CHAPTER IV

The influence of the experimental data quality and quantity on parameter estimation precision

This chapter is in preparation for submission to Environmental Modelling and Software

ABSTRACT

Model parameters are usually estimated through minimization algorithms with respect to experimental data. However, the values obtained in this classical minimization approach are not always reliable and need critical evaluation (though the minimum of the cost function is attained). A new methodology based on the Fisher Information Matrix (FIM) is gaining importance in view of studying the model identifiability and for parameter estimation issues. The FIM matrix summarizes the amount of information obtained in each experiment because it considers the output sensitivity functions and the measurement errors of the experimental data. The information contained in this matrix depends on the quantity and quality of the experimental data. This work shows that parameter estimation values (and errors) are strongly dependent on the number of data available and on the measurement error implicit on this data.

IV.1 Motivation of this work

IV.1.1 PROBLEM STATEMENT

The utilization of modeling tools in view of process design and characterization has become very widespread, for example in the environmental engineering field. Model parameters are estimated through minimization algorithms with respect to experimental data and, afterwards, the calibrated model can be used for the process improvement (e.g. in process design or process control). However, for a reliable posterior utilization of the estimated parameters, the assessment of their confidence intervals should be as important as the estimation of the parameter value itself (Dochain and Vanrolleghem, 2001 or Brun *et al.*, 2002, among many others).

Recently, a lot of research is being conducted in assessing the precision of the parameters estimated from experimental data as for example in Walter and Pronzato, (1999); Omlin and Reichert, (1999); Dochain and Vanrolleghem, (2001); Brun *et al.*, (2002) or Marsilli-Libelli *et al.*, (2003). Confidence intervals assessment is not a straightforward issue, because many different factors are involved such as the experimental measurements quality, the inherent structure of the model or the minimization approach used (Beck, 1987). The FIM is gaining importance in view of examining the model identifiability and for parameter estimation issues. This matrix integrates the sensitivity of the measured outputs with the estimated parameters with the quantity and quality of the experimental data.

This chapter aims to be a detailed and critical description (step by step) of the methodology based on FIM for the parameter estimation error assessment. For this aim, a typical example of substrate inhibition model in activated sludge processes (Andrews's model) was used to detect the link between parameter estimation error and data quantity and quality. This example was developed so that it be could understood how model parameters were estimated. Moreover, it was thought to show how each parameter had an uncertainty which needs to be considered. The parameter estimation was carried out with the classical simplex Nelder & Mead minimization algorithm (Nelder and Mead, 1965), using as a cost function the norm of the difference between the experimental data and the modeled data. The confidence intervals were assessed through a numerical method based on FIM. Both procedures were implemented in MATLAB® (MATLAB, 1999).

IV.1.2 THEORETICAL BACKGROUND

1. Linear versus non-linear models

Linearity is a basic characteristic of the model which has a very high impact on the properties of the solution. Two different kinds of linearity can be distinguished (for example, Walter and Pronzato, 1999):

- *Input linearity*: a model is linear in its inputs when it satisfies the superposition principle with respect to its inputs (u) [eq IV.1].

$$y_M(t, \theta, \alpha \cdot u_1 + \beta \cdot u_2) = \alpha \cdot y_M(t, \theta, u_1) + \beta \cdot y_M(t, \theta, u_2) \quad \text{(IV.1)}$$

where u = input (u: vector of inputs)
y_M = measured output
θ = parameter (θ: vector of parameters)

- *Parameter linearity:* a model is linear in its parameters when it satisfies the superposition principle with respect to its parameters (θ) [eq IV.2].

$$y_M(t, \alpha \cdot \theta_1 + \beta \cdot \theta_2, u) = \alpha \cdot y_M(t, \theta_1, u) + \beta \cdot y_M(t, \theta_2, u) \quad \text{(IV.2)}$$

Table IV.1 depicts examples of all the linear possibilities:

Table IV.1 Description of the different existing linearities

Model example	Input linearity	Parameter linearity
$y_M(t+1, \theta, u) = \theta_1 \cdot u_1(t)$	YES	YES
$y_M(t+1, \theta, u) = \theta_1 \cdot y_M(t) + u_1(t)$	YES	NO
$y_M(t+1, \theta, u) = \theta_1 \cdot u_1^2(t)$	NO	YES
$y_M(t+1, \theta, u) = \theta_1 \cdot y_M^2(t) + u_1(t)$	NO	NO

Linear models are preferred since an analytical solution can be found and many mathematical tools for solving them have already been developed. On the other hand, non-linear systems may require complex numerical solutions and that is why they tend to be linearised. This linearization is performed either because the model can be considered as linear in the context of study or because one has been able to transform it as linear by a proper variable manipulation. For example, the system can be linearised around some equilibrium point (or steady state). However, the results obtained at this point will only be valid close to these parameter and variables values (Walter and Pronzato, 1999; Dochain and Vanrolleghem, 2001). Therefore, the utilization of non-linear models (where classical parameter uncertainty assessment tools can not be used) is increasing because of the narrow range of applicability of linearised non-linear models.

2. Normal distribution

The normal distribution [eq IV.3] is a two parameter mathematical distribution that describes the distribution of the population in terms of frequency versus value. In short, for a certain population with a certain mean (λ) and a certain standard deviation (σ), describes the frequency of appearance of certain value around μ.

$$f(\lambda, \sigma) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\lambda)^2}{2\sigma^2}} \quad \text{(IV.3)}$$

3. Andrews substrate inhibition model

The response of a microbial population to external substrate presence is generally modeled using Monod kinetics (see Henze *et al.*, 2000 for example). However, different patterns of dependence on substrate concentration have been described in the literature. These different responses are commonly modeled using variants of the Monod kinetics. Substrate inhibition is a frequent phenomenon observed in the literature and several modifications of Monod have appeared for its description (Luong, 1987; Han and Levenspiel, 1988 or Meriç *et al.*, 2002). Andrews modification (Andrews, 1968) is probably the most often used among all the others [eq. IV.4].

$$\mu = \frac{\mu_{MAX} \cdot S}{(S + K_S)(1 + S/K_I)} \qquad \text{(IV.4)}$$

where K_I = Inhibition constant (g/L)
K_S = Substrate affinity constant (g/L)
S = Substrate concentration (g/L)
μ_{MAX} = maximum specific growth rate (1/d)

This model derives from the Haldane equation (Haldane, 1965), which described enzyme inhibition by the formation of an inactive complex of the enzyme with two substrate molecules. In most of the cases, the ratio K_S/K_I is considered to be negligible in front of 1 and, hence, the equation 4 can be simplified in terms of equation IV.5.

$$\mu = \frac{\mu_{MAX} \cdot S}{S + K_S + S^2/K_I} \qquad \text{(IV.5)}$$

A lot of applications of Andrews inhibition model can be found in the literature. An example is the nitrification process, which is a two-step process where ammonia is first oxidized to nitrite and subsequently to nitrate. Both steps are considered to be inhibited by their own substrate: ammonia and nitrite respectively: for example Anthonisen *et al.*, (1976) or Carrera *et al.*, (2004a).

4. Parameter estimation and confidence interval assessment

The parameters can be estimated through a minimization algorithm where the weighed sum J (equation IV.6) of squared errors between model outputs $y(k,\theta)$ and the measured outputs $y_M(k)$ is minimized (k represented a certain sampling point). Q_k is a weighting matrix to balance the effect of each kind of measurement.

$$J = \sum_{k=1}^{N} \left[y(k,\theta) - y_M(k)\right]^T Q_k \left[y(k,\theta) - y_M(k)\right] \qquad \text{(IV.6)}$$

where N is the number of measurements and θ are the values of set of parameters used to calculate the model outputs.

The presented confidence interval assessment procedure is based on the Fisher Information Matrix (Dochain and Vanrolleghem, 2001). This matrix is regarded as an indicator of the amount of information contained in the experimental data. The FIM is calculated using a linearization of each one of the output signals in the neighborhood of the optimal vector of parameters θ_O. The linearization is conducted for each parameter as it is expressed in equation IV.7.

$$y(t, \theta_O + \delta\theta) = y(t, \theta_O) + \left[\frac{\delta y(t, \theta)}{\delta\theta}\right]_{\theta_O} \cdot \delta\theta = y(t, \theta_O) + Y_\theta^y(t)\delta\theta \qquad \text{(IV.7)}$$

$Y_\theta^y(t)$ stands for to the so called output sensitivity function, which are a numerical approach to the derivate of the output variable with respect to one parameter [IV.8]:

$$Y_\theta^y = \frac{\partial y}{\partial \theta}(k) = \lim_{\Delta\theta \to 0} \frac{y(t, \theta + \Delta\theta) - y\ (t, \theta - \Delta\theta)}{2\Delta\theta} \qquad \text{(IV.8)}$$

The information related to the dependencies among parameters can also be summarized in the FIM (Mehra, 1974). If Q_k is calculated as the inverse of the covariance matrix of the measurement noise (error), the FIM is defined as equation IV.9.

$$FIM = \sum_{k=1}^{N} Y_\theta^T(k) Q_k Y_\theta(k) \qquad \text{(IV.9)}$$

The FIM is a square matrix with the same number of columns (and rows) as parameters to estimate. For example, for a system with three estimated parameters (θ_1, θ_2 and θ_3) and two output variables (y1, y2), the FIM is calculated as given below in equation IV.10:

$$\begin{bmatrix} Y_{\theta 1}^{y1}(k, \theta_{OPT}) & Y_{\theta 1}^{y2}(k, \theta_{OPT}) \\ Y_{\theta 2}^{y1}(k, \theta_{OPT}) & Y_{\theta 2}^{y2}(k, \theta_{OPT}) \\ Y_{\theta 3}^{y1}(k, \theta_{OPT}) & Y_{\theta 3}^{y2}(k, \theta_{OPT}) \end{bmatrix} \cdot \begin{bmatrix} Q_1(k) & 0 \\ 0 & Q_2(k) \end{bmatrix} \cdot \begin{bmatrix} Y_{\theta 1}^{y1}(k, \theta_{OPT}) & Y_{\theta 2}^{y1}(k, \theta_{OPT}) & Y_{\theta 3}^{y1}(k, \theta_{OPT}) \\ Y_{\theta 1}^{y2}(k, \theta_{OPT}) & Y_{\theta 2}^{y2}(k, \theta_{OPT}) & Y_{\theta 3}^{y2}(k, \theta_{OPT}) \end{bmatrix} = \begin{bmatrix} a(k) & b(k) & c(k) \\ d(k) & e(k) & f(k) \\ g(k) & h(k) & i(k) \end{bmatrix}$$

$$FIM = \sum_{k=1}^{N} \begin{bmatrix} a(k) & b(k) & c(k) \\ d(k) & e(k) & f(k) \\ g(k) & h(k) & i(k) \end{bmatrix} \qquad \text{(IV.10)}$$

Where $Q_1(_k)$ corresponds to the inverse of the covariance vector of the measurement noise of the output variable y_1 for each sampling point (k). When the measurement error is considered constant along the experiment, Q(k) becomes a number instead of a vector. The FIM matrix summarizes the quantity and quality of information obtained in each experiment because it considers the output sensitivity functions and the measurement errors of the experimental data (i.e. precision of an experiment). Assuming no model mismatch, measurement noise white (i.e. independent and normally distributed with zero mean) and uncorrelated (i.e. the measurement error covariance matrix is a diagonal matrix), the inverse of the FIM provides the lower bound of the parameter estimation error covariance matrix, which can be used for assessing the estimation uncertainty of θ_O (equation IV.11).

$$COV(\theta_O) \geq FIM^{-1} \qquad \text{(IV.11)}$$

Moreover, since output sensitivities of parameters with respect to measurement(s) are calculated using a model, the FIM also depends on the structure of the model. This property of FIM can be used to study the practical identifiability (local) of the model under the available experimental data (Dochain and Vanrolleghem, 2001). The model structure is a very important issue to take into account; since the FIM procedure aforementioned is based in no model mismatch (i.e. the model can describe correctly the experimental observations with the correct parameters). This assumption of “correct model” should not be always instantaneously accepted without examination.

IV.2 Andrews model as a case study

IV.2.1 KINETIC MODEL

The modeling software used is MATLAB® 6.5. The differential equations are solved using the internal function *ode45*. This solver is based on an explicit Runge-Kutta formula, the Dormand-Prince pair. Parameter optimization to fit to experimental data is carried out by using the heuristic method implemented in the MATLAB® function *fminsearch* (Nelder-Mead-Simplex) using equation 6 as the cost function.

The output measurement considered in this study was not the substrate concentration but the oxygen uptake rate (OUR) profile. This output variable is widely used for modeling purposes because it is relatively easy to measure meanwhile provides a lot of information. The amount of oxygen taken up is stoichiometrically linked to the substrate consumption rate by means of the biomass substrate yield. Equation IV.12 shows (in terms of mass basis of COD and nitrogen) the process occurring in a classical biomass growth for a certain organic substrate.

$$\frac{1}{Y_H} CH_yO_z + \left(1 - \frac{1}{Y_H}\right) O_2 + i_{NB} \cdot NH_3 \rightarrow CH_aO_bN_c + (\ldots)\ CO_2 + (\ldots)\ H_2O \qquad \text{(IV.12)}$$

where CH_yO_z is the substrate (mg COD/L)
$CH_aO_bN_c$ is the biomass (mg COD/L)
Y_H is the biomass growth yield (g VSS/g COD)
i_{NB} is the percentage of nitrogen in biomass (g N/ g COD).

Equations IV.13a-c show the set of ordinary differential equations to describe the kinetics of the system and Table IV.2 shows the parameters involved.

$$\frac{dX}{dt} = \frac{\mu_{MAX} \cdot S}{(S + K_S + S^2/K_I)} \cdot X_H \qquad \text{(IV.13a)}$$

$$\frac{dS}{dt} = -\frac{1}{Y} \cdot \frac{dX}{dt} \qquad \text{(IV.13b)}$$

$$\left.\frac{dO}{dt}\right]_{reaction} = OUR(t) = \frac{(1 - Y)}{Y} \frac{dX}{dt} \qquad \text{(IV.13c)}$$

Table IV.2 Model parameters

Known parameters		Unknown parameters
Biomass yield: Y	0.67 (g COD_X/ g COD_S)	μ_{MAX} (1/d)
Initial heterotrophic biomass: X(0)	2000 (mg COD_X/L)	K_S (mg COD_S /L)
Initial substrate S(0)	200 (mg COD_S/L)	K_I (mg COD_I /L)
Final time	50 min	

As not all the parameters of the model depicted in equations IV.13a-c could be reliably estimated using only OUR measurements because of identifiability problems (Petersen *et al.*, 2001 or Gernaey *et al.* 2002a), the biomass yield and the initial values of substrate and biomass were assumed as known parameters whereas the kinetic parameters (μ_{MAX}, K_S and K_I) were considered the unknown parameters.

IV.2.2 EXPERIMENTAL DATA GENERATION

The "experimental data" used in this work was previously generated with the model described in equations IV.13 with

μ_{MAX} = 6 1/d
K_S = 20 mg COD_S/L
K_I = 100 mg COD_I/L.

The OUR profile obtained (from this point on reference OUR) is depicted in Figure IV.1. An error was added to each of the experimental measurements so that the data became more realistic. The measurement error noise was considered to be white, so it followed a normal distribution with zero mean. Hence, the MATLAB® function *normrnd* was used to generate random experimental error. This function returned a matrix of random numbers chosen from the normal distribution with the input parameters λ (reference OUR) and σ (measurement error). This function is only available with the statistics toolbox. However, the software provided also contains a similar version with the MATLAB® function *randn* which does not require any special toolbox, but the values generated do not follow a normal distribution. The experimental data obtained (using 3 % of measurement error) is also depicted on Figure IV.1.

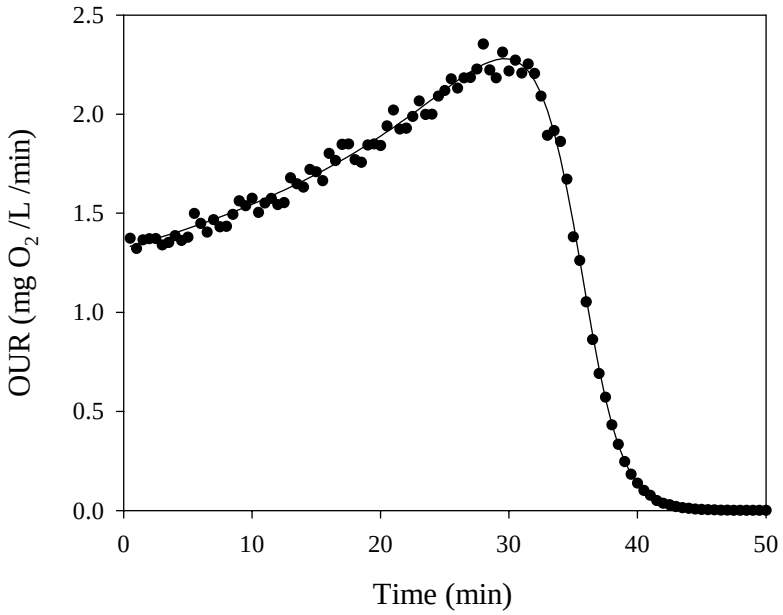


Figure IV.1 OUR profiles: Reference (solid line) and experimental (•) (3 % of meas. error)

IV.2.3 IMPLEMENTATION OF THE PROCEDURE

The structure of the software implemented is described next. First of all, the "experimental" OUR measurements were generated as indicated in the paragraph above. Then, μ_{MAX}, K_S and K_I were estimated through *fminsearch.* The cost function was defined as the norm of the vector resulting of the difference between the experimental and the modeled data [eq. IV.6]. The OUR profile obtained using the optimal parameters (which minimize the cost function) was used for the calculation of the sensitivity functions, which are the basis for the calculation of the FIM. The sensitivity function of each parameter with respect to the output measurement (OUR) was calculated as described in equation IV.8. The FIM was calculated by means of equation IV. 9.

Q_K is usually chosen as the inverse of the measurement error covariance matrix, and it is a square matrix with the same number of files (and columns) as output measurements used. In this study, the measurement error covariance matrix for the respirometric measurements (OUR) was calculated following the Petersen's method (Petersen *et al.*,

2001). The Petersen's method used the data from a period when the value of the output variable was perfectly known. For example, Q_{OUR} can be estimated using a phase without external substrate (i.e. endogenous respiration phase) in which OUR is assumed to be relatively constant during a short-term period e.g. 30 min (Figure IV.2). In this OUR constant period, the average of the data and the resulting residuals (particularly the difference between the average OUR value and the experimental data) are calculated. The measurement error (s^2) is then calculated as follows [eq. IV. 14]:

$$s^2 = \frac{SSE}{N - p} \qquad \text{(IV.14)}$$

where SSE = Sum Squared Errors
N = Number of OUR measurements
p = Number of parameters

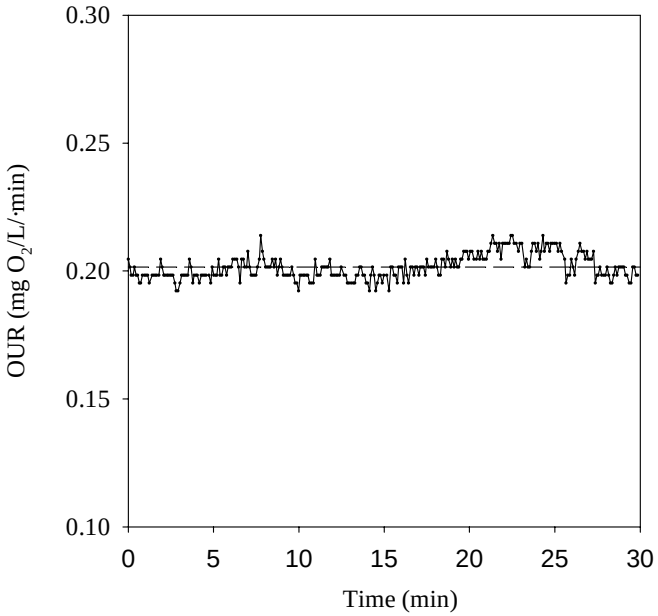


Figure IV.2 Example of Q_{OUR} estimation

Then, the measurement error weighting matrix, Q_k, is calculated taking the inverse of s^2. This choice of Q_k means that the more a measurement error is noise corrupted, the less it will count in the FIM. Finally, approximate standard errors for the estimated parameters can be calculated as the square root of the diagonal elements of the inverse of the FIM (Dochain and Vanrolleghem, 2001):

$$\sigma(\theta_i) = \sqrt{V(i,i)} \quad \text{where } V = 1/FIM \qquad \text{(IV.15)}$$

Hence, the higher the FIM elements, the lower the standard errors estimated. This is understandable if one analyzes which factors make the FIM be higher [eq. IV.9]: high parameter sensitivity (high Y^y_θ) and low measurement errors (low s^2, high Q_K). Once the procedure is implemented, several changes can be developed to study their influence on the parameter estimation and the confidence interval assessment. These changes can be related to the model itself (variations in the model equations), to the experimental data used (number and quality) or to the initial guesses in the minimization algorithm. In this study, the influence of the number of data used for parameter estimation and the influence of the measurement error of this data were analyzed.

IV.3 Influence of quantity of experimental data

In order to analyze the influence of the data quantity, the whole procedure was run several times varying the "experimental" measurement interval (sampling frequency): this was developed by varying the number of elements of the same "experimental" OUR vector. The parameter estimation results and standard deviation assessment are depicted in Figure IV.3. On the one hand, as can be seen in Figure IV. 3-left, there is a high variability in the parameter values depending on the measurement interval used. The general trend observed is the lower the sampling frequency (the higher the number of measurements), the closer the final estimated values to the real ones. When the measurement interval used was higher than one measurement each two minutes (in this case, less than twenty measurements), the values of the parameter estimation were not reliable and differ a lot from the real ones. However, there is also the possibility to obtain a good set of parameters with a fortunate set of experimental data, as for example the obtained with a measurement interval of 3.5 minutes. In addition, if the parameter fitting procedure was repeated with another experimental data set, different parameter values could be obtained although the same measurement interval was used. On the other hand, the Figure 3-right shows the ratio of the standard deviation of each parameter (calculated through equation IV. 15) over the value of the parameter versus the measurement interval. The graphic shows that the lower the measurement interval is, the more reliable the values of the estimated parameters become.

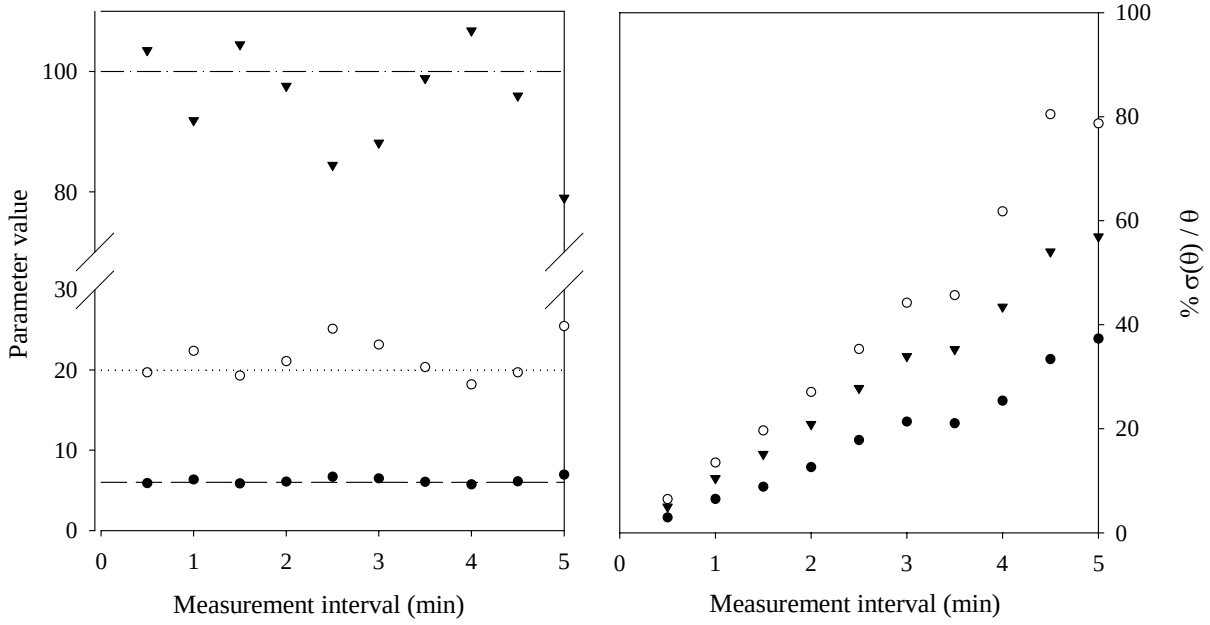


Figure IV.3 Influence of the measurement interval on the estimated parameter value (LEFT) and on the confidence interval assessment (RIGHT).
μ_{MAX}: filled circle (estimated values), dashed line (mean)
K_S: white circle (estimated values) and dotted line (mean)
K_I: triangle (estimated values) and dash-dotted line (mean)

An important fact to highlight is that even though the measurement error was unique, the parameter estimation errors obtained are different among the three parameters. The reason can be found when analyzing the sensitivity of each parameter with respect to the OUR profile. As an example, let us compare μ_{MAX} and K_I. As can be seen in Figure IV.4, in the experiment proposed in this work, the sensitivity of the μ_{MAX} was higher than the sensibility of the K_I in terms of absolute value. In short, small variations on the μ_{MAX} value would have stronger effects on the OUR profile than small variations on the K_I value.

Hence, the value of μ_{MAX} would always be more easily assessed and the relative estimation error of μ_{MAX} would be always lower than the relative estimation error of K_I.

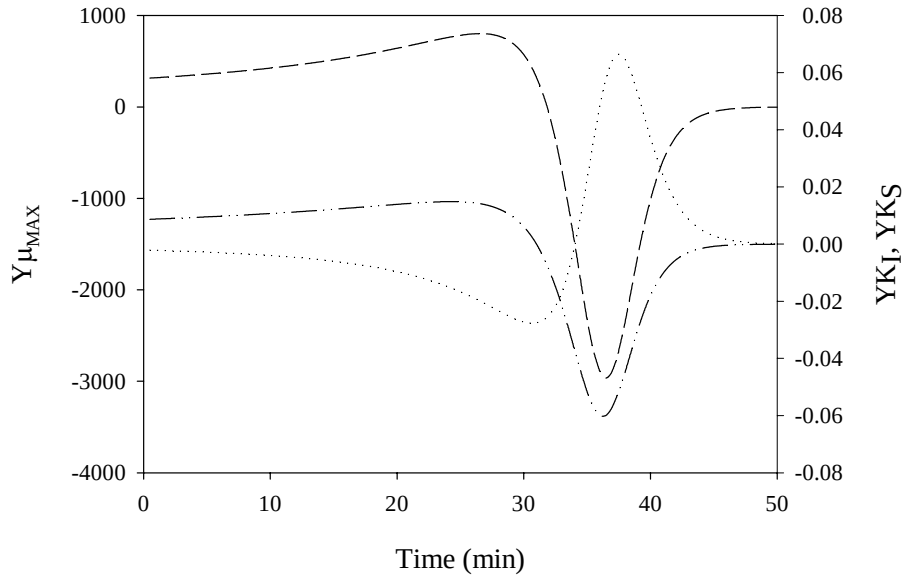


Figure IV.4 Sensitivity functions of OUR with respect to μ_{MAX} (dashed), K_S (dotted) and K_I (dash-dotted).

IV.4 Influence of quality of experimental data

The influence of the data quality was studied through the variation of the measurement error of the experimental data. This measurement error appeared twice in the exercise: firstly when generating the "experimental" data and secondly when calculating the measurement error covariance matrix. The results obtained using different measurement errors are depicted in Figure IV.5. At first glance, the results obtained agreed with what should be expected: the increase of measurement error implied high dispersion on the parameter estimation (Figure IV.5-left) and a decrease in the reliability of the confidence intervals (Figure IV.5-right). Although a good set estimated parameter values can be obtained with a high measurement error (see for example error 0.06), they have a high uncertainty around 10-20 % (Figure IV.5 right). In addition, the repetition of the experiment with the same measurement error would probably lead to different estimated parameter values.

The main idea derived form these two experiments is that parameter estimation should be a two-step process: parameter optimization + parameter error assessment. Otherwise, if only the first step was developed, it is not possible to know whether the obtained parameters were far from the reality or not. For example, one should be able to know that the parameter estimation values obtained with the higher measurement error and the lower frequency sampling are not as reliable as the ones obtained with lower measurement error and higher sampling frequency. As it is shown on Figures IV.3 and IV.4, these parameter estimation values obtained in both cases are very different.

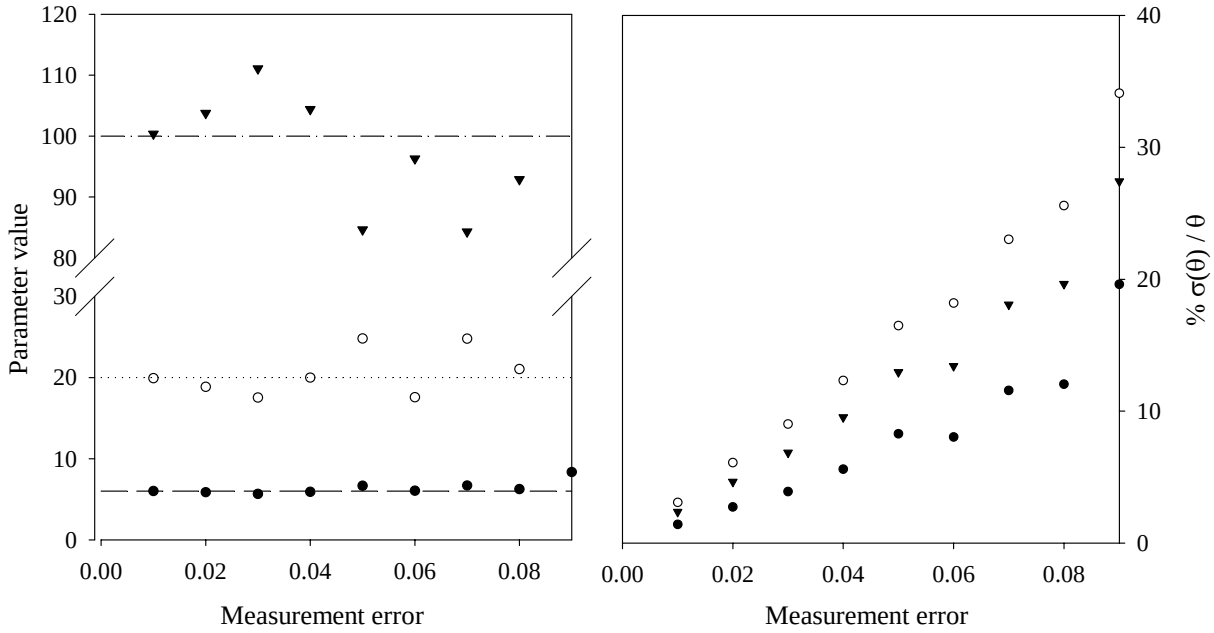


Figure IV.5 Influence of measurements error on the parameter estimation (LEFT) and on the confidence interval assessment (RIGHT)
μ_{MAX}: filled circle (estimated values), dashed line (mean)
K_S: white circle (estimated values) and dotted line (mean)
K_I: triangle (estimated values) and dash-dotted line (mean)

CHAPTER IV Conclusions

- An easily implementable procedure for parameter estimation and confidence interval assessment is described using the common Andrews substrate inhibition model as an example.

- In addition, to understand the parameter estimation and confidence interval assessment, it is shown that a critical evaluation on the parameter estimation values is always required. The importance of the confidence interval assessment in view of future usage of the parameter estimation values (for example for process design or control) is also highlighted.

- The effect of the quantity (number of measurements) and quality (measurement error) of data on the parameter estimation and confidence interval assessment is also clearly depicted. In general, the results show that the increase of the measurement frequency and the decrease of the measurement error imply a more accurate parameter estimation (in terms of proximity of the estimated value to the real one and in terms of confidence interval assessment).

CHAPTER V

BIOLOGICAL COD REMOVAL

CHAPTER V.A

Limitations of ASM1 and ASM3: a comparison based on batch OUR profiles from different full-scale WWTP

This chapter is published as:

Guisasola A., Sin G., Carrera J., Baeza J.A. and Vanrolleghem P.A. (2004b) Limitations of ASM1 and ASM3: a comparison based on batch OUR profiles from different full-scale WWTP. In: Proceedings 4th IWA Word Water Congress and Exhibition. Marrakech. Morocco. 19-24 September 2004.

ABSTRACT

This chapter deals with the two most popular models for the description of the biological COD removal: ASM1 and ASM3. However, some numerical inconsistencies arise when using these models to interpret the data obtained in short-term respirometric batch experiments. Both models were fitted to four different respirometric batch profiles obtained with biomass from different WWTPs. The parameter estimation results of both models are interpreted and discussed in view of their possible (mechanistic) biological meaning. Further, the practical (local) identifiability of both models is compared in view of unique parameter estimations. Improvements to modelling substrate conversion processes are discussed based on the mechanistic meaning and identifiability of the parameter estimates.

V.A.1 Modelling COD removal: ASM models

In 1987, the International Water Association (IWA) introduced the Activated Sludge Model nº1 (ASM1) for the description of the biological COD and nitrogen removal (Henze *et al*., 2000). In this model, biomass was considered to grow solely on the external substrate present and the oxygen consumption after the external substrate depletion was explained with the decay of biomass. In the conventional activated sludge processes, the feed regime is highly variable and biomass is subjected to alternating conditions of external substrate availability (feast phase) and absence of external substrate (famine phase). Under these dynamic conditions, internal storage polymers play an important role in the substrate consumption (van Loosdrecht *et al.,* 1997).

Recently, a new model for the COD removal (ASM3) has been developed mainly to take this storage phenomenon into account (Henze *et al*., 2000). The main innovation of this model is the assumption that all the readily biodegradable organic substrates taken up under feast conditions are directly converted into stored material. These stored compounds become the carbon and energy source for growth purposes in the subsequent famine period. In ASM3, the decay processes are replaced with the endogenous processes. The conceptual basis of ASM3 has been largely criticized and alternative models taking into account simultaneous storage and growth processes were proposed (e.g. van Aalst-van Leeuwen *et al.,* 1997;, Krishna and Van Loosdrecht, 1999; Beccari *et al.,* 2002; van Loosdrecht and Heijnen, 2002 or Karahan Gül *et al.,* 2003).

V.A.2 Experimental design

In this study, parameter estimation and identifiability issues of ASM3 in view of model calibration are addressed and compared with the well-studied ASM1 model. To this aim, oxygen uptake rate (OUR) measurements of biomass sampled from three different full-scale WWTPs were used. This work was derived from cooperation between the BIOMATH group (Ghent University) and the Environmental Engineering Group of UAB. For this reason, the experimental task was developed in two different laboratories with two different experimental set-ups for the OUR measurement.

V.A.2.1 EXPERIMENTAL SET-UPS

The experimental work was performed in two different set-ups. On the one hand, tests A and B were performed in the LFS-respirometer. On the other hand, tests B and C were performed using the hybrid-respirometric set-up. Both equipments are detailed in the equipment section (Chapter III.1.1 and III.1.3 respectively). Although based on different concepts, both set-ups allowed the measurement of a continuous OUR profile.

In both set-ups, the biomass was first aerated overnight to reach the endogenous-state. Then, a first pulse of acetate was added to induce a “wake-up” effect on the biomass activity (Vanrolleghem *et al.,* 1998). At the same time, ammonia in excess and ATU (30

mg/l) was added to avoid growth-limitation and nitrification, respectively. Activated sludge sampled from three different WWTP was used during experimental work: experiment D used biomass from the Maria Middlelares WWTP (Gent, Belgium), which performs COD removal and nitrification. Experiment C used biomass from Ossemeersen WWTP (Gent, Belgium), which performs COD removal, nitrification and denitrification the same way as Granollers WWTP (Catalonia, Spain) whose biomass was used for experiments A and B. These biomass samples were analysed for TSS and VSS according standard methods (APHA, 1995).

V.A.2.2 PARAMETER ESTIMATION

Modelling, simulation and parameter estimation were performed using MATLAB 6.5 (The MathWorks, Natick, MA). The differential equations were solved using an explicit Runge-Kutta (4,5) formula. Parameter estimation was carried out by using the Nelder-Mead Simplex search method, where the weighed sum J [eq. V.A1] of squared errors between model outputs $y(t_k,\theta)$ and the measured outputs $y_M(k)$, with Q_k as weighting matrix (equal to the inverse of the measurement error covariance matrix), is minimised:

$$J = \sum_{k=1}^{N} \left[y(t_k,\theta) - y_M(k)\right]^T Q_k \left[y(t_k,\theta) - y_M(k)\right] \qquad \text{(V.A1)}$$

where N is the number of measurements. Each of the output signals can be linearised in the neighbourhood of the optimal vector of parameters θ_O (Dochain and Vanrolleghem, 2001):

$$y(t,\theta_O + \delta\theta) = y(t,\theta_O) + \left[\frac{\delta y(t,\theta_O)}{\delta\theta^T}\right]_{\theta_O} \cdot \delta\theta = y(t,\theta_O) + Y_\theta^T(t)\delta\theta \qquad \text{(V.A2)}$$

where $Y_\theta(t)$ is the so called output sensitivity function. If Q_k is the covariance matrix of the measurement noise, the Fisher Information Matrix (FIM) is defined as:

$$FIM = \sum_{k=1}^{N} Y_\theta^T(t_k) Q_k Y_\theta(t_k) \qquad \text{(V.A3)}$$

The FIM matrix summarises the quantity and quality of information obtained in each experiment because it considers the output sensitivity functions and the measurement errors of an experimental data (i.e. accuracy of an experiment). Assuming white measurement noise and no model mismatch, the inverse of the FIM provides the lower bound of the parameter estimation error covariance matrix, which can be used for assessing the estimation uncertainty of θ_O [eq. V.A4].

$$COV(\theta_O) \geq FIM^{-1} \qquad \text{(V.A4)}$$

Moreover, since output sensitivities of parameters with respect to measurement(s) are calculated using a model, the FIM also depends on the structure of the model. This property of FIM can be used to study the practical identifiability (local) of the model under the available experimental data (Dochain and Vanrolleghem, 2001).

V.A.2.3 ASM MODELS SIMPLIFICATION

The mathematical models used (see Table V.A1) to interpret the experimental data are simplified versions of ASM1 and ASM3 respectively: aerobic degradation of COD as substrate. The processes included in Table V.A1 are described in detail in Henze *et al.,* (2000). An empirical factor was added in the kinetics of two processes (processes nº 1 and 4) to describe the fast transient period (1-3 minutes) in reaching the maximum OUR observed after the substrate addition. This phenomenon, known as "start-up", can be mathematically described by a first order model (Guisasola *et al*., 2003; Vanrolleghem *et al.* 2004*)*.

For the parameter estimation, the initial concentration of biomass, $X_H(0)$ is estimated using the baseline endogenous OUR level prior to substrate addition, while fixing the decay rate coefficient (b_H) to its default value assigned in the corresponding model. This approach was adopted since it is not possible to obtain unique values of both b_H and $X_H(0)$ using OUR measurements alone. Hence, only one of the two parameters can be estimated and the other one should be fixed. In this study, b_H was fixed to its default value since it does not vary significantly among different WWTPs.

Table V.A1. The simplified ASM1 and ASM3 models used in this work (M stands for the Monod kinetics of the corresponding parameter: e.g. $M_O = \frac{S_O}{S_O + K_O}$)

ASM1 Processes	X_H	X_{STO}	X_S	S_S	S_O	Kinetics
1. Growth on S_S	1			$-\frac{1}{Y_{H,S}}$	$-\frac{1-Y_{H,S}}{Y_{H,S}}$	$\mu_H M_S \cdot M_O X_H \cdot (1 - e^{-\frac{t}{\tau}})$
2. Biomass decay	-1		$(1-f_{XS})$			$b_H X_H$
3. Hydrolysis			-1	1		$k_H M_{X_S / X_H} \cdot M_O \cdot X_H$
ASM3 Processes	X_H	X_{STO}	X_S	S_S	S_O	Kinetics
4. S_S Storage		1		$-\frac{1}{Y_{STO}}$	$-\frac{1-Y_{STO}}{Y_{STO}}$	$k_{STO} \cdot M_S \cdot M_O \cdot X_H \cdot (1 - e^{-\frac{t}{\tau}})$
5. Growth on X_{STO}	1	$-\frac{1}{Y_{H,STO}}$			$-\frac{1-Y_{H,STO}}{Y_{H,STO}}$	$\mu_H \cdot M_{X_{STO} / X_H} \cdot M_O \cdot X_H$
6. Endogenous respiration	-1				-1	$b_H \cdot M_O \cdot X_H$
7. X_{STO} respiration		-1			-1	$b_{STO} \cdot M_O \cdot X_H$

where b_{STO} = X_{STO} respiration rate (1/d)
f_{XS} = fraction of slowly biodegradable products from lysis (g COD/ g COD)
k_H = hydrolysis maximum rate (1/d)
K_O = oxygen affinity constant (mg O_2/L)
k_{STO} = maximum storage rate (1/d)
S_O = dissolved oxygen (mg O_2/L)
S_S = external substrate (mg COD_S/L)
X_H = heterotrophic biomass (mg COD_X/L)
X_S = slowly biodegradable products (mg COD/L)
X_{STO} = storage products (mg COD_{STO}/L)
$Y_{H,S}$ = heterotrophic growth yield on external substrate (g COD_X/g COD_S)
$Y_{H,STO}$ = heterotrophic growth on storage products (g COD_X/ g COD_{STO})
Y_{STO} = heterotrophic storage yield (g COD_{STO}/g COD_S)
$\mu_{H,S}$ = maximum growth rate on external substrate (1/d)
$\mu_{H,STO}$ = maximum growth rate on storage products(1/d)
τ = time constant (start-up) (min)

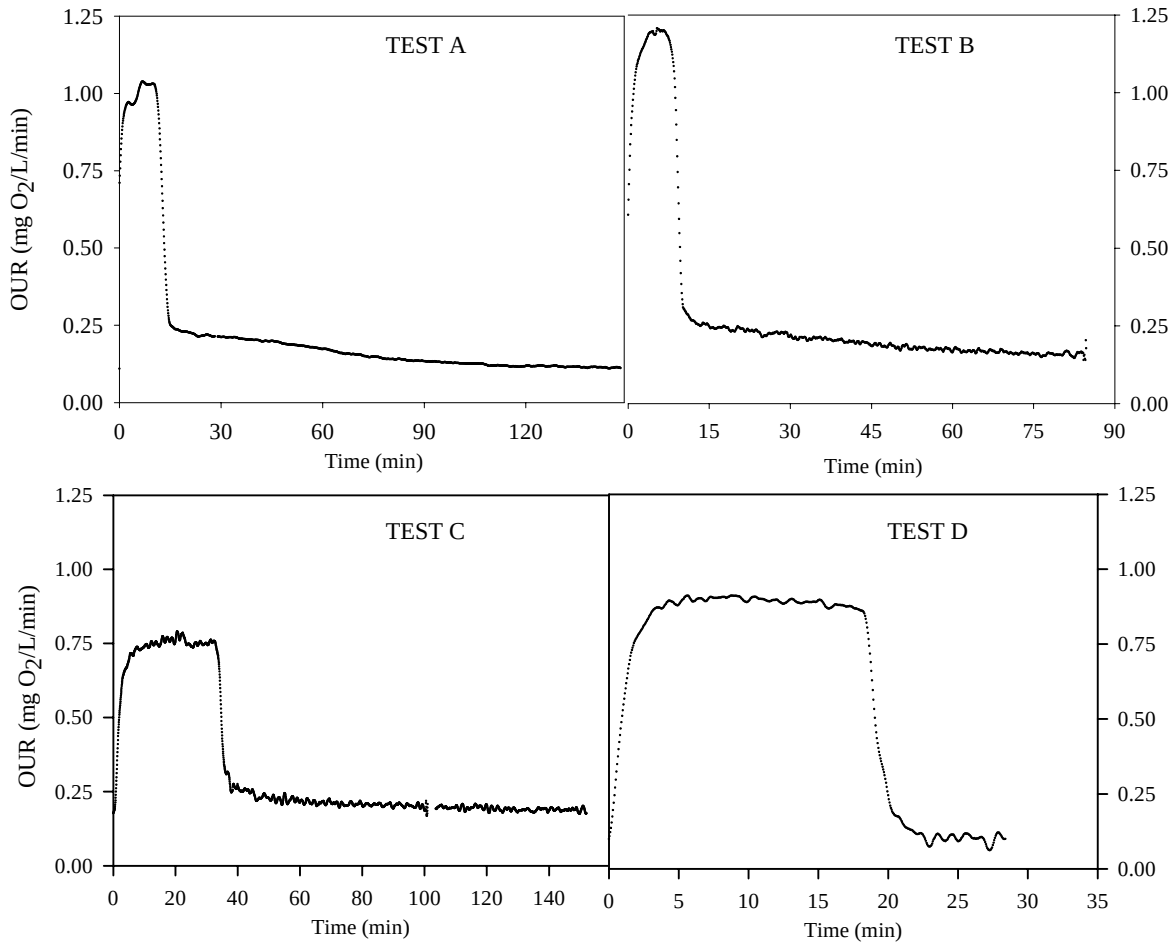
V.A.3 Results and discussion

V.A.3.1 EXPERIMENTAL RESULTS

Experiment V.A1 (Table V.A2) consisted of four different respirometric batch tests obtained from three different full-scale WWTPs (Figures V.A1a-d). These OUR profiles were used to investigate the model-fit performance and identifiability for ASM1 and ASM3 and showed different trends despite a pulse of the same substrate (acetate) was added, because the biomass used in each experiment was withdrawn from different WWTPs. The differences appreciated among the OUR profiles were probably linked to the operational conditions of the plant (i.e. SRT or alternating feed and famine conditions).

Table V.A2 Experiment V.A1

EXPERIMENT V.A1	Respirometric batch experiments from different WWTP
Equipments	LFS respirometer (V_0 = 1 L)
	Hybrid respirometer (V_0 = 3 L)
pH	7.8
Temperature	25 ºC
Pulses	TEST A and B (Granollers) - 50 mg COD/L
	TEST C (Ossemeersen) – 62.2 mg COD/L
	TEST D (Maria Meddelares) – 58 mg COD/L



Figures V.A1 (a-d) Experimental OUR profiles obtained.

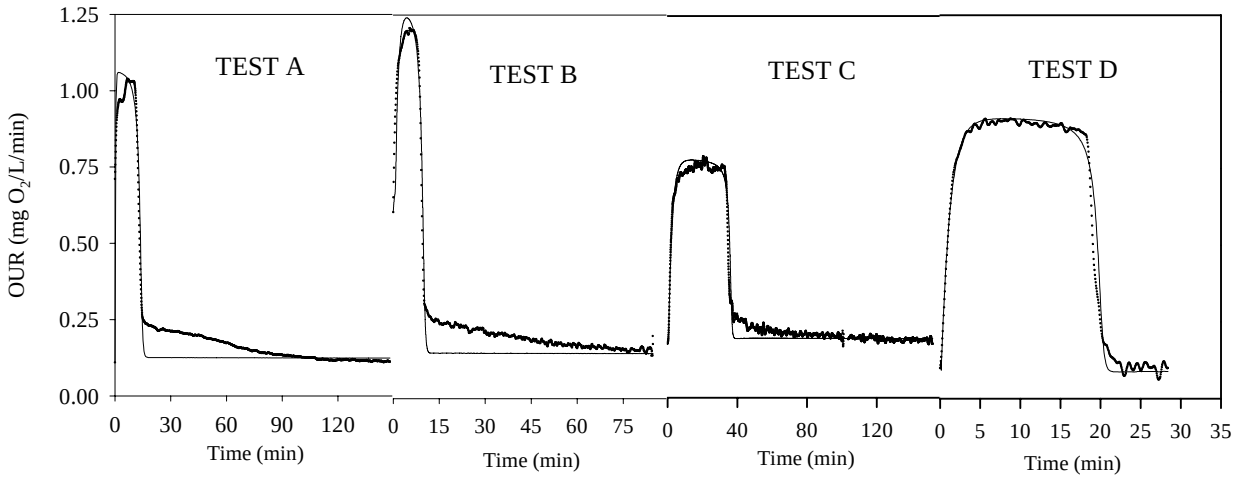
For instance, in tests A, B and C two different phases could be easily distinguished. The first phase was related to the external substrate consumption, while the second phase corresponded to the consumption of the previously stored internal polymer. In Granollers and Ossemeersen WWTPs (tests A, B and C), both nitrification and denitrification took

place and the biomass was subjected to alternating anoxic and aerobic conditions under the dynamic influent substrate/wastewater pattern. Under these conditions of alternating external substrate availability, bacteria capable of storing substrate have a competitive advantage because they are able to balance their growth rate under continuously changing conditions as described in van Loosdrecht *et al.* (1997). They argued that under the periods of excess substrate, non-storage bacteria have to invest extra energy to grow faster in the presence of substrate and will deteriorate in the periods without substrate.

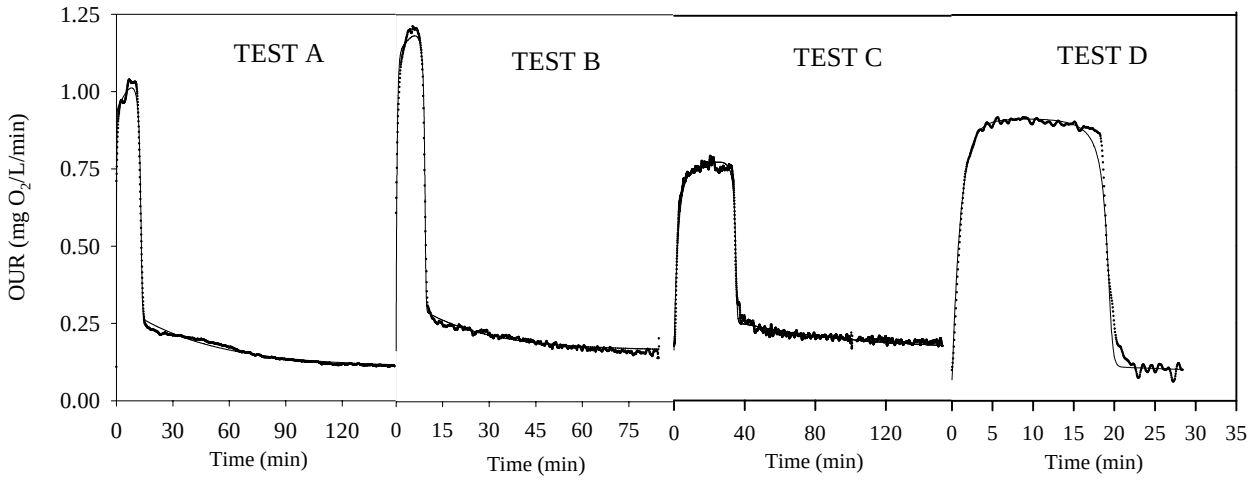
On the other hand, one unique shoulder can be distinguished in test D which is a typical ASM1 type OUR profile. The biomass used in this experiment was withdrawn from the Maria Middlelares WWTP, which is continuously aerated, and is probably subjected to rather stable influent dynamics. In other words, the feast and famine phases are probably less pronounced in this WWTP. A deeper study in the effect of WWTP operational conditions on respirometric batch tests is done on Chapter V.D.

V.A.3.2 ASM1 AND ASM3 FITTINGS

The fittings of the models (ASM1 and ASM3) are depicted in Figures V.A2 and V.A3 respectively and the results of the parameter estimation are given in Table V.A3. The parameter estimation errors obtained are quite small, in part because the method used is known to give too optimistic results due to autocorrelation in the OUR data (Dochain and Vanrolleghem, 2001).



Figures V.A2 (a-d) Experimental OUR profiles (dotted) versus simulated OUR of ASM1 (solid).



Figures V.A3 (a-d) Experimental OUR profiles (dotted) versus simulated OUR of ASM3 (solid).

A first glance at Figure V.A2 shows that ASM1 is not able to describe the tail observed in tests A and B, where the storage effect is emphasized. Many respirograms can be found in the literature with this tail, and the main criticism that ASM1 may receive is that these tails are not predicted when the feed solely contains readily biodegradable substrate. In contrast, when using typical raw wastewater the effect of storage would be lumped in the hydrolysis process and, hence, ASM1 could describe correctly the experimental OUR profile. As shown in Table V.A3, ASM3 better describes all the experimental profiles when comparing the sum of squared errors (SSE). This fact was expected since more parameters are estimated in ASM3 (seven versus five). The more parameters to be estimated, the more chances to obtain better fittings.

Hence, the clearer the storage effect is, the higher the improvement of using ASM3 instead of ASM1. This improvement is even observed in the experiment D, where no storage can be appreciated. However, once a good fitting is obtained, an analysis on the mechanistic meaning of the parameter estimation results is required. In the following, the analysis of the parameter estimation results of both models is developed.

Table V.A3 Parameter estimation results and confidence intervals
(COD_X – COD biomass, COD_S – COD external substrate, COD_P – COD PHA)

ASM1 fittings	TEST A	TEST B	TEST C	TEST D
$\mu_{H,S}$ (1/d)	3.876±0.003	4.112±0.009	1.020±0.001	2.951±0.001
$Y_{H,S}$ (g COD_X/g COD_S)	0.757±0.001	0.792±0.001	0.666±0.001	0.726±0.001
K_S (mg COD/L)	1.789±0.005	1.63±0.02	0.558±0.008	0.718±0.005
τ (min)	0.240±0.007	0.95±0.05	2.072±0.006	1.065±0.007
$X_H(0)$ (mg COD_X/L)	1250	1800	2300	1250
SSE	2.386	2.192	1.673	0.966
ASM3 fittings	**TEST A**	**TEST B**	**TEST C**	**TEST D**
k_{STO} (1/d)	4.88±0.009	4.679±0.009	1.056±0.002	3.027±0.007
Y_{STO} (g COD_P/g COD_S)	0.796±0.006	0.831±0.006	0.715±0.005	0.75±0.01
K_S (mg COD/L)	0.80±0.02	0.91±0.02	0.69±0.02	0.79±0.02
$\mu_{H,STO}$ (1/d)	28.1±0.5	64±2	19.8±0.4	51± 32
$Y_{H,STO}$ (g COD_X/g COD_P)	0.804±0.002	0.921±0.002	0.838±0.002	0.96±0.01
τ (min)	0.123±0.005	0.34±0.01	2.21±0.03	1.02±0.03
$X_H(0)$ (mg COD_X·L)	1000	1500	2000	1000
SSE	0.560	0.744	0.999	0.755

where K_S: substrate affinity constant (mg COD_S/L)

V.A.3.3 EVALUATION OF THE PARAMETER ESTIMATION RESULTS

In the experiments with apparent storage (tests A and B) two different shoulders can be easily distinguished. According to ASM1, the direct growth on external substrate is the cause of the first shoulder, whereas the ASM3 model links this first consumption to the storage of substrate into internal polymer. These processes have different default yield values: 0.67 for the growth yield in ASM1 and 0.85 for the storage yield in ASM3, because less energy is required to store external substrate than to produce new cells. When fitting experimental data to ASM1, the growth yields obtained (0.76 and 0.79) were higher than 0.67. This finding indicated the storage presence because less oxygen consumption was observed while the majority of the substrate flux was incorporated into biomass (e.g. as new cells in ASM1 or internal storage products + new cells ASM3).

On the other hand, the storage yields obtained by fitting ASM3 (0.79 and 0.83) were somehow lower than the default ASM3 (0.85), probably reflecting that not all the acetate consumed is stored. Yield values for storage with acetate in this range were also experimentally observed in other similar works: van Aalst-van Leeuwen *et al.,* 1997 (0.75); Krishna and van Loosdrecht, 1999 (0.73); Koch *et al.,* 2000 (0.72); Karahan-Gül *et al*., 2003 (0.78).

These observations (i.e. higher growth yield in ASM1 and lower storage yields in ASM3) agreed with the fact that both growth and storage processes occurred simultaneously and part of acetate was used for growth, while the rest was stored.

Although the tail is accurately fitted by ASM3 (Figure V.A3), the mechanistic meaning of the parameters related to this tail was highly questionable. First of all, the parameter estimation error of μ_H was the highest of all the parameters (especially in test D, where no storage effect was observed). Moreover, both the maximum growth rate (μ_H) and the growth yield ($Y_{H,STO}$) estimated by ASM3 (see Table V.A2) are noticeably higher than the default ones, 2 d^{-1} and 0.63 respectively. The reason for these high values could be that the real production of X_{STO} (e.g. PHA) during the experiment was less than the one predicted by the model. This is not surprising since ASM3 considers that all the acetate is stored. Hence, the experimentally observed tail (see Figure V.A3) is much smaller than the one predicted by the ASM3 with its default values. From a model-fit point of view, the value of μ_H must be increased so that the endpoint of PHA consumption can be correctly predicted. From a parameter identifiability point of view, however, $Y_{H,STO}$ is correlated with μ_H (See below). Therefore an increase in μ_H is compensated by an increase in the estimate of Y_{GSTO} so that the total oxygen consumed is correctly predicted.

High values of the ASM3 growth yield, $Y_{H,STO}$, were also observed in the literature when fitting ASM3 to experimental data (Koch *et al.,* 2000; Karahan-Gül *et al.,* 2003; Beccari *et al.,* 2002). These high values were contradicting the conceptual basis of ASM3 since the predicted growth yield values did not have any longer mechanistic meaning. On the other hand ASM1 was not able to predict the tail often observed in OUR obtained from batch experiments. Nevertheless, for a profile with low storage effect (as test C) an increase in the b_H value could result in better model-fit.

Test D seemed to be the only OUR profile which was in agreement with ASM1 because the typical storage tail was not observed. However, the estimated growth yield (0.73) (see Table V.A3) was still higher than the default value in ASM1 (0.67). This observation strongly suggested the presence of storage phenomenon and as such, it again supported the aforementioned observation of simultaneous storage and growth. Concerning the fit of ASM3 to test D, non-reliable/non-mechanistic parameter estimates were obtained (see Table V.A3), particularly the values referring to the growth on storage product. For example, μ_H was around 50 d^{-1} for the same reason explained below: the actually experimental produced X_{STO} is lower than what the ASM3 predicts.

V.A.3.4 SIMULTANEOUS GROWTH AND STORAGE HYPOTHESIS

In general, more reliable parameter values would be obtained if the ASM3 model could describe that part of the acetate was used directly for growth. In this case, the model would predict less PHA production and the predicted tail would be lower and, then, closer to the experimental data. Moreover, a decrease on the values of μ_H and $Y_{H,STO}$ would be necessary to describe the tail. The reduction of the tail (i.e. the reduction of the oxygen consumption due to the storage process) as a function of a percentage of the acetate used directly for growth is depicted in Figure V.A4. In this figure, four simulations with a model coming from a combination of ASM1 and ASM3 are performed.

This observation of simultaneous growth and storage on the external substrate has already been developed in some metabolic models such as: van Aalst-van Leeuwen *et al.* (1997) and, recently, in the works of van Loosdrecht and Heijnen (2002) and Karahan-Gül *et al.* (2003). Apart from a more reliable description of the reality, considering the growth on external substrate can help to overcome another described failure of ASM3: ASM3 fails in predicting the maximum growth rate profile on short-term respirometric batch experiments. Krishna and van Loosdrecht (1999) pointed out the presence of a discontinuity on this profile. In other words, the growth rate observed in the feast phase is higher than the one observed in the famine phase. ASM1 correctly describes this observation, because the oxygen consumption is solely related to the growth process, so

both the OUR and the growth rate profiles have the same trend. In contrast, ASM3 predicts a constant growth rate and continuous along the experiment corresponding to the maximum growth rate on storage product. Finally, two different growth rates are predicted when simultaneous growth and storage on external substrate is considered and, hence, the model describes more accurately the reality. Chapter V.B of this thesis develops a new model for biological COD removal including simultaneous growth and storage, which is upgraded with titrimetric measurements. This model is calibrated and validated in chapters V.C and V.D.

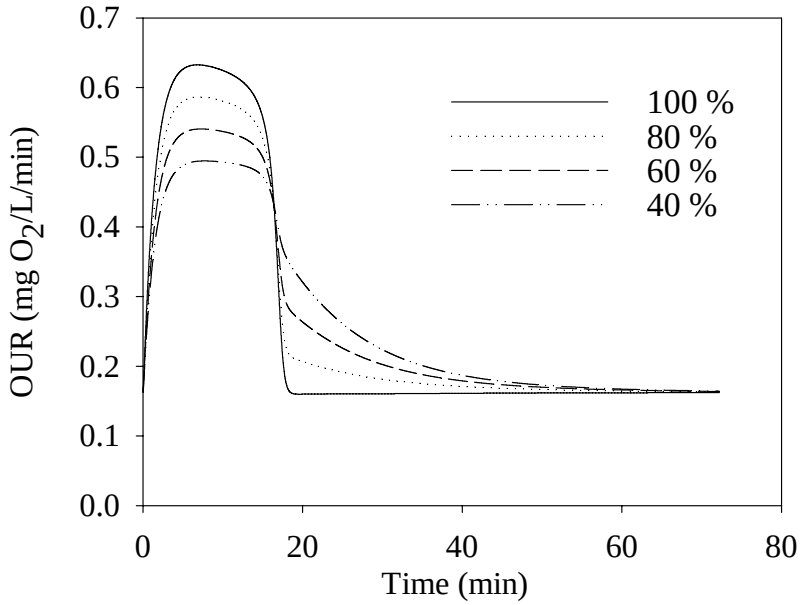


Figure V.A4 Simulation of the effect of a percentage of acetate being directly used for growth

V.A.3.5 PRACTICAL (LOCAL) IDENTIFIABILITY ASM3 USING OUR

An important issue that should be considered in modeling, particularly in view of calibration, is the identifiability of the models. The identifiability of the ASM1 model based on short-term respirometric profiles is already discussed in detail in Dochain and Vanrolleghem (2001). ASM3 introduces the storage process in addition to the growth process for the description of the tail.

The growth process on X_{STO} of ASM3 is a totally different model structure, which contains three parameters: μ_H, K_{STO} and $Y_{H,STO}$ (see Table V.A2). As shown in Figure V.5-upleft, the output sensitivity functions of μ_H and K_{STO} (calculated for test C) are correlated with each other. This implies that both parameters cannot be uniquely identified. The correlation between these two parameters became clear when the shape of the objective function, J, [eq. V.A1] was calculated around an optimum as a function of μ_H and $K_{STO.}$ The shape of objective function (see Fig V.A5-upright) showed a flat valley with a certain direction in the plane (μ_H and K_{STO}). This has often been observed in Monod-type models (e.g. Dochain and Vanrolleghem, 2001). This means that several different combinations of μ_H and K_{STO} can fit the experimental data equally well. This observation was also confirmed when both parameters were considered for parameter estimation. In that case, the parameter estimation error (uncertainty in parameter estimation) of μ_H and K_{STO} increased considerably, up to 300% of relative errors, indicating no reliable estimates for both parameters are possible.

On the other hand, the shape of the objective function as a function of μ_H and $Y_{H,STO}$ depicted in Figure V.A5-downleft does not show linearity in the plane (μ_H and $Y_{H,STO}$). However, the contour plots of the objective function are rather large which indicates that still a high correlation exists between these two parameters.

The same conclusion can be obtained from the plot of the objective function as a function of K_{STO} and $Y_{H,STO}$ (Figure V.A5 – downright). This implies the existence of a severe correlation between K_{STO} and $Y_{GSTO.}$ In this study, K_{STO} was not estimated together with μ_H and Y_{GSTO} and it was fixed to its default value in ASM3 i.e. 1 g COD_{STO} / g COD_X.

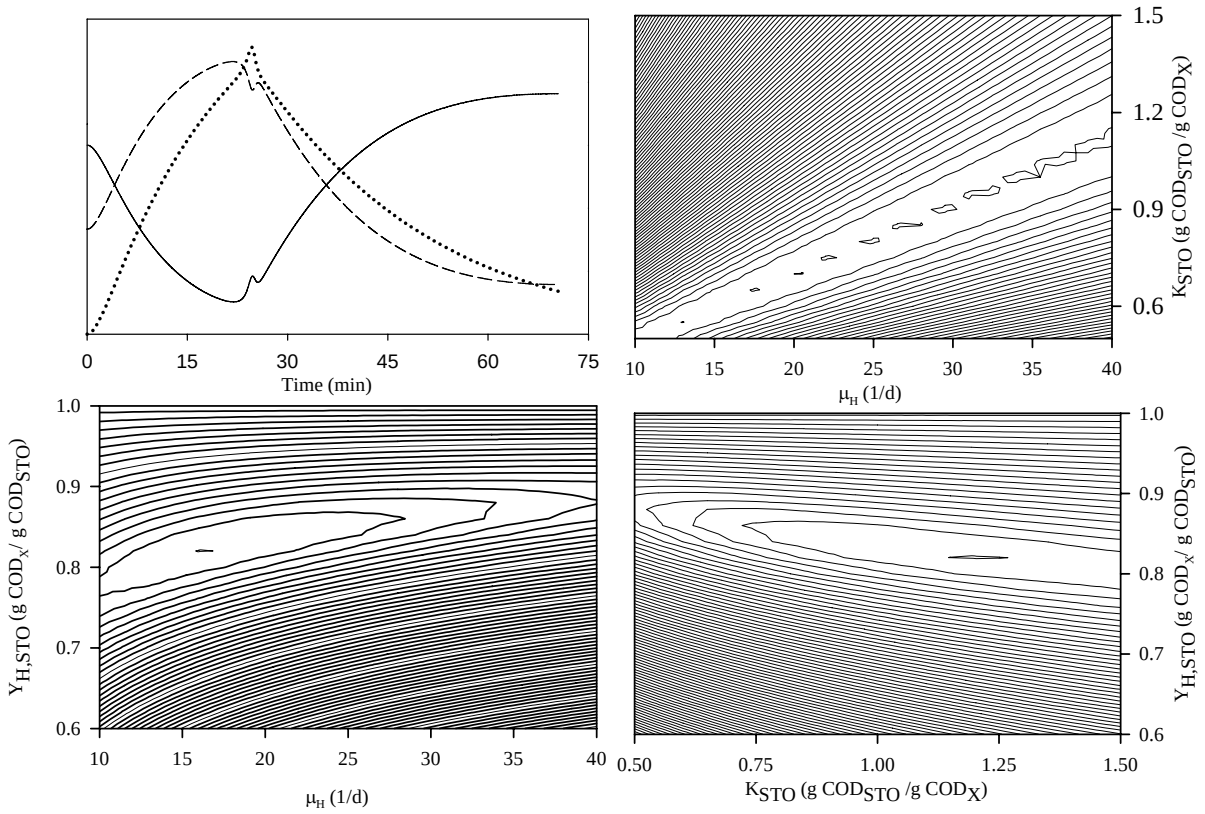


Figure V.A5 Sensitivity functions of μ_H (solid line), K_{STO} (short-dashed line) and $Y_{H,STO}$ (dotted line) (UPLEFT) and correlation of μ_H and K_{STO} (UPRIGHT), μ_H and $Y_{H,STO}$ (DOWNLEFT) and K_{STO} and $Y_{H,STO}$ (DOWNRIGHT).

CHAPTER V.A Conclusions

- ASM3 better describes all the experimental profiles when comparing the sum of squared errors. However, it has to be taken into account that seven parameters are estimated in this model in contrast with ASM1, where only five parameters are estimated.

- In experiments with considerable storage, ASM1 is not able to predict the tail observed due to the internal polymer consumption. In contrast, ASM3 can describe this second tail accurately, but non-mechanistic parameters are obtained.

- The growth yield (Y_H) obtained by fitting ASM1 to the short-term respirometric batch profiles is higher than the default one (0.67) and the storage yield (Y_{STO}) obtained by fitting ASM3 is lower than the default one (0.85). These values agree with the observation of simultaneous storage and growth on external substrate already developed in other works (e.g. van Loosdrecht and Heijnen, 2002). The introduction of this hypothesis would also help to improve the mechanistic meaning of the estimated parameters.

- From a practical identifiability point of view, this study shows the difficulty to obtain reliable values of the parameters related to the ASM3-growth process because μ_H and K_{STO} are not identifiable, and high correlation exists between Y_{GSTO} and μ_H, and Y_{GSTO} and K_{STO}.

- Future model developments should take into account the identifiability issues. Non-identifiable model structures should be avoided to improve the mechanistic meaning of model parameters thereby facilitating model validation tasks.

CHAPTER V.B

Model development for the simultaneous storage and growth processes under aerobic conditions

ABSTRACT

This research aims to further improve the mechanistic modelling of the simultaneous storage and growth processes occurring in activated sludge systems as described in the previous chapter. For this aim, a new model based on an extensive and critical review on previous works is developed. This model is particularly focused in describing the natural conditions of a WWTP (i.e. slowly growing biomass under low and alternating F/M ratio with low PHB content). This chapter includes the elementary mass and reduction balances to obtain the corresponding stoichiometric coefficients and the deduction of the kinetics of each process. This model is designed to be calibrated using both respirometric and titrimetric data, hence, the prediction of the proton production/consumption is also developed for each of the processes.

V.B.1 Motivation of this work

The modelling of activated sludge processes, particularly the biological substrate conversions, has evolved fundamentally in the last two decades from simple growth-based kinetics, ASM1, to more complicated models involving the description of storage phenomena, ASM3 (Henze *et al*., 2000). In the case of biological COD removal, the major driving force behind this modelling trend was the increased understanding of storage polymers to play an essential role as intermediate in the substrate removal processes. ASM3 was one of the first models to address the storage phenomenon. To keep the modelling exercise simple, ASM3 assumes that all readily biodegradable substrate (S_S) is first stored as internal storage products (X_{STO}) before it is used for growth during the famine phase. Being the first attempt to evaluate ASM3 using experimental data, Krishna and van Loosdrecht (1999) had observed that ASM3 failed to model two significant experimental observations:

- The discontinuity in the growth rate of biomass observed experimentally between feast and famine phases
- It required prediction of higher levels of internal storage polymers than measured to fit the oxygen consumption during feast and famine phases.

As described in Chapter V.A, the major cause of this failure was the experimentally observed fact that storage and growth on external substrate occur simultaneously as opposed to the assumption of ASM3 that only storage occurs during the feast phase. Moreover, the ASM3 approach also causes severe practical identifiability problems that resulted in unrealistic and non-mechanistic parameter estimates when using batch OUR data (a common methodology in ASM models calibration). From a mechanistic modelling point of view, it becomes clear that ASM3 should be extended to account for simultaneous storage and growth process (see Chapter V.A).

For this aim, several models in addition to Krishna and van Loosdrecht (1999) have been proposed to improve the mechanistic modelling of simultaneous storage and growth processes in activated sludge systems. However, there is no commonly agreed model yet. The identifiability issues are very important in this case, since once the simultaneous growth and storage theory is admitted, a new model with at least three processes with different rates and yields should be proposed:

1. The substrate uptake for storage purposes (Y_{STO} : g COD_{STO}/g COD_S)
2. The substrate uptake for growth purposes ($Y_{H,S}$: g COD_X/g COD_S)
3. The growth on stored product ($Y_{H,STO}$: g COD_H/g COD_{STO})

Figure V.B1 shows a schematic view of these processes:

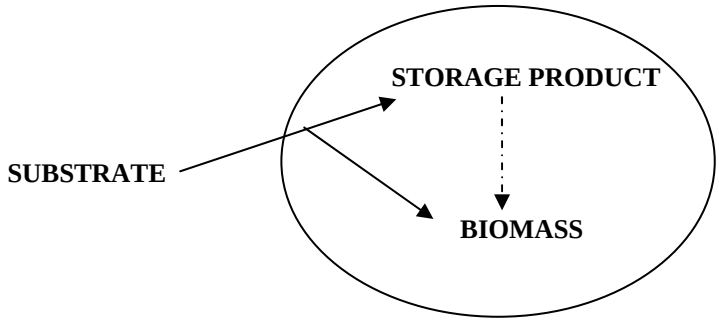


Figure V.B1 Schematic representation of the simultaneous growth and storage process. Feast phase (solid), famine phase (dash-dotted).

The endogenous decay of the biomass and the lysis of the stored product should be added to the model for a reliable description of the system. Hence, the model is rather complex and requires five processes to describe the system. These models are generally calibrated with respirometric batch tests and the OUR measurement alone does not provide enough information for a reliable estimation of all the stoichiometric and kinetic parameters. For this reason, some works include an extra measurement as substrate or storage product in view of solving the identifiability problem among other reasons.

As the model contain ordinary differential equations, at least and initial value of each compound is necessary. This fact may be significant particularly in the case of storage product and heterotrophic biomass fraction because these compounds are not commonly measured. