A logistic model for the remediation of filamentous bulking in a biological nutrient removal wastewater treatment plant

Nashia Deepnarain, Sheena Kumari, Jordache Ramjith, Feroz Mahomed Swalaha, Valter Tandoi, Kriveshin Pillay and Faizal Bux

ABSTRACT

Biological nutrient removal (BNR) systems across the globe frequently experience bulking and foaming episodes, which present operational challenges such as poor sludge settling due to excessive filamentous bacteria. A full-scale BNR plant treating primarily domestic wastewater was monitored over a period of 1 year to investigate filamentous bacterial growth response under various plant operating parameters. Identification of filamentous bacteria by conventional microscopy and fluorescent in situ hybridisation indicated the dominance of Eikelboom Type021N, Thiothrix spp., Eikelboom Type 1851 and Eikelboom Type 0092. A cumulative logit model (CLM) was applied to elucidate significant relationships between the filamentous bacteria and plant operational parameters. The model could predict the potential abundance of dominant filamentous bacteria in relation to wastewater treatment plant operational parameters. Data obtained from the model corroborated with previous findings on the dominance of most filaments identified, except for Type 0092, which exhibited some unique traits. With further validation, the model could be successfully applied for identifying specific parameters which could contribute towards filamentous bulking, thus, providing a useful tool for regulating specific filamentous growth in full-scale wastewater treatment plants.

Key words | activated sludge, bulking, cumulative logit model, Eikelboom Type 0092, filamentous bacteria, Type 021N

INTRODUCTION

The activated sludge (AS) process with several modifications in its design and configuration, continues to be the most widely used technology for the treatment of industrial and domestic wastewater (Banadda et al. 2005). The AS process was gradually engineered into various existing biological nutrient removal (BNR) processes which further incorporated anoxic and/or anaerobic selectors to reduce the organic load, nitrogen and/or phosphorus (Grady et al. 2012). The selective tanks under optimum operating conditions facilitate carbon removal, nitrification, denitrification and phosphorus removal via the functional microorganisms that form a floc. An AS floc can be defined as a complex microbial community comprising the different functional groups of micro-organisms that form microcolonies attached to a filamentous backbone. Filamentous bacteria therefore play a key role in the AS floc formation by providing the necessary structural support for settling (Madoni et al. 2000; Gerardi 2006). However, since the inception of the AS process, many studies have shown that the excessive growth of filamentous bacteria had resulted in bulking and/or foaming conditions which have plagued wastewater treatment plant (WWTP) operations for decades (Muvvoto et al. 1999; Richard et al. 2003; Tandoi et al. 2006; Mielczarek et al. 2002a).

Thereby, bulking can be described as sludge which either settles very slowly or is unable to settle at all, hence causing sludge to float onto the surface of final settling tanks and escape with the effluent water into receiving environments (Jenkins et al. 2004). A sludge volume index (SVI) greater than 150 mL g⁻¹ was often used to determine bulking sludge (Jenkins et al. 2004; Liu & Fang 2010; Seviour 2010).
However, SVI was often criticised as unreliable and not a clear representation of settled sludge, usually when the sludge was highly concentrated (Martins et al. 2004; Seviour 2010; van Haandel & van der Lubbe 2012). Subsequently, diluted SVI (DSVI) became a method of choice in many countries which was determined by diluted sludge in batches until the diluted suspension after settling measured <200 mL (Seviour 2010). Using this method, sludge concentration can be kept constant thereby eliminating the influence of sludge concentration on the measured value (Jenkins et al. 2004; Seviour 2010; van Haandel & van der Lubbe 2012).

Proliferation of filamentous bacteria has been associated with various operational and environmental parameters, some of which include substrate concentration of the inflowing wastewater, low dissolved oxygen (DO), low food to micro-organism ratio (F/M), the presence of sulphur, septic sludge, high grease and oil content, nutrient deficiency, sludge age and temperature (Jenkins et al. 2004; Martins et al. 2004; Rossetti et al. 2005; Mielczarek et al. 2012a; Aygun et al. 2013). Some of these filamentous organisms have been studied at the laboratory scale; however, owing to their complex behaviour, a thorough evaluation of these organisms in full-scale plants has proved to be very difficult and cumbersome (Kanagawa et al. 2000; Rossetti et al. 2005; Naidoo 2005; Ramothokang et al. 2006; Kragelund et al. 2008; Mielczarek et al. 2012b).

Many of these organisms cannot be cultured using current laboratory techniques and thus understanding their ecophysiological traits had most often been ambiguous and unsuccessful (Richard et al. 2005; Martins et al. 2004; Mielczarek et al. 2012b). Thus far, there is no universal strategy to overcome the proliferation of filamentous bacteria in full-scale WWTPs (Da Motta et al. 2003; Mielczarek et al. 2012b). Knowledge on specific nutrient requirements and operational parameters that enhance the activity of a single filamentous bacterial type may be useful in controlling their growth and proliferation (Martins et al. 2004).

Various statistical models are currently being used to explain the complexities of biological systems in relation to the function of micro-organisms (Belanche et al. 2000; Lou & Zhao 2012). Lou & 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 thoroughly explained and the mechanisms of the inner signal processing of the model were unknown. Moreover, 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 (Singh et al. 2008). All three models are data-driven (based on input and the output data generated), which analyse extensive data sets and it is formed by computational programming. However, due to the complexity of understanding and interpreting these models, its application on filamentous bacteria was found to be limited.

