Risk model for microbiology-related settling problems implementation

Proposal for the risk model development (Using standard modeling descriptions and equations)

Dalmau, J., Rodríguez-Roda, I. and Comas, J.

LEQUIA (University of Girona) and IEA-LTH (Lund University)

December 2005 (updated April 2007; December 2008)

Summary

This technical report describes the implementation of the risk assessment model for microbiology-related settling problems during simulation of wastewater treatment plants using standard modelling descriptions and equations. This report enables to implement the risk model described in the Technical report #12 for any software platform. The report is structured as follows: first section goes through the different steps required to estimate in general in a fuzzy logic way the risk for any settling problem. Second section illustrates a specific numerical example of the calculations detailed in section 1. The whole knowledge base, including input and output variables, fuzzy membership functions and the decision matrices used to estimate the risk of each individual microbiology-related activated sludge solids separation problem i.e. risk of filamentous bulking due to low dissolved oxygen, risk of filamentous bulking due to nutrient deficiency, risk of filamentous bulking due to low organic loading, risk of foaming due to low F/M ratio, risk of foaming due to high readily biodegradable organic matter fraction and risk of rising sludge, are fully described in section 3. Next section specifies how to integrate the six individual risks into an overall index for microbiology-related solids separation problems. Then, the response surfaces of the risk assessment model are provided and, finally, some key references are provided.

Copyright © 2008 Jordi Dalmau, Ignasi Rodríguez-Roda and Joaquim Comas

Table of Contents

1. A FUZZY IMPLEMENTATION

The collection of rules extracted from the previous flow diagrams constitutes the knowledge base of the risk model. The uncertainty associated with the description of heuristic knowledge is tackled by using fuzzy logic in the development of the heuristic IF-THEN rules. Membership functions were developed for each variable considered in the flow diagrams while a fuzzy rule base or decision matrix must be defined for each operational problem.

Figure 1 illustrates how this new knowledge-based model works to be able to obtain the risk of microbiologically related settling problems. Some data from the simulation output is fuzzified, then the fuzzy rule-based system launches those rules whose antecedents (IF part) are satisfied and finally the defuzzification takes place to provide the user with a numerical value (between 0 and 1, or 0 and 100%) for the risk of each one of the operational problems considered.

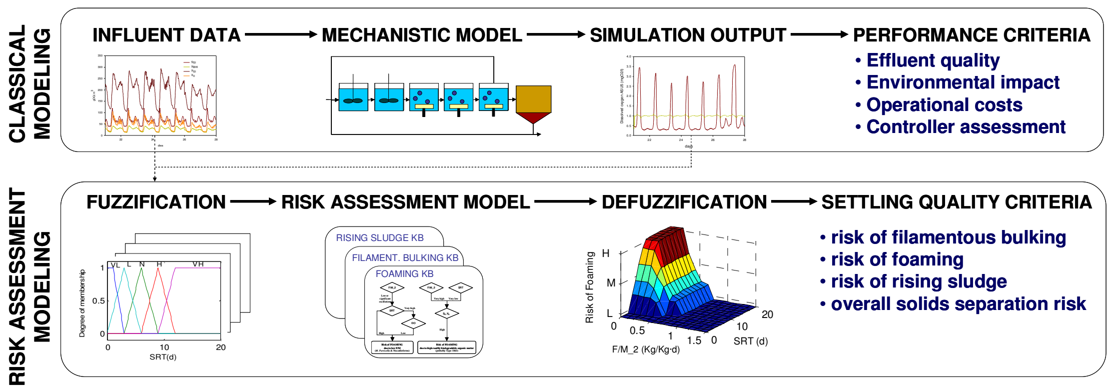

Figure 1: Scheme of the operations performed by the risk model.

Next the fuzzification, rule base definition and defuzzification steps are detailed using standard modelling descriptions and equations.

1.1. Fuzzification

For each input and output variable selected, we define two or more membership functions (MF), normally three but can be more. We have to define a qualitative category for each one of them, for example: low, normal or high. The shape of these functions can be diverse but we will usually work with triangles and trapezoids (actually usually pseudo-trapezoids) (see Figure 2). For this reason we need at least three (for triangles) or four (for trapezoids) points to define one MF of one variable.

> ***Example 1:*** If we take **x** like a variable and low, normal and high as trapeizodal, triangle and trapezoidal MFs, respectively (Figure 2),
>
> - in fact, the MF **low** will be defined by three points: (x_1, x_2, x_3) since x_1 will always be 0. However, in order to define a real trapezoid a fourth point at the left of x_1 (any negative one, e.g x_0) have to be defined. Despite this, any x value lower than x_1 will have a degree of membership to the low MF of 1.
>
> - following the same reasoning, the MF **high** have to be defined by four points: (x_3, x_4, x_5, x_6) (x_6 any positive > x_5. Despite this, any x value higher than x_5 will have a degree of membership to the high MF of 1.
>
> - finally, the MF **normal** (like any other triangular MF) will be defined by three points: (x_2, x_3, x_4).

In case the MF are trapezoids (or pseudo-trapezoids) (in this case 'low' and 'high'), the MF can be defined as:

$$y^{low}_{(trap)}(x; x_0, x_1, x_2, x_3) = \begin{cases} \max\left(\min\left(\frac{x - x_0}{x_1 - x_0}, 1, \frac{x_3 - x}{x_3 - x_2}\right), 0\right) & \text{for } x_3 \geq x \geq x_1 \\ 1 & \text{for } x < x_1 \end{cases} \quad (1)$$

$$y^{high}_{(trap)}(x; x_3, x_4, x_5, x_6) = \begin{cases} \max\left(\min\left(\frac{x - x_3}{x_4 - x_3}, 1, \frac{x_6 - x}{x_6 - x_5}\right), 0\right) & \text{for } x_5 \geq x \geq x_3 \\ 1 & \text{for } x > x_5 \end{cases} \quad (2)$$

In case the MF are triangles (in this case 'normal'), the MF can be defined as:

$$y^{normal}_{(tri)}(x; x_2, x_3, x_4) = \max\left(\min\left(\frac{x - x_2}{x_3 - x_2}, \frac{x_4 - x}{x_4 - x_3}\right), 0\right) \quad (3)$$

It is important to emphasize that the computation of all the functions/equations for all the MFs of all variables has to be done every time the shape and interval of the MFs are changed (on the contrary, once computed for the first time, these computations do not have to be done if MFs are not changed).

How the fuzzification step works

Next question to be solved is how to fuzzificate all the real values of the variable **x**. First, for a given value of **x**, for example $\mathbf{x_n}$ which can belong to one or more MF we calculate the **y** value for each of the MF/s which $\mathbf{x_n}$ belong. This **y** value has to be between 0 and 1. For example: Consider three MF: low, normal and high and a given value of $\mathbf{x_n}$, then the degrees of membership to each MF (**y** values) for $\mathbf{x_n}$ can be, for example: 0.6 for the MF low and 0.4 for the MF normal (see Figure 2). Likewise, we can fuzzificate all the values of any variable. Any of the values will belong to at least one MF with a certain degree of membership.

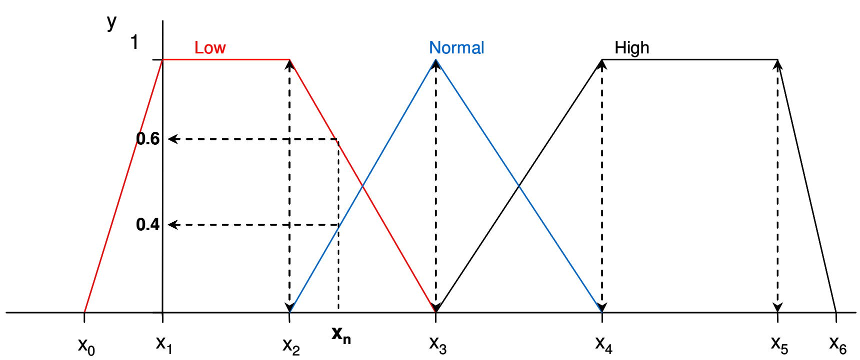

Figure 2: Example of the three MF for a given input.

1.2. Rule base (decision matrix) definition

Once the input and output variables and MF are defined, we have to design the rule-base (or decision matrix of the fuzzy knowledge-base) composed of expert *IF <antecedents> THEN <conclusions>* rules. These rules transform the input variables to an output that will tell us the risk of operational problems (this output variable, risk of a problem, also have to be defined with MF, usually low, normal and high risk). Depending on the number of MF for the input and output variables, we will be able to define more or less potential rules. The easier case is a rule-base concerning only one input and one output variable.

Example 2: For a given variable **x** involved in the development of a problem, we could have this kind of “theoretical” rule:

IF **x** is **normal** THEN risk of problem is **medium**.

The more variables we have, the more rules we have to define in order to make the inference reliable.

Example 3: Let’s suppose that we have these variables, **x** and **y**, both having three MF, low, normal and high. Each variable can belong to a different MF. Depending on the expert knowledge, we can have several inputs and several outputs.

IF **x** is **normal** and **y** is **very high** THEN risk of problems is **high**.

Once we have defined the realistic rules according to the expert knowledge, these rules will become the knowledge base of each of the problems considered in the risk model. It is necessary to say that the whole knowledge does not necessarily have to be translated in rules, sometimes some of the rules can be redundant. Let’s suppose the following *invented* decision matrix that contains the expert knowledge to detect the risk of a problem from the inputs X and Y:

Table 1: Membership functions for each variable considered in the risk assessment model.

		Input Y			
		LOW	NORMAL	HIGH	VERY HIGH
Input X	LOW	Low	High	High	High
	NORMAL	Low	Low	Medium	Medium
	HIGH	Low	Low	Low	Low
	VERY HIGH	Low	Low	Low	Low

How the rule base works

Next raising question is to compute the degree of membership to the MF (low, normal or high) of the output (the risk of the problem). As explained in the fuzzification part, once a variable is fuzzified it takes a value between 0 and 1 indicating degree of membership to a given MF of that specific variable. The degrees of membership of the input variables have to be combined to get the degree of membership of the output variable.

Example 4: For a given variable **x** involved in the development of a problem (the risk-output have its own MF, low, normal and high risk), we can for example have a rule-base “saying” that:

IF **x** is **low** THEN risk of problem is **low**.

IF **x** is **normal** THEN risk of problem is **medium**.

IF **x** is **high** THEN risk of problem is **high**.

According to these rules, if we suppose that the degree of membership for **x** is 0.6 to the MF low, then the risk of the problem will be 0.6 low, too.

In case we have more than one input variable (in fact, the normal case), the degree of membership for the output value will be the **minimum value of the degree of membership for the different inputs**.

