full-scale water reclamation plant. *Applied Microbiology and Biotechnology,* **99**(9)**:** 4045-57.

AWOLUSI, O. O., NASR, M., KUMARI, S. & BUX, F. 2018. Principal component analysis for interaction of nitrifiers and wastewater environments at a full-scale activated sludge plant. *International Journal of Environmental Science and Technology,* **15**(7)**:** 1477-90.

AYAZ, N. D., GENCAY, Y. E. & EROL, I. 2014. Prevalence and molecular characterization of sorbitol fermenting and non-fermenting *Escherichia coli* O157:H7(+)/H7(-) isolated from cattle at slaughterhouse and slaughterhouse wastewater. *International Journal of Food Microbiology,* **174**(1)**:** 31-8.

AZIMI, A. A. & ZAMANZADEH, M. 2006. The effect of selectors and reactor configuration on filamentous sludge bulking control in activated sludge. *Pakistan Journal of Biological Sciences,* **9**(3)**:** 345-49.

BAGHERI, M., MIRBAGHERI, S. A., BAGHERI, Z. & KAMARKHANI, A. M. 2015. Modeling and optimization of activated sludge bulking for a real wastewater treatment plant using hybrid artificial neural networks-genetic algorithm approach. *Process Safety and Environmental Protection,* **95**(1)**:** 12-25.

BALKHAIR, K. S. 2016. Microbial contamination of vegetable crop and soil profile in arid regions under controlled application of domestic wastewater. *Saudi Journal of Biological Sciences,* **23**(1)**:** 83-92.

BANADDA, E. N., SMETS, I. Y., JENNE, R. & VAN IMPE, J. F. 2005. Predicting the onset of filamentous bulking in biological wastewater treatment systems by exploiting image analysis information. *Bioprocess and Biosystems Engineering,* **27**(5)**:** 339-48.

BANADDA, N., NHAPI, I. & KIMWAGA, R. 2011. A review of modeling approaches in activated sludge systems. *African Journal of Environmental Science and Technology* **5**(6)**:** 397-408.

BARAK, J. D., SANANIKONE, K. & DELWICHE, M. J. 2005. Comparison of primers for the detection of pathogenic *Escherichia coli* using real-time PCR. *Letters in Applied Microbiology,* **41**(2)**:** 112-8.

BARNARD, J. L. 1983. Background to biological phosphorus removal. *Water Science and Technology,* **15**(1)**:** 1-13.

BARNARD, J. L. & COMEAU, Y. 2014. Macro-nutrient removal (phosphorus). *In:* JENKINS, D. & WANNER, J. (eds.) *Activated Sludge - 100 Years and Counting.* London: IWA publishing.

BEAUDEQUIN, D., HARDEN, F., ROIKO, A. & MENGERSEN, K. 2016. Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. *Science of the Total Environment,* **541**(1)**:** 1393-1409.

BEER, M., SEVIOUR, E. M., KONG, Y., CUNNINGHAM, M., BLACKALL, L. L. & SEVIOUR, R. J. 2002. Phylogeny of the filamentous bacterium Eikelboom Type 1851, and design and application of a 16S rRNA targeted oligonucleotide probe for its fluorescence *in situ* identification in activated sludge. *FEMS Microbiology Letters,* **207**(2)**:** 179-83.

BELANCHE, L., VALDES, J. J., COMAS, J., RODA, I. R. & POCH, M. 2000. Prediction of the bulking phenomenon in wastewater treatment plants. *Artificial Intelligence in Engeneering,* **14**(1)**:** 307-17.

BESHA, A. T., GEBREYOHANNES, A. Y., TUFA, R. A., BEKELE, D. N., CURCIO, E. & GIORNO, L. 2017. Removal of emerging micropollutants by activated sludge process and membrane bioreactors and the effects of micropollutants on membrane fouling: A review. *Journal of Environmental Chemical Engineering,* **5**(3)**:** 2395-414.

BLACKALL, L. L., STRATTON, H., BRADFORD, D., DOT, T. D., SJORUP, C., SEVIOUR, E. M. & SEVIOUR, R. J. 1996. "*Candidatus Microthrix parvicella*", a filamentous bacterium from activated sewage treatment plants. *International Journal of Systematic Bacteriology,* **46**(1)**:** 233-46.

BLACKBEARD, J., EKAMA, G. & MARAIS, G. V. R. 1986. A survey of filamentous bulking and foaming in activated sludge plants in South Africa. *Water Pollution and Control,* **85**(1)**:** 90-100.

BLACKBEARD, J. R., GABB, D. M. D., EKAMA, G. A. & MARAIS, G. V. R. 1988. Identification of filamentous organisms in nutrient removal activated sludge plants in South Africa. *Water SA,* **14**(1)**:** 29-33.

BOKULICH, N. A., BAMFORTH, C. W. & MILLS, D. A. 2012. A review of molecular methods for microbial community profiling of beer and wine. *Journal of the American Society of Brewing Chemists,* **70**(3)**:** 150-62.

BORETTI, A. & ROSA, L. 2019. Reassessing the projections of the World Water Development Report. *Clean Water,* **2**(15).

BOZTOPRAK, H., ÖZBAY, Y., GÜÇLÜ, D. & KÜÇÜKHEMEK, M. 2015. Prediction of sludge volume index bulking using image analysis and neural network at a full-scale activated sludge plant. *Desalination and Water Treatment,* **57**(37)**:** 17195-205.

BREIMAN, L., FRIEDMAN, J., OLSHEN, R. & STONE, C. 1984. *Classification and Regression Trees,* California, USA: Boca Raton, FL: CRC Press.

BROSCH, R., GORDON, S. V., MARMIESSE, M., BRODIN, P., BUCHRIESER, C., EIGLMEIER, K., GARNIER, T., GUTIERREZ, C., HEWINSON, G., KREMER, K., PARSONS, L. M., PYM, A. S., SAMPER, S., VAN SOOLINGEN, D. & COLE, S. T. 2002. A new evolutionary scenario for the *Mycobacterium* tuberculosis complex. *Proceedings of the National Academy of Sciences of the United States of America,* **99**(6)**:** 3684-9.

BURGER, W., KRYSIAK-BALTYN, K., SCALES, P. J., MARTIN, G. J. O., STICKLAND, A. D. & GRAS, S. L. 2017. The influence of protruding filamentous bacteria on floc stability and solid-liquid separation in the activated sludge process. *Water Research,* **123**(1)**:** 578-85.

BUX, F. & KASAN, H. C. 1994. A microbiological survey of ten activated sludge plants. *Water SA,* **20**(1)**:** 61-72.

CAI, L. & ZHANG, T. 2013. Detecting human bacterial pathogens in wastewater treatment plants by a high-throughput shotgun sequencing technique. *Environmental Science and Technology* **47**(10)**:** 5433-41.

CAO, C. & LOU, I. 2015. Analysis of environmental variables on population dynamic change of *Haliscomenobacter hydrossis*, the bulking causative filament in Macau wastewater treatment plant. *Desalination and Water Treatment,* **57**(1)**:** 1-14.

CAO, C., LOU, I., HUANG, C. & LEE, M.-Y. 2016. Metagenomic sequencing of activated sludge filamentous bacteria community using the Ion Torrent platform. *Desalination and Water Treatment,* **57**(5)**:** 2175-83.

CARAVELLI, A., GIANNUZZI, L. & ZARITZKY, N. 2006. Effectiveness of chlorination and ozonation methods on pure cultures of floc-forming microorganisms and activated sludge: A comparative study. *Water SA,* **32**(4)**:** 585-96.

CARR, E. L., EALES, K., SODDELL, J. & SEVIOUR, R. J. 2005. Improved permeabilization protocols for fluorescence *in situ* hybridization (FISH) of mycolic-acid-containing bacteria found in foams. *Journal of Microbiological Methods,* **61**(1)**:** 47-54.

CASEY, T. G., EKAMA, G. A., WENTZEL, M. C. & MARAIS, G. V. R. 1995. Filamentous organism bulking in nutrient removal activated sludge systems. Paper 1: An historical overview of causes and control. *Water SA,* **21**(3)**:** 321-38.

CASEY, T. G., WENTZEL, M. C., EKAMA, G. A. & MARAIS, G. V. R. 1994a. Causes and control of low F/M filamentous bulking in long sludge age nutrient removal activated sludge systems. WRC Report No. 286/2/93. South Africa: Water Research Commission.

CASEY, T. G., WENTZEL, M. C., EKAMA, G. A. & MARAIS, G. V. R. 1994b. Development and Evaluation of Specific Control Methods for Ameliorating Low F/M Filament Bulking. WRC Report No. 286/1/93. South Africa: Water Research Commission.

CENENS, C., SMETS, I. Y., RYCKAERT, V. G. & VAN IMPE, J. F. 2000. Modeling the competition between floc- forming and filamentous bacteria in activated sludge waste water treatment-annotated. *Water Research,* **34**(9)**:** 2525-34.

CHING, T., HIMMELSTEIN, D. S., BEAULIEU-JONES, B. K., KALININ, A. A., DO, B. T., WAY, G. P., FERRERO, E., AGAPOW, P.-M., ZIETZ, M., HOFFMAN, M. M., XIE, W., ROSEN, G. L., LENGERICH, B. J., ISRAELI, J., LANCHANTIN, J., WOLOSZYNEK, S., CARPENTER, A. E., SHRIKUMAR, A., XU, J., COFER, E. M., LAVENDER, C. A., TURAGA, S. C., ALEXANDARI, A. M., LU, Z., HARRIS, D. J., DECAPRIO, D., QI, Y., KUNDAJE, A., PENG, Y., WILEY, L. K., SEGLER, M. H. S., BOCA, S. M., SWAMIDASS, S. J., HUANG, A., GITTER, A. & GREENE, C. S. 2018. Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society, Interface,* **15**(141)**:** 1-47.

CHMIELOWSKI, K., CZEKAŃSKI, A. & LEŚNIAŃSKA, A. 2019. Using data mining to predict sludge and filamentous microorganism sedimentation. *Polish Journal of Environmental Studies,* **28**(5)**:** 3105-13.

CHUDOBA, J., GRAU, P. & OTTOVÁ, V. 1973. Control of activated-sludge filamentous bulking–II. Selection of microorganisms by means of a selector. *Water Research,* **7**(10)**:** 1389-98.

CHUN, T. S., MALEK, M. A. & ISMAIL, A. R. 2017. A Review of Wastewater Treatment Plant Modelling: Revolution on Modelling Technology. *American Journal of Environmental and Resource Economics,* **2**(1)**:** 22-26

COLLIVIGNARELLI, M., ABBÀ, A., BENIGNA, I., SORLINI, S. & TORRETTA, V. 2017. Overview of the main disinfection processes for wastewater and drinking water treatment plants. *Sustainability,* **10**(2)**:** 1-21.

COMAS, J., RODA, I. R., GERNAEY, K. V., ROSEN, C., JEPPSSON, U. & POCH, M. 2008. Risk assessment modelling of microbiology-related solids separation problems in activated sludge systems. *Environmental Modelling and Software,* **23**(10)**:** 1250-61.

COSKUNER, G. 2002. A new molecular technique for the identification of microorganisms in biological treatment plants: Fluorescent *in situ* hybridization. *Turkish Journal of Biology,* **26**(1)**:** 57-63.

COX, D. R. & D.V., H. 1974. Interval Estimation *In:* HALL, C. A. (ed.) *Theoretical Statisitics.* 1 ed. Boca Raton, Florida CRC Press.

CUI, Q., FANG, T., HUANG, Y., DONG, P. & WANG, H. 2017. Evaluation of bacterial pathogen diversity, abundance and health risks in urban recreational water by amplicon next-generation sequencing and quantitative PCR. *Journal of Environmental Sciences,* **57**(1)**:** 137-49.

CYDZIK-KWIATKOWSKA, A. & ZIELINSKA, M. 2016. Bacterial communities in full-scale wastewater treatment systems. *World Journal of Microbiology and Biotechnology,* **32**(4)**:** 1-8.

CYDZIK-KWIATKOWSKA, A., ZIELINSKA, M. & WOJNOWSKA-BARYLA, I. 2012. Impact of operational parameters on bacterial community in a full-scale municipal wastewater treatment plant. *Polish Journal of Microbiology,* **61**(1)**:** 41-9.

DA MOTTA, M., PONS, M. N. & ROCHE, N. 2003. Monitoring filamentous bulking in activated sludge systems fed by synthetic or municipal wastewater. *Bioprocess and Biosystems Engineering,* **25**(6)**:** 387-93.

DAIMS, H., BRUHL, A., AMANN, R., SCHLEIFER, K. H. & WAGNER, M. 1999. The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: Development and evaluation of a more comprehensive probe set. *Systematic and Applied Microbiology,* **22**(3)**:** 434-44.

DAVENPORT, R. J., CURTIS, T. P., GOODFELLOW, M., STAINSBY, F. M. & BINGLEY, M. 2000. Quantitative use of fluorescent *in situ* hybridization to examine relationships between mycolic acid-containing actinomycetes and foaming in activated sludge plants. *Applied and Environmental Microbiology,* **66**(3)**:** 1158-66.

DAVIES, P. S. 2005. The Biological Basis of Wastewater Treatment. *Strathkelvin Instruments Ltd*.

DE CONINCK, R., MAGNUSSON, F., AKESSON, J. & HELSEN, L. 2015. Toolbox for development and validation of grey-box building models for forecasting and control. *Journal of Building Performance Simulation,* **9**(3)**:** 288-303.

DE LOS REYES, F. L., RITTER, W. & RASKIN, L. 1997. Group-specific small-subunit rRNA hybridization probes to characterize filamentous foaming in activated sludge systems. *Applied and Environmental Microbiology,* **63**(3)**:** 1107-17.

DE LOS REYES, M. F., DE LOS REYES, F. L., HERNANDEZ, M. & RASKIN, L. 1998. Quantification of *Gordona amarae* strains in foaming activated sludge and anaerobic digester systems with oligonucleotide hybridization probes. *Applied and Environmental Microbiology,* **64**(7)**:** 2503-12.

DEEPNARAIN, N., KUMARI, S., RAMJITH, J., SWALAHA, F. M., TANDOI, V., PILLAY, K. & BUX, F. 2015. A logistic model for the remediation of filamentous bulking in a biological nutrient removal wastewater treatment. *Water Science and Technology,* **72**(3)**:** 391-404.

DEEPNARAIN, N., NASR, M., KUMARI, S., STENSTRÖM, T. A., REDDY, P., PILLAY, K. & BUX, F. 2019. Decision Tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant. *Process Safety and Environmental Protection,* **126**(1)**:** 25-34.

DIAS, E., JAMES, E. & TAYLOR, H. 2019. Estimating the concentration of viral pathogens and indicator organisms in the final effluent of wastewater treatment processes using stochastic modelling. *Microbial Risk Analysis,* **11**(1)**:** 47-56.

DIAS, P. A., DUNKEL, T., FAJADO, D. A. S., GALLEGOS, E. D. L., DENECKE, M., WIEDEMANN, P., SCHNEIDER, F. K. & SUHR, H. 2016. Image processing for identification and quantification of filamentous bacteria in *in situ* acquired images. *BioMedical Engineering OnLine,* **15**(1)**:** 64.

DICKIN, S. K., SCHUSTER-WALLACE, C. J., QADIR, M. & PIZZACALLA, K. 2016. A review of health risks and pathways for exposure to wastewater use in agriculture. *Environmental Health Perspectives,* **124**(7)**:** 900-9.

DIEHL, S. & FARAS, S. 2013. Control of an ideal activated sludge process in wastewater treatment via an ODE–PDE model. *Journal of Process Control,* **23**(3)**:** 359-81.

DONG, S., LI, J., KIM, M.-H., PARK, S.-J., EDEN, J. G., GUEST, J. S. & NGUYEN, T. H. 2017. Human health trade-offs in the disinfection of wastewater for landscape irrigation: Microplasma ozonation vs. chlorination. *Environmental Science: Water Research and Technology,* **3**(1)**:** 106-18.

DOS SANTOS, L. A., FERREIRA, V., NETO, M. M., PEREIRAM, M. A., MOTA, M. & NICOLAU, A. 2015. Study of 16 Portuguese activated sludge systems based on filamentous bacteria populations and their relationships with environmental parameters. *Applied Microbiology and Biotechnology,* **99:** 5307–16.

DUNKEL, T., DE LEON, E., BOCK, C., LANGE, A., HOFFMANN, D., BOENIGK, J. & DENECKE, M. 2017. Illumina sequencing for the identification of filamentous bulking and foaming bacteria in industrial activated sludge plants. *International Journal of Environmental Science and Technology,* **15**(1)**:** 1139-58.

DUNKEL, T., DE LEON GALLEGOS, E. L., SCHONSEE, C. D., HESSE, T., JOCHMANN, M., WINGENDER, J. & DENECKE, M. 2016. Evaluating the influence of wastewater composition on the growth of *Microthrix parvicella* by GCxGC/qMS and real-time PCR. *Water Research,* **88**(1)**:** 510-23.

DWA 1998. National Water Act No. 36, 1998. Republic of South Africa: DWA (Department of Water Affairs) Government Gazette Staatskoerant.

DWA 2012a. 2012/13 Annual Report. South Africa: Department: Water Affairs.

DWA 2012b. Green Drop Progress Report, National Overview. South Africa: Department of Water Affairs.

EBRAHIMI, M., BADALIANS GHOLIKANDI, G., JAMSHIDI, S. & EZZO, H. 2016. Dolomite reactor, a retrofitting approach for activated sludge against bulking. *Iranian Journal of Science and Technology, Transactions A: Science,* **42**(3)**:** 1215-21.

EDOKPAYI, J. N., ODIYO, J. O. & DUROWOJU, O. S. 2017. Impact of wastewater on surface water quality in developing countries: A case study of south africa. *In:* TUTU, H. (ed.) *Water Quality* Inteck.

EIKELBOOM, D. H. 2000. *Process Control of Activated Sludge Plants by Microscopic Investigation* London, IWA Publishing.

EIKELBOOM, D. H., ANDREADAKIS, A. & ANDREASEN, K. 1998. Survey of filamentous populations in nutrient removal plants in four European countries. *Water Science and Technology,* **37**(4-5)**:** 281-289.

EKAMA, G. A. 2010. The role and control of sludge age in biological nutrient removal activated sludge systems. *Water Science and Technology,* **61**(7)**:** 1645-52.

EKAMA, G. A. 2015. Review: Recent developments in biological nutrient removal. *Water SA,* **41**(4)**:** 515-24.

EKWANZALA, M. D., DEWAR, J. B., KAMIKA, I. & MOMBA, M. N. B. 2018. Systematic review in South Africa reveals antibiotic resistance genes shared between clinical and environmental settings. *Infection and Drug Resistance,* **11**(1)**:** 1907-20.

ENITAN, A. M., KUMARI, S., SWALAHA, F. M., ADEYEMO, J., RAMDHANI, N. & BUX, F. 2014. Kinetic modelling and characterization of microbial community present in a full-scale UASB reactor treating brewery effluent. *Microbial Ecology,* **67**(2)**:** 358-68.

EPA 1999. Ozone Disinfection. *In:* AGENCY, U. S. E. P. (ed.) *CX824652.* Cincinnati, Ohio.

ERHART, R., BRADFORD, D., SEVIOUR, R. J., AMANN, R. & BLACKALL, L. L. 1997. Development and use of fluorescent *in situ* hybridization probes for the detection and identification of "*Microthrix parvicella*" in activated sludge. *Systematic and Applied Microbiology,* **20**(2)**:** 310-18.

ERIKSSON, L., STEEN, I. & TENDAJ, M. 1992. Evaluation of sludge properties at an activated sludge plant. *Water Science and Technology,* **25**(6)**:** 251-65.

ESPIGARES, E., BUENO, A., ESPIGARES, M. & GALVEZ, R. 2006. Isolation of *Salmonella* serotypes in wastewater and effluent: Effect of treatment and potential risk. *International Journal of Hygiene and Environmental Health,* **209**(1)**:** 103-7.

ETTERER, T. J. 2006. *Formation, Structure and Function of Aerobic Granular Sludge.* Doctor of Natural Sciences, Technical University of Munich

FALKINHAM, J. O. 2009. Surrounded by *Mycobacteria*: Nontuberculous *Mycobacteria* in the human environment. *Journal of Applied Microbiology,* **107**(2)**:** 356-67.

FAN, H., LIU, X., WANG, H., HAN, Y., QI, L. & WANG, H. 2017a. Oxygen transfer dynamics and activated sludge floc structure under different sludge retention times at low dissolved oxygen concentrations. *Chemosphere,* **169**(1)**:** 586-95.

FAN, N., QI, R., HUANG, B., JIN, R. & YANG, M. 2020. Factors influencing *Candidatus Microthrix parvicella* growth and specific filamentous bulking control: A review. *Chemosphere,* **244**(1)**:** 1-13.

FAN, N., QI, R., ROSSETTI, S., TANDOI, V., GAO, Y. & YANG, M. 2017b. Factors affecting the growth of *Microthrix parvicella*: Batch tests using bulking sludge as seed sludge. *The Science of the Total Environment,* **609**(1)**:** 1192-9.

FAN, N., WANG, R., QI, R., GAO, Y., ROSSETTI, S., TANDOI, V. & YANG, M. 2018. Control strategy for filamentous sludge bulking: Bench-scale test and full-scale application. *Chemosphere,* **210**(1)**:** 709-16.

FAO 2011. The state of the world's land and water resources for food and agriculture (SOLAW)—managing systems at risk. New York: Food and Agriculture. The Food and Agriculture Organization of the United Nations and Earthscan.

FAWZY, M., NASR, M., NAGY, H. & HELMI, S. 2018. Artificial intelligence and regression analysis for Cd(II) ion biosorption from aqueous solution by gossypium barbadense waste. *Environmental Science and Pollution Research International,* **25**(6)**:** 5875-88.

FIGUEROA, M., VAL DEL RÍO, Á., CAMPOS, J., MENDEZ, R. & MOSQUERA-CORRAL, A. 2015. Filamentous bacteria existence in aerobic granular reactors. *Bioprocess and biosystems engineering,* **38**(5)**:** 841-51.

FLORES-ALSINA, X., COMAS, J., RODRIGUEZ-RODA, I., GERNAEY, K. V. & ROSEN, C. 2009. Including the effects of filamentous bulking sludge during the simulation of wastewater treatment plants using a risk assessment model. *Water Research,* **43**(18)**:** 4527-38.

GAR ALALM, M. & NASR, M. 2018. Artificial intelligence, regression model, and cost estimation for removal of chlorothalonil pesticide by activated carbon prepared from casuarina charcoal. *Sustainable Environment Research,* **28**(3)**:** 101-10.

GARSON, G. D. 1991. Interpreting neural-network connection weights. *AI Expert,* **6**(4)**:** 46–51.

GENTHE, B., LE ROUX, W. J., SCHACHTSCHNEIDER, K., OBERHOLSTER, P. J., ANECK-HAHN, N. H. & CHAMIER, J. 2013. Health risk implications from simultaneous exposure to multiple environmental contaminants. *Ecotoxicology and Environmental Safety,* **93**(1)**:** 171-9.

GERARDI, M. H. 2006. Floc-forming bacteria. *In:* GERARDI, M. H. (ed.) *Wastewater microbiology.* Canada: John Wiley and Sons.

GERNAEY, K. V., VAN LOOSDRECHT, M. C., HENZE, M., LIND, M. & JORGENSEN, B. 2004. Activated sludge wastewater treatment plant modelling and simulation state of the art. *Environmental Modelling and Software,* **19**(1)**:** 773-83.

GICH, F., GARCIA-GIL, J. & OVERMANN, J. 2001. Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. *Archives of Microbiology,* **177**(1)**:** 1-10.

GOVOREANU, R., SEGHERS, D., NOPENS, I., DE CLERCQ, B., SAVEYN, H., CAPALOZZA, C., VAN DER MEEREN, P., VERSTRAETE, W., TOP, E. & VANROLLEGHEM, P. 2003. Linking floc structure and settling properties to activated sludge population dynamics in an SBR. *Water Science and Technology,* **47**(12)**:** 9-18.

GRADY, C. P. L., DAIGGER, G. T. & LIM, H. C. 1999. *Biological Wastewater Treatment,* New York, USA., Marcel Dekker, Inc.

GRADY, C. P. L., DAIGGER, G. T., NANCY, G. L. & FILIPE, C. D. M. 2011. *Biological Wastewater Treatment* UK, IWA.

GRAVELEAU, L., COTTEUX, E. & DUCHÈNE, P. 2005. Bulking and foaming in France: The 1999–2001 survey. *Acta hydrochimica et Hydrobiologica,* **33**(1)**:** 223-31.

GUO, F. & ZHANG, T. 2012. Profiling bulking and foaming bacteria in activated sludge by high throughput sequencing. *Water Research,* **46**(8)**:** 2772-82.

GUO, J., PENG, Y., WANG, Z., YUAN, Z., YANG, X. & WANG, S. 2012. Control filamentous bulking caused by chlorine-resistant Type 021N bacteria through adding a biocide CTAB. *Water Research,* **46**(19)**:** 6531-42.

HAAS, C., ROSE, J. & GERBA, C. 2014. *Quantitative microbial risk assessment* New York Jhon Wiley and Sons.

HAMDY, A., MOSTAFA, M. & NASR, M. 2018. Regression analysis and artificial intelligence for removal of methylene blue from aqueous solutions using nanoscale zero-valent iron. *International Journal of Environmental Science and Technology,* **16**(1)**:** 357-72.

HAMILTON, A. J., STAGNITTI, F., PREMIER, R., BOLAND, A. M. & HALE, G. 2006. Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. *Applied and Environmental Microbiology,* **72**(5)**:** 3284-90.

HAN, H., WU, X., GE, L. & QIAO, J. 2018a. A sludge volume index (SVI) model based on the multivariate local quadratic polynomial regression method. *Chinese Journal of Chemical Engineering,* **26**(5)**:** 1071-77.

HAN, H. G., LI, Y., GUO, Y. N. & QIAO, J. F. 2016. A soft computing method to predict sludge volume index based on a recurrent self-organizing neural network. *Applied Soft Computing,* **38**(1)**:** 477–86.

HAN, H. G., LIU, H. X., LIU, Z. & QIAO, J. F. 2019. Fault detection of sludge bulking using a self-organizing type-2 fuzzy-neural-network. *Control Engineering Practice,* **90**(1)**:** 27-37.

HAN, H. G., LIU, Z., GUO, Y. N. & QIAO, J. F. 2018b. An intelligent detection method for bulking sludge of wastewater treatment process. *Journal of Process Control,* **68:** 118-28.

HAN, H. G. & QIAO, J. 2013. Hierarchical neural network modeling approach to predict sludge volume index of wastewater treatment process. *IEEE Transactions on Control Systems Technology,* **21**(6)**:** 2423-31.

HAN, W., PENG, Z., LI, T., FAN, P. & YU, L. 2018c. Control of sludge settleability based on organic load and ammonia nitrogen load under low dissolved oxygen. *Water Science and Technology,* **78**(10)**:** 2113-18.

HARTMANN, L. 1999. Historical Development of Wastewater Treatment Processes. *In:* REHM, H. J., REED, G., PUHLER, A. & STADLER, P. (eds.) *Biotechnology: Environmental Processes I.* Second Edition ed. Germany: Wiley.

HAVELAAR, A. H. QMRA - A framework for assessing microbiological public health risks. Dutch National Institute for Public Health and the Environment (RIVM) and Utrecht University. The European College of Veterinary Public Health Annual Scientific Conference, Maastricht, 23 August 2012., 2012.

HENRIET, O., MEUNIER, C., HENRY, P. & MAHILLON, J. 2017. Filamentous bulking caused by *Thiothrix* species is efficiently controlled in full-scale wastewater treatment plants by implementing a sludge densification strategy. *Scientific Reports,* **7**(1)**:** 1430-40.

HENZE, M., GRADY, C., JR, G., W, , GR, M. & MATSUO, T. 1987. Activated Sludge Model No. 1, IAWPRC Scientific and Technical Reports 1, IAWPRC, London.

HENZE, M., VAN LOOSDRECHT, M. C., EKAMA, G. A. & BRDJANOVIC, D. 2008. *Biological wastewater treatment: Principles modelling and design,* London, IWA Publishing

HORAN, N. J., BU'ALI, A. M. & ECCLES, C. R. 1988. Isolation, identification and characterisation of filamentous and floc-forming bacteria from activated sludge flocs. *Environmental Technology Letters,* **9**(5)**:** 449-57.

HREIZ, R., LATIFI, M. A. & ROCHE, N. 2015. Optimal design and operation of activated sludge processes: State-of-the-art. *Chemical Engineering Journal,* **281**(1)**:** 900-20.

HU, Z., HOUWELING, D. & DOLD, P. 2012. Biological nutrient removal in municipal wastewater treatment: New directions in sustainability. *Journal of Environmental Engineering,* **138**(3)**:** 307-17.

HU, Z. R., SOTEMANN, S., MOODLEY, R., WENTZEL, M. C. & EKAMA, G. A. 2003. Experimental investigation of the external nitrification biological nutrient removal activated sludge (ENBNRAS) system. *Biotechnology and Bioengineering* **83**(3)**:** 260-73.

HUBER, D., VOITH VON VOITHENBERG, L. & KAIGALA, G. V. 2018. Fluorescence *in situ* hybridization (FISH): History, limitations and what to expect from micro-scale FISH? *Micro and Nano Engineering,* **1:** 15-24.

HUG, T., GUJER, W. & SIEGRIST, H. 2006. Modelling seasonal dynamics of *Microthrix parvicella*. *Water Science and Technology,* **54**(1)**:** 189-98.

HUGENHOLTZ, P., TYSON, G. W., WEBB, R. I., WAGNER, A. M. & BLACKALL, L. L. 2001. Investigation of candidate division TM7, a recently recognized major lineage of the domain bacteria with no known pure-culture representatives. *Applied and Environmental Microbiology,* **67:** 411-9.

INSEL, G., EROL, S. & OVEZ, S. 2014. Effect of simultaneous nitrification and denitrification on nitrogen removal performance and filamentous microorganism diversity of a full-scale MBR plant. *Bioprocess and Biosystems Engineering,* **37**(11)**:** 2163-73.

JENKINS, D., RICHARD, M. G. & DAIGGER, G. T. 2003. *Manual on the Causes and Control of Activated Sludge Bulking, Foaming, and Other Solids Separation Problems,* Michigan Lewis Publishers.

JEPPSSON, U. 1996. *Modelling Aspects of Wastewater Treatment Processes.* Doctor of Philosophy Lund Institute of Technology.

JIANG, X., GUO, F. & ZANG, T. 2015. Population dynamics of bulking and foaming bacteria in a full-scale wastewater treatment plant over five years. *Scientific Reports,* **6**(1)**:** 1-9.

JIN, B., WILÉN, B.-M. & LANT, P. 2003. A comprehensive insight into floc characteristics and their impact on compressibility and settleability of activated sludge. *Chemical Engineering Journal,* **95**(1)**:** 221-234.

JOLLIFFE, I. 2002. *Principal component analysis* Verlag New York, Springer.

JONES, P. & SCHULER, A. 2010. Seasonal variability of biomass density and activated sludge settleability in full-scale wastewater treatment systems. *Chemical Engineering Journal,* **164:** 16-22.

JYOTI, A., RAM, S., VAJPAYEE, P., SINGH, G., DWIVEDI, P. D., JAIN, S. K. & SHANKER, R. 2010. Contamination of surface and potable water in South Asia by *Salmonellae*: culture-independent quantification with molecular beacon real-time PCR. *Science of the Total Environment,* **408**(6)**:** 1256-63.

KAETZKE, A., JENTZSCH, D. & ESCHRICH, K. 2005. Quantification of *Microthrix parvicella* in activated sludge bacterial communities by real-time PCR. *Letters in Applied Microbiology,* **40**(3)**:** 207-11.

KAMINSKI, B., JAKUBCZYK, M. & SZUFEL, P. 2018. A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research,* **26**(1)**:** 135-159.

KANAGAWA, T., KAMAGATA, Y., ARUGA, S., KOHNO, T., HORN, M. & WAGNER, M. 2000. Phylogenetic analysis of oligonucleotide probe development for Eikelboom Type 021N filamentous bacteria isolated from bulking activated sludge. *Applied and Environmental Microbiology,* **66**(11)**:** 5043–52.

KARPINSKA, A. M. & BRIDGEMAN, J. 2016. CFD-aided modelling of activated sludge systems - A critical review. *Water Research,* **88**(1)**:** 861-79.

KELLER, G. H. & MANAK, M. M. 1989. *DNA Probes*, Macmillan Publishers Ltd.

KHOSRAVI, K., PHAM, B. T., CHAPI, K., SHIRZADI, A., SHAHABI, H., REVHAUG, I., PRAKASH, I. & TIEN BUI, D. 2018. A comparative assessment of Decision Trees algorithms for flash flood susceptibility modeling at Haraz watershed, Northern Iran. *Science of the Total Environment,* **627**(1)**:** 744-55.

KHUNJAR, W. O., PITT, P. A., BOTT, C. B. & CHANDRAN, K. 2014. Macro-nutrient removal (nitrogen). *In:* JENKINS, D. & WANNER, J. (eds.) *Activated Sludge - 100 Years and Counting.* London: IWA publishing.

KIM, J. W., LEE, B. H., SHAW, M. J., CHANG, H.-L. & NELSON, M. 2001. Application of Decision Tree induction techniques to personalized advertisements on internet storefronts. *International Journal of Electronic Commerce,* **5**(3)**:** 45-62.

KIM, J. W., LEE, B. H., SHAW, M. J., CHANG, H.-L. & NELSON, M. 2014. Application of Decision Tree induction techniques to personalized advertisements on internet storefronts. *International Journal of Electronic Commerce,* **5**(3)**:** 45-62.

KIRANMAYI, C. & MALLIKA, N. 2010. *Escherichia coli* O157:H7 - An emerging pathogen in foods of animal origin. *Veterinary World,* **3**(9)**:** 382-9.

KOLLU, K. & ORMECI, B. 2012. Effect of particles and bioflocculation on ultraviolet disinfection of *Escherichia coli*. *Water Research,* **46**(3)**:** 750-60.

KOUAME, P. K., NGUYEN-VIET, H., DONGO, K., ZURBRUGG, C., BIEMI, J. & BONFOH, B. 2017. Microbiological risk infection assessment using QMRA in agriculture systems in Cote d'Ivoire, West Africa. *Environmental Monitoring and Assessment,* **189**(11)**:** 587.

KOUGIAS, P. G., BOE, K., S, O. T., KRISTENSEN, L. A. & ANGELIDAKI, I. 2014. Anaerobic digestion foaming in full-scale biogas plants: A survey on causes and solutions. *Water Science and Technology,* **69**(4)**:** 889-95.

KOWALSKA, E., PATUREJ, E. & ZIELIŃSKA, M. 2016. Use of *Lecane inermis* for control of sludge bulking caused by the *Haliscomenobacter* genus. *Desalination and Water Treatment,* **57**(23)**:** 10916-23.

KRUIT, J., HULSBEEK, J. & VISSER, A. 2002. Bulking sludge solved? *Water Science and Technology,* **46**(1-2)**:** 457-64.

KUMARI, S. K., MARRENGANE, Z. & BUX, F. 2009. Application of quantitative RT-PCR to determine the distribution of *Microthrix parvicella* in full-scale activated sludge treatment systems. *Applied Microbiology and Biotechnology,* **83**(6)**:** 1135-41.

LACKO, N., BUX, F. & KASAN, H. C. 1999. Survey of filamentous bacteria in activated sludge plants in KwaZulu-Natal. *Water SA,* **25:** 63-68.

LAKAY, M. T., WENTZEL, M. C., EKAMA, G. A. & MARAIS, G. V. R. 1988. Bulking control with chlorination in a nutrient removal activated sludge system. *Water SA,* **14:** 35-42.

LAU, A. O., STROM, P. F. & JENKINS, D. 1984. The competitive growth of floc-forming and filamentous bacteria: A model for activated sludge bulking. *Water Pollution Control Federation,* **56**(1)**:** 52-61.

LEE, S., BASU, S., TYLER, C. W. & PITT, P. A. 2003. A survey of filamentous organisms at the Deer Island treatment plant. *Environmental Technology* **24**(7)**:** 855-65.

LEMMER, H., LIND, G., MÜLLER, E. & SCHADE, M. 2005. Non-famous scum bacteria: Biological characterization and troubleshooting. *Acta Hydrochimica et Hydrobiologica,* **33**(3)**:** 197-202.