Filamentous quantification in full-scale WWTPs is usually performed based on filamentous indexing (FI) using a scale from 0 (no filaments present) to 6 (excessive filaments, present in all flocs), which are actually categorical or ordinal data (Jenkins et al. 2004). This type of data can be used in cumulative link models with a logit link and are now exploited as a flexible regression framework in order to allow for in-depth data analysis (Christensen 2012). Among various statistical approaches, the cumulative logit model (CLM) had been successfully applied in social and biomedical sciences particularly dealing with categorical data (Park 2005; Agresti 2007). However, the model has not been implemented in AS systems, relating specifically to filamentous organisms and sludge bulking. Hence, this study enables researchers to further expand the surveys of filamentous bacteria using the CLM, which could be used as a tool to understand factors/conditions which may influence the growth of these organisms.

The CLM can be described as a statistical 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 aim of this study was to evaluate specific plant operational parameters which could strongly affect and predict the dominance of filamentous bacteria in a full-scale WWTP using a simple logistic model approach. The method of selection exploited the dynamics of the model itself, such that it could be used in surveys across many WWTPs.

**MATERIALS AND METHODS**

**Wastewater treatment plant description**

A full-scale WWTP treating domestic wastes in the KwaZulu-Natal region was selected for the current study. The plant treats 3.4 ML d⁻¹ average dry weather flow of mainly domestic wastewater (95%) with an average influent flow rate and organic loading rate of 3,100 m³ d⁻¹ and 2,980 kg COD m⁻³ d⁻¹, respectively. The plant was designed
according to the University of Cape Town (UCT) process configuration which currently operates in a simple restricted aeration mode without an internal ‘a’ recycle (Figure 1). The plant complied to the Department of Water Affairs and Forestry (Government Gazette Act No. 36, 1998), in terms of the South African water quality standards, however, the AS system often experienced problems of poor sludge settling due to the presence of excessive filamentous bacteria.

**Sampling and plant data**

Mixed liquor sludge grab samples (1 L) were analysed biweekly from various compartments viz., anaerobic (760 m³), anoxic (1,060 m³), aeration (4,380 m³), secondary clarifier and the return activated sludge stream, over a 1-year period. Ammonia (NH₄-N), nitrates (NO₃), nitrites (NO₂) and phosphorus (PO₄-P) were analysed according to protocols outlined by the US Environmental Protection Agency (USEPA 2010), using an Aquakem Gallery Photometric Auto-analyser (Thermo Scientific, Germany). Chemical oxygen demand (COD) and mixed liquor suspended solids (MLSS) were measured according to Standard Methods (APHA 1998). The SVI and F/M ratio were calculated according to Jenkins et al. (2004). Temperature and DO were measured using a multi-parameter-YSI model, 556 MPS system (Yellow Spring Systems, USA).

**Presumptive identification and semi-quantification of filamentous bacteria**

Filamentous bacteria were presumptively identified by a group of trained individuals using microscopic analyses. The AS sludge samples were assessed using the morphological classification system according to methods described by Eikelboom (2000) and Jenkins et al. (2004). Wet mounts, Gram and Neisser stains, PHB and sulphur tests were prepared on the sludge samples to assess morphological traits (floc size, shape, filamentous structure) and filamentous abundance (Jenkins et al. 2004). The abundance of filamentous bacteria was determined using a subjective scoring method, where the assessment of filaments were semi-quantified using a seven point FI scale (FI: 0–6) (Eikelboom 2000).

**Molecular identification of filamentous bacteria**

The identification of filamentous bacteria was further confirmed by fluorescent in situ hybridisation (FISH) analysis, using 16S rRNA-targeted oligonucleotide probes which targeted specific filamentous bacteria (Table 1). These probes were labelled with 5(6)-carboxyfluorescein (FAM) ester dye (Inqaba Biotech, South Africa). The FISH procedure was followed according to guidelines by Nielson et al. (2009), and the slides were performed using an Axiolab Apotome microscope (Carl Zeiss, Germany) containing the FAM fluorochrome filter set (Manz et al. 1992). Image analyses were examined with Zeiss Axio Vision Release 4.6 imaging software.

**Application of the CLM using statistical analysis software**

Logistic regression analysis was performed using the statistical analysis software package, (PROC LOGISTIC, SAS V 9.2). The logistic regression analysis was selected to compare against plant operational parameters which were also grouped accordingly into low (0) and high (1) levels (Table 2). The model was initially screened across the selected WWTP for operational parameters which illustrated the most significant effect on the dominant filamentous populations (data not shown). Among these operational parameters viz., pH, COD and NH₄-N were incorporated into the model and were categorised into low and high levels based on median values and whereas others viz., DO, F/M and temperature were selected and

![Figure 1](https://iwaponline.com/wst/article-pdf/72/3/391/467368/wst072030391.pdf)
Table 1 | Oligonucleotide probes used for the detection of filamentous bacteria in AS

<table>
<thead>
<tr>
<th>Probe</th>
<th>Sequence</th>
<th>Filamentous bacteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>G123 T</td>
<td>5'-CCT TCC GAT CTC TAT GCA-3'</td>
<td><em>Thiobrix</em> spp., T. eikelboomii, <em>T. nivea</em>, <em>T. unzii</em>, <em>T. fructosivorans</em>, <em>T. defluvi</em>, Eikelboom Type 021N group I, II, III</td>
</tr>
<tr>
<td>G1B</td>
<td>5'-TGT GTT CGA GTT CCT TGC-3'</td>
<td>Eikelboom Type 021N group I</td>
</tr>
<tr>
<td>GNS B94</td>
<td>5'-AAA CCA CAC GCT CCG CT-3'</td>
<td>Type 0041, Type 0675, Type 1851</td>
</tr>
<tr>
<td>CFX 1223</td>
<td>5'-CCA TTG TAG CGT GTG TGT MG-3'</td>
<td>Type 0041, Type 0675, Type 1851</td>
</tr>
<tr>
<td>SNA</td>
<td>5'-CAT CCC CCT CTA CCG TAC-3'</td>
<td><em>Sphaerotilus natans</em></td>
</tr>
<tr>
<td>Type CHL 1851</td>
<td>5'-AAT TCC ACG AAC CTC TGC CA-3'</td>
<td>Type 1851</td>
</tr>
<tr>
<td>MPA60</td>
<td>5'-GGA TGG CCG TCG ACT-3'</td>
<td><em>M. parvicella</em></td>
</tr>
<tr>
<td>Goam 192</td>
<td>5'-CAC CCA CCC CCA TGC AGG-3'</td>
<td><em>Gordonia amarae</em></td>
</tr>
<tr>
<td>NLI M1 91</td>
<td>5'-CGC CAC TAT CTT CTC AGT-3'</td>
<td><em>Nostocoida limicola</em></td>
</tr>
<tr>
<td>NLI M1111 301</td>
<td>5'-CCC AGT GTG CCG GCC CAC-3'</td>
<td><em>Nostocoida limicola</em> III strains</td>
</tr>
</tbody>
</table>

Table 2 | Plant operational parameters represented as two groups: low (0) and high (1), for the ordinal logistic regression model analysis

<table>
<thead>
<tr>
<th>Factors</th>
<th>Group 0 (low)</th>
<th>Group 1 (high)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>&lt;20 °C</td>
<td>≥20 °C</td>
<td>Jenkins <em>et al.</em></td>
</tr>
<tr>
<td>pH</td>
<td>&lt;7.15</td>
<td>≥7.15</td>
<td>Martin <em>et al.</em></td>
</tr>
<tr>
<td>DO</td>
<td>&lt;1.1 mg O₂L⁻¹</td>
<td>≥1.1 mg O₂L⁻¹</td>
<td>Martin <em>et al.</em></td>
</tr>
<tr>
<td>F/M</td>
<td>&lt;0.1 kg COD/ kg MLSS d⁻¹</td>
<td>≥0.1 kg COD/ kg MLSS d⁻¹</td>
<td>Jenkins <em>et al.</em></td>
</tr>
<tr>
<td>Influent COD</td>
<td>&lt;750 mg CODL⁻¹</td>
<td>≥750 mg CODL⁻¹</td>
<td></td>
</tr>
<tr>
<td>Influent NH₄-N</td>
<td>&lt;32 mgL⁻¹</td>
<td>≥32 mgL⁻¹</td>
<td></td>
</tr>
</tbody>
</table>

categorised based on the literature (Martins *et al.* 2003; Jenkins *et al.* 2004).

**Ordinal logistic regression**

The Akaike information criterion (AIC) was used to evaluate the fit of the model and was calculated as 

\[ \text{AIC} = -2 \log L + 2((k-1) + s) \]

where \( k \) is the number of levels of the dependent variable. \( \log L \) is the log-likelihood function and \( s \) is the number of predictors in the model. AIC is used for the comparison of non-nested models on the same sample. Ultimately, the model with the smallest AIC is considered the best, although the AIC value itself is not meaningful. The Schwarz criterion (SC) is defined as 

\[ \text{SC} = -2 \log L + (k-1) + s \log(\sum f_i) \]

where \( f_i \) is the frequency value of the \( i \)th observation, and \( k \) and \( s \) are defined above. Like AIC, the SC value penalises for the number of predictors in the model and the smallest SC is most desirable, the SC value is not meaningful by itself (SAS Annotated output 2012).

The cumulative logits are defined as

\[ L_j = \logit \left( \frac{F_j(x)}{1 - F_j(x)} \right) = \log \left( \frac{\pi_1(x) + \cdots + \pi_j(x)}{\pi_{j+1}(x) + \cdots + \pi_J(x)} \right), j = 1, \ldots, J - 1. \]

Each cumulative logit uses all J response categories. The cumulative logit \( L_j \) is an ordinary logit model for a binary response in which categories 1 to \( j \) form a single category, and categories \( j + 1 \) to \( J \) form the second category. The CLM incorporates all \( J - 1 \) cumulative logits for a \( J \)-category response into a single, parsimonious model (Agresti 2007).

Let \( L_j(x) = \logit(F_j(x)), j = 1, \ldots, J - 1 \) where \( F_j(x) = P(Y \leq j|x) \) is the cumulative probability for response category \( j \), when the explanatory variables take value \( x \) (Yay & Akinci 2009; Gardiner & Luo 2011). To include effects of the explanatory variables, proportional odds are expressed as follows:

\[ L_j(x) = a_j + \beta' x, j = 1, \ldots, J - 1 \]

The \( a_1, \ldots, a_{j-1} \) parameters are non-decreasing in \( j \) and are known as the intercepts or the ‘cut-off points’ (Yay & Akinci 2009). The odds ratio (OR) of cumulative probabilities is called a cumulative OR. The log of the cumulative OR is proportional to the distance between the values of the explanatory variables, with the same proportionality constant applying to each cut-off point. The
interpretation of the proportional odds model is that the odds of making response \( \leq j \) are \( \exp[\beta'(x_1 - x_2)] \) times higher at \( x = x_1 \) than at \( x = x_2 \). The model assumes the effect of the variable on the odds of response below category \( j \) is the same for all \( j \). Proportional odds model has been found to be the most popular model for ordinal data (Gameroff 2005). This model describes the log-odds of two cumulative probabilities, one less-than and the other greater-than type. Proportional odds measure how likely the response is to be ranked \( j \) or below versus being ranked higher than \( j \). OR considers the effects of the independent variable (PENNSTATE 2013).

### RESULTS AND DISCUSSION

#### Floc structure, SVI and the dominant filamentous bacteria

The impact of filamentous bacteria on the sludge floc structure was determined using conventional microscopic and staining techniques. Under high bulking conditions, the sludge flocs were dominated by excess filamentous bacteria leading to open floc structures (Fl: 4–6). However, under non-bulking conditions, filaments were more intact within the floc (Fl: 3). These flocs were found irregular and diffuse during the winter period and slightly compact during the summer and remaining sampling periods (data not shown).

Several investigations had reported that high SVI and DSVI values were associated with a high abundance of filamentous bacteria (Da Motta et al. 2002; Jenkins et al. 2004; Banadda et al. 2005; Tandoi et al. 2006), and that filamentous bulking most often occurred during winter and spring seasons (Kruit et al. 2002; Lou & Zhao 2012; Mielczarek et al. 202b). In this research study, bulking conditions were relatively high (average SVI: 206.22 mL g\(^{-1}\)) during autumn and winter (i.e. March–August), at an average temperature of 20.40 °C (Table 3). Alternatively, SVI values were occasionally low (average SVI: 144.50 mL g\(^{-1}\)) during spring and summer (average temperature: 27.05 mL g\(^{-1}\)). However, during the warmer months, when filamentous bacteria were highly abundant (Fl: 4–6), the observed SVI values were much lower than expected (120–130 mL g\(^{-1}\), (Table 3). During this phase, the MLSS concentrations were exceptionally high and possibly a DSVI should have been a method of choice to analyse sludge settling (Table 3). Owing to the discrepancy of SVI values in this

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Average plant operating conditions for Kingsburgh WWTP during the study period (January-December 2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Influent COD (mg COD L(^{-1}))</td>
<td>624.84 ± 35.12</td>
</tr>
<tr>
<td>CRD removal efficiency (%)</td>
<td>92.44 ± 3.55</td>
</tr>
<tr>
<td>Influent NH(_4)-N (mg L(^{-1}))</td>
<td>94.48 ± 3.68</td>
</tr>
<tr>
<td>CRD removal efficiency (%)</td>
<td>92.50 ± 3.12</td>
</tr>
<tr>
<td>Influent MLSS (mg L(^{-1}))</td>
<td>29.02 ± 1.76</td>
</tr>
<tr>
<td>MLSS mg L(^{-1})</td>
<td>28.02 ± 1.60</td>
</tr>
<tr>
<td>SVI (mL g(^{-1}))</td>
<td>5.30 ± 1.10</td>
</tr>
<tr>
<td>Aeration PO(_4)-P (mg PL(^{-1}))</td>
<td>0.67 ± 0.39</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>19.86 ± 4.52</td>
</tr>
<tr>
<td>pH</td>
<td>7.09 ± 0.02</td>
</tr>
</tbody>
</table>

by guest

Downloaded from https://waponline.com/wst/article-pdf/72/3/391/467368/wst072030391.pdf
In the current study, the most frequently occurring filamentous bacteria in the aeration tank in decreasing order were Type 021N, Thiothrix spp., Type 1851, Type 0092, Type 0041 and S. natans. The FISH technique confirmed identification on leading filamentous types, however, due to the sensitivity and difficulties in quantification analysis, FISH techniques were employed only once every four months. Moreover, with vigilant monitoring via simple staining methods by trained individuals, the FI scale proved to be a relatively quick and robust technique to semi-quantify filamentous bacteria. Thus, the sludge samples could be monitored on a weekly basis without difficulty. The composition of the filamentous bacterial population showed significant variations in their individual dominance across the study period (Figure 2). Results showed that some filaments (Type 1851 and Thiothrix spp.) dominated during spring and summer, while a few others (S. natans and Type 0041) showed a slight drop during summer (Figure 2). The dominance of these organisms can also be influenced by numerous other factors, including wastewater composition and plant operational parameters apart from temperature (Martins et al. 2004; Mielczarek et al. 2012).

A logistic approach to analyse the significant relationships between filamentous bacterial dominance and plant operational parameters

An ordinal logistic model (PROC LOGISTIC, SAS V 9.2) with a cumulative logit link function was used to determine the significant relationships of individual filamentous bacteria with plant operational parameters (Table 4). Results

---

**Table 4** Cumulative logistic output, illustrating significant and non-significant relationships between organisms and operational parameters

<table>
<thead>
<tr>
<th>Filamentous bacteria</th>
<th>Influent COD</th>
<th>Influent NH₄-N</th>
<th>DO</th>
<th>F/M</th>
<th>Temperature</th>
<th>pH</th>
<th>PO₄-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1851</td>
<td>*0.0059</td>
<td>*0.0077</td>
<td>*0.014</td>
<td>*0.0001</td>
<td>0.603</td>
<td>0.1939</td>
<td>*0.0163</td>
</tr>
<tr>
<td>Type 021N</td>
<td>*0.0380</td>
<td>*0.0604</td>
<td>0.2277</td>
<td>*0.0064</td>
<td>0.2131</td>
<td>0.6833</td>
<td>0.3668</td>
</tr>
<tr>
<td>Type 0092</td>
<td>*0.0041</td>
<td>*0.0043</td>
<td>*0.008</td>
<td>*0.0160</td>
<td>*0.0028</td>
<td>0.3440</td>
<td>0.7628</td>
</tr>
<tr>
<td>Thiothrix spp.</td>
<td>*0.0172</td>
<td>0.3970</td>
<td>0.1844</td>
<td>*0.0035</td>
<td>0.7691</td>
<td>0.1860</td>
<td>0.1158</td>
</tr>
<tr>
<td>Type 0041</td>
<td>0.7663</td>
<td>*0.0123</td>
<td>0.7880</td>
<td>0.3850</td>
<td>0.5441</td>
<td>*0.0175</td>
<td>0.5086</td>
</tr>
<tr>
<td>S. natans</td>
<td>*0.0400</td>
<td>*0.0788</td>
<td>*0.0778</td>
<td>0.0312</td>
<td>0.2280</td>
<td>0.5434</td>
<td>0.1732</td>
</tr>
</tbody>
</table>

*<sup>*</sup>Significant relationships (p < 0.05).

**<sup>**Significant relationships (p < 0.1), indicating that the filamentous bacteria were significant to some of the operational parameters (i.e. COD, temperature, F/M, DO, NH₄-N and P).
showed that AIC and SC values of the model decreased when the independent variables were included in the model. A lower AIC and SC value indicated a better fit of the model to the selected set of data analysed, hence the model was significantly fit for the current data input (data not shown).

Table 4 indicates the significant relationships ($p < 0.1$ and $p < 0.05$) between the selected operational parameters and the dominant filamentous bacteria. A maximum likelihood estimate (MLE) and proportional OR were calculated for each of the dominant filamentous bacteria against the selected operational parameters (Tables 5 and 6). The output generated from the MLE analyses showed a positive coefficient indicating an increased chance that a subject (in this case filamentous bacteria), with a higher score on the independent variables (operational factors), will most likely be observed at a lower indexed scale (FI: $\leq 3$); (Tables 5 and 6) and vice versa. This was because the probabilities modelled were cumulated over the lower ordered values.

### Application of a CLM to predict filament abundance under different plant operational parameters

**Influent chemical level vs. filamentous bacterial index**

A great deal of work had been done previously on the relationships between filamentous bacteria and flocc forming bacteria at different nutrient concentrations and plant operational parameters (Chudoba et al. 1973; Lau et al. 1984; Van Loosdrecht et al. 2008). The modelled MLE showed that at a higher COD level (>750 mg CODL$^{-1}$), the dominant filamentous bacteria viz., Type 1851, Type 021N, Thiothrix spp. and S. natans, were at a lower ranking (FI: $\leq 3$) except Type 0092 (Tables 5 and 6). Previous research showed that in completely mixed systems when substrate concentrations were low, filamentous bacteria was found to have a higher specific growth rate and outcompete flocc-forming bacteria (Madoni et al. 2000; Martins et al. 2004), as evident in this study.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Estimated regression coefficients with standard errors and ORs (Type 1851, Type 021N, Type 0092)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
<td>Type 1851</td>
</tr>
<tr>
<td></td>
<td>Max likelihood</td>
</tr>
<tr>
<td>Influent COD</td>
<td>1.3884</td>
</tr>
<tr>
<td>Temperature</td>
<td>-</td>
</tr>
<tr>
<td>F/M</td>
<td>3.4963</td>
</tr>
<tr>
<td>DO</td>
<td>1.7498</td>
</tr>
<tr>
<td>Influent NH$_4$-N</td>
<td>1.4931</td>
</tr>
<tr>
<td>PO$_4$-P</td>
<td>1.8526</td>
</tr>
</tbody>
</table>

*OR indicates the relative differences among each of the operational parameters (i.e. COD, temperature, F/M, DO, NH$_4$-N and PO$_4$-P) at high and low levels. For instance, the OR analysis for influent COD ‘high’ versus ‘low’ (for Type 1851) showed that, with high COD levels the organism has *4.009 times the odds of being at a lower ranking, than low levels of COD.*

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Estimated regression coefficients with standard errors and ORs (Thiothrix spp., Type 0041, S. natans)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
<td>Thiothrix spp.</td>
</tr>
<tr>
<td></td>
<td>Max likelihood</td>
</tr>
<tr>
<td>Influent COD</td>
<td>1.2614</td>
</tr>
<tr>
<td>Temperature</td>
<td>-</td>
</tr>
<tr>
<td>F/M</td>
<td>1.9805</td>
</tr>
<tr>
<td>DO</td>
<td>-</td>
</tr>
<tr>
<td>Influent NH$_4$-N</td>
<td>-</td>
</tr>
<tr>
<td>PO$_4$-P</td>
<td>1.1415</td>
</tr>
</tbody>
</table>
The OR from the model corroborated that at a high COD level (≥750 mg CODL⁻¹), Type 1851 has 4.00 times the odds of being at a lower ranking (FI: ≤3), than low COD levels (<750 mg CODL⁻¹) (Table 5). The same relationships could be seen with Type 021N, *Thiothrix* spp. and *S. natans* with an OR of 2.77, 3.53 and 2.78, respectively (Tables 5 and 6). However, at a high COD level, Type 0092 showed an OR of less than 1 (OR: 0.20), which indicated that the odds of Type 0092 ranked on the higher index scale (FI: ≥4), when the COD was high as opposed to a low COD, is within a ratio of 1:5. This means that there was a greater chance of Type 0092 ranked high at higher COD level (Table 5).

In addition, for each of the dominant filamentous organisms identified, a probability graph (polybar plot) was also examined in the logit model. From the polybar plots in Figures 3(a)–3(c), it can be observed that filamentous bacteria Type 021N, Type 1851 and *Thiothrix* spp., showed an increased (>50%) chance of being in a lower ranking (FI: ≤3) when the influent COD was high (1), whereas Type 0092 showed a higher (>80%) chance of being in a lower ranking (FI: ≤3) when the influent COD was low (0) (Figure 3(d)). This could be because of their COD preference for their growth. It was reported that Type 0092 prefers slowly biodegradable COD (SBCOD) as opposed to the other filamentous organisms which favour readily biodegradable COD (RBCOD) (Jenkins et al. 2004; Martins et al. 2004). Thus Type 0092 possibly had a competitive advantage over other filamentous bacteria at a higher concentration of SBCOD. However, this needs further validation based on influent COD characterisation.

Figure 3 | Polybar plots (a)–(d) illustrate the probability of each filamentous bacterial ranking y (FI: 2–6) at low (0) and high (1) levels of influent COD. The plots (e)–(g) indicate the effect of low (0) and high (1) influent NH₄-N levels. The y-axis indicates the probability percentage at different ranking levels.
Similarly, Figures 3(e)–3(g) represent filamentous abundance at low (0) and high (1) NH₄-N levels indicating the influence of NH₄-N on Type 1851, Type 021N and Type 0092, respectively. From the MLE and OR analysis, high NH₄-N (≥32 mg L⁻¹) reduced the likelihood of Type 021N being in a lower ranking (MLE: −1.0493; OR: 0.350) (Tables 5 and 6). It was observed that there was higher chance of Type 021N ranked at a higher filament index (FI: ≥4) when the NH₄-N is high. It has been previously reported that Type 021N showed a rapid uptake of ammonia than flocc formers when the supply of nutrients was intermittent. The growth rate of Type 021N increased and the organism became dominant (Jenkins et al. 2004; Martins et al. 2004; Vaiopoulou et al. 2007). The MLE and OR analysis (Tables 5 and 6) showed that a higher DO (≥1.1 mg O₂ L⁻¹) increased the probability of Type 1851 and S. natans being ranked at lower levels (estimated coefficients: 1.7498 and 1.1529, respectively) and decreased the odds of Type 0092 (estimated coefficient: −1.9217) being at a lower level (Table 5). This can further be corroborated using the probability plot (Figure 4(b)) which showed that Type 0092 had a greater (>60%) chance of being at a higher ranking (FI: >3) when the DO was high (≥1.1 mg O₂ L⁻¹).

The F/M ratio of the selected WWTP ranged between 0.08 and 0.18 kg COD/kg MLSS d⁻¹ which is well within the range for most filamentous bacteria to proliferate (Table 3) (Jenkins et al. 2004). The CLM showed significant relationships (p < 0.1) between F/M ratio and filamentous bacteria Type 1851, Type 0092, Thiothrix spp. and S. natans. The model showed that low F/M ratios (<0.1 kg COD/kg MLSS d⁻¹) increased the probability of Type 1851, Thiothrix spp. and S. natans (estimated likelihood: 3.49, 1.98 and 1.30, respectively) of being in a higher ranking as opposed to Type 0092 (Table 5) provided that all other variables were held constant. The polybar plots predicted higher chances (>70%) of these organisms (Type

**Operational parameters vs. filamentous bacterial index**

Many studies reported that low DO and F/M ratios could be responsible for the excessive filamentous growth in AS systems (Wilen & Balmer 1999; Beer et al. 2002; Jenkins et al.

---

**Figure 4** | Polybar plots (a) and (b) illustrate the probability of each filamentous bacterial ranking y (FI: 2–6) at low (0) and high (1) levels of DO. The plots (c)–(f) indicate the effect of low (0) and high (1) F/M ratio levels. The y-axis indicates the probability percentage at different ranking levels.
1851, Type 021N and *Thiothrix* spp.) being ranked at higher levels (FI: >4) when the F/M was low (<0.1 kg COD/kg MLSS d<sup>-1</sup>) (Figures 4(c)–4(e), respectively). Previous studies have also shown that these filamentous organisms proliferated in the WWTP with low F/M and long sludge age conditions (Beer et al. 2002; Jenkins et al. 2004; Speirs et al. 2009). However, in contrast to previous research findings, in this logit model, Type 0092 showed an increased (>90%) chance of being ranked at low levels (FI: <3) when the F/M was low (Figure 4(f)). It has been reported that Type 0092 was slow growing and these organisms have gained a competitive advantage over simultaneous nitrification-denitrification (SND) conditions (Jenkins et al. 2004). In this study, since the WWTP had operated via a SND mode under limited DO concentrations, it can be postulated that during low F/M conditions the remaining filamentous organisms may have won the competition over Type 0092 for the remaining available substrates.

**Environmental factors vs. filamentous bacterial index**

The influence of environmental factors such as temperature and pH on filamentous dominance was also investigated. Temperature and pH of the selected WWTP ranged between 17.50–24.80 °C and 7.07–7.26, respectively (Table 3). Type 0092 showed a significant relationship (p < 0.1) with temperature. The results indicated that high temperatures (>20 °C) increased the probability of Type 0092 being in a lower ranking (FI: ≤3). Furthermore, Type 0092 displayed an OR of 6.60 also indicating the high odds of Type 0092 being ranked (FI: 1–3) at a lower scale when the temperature was found to be >20 °C provided that all the other variables were held constant (Table 5, Figure 5(a)). Although Type 0041 showed a significant relationship with pH (p < 0.1), the change in pH were found to be within neutral range (Table 4, Figure 5(b)).

**A probability graph model to predict specific filamentous growth based on multiple factors**

A probability plot was examined in this study which included the cumulative effect of plant operating parameters relative to each of the dominant filamentous bacteria. The plot showed a visual representation of the significant variables which could have resulted in the proliferation of filamentous bacteria, as well as possible control measures once these organisms were established in the AS system.

Filamentous bacteria viz., Type 1851, Type 021N, *Thiothrix* spp. essentially utilise readily metabolisable substrates and were found dominant in WWTPs operating at moderate to high sludge age/mean cell residence time (MCRT) (Jenkins et al. 2004). It was further reported that the growth of these organisms in AS was generally encouraged by the use of uniformly aerated, completely mixed aerated basins (Vaiopoulou et al. 2007). The introduction of anaerobic/anoxic basins before aeration apparently reduced its growth. Since Kingsburgh WWTPs was initially designed as a UCT process but without an ‘a’ recycle (Figure 1), this modification could have resulted in a longer MCRT to facilitate nitrification and hence influenced the growth of these filaments.

**Predicted cumulative probabilities for Type 1851**

Type 1851 showed significant relationships (p < 0.1) with COD, DO, NH<sub>4</sub>-N and F/M ratios (Table 4). When all the
above operational parameters were grouped together in a single probability plot, the organism illustrated a greater than 90% chance of being at a lower filament index (FI: ≤3) when all these parameters were at a high level (1) (Figure 6(a), see *). In addition, the graph showed a maximum predicted cumulative probability (>90% chance) of Type 1851 being at a low ranking (FI: ≤3), when F/M and DO were high (1) regardless of the other two operational parameters (COD and NH₄-N) being at low or high levels (Figure 6(a), see * and **). The plot also illustrated the distribution of Type 1851 when the parameters were reversed. In all cases when both F/M and DO were low, the plot showed a greater than 75% chance of the organism being at a higher scale (FI: ≥4). This occurrence was also in agreement with other studies indicating that these conditions had an effect on sludge bulking (Beer et al. 2002; Martins et al. 2003). As a result, the distribution of probabilities investigated clearly indicated the role of DO and F/M over the other operational parameters mentioned in controlling Type 1851.

Figure 6 | (a) Plot indicating a computational output with the cumulative low (0) and high (1) levels of four measured parameters, in the order of COD, F/M, DO and NH₄-N for Eikelboom Type 1851. (N.B. *, ** indicates the maximum probabilities of the organism ranked at low levels (FI: ≤3) which had occurred in all cases when DO and F/M ratio were high, regardless of the other factors being at high or low levels). (b) Plot indicating a computational output with the cumulative low (0) and high (1) levels of four measured parameters in the order of COD, F/M, DO and NH₄-N for Eikelboom Type 021N. (N.B. * indicates the maximum occurrence of Eikelboom Type 021N at a lower scale (FI: ≤3) which was found when COD was high and NH₄-N was low, ** indicated a greater than 90% chance of the organism ranked at a higher FI scale (FI: ≥4), which occurred in all cases when COD was low and NH₄-N was high). (c) Plot indicating a computational output with the cumulative low (0) and high (1) levels of four measured parameters in the order of COD, F/M, DO and NH₄-N for Thiothrix spp. (d) Plot indicating a computational output with the cumulative low (0) and high (1) levels of five measured parameters in the order of COD, DO, NH₄-N, Temperature and F/M for Eikelboom Type 0092.
Predicted cumulative probabilities for Type 021N

Type 021N showed significant relationships \((p < 0.1)\) with COD, F/M and NH\(_4\)-N. The maximum occurrence of Type 021N predicted at a lower scale (FI: ≤3) was found when COD was high and NH\(_4\)-N was low (Figure 6(b), see 6). Conversely, in all cases when influent COD was low (<750 mg CODL\(^{-1}\)) and influent NH\(_4\)-N was high (≥32 mgL\(^{-1}\)), there was a 90% chance that the organism would be at a higher FI scale (FI: 4–6); (Figure 6(b), see **). Moreover, in the four cases when COD and ammonia where high, regardless of F/M and DO being at low or high levels, the model showed a greater than 80% chance of the organism being at higher FI level (FI: ≥4). Hence, indicating a strong influence of ammonia on Type 021N. The uptake of NH\(_4\)-N and nitrite by filamentous bacteria has been previously reported (Jenkins et al. 2004; Tian et al. 2011). A study on filamentous bulking in South African WWTPs on nutrient removal AS systems showed an increase in low F/M filaments with the effect of increased nitrates and nitrites at low COD levels (Musvoto et al. 1999). Results from their study have confirmed an increase in filamentous bacteria population in a modified UCT process as a result of addition of NH\(_4\)-N into the system, which also resulted in an increased DSVI. Similar observations were derived from the logit model in this study for Type 021N.

Predicted cumulative probabilities for Thiothrix spp.

Filamentous bacteria Thiothrix spp. showed significant relationships \((p < 0.1)\) with influent COD and F/M ratio (Table 4). However, the probability plot indicated a greater than 70% chance of this organism being on a higher scale (FI: ≥4) when the F/M ratio was low (<0.1 kg COD/kg MLSS d\(^{-1}\)) and NH\(_4\)-N levels were high (≥32 mgL\(^{-1}\)) (Figure 6(c), see 6). In contrast to the observations from the plot, in all cases when F/M was high and NH\(_4\)-N levels were low (regardless of COD and DO), the organism showed a higher chance (>60%) of being on a lower scale (FI: ≤3) (Figure 6(c), see **). From the CLM model, it was found that in an increase in COD and/or DO levels, Thiothrix spp. showed an increased chance (>75%) of being ranked at lower levels (FI: 1–3). Many studies reported that Thiothrix spp. can use readily biodegradable substrates and is capable of sulphide oxidising by storing sulphur granules. Moreover, it was found that this organism rapidly take up nutrients under nutrient deficient conditions (low F/M) (Aruga et al. 2002; Rossetti et al. 2003; Jenkins et al. 2004).

The mutual interaction of Type 021N and Thiothrix spp. in this system also needs further evaluation as their dominance followed a similar pattern. Both filamentous groups were the major dominant species in the current study. Some studies reported that both Thiothrix group 1 and Type 021N can utilise hydrogen sulphide as a source of energy, which together with the readily biodegradable compounds allow for their excessive growth in ‘septic’ wastewater (Eikelboom 2000; Rossetti et al. 2003; Martins et al. 2004; Jenkins et al. 2004). Although the concentrations of sulphur compounds in the aeration tank were not measured throughout the year and were not included in the model due to the small sample size, results strongly suggest the septicity of wastewater. The longer retention time in the aeration tank could also have resulted in the formation of septic sludge thereby influencing the dominance of these two filaments throughout the study period.

Predicted cumulative probabilities for Type 0092

Type 0092 showed significant relationships \((p < 0.1)\) with the operational parameters including; COD, NH\(_4\)-N, DO, F/M and temperature (Table 4). However, as opposed to the other filamentous bacteria, the model predicted a greater than 80% chance of the organism being at lower levels (FI: 2–3) when the preceding parameters were low except for temperature (Figures 3–5). Aygun et al. (2013) found that Type 0092 increased in warmer temperatures during summer (>15 °C) and decreased in colder temperatures. However, from the proportional OR analysis, it was found that higher temperatures (>20 °C), increased the likelihood of Type 0092 being in a lower ranking and lower temperatures increased the likelihood of this organism being at a higher ranking. When all the factors were grouped together (Figure 6(d), see 6), low temperatures (<20 °C) and high DO (>1.1 mg O\(_2\)L\(^{-1}\)) predicted that Type 0092 has a greater than 90% chance of being at a higher level (FI: >4). During low temperatures, the metabolic activity of microorganisms was slow, Type 0092 could have had a competitive advantage over these slow growing micro-organisms since these morphotypes also prefer SBCOD instead of RBCOD (Jenkins et al. 2004). As a result, during these conditions, Type 0092 could have possibly out-competed heterotrophic bacteria. However, these findings need further evaluation based on pilot-scale studies with the enrichment of filamentous organisms in a continuous or sequencing batch reactor system. Moreover, a survey across numerous BNR AS processes using advanced techniques would add great value to the applied model.
CONCLUSIONS

- The dominant filamentous bacteria identified from bulking sludge samples were Type 021N, *Thiothrix* spp., Type 1851 and Type 0092.
- The CLM proved to be an efficient tool to predict the abundance of filamentous bacteria under specific plant operational parameters.
- The CLM predicted the dominance of Type 1851 (FI: ≥4) under low DO (<1 mg O₂ L⁻¹) and F/M (<0.1 kg COD/kg MLSS d⁻¹) ratios.
- A low F/M ratio and high ammonia levels predicted the occurrence and dominance of *Thiothrix* spp.
- The CLM illustrated the dominance of Type 021N during low COD (<750 mg COD L⁻¹) and high NH₄⁻N levels (>32 mg L⁻¹), however, these findings need further evaluation based on the total nitrogen level.
- High DO (≥1.1 mg O₂ L⁻¹) and low temperatures (<20 °C) were favourable conditions predicting the common occurrence of Type 0092.
- Future studies will be focussed on evaluating multiple WWTs (domestic and industrial) to understand the behaviour of filamentous bacteria under different operational parameters using molecular techniques, and the CLM model will be compared to other existing models such as ANN and SVR.

ACKNOWLEDGEMENTS

The support of the Water Research Commission (WRC), South Africa, both technically and financially for the successful completion of this project is gratefully acknowledged. The authors gratefully acknowledge King’sburgh WWTP, KwaZulu-Natal, for their continuous support during this study. A special thank you to Professor Glenda Matthews (School of Mathematics, Statistics and Computer Science, UKZN), and the anonymous reviewers for their valuable comments and suggestions.

REFERENCES


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, 1309–1316.


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), 387–393.


Kanagawa, T., Kamagata, Y., Aruga, S., Kohno, T., Horn, M. & Wagner, A. M. 2000 *Phylogenetic analysis of and*
oligonucleotide probe Eikelboom Type 021N filamentous bacteria isolated from bulking activated sludge. *Applied and Environmental Microbiology* **66** (11), 5043–5052.


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** (8), 2446–2452.


First received 9 February 2015; accepted in revised form 1 April 2015. Available online 12 May 2015.