Example 5: Looking at the Figure 3, for a set of 9 rules resulting from the decision matrix (see Table 1, above), let’s say that **Input X** = 0.55 has a membership degree of 0.8 to the MF ‘normal’ (rules 4, 6 and 7), and a membership degree of 0.2 to the MF ‘high’ (rule 8). On the other hand, **Input Y** = 6.5 has a membership degree of 0.2 to the MF ‘high’ (rules 1 and 7) and a membership degree of 0.9 the MF ‘normal’ (rules 3 and 4). When a rule is totally satisfied (the antecedent is satisfied, those with (1) in

Figure 3, rules: 4, 7 and 8), it will have an output with a certain membership degree to an output MF. These are the rules satisfied in this example:

IF **Input X** is **normal** (degree of 0.8) and **Input Y** is **normal** (degree of 0.9) THEN Risk of problem is **low** (degree of 0.8) **(Rule 4)**

IF **Input X** is **normal** (degree of 0.8) and **Input Y** is **high** (degree of 0.2) THEN Risk of problem are **medium** (degree of 0.2) **(Rule 7)**

IF **Input X** is **high** (degree of 0.2) THEN Risk of problem is **low** (degree of 0.2) **(Rule 8)**

The MF of the output will have a degree of membership equal to the lower among the inputs.

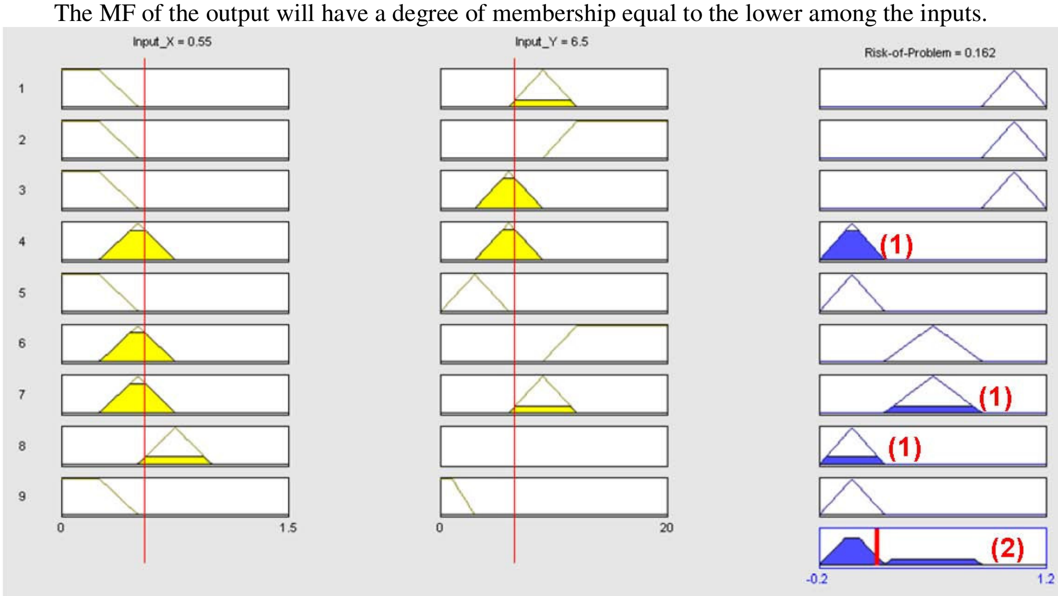

Figure 3: Example of the rules for the determination of a hypothetic risk of problem (detected by the rules of table 1).

From here, we will only look at those satisfied rules (4, 7 and 8). The resulting figure for the output (2) has the MF 'low' due to rules 4 and 8. Moreover, it has the MF 'normal' due to rule 7. To sum up, the final output figure (2) is the integration (sum) of the MF from the satisfied rules (1). Among the satisfied rules, the membership degree of each output MF will be the higher among the rules that have as a result that MF. It means that the degree of membership of the MF 'normal' (0.2) (in (2)) is due to rule 7 and that the degree of membership of the MF 'low' (0.8) (in (2)) is due to the higher among rules 4 and 8 (those that have as a conclusion the MF 'low').

1.3. Defuzzyfication

The MFs of the output have always the same shape and configuration in our risk model: the risk of any problem has the same ranks for the MFs of the output: low, normal and high, and always without overlapping. Figure 4 shows the shape of each MF of the output variable (risk on any problem considered in the risk model).

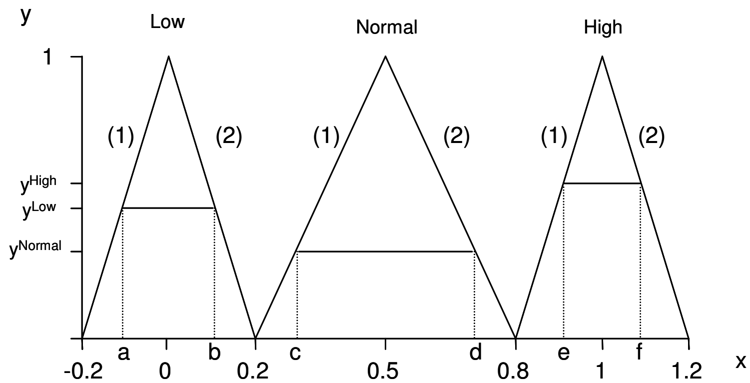

Figure 4: MFs of the risk of any problem within the risk model.

In order to obtain a percentage of risk of a problem, the output must be defuzzyfied. The equations of the straight lines of each MF of the output have to be calculated. The calculations for each of the MFs are presented next:

For the MF '*Low*':

$$y_1^{low} = m_1 \cdot x_1 + n_1$$
$$y_2^{low} = m_2 \cdot x_2 + n_2$$
$$m_1 = (1-0)/(0.2-0) = 5$$
$$m_2 = (1-0)/(0-0.2) = -5$$

Now, to find the n, the point (0, 1) is substituted on both straight lines to obtain the two equations:

$$\begin{aligned} y_1^{Low} &= 5 \cdot x + 1 \\ y_2^{low} &= -5 \cdot x + 1 \end{aligned} \quad (4)$$

A similar calculation is done for the '*Medium*' (or '*Normal*') and '*High*' MFs, obtaining finally:

$$\begin{aligned} y_1^{Medium} &= \frac{1}{0.3} \cdot x - 0.665 \\ y_2^{Medium} &= -\frac{1}{0.3} \cdot x + 2.665 \end{aligned} \quad (5) \qquad \begin{aligned} y_1^{High} &= 5 \cdot x - 4 \\ y_2^{High} &= -5 \cdot x + 6 \end{aligned} \quad (6)$$

The value of each MF between each letter (i.e., between a and b, c and d, and e and f), it corresponds to the value of the degree of membership for that function.

To calculate a, b, c, d, e and f the degree of membership has to be substituted in the y of the corresponding function whereas x will be the corresponding letter (i.e. a, b, c, d, e or f).

Example 6: From example 5 we have obtained finally two output MF, 'low' and 'medium' with a degree of membership of 0.8 and 0.2 respectively. Thus, to find points *a* and *b* 0.8 has to be substituted in:

$$\begin{cases} y_1^{Low} = 5 \cdot a + 1 \\ y_2^{low} = -5 \cdot b + 1 \end{cases}$$

Where, is the degree of membership of 0.8 like this:

$$\begin{cases} 0.8 = 5 \cdot a + 1 \\ 0.8 = -5 \cdot b + 1 \end{cases}$$

And finally,

$$\begin{cases} a = -0.04 \\ b = 0.04 \end{cases}$$

The same is true for the 'medium' MF,

$$\begin{cases} y_1^{Medium} = \frac{1}{0.3} \cdot c - 0.665 \\ y_2^{Medium} = -\frac{1}{0.3} \cdot d + 2.665 \end{cases}$$

$$\begin{cases} 0.2 = \frac{1}{0.3} \cdot c - 0.665 \\ 0.2 = -\frac{1}{0.3} \cdot d + 2.665 \end{cases}$$

$$\begin{cases} c = 0.2595 \\ d = 0.7395 \end{cases}$$

How the defuzzyfication step works

The next step involves calculating the area of the resulting figures for each MF, taking into account that these areas will not be triangle in all cases (most of the times they are triangles and trapezoids).

The basic idea is to evaluate each output activated MF in intervals of 0.014. Those output MFs that have not been activated by the rules in the step 1.2- Rules base, take 0 as value in all its range (see example 7).

It will be necessary to store each pair (x, y) for later calculation of the centroid.

So, for the low MF should be like follows.

Table 2: Intervals to be evaluated in steps of 0.014 for low output MF. Letters are referred to Figure 4.

from -0.2 *to* a	$y_1^{Low} = 5 \cdot x + 1$
from a *to* b	The value in this interval is the degree of membership of the resulting MF of the step 1.2 – Rule base
from b *to* 0.192	$y_1^{Low} = -5 \cdot x + 1$

The same is true for the remaining MF with their respective functions.

Table 3: Intervals to be evaluated in steps of 0.014 for medium output MF. Letters are referred to Figure 4.

from 0.206 *to* c	$y_1^{Medium} = \frac{1}{0.3} \cdot x - 0.665$
from c *to* d	The value in this interval is the degree of membership of the resulting MF of the step 1.2 – Rule base
from d *to* 0.794	$y_2^{Medium} = -\frac{1}{0.3} \cdot x + 2.665$

Table 4: Intervals to be evaluated in steps of 0.014 for high output MF. Letters are referred to Figure 4.

from 0.808 *to* e	$y_1^{High} = 5 \cdot x - 4$
from e *to* f	The value in this interval is the degree of membership of the resulting MF of the step 1.2 – Rule base
from f *to* 1.2	$y_2^{High} = -5 \cdot x + 6$

As said above, if any of these three MFs have not been activated by the rules, it will take 0 in all its evaluation range.

All this calculations will end up with 101 (x, y) pairs, with x going from -0.2 to 1.2 in steps of 0.014. To calculate the final output of the Risk Model, centroid has to be calculated as follows:

$$x_c = \frac{\sum_{i=1}^{i=101} x_i \cdot y_i}{\sum_{i=1}^{i=101} y_i} \qquad (7)$$

Concerning the Risk Model, after multiplying by 100, the result will indicate the percentage of risk to experience a problem related to activated sludge suspended solids separation.

Example 7: From example 6 we know a, b and c, d. If we evaluate the MF as stated above, we obtain:

Table 5: Example of centroid calculation.

i	**x**	**y**
1	-0.2 (-0.2 ➔a)	0 ($y_1^{Low} = 5 \cdot x + 1$)
…	…	…
17	0.024 (a➔b)	0.8 (degree of membership 'low')
…	…	…
28	0.178 (b➔0.192)	0.11 ($y_1^{Low} = -5 \cdot x + 1$)
…	…	…
31	0.22 (0.206➔c)	0.068 ($y_1^{Medium} = \frac{1}{0.3} \cdot x - 0.665$)
…	…	…
44	0.402 (c➔d)	0.2 (degree of membership 'medium')
…	…	…
71	0.78 (d➔0.794)	0.065 ($y_2^{Mediuml} = -\frac{1}{0.3} \cdot x + 2.665$)
…	…	…
73➔101	(0.808➔1.2)	0 (The 'high' MF is not active for this example)
	$\sum_{i=1}^{i=101} y_i$	21.47

$\sum_{i=1}^{i=101} x_i \cdot y_i$	3.87

Applying the centroid equation for the 101 (x, y) pairs:

$$x_c = \frac{\sum_{i=1}^{i=101} x_i \cdot y_i}{\sum_{i=1}^{i=101} y_i}$$

$$x_c = \frac{3.87}{21.47}$$

$$x_c = 0.18$$

1.4. Temperature effect

A literature review revealed that most of the studies regarding the effect of the temperature on the settling properties were linked to *Microthrix Parvicella*. In Eikelboom *et al.* (1998) and Spering *et al.* (2008) is stated that bulking and foaming problems due to the abundance of *M. Parvicella* follows a typical seasonal pattern with growth favoured during winter and early spring.