LEVANTESI, C., ROSSETTI, S., THELEN, K., KRAGELUND, C., KROONEMAN, J., EIKELBOOM, D., NIELSEN, P. H. & TANDOI, V. 2006. Phylogeny, physiology and distribution of '*Candidatus Microthrix calida*', a new *Microthrix* species isolated from industrial activated sludge wastewater treatment plants. *Environmental Microbiology* **8**(9)**:** 1552-63.

LEVER, M. 2015. Review Filamentous Bulking in BNR Plants , Queensland, Austrialia. LEVEREDGE Water Services.

LI, B. 2016. *One-Dimensional Modeling of Secondary Settling Tanks.* Doctor of Philosophy in Civil Engineering, University of California, Los Angeles,.

LIAO, J., LOU, I. & DE LOS REYES, F. L. 2004. Relationship of species-specific filament levels to filamentous bulking in activated sludge. *Applied and Environmental Microbiology,* **70**(4)**:** 2420-8.

LIU, J. R., BURRELL, P., SEVIOUR, E. M., SODDELL, J. A., BLACKALL, L. L. & SEVIOUR, R. J. 2000. The filamentous bacterial morphotype '*Nostocoida limicola*' I contains at least two previously described genera in the low G+C Gram positive bacteria. *Systematic and Applied Microbiology,* **23**(4)**:** 528-34.

LIU, J. R. & SEVIOUR, R. J. 2001. Design and application of oligonucleotide probes for fluorescent *in situ* identification of the filamentous bacterial morphotype *Nostocoida limicola* in activated sludge. *Environmental Microbiology* **3**(9)**:** 551-60.

LIU, M., GILL, J. J., YOUNG, R. & SUMMER, E. J. 2015. Bacteriophages of wastewater foaming-associated filamentous *Gordonia* reduce host levels in raw activated sludge. *Scientific Reports,* **5**(1)**:** 1-13.

LIU, Y., PAN, Y., HUANG, D. & WANG, Q. 2017. Fault prognosis of filamentous sludge bulking using an enhanced multi-output gaussian processes regression. *Control Engineering Practice,* **62**(1)**:** 46-54.

LOH, W. & SHIH, Y. 1997. Split selection methods for classification trees. *Statistica Sinica,* **7:** 815-40.

LORET, J. F. & DUMOUTIER, N. 2019. Non-tuberculous *Mycobacteria* in drinking water systems: A review of prevalence data and control means. *International Journal of Hygiene and Environmental Health,* **222**(4)**:** 628-34.

LOU, I. & ZHAO, Y. 2012. Sludge bulking using principle component regression and artificial neural network. *Mathematical Problems in Engineering,* **2012:** 1-17.

LOU, I. C. & DE LOS REYES, F. L. 2005. Integrating decay, storage, kinetic selection, and filamentous backbone factors in a bacterial competition model. *Water Environment Research,* **77**(3)**:** 287-96.

LUUKKONEN, T., HEYNINCK, T., RÄMÖ, J. & LASSI, U. 2015. Comparison of organic peracids in wastewater treatment: Disinfection, oxidation and corrosion. *Water Research,* **85**(1)**:** 275-85.

MACAULEY, J. J., QIANG, Z., ADAMS, C. D., SURAMPALLI, R. & MORMILE, M. R. 2006. Disinfection of swine wastewater using chlorine, ultraviolet light and ozone. *Water Research,* **40**(1)**:** 2017-26.

MADONI, P., DAVOLI, D. & GIBIN, G. 2000. Survey of filamentous microorganisms from bulking and foaming activated sludge plants in Italy. *Water Research,* **34**(6)**:** 1767-72.

MANZ, W., AMANN, R., LUDWIG, W., WAGNER, M. & SCHLEIFER, K.-H. 1992. Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. *Systematic and Applied Microbiology,* **15**(4)**:** 593-600.

MARRENGANE, Z., KUMAR, S. K., PILLAY, L. & BUX, F. 2011. Rapid quantification and analysis of genetic diversity among *Gordonia* populations in foaming activated sludge plants. *Journal of Basic Microbiology* **51**(4)**:** 415-23.

MARTÍNEZ, M., SÀNCHEZ-MARRÈ, M., COMAS, J. & RODRÍGUEZ-RODA, I. 2003. Case-Based Reasoning, a promising tool to face solids separation problems in the activated sludge process. Spain: Universitat de Girona.

MARTINS, A. M. 2004. *Bulking sludge control: Kinetics, substrate storage, and process design aspects.* PhD, Delft University of Technology.

MARTINS, A. M., HEIJNEN, J. J. & VAN LOOSDRECHT, M. C. 2003. Effect of dissolved oxygen concentration on sludge settleability. *Applied Microbiology Biotechnology,* **62**(5)**:** 586-593.

MARTINS, A. M., PAGILLA, K., HEIJNEN, J. J. & VAN LOOSDRECHT, M. C. 2004. Filamentous bulking sludge-a critical review. *Water Research,* **38**(4)**:** 793-817.

MCILROY, S. J., LAPIDUS, A., THOMSEN, T. R., HAN, J., HAYNES, M., LOBOS, E., HUNTEMANN, M., PATI, A., IVANOVA, N. N., MARKOWITZ, V., VERBARG, S., WOYKE, T., KLENK, H. P., KYRPIDES, N. & NIELSEN, P. H. 2015a. High quality draft genome sequence of *Meganema perideroedes* str. Gr1(T) and a proposal for its reclassification to the family *Meganemaceae* fam. nov. *Standards in Genomic Sciences,* **10**(1)**:** 23-33.

MCILROY, S. J., SAUNDERS, A. M., ALBERTSEN, M., NIERYCHLO, M., MCILROY, B., HANSEN, A. A., KARST, S. M., NIELSEN, J. L. & NIELSEN, P. H. 2015b. MiDAS: The field guide to the microbes of activated sludge. *Database : The Journal of Biological Databases and Curation,* **2015**(1)**:** 1-8.

MCKENZIE, C. M., SEVIOUR, E. M., SCHUMANN, P., MASZENAN, A. M., LIU, J.-R., WEBB, R. I., MONIS, P., SAINT, C. P., STEINER, U. & SEVIOUR, R. J. 2006. Isolates of '*Candidatus Nostocoida limicola*' Blackall et al. 2000 should be described as three novel species of the genus *Tetrasphaera*, as

Tetrasphaera jenkinsii sp. nov., *Tetrasphaera vanveenii* sp. nov. and *Tetrasphaera veronensis* sp. nov. *International Journal of Systematic and Evolutionary Microbiology,* **56**(10)**:** 2279-90.

MCSWAIN, B. S., IRVINE, R. L., HAUSNER, M. & WILDERER, P. A. 2005a. Composition and distribution of extracellular polymeric substances in aerobic flocs and granular sludge. *Applied and Environmental Microbiology,* **71**(2)**:** 1051-1057.

MCSWAIN, B. S., IRVINE, R. L., HAUSNER, M. & WILDERER, P. A. 2005b. Composition and distribution of extracellular polymeric substances in aerobic flocs and granular sludge. *Applied and Environmental Microbiology,* **71**(2)**:** 1051-57.

MESQUITA, D. P., AMARAL, A. L. & FERREIRA, E. C. 2011. Identifying different types of bulking in an activated sludge system through quantitative image analysis. *Chemosphere,* **85**(4)**:** 643-52.

MIELCZAREK, A. T., KRAGELUND, C., ERIKSEN, P. S. & NIELSEN, P. H. 2012. Population dynamics of filamentous bacteria in Danish wastewater treatment plants with nutrient removal. *Water Research,* **46**(1)**:** 3781-95.

MIELCZAREK, A. T., SAUNDERS, A. M., AARON, M., LARSEN, P., ALBERTSEN, M., STEVENSON, M. A., MARIANNE, A., NIELSEN, J. L. & NIELSEN, P. H. 2013. The Microbial Database for Danish wastewater treatment plants with nutrient removal (MiDas-DK) - a tool for understanding activated sludge population dynamics and community stability. *Water Science and Technology,* **67**(11)**:** 2519-26.

MILOBEDZKA, A., WITESKA, A. & MUSZYNSKI, A. 2016. Factors affecting population of filamentous bacteria in wastewater treatment plants with nutrients removal. *Water Science and Technology,* **73**(4)**:** 790-7.

MOAZENI, M., NIKAEEN, M., HADI, M., MOGHIM, S., MOUHEBAT, L., HATAMZADEH, M. & HASSANZADEH, A. 2017. Estimation of health risks caused by exposure to enteroviruses from agricultural application of wastewater effluents. *Water Research,* **125**(1)**:** 104-13.

MOHAMMED, H. & SEIDU, R. 2019. Climate-driven QMRA model for selected water supply systems in Norway accounting for raw water sources and treatment processes. *Science of The Total Environment,* **660**(1)**:** 306-20.

MOMBA, M. N. B., OSODE, A. N. & SIWEBU, M. 2006. The impact of inadequate wastewater treatment on the receiving water bodies case sudy: Buffalo city and Nkokonbe municipalitites of the Eastern Cape province. *Water SA,* **32**(5)**:** 687-92.

MOTER, A. & GÖBEL, U. B. 2000. Fluorescence *in situ* hybridization (FISH) for direct visualization of microorganisms. *Journal of Microbiological Methods,* **41**(2)**:** 85-112.

MULAS, M., TRONCI, S., CORONA, F., HAIMI, H., LINDELL, P., HEINONEN, M., VAHALA, R. & BARATTI, R. 2015. Predictive control of an activated sludge process: An application to the Viikinmäki wastewater treatment plant. *Journal of Process Control,* **35**(1)**:** 89-100.

MUSVOTO, E. V., LAKAY, M. T., CASEY, T. G., WENTZEL, M. C. & EKAMA, G. A. 1999. Filamentous organism bulking in nutrient removal activated sludge systems paper 8: The effect of nitrate and nitrite. *Water SA,* **25**(4)**:** 397-408.

NASR, M. 2018. Phytobiont and Ecosystem Restitution. *In:* KUMAR, V., KUMAR, M. & PRASAD, R. (eds.) *Modeling Applications in Environmental Bioremediation Studies.* Singapore: Springer.

NASR, M. & ZAHRAN, H. F. 2016. Performance evaluation of agricultural drainage water using modeling and statistical approaches. *The Egyptian Journal of Aquatic Research,* **42**(2)**:** 141-48.

NEFESLIOGLU, H. A., SEZER, E., GOKCEOGLU, C., BOZKIR, A. S. & DUMAN, T. Y. 2010. Assessment of landslide susceptibility by Decision Trees in the metropolitan area of Istanbul, Turkey. *Mathematical Problems in Engineering,* **10**(1)**:** 1-15.

NEWHART, K. B., HOLLOWAY, R. W., HERING, A. S. & CATH, T. Y. 2019. Data-driven performance analyses of wastewater treatment plants: A review. *Water Research,* **157**(1)**:** 498-513.

NGUYEN, L. N., COMMAULT, A. S., JOHIR, M. A. H., BUSTAMANTE, H., AURISCH, R., LOWRIE, R. & NGHIEM, L. D. 2019. Application of a novel molecular technique to characterise the effect of settling on microbial community composition of activated sludge. *Journal of Environmental Management,* **251**(1)**:** 1095-9.

NGUYEN, V. L., HE, X. & DE LOS REYES, F. L. 2016. Quantifying *in situ* growth rate of a filamentous bacterial species in activated sludge using rRNA:rDNA ratio. *FEMS Microbiology Letters,* **363**(22)**:** 1-6.

NIELSEN, P. H., KRAGELUND, C., SEVIOUR, R. J. & NIELSEN, J. L. 2009a. Identity and ecophysiology of filamentous bacteria in activated sludge. *FEMS Microbiology Reviews,* **33**(6)**:** 969-98.

NIELSEN, P. H., LEMMER, H. & DAIMS, H. 2009b. *FISH handbook for biological wastewater treatment,* UK, IWA Publishing.

NIERYCHLO, M., MILOBEDZKA, A., PETRIGLIERI, F., MCILROY, B., NIELSEN, P. H. & MCILROY, S. J. 2018. The morphology and metabolic potential of the *Chloroflexi* in full-scale activated sludge wastewater treatment plants. *FEMS Microbiology Ecology,* **95**(2)**:** 1-11.

NITTAMI, T., SHOJI, T., KOSHIBA, Y., NOGUCHI, M., OSHIKI, M., KURODA, M., KINDAICHI, T., FUKUDA, J. & KURISU, F. 2019. Investigation of prospective factors that control *Kouleothrix* (Type 1851) filamentous bacterial abundance and their correlation with sludge settleability in full-scale wastewater treatment plants. *Process Safety and Environmental Protection,* **124**(1)**:** 137-42.

NITTAMI, T., SPEIRS, L. B., FUKUDA, J., WATANABE, M. & SEVIOUR, R. J. 2014. Fluorescence *in situ* hybridization probes targeting members of the phylum *Candidatus Saccharibacteria* falsely target Eikelboom type 1851 filaments and other *Chloroflexi* members. *Environmental Microbiology Reports,* **6**(6)**:** 611-7.

NITTAMI, T., SPEIRS, L. B. M., YAMADA, T., SUZUKI, I., FUKUDA, J., KURISU, F. & SEVIOUR, R. J. 2017. Quantification of *Chloroflexi* Eikelboom morphotype 1851 for prediction and control of bulking events in municipal activated sludge plants in Japan. *Applied Microbiology and Biotechnology,* **101**(9)**:** 3861-9.

NOORI, R., KHAKPOUR, A., OMIDVAR, B. & FAROKHNIA, A. 2010. Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications,* **37**(1)**:** 5856-62.

NOWAK, G. & BROWN, G. D. 1990. Characteristics of *Nostocoida limicola* and its activity in activate sludge suspension. *Journal of Water Pollution Control Federation,* **62**(2)**:** 137-42.

NTOMBELA, C., FUNKE, N., MEISSNER, R., STEYN, M. & MASANGANE, W. 2016. A critical look at South Africa's Green Drop Programme. *Water SA,* **42**(4)**:** 703.

OKUBO, T., KUBO, K., HOSOMI, M. & MURAKAMI, A. 1994. A knowledge-based decision support system for selecting small-scale wastewater treatment processes. *Water Science and Technology,* **30**(2)**:** 175-84.

OLSON, B. H. & ASVAPATHANAGUL, P. 2017. *Detection of foaming and bulking bacteria in wastewater*. USA patent application.

OVEZ, S., ORS, C., MURAT, S. & ORHON, D. 2006. Effect of hypochloride on microbial ecology of bulking and foaming activated sludge treatment for tannery wastewater. *Environmental Science Health. Part A, Toxic/Hazard Substances and Environmental Engineering,* **41**(10)**:** 2163-74.

OWENS, C. E. L., ANGLES, M. L., COX, P. T., BYLEVELD, P. M., OSBORNE, N. J. & RAHMAN, M. B. 2020. Implementation of quantitative microbial risk assessment (QMRA) for public drinking water supplies: Systematic review. *Water Research,* **174**(1)**:** 1-18.

PAL, P., KHAIRNAR, K. & PAUNIKAR, W. N. 2014. Causes and remedies for filamentous foaming in activated sludge treatment plant. *Global Nest,* **16**(1)**:** 1-11.

PARK, H. 2005. Categorical Dependent Variable Regression Models Using STATA, SAS, and SPSS. UITS Center for Statistical and Mathematical Computing.

PARKER, D., APPLETON, R., BRATBY, J. & MELCER, H. 2004. North American performance experience with anoxic and anaerobic selectors for activated sludge bulking control. *Water Science and Technology,* **50**(1)**:** 221-8.

PEDRAZZANI, R., MENONI, L., NEMBRINI, S., MANILI, L. & BERTANZA, G. 2016. Suitability of sludge biotic index (SBI), sludge index (SI) and filamentous bacteria analysis for assessing activated sludge process performance: The case of piggery slaughterhouse wastewater. *Industrial Microbiology and Biotechnology,* **43**(7)**:** 953-64.

PEI, M., ZHANG, B., HE, Y., SU, J., GIN, K., LEV, O., SHEN, G. & HU, S. 2019. State of the art of tertiary treatment technologies for controlling antibiotic resistance in wastewater treatment plants. *Environment International,* **131**(1)**:** 1-14.

PEREIRA, A., PINHO, J. L. S., FARIA, R., VIEIRA, J. M. P. & COSTA, C. 2019. Improving operational management of wastewater systems. A case study. *Water Science and Technology,* **80**(1)**:** 173–83.

PEREZ, Y. G., LEITE, S. G. F. & COELHO, M. A. Z. 2018. Activated sludge morphology characterization through an image analysis procedure. *Brazilian Journal of Chemical Engineering,* **23**(1)**:** 319-30.

PERNTHALER, J., GLÖCKNER, F.-O., SCHÖNHUBER, W. & AMANN, R. 2001. Fluorescence *in situ* hybridization with rRNA-targeted oligonucleotide probes. *Methods in Microbiology.* Academic Press.

PETROVSKI, S., TILLETT, D. & SEVIOUR, R. J. 2012. Isolation and complete genome sequence of a bacteriophage lysing *Tetrasphaera jenkinsii*, a filamentous bacteria responsible for bulking in activated sludge. *Virus Genes,* **45**(2)**:** 380-8.

PETTERSON, S. R. & STENSTRÖM, T. A. 2015. Quantification of pathogen inactivation efficacy by free chlorine disinfection of drinking water for QMRA. *Journal of Water Health,* **13**(3)**:** 625-44.

PETTERSON, S. R., STENSTRÖM, T. A. & OTTOSON, J. 2016. A theoretical approach to using faecal indicator data to model norovirus concentration in surface water for QMRA: Glomma River, Norway. *Water Research,* **91**(1)**:** 31-37.

PISA, I., SANTÍN, I., VICARIO, J. L., MORELL, A. & VILANOVA, R. 2019. ANN-based soft sensor to predict effluent violations in wastewater treatment plants. *Sensors,* **19**(1280)**:** 1-26.

PLANTE, T. R. 1990. *Solving sludge bulking problems through filamentous organism identification: Case studies in Massachusetts.* Master of Science in Environmental Engineering, University of Massachusetts.

PUYOL, D., BATSTONE, D. J., HÜLSEN, T., ASTALS, S., PECES, M. & KRÖMER, J. O. 2017. Resource recovery from wastewater by biological technologies: Opportunities, challenges, and prospects. *Frontiers in Microbiology,* **7**(2106)**:** 1-23.

QADIR, M., DRECHSEL, P., JIMÉNEZ CISNEROS, B., KIM, Y., PRAMANIK, A., MEHTA, P. & OLANIYAN, O. 2020. Global and regional potential of wastewater as a water, nutrient and energy source. *Natural Resources Forum,* **44**(1)**:** 40-51.

QUACH-CU, J., HERRERA-LYNCH, B., MARCINIAK, C., ADAMS, S., SIMMERMAN, A. & REINKE, R. 2018. The effect of primary, secondary, and tertiary wastewater treatment processes on antibiotic resistance gene (ARG) concentrations in solid and dissolved wastewater fractions. *Water,* **10**(1)**:** 37.

RADOMSKI, N., BETELLI, L., MOILLERON, R., HAENN, S., MOULIN, L., CAMBAU, E., ROCHER, V., GONCALVES, A. & LUCAS, F. S. 2011. *Mycobacterium* behavior in wastewater treatment plant, a bacterial model distinct from *Escherichia coli* and Enterococci. *Environmental Science and Technology,* **45**(12)**:** 5380-6.

RAMAN, E. 2014. *Modeling SST and filamentous bulking conditions.* PhD, Technical University of Denmark.

RAMIN, E., WÁGNER, D. S., YDE, L., SZABO, P., RASMUSSEN, M. R., DECHESNE, A., SMETS, B. F., MIKKELSEN, P. S. & PLÓSZ, B. G. Modelling the impact of filamentous bacteria abundance in a secondary settling tank: CFD sub-models optimization using long-term experimental data Wastewater Treatment Modelling Seminar 2014 Denmark

RAMIREZ-CASTILLO, F. Y., LOERA-MURO, A., JACQUES, M., GARNEAU, P., AVELAR-GONZALEZ, F. J., HAREL, J. & GUERRERO-BARRERA, A. L. 2015. Waterborne pathogens: detection methods and challenges. *Pathogens,* **4**(2)**:** 307-34.

RAMOTHOKANG, T. R., DRYSDALE, G. D. & BUX, F. 2003. Isolation and cultivation of filamentous bacteria implicated in activated sludge bulking. *Water S.A.,* **29**(4)**:** 405-9.

RAMOTHOKANG, T. R., MTHEMBU, N. N. & BUX, F. 2006a. Evaluation of growth characteristics of filamentous bacteria using optimised isolation techniques. Institute for Water and Wastewater Technology. . South Africa.

RAMOTHOKANG, T. R., NAIDOO, D. & BUX, F. 2006b. 'Morphological shifts' in filamentous bacteria isolated from activated sludge processes. *World Journal of Microbiology and Biotechnology,* **22**(8)**:** 845-850.

RANMIN, E., SIN, G., MIKKELSEN, P. S. & PLOSZ, B. G. 2014. Significance of settling model structures and parameter subsets in modelling WWTPs under wet-weather flow and filamentous bulking conditions. *Water Research,* **63:** 209-221.

RICHARD, B., BROWN, S. & COLLINS, F. 2003. Activated sludge microbiology and their control. *USEPA National Operator Trainers Conference* New York, USA.

ROSSETTI, S., TANDOI, V. & WANNER, J. 2017. *Activated Sludge Separation Problems: Theory, Control Measures, Practical Experiences - Second Edition.* .

ROSSETTI, S., TOMEI, M. C., NIELSEN, P. H. & TANDOI, V. 2005. "*Microthrix parvicella*", a filamentous bacterium causing bulking and foaming in activated sludge systems: A review of current knowledge. *FEMS Microbiology Reviews,* **29**(1)**:** 49-64.

RUSTUM, R. 2009. *Modelling Activated Sludge Wastewater Treatment Plants Using Artificial Intelligence Techniques (Fuzzy Logic and Neural Networks).* Doctor of Philosophy, Heriot-Watt University.

SAKAJI, R. & FUNAMIZU, N. 1998. Microbial risk assessment and its role in the development of wastewater reclamation policy. *In:* ASANO, T. (ed.) *Wastewater Reclamation and Reuse.* Boca Raton, Fla, USA: CRC Press.

SAKTAYWIN, W., TSUNO, H., NAGARE, H., SOYAMA, T. & WEERAPAKKAROON, J. 2005. Advanced sewage treatment process with excess sludge reduction and phosphorus recovery. *Water Research,* **39**(5)**:** 902-910.

SALGOT, M. & FOLCH, M. 2018. Wastewater treatment and water reuse. *Current Opinion in Environmental Science & Health,* **2:** 64-74.

SAMPSON, A., OWUSU-ANSAH, E. D.-G. J., MILLS-ROBERTSON, F. C., AYI, I., ABAIDOO, R. C., HALD, T. & PERMIN, A. 2017. Probabilistic quantitative microbial risk assessment model of farmer exposure to *Cryptosporidium* spp. in irrigation water within Kumasi Metropolis-Ghana. *Microbial Risk Analysis,* **6:** 1-8.

SANAWAR, H., XIONG, Y., ALAM, A., CROUÉ, J.-P. & HONG, P.-Y. 2017. Chlorination or monochloramination: Balancing the regulated trihalomethane formation and microbial inactivation in marine aquaculture waters. *Aquaculture,* **480:** 94-102.

SANKARAN, S., KHANAL, S. K., POMETTO, A. L., 3RD & VAN LEEUWEN, J. H. 2008. Ozone as a selective disinfectant for nonaseptic fungal cultivation on corn-processing wastewater. *Bioresour Technol,* **99**(17)**:** 8265-72.

SANT'ANA, A. S., FRANCO, B. D. G. M. & SCHAFFNER, D. W. 2014. Risk of infection with *Salmonella* and *Listeria monocytogenes* due to consumption of ready-to-eat leafy vegetables in Brazil. *Food Control,* **42:** 1-8.

SCHAUER, M. & HAHN, M. W. 2005. Diversity and phylogenetic affiliations of morphologically conspicuous large filamentous bacteria occurring in the pelagic zones of a broad spectrum of freshwater habitats. *Applied and Environmental Microbiology,* **71**(4)**:** 1931-40.

SCHETS, F. M., SCHIJVEN, J. F. & DE RODA HUSMAN, A. M. 2011. Exposure assessment for swimmers in bathing waters and swimming pools. *Water Research,* **45**(7)**:** 2392-400.

SCHIJVEN, J., TEUNIS, P., SUYLEN, T., KETELAARS, H., HORNSTRA, L. & RUTJES, S. 2019. QMRA of adenovirus in drinking water at a drinking water treatment plant using UV and chlorine dioxide disinfection. *Water Research,* **158**(1)**:** 34-45.

SDIRI, A., PINHO, J. & RATANATAMSKUL, C. 2018. Water resource management for sustainable development. *Arabian Journal of Geosciences* **11**(124)**:** 1-2.

SEKI, K., THULLNER, M. & BAVEYE, P. 2004. Nutrient uptake kinetics of filamentous microorganisms: Comparison of cubic, exponential, and Monod models. *Annals of Microbiology,* **54**(2)**:** 181-8.

SEKIGUCHI, Y., KAMAGATA, Y., SYUTSUBO, K., OHASHI, A., HARADA, H. & NAKAMURA, K. 1998. Phylogenetic diversity of mesophilic and thermophilic granular sludges determined by 16S rRNA gene analysis. *Microbiology,* **144**(9)**:** 2655-65.

SERRA, P., SÀNCHEZ, M., LAFUENTE, J., CORTÉS, U. & POCH, M. 1997. ISCWAP: A knowledge-based system for supervising activated sludge processes. *Computers & Chemical Engineering,* **21**(2)**:** 211-21.

SERUGA, P., KRZYWONOS, M., PYZANOWSKA, J., URBANOWSKA, A., PAWLAK-KRUCZEK, H. & NIEDZWIECKI, L. 2019. Removal of Ammonia from the Municipal Waste Treatment Effluents using Natural Minerals. *Molecules,* **24**(20)**:** 1-13.

SETO, E. Y., KONNAN, J., OLIVIERI, A. W., DANIELSON, R. E. & GRAY, D. M. D. 2016. A quantitative microbial risk assessment of wastewater treatment plant blending: Case study in San Francisco Bay. *Environmental Science: Water Research and Technology,* **2**(1)**:** 134-145.

SEVIOUR, E. M., EALES, K., IZZARD, L., BEER, M., CARR, E. L. & SEVIOUR, R. J. 2006. The *in situ* physiology of "*Nostocoida limicola*" II, a filamentous bacterial morphotype in bulking activated sludge, using fluorescence *in situ* hybridization and microautoradiography. *Water Science and Technology* **54**(1)**:** 47-53.

SEVIOUR, E. M., WILLIAMS, C., DEGREY, B., SODDELL, J. A., SEVIOUR, R. J. & LINDREA, K. C. 1994. Studies on filamentous bacteria from australian activated sludge plants. *Water Research,* **28**(11)**:** 2335-42.

SEVIOUR, R. & NIELSEN, P. H. 2010. *Microbial ecology of activated sludge,* London, UK, IWA Publishing.

SEVIOUR, R. J., LINDREA, K. C. & OEHMEN, A. 2010. The activated sludge process. *In:* SEVIOUR, R. J. & NIELSEN, P. H. (eds.) *Microbial ecology of activated sludge.* London: IWA Publishing.

SEZGIN, M., JENKINS, D. & PARKER, D. S. 1978. A unified theory of filamentous activated sludge bulking. *Water Pollution Control Federation,* **50**(2)**:** 362-381.

SEZGIN, M., JENKINS, D. & PARKER, D. S. 1982. Variation of sludge volume index with activated sludge characteristics *Water Research,* **16**(1)**:** 83-88.

SHAHZAD, M., KHAN, S. & PAUL, P. 2015. Influence of temperature on the performance of a full-scale activated sludge process operated at varying solids retention times whilst treating municipal sewage. *Water,* **7**(12)**:** 855-67.

SHWE, S. H. M. & OO, H. 2016. Fever Classification and Remedial Recommendation System. *International Journal of Open Information Technologies,* **4**(12)**:** 2307-8162.

SINGH, K. K., PAL, M., OJHA, C. S. P. & SINGH, V. P. 2008. Estimation of Removal Efficiency for Settling Basins Using Neural Networks and Support Vector Machines. *Journal of Hydrologic Engineering,* **13**(3)**:** 146-155.

SMETS, I. Y., BANADDA, N., DEURINCK, J., RENDERS, N., JENNE, R. & VAN IMPE, J. F. 2006. Dynamic modeling of filamentous bulking in lab-scale activated sludge processes. *Journal of Process Control,* **16:** 313-9.

SOBSEY, M. D. 1989. Inactivation of health-related microorganisms in water by disinfection processes. *Water Science and Technology* **21**(3)**:** 179-195.

SOLLER, J. A., EFTIM, S. E. & NAPPIER, S. P. 2018. Direct potable reuse microbial risk assessment methodology: Sensitivity analysis and application to State log credit allocations. *Water Research,* **128**(1)**:** 286-92.

SOLLER, J. A., SCHOEN, M. E., BARTRAND, T., RAVENSCROFT, J. E. & ASHBOLT, N. J. 2010. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Research,* **44**(16)**:** 4674-91.

SPEIRS, L., NITTAMI, T., MCILROY, S., SCHROEDER, S. & SEVIOUR, R. J. 2009. Filamentous bacterium Eikelboom Type 0092 in activated sludge plants in Australia is a member of the phylum Chloroflexi. *Applied and Environmental Microbiology,* **75:** 2446-52.

SPEIRS, L. B. M., DYSON, Z. A., TUCCI, J. & SEVIOUR, R. J. 2017. Eikelboom filamentous morphotypes 0675 and 0041 embrace members of the *Chloroflexi*: resolving their phylogeny, and design of fluorescence *in situ* hybridisation probes for their identification. *FEMS Microbiology Ecology* **93**(10)**:** 1-13.

SPEIRS, L. B. M., MCILROY, S. J., PETROVSKI, S. & SEVIOUR, R. J. 2011. The activated sludge bulking filament Eikelboom morphotype 0914 is a member of the *Chloroflexi Environmental Microbiology Reports,* **3**(2)**:** 159-64.

SPEIRS, L. B. M., RICE, D. T. F., PETROVSKI, S. & SEVIOUR, R. J. 2019. The Phylogeny, Biodiversity, and Ecology of the *Chloroflexi* in Activated Sludge. *Frontiers in Microbiology,* **10:** 1-28.

SPELLMAN, F. R. 2003. *Handbook of Water and Wastewater Treatment Plant Operations,* Boca Raton London, New York. , Lewis Publishers.

SUNGER, N., HAMILTON, K. A., MORGAN, P. M. & HAAS, C. N. 2018. Comparison of pathogen-derived 'total risk' with indicator-based correlations for recreational (swimming) exposure. *Environmental Science and Pollution Research* **26**(30)**:** 30614-24.

SUPPES, L. M., CANALES, R. A., GERBA, C. P. & REYNOLDS, K. A. 2016. *Cryptosporidium* risk from swimming pool exposures. *International Journal of Hygiene and Environmental Health,* **219**(8)**:** 915-19.

SVEC, D., TICHOPAD, A., NOVOSADOVA, V., PFAFFL, M. W. & KUBISTA, M. 2015. How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments. *Biomolecular Detection and Quantification,* **3:** 9-16.

SZELĄG, B., GAWDZIK, A. & GAWDZIK, A. 2017. Application of selected methods of black box for modelling the settleability process in wastewater treatment plant. *Ecological Chemistry and Engineering,* **24**(1)**:** 119-27.

TAK, S. & KUMAR, A. 2017. Chlorination disinfection by-products and comparative cost analysis of chlorination and UV disinfection in sewage treatment plants: Indian scenario. *Environmental Science and Pollution Research International,* **24**(34)**:** 26269-26278.

TAKACS, I. 2008. *Experiments in activated sludge modelling.* Doctor of Philosophy Ghent University.

TAKENAKA, R., AOI, Y., OZAKI, N., OHASHI, A. & KINDAICHI, T. 2018. Specificities and Efficiencies of Primers Targeting Candidatus Phylum *Saccharibacteria* in Activated Sludge. *Materials,* **11**(7)**:** 1129.

TANDOI, V., JENKINS, D. & WANNER, J. 2006. *Activated Sludge Separation Problems.,* London, IWA Publishing.

TAYLOR, M. J., BENTHAM, R. H. & ROSS, K. E. 2014. Limitations of using propidium monoazide with qPCR to discriminate between live and dead legionella in biofilm samples. *Microbiology Insights,* **7:** 15-24.

TCHOBANOGLOUS, G., BURTON, F. L. & STENSEL, H. D. 2003. *Wastewater Engineering: Treatment and Reuse,* McGraw-Hill, New York, Metcalf and Eddy, Inc.

TEKLEHAIMANOT, G. Z., GENTHE, B., KAMIKA, I. & MOMBA, M. N. 2015. Prevalence of enteropathogenic bacteria in treated effluents and receiving water bodies and their potential health risks. *Science and the Total Environment,* **518-19:** 441-9.

THOMSEN, T. R., BLACKALL, L. L., DE MURO, M. A., NIELSEN, J. L. & NIELSEN, P. H. 2006. *Meganema perideroedes* gen. nov., sp. nov., a filamentous alphaproteobacterium from activated sludge. *International Journal of Systematic and Evolutionary Microbiology,* **56**(8)**:** 1865-68.

THOMSEN, T. R., KJELLERUP, B. V., NIELSEN, J. L., HUGENHOLTZ, P. & NIELSEN, P. H. 2002. *In situ* studies of the phylogeny and physiology of filamentous bacteria with attached growth. *Environmental Microbiology* **4**(7)**:** 383-91.

TIAN, Y., ZHENG, L. & SUN, D. 2006. Functions and behaviours of activated sludge extracellular polymeric substances (EPS): a promising environmental interest. *Journal of Environmental Sciences,* **18**(3)**:** 420-427.

TREBESIUS, K., LEITRITZ, L., ADLER, K., SCHUBERT, S., AUTENRIETH, I. B. & HEESEMANN, J. 2000. Culture independent and rapid identification of bacterial pathogens in necrotising fasciitis and

streptococcal toxic shock syndrome by fluorescence *in situ* hybridisation. *Medical Microbiology and Immunology,* **188**(4)**:** 169-75.
UNESCO 2019. The United Nations World Water Development Report 2019.
USEPA 1983. Methods for Chemical Analysis of Water and Wastes. Washington, DC: United States Environmental Protection Agency.
USEPA 2012. *Recreational Water Quality Criteria (Office of Water 820-F-12-058),* Washington, DC: U.S.
VAEZI, F., NADDAFI, K., KARIMI, F. & ALIMOHAMMADI, M. 2004. Application of chlorine dioxide for secondary effluent polishing. *IJEST*.
VAIOPOULOU, E., MELIDIS, P. & AIVASIDIS, A. 2007. Growth of filamentous bacteria in an enhanced biological phosphorus removal system. *Desalination,* **213**(1)**:** 288-96.
VAN HAANDEL, A. C. & VAN DER LUBBE, J. G. M. 2012. *Handbook of biological wastewater treatment,* London, IWA Publishing.
VAN LEEUWEN, J. 1988. Bulking control with ozonation in a nutrient removal activated sludge system. *Water SA,* **14**(3)**:** 119-24.
VAN LEEUWEN, J. 1992. A review of the potential application of non-specific activated sludge bulking. *Water SA,* **18:** 101-6.
VAN LOOSDRECHT, M. C. M., LOPEZ-VAZQUEZ, C. M., MEIJER, S. C. F., HOOIJMANS, C. M. & BRDJANOVIC, D. 2015. Twenty-five years of ASM1: past, present and future of wastewater treatment modelling. *Journal of Hydroinformatics,* **17**(5)**:** 697-718.
VAN VU, T., PHAM, P. D., WINKLER, M. S., ZURBRÜGG, C., ZINSSTAG, J., TRAN, B. H. & NGUYEN-VIET, H. 2018. Estimation of involuntary excreta ingestion rates in farmers during agricultural practices in Vietnam. *Human and Ecological Risk Assessment: An International Journal,* **25**(8)**:** 1942-52.
VANYSACKER, L., DENIS, C., ROELS, J., VERHAEGHE, K. & VANKELECOM, I. F. J. 2014. Development and evaluation of a TaqMan duplex real-time PCR quantification method for reliable enumeration of *Candidatus Microthrix*. *Journal of Microbiological Methods,* **97:** 6-14.
VERGARA, G., ROSE, J. B. & GIN, K. Y. H. 2016. Risk assessment of noroviruses and human adenoviruses in recreational surface waters. *Water Research,* **103:** 276-82.
VERVAEREN, H., WILDE, K., MATTHYS, J., BOON, N., RASKIN, L. & VERSTRAETE, W. 2005. Quantification of an Eikelboom Type 021N bulking event with fluorescence *in situ* hybridization and real-time PCR. *Applied Microbiology and Biotechnology,* **68:** 695-704.
VOULVOULIS, N. 2018. Water reuse from a circular economy perspective and potential risks from an unregulated approach. *Current Opinion in Environmental Science & Health,* **2:** 32-45.
VREČKO, D., HVALA, N. & CARLSSON, B. 2003. Feedforward-feedback control of an activated sludge process: a simulation study. *Water Science and Technology,* **47**(12)**:** 19-26.
WAGNER, D. S., RAMIN, E., SZABO, P., DECHESNE, A. & PLOSZ, B. G. 2015. *Microthrix parvicella* abundance associates with activated sludge settling velocity and rheology: Quantifying and modelling filamentous bulking. *Water Research,* **78**(1)**:** 121-32.
WAGNER, M., AMANN, R., KÄMPFER, P., ASSMUS, B., HARTMANN, A., HUTZLER, P., SPRINGER, N. & SCHLEIFER, K.-H. 1994a. Identification and *in situ* detection of gram-negative filamentous bacteria in activated sludge. *Systematic and Applied Microbiology,* **17**(3)**:** 405-17.
WAGNER, M., ERHART, R., MANZ, W., AMANN, R., LEMMER, H., WEDI, D. & SCHLEIFER, K. H. 1994b. Development of an rRNA-targeted oligonucleotide probe specific for the genus Acinetobacter and its application for in situ monitoring in activated sludge. *Applied and Environmental Microbiology,* **60**(3)**:** 792-800.
WANG, H. F., HU, H., YANG, H. Y. & ZENG, R. J. 2016a. Characterization of anaerobic granular sludge using a rheological approach. *Water Research,* **106:** 116e125.
WANG, J., LI, S. Y., JIANG, F., WU, K., LIU, G. L., LU, H. & CHEN, G. H. 2015. A modified oxic-settling-anaerobic activated sludge process using gravity thickening for excess sludge reduction. *Scientific Reports,* **5:** 1-10.

WANG, P., YU, Z., QI, R. & ZHANG, H. 2016b. Detailed comparison of bacterial communities during seasonal sludge bulking in a municipal wastewater treatment plant. *Water Research,* **105:** 157-66.

WANG, X., RATNAWEERA, H., HOLM, J. A. & 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.

WANNER, J. 1994. *Activated sludge: Bulking and foaming control,* Lancaster, PA, Lewis Publishers.

WANNER, J. & GRAU, P. 1989. Identification of filamentous microorganisms from activated sludge: A compromise between wishes, needs and possibilities. *Water Research,* **23**(7)**:** 883-91.

WELZ, P. J., ESTERHUYSEN, A., VULINDLU, M. & BEZUIDENHOUT, C. 2014. Filament identification and dominance of Eikelboom Type 0092 in activated sludge from wastewater treatment facilities in Cape Town, South Africa. *Water SA,* **40**(4)**:** 649-57.

WESTRELL, T. 2004. *Microbial risk assessment and its implications for risk management in urban water systems.* PhD Thesis, Linköpings University.

WHO 2001. *Water Quality: Guidelines, Standards and Health. Assessment of Risk and Risk Management for Water Related Infectious Diseases,* Geneva: World Health Organisation.

WHO 2006. *Guidelines for the safe use of wastewater, excreta and greywater,* Geneva: World Health Organisation

WHO 2011. *Guidelines for Drinking-Water Quality.,* Geneva: World Health Organisation.

WHO 2016. *Quantitative Microbial Risk Assessment: application for water safety management.,* Geneva: World Health Organisation

WIESE, J., SCHMITT, S., STAHL, A., HANSEN, J. & SCHMITT, T. G. Experience management for wastewater treatment. Proc. of the GI-Workshopwoche "Lehren-Lernen-Wissen-Adaptivität", Workshop on Knowledge and Experience Management, FGWM03, 2003.

WIJNBLADH, E. 2007. *Ozone technology for sludge.* Master of Science, University of Agricultural Sciences.

WILLIAMS, T. & UNZ, R. 1985. Isolation and characterization of filamentous bacteria present in bulking activated sludge. *Applied Microbiology and Biotechnology,* **22:** 273-82.

WUKASCH, R. F. 1993. Proceedings of the 48th Industrial Waste Conference Purdue University. USA: CRC Press - Taylor & Francis.

XIE, G., ROIKO, A., STRATTON, H., LEMCKERT, C., DUNN, P. K. & MENGERSEN, K. 2016. A generalized QMRA beta-poisson dose-response model. *Risk Analysis,* **36**(10)**:** 1948-58.

XIN, G., GOUGH, H. L. & STENSEL, H. D. 2008. Effect of Anoxic Selector Configuration on Sludge Volume Index Control and Bacterial Population Fingerprinting. *Water Environment Research,* **80**(12)**:** 2228-40.

XU, S., SUN, M., ZHANG, C., SURAMPALLI, R. & HU, Z. 2014. Filamentous sludge bulking control by nano zero-valent iron in activated sludge treatment systems. *Environmental Science: Processes and Impacts,* **16**(12)**:** 2721-8.

YANG, H., YAO, Q., HUANG, C. L., DENG, J. C., ZHANG, J. H. & ZHANG, J. 2009. Control of filamentous bulking in the A/O biological phosphorus removal process. *Beijing Gongye Daxue Xuebao / Journal of Beijing University of Technology,* **35:** 1663-9.

YEO, B. & GRANT, D. 2018. Predicting service industry performance using Decision Tree analysis. *International Journal of Information Management,* **38:** 288-300.

YONG, M., YONGZHEN, P. & SHUYING, W. 2005. Feedforward-feedback control of dissolved oxygen concentration in a predenitrification system. *Bioprocess and Biosystems Engineering* **27**(4)**:** 223-8.

YOU, S. J. & SUE, W. M. 2009. Filamentous bacteria in a foaming membrane bioreactor. *Journal of Membrane Science,* **342**(1)**:** 42-49.

YU, G. P., WANG, J. Y., YUAN, M. Z. & YU, Y. Research on prediction method of sludge bulking based on ANN and grey Markov model. 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 8-12 June 2015 2015. 1622-1627.

YU, Z., HAGHIGHAT, F., FUNG, B. C. M. & YOSHINO, H. 2010. A Decision Tree method for building energy demand modeling. *Energy and Buildings,* **42**(10)**:** 1637-46.

ZHANG, M., YAO, J., WANG, X., HONG, Y. & CHEN, Y. 2019. The microbial community in filamentous bulking sludge with the ultra-low sludge loading and long sludge retention time in oxidation ditch. *Scientific Reports,* **9**(1)**:** 136-93.

ZHANG, P., YUAN, M. & WANG, H. 2008. Improvement of nitrogen removal and reduction of operating costs in an activated sludge process with feedforward–cascade control strategy. *Biochemical Engineering Journal,* **41**(1)**:** 53-58.

ZHANG, T., SHAO, M.-F. & YE, L. 2012. 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *Multidisciplinary Journal of Microbial Ecology,* **6**(6)**:** 1137-47.

ZHOU, L., ECHIGO, S., NAKANISHI, T., YAMASAKI, S. & ITOH, S. 2017. Development of a Multiphase Inactivation Model for an Advanced Oxidation Process and Uncertainty Analysis in Quantitative Microbial Risk Assessment. *Ozone: Science & Engineering,* **40**(2)**:** 79-92.

ZHU, X. X. & SIMPSON, A. R. 1996. Expert System for Water Treatment Plant Operation. *Journal of Environmental Engineering,* **122**(9)**:** 822-829.

ZHUANG, Y., REN, H., GENG, J., ZHANG, Y., ZHANG, Y., DING, L. & XU, K. 2015. Inactivation of antibiotic resistance genes in municipal wastewater by chlorination, ultraviolet, and ozonation disinfection. *Environmental Science and Pollution Research International,* **22**(9)**:** 7037-44.

ZOU, J., HAN, Y. & SO, S.-S. 2009. Overview of Artificial Neural Networks. *Methods in molecular biology (Clifton, N.J.),* **458:** 14-22.

ZVIMBA, J. N. & MUSVOTO, E. V. 2020. Modelling energy efficiency and generation potential in the South African wastewater services sector. *Water Science and Technolology,* **81**(5)**:** 876-890.

APPENDICES

APPENDIX ONE: COMMON FILAMENTOUS BACTERIA IDENTIFIED VIA FISH METHODS AND REPORTED IN THIS STUDY

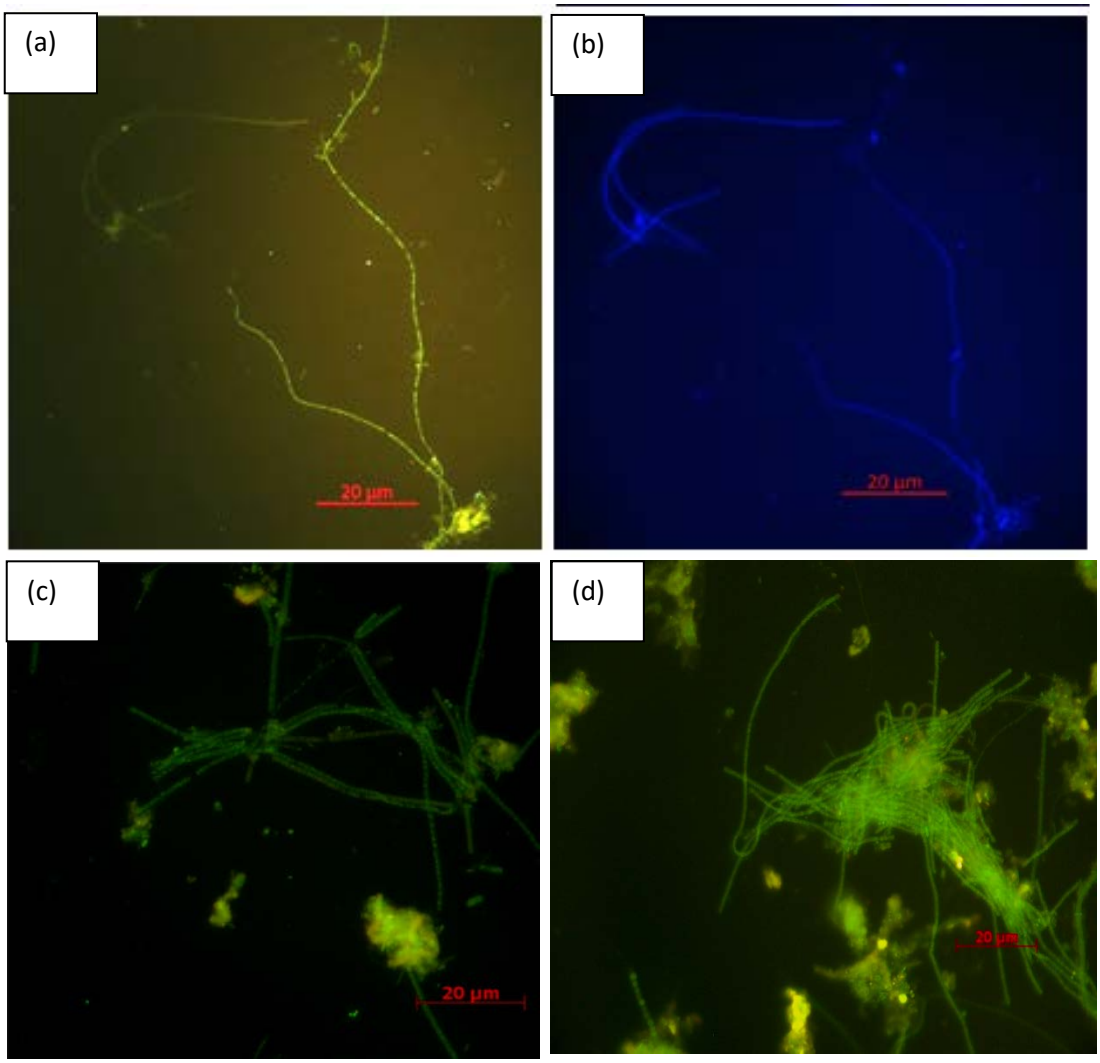


Figure 1. Fluorescence *in situ* hybridisation images at 20 μm: (a) Type 1851 confirmed with CHL 1851 probe; b) DNA staining with DAPI of Type 1851, (c) *Thiothrix* spp with G123T probe; (d) Type 021N with 21N probe.

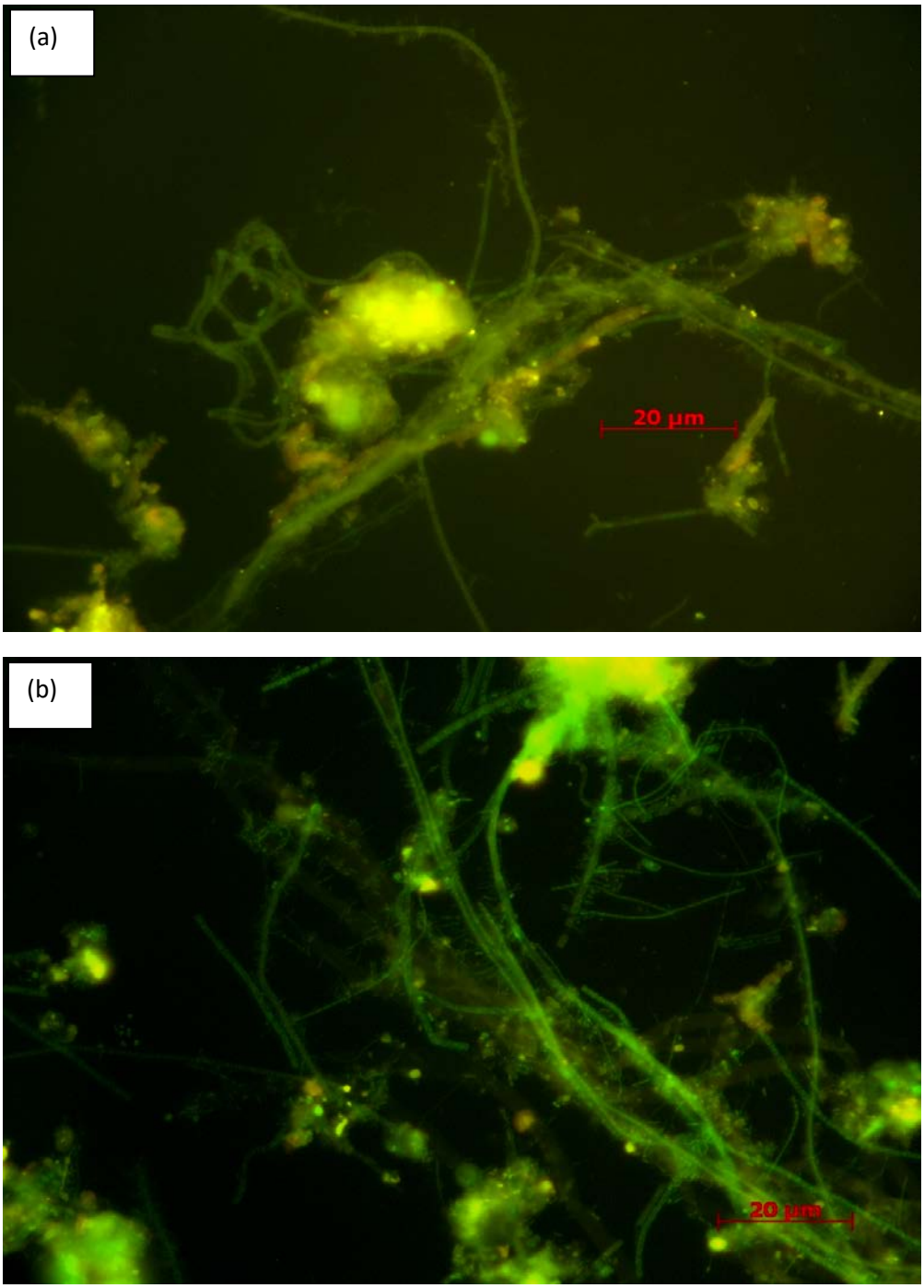


Figure 2. Fluorescence *in situ* hybridisation images images of EUB mix (EUB338 I, EUB 338 II and EUB 338 III) probes targeting most bacteria (images at 20 μm).

APPENDIX TWO: PRINCIPAL COMPONENT ANALYSIS STANDARDIZATION AND PERFORMANCE

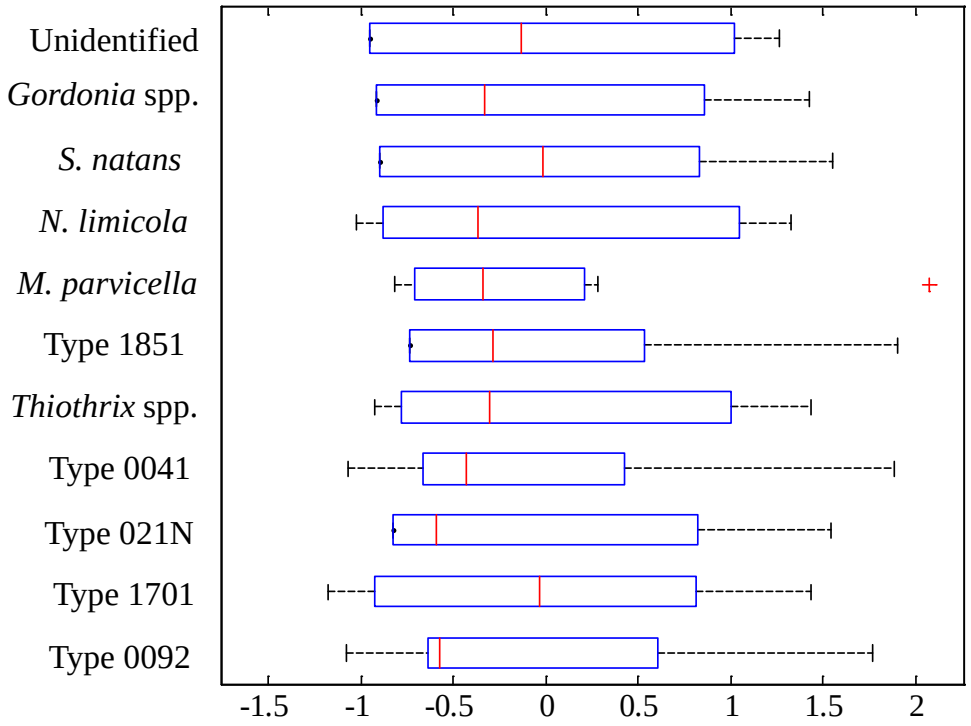


Figure 1. Standardization process to obtain variables with a mean of 0 and a standard deviation of 1 for principal component analysis.

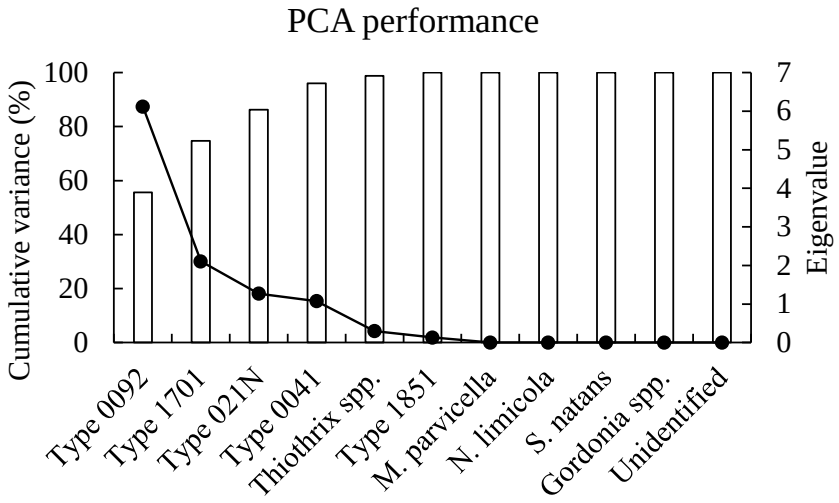


Figure 2. Percentages of the explained variances and eigenvalues for principal component analysis.

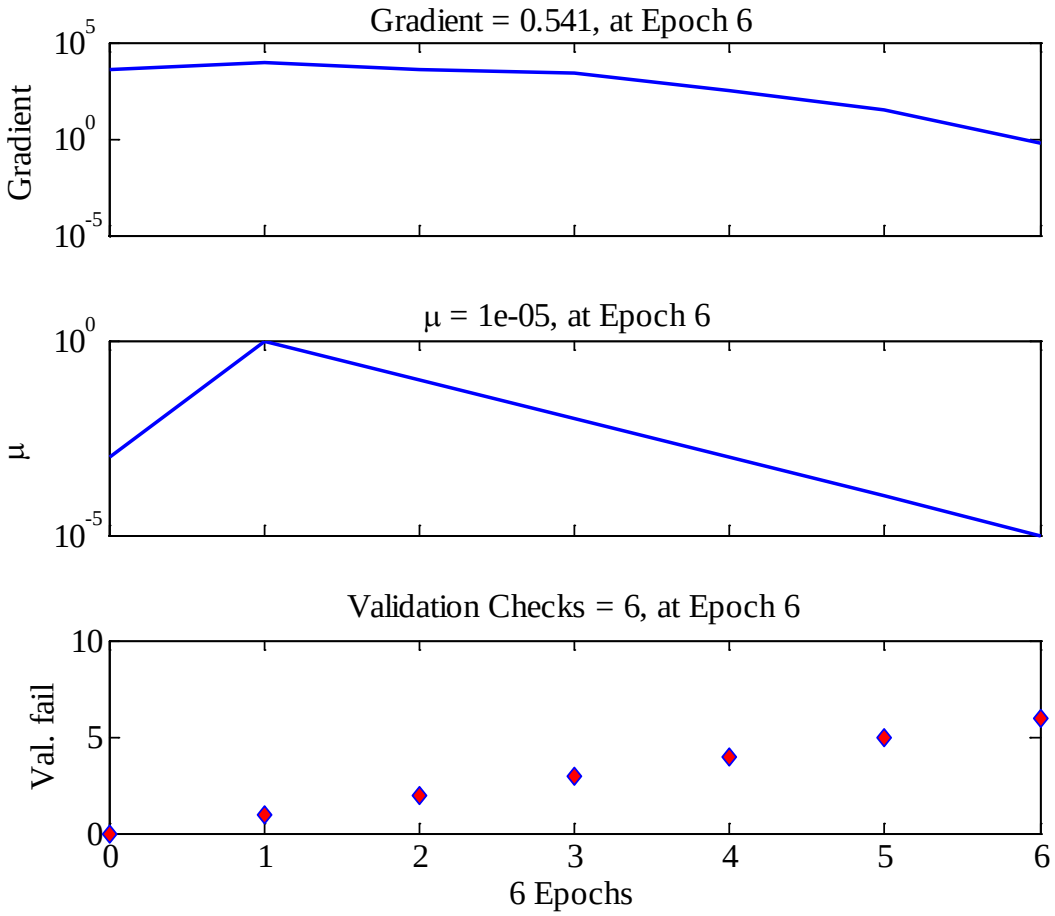


Figure 3. Progress of the training process terminated based on the maximum validation failures, achieving 6 consecutive epochs (trials).

APPENDIX THREE: ARTIFICIAL INTELLIGENCE AND MULTIVARIATE STATISTICS FOR COMPREHENSIVE ASSESSMENT OF FILAMENTOUS BACTERIA IN WASTEWATER TREATMENT PLANTS EXPERIENCING SLUDGE BULKING

Environmental Technology & Innovation 19 (2020) 100853

Contents lists available at ScienceDirect

Environmental Technology & Innovation

journal homepage: www.elsevier.com/locate/eti

Artificial intelligence and multivariate statistics for comprehensive assessment of filamentous bacteria in wastewater treatment plants experiencing sludge bulking

Nashia Deepnarain [a], Mahmoud Nasr [b], Sheena Kumari [a], Thor A. Stenström [a], Poovendhree Reddy [c], Kriveshin Pillay [a], Faizal Bux [a,*]

[a] Institute for Water and Wastewater Technology, Durban University of Technology, Durban, 4000, South Africa
[b] Sanitary Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, 21544, Egypt
[c] Department of Community Health Studies, Faculty of Health Sciences, Durban University of Technology, South Africa

ARTICLE INFO

Article history:
Received 6 February 2020
Received in revised form 27 April 2020
Accepted 27 April 2020
Available online 1 May 2020

Keywords:
Artificial neural network
Filament index scale
Filaments images
Full-scale WWTPs
Principal component analysis

ABSTRACT

Sludge bulking is an operational hurdle that affects the solid–liquid separation in wastewater treatment plants (WWTPs) worldwide. In this study, filamentous bulking issues were investigated in seven WWTPs located in South Africa using artificial neural network (ANN) and multivariate statistics. The microbial community belonging to sludge bulking was determined using staining and microscopic methods, with a further confirmed identification of selected species via fluorescent *in situ* hybridization (FISH). Based on a filament index scale from 1 (None filament) to 7 (Excessive filament), the developed ANN could predict the sludge volume index (SVI) using the abundances of ten inputs of filamentous species. Eikelboom Type 0041 attained the highest impact on SVI, followed by *Gordonia* spp., *Nostocoida limicola*, and *Thiothrix* spp. Principal component analysis (PCA) combined with FISH images showed that most WWTPs experienced inadequate sludge settling properties; however, the application of an efficient aeration system (i.e., diffusion) in the three-stage Phoredox process improved the settling characteristics of bio-flocs. Operational conditions that caused filament overgrowth in each WWTP were also determined. The study outputs would provide a scientific basis to control the proliferation of filamentous bacteria in other WWTPs located in similar environmental conditions to South Africa.

1. Introduction

The activated sludge process is widely employed in wastewater treatment plants (WWTPs), where aeration units are used for the biological conversion of soluble organic compounds into settleable solids followed by clarification tanks for solid–liquid separation (Björnsson et al., 2002). The efficiency of the sedimentation process is critical in activated sludge WWTPs, and it governs the capability and potentiality of the entire treatment system (Nasr, 2018). Generally, the settling properties of sludge is described in terms of sludge volume index (SVI), expressing the amount (in mL) occupied by 1 gram of suspended solids in the mixed liquor (usually 1 L sample) for 30 min (Han et al., 2016). For instance, activated sludge having SVI over 150 mL/g could be subjected to sludge bulking, which hinders the operation of the activated sludge process

* Correspondence to: Institute for Water and Wastewater Technology, Durban University of Technology, P.O. Box 1334, Durban, South Africa.
E-mail address: faizalb@dut.ac.za (F. Bux).

https://doi.org/10.1016/j.eti.2020.100853
2352-1864/

(Deepnarain et al., 2019). One of the main reasons of sludge bulking is the mass proliferation of filamentous bacteria (Fan et al., 2020). In this context, monitoring and assessment of the filament abundance are important domains of research to attain proper management of WWTPs.

The conventional techniques of filament identification are mainly carried out based on the morphological and staining characteristics of microorganisms (Deepnarain et al., 2015). These features include filament shape and length, trichome cell shape, Gram and Neisser staining, and the detection of intracellular energy storage products such as poly(3-hydroxybutyrate) (PHB) (Zhang et al., 2019). However, it has been reported that traditional identification techniques are still insufficient to distinguish between different filaments having the same morphological and phenotypic characteristics (Abusam et al., 2017). For instance, the filaments *Nostocoida limicola* types I, II, and III vary only by their general appearance and distinct cell dimensions; however, they retain the same Neisser staining reaction (Snaidr et al., 2002). Alternatively, fluorescence *in situ* hybridization (FISH) has been successfully used as an effective technique to identify filamentous organisms in bulking activated sludge according to the matching nature of deoxyribonucleic acid (DNA) or DNA/RNA double strands (Nielsen et al., 2009). The FISH technique has also become a powerful tool for obtaining accurate quantitative data on filamentous bacteria (Nittami et al., 2019). It has also been demonstrated that the FISH method could be employed to evaluate the effects of operating conditions and physicochemical factors on the growth of filaments in WWTPs (Speirs et al., 2017).

Various filamentous species, including *Microthrix parvicella*, Type 1701, *Thiothrix* spp., Type 021N, *Gordonia* spp., and Type 0041, have been observed to cause severe bulking events in activated sludge WWTPs (Graveleau et al., 2005). Previous studies have reported that the Eikelboom Types 0041, 1851, and 1701 filamentous species are linked to severe sludge bulking episodes in WWTPs treating domestic and industrial sources (Thomsen et al., 2002; Lee et al., 2003). In another study, Insel et al. (2014) detected the filamentous microorganisms Type 0092, Type 0041, and Type 0803 in a membrane bioreactor (MBR) WWTP. The growth of each filamentous type is influenced by various operating conditions, *viz*., dissolved oxygen (DO) concentration, food-to-microorganisms (F/M) ratio, sludge retention time (SRT), and/or sulfur concentration (Fan et al., 2020). In addition, the influent wastewater characteristics (e.g., pH and nitrogen and phosphorus concentrations) impact the dominance of the filamentous bacterial populations responsible for sludge bulking in WWTPs (Martins et al., 2004). Accordingly, the relative abundance of each filament is considered a "fingerprint" influenced by certain environmental conditions in the treatment units (Lee et al., 2003). Moreover, there are no universal or common strategies that could be used to control the excessive growth of filamentous populations in most WWTPs (Han et al., 2018). Hence, the function and physiological behavior of filamentous bacteria still need comprehensive studies. This objective can be achieved with the assistance of advanced modeling and statistical methods.

Full-scale WWTPs comprise various physical, chemical, and biological processes (Nguyen et al., 2019), which are complex and difficult to describe using conventional linear models. Artificial neural network (ANN) is an effective and powerful tool able to simulate these non-linear processes, even under fluctuating environmental conditions (Han et al., 2019). Large historical data obtained from real WWTPs can be imported into an ANN model for the prediction of the non-linear behaviors with high accuracy and adequacy (Nasr, 2018). Moreover, the ANN model can be employed for process optimization, giving more robust and accurate results compared to regression-based mathematical models (Fawzy et al., 2018). ANN has a parallel processing structure, in which a number of neurons or nodes is operated concurrently to mimic the biological nervous systems of humans (Bagheri et al., 2015). This black-box model has an adaptive and self-learning ability to reveal the influence of each independent input on a dependent response (output) (Boztoprak et al., 2016). Large data retrieved from WWTPs can also be interpreted and classified using principal component analysis (PCA) to extract valuable information (Abusam et al., 2017). This multivariate approach is suitable for identifying correlations between random variables and describing the essential features of WWTPs via data dimensionality reduction (Awolusi et al., 2018). In PCA, the dimensionality of such datasets is reduced based on two axes, known as principal components (PCs). The first principal component (PC1) retains the largest variance direction in the data, whereas the second principal component (PC2) is orthogonal to PC1, and it represents the second maximum source of variation in the data (Nasr, 2018). The PCA technique can also be used to compress a set of data-points into specific numbers of clusters. These methods have recently found several applications for the assessment of bacterial abundance in full-scale WWTPs (Awolusi et al., 2018).

According to the aforementioned information, this work aims at investigating the dominance of filamentous bacteria in seven WWTPs located in South Africa using ANN and PCA tools. For this purpose, wastewater samples were collected from the aeration tanks and subjected to the SVI examination. The microbial community associated with sludge bulking was assessed using staining and microscopic methods, and then selected species were confirmed via FISH using 16S rRNA oligonucleotide probes. An ANN model was proposed to predict the SVI and to rank the filament species according to the order of importance. Further, PCA was employed to correlate the main features of each WWTP with the abundance of filamentous bacteria. The outcomes of this work are expected to provide valuable knowledge on the control of sludge bulking in WWTPs.

2. Materials and methods

2.1. Sampling from full-scale WWTPs

The selected WWTPs in this study are located in three South African provinces, i.e., Gauteng, KwaZulu-Natal, and Western Cape. The plants were designed for the requirements of biological nutrient removal (BNR), *viz*., University of Cape

Table 1
Descriptive overview of seven wastewater treatment works (WWTWs) used for monitoring filamentous bulking.

WWTW	WWTW-I	WWTW-II	WWTW-III	WWTW-IV	WWTW-V	WWTW-VI	WWTW-VII
Province	Western Cape	Western Cape	Western Cape	Gauteng	Gauteng	KZN	KZN
Configuration	Five-stage Phoredox Process (Modified Bardenpho)	Modified University of Cape Town (UCT)	Modified University of Cape Town (UCT)	Three-stage Phoredox process	Three-stage Phoredox process	Modified Ludzack–Ettinger (MLE)	Modified University of Cape Town (UCT)
Wastewater source	Domestic	Domestic	Domestic with 5% industrial	Domestic with 5% industrial	Domestic with 5% industrial	Domestic	Domestic with 5% industrial
Aeration technique	Surface aeration	Surface aeration	Surface aeration	Fine and course bubbles using new aeration diffusers	Fine and course bubbles using old aeration-blowers	Surface aeration	Surface aeration
Primary settling tank	No	No	No	Yes	Yes	Yes	No
Balancing tank	No	No	Yes	Yes	Yes	No	No
Anaerobic digester	Yes	Yes	Yes	No	No	Yes	No
Chemical dosing	None	None	None	Ferric chloride	Ferric chloride	None	Aluminum chloride
Annual dry weather flow (ML/d)	12.5	7.5	14.0	45.0	40.0	25.0	7.2
SVI (mL/g)	155.85 ± 7.20	157.05 ± 2.27	167.62 ± 12.15	50.13 ± 13.52	98.09 ± 3.27	167.95 ± 7.18	178.33 ± 6.19
MLSS (mg/L)	4837 ± 638	5346 ± 752	5465 ± 657	4384 ± 482	4514 ± 526	5105 ± 750	5071 ± 597
SRT (days)	30	25	15	20	20	25	30
F/M ratio (1/day)	0.13	0.11	0.14	0.11	0.11	0.14	0.18
Settling efficiency	Poor	Poor	Poor	Good	Good	Poor	Poor

SVI is sludge volume index (mL/g); MLSS is mixed liquor suspended solids (mg/L); SRT is sludge retention time (days); F/M ratio is food-to-microorganism ratio (kg_{COD}/kg_{VSS}/day).

Town (UCT) process configuration, three-stage Bardenpho process, modified Ludzack–Ettinger (MLE) process, and five-stage Bardenpho system. Table 1 lists the specifications and operational features of the seven BNR wastewater treatment works (WWTWs). The wastewater samples were collected on a monthly basis from the aeration tanks to determine sludge settleability, bio-floc structure, and filamentous abundance.

SVI is sludge volume index (mL/g); MLSS is mixed liquor suspended solids (mg/L); SRT is sludge retention time (days); F/M ratio is food-to-microorganism ratio (kg_{COD}/kg_{VSS}/day)

2.2. Filamentous bacteria identification

2.2.1. Conventional staining and microscopic methods

Initially, the morphotype and staining reactions of filamentous bacteria were performed using the classification system of Eikelboom et al. (1998). Wet mounts, and Gram and Neisser staining reactions were conducted to evaluate the physical factors (e.g., floc shape, size, and morphology), filament abundance, and the impact of filaments on floc quality. In addition, cell shape, either the presence or absence of a sheath, and bacteria with attached growth were used to distinguish the filamentous organisms (Jenkins et al., 1986). Further, the filament index was rated on a scale that ranged from 1 to 7, where 1 denoted no filamentous organisms, and 7 represented excess growth of filamentous organisms (Deepnarain et al., 2019). For instance, the filamentous abundance was categorically ranked as "None", "Few", "Some", "Common", "Very common", "Abundant", and "Excessive" for filament index records of "1", "2", "3", "4", "5", "6", and "7", respectively.

2.2.2. Fluorescent in situ hybridization (FISH)

First, mixed-liquor samples were pre-treated, fixed, and dehydrated. The pre-treatment step was conducted using 1× Phosphate-buffered saline (PBS) and 8% paraformaldehyde, and then the samples were washed in PBS and stored in PBS:ethanol (1:1, v:v) at −20 °C until further use (Nielsen et al., 2009). The fixed samples were sonicated for 45 s at 2 watts using an ultrasonic liquid processor (XL-2000; Tri-Lab Support, USA) to shear and break up the floc structure into smaller pieces. The selected filamentous bacteria were targeted using 16S rRNA-specific oligonucleotide probes (Table 2). The probes were labeled with a fluorescent dye 5-(and-6)-Carboxyfluorescein diacetate N-succinimidyl ester (CFSE; MWG-Biotech, Ebersberg, Germany; A subsidiary company of Roche Products (Pty) Ltd. South Africa).

Table 2
Oligonucleotide probes and target microorganisms for detection of filamentous bacteria.

Probe	Sequence (5′–3′)	FA (%)	Target microorganism	Reference
EUB338 mix (I, II and III)	GCT GCC TCC CGT AGG AGT	35	Most bacteria, *Plantomycetales*, and *Verrucomicrobiales*	Daims et al. (1999)
G123T	CCT TCC GAT CTC TAT GCA	40	*Thiothrix* species: *T. nivea, T. unzii, T. fructosivorans, and T. defluvii.* Type 021N group I, II, III	Kanagawa et al. (2000)
G1B	TGT GTT CGA GTT CCT TGC	30	Type 021N group I	Kanagawa et al. (2000)
G2M	GCACCACCGACCCCTTAG	35	Type 021N group II	Kanagawa et al. (2000)
G3M	CTCAGGGATTCCTGCCAT	30	Type 021N group III	Kanagawa et al. (2000)
21N	TCCCTCTCCCAAATTCTA	30	Type 021N	Kanagawa et al. (2000)
CFX1223	CCATTGTAGCGTGTGTGTMG	35	All *Chloroflexi*	Björnsson et al. (2002)
CFX197	TCCCGGAGCGCCTGAACT	40	*Chloroflexi* OTU A	Speirs et al. (2009)
CFX223	GGTGCTGGCTCCTCCCAG	35	*Chloroflexi* OTU B	Speirs et al. (2009)
SNA	CAT CCC CCT CTA CCG TAC	45	*Sphaerotilus natans*	Wagner et al. (1994)
NLIMI 91	CGC CAC TAT CTT CTC AGT	20	*Nostocoida limicola I*	Liu and Seviour (2001)
NLIMII 175	GGC TCC GTC TCG TAT CCG	40	*N. limicola II*	Liu and Seviour (2001)
NLIMIII 301	CCC AGT GTG CCG GGC CAC	20	*N. limicola III*	Liu and Seviour (2001)
CHL 1851	AATTCCACGAACCTCTGCCA	20	Eikelboom Type 1851	Beer et al. (2002)
MPA60	GGATGGCCGCGTTCGACT	20	*Microthrix parvicella*	Erhart et al. (1997)
MPA223	GCCGCGAGACCCTCCTAG			
MPA645	CCGGACTCTAGTCAGAGC			
MPA650	CCCTACCGGACTCTAGTC			

FA: Formamide range of 20–45% for proper hybridization assays.

After sonication, 10 μL from the sample were plated on wells of Teflon-coated slides and dried at 48 °C for 10 min. The slides were subjected to a dehydration step using an ethanol series in 50 mL polypropylene tubes (3 min each in 50%, 80%, and 100% ethanol). This step is used to eliminate the excess water from the bacterial cells, improving the image resolution during microscopy.

The hybridization step was carried out under the conditions of high stringency (Nielsen et al., 2009). Hybridization buffer (Milli-Q water; 5 M NaCl; 1 M Tris–HCl; 10% sodium dodecylsulfate; 20%–45% formamide) was prepared for each percentage of formamide. The prepared buffer, together with the specific gene probe (50 ng/μL working solution), was added onto each well and incubated overnight at 46 °C. Following incubation, the slides were rinsed with warm distilled water and transferred onto a pre-warmed wash buffer (1M Tris/HCL; 10% SDS; 5 M NaCl; 0.5 M EDTA) in 50 mL polypropylene tubes. The tubes were incubated at 48 °C for 45 min. After incubation, the slides were counter-stained with the DNA stain 4′, 6-diamidino-2-phenylindol (DAPI) (0.25 μg/mL). Further, each well was covered by a volume of 10 μL DAPI and maintained under the dark condition for 10 min; i.e., the slides were rinsed with distilled water and left to air-dry overnight. A drop of Vector Shield (mounting agent containing anti-bleaching) was added onto the slides and enclosed by a cover-slip.

For image visualization, the slides were examined using an AxioLab ApoTome microscope (Carl Zeiss; Göttingen, Germany) comprising FLUOS Fluorochrome/Filter Set. Zeiss Axio Vision (Release 4.6.3) was used as an imaging software to create images.

2.3. Sludge settling test

For the sludge volume index (SVI in mL/g) test, a mixed liquor sample with a known volume of 1 L was allowed to precipitate in a measuring cylinder for 30 min. Subsequently, the volume of settled sludge per gram of solids was recorded, as reported by Boztoprak et al. (2016).

2.4. Artificial neural network (ANN) modeling

ANN is built from a set of interconnected elements (i.e., units, neurons, or nodes), which are arranged in multiple layers. In this study, a feed-forward ANN model was used to predict the SVI data using ten inputs of filaments bacteria, namely Type 0092, *Thiothrix* spp., Type 1851, *M. parvicella*, Type 1701, *N. limicola*, Type 0041, *Sphaerotilus natans*, *Gordonia* spp., and Type 021N (Fig. 1a). The ANN structure is composed of (a) an input layer that collects the records from the 10 input attributes, (b) a hidden layer containing 15 neurons, and (c) an output layer designating the model output (i.e., SVI). Hence, the network configuration is defined as 10–15–1. This configuration was chosen based on several trial-and-error procedures. The weights and activation functions were used to transform the data, in which the Levenberg–Marquardt optimization was selected for the training stage to adjust the weight and bias values. The data points were distributed into 70% for training, 15% for cross-validation, and 15% for testing the ANN model.

The ANN computations were performed using MATLAB (R2015a) software, according to the following steps:

The input layer comprised ten neurons that connected the input signals to fifteen neurons in the hidden layer. The neurons in the input layer ($x_{10\times1}$) were multiplied by a weight matrix ($w_{15\times10}$) and added to a constant bias ($b_{15\times1}$). The

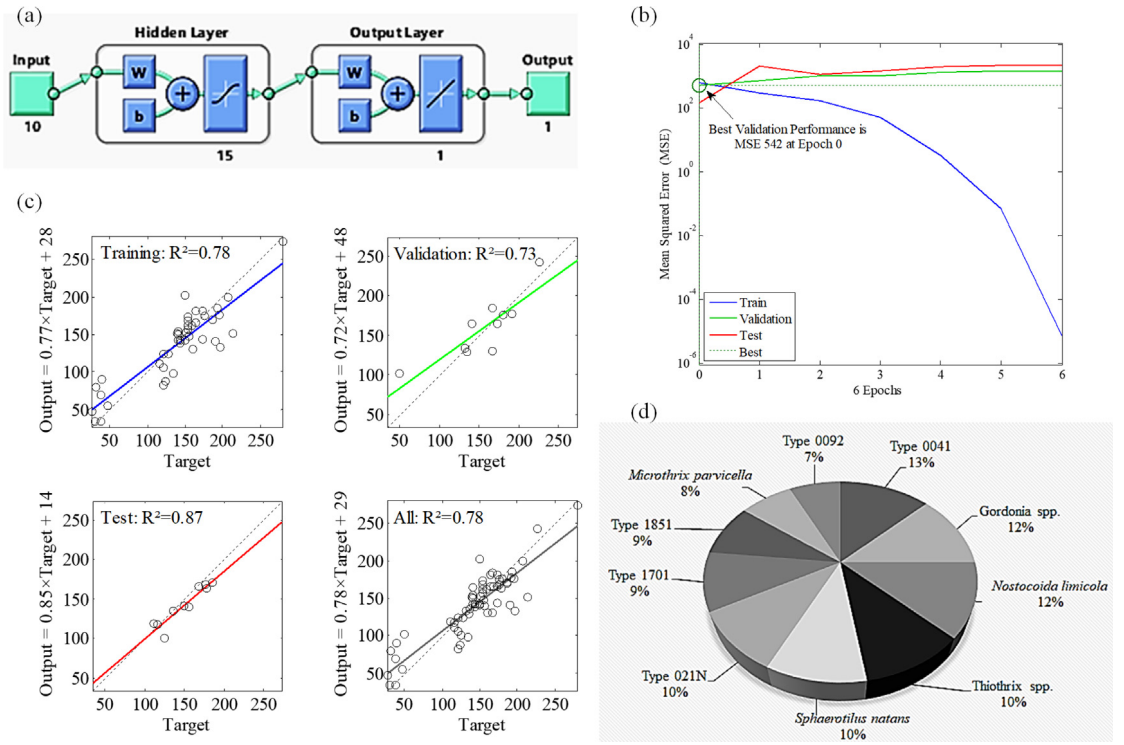


Fig. 1. ANN simulation for prediction of SVI (an output) using abundance of filamentous bacteria (ten inputs) Type 0041, *Gordonia* spp., *N. limicola*, *Thiothrix* spp., *S. natans*, Type 021N, Type 1701, Type 1851, *M. parvicella*, and Type 0092: (a) ANN model configuration 10–15–1, (b) ANN performance, (c) ANN prediction accuracy for training, validation, and test, and (d) ANN relative importance of each filament on SVI.

results were summed up and adapted by a tan-sigmoid "tansig" activation function (Supplementary Eq. S1) to limit the product between −1 and 1. The outputs of the hidden layer ($y_{15\times1}$) were estimated by Eq. (1).

$$y_{15\times1} = \text{tansig}\left(\sum w_{15\times10}\cdot x_{10\times1} + b_{15\times1}\right) \tag{1}$$

Similar to the previous step, each node in the middle layer was interconnected to that of the last layer. The responses of the hidden layer ($y_{15\times1}$) were multiplied by weights ($w_{1\times15}$) and added to a constant bias ($b_{1\times1}$). The product was summed up and transformed into an output ($z_{1\times1}$) via a linear "purelin" activation function (Supplementary Eq. S2). The output of this stage is given by Eq. (2).

$$z_{1\times1} = \text{purelin}\left(\sum w_{1\times15}\cdot y_{15\times1} + b_{1\times1}\right) \tag{2}$$

2.5. Statistical analysis

Filamentous bacteria were monitored in seven WWTPs and then described as categorical data using multivariate analysis. PCA was conducted using the following sequential steps (Awolusi et al., 2018): (a) data standardization into standard deviation = 1 and mean = 0 (Supplementary Fig. S1), (b) computation of covariance matrix, (c) estimation of eigenvectors and eigenvalues (Supplementary Fig. S2), (d) calculation of principal components (i.e., PC1 corresponds to the largest eigenvalues), and (e) reduction of dataset dimensions to enhance the quality of representation. The results of PCA were then used for clustering via the *k*-means algorithm. These steps were performed using the function "*princomp*" in MATLAB (R2015a) software.

3. Results and discussion

3.1. Subjective scoring of filamentous bacteria abundance

Fig. 2 shows the wet mount observations of the floc structural characteristics during different bulking conditions. It was observed that the filaments formed bridging among the flocs and enlarged the surface area, leading to a loose and

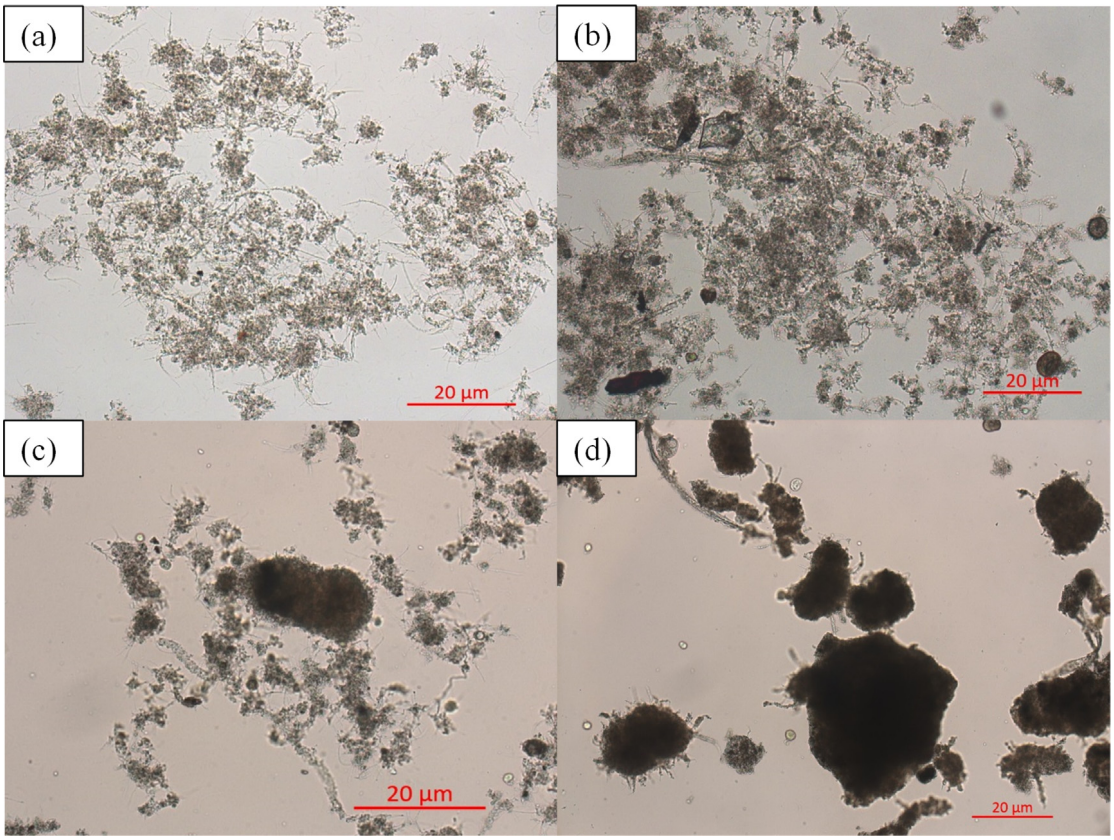


Fig. 2. Wet mount observation of floc structural characteristics (Images at 40× magnification, direct illumination, and phase contrast): (a) Subjective scores of 5 and 6 for "Most dominant" and "Excessive" filamentous bacteria, (b) Subjective score of 4 for "Dominant" filamentous bacteria, (c) Subjective scores of 2 and 3 for "Few" and "Common" filamentous bacteria, and (d) Subjective scores of <2 for "Rare" filamentous bacteria.

less compacted structure. For a better illustration of filament observations, the subjective scores for the occurrence and dominance of filamentous bacteria were scaled, as adapted from Lee et al. (2003).

Fig. 2(a) shows the presence of filamentous bacteria with the classifications of "Most dominant" and "Excessive". The shape and strength of flocs were irregular, diffuse, and weak, corresponding to subjective scores of 5 and 6. During "Excessive" bulking conditions, filamentous microorganisms tend to protrude from the activated sludge flocs, forming "filament-to-filament" or "floc-to-filament" aggregates with low density (Jin et al., 2003). Fig. 2(b) represents the "Dominant" filamentous bacteria, having irregular and open flocs. This classification could be referred to as a subjective score of 4, and the filaments were observed in all flocs at medium density (i.e., 5–20 per floc). Fig. 2(c) displays the filaments with subjective scores of 2 and 3 and a low density of 1–5 filaments per floc. This finding could describe the "Few" and "Common" filamentous scale, where the floc is firmer, more compact, and round in shape. Fig. 2(d) illustrates the "Rare" filamentous sector, having a subjective score below <2. In this group, the flocs are more firm, compact, and spherical in shape, and the filaments are observed in an occasional floc (Lee et al., 2003).

3.2. Standard staining techniques (Gram and Neisser) for filamentous detection

Fig. 3 shows Neisser and Gram staining of filamentous bacteria. Fig. 3(a) demonstrates Neisser positive of the dominant filamentous bacteria Type 0092 detected mainly within flocs, in which both filament branching and attached growth were absent. Moreover, Fig. 3(b) displays Gram-positive culture of filament *Candidatus Microthrix parvicella* with coiled and twisted morphology. Fig. 3(c) indicated that Type 0041 and Type 1701 appeared with abundant attached growth; i.e., the bacterium Type 0041 is defined as a Gram variable, sheathed, curved, and unbranched filament (Williams and Unz, 1985). In another work, Thomsen et al. (2002) found that Type 0041 appeared in about 88% of the WWTP samples, indicating the "Most abundant" filamentous bacterium with attached growth in activated sludge. In Fig. 3(d), Type 021N appeared as Gram-negative and unbranched, and the septa were clearly visible with the absence or very few of the attached growth forms. The features of Type 021N have also been reported in a previous study by Guo et al. (2012) in a continuous activated sludge system.

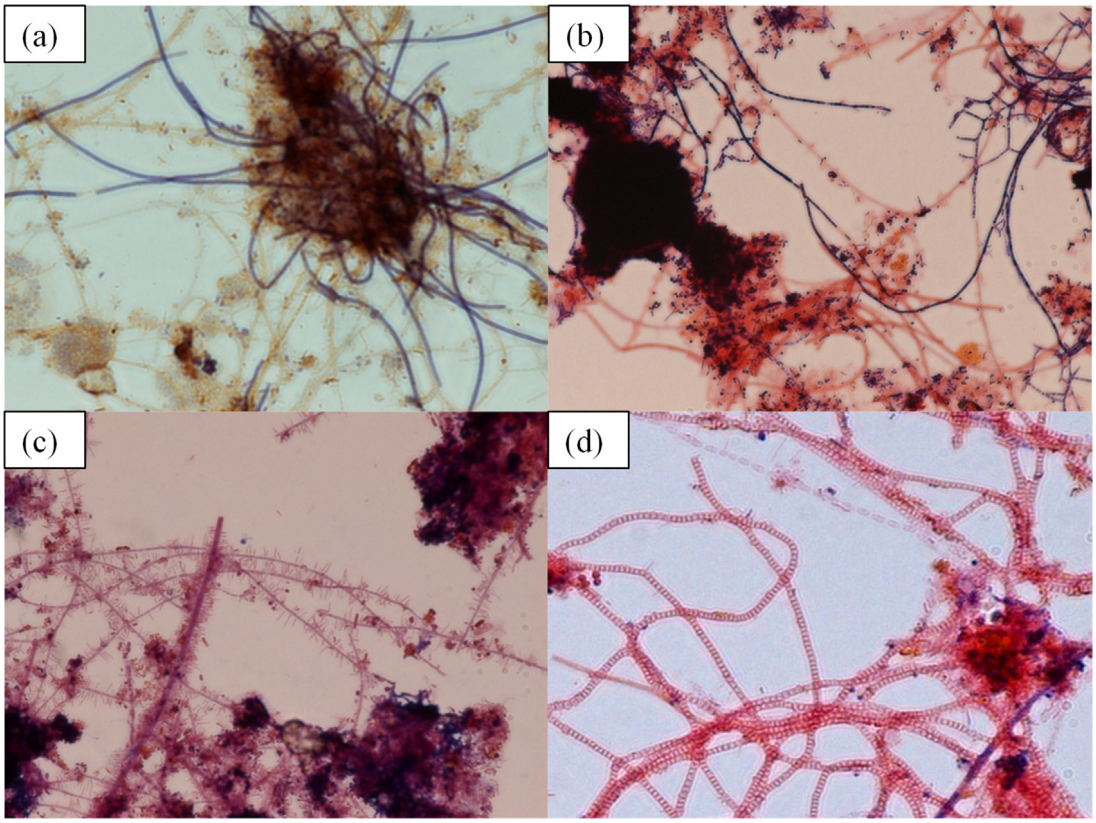


Fig. 3. Images of filamentous bacteria: (a) Neisser positive of dominant filamentous bacteria Type 0092 found mainly within flocs, (b) Gram positive culture of filament *M. parvicella*, (c) Type 0041 and Type 1701, and (d) Type 021N: Gram negative, unbranched, and septa clearly visible.

3.3. Fluorescence in situ hybridization (FISH) images of filamentous bacteria

The FISH images were used to ensure the identification and abundance of specific filamentous species (Fig. 4). For instance, the presence of Type 0092 was confirmed using the Phylum-targeted CFX-MIX probes (Fig. 4a) and DNA staining with DAPI (Fig. 4b). Moreover, the filament *M. parvicella* was observed using the MPA-MIX probe (Fig. 4c) and DNA staining with DAPI (Fig. 4d).

Similarly, Lemmer et al. (2005) applied the FISH technique to detect Type 0092 filaments using CF319 probe and *M. parvicella* using either single probes (i.e., MPA60, MPA223, and MPA645) or MPA-mix. In addition, Graveleau et al. (2005) detected Type 0092 and *M. parvicella* using both standard staining techniques (Gram, Neisser, and PHB) and FISH probes. Similar observations have also been reported by You and Sue (2009), indicating that the FISH and DNA sequencing methods could be more appropriate for estimating the filamentous microbial diversity in activated sludge WWTPs.

3.4. Relative importance of each filament using ANN simulation

3.4.1. ANN performance

The mean squared error (MSE) between the experimental data of SVI and the model output (predicted) was used to evaluate the performances of training, validation, and testing (Fig. 1b):

During the training stage, the ANN parameters (i.e., weights and biases) were altered to obtain a model output congruent to the target data. The training step is stopped based on either of the following criteria, whichever occurs first (Fawzy et al., 2018): (1) the magnitude of the performance gradient reaches a minimum threshold of 10^{-5}, or (2) the number of validation checks becomes 6. As shown in Fig. 1(b), the initial MSE of the training dataset was 542, which tended to decrease along with a serial number of trials until reaching the maximum validation checks of 6. Hence, the progress of the training process was terminated according to the maximum validation failures, achieving 6 consecutive epochs (trials) (See also Supplementary Fig. S3). This pattern revealed that the learning procedure successfully improved the ANN model performance during training.

The performance function of the validation dataset was also examined, along with the training phase, to eliminate any inaccuracies from the network. The validation step is employed to evaluate the generalization ability of the model as a

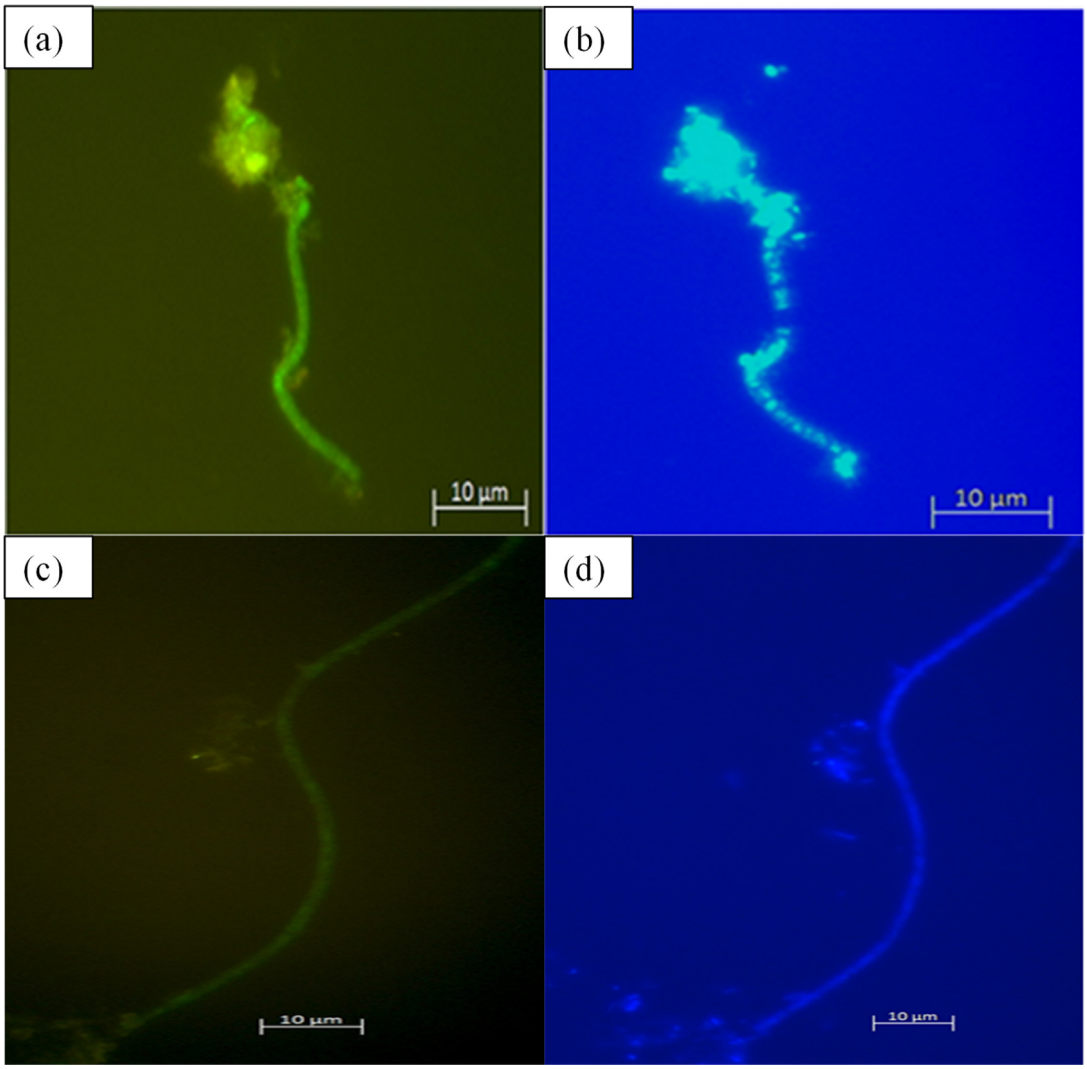


Fig. 4. Fluorescence *in situ* hybridization (FISH) images for the dominant filamentous bacteria at 10 µm: (a) Phylum-targeted CFX-MIX probes, (b) DNA staining with DAPI of Type 0092, (c) MPA-MIX probe, and (d) DNA staining with DAPI of *Microthrix parvicella*.

predictive tool within the range of the data records used for training (Bagheri et al., 2015). The validation process stopped earlier at an epoch number 0, i.e., once the MSE began to increase (Fig. 1b), signifying that the network initiated to overfit the data. At the point of overfitting, the model generates errors and/or random noises while importing new input data. Hence, this early stopping criterion is used to avoid overfitting problems and improve generalization capacity.

Finally, the testing phase was undertaken using new independent records that were not considered for either training or validation. The error (in terms of MSE) of the test set enlarged after the initial iteration, indicating that the validation and testing curves were comparable. This observation implied that the proposed ANN model had satisfactory robustness and generalization capabilities (Nasr, 2018).

Based on the above findings, the ANN model displayed the highest performance at epoch 0 with MSE (the performance function) of 542. At this point, the adjustable weights and biases were estimated, which were then used to determine the relative importance of each input (See section "3.4.3. Sensitivity analysis").

3.4.2. Regression plot

Fig. 1(c) shows the coefficient of determination (R^2 value) between the measured SVI and the predicted outputs for training, validation, and test. In each figure, the dashed and solid lines represent the actual and perfect fits, respectively. The overall R^2 value of the entire dataset was 0.78, revealing that the designed network could describe 78.0% of the variability of SVI episodes. In this context, the ANN model could be employed to offer acceptable predictive accuracy for the SVI data due to filamentous bacteria in WWTPs.

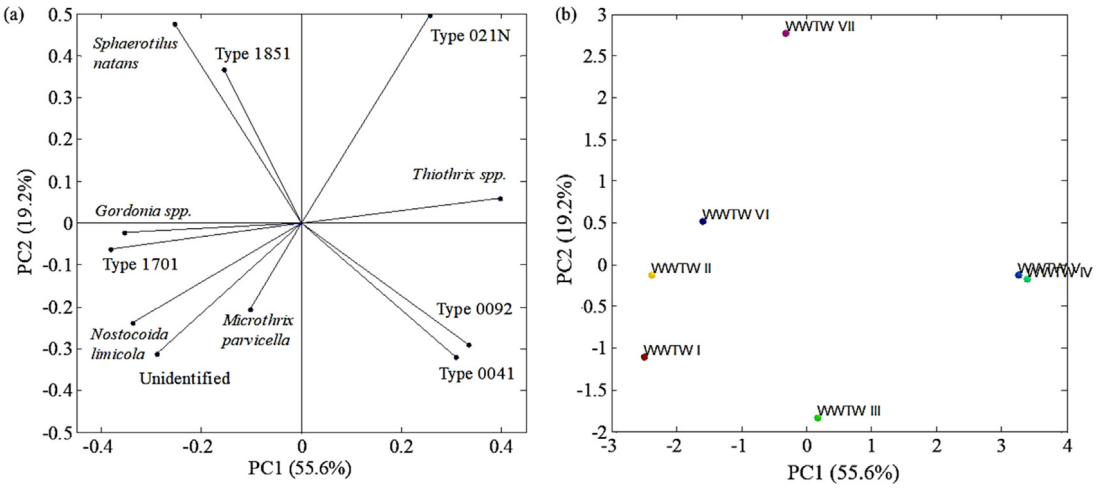


Fig. 5. PCA results for distribution of filamentous bacteria in studied wastewater treatment works (WWTWs): (a) Loading plot, and (b) Cluster analysis.

3.4.3. Sensitivity analysis

The connection weights estimated from the ANN model at the optimum performance could describe the relative importance of each independent input. Hence, the partitioning step of the input-to-hidden and hidden-to-output weights was applied using Eq. (3) to arrange the input attributes by order of importance (Garson, 1991).

$$I_j = \frac{\sum_{m=1}^{m=N_h}\left(\left(\frac{\left|W_{jm}^{ih}\right|}{\sum_{k=1}^{N_i}\left|W_{km}^{ih}\right|}\right)\times\left|W_{mn}^{ho}\right|\right)}{\sum_{k=1}^{k=N_i}\left\{\sum_{m=1}^{m=N_h}\left(\frac{\left|W_{km}^{ih}\right|}{\sum_{k=1}^{N_i}\left|W_{km}^{ih}\right|}\right)\times\left|W_{mn}^{ho}\right|\right\}} \tag{3}$$

where, I_j is the relative importance of jth input factor, N_i is the number of the inputs ($N_i = 10$), N_h is the number of hidden neurons ($N_h = 15$), and W is the connection weight, in which the subscripts "k", "m" and "n" refer to input, hidden and output neurons, respectively, and the superscripts "i", "h" and "o" denote input, hidden and output layers, respectively.

The relative importance of every input regarding its contribution to the model output was also confirmed using the "Weights method", as reported by Gar Alalm and Nasr (2018). Fig. 1(d) indicated that Type 0041 had the highest impact (12.91%) on SVI, followed by *Gordonia* spp. (12.09%). This observation was followed by *N. limicola* (11.98%) and *Thiothrix* spp. (10.47%). The importance values of *S. natans*, Type 021N, and Type 1701 were almost similar. Moreover, all the percentages for the contribution of filaments to SVI were relevant, and thus, no input could be excluded during the multivariate statistical procedures.

3.5. Spatial variation of filaments using PCA

Fig. 5 shows a detailed investigation of the spatial distribution of filamentous bacteria in relation to the monitored WWTWs.

3.5.1. Dominance of N. limicola

As shown in Fig. 1(d), *N. limicola* had the third contribution to SVI, triggering serious operational problems. The filamentous bacterial morphotype "*N. limicola*" appeared on both -PC1 and -PC2 directions (Fig. 5a). Hence, the *N. limicola* filaments showed a high relative abundance in WWTPs I, II, III, and VI, and they increased due to either nutrient deficit or low F/M conditions (Table 3). Similar to previous works (Liu et al., 2000; Petrovski et al., 2012), the *N. limicola* morphotype has caused incidents of bulking and foaming in activated sludge systems. The coiled filament "*N. limicola*" contains coccal/discoid cells, splitting into three morphotypes with different cell dimensions, *viz*., I (small cells), II (medium cells), and III (large cells) (Seviour et al., 2006). Pajdak-Stós and Fiałkowska (2012) reported that *N. limicola*-like bacteria could be controlled through ingestion by *Lecane inermis* rotifers.

3.5.2. Dominance of Gordonia spp.

Fig. 1(d) indicated that *Gordonia* spp. were the second contributor to SVI, with a relative importance of 12.09%. These bulking-related species tend to reduce the specific gravity of the flocs and increase the SVI records (Insel et al., 2014). They

Table 3
Operating conditions causing filament overgrowth in WWTPs.

	Low DO	Nitrogen deficit	Low F/M	High SRT
Type 0092		✓	✓	
Type 1701	✓		✓	
Type 021N		✓		
Type 0041	✓		✓	✓
Thiothrix spp.	✓	✓		
Type 1851	✓			
Microthrix parvicella	✓		✓	
Nostocoida limicola		✓	✓	
Sphaerotilus natans	✓	✓		
Gordonia spp.				✓

could also entrap the gas bubbles and cover the surface of the biological reactors, known as hydrophobic characteristics (Kougias et al., 2014). As shown in the loading plot of PCA (Fig. 5a), *Gordonia* spp. located on the extreme direction of -PC1; hence, they were detected mostly in WWTWs I and II. This result could be linked to the longer sludge ages compared to other treatment plants, as previously reported by Richard (2003). In a similar study, Dunkel et al. (2018) found that Nocardiaform-like organisms (*Gordonia*) were the dominant bulking and foaming bacteria in industrial activated sludge plants. In addition, *Gordonia (Nocardia) amarae* could be excessively abundant in membrane bioreactor (MBR) systems operated for nutrient removal (Insel et al., 2014). *Gordonia* species could also physiologically thrive under anaerobic, anoxic, and aerobic environments at low DO levels (Insel et al., 2014). However, Zhang et al. (2019) found that the filamentous bacteria *Gordonia* were less dominant in oxidation ditch bulking sludge. In another study, Liu et al. (2015) revealed that phages could be applied to control foaming and bulking events caused by *Gordonia*, leading to improved sludge settling capabilities.

3.5.3. Dominance of Type 0041

The bacterium Type 0041 was the first contributor to SVI with a relative importance of 12.91% (Fig. 1d). In addition, this bacterium appeared in all WWTPs with various dominance levels, suggesting its ability to sustain under different environmental conditions. This finding complied with a previous study by Seviour et al. (1994), which indicated that the filament morphotype Type 0041 could be defined as "All zone growers". Similarly, Thomsen et al. (2002) suggested that Type 0041 was associated with the bulking and foaming events in anaerobic–anoxic–aerobic biological processes at nutrient-removing WWTPs. Moreover, the bacterium Type 0041 could favor high SRT, low DO, and low F/M, as depicted in WWTWs III, V, and IV (Fig. 5b).

3.5.4. Dominance of Type 1701

As shown in the PCA findings (Fig. 5a), Eikelboom Type 1701 distributed on the -PC1 axis. Hence, this filament was dominant in WWTPs I, II, and VI with the conditions of low DO by surface aerators and/or low F/M (Table 3). The filament Type 1701 is Neisser negative and Gram-negative (Thomsen et al., 2002), which has been detected with attached growth in activated sludge systems (see Fig. 3c).

3.5.5. Dominance of Thiothrix spp.

Thiothrix was detected in all WWTPs; however, it was dominated by the low TN condition in WWTWs IV and V (Fig. 5b). Moreover, *Thiothrix* recorded the fourth contributor to SVI with a relative importance of 10.47% (Fig. 1d). *Thiothrix* spp. have been linked to high SVI > 200 mL/g in activated sludge WWTPs treating a domestic wastewater source (Williams and Unz, 1985). In another work, Henriet et al. (2017) reported that the growth of *Thiothrix* was highly associated with a low DO condition of 1.4–4.0 mg/L. The existence of *Thiothrix* strains would cause sludge washout and a reduction in the biological performance of activated sludge systems. The application of some strategies such as (a) mixing tanks (anoxic or aerobic stages), (b) contact units with no aeration, or (c) aerobic selectors could limit the excessive growth of *Thiothrix* (Martins et al., 2004).

3.5.6. Dominance of S. natans

S. natans have Neisser negative, Gram-negative, and unbranched characteristics. In this study, *S. natans* were related to a subjective scoring <2 in WWTPs, suggesting "Few abundance" levels. This result could be because the WWTPs were operated at low F/M ratios of 0.11–0.18 kg_{COD}/kg_{VSS}/day; however, an F/M level about 1 kg_{COD}/kg_{VSS}/day would favor the growth of *S. natans* (Figueroa et al., 2015). Guo et al. (2012) found that *S. natans* were controlled and deactivated by adding a chlorine dosage over 20 mg of Cl_2 per g of suspended solids.

3.5.7. Dominance of Type 021N

Type 021N existed in WWTWs IV, V, and VII, which could be because this filamentous bacterium tends to grow under a nutrient deficit condition (i.e., Low TN) (Han et al., 2019). Type 021N has the ability to consume nitrate as an electron acceptor (Martins et al., 2004). Moreover, the dominance of Type 021N in these plants could probably be linked to the direct addition of septic sludge to the preliminarily treated wastewater (i.e., this claim requires further investigations). Guo et al. (2012) found that cetyltrimethyl ammonium bromide (CTAB) could be used as a biocidal agent to inactivate the chlorine-resistant Type 021N bacteria.

3.5.8. Dominance of Type 1851

Type 1851 bacterium is recognized as a Neisser negative and Gram-positive strain (Thomsen et al., 2002). In this study, Type 1851 appeared as "Few" filamentous in most WWTPs during high SVI data. Similarly, Thomsen et al. (2002) reported that the Eikelboom morphotype Type 1851 existed in only 4%–5% of WWTP samples. The proliferation of filamentous bacterium Eikelboom type 1851 could be largely associated with low DO levels (i.e., below 1.1 mg/L) (Nittami et al., 2019).

3.5.9. Dominance of Type 0092

Type 0092 (Gram-negative) exhibited "Abundant" and "Excessive" scores in most WWTPs (see images of Type 0092 in Fig. 3a). The excessive growth of Type 0092 has been linked to severe issues for solids separation in biological nutrient removal WWTPs (Madoni et al., 2000; Liu et al., 2015). Moreover, the Eikelboom morphotype Type 0092 has been defined as "All zone grower" due to its existence in anaerobic, anoxic, and aerobic zones (Speirs et al., 2009). Insel et al. (2014) suggested that limited aeration capacity (DO = 0.2–0.3 mg/L) could allow the growth of Type 0092 in aerobic units. However, the "Some" and "Common" abundance of Type 0092 in WWTPs I and II could be because these plants were treating COD using activated sludge configurations (i.e., this observation needs further studies).

3.5.10. Dominance of Candidatus Microthrix parvicella

In this work, *M. parvicella* appeared in most WWTP samples during high SVI events, especially during the winter season. *M. parvicella* is a gram-positive and unsheathed filamentous bacterium (Fig. 3b), which has been detected in several WWTPs in South Africa (Deepnarain et al., 2019). The growth of *M. parvicella* could be due to the low F/M (< 0.2 1/d) in nutrient removal WWTPs treating domestic wastewater (Eikelboom, 2000).

3.6. Future considerations for mitigating filamentous bulking sludge

In this study, an ANN model was employed for the prediction of SVI, which could be used as a reliable indicator to diminish and control the filamentous bulking problems. Additional input variables, e.g., operating conditions (temperature, sludge age, and hydraulic retention time), sludge morphological properties, and wastewater characteristics (BOD, N, and P), would be imported into the ANN model to improve its predictive usefulness. Further, the input variables of the modified ANN model would be adjusted to control the sludge bulking problem occurrence by limiting the SVI output to be lower than 150 mL/g. This objective can be efficiently attained by the ANN approach without having a deep understanding of the complete mechanisms and the relationships among the input attributes (Nasr, 2018).

Furthermore, in this work, different species of problematic filamentous bacteria were assessed in seven BNR systems treating primarily domestic wastewater. For this purpose, the PCA approach was employed to give comprehensive information on the causes of filament overgrowth in WWTPs, i.e., also summarized in Table 3. The clustering statistics depicted that floc-formers attempted to outcompete filamentous bacteria in the WWTWs IV and V, which could be associated with the use of an efficient aeration system (i.e., diffusion) in the three-stage Phoredox process. In a similar study, Abusam et al. (2017) applied the PCA and clustering techniques to identify the environmental conditions associated with the abundance of the filaments.

Interestingly, the application of multivariate statistical analysis along with FISH images in our study would provide a scientific basis to control the proliferation of filamentous bacteria in other WWTPs situated in similar climates to South Africa. Although the FISH method was able to visualize the filamentous bacteria, it is considered an expensive and time-consuming technique that employs targeted probes and requires a fluorescent microscope (Bokulich et al., 2012). Other molecular analysis methods/techniques, such as polymerase chain reaction (PCR), PCR followed by denaturing gradient gel electrophoresis (DGGE), qPCR, Illumine-MiSeq sequencing, and next generation sequencing (NGS), could be used for obtaining accurate and complete characterization of the microbial community. These techniques would be employed in future studies to improve the predictive ability of the mathematical models, leading to a better understanding and control of sludge bulking at full-scale WWTPs.

In a comparable study, Bagheri et al. (2015) applied ANN to predict SVI using 7 operating parameters, and they found that the artificial intelligence techniques could be used to optimize the weights and thresholds of the model. Similarly, Boztoprak et al. (2016) revealed that ANN had a high prediction accuracy of a SVI output using 33 input parameters obtained from automated image acquisition. Han et al. (2016) indicated that the inputs of MLSS, COD, TN, pH, and DO could be used to predict SVI values efficiently. Han et al. (2018) used a fuzzy neural network model for the prediction of sludge bulking, which would be a valuable option for the cases of poor sensor sensitivity and disturbance environment.

Other modeling techniques such as a logistic model (Deepnarain et al., 2015) and a regression tree model (Deepnarain et al., 2019) have also been employed to predict the proliferation of filamentous microorganisms at full-scale WWTPs.

Based on these studies, the advantage of our model is the determination of the most influential filament using the "Weights method" (see Fig. 1d), which could be used to improve the knowledge of sludge characteristics for further control of filamentous bulking. Moreover, the proposed models would provide an efficient tool to predict, simulate, and control sludge bulking in WWTWs having approximately similar influent wastewater properties and operational conditions. Moreover, the presented models were developed based on monitoring the sludge bulking events in seven WWTPs, giving a broad range of applicability in other wastewater treatment systems.

4. Conclusions

This study analyzed and assessed the microbial community associated with sludge bulking in seven WWTPs operated under different environmental conditions. The main conclusions are:

- Staining and microscopic methods, along with FISH images, revealed that filaments formed bridging among the flocs and enlarged the surface area, leading to a loose and poorly compacted structure.
- Subjective scores, based on the microbial investigation, succeeded in determining the occurrence, relative quantities, and dominance of filamentous bacteria.
- ANN with a structure of 10–15–1 was able to predict the SVI data using ten attributes of filament bacteria. The ANN results depicted that Eikelboom Type 0041 contributed to the highest impact on SVI, followed by *Gordonia* spp., *N. limicola*, and *Thiothrix* spp.
- Based on multivariate analysis, the distribution of filamentous species among seven WWTPs was adequately reported, signifying that the implementation of an efficient aeration system (i.e., diffusion) in the three-stage Phoredox process enhanced the settling properties of flocs. Clustering analysis also revealed the operating conditions that caused filament overgrowth in WWTPs.
- The outputs of this study are of growing importance to produce feasible solutions and ease of control for the operation of WWTPs experiencing frequent bulking problems.

CRediT authorship contribution statement

Nashia Deepnarain: Investigation, Data curation. **Mahmoud Nasr:** Formal analysis, Writing - original draft. **Sheena Kumari:** Validation, Writing - original draft. **Thor A. Stenström:** Writing - review & editing. **Poovendhree Reddy:** Methodology, Writing - original draft. **Kriveshin Pillay:** Investigation, Data curation. **Faizal Bux:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The support of the Water Research Commission (WRC), South Africa, for the successful completion of this project is gratefully acknowledged. The second author would like to acknowledge Nasr Academy for Sustainable Environment (NASE).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eti.2020.100853.

References

Abusam, A., Mydlarczyk, A., Al-Salameen, F., Ahmed, M.I., 2017. Identification of the most probable causes for filamentous bacteria over-proliferation in Riqqa wastewater treatment plant, Kuwait. Desalin. Water Treat. 72, 78–84.

Awolusi, O.O., Nasr, M., Kumari, S., Bux, F., 2018. Principal component analysis for interaction of nitrifiers and wastewater environments at a full-scale activated sludge plant. Int. J. Environ. Sci. Technol. 15 (7), 1477–1490.

Bagheri, M., Mirbagheri, S.A., Bagheri, Z., Kamarkhani, A.M., 2015. Modeling and optimization of activated sludge bulking for a real wastewater treatment plant using hybrid artificial neural networks-genetic algorithm approach. Process Saf. Environ. Prot. 95, 12–25.

Beer, M., Seviour, E.M., Kong, Y., Cunningham, M., Blackall, L.L., Seviour, R.J., 2002. Phylogeny of the filamentous bacterium Eikelboom Type 1851, and design and application of a 16S rRNA targeted oligonucleotide probe for its fluorescence in situ identification in activated sludge. FEMS Microbiol. Lett. 207 (2), 179–183.

Björnsson, L., Hugenholtz, P., Tyson, G.W., Blackall, L.L., 2002. Filamentous Chloroflexi (green non-sulfur bacteria) are abundant in wastewater treatment processes with biological nutrient removal. Microbiology 148 (8), 2309–2318.

Bokulich, N.A., Bamforth, C.W., Mills, D.A., 2012. A review of molecular methods for microbial community profiling of beer and wine. J. Am. Soc. Brew. Chem. 70 (3), 150–162.

Boztoprak, H., Özbay, Y., Güçlü, D., Küçükhemek, M., 2016. Prediction of sludge volume index bulking using image analysis and neural network at a full-scale activated sludge plant. Desalin. Water Treat. 57 (37), 17195–17205.
Daims, H., Brühl, A., Amann, R., Schleifer, K.-H., Wagner, M., 1999. The domain-specific probe EUB338 is insufficient for the detection of all bacteria: Development and evaluation of a more comprehensive probe set. Syst. Appl. Microbiol. 22 (3), 434–444.
Deepnarain, N., Kumari, S., Ramjith, J., Swalaha, F.M., Tandoi, V., Pillay, K., Bux, F., 2015. A logistic model for the remediation of filamentous bulking in a biological nutrient removal wastewater treatment plant. Water Sci. Technol. 72 (3), 391–405.
Deepnarain, N., Nasr, M., Kumari, S., Stenström, T.A., Reddy, P., Pillay, K., Bux, F., 2019. Decision tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant. Process Saf. Environ. Prot. 126, 25–34.
Dunkel, T., de León Gallegos, E.L., Bock, C., Lange, A., Hoffmann, D., Boenigk, J., Denecke, M., 2018. Illumina sequencing for the identification of filamentous bulking and foaming bacteria in industrial activated sludge plants. Int. J. Environ. Sci. Technol. 15 (6), 1139–1158.
Eikelboom, D., 2000. Process Control of Activated Sludge Plants by Microscopic Investigation. IWA publishing.
Eikelboom, D.H., Andreadakis, A., Andreasen, K., 1998. Survey of filamentous populations in nutrient removal plants in four european countries. Water Sci. Technol. 37, 281–289.
Erhart, R., Bradford, D., Seviour, R.J., Amann, R., Blackall, L.L., 1997. Development and use of fluorescent in situ hybridization probes for the detection and identification of '*Microthrix parvicella*' in activated sludge. Syst. Appl. Microbiol. 20 (2), 310–318.
Fan, N.-S., Qi, R., Huang, B.-C., Jin, R.-C., Yang, M., 2020. Factors influencing *Candidatus Microthrix parvicella* growth and specific filamentous bulking control: A review. Chemosphere 244, 125371.
Fawzy, M., Nasr, M., Nagy, H., Helmi, S., 2018. Artificial intelligence and regression analysis for Cd(II) ion biosorption from aqueous solution by Gossypium barbadense waste. Environ. Sci. Pollut. Res. 25 (6), 5875–5888.
Figueroa, M., Val del Río, A., Campos, J.L., Méndez, R., Mosquera-Corral, A., 2015. Filamentous bacteria existence in aerobic granular reactors. Biotechnol. Biotechnol. Equip. 38 (5), 841–851.
Gar Alalm, M., Nasr, M., 2018. Artificial intelligence, regression model, and cost estimation for removal of chlorothalonil pesticide by activated carbon prepared from casuarina charcoal. Sustain. Environ. Res. 28 (3), 101–110.
Garson, G., 1991. Interpreting neural network connection weights. AI Expert 6 (4), 46–51.
Graveleau, L., Cotteux, E., Duchène, P., 2005. Bulking and foaming in France: The 1999-2001 survey. Acta Hydrochim. Hydrobiol. 33 (3), 223–231.
Guo, J., Peng, Y., Wang, Z., Yuan, Z., Yang, X., Wang, S., 2012. Control filamentous bulking caused by chlorine-resistant Type 021N bacteria through adding a biocide CTAB. Water Res. 46 (19), 6531–6542.
Han, H.-G., Li, Y., Guo, Y.-N., Qiao, J.-F., 2016. A soft computing method to predict sludge volume index based on a recurrent self-organizing neural network. Appl. Soft Comput. J. 38, 477–486.
Han, H., Liu, z., Ge, L., Qiao, J., 2018. Prediction of sludge bulking using the knowledge-leverage-based fuzzy neural network. Water Sci. Technol. 77 (3), 617–627.
Han, H.-G., Liu, H.-X., Liu, Z., Qiao, J.-F., 2019. Fault detection of sludge bulking using a self-organizing type-2 fuzzy-neural-network. Control Eng. Pract. 90, 27–37.
Henriet, O., Meunier, C., Henry, P., Mahillon, J., 2017. Filamentous bulking caused by Thiothrix species is efficiently controlled in full-scale wastewater treatment plants by implementing a sludge densification strategy. Sci. Rep. 7 (1430).
Insel, G., Erol, S., Övez, S., 2014. Effect of simultaneous nitrification and denitrification on nitrogen removal performance and filamentous microorganism diversity of a full-scale MBR plant. Bioprocess. Biosyst. Eng. 37 (11), 2163–2173.
Jenkins, D., Richard, M.G., Diagger, G.T., 1986. Manual on the Causes and Control of Activated Sludge Bulking and Foaming. Water Research Commission, Pretoria, Republic of South Africa.
Jin, B., Wilén, B., Lant, P., 2003. A comprehensive insight into floc characteristics and their impact on compressibility and settleability of activated sludge. Chem. Eng. J. 95 (1–3), 221–234.
Kanagawa, T., Kamagata, Y., Aruga, S., Kohno, T., Horn, M., Wagner, M., 2000. Phylogenetic analysis of and oligonucleotide probe development for Eikelboom type 021N filamentous bacteria isolated from bulking activated sludge. Appl. Environ. Microbiol. 66 (11), 5043–5052.
Kougias, P.G., Boe, K., O-Thong, S., Kristensen, L.A., Angelidaki, I., 2014. Anaerobic digestion foaming in full-scale biogas plants: A survey on causes and solutions. Water Sci. Technol. 69 (4), 889–895.
Lee, S., Basu, S., Tyler, C.W., Pitt, P.A., 2003. A survey of filamentous organisms at the deer island treatment plant. Environ. Technol. 24 (7), 855–865.
Lemmer, H., Lind, G., Müller, E., Schade, M., 2005. Non-famous scum bacteria: Biological characterization and troubleshooting. Acta Hydrochim. Hydrobiol. 33 (3), 197–202.
Liu, J.R., Burrell, P., Seviour, E.M., Soddell, J.A., L.L. Blackall Seviour, R.J., 2000. The filamentous bacterial morphotype ' Nostocoida limicola' I contains at least two previously described genera in the low G+C gram positive bacteria. Syst. Appl. Microbiol. 23 (4), 528–534.
Liu, M., Gill, J.J., Young, R., Summer, E.J., 2015. Bacteriophages of wastewater foaming-associated filamentous Gordonia reduce host levels in raw activated sludge. Sci. Rep. 5 (13754).
Liu, J.R., Seviour, R.J., 2001. Design and application of oligonucleotide probes for fluorescent in situ identification of the filamentous bacterial morphotype *Nostocoida limicola* in activated sludge. Environ. Microbiol. 3 (9), 551–560.
Madoni, P., Davoli, D., Gibin, G., 2000. Survey of filamentous microorganisms from bulking and foaming activated- sludge plants in Italy. Water Res. 34 (6), 1767–1772.
Martins, A.M.P., Pagilla, K., Heijnen, J.J., Van Loosdrecht, M.C.M., 2004. Filamentous bulking sludge - A critical review. Water Res. 38 (4), 793–817.
Nasr, M., 2018. Modeling applications in environmental bioremediation studies. In: Kumar, V., Kumar, M., R., Prasad (Eds.), Phytobiont and Ecosystem Restitution. Springer, Singapore, pp. 143–160.
Nguyen, L.N., Commault, A.S., Johir, M.A.H., Bustamante, H., Aurisch, R., R. Lowrie Nghiem, L.D., 2019. Application of a novel molecular technique to characterise the effect of settling on microbial community composition of activated sludge. J. Environ. Manag. 251, 109594.
Nielsen, P.H., Daims, H., Lemmer, H., Arslan-Alaton, I., 2009. FISH Handbook for Biological Wastewater Treatment. IWA publishing, London, UK.
Nittami, T., Shoji, T., Koshiba, Y., Noguchi, M., Oshiki, M., Kuroda, M., Kindaichi, T., Fukuda, J., Kurisu, F., 2019. Investigation of prospective factors that control Kouleothrix (Type 1851) filamentous bacterial abundance and their correlation with sludge settleability in full-scale wastewater treatment plants. Process Saf. Environ. Prot. 124, 137–142.
Pajdak-Stós, A., Fiałkowska, E., 2012. The influence of temperature on the effectiveness of filamentous bacteria removal from activated sludge by rotifers. Water Environ. Res. 84 (8), 619–625.
Petrovski, S., Tillett, D., Seviour, R.J., 2012. Isolation and complete genome sequence of a bacteriophage lysing tetrasphaera jenkinsii, a filamentous bacteria responsible for bulking in activated sludge. Virus Genes 45 (2), 380–388.
Richard, M., 2003. Activated sludge microbiology problems and their control. In: Proceedings of the 20th Annual US Environmental Protection Agency National Operator Trainers Conference; nBuffalo, New York, Jun 8. US Environmental Protection Agency, Washington, DC.
Seviour, E.M., Eales, K., Izzard, L., Beer, M., Carr, E.L., Seviour, R.J., 2006. The in situ physiology of Nostocoida limicola II, a filamentous bacterial morphotype in bulking activated sludge, using fluorescence in situ hybridization an microautoradiography. Water Sci. Technol. 54 (1), 47–53.
Seviour, E.M., Williams, C., DeGrey, B., Soddell, J.A., Seviour, R.J., Lindrea, K.C., 1994. Studies on filamentous bacteria from australian activated sludge plants. Water Res. 28 (11), 2335–2342.

Snaidr, I., Beimfohr, C., Levantesi, C., Rossetti, S., van der Waarde, J., Geurkink, B., Elkelboom, D., Lemaitre, M., Tandoi, V., 2002. Phylogenetic analysis and in situ identification of nostocoida limicola-like filamentous bacteria in activated sludge from industrial wastewater treatment plants. Water Sci. Technol. 46 (1–2), 99–104.

Speirs, L.B.M., Dyson, Z.A., Tucci, J., Seviour, R.J., 2017. Eikelboom filamentous morphotypes 0675 and 0041 embrace members of the Chloroflexi: Resolving their phylogeny, and design of fluorescence in situ hybridisation probes for their identification. FEMS Microbiol. Ecol. 93 (10), fix115.

Speirs, L., Nittami, T., McIlroy, S., Schroeder, S., Seviour, R.J., 2009. Filamentous bacterium eikelboom type 0092 in activated sludge plants in Australia is a member of the phylum chloroflexi. Appl. Environ. Microbiol. 75 (8), 2446–2452.

Thomsen, T.R., Kjellerup, B.V., Nielsen, J.L., Hugenholtz, P., Nielsen, P.H., 2002. In situ studies of the phylogeny and physiology of filamentous bacteria with attached growth. Environ. Microbiol. 4 (7), 383–391.

Wagner, M., Amann, R., Kämpfer, P., Assmus, B., Hartmann, A., Hutzler, P., Springer, N., Schleifer, K.-H., 1994. Identification and in situ detection of gram-negative filamentous bacteria in activated sludge. Syst. Appl. Microbiol. 17 (3), 405–417.

Williams, T.M., Unz, R.F., 1985. Isolation and characterization of filamentous bacteria present in bulking activated sludge. Appl. Microbiol. Biotechnol. 22 (4), 273–282.

You, S.J., Sue, W.M., 2009. Filamentous bacteria in a foaming membrane bioreactor. J. Membr. Sci. 342 (1–2), 42–49.

Zhang, M., Yao, J., Wang, X., Hong, Y., Chen, Y., 2019. The microbial community in filamentous bulking sludge with the ultra-low sludge loading and long sludge retention time in oxidation ditch. Sci. Rep. 9 (1), 13693.

APPENDIX FOUR: DECISION TREE FOR IDENTIFICATION AND PREDICTION OF FILAMENTOUS BULKING AT FULL-SCALE ACTIVATED SLUDGE WASTE WATER TREATMENT PLANT

Process Safety and Environmental Protection 126 (2019) 25–34

Contents lists available at ScienceDirect

Process Safety and Environmental Protection

journal homepage: www.elsevier.com/locate/psep

Decision tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant

Nashia Deepnarain[a], Mahmoud Nasr[b], Sheena Kumari[a], Thor A. Stenström[a], Poovendhree Reddy[c], Kriveshin Pillay[a], Faizal Bux[a,*]

[a] Institute for Water and Wastewater Technology, Durban University of Technology, Durban, 4000, South Africa
[b] Sanitary Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, 21544, Egypt
[c] Department of Community Health Studies, Faculty of Health Sciences, Durban University of Technology, South Africa

ARTICLE INFO

Article history:
Received 4 April 2018
Received in revised form 19 February 2019
Accepted 26 February 2019
Available online 12 March 2019

Keywords:
Classification and regression trees
Filamentous detection
Sludge volume index
Sludge bulking
Wastewater characteristics

ABSTRACT

This study attempted to model sludge bulking in a full-scale wastewater treatment plant operated as a 3-stage Phoredox process. Principal component analysis and a regression tree model were employed to describe the correlations between influent wastewater characteristics and operational conditions (as inputs) and sludge volume index (SVI) as an output. A classification tree model was used to determine the environmental factors that affected the proliferation of filamentous microorganisms. Fluorescent *in situ* hybridisation analysis identified filamentous species of *Microthrix parvicella*, *Thiothrix* I & II, and Eikelboom Types 0041, 0092, and 021 N. It was found that SVI increased with an increment in sludge retention time, but it negatively correlated with soluble chemical oxygen demand (sCOD) and ammonium-nitrogen. The dominance of *Microthrix parvicella* was observed with a decline in temperature below 15.5 °C, causing an increase in SVI during the winter and spring seasons. The overgrowth of *Thiothrix* could be linked to the unbalanced ratio between readily biodegradable COD and nutrient species. The filament Type 0092 contributed to high SVI, and it prevailed with a decrease in food-to-microorganisms ratio below 0.08 1/d. Based on the satisfactory training, validation, and generalization procedures, the proposed models could be applied for the prediction of sludge bulking episodes.

1. Introduction

Activated sludge process is a reasonably well-established technology that has been widely applied for the treatment of different sources of domestic and industrial wastewater (Williams and Unz, 1985; Araújo dos Santos et al., 2015). The system is mainly composed of aeration and anoxic biological units, which hold various types of microorganisms, coupled with secondary settling tanks (SSTs) (Madoni et al., 2000). The SSTs are used to separate treated wastewater from the solid particles via gravitational settling so that the clarified effluent can be safely discharged into the receiving environments (Comas et al., 2008). The effectiveness of sludge settleability in SSTs is considered one of the important features that identify the functionality of activated sludge systems (Lou and Zhao, 2012). However, the overgrowth of undesirable microorganisms resulting from inadequate system operation and/or unbalanced carbon and nutrient ratios can reduce the solid precipitation efficiency (Mielczarek et al., 2012). Poor sludge settleability can block the piping system, reduce the oxygen transfer efficiency in mixed liquor, deteriorate the performance of the biological processes, and cause a carryover of solid particles with the final effluent (Martins et al., 2004). Hence, proper statistical models should be employed to assess and evaluate the settling performance in activated sludge systems.

Activated sludge flocs have a heterogeneous structure that contains a diversity of microorganisms in addition to organic and inorganic particles (Ovez et al., 2006). Floc-forming bacteria such as nitrifiers, denitrifiers, and phosphate-accumulating organisms (PAOs) can bond together by extracellular polymeric matters, and form bridges with the neighboring microorganisms (Araújo dos Santos et al., 2015). The biological treatment units can also contain non-floc forming microorganisms, mostly filamentous bacteria, which contribute to over 90% of the sludge bulking scenarios (Lacko et al., 1999). The development of well-settling sludge (dense and strong flocs) is influenced by a balanced growth between floc-forming and filamentous bacteria (Wágner et al., 2015). Sludge bulking occurs when a large number of filamentous microorgan-

* Corresponding author at: Institute for Water and Wastewater Technology, Durban University of Technology, P.O. Box 1334, Durban, 4000, South Africa.
E-mail address: faizalb@dut.ac.za (F. Bux).

https://doi.org/10.1016/j.psep.2019.02.023
0957-5820/

isms generates inter-floc bridging, reduces biomass density, and diffuses bio-flocs having poor precipitation properties (Dunkel et al., 2016). Filamentous foaming is caused by the establishment of stable foam layers or a thick viscous scum on the surface of aerobic basins and clarifiers (Jenkins et al., 2003). Ganidi et al. (2009) defined foaming as a floating biomass layer occurring due to the presence of a high abundance of filamentous bacteria having hydrophobic cell surfaces. The condition of poor sludge settleability is characterized by a high sludge volume index (SVI), which can exceed 120 mL/g (Bagheri et al., 2015).

Recently, various efforts have been exerted to determine the appropriate statistical and computational modelling techniques that can evaluate and predict sludge bulking scenarios (Comas et al., 2008; Lou and Zhao, 2012; Bagheri et al., 2015; Deepnarain et al., 2015; Han et al., 2016). Multivariate statistical tools such as principal component analysis (PCA) can be employed to describe the correlations between wastewater characteristics and SVI in activated sludge systems. For example, Lou and Zhao (2012) employed PCA to predict SVI using several inputs such as temperature, chemical oxygen demand (COD), ammonium-nitrogen (NH_4^+-N), total nitrogen (TN), and total phosphorus (TP). PCA can extract useful information from large datasets and describe complex relationships between input attributes and target variables (Nasr and Zahran, 2016). PCA is used to transform and reduce a complex multi-dimensional system into orthogonal variables, known as principal components (PCs) (Jolliffe, 2002). PCs having eigenvalues greater than 1 can be used to represent a high percentage of total variation in the investigated variables (Lou and Zhao, 2012).

Decision tree technique can also be applied to predict sludge bulking events, and to select the most relevant input variables. Decision tree is considered a data mining method that employs a flowchart-like tree structure for partitioning a set of data into discrete sub-groups or nodes (Yu et al., 2010). The flowchart structure of decision tree consists of root, intermediate, and leaf nodes (Kim et al., 2001). The root node, which is the most important predictor with respect to the target variable, partitions all records into two or multiple subgroups (Breiman et al., 1984). The intermediate node expresses a condition or a binary split test for each independent variable; i.e., the top edge is linked to a parent node, whereas the bottom edge is associated with child or leaf nodes (Kim et al., 2001). The leaf node represents the result of a combination of decisions, being categorical outcomes for classification trees and numerical results for regression trees. Decision trees also contain branches (i.e., chance outcomes), which are emanated from both root and internal nodes (Loh and Shih, 1997). Decision tree provides adequate categorization, generalization, and prediction of a given population using a set of if-then rules. Decision tree can deal with undefined observations in the dataset by classifying the records into independent branches, and hence, the entire dataset is used in the investigation (Yeo and Grant, 2018). Decision tree can also display both linear and non-linear relationships of large datasets in simple-to-interpret visualizations. These benefits tend to improve the predictive accuracy of both continuous and categorical target variables.

Previous studies (Wágner et al., 2015; Dunkel et al., 2016) have investigated the ecophysiological traits and characteristics of filamentous microorganisms using laboratory strategies. However, laboratory scale might be unreliable to cultivate various filamentous species (Mielczarek et al., 2012). Hence, in this study, a full-scale wastewater treatment plant (WWTP) with an activated sludge system situated in Gauteng, South Africa was monitored for two years. Several variables that described the seasonal variation in influent wastewater characteristics, effluent quality, operational conditions, and filamentous bacteria in the aeration tanks were recorded. PCA and decision tree techniques were employed to evaluate and model the filamentous sludge bulking and to determine the factors that controlled the proliferation of filaments. The model input factors contained pH, temperature, dissolved oxygen (DO), sludge retention time (SRT), food-to-microorganisms (F/M) ratio, soluble COD (sCOD), total COD (tCOD), NH_4^+-N, total Kjeldahl nitrogen (TKN), phosphorus as phosphate (PO_4^{3-}-P), TP, and total suspended solids (TSS). These parameters were selected because they significantly varied over time ($p < 0.05$). Fluorescent *in situ* hybridisation (FISH) analysis was used to determine the abundance levels of various filamentous species such as *Microthrix parvicella*, *Thiothrix* I & II, and Eikelboom Types 0041, 0092, and 021 N. The environmental factors that affected the proliferation of these species were also determined.

2. Materials and methods

2.1. Wastewater treatment facilities

A WWTP located in Gauteng, South Africa was monitored on a monthly basis for 2 years (i.e., 2015 and 2016). The WWTP was employed for the treatment of domestic wastewater in addition to 5% industrial effluents. The feed wastewater was screened, degritted, and collected in balancing tanks. The biological treatment system was operated as a 3-stage Phoredox process containing anaerobic (29%), anoxic (29%), and aerobic (42%) zones. The mixed liquor discharged from the bioreactors was subjected to quiescent conditions in the SSTs. The WWTP was operated using return activated sludge (RAS) from the underflow of the SSTs to the head of the biological basins, and internal recycle (IR) of mixed liquor. The waste activated sludge (WAS) was removed from the treatment units and pumped to the dissolved air flotation (DAF) units for thickening.

2.2. Sampling

Through the monitoring campaign, wastewater samples were collected at appropriate points (inlet and outlet) to determine the plant performance. The effluent quality was compared with the allowable discharge standards reported by the Department of Water Affairs of South Africa (DWA, 1998). The samples were preserved at a temperature below 4 °C in a portable icebox containing ice packs and transported to the laboratory within 2 h. Composite samples of the mixed liquor were collected from the beginning, middle, and end of aeration tanks using an aluminum dipper (1 L) connected to a wooden handle. About 1/3 of the sample container was kept empty to ensure the survival of microorganisms under the aerobic conditions. To avoid the growth/decay of microbial cells after harvesting, samples were subjected to fixation, cell wall permeabilization, and hybridization following the protocols of Nielsen et al. (2009). The fixed samples were kept in the freezer (-20 °C) for further analyses.

2.3. Bacterial analysis

2.3.1. Identification of filamentous bacteria

The morphological classification and microscopic examination of filamentous bacteria were determined using staining characteristics, as specified by Eikelboom (2000) and Jenkins et al. (2003). The morphological traits including floc size, shape, and structure were determined using wet mount examination, Gram and Neisser staining techniques, cellular inclusions of poly-β-hydroxybutyrate (PHB), and Sulphur storage test (Jenkins et al., 2003). Gram staining technique (Adamse, 1970) was employed to distinguish between Gram-positive and Gram-negative filamentous bacteria. The filamentous abundance were categorically ranked as follows: “None”, “Few”, “Some”, “Common”, “Very common”, “Abundant”, and “Excessive” (Table 1). For instance, filaments are not found at

Table 1
Classification of the abundance of filamentous microorganisms.

Abundance	Description
None	Filaments are not detected.
Few	Filaments are only detected in an occasional floc.
Some	Filaments are commonly found (not present in all flocs).
Common	Filaments are commonly observed in all flocs, but in low densities that range between 1 to 5 filaments per floc.
Very common	Filaments exist in all flocs at medium density (5 - 20 filaments per floc)
Abundant	Filaments occur in all flocs at high density; greater than 20 filaments per floc.
Excessive	Filaments grow in high abundance in mixed liquor (e.g., filaments are greater than flocs).

the "None" classification, whereas sludge bulking occurs at "Abundant", and "Excessive" categories (Jenkins et al., 2003). This scale is used to provide relevant information about the tendency of bulking in activated sludge systems.

2.3.2. *Fluorescent* in situ *hybridisation (FISH) analysis*

FISH analysis was employed on a monthly basis to confirm the identification of filamentous bacteria. The FISH procedures were performed based on the guidelines described by Nielson et al. (2009). Before FISH analysis, the samples were subjected to sonication (2 W for 45 s) using an XL-2000 ultrasonic liquid processor (Tri-Lab Support, USA) to detach cells from particles or aggregates. Subsequently, the samples were pre-treated with lysozyme (5 mg/L) and incubated at 25 °C for 20 min. Table 2 lists the specific 16S rRNA probes used for targeting the filamentous organisms during hybridisation (Amann et al., 1990; Daims et al., 1999; Kanagawa et al., 2000; Gich et al., 2001; Speirs et al., 2009). The probes were labelled with 5(6)-Carboxyfluorescein ($C_{21}H_{12}O_7$), also known as 5(6)-FAM, and synthesised at Inqaba Biotechnical Industries (Pty) Ltd, South Africa. The EUB 338 probe, which is a general probe for most bacteria, was employed as the positive control, whereas hybridization without probe was applied as the negative control. The procedures of hydration and washing were outlined as per Nielsen et al. (2009). For slides preparation, the DNA was counterstained in 0.25 μg/mL 4′,6-diamidino-2-phenylindole (DAPI) for 5 min. The slides were visualized under an AxioLab ApoTome microscope (Carl Zeiss, Germany) containing FLUOS Fluorochrome/Filter Set. Image analyses were performed by AxioVision imaging software (Version 4.6, Carl Zeiss Microimaging GmbH, Germany).

2.4. *Analytical analysis*

The influent flowrate, pH, temperature, and DO were recorded by the City of Tshwane using a supervisory control and data acquisition (SCADA) system. The nitrogen and phosphorus species were analyzed by Aquakem Gallery Photometric Auto-analyser (Thermo Scientific, Germany). The analyses were conducted according to the protocols outlined by the US Environmental Protection Agency (EPA) (USEPA, 1983). Other parameters such as COD and mixed liquor suspended solids (MLSS) were measured according to the standard methods of the American Public Health Association (APHA) (APHA, 1998).

2.5. *SVI calculation*

Sludge volume index (SVI) was considered in this study as a typical indicator for the assessment of sludge settling performance. SVI is defined as the volume (in mL) occupied by 1 g of activated sludge after precipitating the aerated mixed liquor for 30 min (Martins et al., 2004). Low SVI values indicate that the activated sludge is dense, and thus, the biomass has high settleability in the final clarifier (Lou and Zhao, 2012). A mixed liquor having SVI lower than 120 mL/g is considered satisfactory, whereas 150 mL/g is considered as the threshold for sludge bulking (Deepnarain et al., 2015). SVI was calculated using Eq. 1 according to Jenkins et al. (2003).

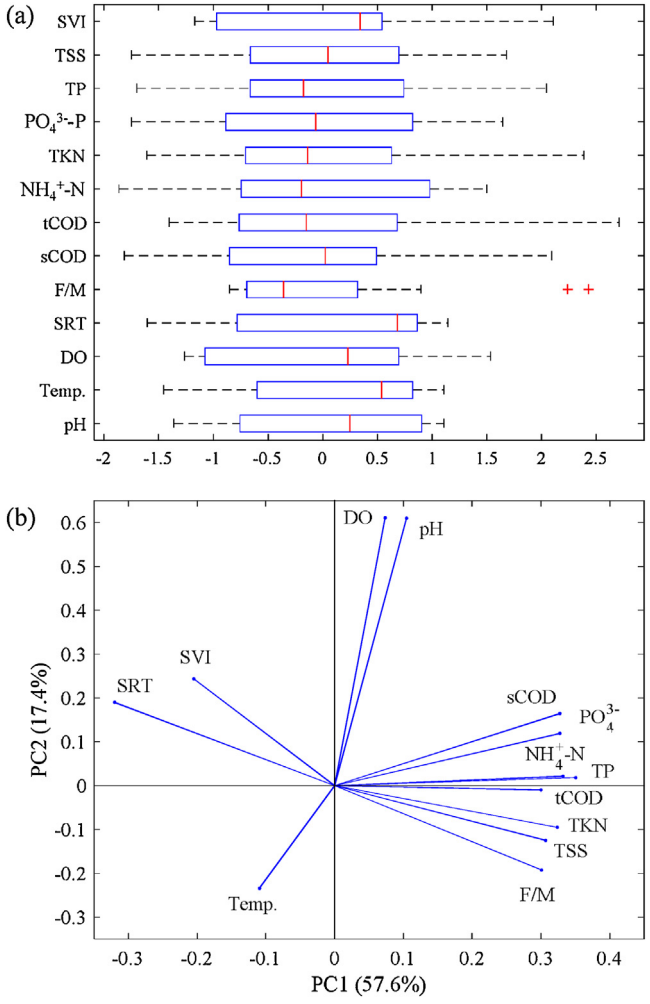


Fig. 1. (a) Normalized variables in Box-and-Whisker plot, and (b) loading plot of PCA.

$$SVI = \frac{V \times 1000}{MLSS} \quad (1)$$

where, V is the settled volume of sludge after 30 min (mL/L), 1000 is a conversion factor (mg/g), and MLSS is the mixed liquor suspended solids (mg/L).

2.6. *Statistical analysis*

2.6.1. *Principal component analysis*

PCA is designed to transform the original variables into new and orthogonal axes, called principal components (PCs) (Nasr and Zahran, 2016). The PCs represent the directions of the maximum variance, and thus, provide information on the most meaningful parameters. The eigenvalues define the associated variance of the PCs, whereas the loadings are the weights of the original variables in each PC. PCA involves the following steps (Jolliffe, 2002): (a) all parameters are z-scale standardized with zero mean and unit variance (see Fig. 1a), that is, to ensure that all the measurements have equal weight in the analysis, (b) the correlation matrix is calculated from Eq. 2,(c) the eigenvectors and eigenvalues are estimated, and the eigenvalues are sorted in a descending order, (d) the eigenvector with the highest eigenvalue represents the most dominant

Table 2
Oligonucleotide probes and target microorganisms for FISH analysis of filamentous bacteria.

Probe	Sequence (5'-3')	Target microorganism	Reference
EUB 338	GCTGCCTCCCGTAGGAGT	Most bacteria	Amann et al. (1990)
EUB338-II	GCAGCCACCCGTAGGTGT	*Plantomycetales*	Daims et al. (1999)
EUB338-III	GCTGCCACC CGTAGGTGT	*Verrucomicrobiales*	Daims et al. (1999)
G123T	CCT TCC GAT CTC TAT GCA	*Thiothrix* species: *T. nivea, T. unzii, T. fructosivorans, and T. defluvii.* Type 021 N group I, II, III	Kanagawa et al. (2000)
G1B	TGTGTTCGAGTTCCTTGC	Type 021 N group I	Kanagawa et al. (2000)
GNS B941	AAACCACACGCTCCGCT	Type 0041, and Type 0675	Gich et al. (2001)
CFX 1223	GGTGCTGGCTCC TCCCAG	Type 0092	Speirs et al., 2009

principal component of the data set, namely PC1, and (e) the second component (PC2) is computed under the constraint of being orthogonal to PC1 and having the second largest variance. The PCA calculations were performed using the functions "pca" and "pcacov" in MATLAB software (R2015a).

$$r_{x,y} = \frac{\sum_{i=1}^{n} (x_i - \mu_x)\left(y_i - \mu_y\right)}{(n-1)\,\sigma_x \sigma_y} \quad (2)$$

where, μ_x and μ_y are the sample means of X and Y, respectively, and σ_x and σ_y are the sample standard deviations of X and Y, respectively.

2.6.2. Decision tree method

Decision trees classify instances or examples into branch-like segments by taking paths from the root node through internal nodes to leaf nodes. In this method, the root node implies the predictor that gives the best split of the target class values (Yu et al., 2010). Further, splitting is applied iteratively to the subgroups (internal nodes or child partitions) until leaf nodes are obtained. Each internal node contains splits and holds two or more child nodes (Loh and Shih, 1997). Pruning is used to avoid overfitting problems by removing unnecessary nodes and optimizing the tree size. The pruning process can either be pre-pruning by preventing the generation of unimportant branches or post-pruning by removing branches from the established tree (Breiman et al., 1984). The decision tree process is completed when (a) the class label of the leaf node has the same target class value, (b) every predictor has already been used to split a partition, and (c) no more records for a particular value of a predictor variable.

The learning procedure of decision trees is achieved via the training and validation steps. The training process is used to build and evaluate the decision tree model by minimizing the difference between the measured and predicted outputs. The validation step is employed to determine the suitable tree size that avoids data overfitting and the loss of generalization ability. These procedures are applied to estimate the accuracy of the decision tree by comparing predicted outputs with actual data. The optimal final model is achieved when the accuracy becomes acceptable, and thus, the decision tree can be employed for classification and prediction purposes using a new dataset (Yu et al., 2010).

In this study, the decision tree model was used to (a) elucidate the relationships between influent wastewater characteristics and operational conditions (as inputs) and SVI (as an output), and (b) define the environmental factors that mainly influenced the dominance of filamentous species. The function "fitctree" in MATLAB (R2015a) was employed to estimate the decision trees based on the input variables (i.e., predictors). The function employs the Classification and Regression Trees (CART) algorithm for building the decision tree by the binary split at each node (Breiman et al., 1984). The algorithm was used due to its flexibility and applicability to various sorts of data. The input attributes that obtain a high degree of purity of the child nodes are selected using the Gini index method for classification trees and mean squared error for regression trees (Loh and Shih, 1997). This function determines the optimal sequence of pruned subtrees, namely pre-pruning.

Table 3
Operational conditions, influent wastewater characteristics, and corresponding removal efficiencies for WWTP located in Gauteng in South Africa.

	Minimum	Maximum	Average	Std.
pH	7.3	7.6	7.5	0.1
Temp. (°C)	14.0	23.0	19.1	3.5
DO (mg/L)	2.5	4.0	3.1	0.5
SRT (d)	9.8	21.6	16.7	4.3
F/M (1/d)	0.1	0.4	0.2	0.1
sCOD (mg/L)	27.3	105.6	63.7	20.0
tCOD (mg/L)	71.3	301.8	149.9	56.1
NH_4^+-N (mg/L)	6.0	18.0	12.7	3.6
TKN (mg/L)	9.4	30.7	18.0	5.3
PO_4^{3-}-P (mg/L)	0.4	1.2	0.8	0.3
TP (mg/L)	0.8	3.2	1.9	0.6
TSS (mg/L)	20.5	112.1	67.2	26.7
tCOD removal (%)	60.7	89.2	76.7	8.5
NH_4^+-N removal (%)	93.4	99.2	97.5	1.5
TKN removal (%)	48.3	95.0	85.7	11.5
PO_4^{3-}-P removal (%)	6.2	94.1	75.9	20.5
TP removal (%)	14.6	90.8	73.8	18.8
TSS removal (%)	69.8	97.5	88.6	7.2

3. Results and discussion

3.1. WWTP performance

The investigated activated sludge system was designed for the treatment of a dry weather flowrate that varied from 20.8×10^3 to 35.2×10^3 m^3/d and operated at a SRT of 9.8–21.6 days. Table 3 lists the measured ranges of wastewater characteristics and the corresponding removal efficiencies obtained during the sampling period. These variables were selected from the database as they covered the organic and nutrient properties of the influent wastewater. The high variation in wastewater characteristics could be because the samples were collected throughout the year (Dunkel et al., 2016). In addition, the study period of two years was satisfactory as it represented all the probable seasonal variations in the wastewater parameters.

The influent wastewater contained tCOD of 149.9 ± 56.1 mg/L, resulting in an organic load of $(1.7–10.0) \times 10^3$ kg/d. Under this condition, the treatment facilities achieved a tCOD removal efficiency of 60.7–89.2%. The effluent tCOD fluctuated between 24.5 and 39.5 mg/L, which complied with the National Water Act (Act No 36 of 1998) of 75 mg-COD/L as general standards and 30 mg-COD/L as specific standards (DWA, 1998). The sCOD/tCOD ratios ranged between 0.31 and 0.65, suggesting that about 85% of data had a readily biodegradable substrate = 54%. The nitrogen loads recorded 364 ± 158 kg-NH_4^+/d and 517 ± 233 kg-TKN/d and the removal effi-

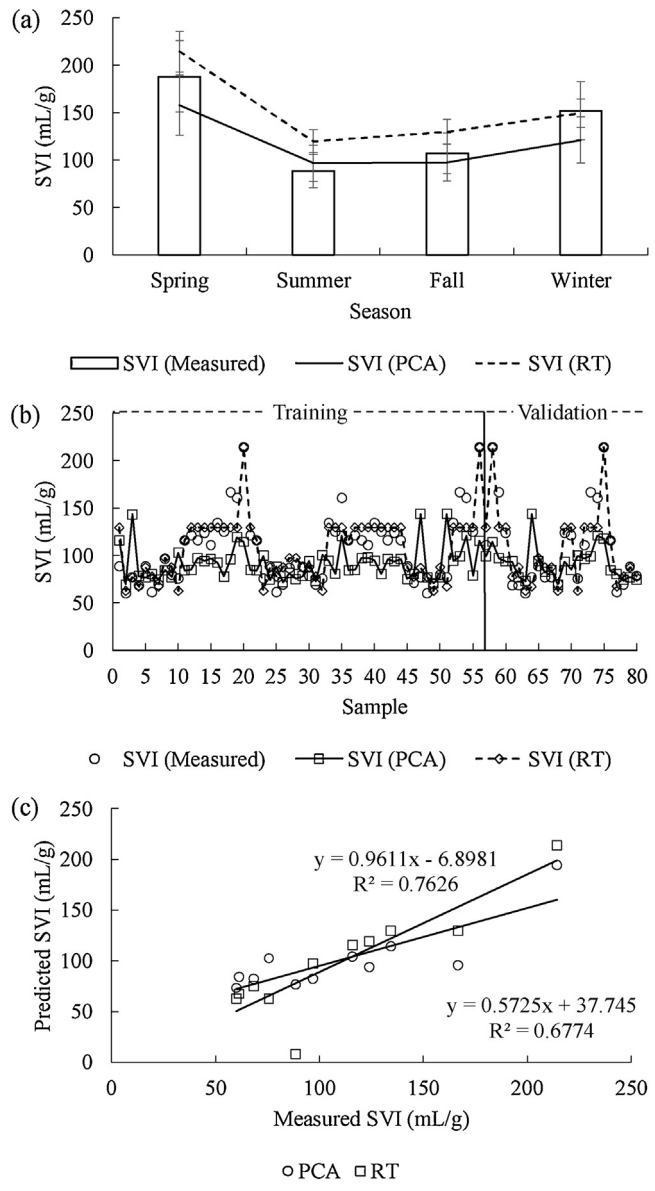


Fig. 2. (a) Seasonal variation of SVI fitted to PCA and regression tree (RT) models, (b) training and validation procures for the prediction of SVI via PCA (R^2 = 0.71 for training and 0.68 for validation) and RT (R^2 = 0.84 for training and 0.81 for validation), and (c) model generalization with R^2 = 0.68 and 0.76 for PCA and RT, respectively.

ciencies were 93.4–99.2% and 48.3–95.0%, respectively. The effluent NH_4^+-N was acceptable according to the allowable regulations of 3.0 and 2.0 mg/L as general and special standards, respectively (DWA, 1998). The biological units were also subjected to PO_4^{3-}-P and TP loads of 23 ± 11 and 55 ± 26 kg/d, respectively. The removal efficiencies were 75.9 ± 20.5% and 73.8 ± 18.8% for PO_4^{3-}-P and TP, respectively. The effluent limits of phosphorus that are applicable for discharge into a water resource in South Africa are 10 mg/L as general standards and 1 mg/L as special standards (DWA, 1998). The influent TSS concentration was 67.2 ± 26.7 mg/L, providing a TSS load of (1.9 ± 1.0) × 10^3 kg/d. The effluent TSS ranged between 2.8 and 163.6 mg/L, in which several observations exceeded the allowable limit of 10 mg-TSS/L (DWA, 1998). About 38% of the total number of SVI observations exceeded the threshold level for sludge bulking (SVI > 120 mL/g). Moreover, the seasonal variation of SVIs ($p < 0.05$) implied that the bulking episodes occurred during the low-temperature seasons (Fig. 2a). This finding might be due to the presence of filamentous bacteria such as *Microthrix parvicella*, and Eikelboom Type 0041 and 0092. In a similar study, Madoni et al. (2000) found that the main filamentous species involved in bulking problems were *Microthrix parvicella* (53.2%), Type 0041 (11.3%), and Type 021 N (9.7%).

The aforementioned results revealed efficient capabilities for the removal of organic and nutrient pollutants, in which the treated wastewater complied with the department of water and sanitation (DWS) regulations during the study period. However, the decrease in settling properties of solids could be associated with the sludge bulking events and the overgrowth of filamentous organisms (Seviour et al., 1994). Under this condition, the tCOD removal efficiency reduced to 63.4%, indicating the detrimental effect of bulking on the effluent quality. The growth of filamentous microorganisms could retard the biodegradation of organic matter due to slower kinetics involved (Ovez et al., 2006).

3.2. PCA application for prediction of SVI

PCA was applied to group the wastewater characteristics and operational conditions that influenced sludge bulking according to common features. Since the input attributes had different units, variances, and wide ranges of measurements, they were subjected to a standardization step. Fig. 1(a) shows the normalized variables displayed by Box-and-Whisker plot. In this figure, each variable extended from the 25th to 75th percentiles and had a mean of zero and unit standard deviation. This procedure prevents certain factors to dominate the analysis because of their large numerical values (Nasr and Zahran, 2016).

The first principal component (PC1) had an eigenvalue > 1.0 and captured the highest variance of the dataset with a value of 57.6%. The second and third principal components (PC2 and PC3) explained only 17.4% and 11.1% of the total variance and provided low loadings for organic and nutrient concentrations. Accordingly, PC1 was used in this study to describe the correlations among the investigated parameters and to identify the most important information in the dataset.

Fig. 1(b) shows the PCA results identified by PC1 (57.6%) and PC2 (17.4%). PC1 had high positive loadings for F/M (0.30), sCOD (0.33), tCOD (0.30), NH_4^+-N (0.35), TKN (0.32), PO_4^{3-}-P (0.33), TP (0.33), and TSS (0.31). This result indicated that PC1 was positively influenced by the studied wastewater characteristics. Moreover, both SVI and SRT had negative loadings on the PC1 coordinate, suggesting that an increase in SVI could be attributed to the increment in SRT. In a similar study, Comas et al. (2008) reported that the risk of foaming in activated sludge systems was obtained at high SRT. Flores-Alsina et al. (2009) also demonstrated that the high SRT values increased the occasion of sludge bulking. Moreover, SVI was adversely correlated with NH_4^+-N and sCOD, indicating that the risk of bulking could be due to the deficiency of organic components and nutrient species. Similarly, Comas et al. (2008) reported that filamentous bulking could occur due to substrate limiting conditions or nutrient deficiency. However, the optimum nutrient concentration required for bacterial growth varies with each activated sludge system (Wang et al., 2016).

The second principal component (PC2) depicted that SVI positioned on the opposite side of temperature, suggesting that SVI increased at low-temperature environments. This observation was also explored in the seasonal variation of SVI, as shown in Fig. 2(a). This result could be due to the decrease in lipid solubility at low temperatures, causing specific hydrophobic microorganisms to outcompete other lipid-utilizing microorganisms for oleic acid (Wang et al., 2016). Moreover, SVI was observed in the opposite direction of F/M, implying that the risk of filamentous bulking could be linked to low F/M ratios. Similarly, Liu et al. (2016) found that low F/M ratios (< 0.25 1/d) encouraged the occurrence of filamentous bulking.

Based on the PCA results, SVI could be predicted using Eq. 3 with R^2 values of 0.71 and 0.68 for the training (n = 56) and validation (n = 24) datasets, respectively (Fig. 2b). In this model, SRT, COD, and NH_4^+-N were selected as predictors to avoid overfitting. Due to the complexity of the full-scale activated sludge system, the prediction accuracy of PCA with R^2 over 0.7 is acceptable (Wang et al., 2017).

$$SVI = 91.05 + 2.56SRT\text{-}0.19sCOD\text{-}2.53NH_4^+\text{-}N \quad (3)$$

In a similar study, Lou and Zhao (2012) employed a PCA technique to predict SVI using several inputs including carbon and nutrient species. Their study (Lou and Zhao, 2012) depicted that SVI directly associated with COD, biological oxygen demand (BOD), TP, and TN, whereas it had negative correlations with temperature, and pH.

3.3. Modelling of SVI using regression tree

The decision tree model was developed using wastewater characteristics and operational conditions (as inputs) and SVI (as an output). It was found that the R^2 value of the training observations was 0.84 (n = 56; Fig. 2b), and hence, the regression tree fitted the training data well. In addition, the R^2 of the validation dataset was 0.81 (n = 24; Fig. 2b), suggesting that the developed model would be able to predict SVI values based on new records of the independent variables.

Fig. 2(c) shows the generalization procedure of the developed models using test data that were not employed during the training and validation steps. The R^2 values of PCA and regression tree models were 0.68 and 0.76, respectively. In addition, the average absolute error between the real records and the predicted outputs was 23.3 mL/g for PCA compared with 14.2 mL/g for the regression tree model. Hence, regression tree could verify a higher predictive ability of SVI than PCA across a wide range of inputs and applications.

The regression tree in Fig. 3 demonstrated that the decision node was NH_4^+-N, which provided the main impact on SVI. Moreover, the intermediate nodes were SRT and sCOD. This observation complied with the PCA results, indicating that NH_4^+-N, SRT, and sCOD were amongst the investigated factors that highly influenced SVI. The decrease in NH_4^+-N below 12.3 mg/L resulted in an increase in SVI to 130 mL/g (using the regression tree at a pruning level 3 of 4). This observation corresponded to a ratio of C (18)/N (<1), which was lower than the optimum value of C/N = 20/1 for a balanced bacterial growth system (Wukasch, 1993). Hence, insufficient nitrogen concentrations in the influent wastewater could favor the overgrowth of filamentous species and cause possible risks of bulking. Similar to the PCA results, an increase in SVI was noticed at either an increase in SRT (> 20.6 d) or a decrease in sCOD (< 99.0 mg/L); i.e., see Eq. 3. However, in comparison to the PCA model, the decision tree method provided higher performances for both training and validation procedures (Fig. 2b). In addition, the decision tree model adequately interpreted and displayed the nonlinear correlations between wastewater characteristic and SVI in a simple graphical interface. Bagheri et al. (2015) applied an artificial neural network (ANN) model with 76 data points to predict SVI (i.e., output) using several inputs such as DO, TSS, COD, and TN. Their study (Bagheri et al., 2015) also found that SVI was considerably influenced by the influent TN (as a single input attribute), and TN and COD (as a group of two inputs). The usefulness of the decision tree model derived in this application for the prediction of bulking events, with respect to other existing models, is listed in Table 4.

3.4. Effect of wastewater characteristics on filamentous bacteria using classification tree

In this investigation, six filamentous bacteria were identified and confirmed using FISH analysis. These microorganisms were *Microthrix parvicella*, *Thiothrix* I and II, and Types 0041, 0092, and 021 N. Table 5 lists the dominance of filamentous species throughout the study period. The decision tree method was employed to classify the proliferation of filamentous microorganisms based on the influent wastewater characteristics and operational conditions. The training dataset (n = 56) was used to build the classification tree model, whereas the validation dataset (n = 24) was employed to determine the suitable tree size required to attain the optimal classifications. About 86% of the training observations were correctly classified, and hence, the classification tree fitted the training data well. In addition, 83% of the validating set records were appropriately classified, and thus, the decisions were acceptable. The classification tree models were able to handle categorical attributes and to assign filamentous bacteria to specific classes, such as, “Few”, “Common”, and “Abundant”. The determination of filamentous classification can be illustrated as follows (Fig. 4).

3.4.1. Microthrix parvicella

Microthrix parvicella is a gram-positive bacterium (Williams and Unz, 1985), which has been detected in several WWTPs in South Africa (Blackbeard et al., 1988). It belongs to the class *Actinobacteria* and has a length of 200–400 μm with a diameter of 0.6 – 0.8 μm (Blackall et al., 1995). Domestic nutrient removal plants with low F/M (< 0.2 1/d) are favorable for the growth of *Microthrix parvicella* (Eikelboom, 2000). In this study, *Microthrix parvicella* appeared in the collected samples as “None” (11%), “Few” (67%), and “Some” (22%). The limited proliferation of *Microthrix parvicella* could be linked to the deficiency of long chain fatty acid and lipids content in wastewater influents (Ovez et al., 2006). In addition, the growth of *Microthrix parvicella* cannot occur in anoxic and anaerobic units (Rossetti et al., 2002). *Microthrix parvicella* causes serious bulking, and its hydrophobic cell surface can adsorb fats and fatty acids and stimulate the formation of stable scum (Hug et al., 2006). Madoni et al. (2000) reported that *Microthrix parvicella* was the most common filamentous microorganism involved in bulking and foaming events.

Fig. 4(a) shows the decision tree for the classification of *Microthrix parvicella*, in which the generated tree included three terminal nodes, i.e., “None”, “Few”, and “Some”. *Microthrix parvicella* was mainly affected by TSS (the root node), and its dominance became “None” at lower TSS values (TSS < 31.9 mg/L). Under this condition, the filament attained a minor influence on the settling velocity of the sludge. Foaming induced by *Microthrix parvicella* can be controlled by reducing the suspended solids concentration in the aeration tank (Madoni et al., 2000). Temperature, i.e., the decision node of sub-tree, had the second impact on the dominance of *Microthrix parvicella*. A decrease in temperature below 15.5 °C (corresponding to winter) observed the appearance of *Microthrix parvicella* as “Some”. Under this environment, the settling properties of the sludge were possibly deteriorated. However, the occurrence of *Microthrix parvicella* as “Few” was noticed in summer with temperature > 15.5 °C. Hence, the development of *Microthrix parvicella* increased at the low-temperature conditions (e.g., winter), equivalent to high SVI events during winter and spring (see Fig. 2a).

Similarly, Eikelboom (2000) revealed that *Microthrix parvicella* grew largely at the end of the winter season, resulting in the occurrence of serious operational problems. Fan et al. (2017) demonstrated that cold climate in winter and spring favored the dominant growth of *Microthrix parvicella*, suggesting that at temperatures over 20 °C, some enzymes are produced and destroy the

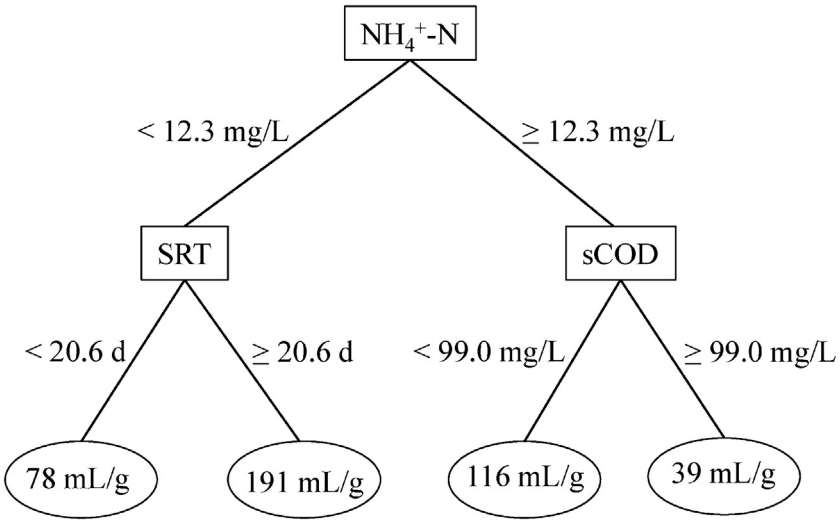


Fig. 3. Regression tree for prediction of SVI using wastewater characteristics and operational conditions.

Table 4
Models used for prediction of filamentous bulking, as reported in the literature.

Model type	Purpose	Benefit	Limitation	Reference
PCA	Predicts SVI using temperature, COD, NH_4^+-N, TN, MLSS, SS, BOD, pH, and TP.	Extracts useful information from large datasets.	Requires a data normalization step.	Lou and Zhao (2012)
Three-layered feedforward ANN	Predicts SVI using temperature, COD, NH_4^+-N, TN, MLSS, SS, BOD, pH, and TP.	Determines complicated and nonlinear correlations.	Requires long training times due to a large number of weights and neurons between layers.	Lou and Zhao (2012)
Cumulative logistic model	Determines the effects of pH, COD, temperature, F/M, NH_4-N, DO, and PO_4-P on the abundance of filamentous bacteria.	Efficient for predicting discrete functions. It does not require data tuning.	Important and independent variables should be used.	Deepnarain et al. (2015)
Genetic algorithm based ANN	Predicts SVI using pH, DO, MLVSS, TSS, COD, TN, and temperature.	Simulates complex functional relationships using an unrestricted number of inputs and outputs.	Requires adjustment of the network parameters and connecting weights.	Bagheri et al. (2015)
Sensitivity analysis based ANN	Predicts SVI using MLSS, COD, TN, pH, and DO.	Attains long-range predictions even in the occurrence of data noise.	Overfitting problem can occur if the error of testing is larger than that of training.	Han et al. (2016)
Decision tree	Predicts SVI using pH, temperature, DO, SRT, F/M ratio, sCOD, tCOD, NH_4^+-N, TKN, PO_4^{3-}-P, TP, and TSS.	Improves the prediction accuracy by considering the growing and pruning strategies.	Requires large amounts of data records.	This study

Table 5
Percentages of filamentous bacteria proliferating in all samples.

	M. parvicella	*Thiothrix* I	*Thiothrix* II	Type 0041	Type 0092	Type 021 N
None	11	0	0	0	0	0
Few	67	50	22	0	0	33
Some	22	28	44	66	17	45
Common	0	22	34	28	28	22
Very common	0	0	0	6	6	0
Abundant	0	0	0	0	33	0
Excessive	0	0	0	0	16	0

cell wall of *Microthrix parvicella*. In another study, Araújo dos Santos et al. (2015) found that *Microthrix parvicella* was directly correlated with pH in the aeration tank and it contributed to the biomass washout along with the decline in the treatment performance.

3.4.2. Thiothrix *I & II*

Thiothrix strains are gram-negative filamentous bacteria that form sulphur granules (Eikelboom, 2000). This type is characterized by filament length < 500 μm, round ended rod and cylindrical cell shape, and colony size of 1–2 mm (Williams and Unz, 1985). Henriet et al. (2017) reported that filamentous bulking caused by *Thiothrix* species in dairy WWTPs caused sludge washout and a decrease in the biological performance. Moreover, their study (Henriet et al., 2017) demonstrated that the growth of *Thiothrix* is mainly influenced by the presence of volatile fatty acids (VFA) and a relatively low DO concentrations (1.4–4.0 mg/L).

The classifications of *Thiothrix* I during the investigation were "Few" (50%), "Some" (28%), and "Common" (22%). As can be seen in Fig. 4(b), the root of the classification tree was F/M, whereas the internal node was TP. Moreover, the end nodes were "Few", "Some", and "Common", and the tree height was 2. It was demonstrated that *Thiothrix* I was "Few" at F/M ≥ 0.18 1/d. However, when the F/M ratio decreased below 0.18 1/d, the existence of *Thiothrix* I was either "Common" or "Some", based on the TP concentrations.

Thiothrix II appeared during the study period as "Few" (22%), "Some" (44%), and "Common" (34%). As can be inferred in Fig. 4(c),

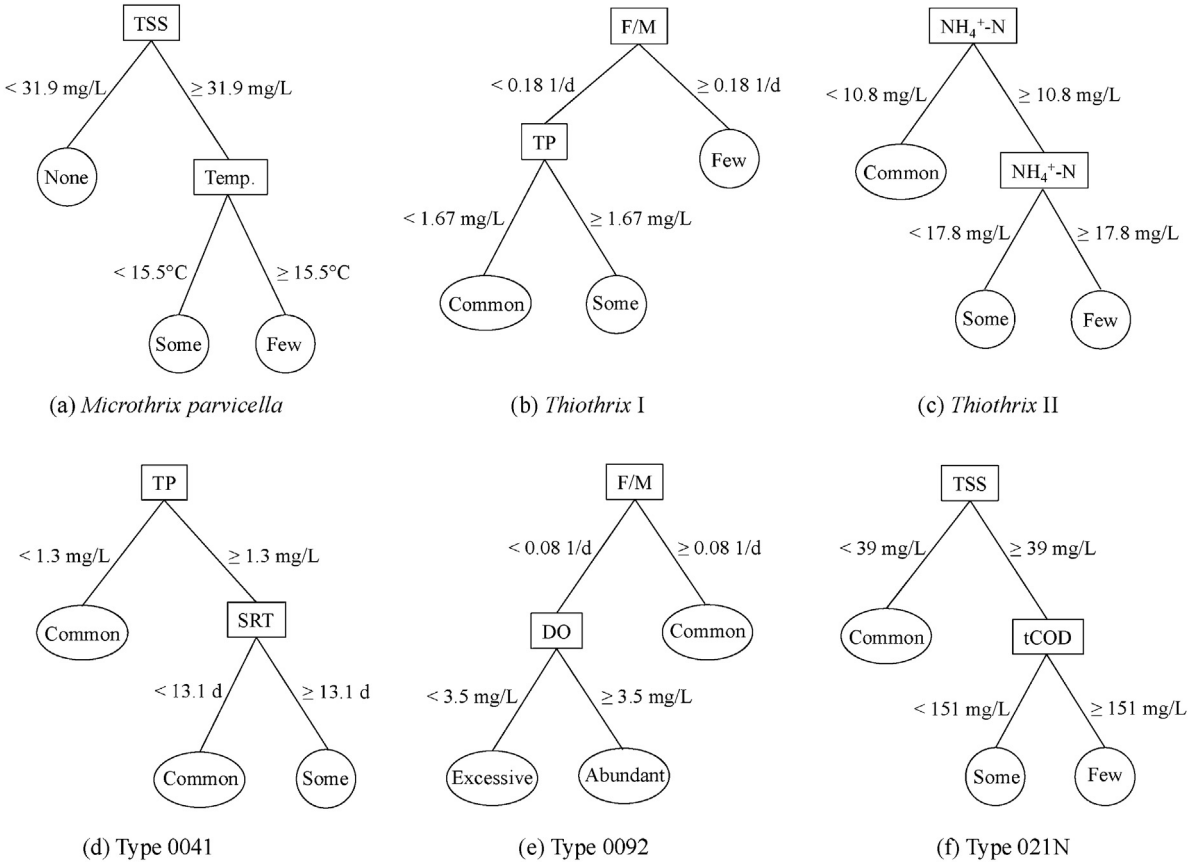


Fig. 4. Decision tree for classifications of (a) *Microthrix parvicella*, (b) *Thiothrix* I, (c) *Thiothrix* II, (d) Type 0041, (e) Type 0092, and (f) Type 021 N.

the influent NH_4^+-N (6.0–18.0 mg/L) was the tree root, and it also appeared as an internal node, indicating that the main factor affecting *Thiothrix* II was NH_4^+-N. The tree terminals were "Few", "Some", and "Common". It was demonstrated that *Thiothrix* II was "Common" at NH_4^+-N < 10.8 mg/L, equivalent to an insufficient nitrogen content of COD (20) to N (<1). The unbalanced ratio between readily biodegradable COD and nutrients favors the overgrowth of *Thiothrix* species (Henriet et al., 2017). Similarly, Eikelboom (2000) reported that the occurrence of *Thiothrix* in activated sludge could be due to the lack of either nitrogen or phosphorus species. Further, *Thiothrix* II became either "Some" or "Few" at NH_4^+-N of 10.8–17.8 mg/L or ≥ 17.8 mg/L, respectively.

Thiothrix spp. have been observed to cause high SVI > 200 mL/g in domestic activated sludge processes (Williams and Unz, 1985). Sludge bulking resulting from the excessive growth of *Thiothrix* spp. can be controlled by the implementation of aerobic selectors, small mixing units (anoxic or aerobic), or contact zone (without aeration) (Martins et al., 2004). They can also be limited by using anaerobic and anoxic stages, such as bio-P and denitrifying activated sludge systems (Wanner and Grau, 1988).

3.4.3. Eikelboom types 0041, 0092, and 021 N

The bacterium Type 0041 is a gram-variable microorganism having a filament length of 200–500 μm and cylindrical, cuboidal, or oval cell shape (Williams and Unz, 1985). Type 0041 is a filament morphotype responsible for the bulking and foaming episodes in biological nutrient removal activated sludge systems, viz., anaerobic–anoxic–aerobic processes (Seviour et al., 1994). The occurrence of bacteria Type 0041 in this study was "Some" (66%), "Common" (28%), and "Very common" (6%). The dominance of Type 0041 could be due to its ability to adapt and thrive in a wide range of environmental conditions in activated sludge systems (Lacko et al., 1999). Similarly, Blackbeard et al. (1988) reported that Type 0041 was the main filamentous microorganisms found in South Africa. TP that varied from 0.8 to 3.2 mg/L was the root node of the tree; viz., the most explanatory variable (Fig. 4d). The leaves of the classification tree were "Some", and "Common". Type 0041 became "Common" with a decrease in TP below 1.3 mg/L. This result corresponded to an unbalanced nutrient and/or substrate-limiting conditions with COD/N/P of 115/9/1. The nutrient requirements for balanced growth of microorganisms and deactivation of filamentous growth during aerobic treatment is C/N/P ratio of 100/5/1 (Wukasch, 1993). Furthermore, an increase in TP over 1.3 mg/L provided a suitable condition for either "Some" at SRT ≥ 13.1 d or "Common" at SRT < 13.1 d.

Type 0092 (Gram-negative) is a strict aerobic bacterium (Horan et al., 1988; Speirs et al., 2009), which has been detected in several nutrient removal activated sludge systems (Madoni et al., 2000). In this study, Type 0092 was the prevalent filamentous bacterium observed during high SVI conditions (SVI > 120 mL/g). Araújo Dos Santos et al. (2015) confirmed the strong relationship between Type 0092 and SVI based on the correlation between filamentous bacteria populations and environmental conditions. Type 0092 was detected throughout the investigated period as "Some" (17%), "Common" (28%), "Very common" (6%), "Abundant" (33%), and "Excessive" (16%). The development of Type 0092 could be because this filament is able to grow under various aerobic, anoxic, and anaerobic conditions; i.e., recognized as "All zone grower" (Araújo dos Santos et al., 2015). The F/M ratio (0.05 – 0.45 1/d) was the most informative input variable, in which Type 0092 became "Common" at F/M ≥ 0.08 1/d (Fig. 4e). With a decrease in F/M below 0.08 1/d, Type 0092 became either "Excessive" at DO < 3.5 mg/L or "Abundant" at DO ≥ 3.5 mg/L, suggesting a poor sludge quality. Madoni et al. (2000) also reported that Type 0092 was dominant in plants

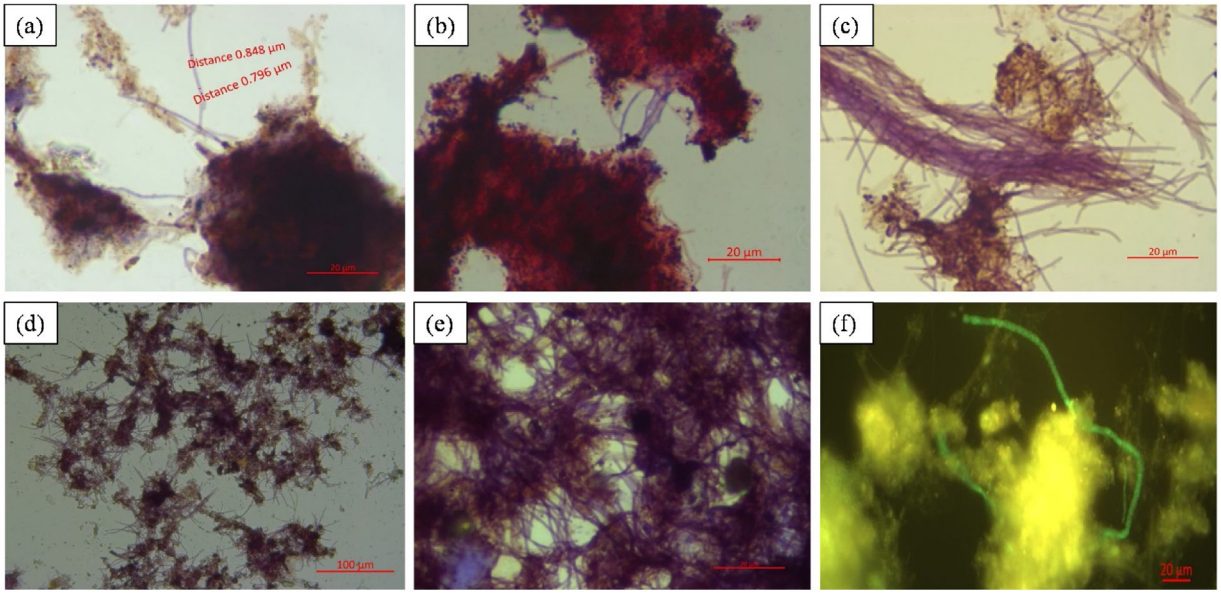


Fig. 5. Microscopic images of filamentous bacterium Type 0092: (a) "Some", (b) "Common", (c) "Very common", (d) "Abundant", (e) "Excessive", and (f) FISH image.

with a low F/M ratio (< 0.1 1/d). These results revealed the major impacts of F/M and DO on the occurrence of Type 0092. Hence, F/M and DO should be optimized to limit the growth of the filament Type 0092.

Type 021 N (Gram-negative) has been recognized as multicellular, sheathless, and curled bacterium that simulates blue-green algae (Aruga et al., 2002). Type 021 N is a filamentous bacterium able to utilize nitrate as an electron acceptor (Martins et al., 2004). The proliferation of Type 021 N was classified as "Few" (33%), "Some" (45%), and "Common" (22%). The influent TSS was the root node, whereas tCOD was the internal node of the classification tree. The terminal nodes of the classification tree were "Few", "Some", and "Common", and the tree height was 2. Type 021 N was "Common" with the decrease in TSS below 39.0 mg/L; however, it recorded either "Few" or "Some" at TSS ≥ 39.0 mg/L (Fig. 4f). The existence of Type 021 N as "Some" at COD < 151 mg/L could be related to the favorability of this filaments to dominate at low F/M activated sludge systems, as described by Plantb et al. (1990). Aerobic selectors and nutrient addition have been found to limit sludge bulking resulted from the growth of Type 021 N (Martins et al., 2004). Similar to *Thiothrix* spp, the dominance of Type 021 N is positively correlated with the soluble readily biodegradable substrates (Jenkins, 1992).

3.5. Filamentous images

Fig. 5 shows the microscopic observations of filamentous bacterium Type 0092. This bacterium was selected because it covered most of the classification categories. During non-bulking conditions (SVI < 120 mL/g), filamentous microorganisms were less common, and the sludge granules were dense and compact. This observation was noticed in Fig. 5(a), which showed medium-sized flocs with "Some" filaments, as identified by the Gram staining method. Moreover, Neisser-stained biomass displaying Eikelboom Type 0092 (violet filaments) indicated "Common" (Fig. 5b) and "Very common" (Fig. 5c) classes. In addition, Fig. 5(d) shows "Abundant" Type 0092 captured during sludge bulking events. Under this condition, filamentous microorganisms reside inside the microbial flocs forming a sponge-like structure that can alter the hindered settling velocity (Wágner et al., 2015). These classifications are also described in Table 1. Neisser staining that indicated "Excessive" Type 0092, obtained at SVI > 120 mL/g, is shown in Fig. 5(e). This bacterium forms bridges among the floc structures and prevents them from aggregation and compaction (Eikelboom, 2000); hence, reduces solid-liquid separation achievements. In addition, the extended filamentous network can capture gas bubbles, which carry the organisms to the surface and stabilize a foam layer. Fig. 5(f) demonstrates a FISH image of Type 0092. This microorganism transferred from the sludge flocs into the bulk solution and reduced the removal efficiency of organics and solids.

4. Conclusions

This study aimed at modelling the sludge bulking events in a full-scale activated sludge system using PCA and decision tree methods. The relationships between wastewater characteristics, operational conditions, and SVI were interpreted. The environmental factors that impacted the proliferation of filamentous microorganisms were classified. The investigated WWTP was capable of producing treated wastewater that complied with the National Water Act (No 36 of 1998) of South Africa; however, some observations of TSS exceeded the permissible limit of 10 mg/L. About 38% of the total number of SVI records were higher than the threshold value of sludge bulking. The results obtained from PCA and regression tree models indicated that the increase in SVI was attributed to the increment in SRT, deficiency of organic components and nutrient species, or a decrease in temperature. The detected filamentous bacteria were *Microthrix parvicella*, *Thiothrix* I & II, and Eikelboom Types 0041, 0092, and 021 N. The classification tree results revealed that *Microthrix parvicella* could dominate with a decrease in temperature below 15.5 °C, implying the bulking occurrence during the winter and spring seasons. Eikelboom Type 0092 was the dominant bacterium responsible for the high SVI events, and it was classified as "Some" (17%), "Common" (28%), "Very common" (6%), "Abundant" (33%), and "Excessive" (16%). Type 021 N was mainly influenced by TSS, whereas the F/M ratio provided the major impact on both *Thiothrix* I and Type 0092. *Thiothrix* II and Type 0041 were primarily impacted by the unbalanced nutrient condition in the feed wastewater. The developed models attained acceptable training and validating steps, and thus, they can be used to predict the sludge bulking episodes. However, future studies will be conducted to (a) investigate the applicability of the developed models for the

prediction of sludge bulking in other WWTPs operated at different environmental conditions, and (b) develop mathematical models that can predict bulking events based on daily measurements.

Acknowledgments

The support of the Water Research Commission (WRC), South Africa, for the successful completion of this project is gratefully acknowledged.

References

Adamse, A., 1970. Some characteristics of arthrobacters from a dairy waste activated sludge. Water Res. 4 (12), 797–803.
Amann, R., et al., 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl. Environ. Microbiol. 56, 1919–1925.
APHA, 1998. Standard Methods for the Examination of Water and Wastewater. American Public Health Association (APHA)-American Water Works Association (AWWA)-Water Environment Federation (WEF), Washington, D.C.
Araújo dos Santos, L., et al., 2015. Study of 16 Portuguese activated sludge systems based on filamentous bacteria populations and their relationships with environmental parameters. Appl. Microbiol. Biotechnol. 99 (12), 5307–5316.
Aruga, S., et al., 2002. Characterization of filamentous Eikelboom type 021N bacteria and description of *Thiothrix disciformis*sp. nov. and Thiothrix flexilis sp. nov. Int. J. Syst. Evol. Microbiol. 52, 1309–1316.
Bagheri, M., Mirbagheri, S., Bagheri, Z., Kamarkhani, A., 2015. Modeling and optimization of activated sludge bulking for a real wastewater treatment plant using hybrid artificial neural networks-genetic algorithm approach. Process Saf. Environ. 95, 12–25.
Blackall, L., et al., 1995. "*Microthrix parvicella*" is a novel, deep branching member of the actinomycetes subphylum. Syst. Appl. Microbiol. 17, 513–518.
Blackbeard, J., Gabb, D., Ekama, G., Marais, G., 1988. Identification of filamentous organisms in nutrient removal activated sludge plants in South Africa. Water SA 14 (1), 1–18.
Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and Regression Trees. CRC Press, California, USA: Boca Raton, FL.
Comas, J., et al., 2008. Risk assessment modelling of microbiology-related solids separation problems in activated sludge systems. Environ. Model. Softw. 23 (10-11), 1250–1261.
Daims, H., et al., 1999. The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. Syst. Appl. Microbiol. 22, 434–444.
Deepnarain, N., et al., 2015. A logistic model for the remediation of filamentous bulking in a biological nutrient removal wastewater treatment plant. Water Sci. Technol. 72 (3), 391–405.
Dunkel, T., et al., 2016. Evaluating the influence of wastewater composition on the growth of Microthrix parvicella by GCxGC/qMS and real-time PCR. Water Res. 88, 510–523.
DWA, 1998. National Water Act No. 36, 1998. Republic of South Africa: DWA (Department of Water Affairs) Government Gazette Staatskoerant.
Eikelboom, D., 2000. Process Control of Activated Sludge Plants by Microscopic Investigation. IWA publishing.
Fan, N., et al., 2017. Factors affecting the growth of Microthrix parvicella: batch tests using bulking sludge as seed sludge. Sci. Total Environ. 609, 1192–1199.
Flores-Alsina, X., et al., 2009. Including the effects of filamentous bulking sludge during the simulation of wastewater treatment plants using a risk assessment model. Water Res. 43 (18), 4527–4538.
Ganidi, N., Tyrrel, S., Cartmell, E., 2009. Anaerobic digestion foaming causes - a review. Bioresour. Technol. 23, 5546–5554.
Gich, F., Garcia-Gil, J., Overmann, J., 2001. Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. Arch. Microbiol. 177 (1), 1–10.
Han, H.-G., Li, Y., Guo, Y.-N., Qiao, J., 2016. A soft computing method to predict sludge volume index based on a recurrent self-organizing neural network. Appl. Soft Comput. 38, 477–486.
Henriet, O., Meunier, C., Henry, P., Mahillon, J., 2017. Filamentous bulking caused by Thiothrix species is efficiently controlled in full-scale wastewater treatment plants by implementing a sludge densification strategy. Sci. Rep. 7, 1430.
Horan, N., Bu'Ali, A., Eccles, C., 1988. Isolation, identification and characterisation of filamentous and floc-forming bacteria from activated sludge flocs. Environ. Technol. Lett. 9, 449–457.
Hug, T., Gujer, W., Siegrist, H., 2006. Modelling seasonal dynamics of M̈icrothrix parvicella. Water Sci. Technol. 54 (1), 189–198.
Jenkins, D., 1992. Towards a comprehensive model of activated sludge bulking and foaming. Water Sci. Technol. 25 (6), 215–230.
Jenkins, D., Richard, M., Daigger, G., 2003. Manual on the Causes and Control of Activated Sludge Bulking, Foaming, and Other Solids Separation Problems, 3rd edition. CRC Press (Lewis Publishers), United States.
Jolliffe, I., 2002. Principal Component Analysis, 2nd ed. Springer, Verlag New York.
Kanagawa, T., et al., 2000. Phylogenetic analysis of and oligonucleotide probe development for eikelboom type 021N filamentous bacteria isolated from bulking activated sludge. Appl. Environ. Microbiol. 66 (11), 5043–5052.
Kim, J.W., et al., 2001. Application of decision-tree induction techniques to personalized advertisements on internet storefronts. Int. J. Electron. Commun. 5 (3), 45–62.
Lacko, N., Bux, F., Kasan, H., 1999. Survey of filamentous bacteria in activated sludge plants in KwaZulu-Natal. Water SA 25, 63–68.
Liu, Y., Guo, J., Wang, Q., Huang, D., 2016. Prediction of filamentous sludge bulking using a state-based gaussian processes regression model. Sci. Rep. 6, 31303.
Loh, W., Shih, Y., 1997. Split selection methods for classification trees. Stat. Sin. 7, 815–840.
Lou, I., Zhao, Y., 2012. Sludge bulking prediction using principle component regression and artificial neural network. Math. Probl. Eng. 2012, 237693, http://dx.doi.org/10.1155/2012/237693, 17 pages.
Madoni, P., Davoli, D., Gibin, G., 2000. Survey of filamentous microorganisms from bulking and foaming activated-sludge plants in Italy. Water Res. 34, 1767–1772.
Martins, A., Pagilla, K., Heijnen, J., van Loosdrecht, M., 2004. Filamentous bulking sludge—a critical review. Water Res. 38 (4), 793–817.
Mielczarek, A., Kragelund, C., Eriksen, P., Nielsen, P., 2012. Population dynamics of filamentous bacteria in Danish wastewater treatment plants with nutrient removal. Water Res. 46 (12), 3781–3795.
Nasr, M., Zahran, H., 2016. Performance evaluation of agricultural drainage water using modeling and statistical approaches. Egypt. J. Aquat. Res. 42 (2), 141–148.
Nielsen, P., Daims, H., Lemmer, H., 2009. FISH Handbook for Biological Wastewater Treatment. UK: IWA Publishing, London.
Ovez, S., Ors, C., Murat, S., Orhon, D., 2006. Effect of hypochloride on microbial ecology of bulking and foaming activated sludge treatment for tannery wastewater. J. Environ. Sci. Health A Tox. Subst. Environ. Eng. 41 (10), 2163–2174.
Plantb, T., PhD Thesis 1990. Solving Sludge Bulking Problems through Filamentous Organism Identificationi Case Studies In Massachusetts. University of Massachusetts.
Rossetti, S., et al., 2002. "*Microthrix parvicella*": a new approach for kinetic and physiological characterization. Water Sci. Technol. 46 (1-2), 65–72.
Seviour, E., et al., 1994. Studies on filamentous bacteria from australian activated sludge plants. Water Res. 28 (11), 2335–2342.
Speirs, L., et al., 2009. Filamentous bacterium eikelboom type 0092 in activated sludge plants in Australia is a member of the phylum Chloroflexi. Appl. Environ. Microbiol. 75 (8), 2446–2452.
USEPA, 1983. Methods for Chemical Analysis of Water and Wastes. United States Environmental Protection Agency, Washmgton, DC.
Wágner, D., et al., 2015. Microthrix parvicella abundance associates with activated sludge settling velocity and rheology - quantifying and modelling filamentous bulking. Water Res. 78, 121–132.
Wang, P., Yu, Z., Qi, R., Zhang, H., 2016. Detailed comparison of bacterial communities during seasonal sludge bulking in a municipal wastewater treatment plant. Water Res. 105, 157–166.
Wang, X., Ratnaweera, H., Holm, J., Olsbu, V., 2017. Statistical monitoring and dynamic simulation of a wastewater treatment plant: a combined approach to achieve model predictive control. J. Environ. Manage. 193, 1–7.
Wanner, J., Grau, P., 1988. Filamentous bulking in nutrient removal activated sludge systems. Water Sci. Technol. 20 (4-5), 1–8.
Williams, T., Unz, R., 1985. Isolation and characterization of filamentous bacteria present in bulking activated sludge. Appl. Microbiol. Biotechnol. 22 (4), 273–282.
Wukasch, R., 1993. Proceedings of the 48th Industrial Waste Conference Purdue University, May 1993. U.S. CRC Press - Taylor & Francis.
Yeo, B., Grant, D., 2018. Predicting service industry performance using decision tree analysis. Int. J. Inf. Manage. 38 (1), 288–300.
Yu, Z., Haghighat, F., Fung, B., Yoshino, H., 2010. A decision tree method for building energy demand modeling. Energy Build. 42, 1637–1646.

APPENDIX FIVE: IMPACT OF SLUDGE BULKING ON RECEIVING ENVIRONMENT USING QUANTITATIVE MICROBIAL RISK ASSESSMENT (QMRA)-BASED MANAGEMENT FOR FULL-SCALE WASTEWATER TREATMENT PLANTS

APPENDIX FIVE

Journal of Environmental Management 267 (2020) 110660

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: http://www.elsevier.com/locate/jenvman

Research article

Impact of sludge bulking on receiving environment using quantitative microbial risk assessment (QMRA)-based management for full-scale wastewater treatment plants

Nashia Deepnarain [a], Mahmoud Nasr [b], Isaac Dennis Amoah [a], Abimbola Motunrayo Enitan-Folami [a], Poovendhree Reddy [c], Thor Axel Stenström [a], Sheena Kumari [a], Faizal Bux [a,*]

[a] Institute for Water and Wastewater Technology, Durban University of Technology, Durban, 4000, South Africa
[b] Sanitary Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, 21544, Egypt
[c] Department of Community Health Studies, Faculty of Health Sciences, Durban University of Technology, South Africa

ARTICLE INFO

Keywords:
Disinfection
Human exposure
Pathogenic bacteria
Quantitative microbial risk assessment (QMRA)
Sludge bulking

ABSTRACT

During sludge bulking in wastewater treatment plants (WWTPs), high amounts of potentially pathogenic bacteria would release into the environment, causing various human-health risks. This is the first study attempting to assess the microbial infections associated with the reuse of WWTP effluents under various bulking conditions. Three common waterborne pathogens, *viz.*, *E. coli* O157:H7, *Salmonella*, and *Mycobacterium*, were quantified from full-scale WWTPs using DNA extraction and qPCR at different sludge volume indices (SVIs). The detected pathogens were incorporated into a quantitative microbial risk assessment (QMRA) model to determine the applicability of WWTP discharge for recreational (bathing) activities and agricultural practices. The QMRA exposures were children, women, and men during swimming, and farmers and vegetable consumers during irrigation. Bacterial abundance in the treated wastewater increased in response to SVIs, and the QMRA values at all bulking events exceeded the tolerable risk of one case of infection per 10,000 people per year. Hence, various disinfection scenarios (chlorination, ultraviolet, and ozonation) were hypothetically tested to control the risks associated with pathogenic bacteria, allowing for safe disposal and reuse of the treated effluent. The ultraviolet application provided the highest ability to inactivate the pathogenic bacteria, except for the case of children exposed to *Salmonella* infection during swimming. The reduction of *Mycobacterium* infection risks with either chlorination or ozonation showed inefficient results. This study would be helpful for the management of human health risks associated with effluent wastewater containing pathogens, i.e., particularly concerning the case of sludge bulking.

1. Introduction

An essential function of wastewater treatment plants (WWTPs) is to avoid environmental deterioration through the reduction of microbiological pollutants and nutrients (Cai and Zhang, 2013). However, the effluents released from WWTPs may contain a wide range of waterborne pathogens, including bacteria, viruses, and parasites (Dickin et al., 2016). Especially during sludge bulking conditions, high amounts of pathogens are expected to escape with the WWTP effluent as a result of inadequate settling (Martins et al., 2004). This trend would cause detrimental health risks (e.g., cholera, amoebiasis, diarrhea, typhoid fever, and gastroenteritis) for the reuse application of treated wastewater (Amha et al., 2017). According to the World Health Organization (WHO) guidelines, the treated wastewater can be provided to consumers concerning the risk level of $\leq 10^{-4}$ infections per year, *viz.*, one case or less per 10,000 people per year (WHO, 2006, 2011; 2016). Hence, more efforts are required to assess the microbial contamination associated with the application of WWTPs effluents for irrigation and other practices, concerning human health risk.

Wastewater treatment facilities contain a diversity of microbial pathogens such as *Campylobacter jejuni*, *Cryptosporidium* spp., *Escherichia coli* O157:H7, Giardia, *Salmonella typhimurium*, *Shigella flexneri*, *Vibrio*

* Corresponding author. at Institute for Water and Wastewater Technology, Durban University of Technology, P.O. Box 1334, Durban, 4000, South Africa.
E-mail address: faizalb@dut.ac.za (F. Bux).

https://doi.org/10.1016/j.jenvman.2020.110660
Received 9 December 2019; Received in revised form 16 April 2020; Accepted 25 April 2020
Available online 3 May 2020
0301-4797/

cholerae, and *Mycobacterium* (Haas et al., 2014; Ekwanzala et al., 2018). Among these contaminants, *E. coli* (especially *E. coli* O157:H7) has emerged as a pathogen of a substantial public health concern (Barak et al., 2005). This microbe can exist in the effluents of WWTPs and resist the treatment and disinfection processes (Ayaz et al., 2014). Moreover, the risk of *Salmonella* infections can be used as a guide of hazard in edible crops irrigated with reclaimed wastewater (Amha et al., 2015). Both *E. coli* O157:H7 and *Salmonella* are considered the most severe foodborne pathogens found, even at a low infective dose, in raw vegetables and fruits irrigated with contaminated water (Ayaz et al., 2014). *Mycobacterium* has been identified as one of the opportunistic environmental pathogens of humans, tending to enrich in foaming bacteria in WWTPs (Amha et al., 2017). Many species of *Mycobacterium* can result in morbidity and mortality, and they exhibit an elevated resistance against disinfectants and antibiotics (Falkinham, 2009). Accordingly, the elimination of these microbial contaminants from WWTPs effluents is a crucial objective to ensure that the treated water is safe for reuse applications.

Quantitative microbial risk assessment (QMRA) is a probabilistic model that integrates data on pathogen abundance, human exposure, and infection to evaluate the potential health impacts related to a polluted environment (Sampson et al., 2017). The components of QMRA include hazard description, exposure estimation, dose-response valuation, and risk classification (Schijven et al., 2019). The QMRA model can provide a clear comparison and assessment of various systems regarding health-based measures (Zhou et al., 2018). This framework can also be used to identify and examine the accessibility of WWTPs effluent for various reuse scenarios (Amha et al., 2015). For instance, farmers in some developing countries having inaccessibility to freshwater resources may use treated wastewater for crop irrigation and cultivation (Van Vu et al., 2018). As a result, these irrigated fruits and vegetables may contain unacceptable pathogenic levels (Dong et al., 2017). Furthermore, communities and households living close to WWTPs have to protect themselves against a variety of waterborne pathogens and related diseases (Sunger et al., 2018).

In this study, the potential health risks associated with the discharge or reuse of effluents from WWTPs experiencing sludge bulking were evaluated. For this objective, a QMRA model was applied using the relative abundance of three common waterborne pathogens, i.e., *E. coli* 0157:H7, *Salmonella*, and *Mycobacterium*. The potential health risks associated with low, moderate, and high sludge bulking conditions in WWTPs were incorporated into the model. The QMRA target was to investigate the safety of treated effluent for (a) children, women, and men during recreational activities, (b) farmers during irrigation practices, and (c) consumers of edible plants (vegetables). The QMRA results were compared to the tolerable risk level of 10^{-4}, and various disinfection treatment scenarios via chlorination, ultraviolet (UV), and ozonation were included in the calculations to control the excess risk. To the best of our knowledge, this is the first study on applying the QMRA approach for the management of risks associated with pathogens in wastewater effluents in bulking WWTPs, i.e., particularly for farming activities and hygiene improvements.

2. Materials and methods

2.1. Sampling

Four full-scale biological WWTPs, situated in Durban, KwaZulu-Natal Province of South Africa, were selected for the sampling study. The treatment facilities were composed of physical unit operations, followed by biological reactors (aerators and secondary settling tanks) and chlorination. Wastewater samples (2 L) were collected over five consecutive months (April–August 2019) from the aeration tank and pre-chlorinated effluent. Composite samples of the mixed liquor were harvested, fixed, and transported to the laboratory within 2 h, following our previous study (Deepnarain et al., 2019). The pre-chlorinated samples were filtered and stored in phosphate-buffered saline (PBS) for DNA extractions. The received samples were maintained in the freezer (−20 °C) for further analyses. The sludge volume index (SVI) data were estimated regarding the wastewater volume (mL) occupied by 1 g of activated sludge after the settlement of mixed liquor for 30 min (Martins et al., 2004).

2.2. Microbial analysis

2.2.1. DNA extraction

The total genomic DNA was isolated and extracted from sludge samples using the phenol extraction method (Sekiguchi et al., 1999; Enitan et al., 2014). The WWTP effluent containing biomass (2 L sample) was centrifuged at 2600×*g* and 4 °C for 20 min, and the supernatant was discarded and pelleted. The concentrated wet biomass (2 mL) were washed with 1X PBS buffer followed by centrifugation. Genomic DNA was recovered by lysing the cells with 700 μL of lysis buffer (50 mM Tris, 5 mM EDTA, 150 mM NaCl, 1% Nonidet P-40, and 1 mM PMSF at pH 8.0). The mixture was then frozen-thawed for 5 min in ice-ethanol, and subsequently for 3 min in water-bath at 65 °C, which was repeated 5 times. Further, samples were pre-treated with 20 μL of Proteinase K (10 mg mL^{-1}) and incubated in a water bath at 37 °C for 30 min. The RNA and proteins were separated from the aqueous solution containing DNA with an equal volume of phenol, chloroform, and isoamyl alcohol (25:24:1) followed by chloroform–isoamyl alcohol (24:1). Further, 1 × volume of isopropanol was added to the genomic DNA precipitate and stored at −20 °C overnight. The DNA precipitate was pelletized at 12000 rpm and 4 °C for 20 min. The DNA was then washed with 70% ethanol and air-dried at room temperature for 20 min. Subsequently, the pellet was dissolved in 100 μL of TE buffer, and the purified DNA was stored at −20 °C for further use. The quantity and purity of the genomic DNA were assessed via Spectrophotometer (NanoDrop Technologies, ND-1000; USA) at an absorbance value of 260 nm, and the A_{260}/A_{280} ratio of about 1.8.

2.2.2. Quantitative polymerase chain reaction (qPCR)

The primer sets developed for the quantitative polymerase chain reaction (qPCR) analysis were selected based on their efficiency and specificity in amplifying *invA*, *eae*, and *IS6110* genes (Supplementary Table S1). For pathogenic *E. coli,* the 0157:H7 *eae* gene was targeted by forward primer 5′-GTAAGTTACACTATAAAAGCACCGTCG-3′ and reverse primer 5′-TCTGTGTGGATGGTAATAAATTTTTG-3' (Barak et al., 2005). For *Salmonella*, the *invA* gene was targeted by forward primer 5′-GTGAAATTATCGCCACGTTCGGGCAA-3′ and reverse primer 5′-TCATCGCACCGTCAAAGGAACC-3' (Jyoti et al., 2010). For *Mycobacterium* detection, the *IS6110* gene was targeted by forward primer 5′-CCTGCGAGCGTAGGCGTCGG-3′ and reverse primer 5′-CTCGTCCAGCGCCGCTTCGG-3' (Brosch et al., 2002). The total numbers of the selected gene copies in the samples were determined based on a real-time PCR analysis (C-1000 Touch, CFX 96, Bio-Rad Laboratories Pty., Ltd.; USA). The qPCR amplification and the thermal cycling parameters were identified, following Brosch et al. (2002), Barak et al. (2005), and Jyoti et al. (2010) (Supplementary Table S2). For the different pathogens, individual standard curves were prepared using the targeted DNA.

To create the copy number standard for the respective target genes, 2 μL of template genomic DNA was amplified using Sso fast green Master Mix (SsoFast™ EvaGreen® Supermix, Bio-Rad Laboratories Pty., Ltd.; USA). For each reaction, 0.5 μL of dNTPs (10 mM), 2 μL of template genomic DNA (10 ng μL^{-1}), 1 μL of each primer (5 μM), and 4 μL of Sso fast green Master Mix were prepared, and then nuclease-free water was added to a final volume of 10 μL. For each primer set tested against wastewater genomic DNA, a negative control was examined, including the selected primers, PCR Master Mix, and sterile water in place of the DNA template. The negative controls were included in the aforementioned procedures to ensure that the recorded fluorescence signals were

indicative of PCR amplification of the DNA template. The copy number of the gene was estimated using Eq. (1) (http://cels.uri.edu/gsc/cndna.html):

$$\text{Number of copies} = \frac{\text{Amount of DNA in nano} - \text{grams (ng)} \times \text{Avogadro's number}}{\text{Length in base pairs (bp)} \times 1 \times 10^9 \times 650} \quad (1)$$

where, Avogadro's number = 6.022×10^{23}, and the average weight of a base pair (bp) is 650 Da. The DNAs of the reference strains were serially diluted to the final concentrations, ranging from 10^1 to 10^8 genomic DNA copies per PCR reactions.

2.3. Quantitative microbial risk assessment (QMRA) framework

The QMRA framework used in this study is shown in Fig. 1, in which the model handled information on sludge bulking and the pathogenic microorganisms evaluated in the selected WWTPs. Data of SVIs at full-scale WWTPs were collected, as reported in our previous study (Deepnarain et al., 2019). Sludge bulking conditions were classified into low (SVI < 100 mL g^{-1}), medium (SVI = 100–200 mL g^{-1}), and high (SVI > 200 mL g^{-1}). The relative abundances of the waterborne microorganisms (*E. coli* O157:H7, *Salmonella*, and *Mycobacterium*) quantified using the qPCR assay (see section 2.2. Microbial analysis) were incorporated into the QMRA model. The risks of infection of these pathogens and their impacts on population and downstream watercourse were modelled using the QMRA method. The model was simulated for 10,000 iterations using Monte Carlo techniques for the probability of infections. The @Risk 7.5 software (Palisade Corporation; USA) add-on to Excel was used for the simulation step (Amoah et al., 2018b).

The estimated risks were compared to the tolerable risk thresholds recommended by (a) WHO for infection risk associated with drinking water, 10^{-4} case per person per year (WHO, 2001, 2006), and (b) Recreational Water Quality Criteria (RWQC) in U.S. Environmental Protection Agency (EPA) for infection risks related to recreational contacts, 32 cases per 1000 events (USEPA, 2012). The estimated risks were further categorized into low, moderate, and high risks using probable risk ranges of 0.05–0.10, 0.10–0.30, and 0.30–0.50, respectively. These categories were proposed based on a qualitative assessment of variation in the human health risk associated with the pathogenicity of *E. coli*, *Salmonella*, and *Mycobacterium*, as previously reported by Soller et al. (2010). The QMRA model was also used to define the treatment targets in relation to the required level of safety and health-based objectives for the local context. This step was done by considering three disinfection scenarios and estimating the corresponding reduction of risks. Similar procedures have been reported by Petterson and Stenström (2015), applying a QMRA model to investigate the inactivation efficiency of several pathogens (e.g., *E. coli* O157 and *Campylobacter*) in drinking water using free chlorine disinfection.

The QMRA framework involved four steps using experimental, practical, and theoretical assessments, which can be explicitly described as follows (Haas et al., 2014; Amoah et al., 2018a and b):

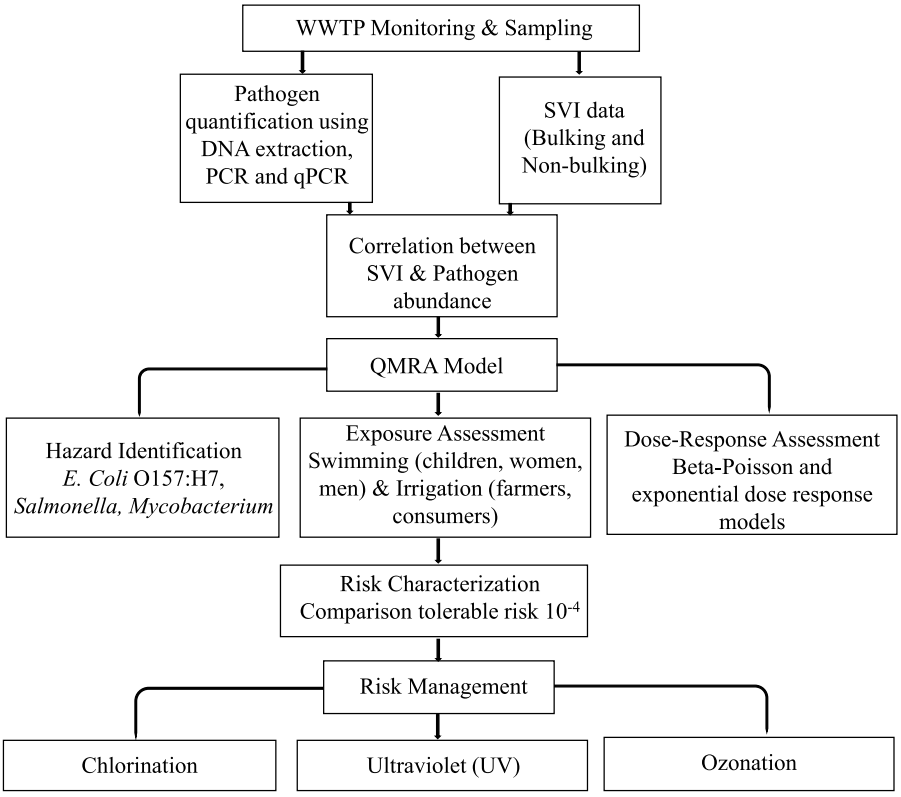


Fig. 1. QMRA framework for evaluation of health risk infection associated with reuse of WWTPs effluents under different sludge bulking events. QMRA model includes hazard identification, exposure assessment, dose-response assessment, and risk management.

2.3.1. Hazard identification

Risks of infections from *E. coli*, *Salmonella*, and *Mycobacterium* were assessed via the QMRA model as a potential hazard for this study. The targeted pathogens were identified and quantified via a set of qPCR experimentation, as reported by Cui et al. (2017).

2.3.2. Exposure assessment

The exposure assessment determines the point of exposure for the populations and the amount of pathogens consumed per scenario. The frequency of exposure per year was considered, and the three exposure scenarios were:

(a) Accidental ingestion of contaminated water during recreational activities such as bathing and swimming. Data for the exposure of swimmers (children, women, and men) to recreational waters were retrieved from a study by Schets et al. (2011), conducting a questionnaire approach to approximately 19000 participants. Exposure data covered the amount of water ingested during bathing and the frequency and duration of swimming events (Supplementary Table S3). Further, categorical information was transformed into numerical data to measure mouthfuls of water during head submersion.
(b) Exposure to surface water during irrigation. This component considered the accidental ingestion of water contaminated with WWTP effluent by farmers or other farmworkers during the agricultural practices (WHO, 2006).
(c) Dietary intake of vegetables irrigated with surface water by local consumers. This scenario was described by the log-normal function ($\mu = 0.108$, $\sigma = 0.019$ mL g^{-1}) (Hamilton et al., 2006). The daily per capita intakes of vegetables for minimum, most likely, and maximum serving portions were assumed as 25, 50, and 75 g, respectively (Sant'Ana et al., 2014) (Supplementary Table S4).

The abundance of total pathogenic bacteria during sludge bulking would be more than that of the non-bulking condition. Hence, the risks associated with these three exposure scenarios were estimated at low, moderate, and high sludge bulking events in WWTPs.

The QMRA model was further used for the disinfection possibility of treated wastewater to eliminate and inactivate the pathogenic bacteria. The doses of Cl_2 residual, UV, and O_3 residual for the inactivation of *E. coli* O157:H7, *Salmonella*, and *Mycobacterium*, as well as the corresponding reduction efficiencies, were adapted from Sobsey (1989); see Supplementary Table S5. Results of the treatment cases were incorporated into the QMRA-based disinfection model by re-estimating the concentration of each pathogen for all exposure scenarios. Further, the new pathogenic abundances were used to evaluate the possible reduced risks due to the implementation of these disinfection treatments.

2.3.3. Dose-response assessment

The dose-response model was employed to assess the risk of infection for the exposed population based on the dose (concentration) of each pathogen ingested (Supplementary Table S6). The models developed for either human feeding trials of particular pathogens or waterborne disease outbreaks were employed. The Beta-Poisson Dose-Response model used for both *E. coli* and *Salmonella* is represented by Eq. (2) (Xie et al., 2016). The exponential model employed for *Mycobacterium* is given by Eq. (3) (Haas et al., 2014).

$$P_I(d) = 1 - \left[1 + \left(\frac{d}{N_{50}}\right)\left(2^{\frac{1}{\alpha}} - 1\right)\right]^{-\alpha} \quad (2)$$

$$P_I(d) = 1 - \exp(-k \cdot d) \quad (3)$$

where, $P_I(d)$ is the risk of infection due to a particular pathogen, *"d"* is the dose (concentration) of pathogen ingested in an identified volume of surface water or crops, N_{50} is the median infection dose equivalent to the number of organisms that will infect 50% of the exposed population ($N_{50} = 2.11 \times 10^6$ for *E. coli* and 1.11×10^6 for *Salmonella*), *"α"* is a dimensionless infectivity constant ($\alpha = 1.55 \times 10^{-1}$ for *E. coli* and 1.75×10^{-1} for *Salmonella*), and *"k"* is the infectivity constant ($k = 6.93 \times 10^{-4}$ for *Mycobacterium*).

2.3.4. Risk characterization

In the risk characterization step, the probability of infection for the exposed population is characterized using all the outputs of the previous actions, i.e., hazard identification, exposure assessment, and dose-response assessment. The risk of infection was determined using the formula in Eq. (4) (Sakaji and Funamizu, 1998):

$$P_I(A) = 1 - [1 - P_I(d)]^n \quad (4)$$

where, $P_I(A)$ represents the risk of infection associated with multiple exposures or annual risk, $P_I(d)$ is the risk of infection from a single exposure to a dose *"d"* of the pathogen, and *"n"* denotes the number of days of exposure to the single-dose *"d"*.

3. Results and discussion

3.1. Pathogenic bacteria detection and quantification

Table 1 lists the results of molecular real-time PCR tests generated from 10-fold dilution series of each target gene (*eae*-, *invA*-, and *IS6110*-gene).

The constructed standard curves of the qPCR assays revealed adequate amplification efficiencies for the selected primers. Both *eae* 0157:H7 and *invA* assays had almost 100% efficiency for *E. coli* and *Salmonella*, respectively. The *IS6110* assay depicted 92% efficiency for *Mycobacterium*. The estimation of these qPCR efficiencies might be influenced by many factors, *viz*., probe design, possibility of sample contamination, and primer and template concentrations, interfering with the normal reaction (Cui et al., 2017). The obtained efficiencies (Table 1) revealed acceptable accuracy due to the low possibility for the presence of inhibitors during the PCR examination, as well as the stability of the DNA extraction process (Barak et al., 2005).

As shown in Fig. 2, *Salmonella* was the most abundant specie during the study period (2270–96733 copies ng^{-1} of DNA) followed by *E. coli* O157:H7 (4133–76847 copies per ng of DNA); whereas, *Mycobacterium* was the least (542–3340 copies ng^{-1} of DNA). These pathogens have been known to cause various potential hazards to human health and receiving environments in South Africa (Ekwanzala et al., 2018). Fig. 2 also depicted that the pathogens in the WWTP effluent elevated mainly during the sludge bulking events. High bulking incidents (SVI > 150 mL g^{-1}) interfere with the compaction and settling properties of the sludge bioflocs, which could further deteriorate the WWTP performance and receiving water bodies (Martins et al., 2004). Moreover, bulking tends to increase the concentrations of TSS in the final effluent to exceed the discharge permit limitation (Deepnarain et al., 2019), which was also evident in this study.

Large quantities of *E. coli* O157:H7 were found in the wastewater effluent during high bulking episodes (Fig. 2). *E. coli* is a common fecal indicator organism of which certain strains can cause diarrheal disease, acute gastrointestinal illness, hemorrhagic colitis, and kidney failure in humans (Ayaz et al., 2014). The detection of *E. coli* in high abundance

Table 1
Estimation of parameters obtained from qPCR assays.

Parameter	*E. coli* 0157:H7	*Salmonella*	*Mycobacterium*
Efficiency (%)	103.0	109.1	92.3
Slope	−3.251	−3.121	−3.522
R^2 of slope	0.994	0.983	0.998

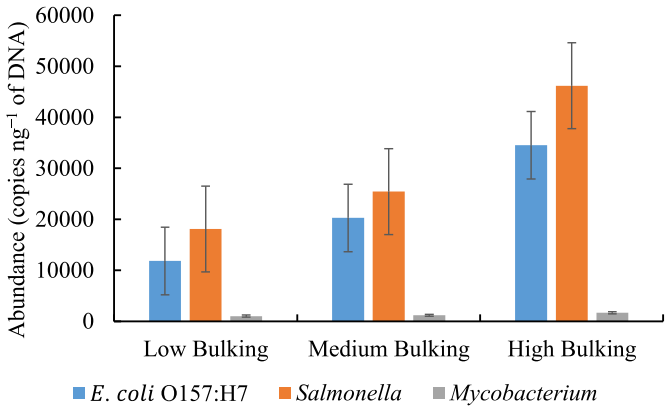


Fig. 2. Correlation between abundances of *E. coli* O157:H7, *Salmonella*, and *Mycobacterium* in WWTP effluent and sludge bulking events.

could be because this pathogenic bacterium is efficiently eliminated via a biological filtration step (Radomski et al., 2011), which was not included in the monitored WWTPs.

A positive correlation ($p < 0.05$) was also found between *Salmonella* in the final effluent and SVI (Fig. 2). During the high bulking events, a large amount of *Salmonella* could pass downstream (i.e., to the receiving environment) through the effluent weir system. Espigares et al. (2006) reported that surface water, contaminated irrigation water, and wastewater pathways could release different serotypes of *Salmonella* to the environment. This gram-negative enteric pathogen could be associated with various human diseases such as typhoid fever, gastrointestinal infection, bacteremia, and salmonellosis (Amha et al., 2017). However, the disinfection option, such as ozonation and UV, has been reported to reduce the negative impacts caused by *Salmonella* (Espigares et al., 2006; Jyoti et al., 2010).

The relative abundance of *Mycobacterium* increased with sludge bulking, recording average values of 1002, 1174, and 1670 copies ng^{-1} of DNA under the low, medium, and high sludge bulking conditions, respectively (Fig. 2). Similarly, Asvapathanagul and Olson (2017) reported that the pathogenic *Mycobacterium* could be linked to bulking in activated sludge WWTPs. Moreover, Cai and Zhang (2013) and Amha et al. (2017) revealed that mycobacteria were identified in treated wastewater and the urban watershed, causing infections to human receptors (e.g., Cervical lymphadenitis in children). Radomski et al. (2011) detected about 5.5 × 10^5 copies L^{-1} of *Mycobacterium* in the influent wastewater, which were mostly eliminated via physical-chemical decantation followed by biofiltration. The incidence of *Mycobacterium* in wastewater samples suggested that the WWTP effluent should be treated before disposal or reuse.

Hence, in the following sections, the microbial risks of these pathogenic bacteria were assessed using the QMRA technique.

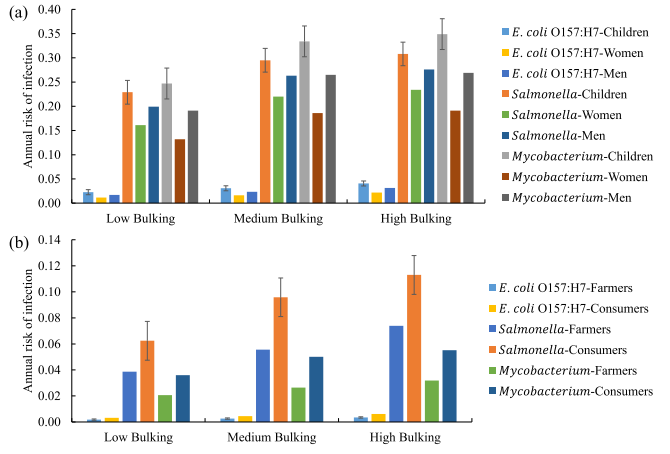


Fig. 3. Annual risk of infection associated with (a) recreational (swimming and bathing) activities for children, women, and men, and (b) irrigational practices for farmers and vegetable consumers. Exposure scenarios include the relative abundance of three waterborne pathogens, i.e., *E. coli* 0157:H7, *Salmonella*, and *Mycobacterium*.

3.2. QMRA assessment and outputs

The data shown in Fig. 3 (also represented in Supplementary Tables S7 and S8) depict the QMRA outcomes after using the wastewater effluent containing bacterial pathogens in (a) recreational activities (swimming) for children, women, and men (Fig. 3a), (b) farming practice (Fig. 3b), and (c) irrigation purposes (Fig. 3b).

3.2.1. Risk of infection during recreational activities

The QMRA values for *E. coli* O157:H7 were in the range of 0.01–0.04, considering the estimated risks associated with swimming-related activities for the three population subgroups, i.e., children, women, and men (Fig. 3a and Supplementary Table S7). The probability of annual infection for children and adults exposed to pathogenic *E. coli* O157:H7 was low during swimming, i.e., compared to the risks estimated from *Salmonella* (0.16–0.31) and *Mycobacterium* (0.13–0.35). Accordingly, medium to high risk findings were noticed for the population exposed to either *Salmonella* or *Mycobacterium* (Supplementary Table S7). In addition, the QMRA values generally increased with the rise in SVIs, suggesting that the high bulking conditions contributed to the human health risks during swimming (Fig. 3a). Moreover, in all bulking scenarios, the 10^{-4} risk level (WHO, 2006) was exceeded.

Children were almost seven-fold more likely to be infected with *Salmonella* and *Mycobacterium* compared to *E. coli* O157:H7. Moreover, children were more susceptible to bacterial infections, representing a risk level of approximately 1.5–2.0-fold higher than that for adults. The elevated risks driven by children could be attributed to the accidental water ingestion during swimming, or the transfer of pathogens from contaminated water to mouth (Sunger et al., 2018). The values are based on Schets et al. (2011), reporting the intake of water while swimming of 31–51 mL for children, as compared to 18–23 mL for women and 27–34 mL for men. According to these estimates, a higher incidence of health risks would be expected among children. Similarly, Suppes et al. (2016) reported that the annual infection risk for adults (>18) and children (≤18) while swimming were 2.2 × 10^{-2} and 2.9 × 10^{-2}, respectively. Their study hypothesized that the higher level of risk for children compared to adults could be attributed to the underdeveloped immune systems, and participation in more swimming activities such as splashing, playing, and diving (Suppes et al., 2016). In addition, Vergara et al. (2016) stated that the illness risk for children swimmers was 46% higher than for adults, suggesting that children were more susceptible to infection. Using the QMRA framework, Cui et al. (2017) found that the incidences of hand-to-mouth/eye contact for children playing with water caused the highest infection risk, regarding three pathogens *Mycobacterium avium*, *Salmonella*, and *Pseudomonas aeruginosa*.

3.2.2. Risk of infection during agricultural practice

Based on the health risks associated with the application of treated wastewater in farming activities (Fig. 3b and Supplementary Table S8), the highest QMRA data were 0.04–0.07 with *Salmonella*, which dropped to 0.02–0.03 with *Mycobacterium*, and below 0.01 with *E. coli* O157:H7. These risk estimates exceeded the WHO recommended limit of 10^{-4} to develop wastewater reuse, indicating possible adverse health impacts (WHO, 2006). Handling, transportation, and storage of contaminated water could be important infection pathways during farming (Amenu et al., 2016). Van Vu et al. (2018) reported that some farmers' actions such as touching mouth with hand, wiping mouth using arm, and other unskillful and untrained practices during work would increase the

possible risks/diseases. Moreover, farmers could accidently ingest water (about 1–5 mL (WHO, 2006)) while working in the wastewater-irrigated fields (Moazeni et al., 2017). Furthermore, nearby populations could be exposed to microbial agents when employing spray irrigation of wastewater. The relative abundance of pathogenic bacteria (*E. coli* O157:H7, *Salmonella*, and *Mycobacterium*) implied that necessary actions should be considered for wastewater reuse in irrigation.

The effect of sludge bulking scenarios on the QMRA levels was also significant ($p < 0.05$). For instance, under the low sludge bulking condition, at least 3 farmers out of 100 (90% Confidence interval: CI) were at risk of being infected due to exposure to *Salmonella*. This value increased to 7 out of 100 (90% CI) at the high bulking condition (Supplementary Table S8). Hence, it can be demonstrated that farmers suffered from higher risks due to irrigation with water contaminated by effluents of WWTPs during moderate to high bulking scenarios. In a similar work, Sampson et al. (2017) found that the annual risk of infection for farmers in frequent contact with wastewater did not meet the stringent risk benchmark of 10^{-6}. They concluded that some proactive mitigations and behavioral strategies such as the application of drip irrigation, wearing face mask, gloves, and boots, and avoiding unhygienic practices would protect the farmers against adverse health risks (Sampson et al., 2017). Moreover, awareness campaigns directed towards the farmer communities would be useful to reveal the potential risks related to the wastewater irrigation strategies.

3.2.3. Risk of infection for vegetable consumers

Treated wastewater can provide nutrients for plant growth; however, the consumption of wastewater-irrigated crops might cause various infections to end-users (Amha et al., 2015). At least 6 consumers out of 100 would be infected due to the intake of vegetables irrigated with wastewater containing *Salmonella* (Supplementary Table S8). This risk increased during high bulking events, in which the risk estimate was $1.13 \times 10^{-1} \pm 8.36 \times 10^{-4}$ (i.e., more than 10 incidences per 100; 90% C.I.) at SVI > 200 mL g^{-1}. Comparable patterns were noticed with *E. coli* O157:H7 and *Mycobacterium* pathogens during all bulking conditions. Hence, the application of wastewater in agriculture would be an important transmission route of infection, tending to increase the pathogen load and microbiological health hazard. Beaudequin et al. (2016) applied the QMRA method to characterize the human health risks associated with wastewater-irrigated lettuce. Their work depicted that some scenarios, such as lettuce washing prior to sending to market and withholding irrigation before harvesting, could reduce the pathogenic effects (Beaudequin et al., 2016). In Saudi Arabia, Balkhair (2016) found that the use of treated wastewater for crop production could have detrimental environmental and health impacts. Other health risks and routes of exposure associated with the wastewater irrigation practices have been reviewed in a previous paper by Dickin et al. (2016). Amha et al. (2017) demonstrated the need for disinfection and continuous monitoring of treated wastewater prior to reuse. These results suggested the necessity of water disinfection before land spreading to avoid potential human infections and food crop contamination.

3.3. Disinfection scenarios using QMRA outcomes

Several studies have reported that disinfection would be a viable option for complying treated wastewater with the health-based target (Kollu and Örmeci, 2012; Schijven et al., 2019). Accordingly, three disinfection scenarios, *viz.*, chlorination, UV, and ozonation, were theoretically tested under the simulated high bulking condition. The reduction efficiencies used in this investigation for the three treatment scenarios were based on data reported by Sobsey (1989) (Supplementary Table S5). The QMRA levels associated with the application of chlorine, as an oxidant, to inactivate the pathogenic bacteria within the effluent discharge are illustrated in Fig. 4(a and b); also, Supplementary Tables S9 and S10. Although chlorination decreased the risks compared to the initial estimates for the direct use of wastewater (Figs. 3 and 4),

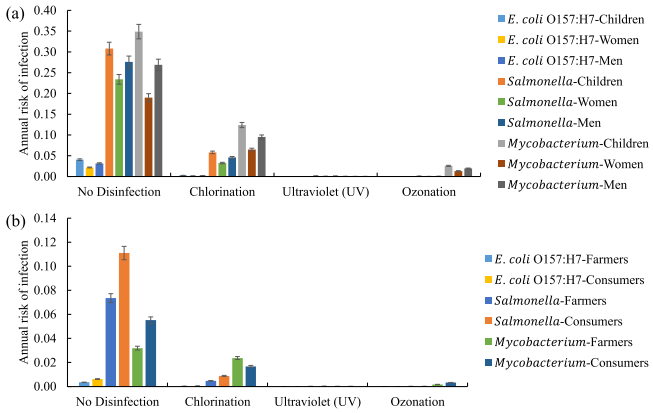


Fig. 4. Annual risk of infection at high sludge bulking condition using disinfection scenarios, *viz.*, chlorination: 0.2 mg L^{-1} (Cl_2 residual) for both *E. coli* 0157:H7 and *Salmonella*, and 0.3–1.0 mg L^{-1} (Cl_2 residual) for *Mycobacterium*. UV: 8.2 mW s cm^{-2} for both *E. coli* 0157:H7 and *Salmonella*, and 60 mW s cm^{-2} for *Mycobacterium*. Ozonation: 0.29–0.36 mg L^{-1} (O_3 residual) for both *E. coli* 0157:H7 and *Salmonella*, and 0.29–1.08 mg L^{-1} (O_3 residual) for *Mycobacterium*. In which: (a) recreational (swimming and bathing) activities with children, women, and men, and (b) irrigational practices with farmers and vegetable consumers.

the reduction of health risks was statistically insignificant ($p > 0.05$). Moreover, the probable risk of infection for both *Salmonella* and *Mycobacterium* exceeded the tolerable risk of 10^{-4} (WHO, 2006). Previous studies have demonstrated that chlorine might generate harmful disinfection by-products, causing a chemical threat to human health (Dong et al., 2017). In this context, finding alternative disinfectants to chlorine would be a suitable solution. Petterson and Stenström (2015) used QMRA to study the infectious risks related to the drinking water systems and depicted that chlorine disinfection inactivated the reference pathogens, e.g., *E. coli* O157:H7, *Giardia*, Rotavirus, *Campylobacter*, and Norovirus. In their study, the Log_{10} reductions of the five pathogens were calculated, assuming initial chlorine residuals of 0.4–1.5 mg L^{-1} (Petterson and Stenström, 2015). Sanawar et al. (2017) assessed the disinfection process to eliminate antibiotic-resistant bacteria (ARB) from marine aquaculture effluents via the QMRA method. Their study revealed that monochloramination could attain an acceptable inactivation efficiency for ARB with the release of a minimum amount of trihalomethanes compared to the case of chlorination (Sanawar et al., 2017).

In Fig. 4(a and b), UV was employed to enhance the disinfection efficacy, according to the theoretical values retrieved from Sobsey (1989) (Supplementary Table S5). The UV treatment reduced the levels of risks to below the tolerable threshold, except for the case of children exposed to *Salmonella* infection during swimming (Fig. 4a and Supplementary Table S9). Kollu and Örmeci (2012) reported that some of the self-aggregated *E. coli* could resist high UV doses such as 90 mJ cm^{-2}. It has been reported that the UV disinfectant could effectively inactivate protozoan parasites within a relatively short contact time (Sobsey, 1989). Furthermore, the UV processes are rarely associated with the release of disinfection by-products. Schijven et al. (2019) applied QMRA to assess the microbial safety of drinking water at Beerenplaat regarding several pathogens, including *Giardia*, *Campylobacter*, and *Cryptosporidium*. Their study depicted that chlorine dioxide was the essential treatment step; however, UV disinfection was more effective with the addition of chlorine dioxide (Schijven et al., 2019).

Ozone is a strong oxidizing agent that can substitute chlorine at some stages of the treatment process to disinfect a number of microorganisms, including *E. coli*, *Giardia*, and chlorine-resistant pathogens (Dong et al., 2017). In this study, ozonation showed an effective reduction of the estimated risks for all pathogens under the exposure scenarios except for

Mycobacterium (Fig. 4 and Supplementary Tables S9 and S10). The risk of *Mycobacterium* infection during ozonation was greater than the tolerable criterion of 10^{-4} recommended by WHO (2006). In a similar QMRA research study, Zhou et al. (2018) found that an O_3/UV contactor attained efficient inactivation of *E. coli* for drinking water systems. Dong et al. (2017) compared between ozonation and chlorination for the inactivation of three waterborne pathogens, namely *Giardia*, *Legionella pneumophila*, and *Cryptosporidium parvum*. Their QMRA work depicted that the treated wastewater could be applied for irrigational landscape reuse (Dong et al., 2017). Moreover, Dong et al. (2017) depicted that the microplasma ozonation technique was adequate for wastewater disinfection, showing reduced human health impacts compared to the chlorination scenario.

Based on the aforementioned risk estimates, the UV and ozonation treatments could provide adequate options for attaining the tolerable risks of infection for the exposure scenarios and populations modelled. However, the case of *Mycobacterium* infection risks after ozonation should be reassessed. This finding could be because *Mycobacteria* tend to naturally aggregate in water, making a shield that protects them against disinfection (Loret and Dumoutier, 2019). Moreover, Amha et al. (2017) reported that various strains of mycobacteria would be over 100-folds resistant to chlorine than *E. coli*.

Although the disinfection option provided essential results for reducing the human health risk associated with sludge bulking, further studies should focus on the economic feasibility of the chlorination, UV, and ozonation systems. For example, chlorination has been recognized as a cheaper option compared to other disinfection processes; however, it tends to generate toxic by-products, comprising trihalomethanes and haloacetic acids. Moreover, a dechlorination step would be required to avoid the negative impact of high chlorine levels on the aquatic environment, resulting in high capital and operating costs. The operational costs of chemical-based disinfection (i.e., Peracid) varied from 0.0114 to 0.0261 € (i.e., about $0.012–0.028 USD using an exchange rate of 1 € = $1.09 USD) for 1 m^3 of treated water (Luukkonen et al., 2015). UV requires a pre-treatment step for the removal of turbidity and suspended solids, increasing the initial capital cost of the entire disinfection process. Moreover, the UV reactors would require high-energy consumption, making it not always the best disinfection option, especially in developing countries. However, for long-term operation, UV can be a cost-effective solution compared to the chlorination-based systems (Tak and Kumar, 2017). Zhuang et al. (2015) found that the disinfection of 1 m^3 of wastewater using chlorination, UV irradiation, and UV/chlorination would cost 0.041, 0.046, and 0.034 Yuan, respectively (i.e., about $0.0058, $0.0065, and $0.0048 USD, respectively using 1 Yuan = $0.14 USD). Ozone is a strong oxidant that diffuses directly through the cell membrane, and deactivates microorganisms; however, its long-run cost is negatively influenced by the presence of organic constituents such as humic substances and fatty acids in the aqueous phase. Accordingly, each disinfection method has its constraints not only in terms of efficiency, but also in terms of feasibility, practicability, and cost. Further comparative studies on a life cycle cost analysis of the disinfection systems, including chemical utilization, energy consumption, UV lamp replacement, land footprint, manpower, plant configuration and size, and miscellaneous equipment repair, are recommended.

3.4. Environmental aspects for improving the precision of QMRA outcomes

This is the first study describing the exposure routes and the risks of infection associated with pathogenic bacteria, i.e., *E. coli* O157:H7, *Salmonella*, and *Mycobacterium*, encountered by different sludge bulking conditions from full-scale WWTPs. These microorganisms exhibit various negative impacts on the environment; e.g., *Mycobacterium* can cause diseases to lymph nodes, skeleton, and skin tissues (Loret and Dumoutier, 2019). The QMRA framework (Fig. 1) was used to examine the utilization of wastewater effluent for recreational and farming activities. Based on laboratory experiments and additional data obtained from literature, either UV or ozonation treatments would be used to avoid the microbiological risks associated with high bulking episodes (Fig. 4). The study objectives are apparent; however, some environmental aspects should be considered for future researches to enhance the QMRA outputs. For instance, either ethidium monoazide or propidium monoazide could be used to avoid the drawbacks of DNA-based methods, including the inability to discriminate between live and dead cells (Taylor et al., 2014). Furthermore, the thresholds of risk (low, moderate, and high) associated with the environmental exposures to pathogens would be identified by quantitative tools using tracer testing and/or computational fluid dynamics (Petterson and Stenström, 2015). Moreover, despite the advantages of UV, the design and scale-up stages would hinder the widespread and consistent practice. The presence of particulate matters would reduce the efficiency of UV disinfection by shielding the pathogenic microorganisms from UV irradiation (Kollu and Örmeci, 2012). Hence, more researches on the measurements of the actual pathogenic reductions are required. Although ozone was suitable during the investigation, it is unstable in water and may react with natural organic matters to form biodegradable oxygenated by-products (Dong et al., 2017). The optimum disinfectant dose for microbial inactivation should be identified to maintain the hygienic environment and economic benefits. Due to sludge bulking conditions in WWTPs, system operators, farmers, and consumers have to be sensitized about the potential occupational health threat during wastewaters reuse. Moreover, the population, especially in developing countries, needs to be educated about hygiene regulations before consuming vegetables. The findings of this work would assist private and public sectors, as well as decision-makers, for the management of risks associated with effluent wastewater pathogens in the case of bulking, i.e., particularly concerning farming activities and hygiene improvements.

4. Conclusions

In this study, a QMRA framework was applied, for the first time, to identify the applicability of treated wastewater for reuse in recreational practices for children, women, and men, and irrigation for farmers and vegetable consumers during sludge bulking in WWTPs. It was concluded that:

- During the sludge bulking events, the QMRA outputs showed unacceptable risk levels due to the high abundance of the pathogens, *viz.*, *E. coli*, *Salmonella*, and *Mycobacterium*, in the WWTP effluent.
- Due to the risk estimates, the bacterial infection among children was 1.5–2.0-fold higher than the case of adults during swimming-related activities, and farmers would suffer from higher microbial risks due to irrigation with effluents of WWTPs experiencing moderate to high bulking scenarios.
- For the disinfection scenarios, the estimated risks of infection after chlorination were still exceeding the tolerable risk level for both *Salmonella* and *Mycobacterium*; UV treatment reduced the levels of risks below the tolerable threshold, except for the case of children exposed to *Salmonella* infection during swimming; for ozonation, the *Mycobacterium* infection resulted in higher risks than the allowable criterion of 10^{-4}.
- Our future work will focus on enhancing the precision of QMRA outcomes using some environmental aspects pointed out in this study.

Declaration of competing interests

The authors declare that there is no conflict of interest.

CRediT authorship contribution statement

Nashia Deepnarain: Validation, Data curation. **Mahmoud Nasr:**

Writing - original draft, Formal analysis. **Isaac Dennis Amoah:** Formal analysis. **Abimbola Motunrayo Enitan-Folami:** Methodology. **Poovendhree Reddy:** Supervision. **Thor Axel Stenström:** Writing - review & editing. **Sheena Kumari:** Data curation. **Faizal Bux:** Funding acquisition.

Acknowledgements

The authors would like to acknowledge the Water Research Commission (WRC), South Africa, and the National Research Foundation of South Africa (UID 84166). The second author would like to acknowledge Nasr Academy for Sustainable Environment (NASE).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2020.110660.

References

Amenu, K., Shitu, D., Abera, M., 2016. Microbial contamination of water intended for milk container washing in smallholder dairy farming and milk retailing houses in southern Ethiopia. SpringerPlus 5, 1195.

Amha, Y., Anwar, M., Kumaraswamy, R., Henschel, A., Ahmad, F., 2017. *Mycobacteria* in municipal wastewater treatment and reuse: microbial diversity for screening the occurrence of clinically and environmentally relevant species in arid regions. Environ. Sci. Technol. 51 (5), 3048–3056.

Amha, Y., Kumaraswamy, R., Ahmad, F., 2015. A probabilistic QMRA of *Salmonella* in direct agricultural reuse of treated municipal wastewater. Water Sci. Technol. 71 (8), 1203–1211.

Amoah, I., Reddy, P., Seidu, R., Stenström, T., 2018a. Concentration of soil-transmitted helminth eggs in sludge from South Africa and Senegal: a probabilistic estimation of infection risks associated with agricultural application. J. Environ. Manag. 206, 1020–1027.

Amoah, I., Reddy, P., Seidu, R., Stenström, T., 2018b. Removal of helminth eggs by centralized and decentralized wastewater treatment plants in South Africa and Lesotho: health implications for direct and indirect exposure to the effluents. Environ. Sci. Pollut. Control Ser. 25 (13), 12883–12895.

Asvapathanagul, P., Olson, B., 2017. Improving qPCR methodology for detection of foaming bacteria by analysis of broad-spectrum primers and a highly specific probe for quantification of *Nocardia* spp. in activated sludge. J. Appl. Microbiol. 122 (1), 97–105.

Ayaz, N., Gencay, Y., Erol, I., 2014. Prevalence and molecular characterization of sorbitol fermenting and non-fermenting Escherichia coli O157:H7+/H7- isolated from cattle at slaughterhouse and slaughterhouse wastewater. Int. J. Food Microbiol. 174, 31–38.

Balkhair, K., 2016. Microbial contamination of vegetable crop and soil profile in arid regions under controlled application of domestic wastewater. Saudi J. Biol. Sci. 23 (1), S83–S92.

Barak, J., Sananikone, K., Delwiche, M., 2005. Comparison of primers for the detection of pathogenic Escherichia coli using real-time PCR. Lett. Appl. Microbiol. 41 (2), 112–118.

Beaudequin, D., Harden, F., Roiko, A., Mengersen, K., 2016. Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. Sci. Total Environ. 541, 1393–1409.

Brosch, R., Gordon, S., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Cole, S., 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc. Natl. Acad. Sci. U.S.A. 99 (6), 3684–3689.

Cai, L., Zhang, T., 2013. Detecting human bacterial pathogens in wastewater treatment plants by a high-throughput shotgun sequencing technique. Environ. Sci. Technol. 47 (10), 5433–5441.

Cui, Q., Fang, T., Huang, Y., Dong, P., Wang, H., 2017. Evaluation of bacterial pathogen diversity, abundance and health risks in urban recreational water by amplicon next-generation sequencing and quantitative PCR. J. Environ. Sci. (China) 57, 137–149.

Deepnarain, N., Nasr, M., Kumari, S., Stenström, T., Reddy, P., Pillay, K., Bux, F., 2019. Decision tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant. Process Saf. Environ. Protect. 126, 25–34.

Dickin, S., Schuster-Wallace, C., Qadir, M., Pizzacalla, K., 2016. A review of health risks and pathways for exposure to wastewater Use in Agriculture. Environ. Health Perspect. 124 (7), 900–909.

Dong, S., Li, J., Kim, M.-H.P.-J., Eden, J., Guest, J.N., 2017. Human health trade-offs in the disinfection of wastewater for landscape irrigation: microplasma ozonation: vs. chlorination. Environ. Sci. J. Integr. Environ. Res.: Water Research and Technology 3 (1), 106–118.

Ekwanzala, M., Dewar, J., Kamika, I., Momba, M., 2018. Systematic review in South Africa reveals antibiotic resistance genes shared between clinical and environmental settings. Infect. Drug Resist. 11, 1907–1920.

Enitan, A., Kumari, S., Swalaha, F., Adeyemo, J., Ramdhani, N., Bux, F., 2014. Kinetic modelling and characterization of microbial community present in a full-scale UASB reactor treating brewery effluent. Microb. Ecol. 67 (2), 358–368.

Espigares, E., Bueno, A., Espigares, M., Gálvez, R., 2006. Isolation of *Salmonella serotypes* in wastewater and effluent: effect of treatment and potential risk. Int. J. Hyg Environ. Health 209 (1), 103–107.

Falkinham, J., 2009. Surrounded by mycobacteria: *Nontuberculous mycobacteria* in the human environment. J. Appl. Microbiol. 107, 356–367.

Haas, C., Rose, J., Gerba, C., 2014. Quantitative Microbial Risk Assessment, second ed. John Wiley and Sons, New York.

Hamilton, A., Stagnitti, F., Premier, R., Boland, A.-M., Hale, G., 2006. Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. Appl. Environ. Microbiol. 72 (5), 3284–3290.

Jyoti, A., Ram, S., Vajpayee, P., Singh, G., Dwivedi, P., Jain, S., Shanker, R., 2010. Contamination of surface and potable water in South Asia by *Salmonellae*: culture-independent quantification with molecular beacon real-time PCR. Sci. Total Environ. 408 (6), 1256–1263.

Kollu, K., Örmeci, B., 2012. Effect of particles and bioflocculation on ultraviolet disinfection of Escherichia coli. Water Res. 46 (3), 750–760.

Loret, J.-F., Dumoutier, N., 2019. Non-tuberculous mycobacteria in drinking water systems: a review of prevalence data and control means. Int. J. Hyg Environ. Health 222 (4), 628–634.

Luukkonen, T., Heyninck, T., Rämö, J., Lassi, U., 2015. Comparison of organic peracids in wastewater treatment: disinfection, oxidation and corrosion. Water Res. 85, 275–285.

Martins, A., Pagilla, K., Heijnen, J., Van Loosdrecht, M., 2004. Filamentous bulking sludge - a critical review. Water Res. 38 (4), 793–817.

Moazeni, M., Nikaeen, M., Hadi, M., Moghim, S., Mouhebat, L., Hatamzadeh, M., Hassanzadeh, A., 2017. Estimation of health risks caused by exposure to enteroviruses from agricultural application of wastewater effluents. Water Res. 125, 104–113.

Petterson, S., Stenström, T., 2015. Quantification of pathogen inactivation efficacy by free chlorine disinfection of drinking water for QMRA. J. Water Health 13 (3), 625–644.

Radomski, N., Betelli, L., Moilleron, R., Haenn, S., Moulin, L., Cambau, E., Lucas, F., 2011. Mycobacterium behavior in wastewater treatment plant, A bacterial model distinct from Escherichia coli and enterococci. Environ. Sci. Technol. 45 (12), 5380–5386.

Sakaji, R., Funamizu, N., 1998. In: Asano, T. (Ed.), Microbial Risk Assessment and its Role in the Development of Wastewater Reclamation Policy, Wastewater Reclamation and Reuse, vol. 10. CRC Press, Boca Raton, Fla, USA, pp. 705–756. *1998*.

Sampson, A., Owusu-Ansah, E.-G., Mills-Robertson, F., Ayi, I., Abaidoo, R., Hald, T., Permin, A., 2017. Probabilistic quantitative microbial risk assessment model of farmer exposure to Cryptosporidium spp. in irrigation water within Kumasi Metropolis-Ghana. Microbial Risk Analysis 6, 1–8.

Sanawar, H., Xiong, Y., Alam, A., Croué, J.-P., Hong, P.-Y., 2017. Chlorination or monochloramination: balancing the regulated trihalomethane formation and microbial inactivation in marine aquaculture waters. Aquaculture 480, 94–102.

Sant'Ana, A., Franco, B., Schaffner, D., 2014. Risk of infection with *Salmonella* and *Listeria monocytogenes* due to consumption of ready-to-eat leafy vegetables in Brazil. Food Contr. 42, 1–8.

Schets, F., Schijven, J., de Roda Husman, A., 2011. Exposure assessment for swimmers in bathing waters and swimming pools. Water Res. 45 (7), 2392–2400.

Schijven, J., Teunis, P., Suylen, T., Ketelaars, H., Hornstra, L.R., 2019. QMRA of adenovirus in drinking water at a drinking water treatment plant using UV and chlorine dioxide disinfection. Water Res. 158, 34–45.

Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A., Harada, H., 1999. Fluorescence in situ hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Appl. Environ. Microbiol. 65 (3), 1280–1288.

Sobsey, M., 1989. Inactivation of health-related microorganisms in water by disinfection processes. Water Sci. Technol. 21 (3), 179–195.

Soller, J., Schoen, M., Bartrand, T., Ravenscroft, J., Ashbolt, N., 2010. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Res. 44 (16), 4674–4691.

Sunger, N., Hamilton, K., Morgan, P., Haas, C., 2018. Comparison of pathogen-derived 'total risk' with indicator-based correlations for recreational (swimming) exposure. Environ. Sci. Pollut. Control Ser. 1–11. https://doi.org/10.1007/s11356-018-1881-x.

Suppes, L., Canales, R., Gerba, C., Reynolds, K., 2016. *Cryptosporidium* risk from swimming pool exposures. Int. J. Hyg Environ. Health 219 (8), 915–919.

Tak, S., Kumar, A., 2017. Chlorination disinfection by-products and comparative cost analysis of chlorination and UV disinfection in sewage treatment plants: Indian scenario. Environ. Sci. Pollut. Control Ser. 24 (34), 26269–26278.

Taylor, M., Bentham, R., Ross, K., 2014. Limitations of using propidium monoazide with qPCR to discriminate between live and dead Legionella in biofilm samples. Microbiol. Insights 7, 15–24.

USEPA, 2012. Recreational Water Quality Criteria (Office of Water 820-F-12-058). U.S. Environmental Protection Agency (U.S. EPA), Washington, DC.

Van Vu, T., Pham, P., Winkler, M., Zurbrügg, C., Zinsstag, J., Tran, B., Nguyen-Viet, H., 2018. Estimation of Involuntary Excreta Ingestion Rates in Farmers during Agricultural Practices in Vietnam. *Human And Ecological Risk Assessment*, pp. 1–11. Article (in press).

Vergara, G., Rose, J., Gin, K., 2016. Risk assessment of noroviruses and human adenoviruses in recreational surface waters. Water Res. 103, 276–282.

WHO, 2001. Water Quality: Guidelines, Standards and Health. Assessment of Risk and Risk Management for Water Related Infectious Diseases. World Health Organization, Geneva.
WHO, 2006. Guidelines for the Safe Use of Wastewater, Excreta and Greywater, vol. 4. World Health Organization (WHO), Geneva.
WHO, 2011. Guidelines for Drinking-Water Quality, fourth ed. ed. World Health Organization, Geneva.
WHO, 2016. Quantitative Microbial Risk Assessment: Application for Water Safety Management. World Health Organization, Geneva.
Xie, G., Roiko, A., Stratton, H., Lemckert, C., Dunn, P., Mengersen, K., 2016. A generalized QMRA beta-Poisson dose-response model. Risk Anal. 36 (10), 1948–1958.
Zhou, L., Echigo, S., Nakanishi, T., Yamasaki, S., Itoh, S., 2018. Development of a multiphase inactivation model for an advanced oxidation process and uncertainty analysis in quantitative microbial risk assessment. Ozone: Sci. Eng. 40 (2), 79–92.
Zhuang, Y., Ren, H., Geng, J., Zhang, Y., Zhang, Y., Ding, L., Xu, K., 2015. Inactivation of antibiotic resistance genes in municipal wastewater by chlorination, ultraviolet, and ozonation disinfection. Environ. Sci. Pollut. Control Ser. 22 (9), 7037–7044.