For example, according to Rossetti *et al.* (2005), *M. Parvicella* have different growth rates at different temperatures: Its optimum growth is at 25º C; it shows some growth around 8ºC, very poor growth towards 35º C and no growth above 35º C. Besides they report some experiences in WWTP which show bulking caused by *M. Parvicella* (due to low F/M and high SRT) at temperatures between 12-15ºC as well as bulking in Danish WWTP during winter.

In order to include the temperature effect on the Risk Assessment Model, two of the individual risks, those related to *M. Parvicella* (**Risk of Foaming due to low F/M ratio and Risk of Bulking due to low organic loading**), are multiplied by a factor provided by the following equation (see also Figure 5):

$$Factor\ Risk = 1.2 \cdot e^{\frac{-(T-5)^2}{625}} \tag{8}$$

(for the case of Risk of Bulking due to low organic loading –decision tree with two branches-, the temperature effect is included after the maximum value of the two branches has been taken, for every time step, i.e. at every time step the maximum value between the risks calculated with the two branches of this decision tree is taken and then the T factor risk is applied).

Once the T effect factor is applied, the obtained risk is limited to 1, i.e. all the values above 1.0 must be set up to 1.0.

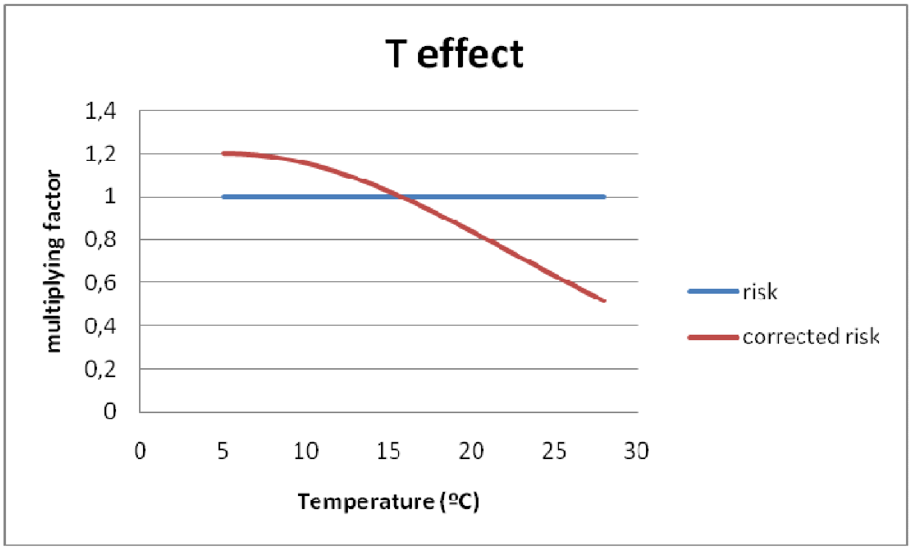

Figure 5: Variation of the factor risk with temperature (at T=15ºC -BSM1 conditions-, Factor=1.0).

Example 8: From the example 7 we obtained a risk =0.18 of, let's suppose, risk of low F/M foaming. If we supose a temperature of 20º C and apply the factor of the equation. We obtain:

$$\text{Factor Risk} = 1.2 \cdot e^{\frac{-(20-5)^2}{625}}$$

$$Factor\ Risk = 0.83$$

$$Risk\ with\ T = 0.83 * 0.18$$

$$Final\ Risk = 0.149\%$$

1.5. Exponential filter (to incorporate dynamics of bulking, foaming and rising)

The appearance of one of these microbiology-related solids separation problems in an activated sludge system requires a high risk for a long and sustained period of time. Therefore for long simulation periods the results of the risk assessment model are smoothed by means of an exponential filter that takes this issue into account. The exponential filter has a time constant related to the dynamics of each specific problem (2 hours for rising sludge and 3 days for filamentous bulking and foaming problems). This filter can be written as

$$y_{filtered}(t) = \alpha \cdot y_{filtered}(t-1) + (1-\alpha) \cdot y(t) \tag{9}$$

where $y_{filtered}$ represents the filtered data, y is the raw data and α is calculated according to

$$\alpha = 1 - \frac{1}{\tau \cdot n_s} \tag{10}$$

Where τ represents the time constant in days and ns is the number of output samples per day in the simulation (here n_s= 96). Moreover, the exponential filter facilitates the visualisation and interpretation of the results.

This exponential filter should be applied for long simulation runs i.e. for simulation runs longer than 1 month. Therefore, for the case of BSM1_LT and BSM2, this filter has been applied to each one of the six individual risks: Bulking due to low DO values, Bulking due to low organic loading, Bulking due to nutrient deficiency, Foaming due to low F/M ratio, Foaming due to high RBOM fraction and Rising sludge. For BSM1 it is not used.

Finally, this type of exponential filter has also been used with a time constant of 3 days to calculate the SRT in the system (in any benchmark model i.e. BSM1, BSM1_LT and BSM2).

1.6. Results of the Risk Model

Results from the risk model allow the degree of truthfulness to be obtained on the risk of microbiologically-related activated sludge solids separation problems (bulking, foaming and/or rising sludge) during the evaluation period (7th to the 14th day in the BSM1, from 265 to 609 in BSM1_LT and BSM2). (The aim of these model is to quantify whether the evaluated control strategies could lead the process to the favourable conditions for them to arise or not).

The results from the risk assessment are reported and quantified in four different ways for each of the six settling problems and causes (filtered for BSM1_LT and BSM2), for the integrated risks of bulking and foaming and for the overall risk index:

i) a time series plot (or average data) showing the evolution of the risk occurrence for a specific settling problem (or for one of the integrated indices) during the evaluation period. In this plot 0 means no risk while 1 indicates the highest possible risk;

Figure 6 presents this type of results obtained when applying the risk model to test and control two different control strategies;

ii) the percentage of time during which the plant is experiencing risk for occurrence of a specific settling problem;

iii) the percentage of time during which the plant is experiencing severe risk of settling problems (an arbitrary but customizable limit value of risk ≥ 0.8 is used for defining a severe problem); and,

iv) the most dangerous situation during the evaluation period, computed as the largest time interval the plant is exposed to an uninterrupted severe risk of experiencing a specific settling problem.

Note that whenever the output results of the risk model indicate that conditions for each specific problem (filamentous bulking, foaming and rising sludge) are satisfied for more than one cause branch), the worst conditions (highest risk) will be selected for this problem.

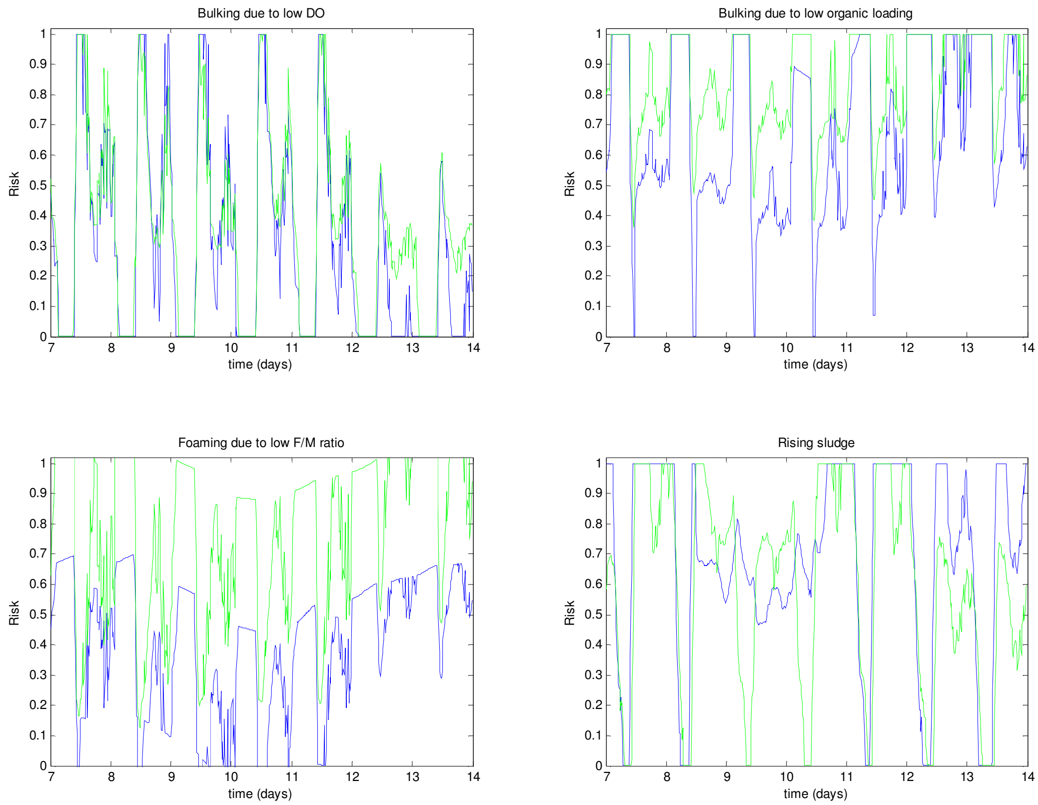

Figure 6: Risk of low DO bulking, low F/M bulking, low F/M foaming and rising under rainy influent conditions (blue: BSM1 default control strategy; green: same as the default one but with RAS flow rate = 2·Inflow rate).

An example of all the computations for the fuzzy implementation (membership functions, fuzzification, rule base, defuzzification) is presented from pages 12 to 21. For further details about the risk model, you can also refer to Comas *et al.* (2008).

2. NUMERIC EXAMPLE

This document presents a complete numerical example for the document "Proposal for the risk model (using standard modelling and equations)". This example will be based on one of the settling problems evaluated by the risk model, i.e. the filamentous bulking caused by low DO concentrations. Three variables are involved in this case: Food to microorgansims ratio (FtoM), dissolved oxygen (DO) and the risk of bulking.

2.1. Fuzzyfication

First of all, the information involving the membership functions (MF) is presented. Figures 7 and 8 illustrate the graphical representation of the MFs.

FtoM MF

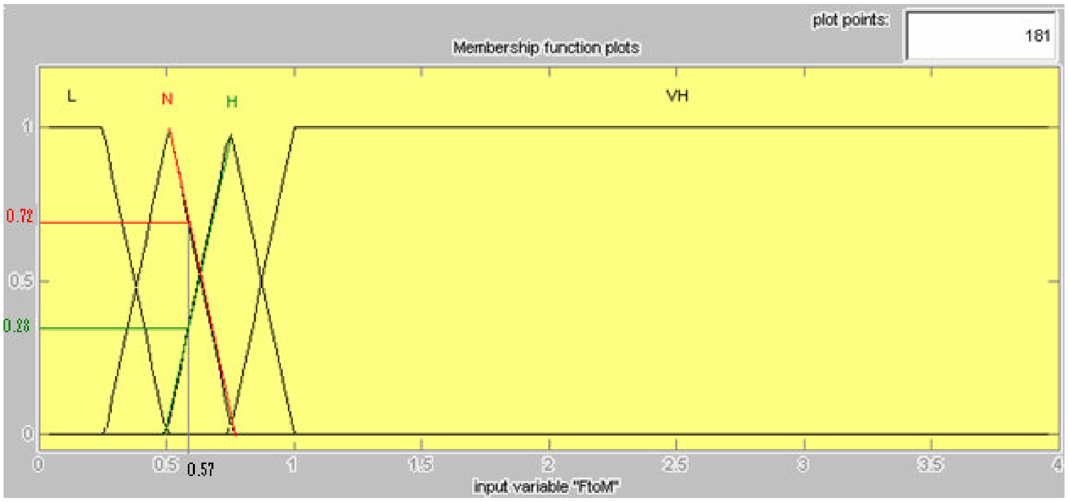

Figure 7: MFs for the input "food to microorganisms ratio" for the risk of Bulking due to low DO. L=Low, N= Normal, H= High and VH=Very High. In green: membership degree to the MF 'High'; In red: membership degree to the MF 'Normal'.

These MFs (Figure 7) are converted to equations as follows:

$$MF_{Low}^{F\,to\,M}\left(FtoM,-0.25,0,0.25,0.5\right)=\max\left(\min\left(\frac{FtoM-(-0.25)}{0-(-0.25)},1,\frac{0.5-FtoM}{0.5-0.25}\right),0\right)$$

$$MF_{Normal}^{F\,to\,M}\left(FtoM,0.25,0.5,0.75\right)=\max\left(\min\left(\frac{FtoM-0.25}{0.5-0.25},\frac{0.75-FtoM}{0.75-0.5}\right),0\right)$$

$$MF_{High}^{F\,to\,M}\left(FtoM,0.5,0.75,1\right)=\max\left(\min\left(\frac{FtoM-0.5}{0.75-0.5},\frac{1-FtoM}{1-0.75}\right),0\right)$$

$$MF_{VeryHigh}^{F\,to\,M}\left(FtoM,0.75,1,4,4.5\right)=\max\left(\min\left(\frac{FtoM-0.75}{1-0.75},1,\frac{4.5-FtoM}{4.5-4}\right),0\right)$$

DO MF

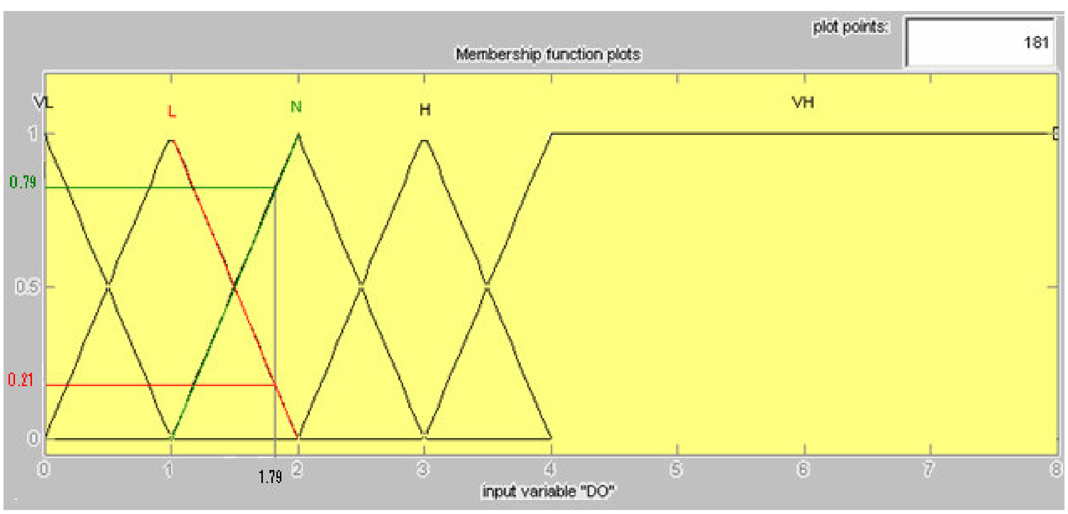

Figure 8: MFs for the input "DO". VL= Very Low, L=Low, N= Normal, H= High and VH=Very High. In green: membership to the MF 'Normal'; In red: membership to the MF 'Low'.

These are the equations of the MF for input 'DO', presented in Figure 8:

$$MF_{Very\,Low}^{DO}(DO,-0.5,-0.25,0,1) = \max\left(\min\left(\frac{DO-(-0.5)}{(-0.25)-(-0.5)},1,\frac{1-DO}{1-0}\right),0\right)$$

$$MF_{Low}^{DO}(DO,0,1,2) = \max\left(\min\left(\frac{DO-0}{1-0},\frac{2-DO}{2-1}\right),0\right)$$

$$MF_{Normal}^{DO}(DO,1,2,3) = \max\left(\min\left(\frac{DO-1}{2-1},\frac{3-DO}{3-2}\right),0\right)$$

$$MF_{High}^{DO}(DO,2,3,4) = \max\left(\min\left(\frac{DO-2}{3-2},\frac{4-DO}{4-3}\right),0\right)$$

$$MF_{Very\,High}^{DO}(DO,3,4,8,8.5) = \max\left(\min\left(\frac{DO-3}{4-3},1,\frac{8.5-DO}{8.5-8}\right),0\right)$$

Step 1:

Now, as an example, the fuzzification will be explained with real values for the two input variables. Imagine that for a given time instant, we have the following values for the two input variables ('FtoM' and 'DO'):

FtoM = 0.57 g COD removed·(g MLVSS)$^{-1}$·d^{-1}

DO = 1.79 g O_2·m^{-3}

The first step is to determine the degree of membership to each MF for each input variable.

Taking the equations for the FtoM input we can calculate,

for the 'Low' MF:

$$MF_{Low}^{F\,to\,M}(FtoM,-0.25,0,0.25,0.5) = \max\left(\min\left(\frac{FtoM-(-0.25)}{0-(-0.25)},1,\frac{0.5-FtoM}{0.5-0.25}\right),0\right)$$
$$for\ FtoM = 0.57$$
$$MF_{Low}^{F\,to\,M}(0.57,-0.25,0,0.25,0.5) = y_{Low}^{F\,to\,M} = \deg ree\,of\ membership = 0$$

for the 'Normal' MF:

$$MF_{Normal}^{F\,to\,M}(FtoM,0.25,0.5,0.75) = \max\left(\min\left(\frac{FtoM-0.25}{0.5-0.25},\frac{0.75-FtoM}{0.75-0.5}\right),0\right)$$
$$forFtoM = 0.57$$
$$MF_{Normal}^{F\,to\,M}(0.57,0.25,0.5,0.75) = y_{Normal}^{F\,to\,M} = \deg ree\ of\ membership = 0.72$$

for the 'High' MF:

$$MF_{High}^{F\,to\,M}\left(FtoM,0.5,0.75,1\right)=\max\left(\min\left(\frac{FtoM-0.5}{0.75-0.5},\frac{1-FtoM}{1-0.75}\right),0\right)$$
$$for\ x=0.57$$
$$MF_{High}^{F\,to\,M}\left(0.57,0.5,0.75,1\right)=y_{High}^{F\,to\,M}=\deg ree\,of\ membership=0.28$$

for the ‘Very High’ MF:

$$MF_{VeryHigh}^{F\,to\,M}\left(FtoM,0.75,1,4,4.5\right)=\max\left(\min\left(\frac{FtoM-0.75}{1-0.75},1,\frac{4.5-FtoM}{4.5-4}\right),0\right)$$
$$for\ x=0.57$$
$$MF_{VeryHigh}^{F\,to\,M}\left(FtoM,0.75,1,4,4.5\right)=y_{VeryHigh}^{F\,to\,M}=\deg ree\,of\ membership=0$$

Taking the equations for the DO input we can calculate, for the ‘Very low’ MF:

$$MF_{Very\,Low}^{DO}\left(DO,-0.5,-0.25,0,1\right)=\max\left(\min\left(\frac{DO-(-0.5)}{(-0.25)-(-0.5)},1,\frac{1-DO}{1-0}\right),0\right)$$
$$for\ x=1.79$$
$$MF_{Very\,Low}^{DO}\left(1.79,-0.5,-0.25,0,1\right)=y_{Verylow}^{DO}=\deg ree\,of\ membership=0$$

for the ‘Low’ MF:

$$MF_{Low}^{DO}\left(DO,0,1,2\right)=\max\left(\min\left(\frac{DO-0}{1-0},\frac{2-DO}{2-1}\right),0\right)$$
$$for\ x=1.79$$
$$MF_{Low}^{DO}\left(1.79,0,1,2\right)=y_{low}^{DO}=\deg ree\,of\ membership=0.21$$

for the ‘Normal’ MF:

$$MF_{Normal}^{DO}\left(DO,1,2,3\right)=\max\left(\min\left(\frac{DO-1}{2-1},\frac{3-DO}{3-2}\right),0\right)$$
$$for\ x=1.79$$
$$MF_{Normal}^{DO}\left(1.79,1,2,3\right)=y_{Nrormal}^{DO}=\deg ree\,of\ membership=0.79$$

for the ‘High’ MF:

$$MF_{High}^{DO}\left(DO,2,3,4\right)=\max\left(\min\left(\frac{DO-2}{3-2},\frac{4-DO}{4-3}\right),0\right)$$
$$for\ x=1.79$$
$$MF_{High}^{DO}\left(1.79,2,3,4\right)=y_{High}^{DO}=\deg ree\,of\ membership=0$$

for the ‘Very high’ MF:

$$MF_{Very\,High}^{DO}(DO,3,4,8,8.5) = \max\left(\min\left(\frac{DO-3}{4-3},1,\frac{8.5-DO}{8.5-8}\right),0\right)$$

$$for\ x = 1.79$$

$$MF_{Very\,High}^{DO}(1.79,3,4,8,8.5) = y_{Veryhigh}^{DO} = \deg ree\,of\ membership = 0$$

To sum up, we can see that for the FtoM value we have a certain degree of membership to the ‘normal’ MF (0.72) and to the ‘high’ MF (0.28). For the DO we can see that we have a certain degree of membership to the ‘low’ MF (0.21) and to the ‘normal’ MF (0.79), i.e.

$$F\ to\ M\begin{cases} Degree\,of\ membership\,of\ 0.72\,to\,the\,'normal'\,MF \\ Degree\,of\ membership\,of\ 0.28\,to\,the\,'high'\,MF \end{cases}$$

$$DO\quad\begin{cases} Degree\,of\ membership\,of\ 0.21\ to\,the\,'low'\,MF \\ Degree\,of\ membership\,of\ 0.79\ to\,the\,'normal'\,MF \end{cases}$$

Figures 7 and 8 illustrate graphically the membership degrees for FtoM and DO, respectively.

2.2. Rule base (decision matrix) definition

From the decision matrix (see Table 6, below) of the risk model knowledge base (pages 22 to 29), the rules are obtained and represented as follows:

Table 6: Representation of a decision matrix.

		F/M_removed (kg COD removed·kg MLVSS^{-1}·d^{-1})			
		L	**N**	**H**	**VH**
DO (mg/L)	**VL**	Low	High	High	High
	L	Low	Medium	High	High
	N	Low	Low	Medium	High
	H	Low	Low	Low	Medium
	VH	Low	Low	Low	Low

For example: From the decision matrix of Table 6 we can obtain the following rules:

IF *F/M_removed* is **Low** and *DO* is **Normal**, THEN *Risk of Low DO bulking* is **Low**.

Step 2:

We have already seen that the input variables have the following memberships:

F/M_removed belongs to both ‘normal’ and ’high’ MF (with different degrees to each one).

DO belongs to both ‘low’ and ‘normal’ MF (also with different degrees to each one).

Now, according to the decision matrix, it is necessary to know which rules will be satisfied. With F/M_removed normal and high (in different degrees) and DO low and normal (in different degrees), the rules satisfied are (for more details see Figure 9 below):

1. *F/M_removed* is **Normal** and *DO* is **Low**, THEN *Risk of Low DO bulking* is **Medium (Rule 12)**
2. IF *F/M_removed* is **Normal** and *DO* is **Normal**, THEN *Risk of Low DO bulking* is **Low (Rule 13)**
3. IF *F/M_removed* is **High** and *DO* is **Low**, THEN *Risk of Low DO bulking* is **High (Rule 7)**
4. IF *F/M_removed* is **High** and *DO* is **Normal**, THEN *Risk of Low DO bulking* is **Medium (Rule 8)**

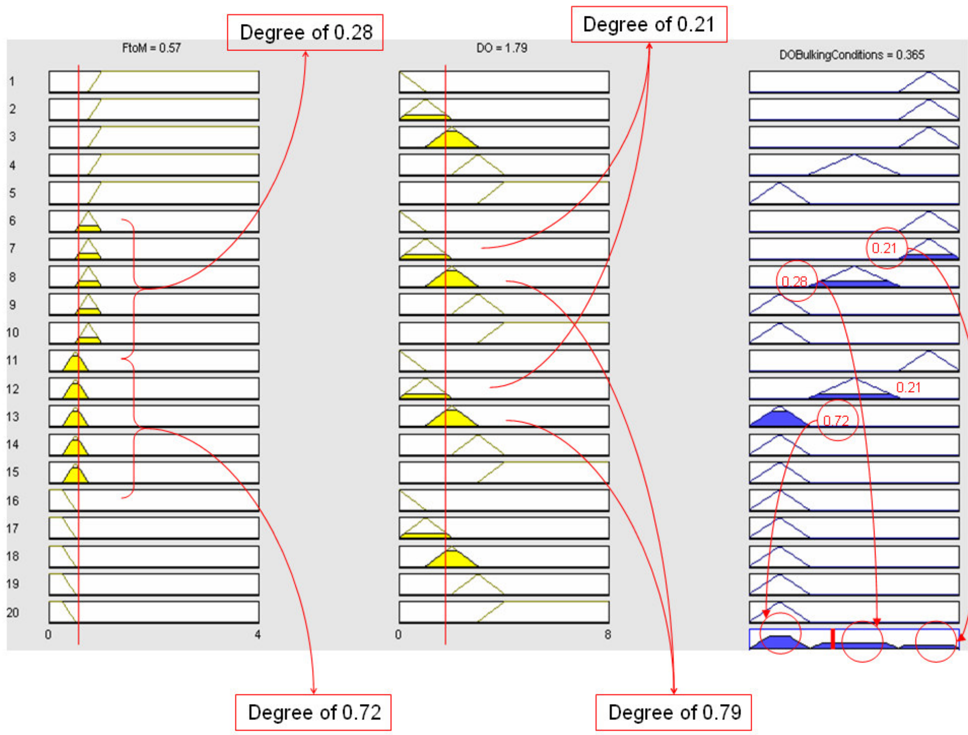

Figure 9: Example of the satisfied rules and chosen degrees of membership for fuzzify and defuzzyfy.

2.3. Defuzzyfication

As said above (see **1.3 Defuzzyfication**), for the Risk model the shape and ranks of the MFs for the output are always the same. Therefore, the MFs for the risk of low DO bulking (as well as for the other risks of problems) can be defined with the following functions:

$$MF_{\text{Low}}^{\text{Risk of problem}}\begin{cases}\text{for}-0.2\leq x<a;\ y=5x+1\\ \text{for}\,a<x<b;\ y=\text{degree of membership}\\ \text{for}\,b<x\leq 0.194;\ y=-5x_2+1\end{cases}$$

$$MF_{\text{Medium}}^{\text{Riskofproblem}}\begin{cases}\text{for}\,0.206<x<c;\ y=\dfrac{1}{0.3}x_1-0.665\\ \text{for}\,c<x<d;\ y=\text{degree of membership}\\ \text{for}\,d<x<0.794;\ y=-\dfrac{1}{0.3}x_2+2.665\end{cases}$$

$$MF_{\text{High}}^{\text{Risk of problem}}\begin{cases}\text{for}\,0.808<x<e;\ y=5x_1-4\\ \text{for}\ e<x<f;\ y=\text{degree of membership}\\ \text{for}\,f<x<1.2;\ y=-5x_2+6\end{cases}$$

These functions will allow calculating the risk of filamentous Bulking due to low DO. We have to look rule by rule, each one of the four satisfied rules (step 2), i.e. to determine the degree of membership of the output for the first satisfied rule:

IF *F/M_removed* is **Normal** and *DO* is **Low**, THEN *Risk of* low *DO bulking* is **Medium**

We have to take the lower degree of membership among the degrees of membership of the inputs, i.e. since the F/M_removed has a degree of membership of 0.72 to the MF ‘Normal’ (remember **step 1**, above) and the DO has a degree of membership of 0.21 to the MF ‘Low’, then, for this first rule, the Risk of low DO bulking will have a degree of membership equal to 0.21 for the Medium MF.

The same must be done for the other 3 rules satisfied of step 2 and the results can be summarized as:

Rule 12: Risk of bulking: MF ‘Medium’ (0.21)

Rule 13: Risk of bulking: MF ‘Low’ (0.72)

Rule 7: Risk of bulking: MF ‘High’ (0.21)

Rule 8: Risk of bulking: MF ‘Medium’ (0.28)

For each Risk of bulking affecting to the same MF of the output (i.e., Rules 12 and 8 affecting to the MF ‘Medium’), we will choose the one with the highest degree of membership. So, we will take:

Rule 13: Risk of bulking: MF ‘Low’ (0.72)

Rule 7: Risk of bulking: MF ‘High’ (0.21)

Rule 8: Risk of bulking: MF ‘Medium’ (0.28)

Now we can start calculating the areas of the figures formed in Figure 9 and calculate the centroid of the “integrated” figure:

Step 3

Now we have to calculate a, b, c, d, e and f points.

Therefore, for the ‘low’ output membership function:

$$\begin{cases} y_1^{Low} = 5 \cdot a + 1 \\ y_2^{low} = -5 \cdot b + 1 \end{cases}$$

Where, y^{low} is the degree of membership of 0.8 like this:

$$\begin{cases} 0.72 = 5 \cdot a + 1 \\ 0.72 = -5 \cdot b + 1 \end{cases}$$

And finally,

$$\begin{cases} a = -0.056 \\ b = 0.056 \end{cases}$$

The same is true for the ‘medium’ MF,

$$\begin{cases} y_1^{Medium} = \dfrac{1}{0.3} \cdot c - 0.665 \\ y_2^{Medium} = -\dfrac{1}{0.3} \cdot d + 2.665 \end{cases}$$

$$\begin{cases} 0.28 = \dfrac{1}{0.3} \cdot c - 0.665 \\ 0.28 = -\dfrac{1}{0.3} \cdot d + 2.665 \end{cases}$$

$$\begin{cases} c = 0.2835 \\ d = 0.7155 \end{cases}$$

Finally for ‘high’ MF,

$$\begin{cases} y_1^{High} = 5 \cdot e - 4 \\ y_2^{High} = -5 \cdot f + 6 \end{cases}$$

$$\begin{cases} 0.21 = 5 \cdot \mathrm{e} - 4 \\ 0.21 = -5 \cdot \mathrm{f} + 6 \end{cases}$$

$$\begin{cases} \mathrm{e} = 0.842 \\ \mathrm{f} = 1.158 \end{cases}$$

i	x	y
1	-0.2	0
2	-0.186	0.07
3	-0.172	0.14
4	-0.158	0.21
5	-0.144	0.28
6	-0.13	0.35
7	-0.116	0.42
8	-0.102	0.49
9	-0.088	0.56
10	-0.074	0.63
11	-0.06	0.7
12	-0.046	0.72
13	-0.032	0.72
14	-0.018	0.72
15	-0.004	0.72
16	0.01	0.72
17	0.024	0.72
18	0.038	0.72
19	0.052	0.72
20	0.066	0.67
21	0.08	0.6
22	0.094	0.53
23	0.108	0.46
24	0.122	0.39
25	0.136	0.32
26	0.15	0.25

27	0.164	0.18
28	0.178	0.11
29	0.192	0.04
30	0.206	0.02166667
31	0.22	0.06833333
32	0.234	0.115
33	0.248	0.16166667
34	0.262	0.20833333
35	0.276	0.255
36	0.29	0.28
37	0.304	0.28
38	0.318	0.28
39	0.332	0.28
40	0.346	0.28
41	0.36	0.28
42	0.374	0.28
43	0.388	0.28
44	0.402	0.28
45	0.416	0.28
46	0.43	0.28
47	0.444	0.28
48	0.458	0.28
49	0.472	0.28
50	0.486	0.28
51	0.5	0.28
52	0.514	0.28
53	0.528	0.28
54	0.542	0.28
55	0.556	0.28
56	0.57	0.28
57	0.584	0.28
58	0.598	0.28
59	0.612	0.28
60	0.626	0.28
61	0.64	0.28
62	0.654	0.28
63	0.668	0.28
64	0.682	0.28
65	0.696	0.28
66	0.71	0.28
67	0.724	0.25166667
68	0.738	0.205
69	0.752	0.15833333
70	0.766	0.11166667
71	0.78	0.065
72	0.794	0.01833333
73	0.808	0.04
74	0.822	0.11
75	0.836	0.18
76	0.85	0.21
77	0.864	0.21
78	0.878	0.21
79	0.892	0.21
80	0.906	0.21

81	0.92	0.21
82	0.934	0.21
83	0.948	0.21
84	0.962	0.21
85	0.976	0.21
86	0.99	0.21
87	1.004	0.21
88	1.018	0.21
89	1.032	0.21
90	1.046	0.21
91	1.06	0.21
92	1.074	0.21
93	1.088	0.21
94	1.102	0.21
95	1.116	0.21
96	1.13	0.21
97	1.144	0.21
98	1.158	0.21
99	1.172	0.14
100	1.186	0.07
101	1.2	0

For centroid calculation,

$$x_c = \frac{\sum_{i=1}^{i=101} x_i \cdot y_i}{\sum_{i=1}^{i=101} y_i}$$

$$x_c = \frac{10.52614}{28.85}$$

$$x_c = 0.365$$

As the example's separation problem does not take any temperature effect the risk will be the centroid.

Risk of Low DO bulking = 0.365

3. KNOWLEDGE BASE

3.1. Fuzzy system inputs (obtained or calculated from simulation outputs)

This section describes how the different input variables for each of the individual fuzzy risks are obtained or calculated from the simulation outputs.

<u>Bulking due to low DO values:</u>

- *Dissolved oxygen (DO)*: The dissolved oxygen (S_O) concentration is obtained from reactor 3 (1st aerobic, $S_{O,as,3}$).

In Matlab for BSM1: reac3part(:,8)

The Food-to-microorganism (F/M) ratio is calculated in two different ways within this risk model even though the membership functions are the same. While F/M_removed (or FtoM_1vec) is calculated based on the daily mass flow rate of COD removed from the whole plant per unit of biomass, F/M_fed (or FtoM_2vec) aims at detecting low organic loading (daily mass flow rate of supplied BOD per unit of biomass).

- *F/M_removed or process loading factor (FtoM_1vec)*(Grady *et al.*, 1999)

$$F / M\_removed = \frac{(CODi - CODe)}{Biomass} \tag{11}$$

Where,

$COD_i = (S_S + S_I + X_S + X_{B,H} + X_{B,A} + X_P + X_I) \cdot Q_i$

(if external carbon source from an additional Qcarb stream, COD from this stream should be included, i.e. COD_{EC})

$COD_e = (S_S + S_I + X_S + X_{B,H} + X_{B,A} + X_P + X_I) \cdot Q_e$

$$\text{Biomass} = 0.75 \cdot (\sum_{n=1}^{n=5} (X_{B,H,as,n} + X_{B,A,as,n}) \cdot V_{as,n}) \tag{12}$$

In Matlab for BSM1:

FtoM_1vec = ((CODinvec+ CODinvec2) - CODevec)./Biomassvec;
*CODinvec=inpart(:,15).*CODin;*
CODin=inpart(:,1)+inpart(:,2)+inpart(:,3)+inpart(:,4)+inpart(:,5)+inpart(:,6)+inpart(:,7); % COD load from influent, different state variables summing COD in the influent

CODin2= CARBONSOURCECONC;
Qcarbonvec = (carbon1vec+carbon2vec+carbon3vec+carbon4vec+carbon5vec); % external carbon flow rate entered to every reactor (for external C source),m3/d
*CODinvec2 = CODin2.*Qcarbonvec; % COD load from external C source,*

CODe=settlerpart(:,17)+settlerpart(:,18)+settlerpart(:,19)+settlerpart(:,20)+settlerpart(:,21)+settlerpart(:,22)+settlerpart(:,23); % different state variables summing COD in the upper layer of the settler
*CODevec=settlerpart(:,31).*CODe; % settlerpart(:,31) effluent flow rate*
(inpart is influent stream; 1, 2… are the state variables ordered according to Copp et al 2002)

Biomassvec=0.75(reac1part(:,5)*VOL1+reac1part(:,6)*VOL1+reac2part(:,5)*VOL2+reac2part(:,6)*VOL2+reac3part(:,5)*VOL3+reac3part(:,6)*VOL3+reac4part(:,5)*VOL4+reac4part(:,6)*VOL4+reac5part(:,5)*VOL5+reac5part(:,6)*VOL5); % 5 and 6 are Xbh and Xba, in TSS units*

Bulking due to low organic loading (low F/M):

- *Readily biodegradable organic matter (S_S):* The readily biodegradable organic matter (S_S) concentration is evaluated in reactor 1 ($S_{S,1}$).

  (if external carbon source from an additional Qcarb stream, S_S from this stream should be included, i.e. $S_{S,EC}$)

In Matlab for BSM1: reac1part(:,2) % S_S in reactor 1 (1st anoxic)

- *SRT:* The sludge residence time (SRT) is calculated as the total mass of TSS within the five reactors divided by the daily mass of TSS removed from the plant via the waste sludge and the effluent (Grady *et al.*, 1999).

$$SRT = \frac{\sum_{i=1}^{5}(TSS_{as,i} \cdot V_{as,i})}{TSS_w + TSS_e} \qquad (13)$$

Where,

$TSS_{as,i}$ is the concentration in each tank,

$V_{as,i}$ is the volume of each tank,

$TSS_e = (0.75 \cdot (X_{S,e} + X_{B,H,e} + X_{B,A,e} + X_{P,e} + X_{I,e}))\cdot Q_e$

$TSS_w = (0.75 \cdot (X_{S,w} + X_{B,H,w} + X_{B,A,w} + X_{P,w} + X_{I,w}))\cdot Q_w$

SRT should be filtered:

alpha=1-1/(3*(1440/samplingtime)); %sampletime=15

SRTvec(t)=alpha*SRTvec(t-1)+(1-alpha)*SRTvec(t); % SRTvec filtered using 3 days as time constant

In Matlab for BSM1:

```
SRTvec =TSSvecreactor./(TSSuvec2+TSSevec2); % TSS in reactor / TSS removed from the system
TSSvecreactor=reac1part(:,14)*VOL1+reac2part(:,14)*VOL2+reac3part(:,14)*VOL3+reac4part(:,14)*VOL4
+reac5part(:,14)*VOL5; %TSS in aeration tanks
TSSuvec2 =TSSwasteconc.*Qwasteflow.*1000; % TSS in WAS
TSSwasteconc=settlerpart(:,41)/1000;  %kg/m3, settlerpart(:,41) is the TSS conc. in the wasting sludge flow
stream; it should be settlerpart(:,53) in BSM1_LT or BSM2 !
Qwasteflow=settlerpart(:,16);        %m3/d; it should be settlerpart(:,22) in BSM1_LT or BSM2 !
TSSevec2=settlerpart(:,30).*settlerpart(:,31); %TSS in the effluent, settlerpart(:,30) is the TSS conc. in the
effluent (it should be settlerpart(:,36) in BSM1_LT or BSM2 !); settlerpart(:,31) is the effluent flow rate (it
should be settlerpart(:,37) in BSM1_LT or BSM2 !).
```

- *F/M_fed (FtoM_2vec)* (WEF, 2002):

$$F / M\_fed = \frac{BOD_{5,i}vec}{Biomass} \qquad (14)$$

Where,

$BOD_{5,i}vec = BOD_{5,i} \cdot Q_i$

(if external carbon source from an additional Qcarb stream, BOD_5 (only S_S!) from this stream should be included, i.e. $S_{S,EC}$)

In Matlab for BSM1:

```
FtoM_2vec = (BOD5invec.+BOD5invec.2) /Biomassvec;

BOD5in=0.65*(inpart(:,2)+inpart(:,4)+(1-f_P)*(inpart(:,5)+inpart(:,6)));
BOD5invec=inpart(:,15).*BOD5in; % BOD load in the influent, kg/d units. inpart(:,15) is the influent flow rate.

BOD5in2=CARBONSOURCECONC;
BOD5invec2= BOD5in2.*Qcarbonvec; % BOD load in the external C source stream, kg/d units.
```

Bulking due to N deficiency:

- *BOD_5toN:* The BOD_5/N ratios is evaluated for the influent wastewater ($BOD_{5,in}/N_{in}$).

$$BOD_5\ to\ N = \frac{BOD_{5,i}}{N_{tot,i}} \tag{15}$$

where,

$BOD_{5,i} = 0.65 \cdot (S_S + X_S + (1\text{-}f_p) \cdot (X_{B,H}+X_{B,A}))$;

(if external carbon source from an additional Qcarb stream, BOD_5 (only S_S!) from this stream should be included, i.e. $S_{S,EC}$)

$N_{tot,i} = S_{NH,i} + S_{ND,i} + X_{ND,i} + i_{XB} \cdot (X_{B,H,i}+X_{B,A,i}) + i_{XP} \cdot (X_{P,i}+X_{I,i})$

In Matlab for BSM1:

BOD5toN = (BOD5in.+ BOD5in2.)/TNin;

TNin = SNKjin+SNOin;
SNKjin=inpart(:,10)+inpart(:,11)+inpart(:,12)+i_XB(inpart(:,5)+inpart(:,6))+i_XP*(inpart(:,3)+inpart(:,7));*
SNOin=inpart(:,9);

*BOD5in2=BOD5in2.*Qcarbonvec*

Foaming due to low F/M ratio:

- *F/M_fed (FtoM_2vec)*
- *SRTvec*

Foaming due to high RBOM fraction:

- *F/M_fed (FtoM_2vec)*
- *S_S/X_S:* The readily biodegradable organic matter (S_S) to slowly biodegradable organic matter (X_S) ratio is evaluated for the influent wastewater ($S_{S,in}/X_{S,in}$).

  (if additional Qcarb stream, S_S from this stream should be included, i.e. $S_{S,EC}$)

In Matlab for BSM1:

(inpart(:,2)+ CARBONSOURCECONC)./inpart(:,4) % Ss/Xs in the influent

- *SRTvec*

Rising sludge:

- *Nitrate concentration (S_{NO})*: The nitrate concentration is obtained from reactor 5 ($S_{NO,as,5}$).

In Matlab: SNOOutReac5(:) % NOx in reactor 5

- *Nitrogen gas production time (NGPT)* (Henze *et al.*, 1993): time for the production of nitrogen gas bubbles:

$$NGPT = \frac{\mathrm{S_{NO}HighLimit}}{\mathrm{Rdn}} + t_{delay} \quad (16)$$

Where,

S_{NO}HighLimit is equal to 8 g N·m^{-3} at 15ºC (in BSM1) and a function of the temperature in BSM1_LT and BSM2, i.e. S_{NO}HighLimit =11.003972*EXP(-0.020295*T);

$$Rdn = \left(\frac{1-Y_H}{2.86 \cdot Y_H}\right) \cdot \left(\mu_H \cdot \frac{S_S Out\,\mathrm{Re}\,ac5}{K_S + S_S Out\,\mathrm{Re}\,ac5}\right) \cdot \left(\frac{S_{NO} Out\,\mathrm{Re}\,ac5}{K_{NO} + S_{NO} Out\,\mathrm{Re}\,ac5}\right) \cdot X_{BH} OutBottomClarifier \cdot \eta_g \quad \text{(denitrification rate, 17)}$$

Where $\mu_H = 4 \cdot e^{((\log(4/3)/5) \cdot (T-15))}$

$$t_{delay} = \frac{S_{O,as,5}}{\dfrac{2.86 \cdot Rdn}{nitrifiers\ fraction}} \quad (18)$$

nitrifiers fraction = 1

In Matlab for BSM1:

NGPT=((SNOHighLimit./Rdn(:))+t_delay(:))
S_{NO}HighLimit=8; % in BSM1_LT and BSM2 this is a function of T !
*Rdn=((1-Y_H)./(2.86.*Y_H)).*(mu_H*(SsOutReac5./(K_S+SsOutReac5))).*(SNOOutReac5./(K_NO+SNOOutReac5)).*XBHOutBottomClarifier.*ny_g;%considering DO=0 mgO2/L*
*t_delay=(SoOutReac5./(2.86*Rdn));*

for BSM1_LT and BSM2 (kinetic parameters are changed according to Arrhenius equations and S_{NO}HighLimit according to Henze et al., 1993!):

*mu_H(i)= 4*exp((log(4/3)/5)*(reac5part(i,16)- 15)); ;%reac5part(i,16) is T in reactor 5!*
*SNOHighLimit(i)=11.003972*EXP(-0.020295*reac5part(i,16)); %according to Henze et al., 1993*

3.2. Membership functions

Table 8 summarizes the number of (triangular and trapezoidal in the extremes) membership functions for each variable considered in the bulking, foaming and rising problems of the risk assessment model and the default limit values corresponding to 100% of certainty for each membership function (see as an example Figure 10).

Table 8: Membership functions, and their values corresponding to 100% certainty, for each variable considered in the risk assessment model (adapted from Comas *et al.*, 2008).

Variable \ Modality		**Very low**	**Low**	**Normal / Medium**	**High**	**Very high**
F/M_removed (g·g^{-1}·d^{-1})	Shape	-	trapezoidal	triangular	triangular	Trapezoidal
	rank	-	[-0.1429 -0.1429	[0.25 0.5	[0.5 0.75 1]	[0.75 1 4.027

			0.25 0.5]	0.75]		4.187]
F/M_fed ($g \cdot g^{-1} \cdot d^{-1}$)	Shape		Trapezoidal	triangular	triangular	Trapezoidal
	rank	-	[-0.0536 -0.0536 0.25 0.5]	[0.25 0.5 0.75]	[0.5 0.75 1]	[0.75 1 1.51 1.57]
$S_{O,as,3}$ ($g \cdot m^{-3}$)	Shape	Trapezoidal	triangular	triangular	triangular	Trapezoidal
	rank	[-0.4488 -0.1164 0 1]	[0 1 2]	[1 2 3]	[2 3.5 5]	[3.5 5 8.021 8.261]
SRT (d)	Shape	Trapezoidal	triangular	triangular	triangular	Trapezoidal
	rank	[-7.2 -0.8 1 3]	[0 3 6]	[3 6 9]	[6 9 12]	[9 12 20.29 23.4]
$BOD_{5,in}/N_{in}$ ($g \cdot g^{-1}$)	Shape	-	Trapezoidal	triangular	Trapezoidal	-
	rank	-	[-7.145 -7.145 10 20]	[10 20 33.33]	[20 33.3 201.3 209.3]	-
$S_{S,as,1}$ ($g \cdot m^{-3}$)	Shape	-	Trapezoidal	triangular	Trapezoidal	-
	rank	-	[-4.645 -4.645 2 10]	[2 10 18]	[10 18 131.6 143]	-
$S_{S,in}/X_{S,in}$ ($g \cdot g^{-1}$)	Shape	-	Trapezoidal	triangular	Trapezoidal	-
	rank	-	[-0.08415 -0.02183 0.1 0.2]	[0.15 0.25 0.35]	[0.3 0.45 1.55 1.56]	-
$S_{NO,as,5}$ ($g \cdot m^{-3}$)	Shape	-	Trapezoidal	triangular	Trapezoidal	-
	rank	-	[-1.429 -1.429 2 5]	[2 5 8]	[5 8 40.27 41.87]	-
Time for nitrogen gas production (d)	Shape	-	Trapezoidal	triangular	Trapezoidal	-
	rank	-	[-0.135 -0.0437 0.046 0.056]*	[0.046 0.056 0.066]*	[0.056 0.066 2.205 2.272]*	-
Risk of filamentous bulking	Shape	-	triangular	triangular	triangular	-
	rank	-	[-0.2 0 0.2]	[0.2 0.5 0.8]	[0.8 1 1.2]	-
Risk of foaming	Shape	-	triangular	triangular	triangular	-
	rank	-	[-0.2 0 0.2]	[0.2 0.5 0.8]	[0.8 1 1.2]	-
Risk of rising sludge	Shape	-	triangular	triangular	triangular	-
	rank	-	[-0.2 0 0.2]	[0.2 0.5 0.8]	[0.8 1 1.2]	-

*These values work for BSM1 control strategies with a constant sludge recycle flow rate of 18446 $m^3 \cdot d^{-1}$.

Food-to-microorganism ratio is calculated in two different ways within this risk model even though the membership functions are the same. While F/M_removed is calculated based on the daily mass flow rate of COD removed on the whole plant per unit of biomass, F/M_fed aims at detecting low organic loading (daily mass flow rate of supplied BOD per unit of biomass).

The dissolved oxygen level is evaluated in reactor 3, the nitrate concentration (S_{NO}) in reactor 5, the readily biodegradable organic matter (S_S) concentration in reactor 1 and BOD_5/N, BOD_5/P and the ratio between S_S and slowly biodegradable organic matter (X_S) are calculated for the influent. The sludge residence time (SRT) is calculated as the total mass of total suspended solids within the five reactors divided by the daily mass of total suspended solids removed from the plant via the waste sludge and the effluent.

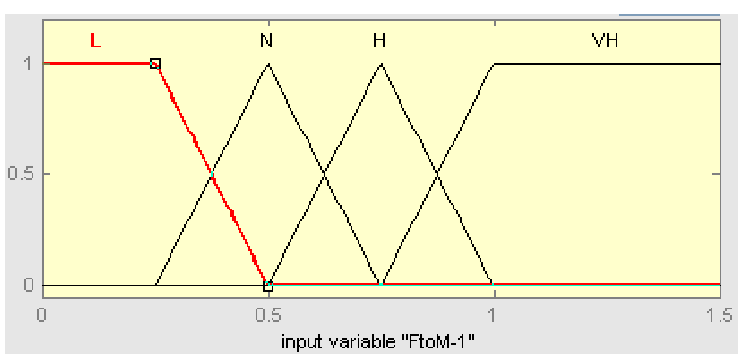

Figure 10: Example of membership function for F/M_1.

If the activated sludge recycle flow rate is not constant but changing along the simulation time, then the limits low, medium and high for the 'Time for nitrogen gas production (d)' should be calculated, everytime the recycle flow rate changes, as:

Limits for membership function ‘Low’: [-0.1292 -0.03783 Limit1(t) Limit2(t)];

Limits for membership function ‘Medium’: [Limit1(t) Limit2(t) Limit3(t)];

Limits for membership function ‘High’: [Limit2(t) Limit3(t) 2.205 2.272];

Where

Limit1(t, in days)=SludgeVolumeInClarifier (m^3 of activated sludge in clarifier)/(Q_r+1);

Limit2(t, in days)=(SludgeVolumeInClarifier (m^3 of activated sludge in clarifier)/(Q_r +1))+0.01;

Limit3(t, in days)=(SludgeVolumeInClarifier (m^3 of activated sludge in clarifier)/(Q_r +1))+0.02;

3.3. Decision matrices (Rule bases)

3.3.1 Decision matrix for Bulking due to N deficiency (middle branch of Figure 1 of TR#12 or Comas et al. 2008).

The following table summarizes the set of rules to infer potential bulking problems caused by N deficiency as a function of the BOD_5/N ratio.

Table 9: Decision matrix to evaluate the risk of bulking due to N deficiency (L: low, N: normal, H: high).

		Risk
BOD_5/N	**L**	Low
	N	Low
	H	High

In total there will be 3x1 (=3) rules.

3.3.2 Decision matrix for Bulking due to low DO (left branch of Figure 1 of TR#12 or Comas et al. 2008).

The following table summarizes the set of rules that describe the relationship between F/M_removed and dissolved oxygen to infer potential bulking problems.

Table 10: Decision matrix to evaluate the risk of bulking due to low DO (L: low, N: normal, H: high).

		F/M_removed (g COD removed·(g MLVSS)$^{-1}$·d^{-1})			
		L	**N**	**H**	**VH**
$S_{O,as,3}$ (g·m^{-3})	**VL**	Low	High	High	High
	L	Low	Medium	High	High
	N	Low	Low	Medium	High
	H	Low	Low	Low	Medium
	VH	Low	Low	Low	Low

It gives a total of 5x4 (=20) rules.

3.3.3 Decision matrices for Bulking due to low F/M ratio (right branch of Figure 1 of TR#12 or Comas et al. 2008).

2 ways of calculation:

The following table summarizes the set of rules that describe the relationship between F/M_fed and SRT to infer potential bulking problems.

Table 11: Decision matrix to evaluate the risk of bulking due to low F/M ratio (relationship between F/M_fed and SRT to infer potential bulking problems; L: low, N: normal, H: high).

		F/M_fed (g BOD supplied·(g MLVSS)$^{-1}$·d^{-1})			
		L	**N**	**H**	**VH**
SRT (d)	**VL**	Low	Low	Low	Low
	L	Low	Low	Low	Low
	N	High	Low	Low	Low
	H	High	Medium	Low	Low
	VH	High	Medium	Low	Low

The following table summarizes the set of rules that describe the relationship between S_S and SRT to infer potential bulking problems.

Table 12: Decision matrix to evaluate the risk of bulking due to low F/M ratio (relationship between S_S and SRT to infer potential bulking problems; L: low, N: normal, H: high).

		$S_{S,as,1}$ (g·m^{-3})		
		L	**N**	**H**
SRT (d)	**VL**	Low	Low	Low
	L	Low	Low	Low
	N	Medium	Low	Low
	H	High	Low	Low
	VH	High	Low	Low

Then the final value for the risk of Bulking due to low F/M simply consists in taking the maximum value of the two risks (calculated in the 2 different ways) at every time step.

3.3.4 Decision matrix for Foaming due to low F/M ratio (left branch of Figure 2 of TR#12 or Comas et al. 2008)

The following table summarizes the set of rules that describe the relationship between F/M_2 and SRT to infer potential foaming problems caused by Nocardioforms and *M. Parvicella*.

Table 13: Decision matrix to evaluate the risk of foaming due to low F/M ratio (L: low, N: normal, H: high).

		F/M_fed (g BOD supplied·(g MLVSS)$^{-1}$·d^{-1})			
		L	**N**	**H**	**VH**
SRT (d)	**VL**	Low	Low	Low	Low
	L	Low	Low	Low	Low
	N	Medium	Low	Low	Low
	H	High	Medium	Low	Low
	VH	High	Medium	Low	Low

A rule is obtained for each 'combination' of SRT and F/M_fed, e.g.:

IF F/M_fed is *low* & SRT is *normal* THEN Risk of foaming is *medium*.

3.3.5 Decision matrices for Foaming due to high readily biodegradable organicmatter (Ss/Xs) fraction (right branch of Figure 2 of TR#12 or Comas et al. 2008)

2 ways of calculation:

The following table summarizes the set of rules that describe the relationship between F/M_2 and S_S/X_S fraction to infer potential bulking problems.

Table 14: Decision matrix to evaluate the risk of foaming due to high readily biodegradable organic matter (relationship between F/M_2 and S_S/X_S; L: low, N: normal, H: high).

		F/M_2 (g BOD supplied·(g MLVSS)$^{-1}$·d^{-1})			
		L	**N**	**H**	**VH**
$S_{S,in}/X_{S,in}$	**L**	Low	Low	Low	Low
	N	Low	Low	Medium	Medium
	H	Low	Low	Medium	High

The following table summarizes the set of rules that describe the relationship between SRT and S_S/X_S to infer potential bulking problems.

Table 15: Decision matrix to evaluate the risk of foaming due to high readily biodegradable organic matter (relationship between SRT and S_S/X_S; L: low, N: normal, H: high).

		SRT (d)				
		VL	**L**	**N**	**H**	**VH**
$S_{S,in}/X_{S,in}$	**L**	Low	Low	Low	Low	Low
	N	Medium	Low	Low	Low	Low
	H	High	Medium	Low	Low	Low

Then the final value for the risk of Foaming due to high S_S/X_S simply consists in taking the maximum value of the two risks (calculated in the 2 different ways) at every time step.

3.3.6 Decision matrix for Rising

The following table summarizes the set of rules that describe the relationship between S_{NO} and the 'Nitrogen gas production time' to infer potential rising problems.

Table 16: Decision matrix to evaluate the risk of rising sludge (L: low, N: normal, H: high).

		Nitrogen gas production time		
		L	**N**	**H**
$S_{NO,as,5}$	**L**	Low	Low	Low
	N	Medium	Low	Low
	H	High	Medium	Low

4. OVERALL RISK INDEX

The risk for all operational problems considered in the risk model can be integrated into only one overall risk index. First, an integrated value for filamentous bulking must be obtained as the maximum value, at each time step, among the risks of the three type of bulking problems (caused by three different causes), at the same time as the maximum value between the risks of low F/M foaming and foaming due to high RBOM fraction provides the integrated foaming index. Then the final aggregation simply consists in picking the maximum filtered value of the integrated bulking, integrated foaming and rising risks at every time step to produce the overall risk. For a specific activated sludge system, these integrated values give an idea of the overall risk of solids separation problems as well as indicates the problem to cope with first.

$$\text{Overall risk (OR)} = \text{Max}_{\text{every t}}\ (\text{Risk}_{\text{Bulking}}, \text{Risk}_{\text{Foaming}}, \text{Risk}_{\text{Rising}})$$

5. RESPONSES (SENSITIVITY) OF THE RISK ASSESSMENT MODEL

The responses (outputs, i.e. risks) of the different problems of the risk assessment model are presented below. In total there are nine different outputs corresponding to the different settling problems and causes of the risk assessment model. These nine surfaces illustrate the sensitivity of the risk indices (outputs of the risk model) with respect to the changes of the inputs.

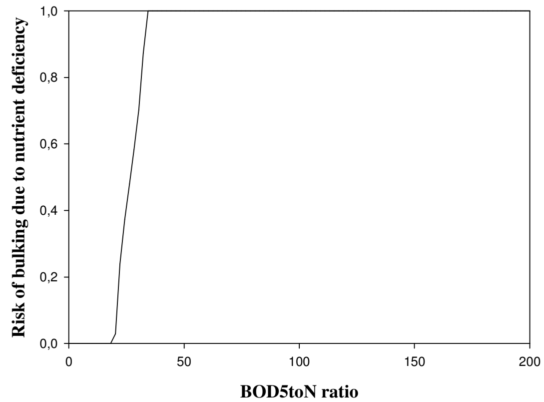

Figure 11: Risk of bulking due to nitrogen deficiency.

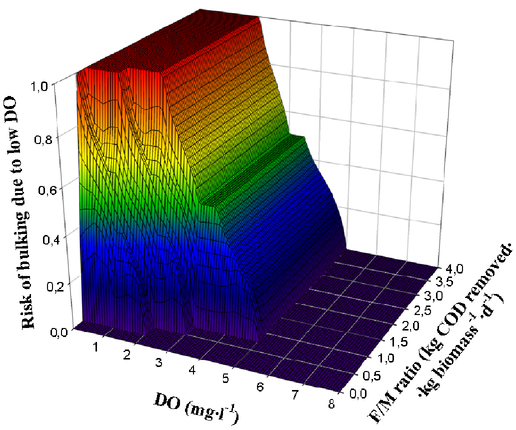

Figure 12: Risk of filamentous bulking due to low DO (F/M ratio –g COD removed·(g biomass)$^{-1}$·d^{-1}- vs. DO –g·m^{-3}-).

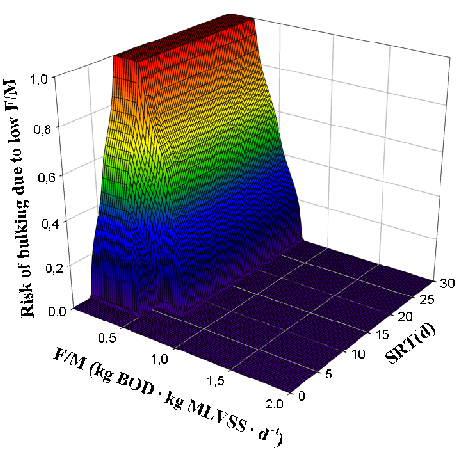

Figure 13: Risk of bulking due to low organic loading (F/M ratio -g BOD_5 supplied·(g biomass)$^{-1}$·d^{-1}- vs. SRT –d-).

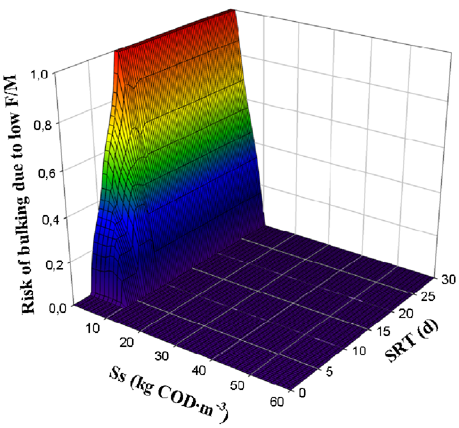

Figure 14: Risk of bulking due to low organic loading (readily biodegradable substrate, Ss vs. SRT -d-).

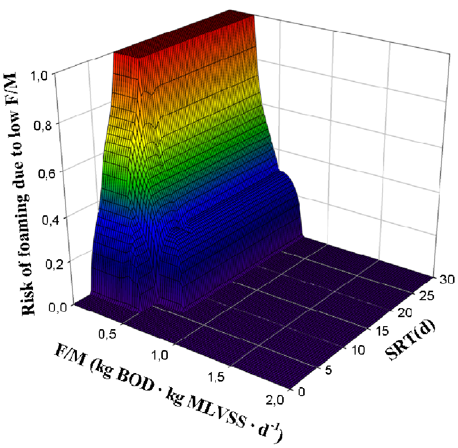

Figure 15: Risk of foaming due to low F/M (F/M ratio -g BOD_5 supplied·(g biomass)$^{-1}$·d^{-1}- vs. SRT -d-).

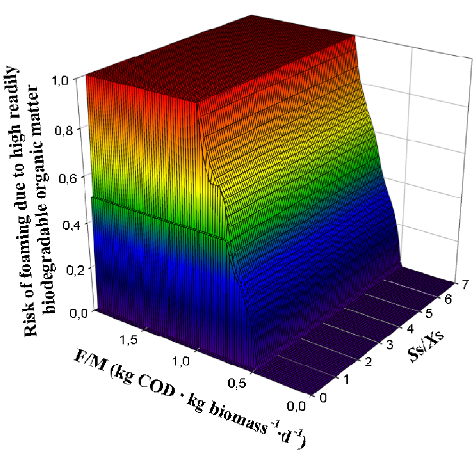

Figure 16: Risk of foaming due to high readily biodegradable organic matter (S_S/X_S) fraction (S_S/X_S vs. F/M ratio -g BOD_5 supplied·(g biomass)$^{-1}$·d^{-1}-)

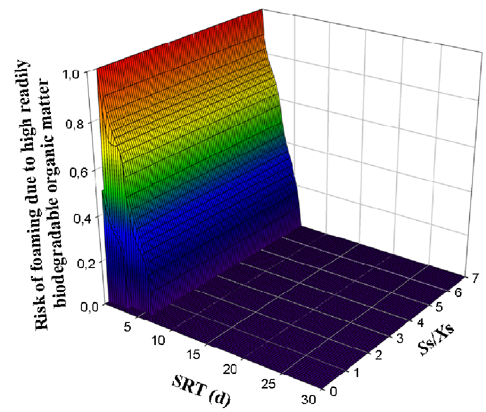

Figure 17: Risk of foaming due to high readily biodegradable organic matter (S_S/X_S) fraction (SRT –d- vs. S_S/X_S).

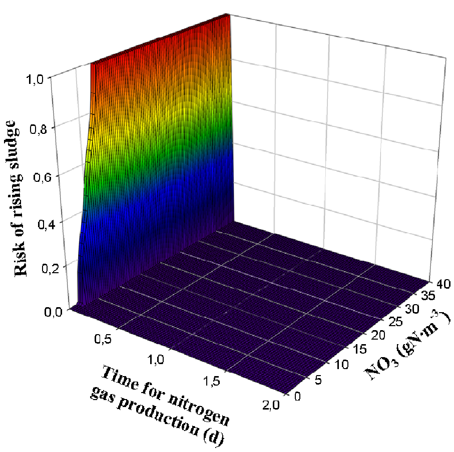

Figure 18: Risk of rising sludge (time for nitrogen gas production vs. S_{NO} concentration).

6. REFERENCES

Comas, J., Rodríguez-Roda, I., Poch, M., Gernaey, K.V., Rosen, C. and Jeppsson, J. (2008). Risk assessment modelling of microbiology-related solids separation problems in activated sludge systems. *Environmental Modelling and Software*, 23, 1250-1261.

Eikelboom D. H., Andreadakis A. and Andreasen K. (1998). Survey of filamentous populations in nutrient removal plants in four European countries. *Water Science and Technology,* 37 (4-5), 281-289.

Grady, C.P.L., Daigger, G.T. and Lim, H.C. (1999). *Biological wastewater treatment*, 2nd edition, Marcel Dekker, New York, USA.

Henze, M., Dupont, R., Grau, P. and De la Sota, A. (1993). Rising sludge in secondary settlers due to denitrification. *Water Research*, 27(2), 231-236.

Rossetti S., Tomei M. C., Nielsen P. H. and 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, 49-64.

Spering, V., Makinia, J. and Rosenwinkel, K.H. (2008). Analyzing and improving a mechanistic model for the Microthrix Parvicella in activated sludge systems. In: *proceeding of the 1st IWA/WEF Wastewater Treatment Modelling Seminar (WWTmod2008)*. Mont-Sainte-Anne, Quebec, 1th-3rd June. Canada.

Wanner, J. (1994). Activated Sludge Bulking & Foaming Control. *Technomic Publishing Company*, Basel, Switzerland.

WEF (2002). Water Environment Federation. *Activated Sludge WEF Manual of Practice No. OM-9*. Secon edition. Water Environment Federation, Alexandria (USA).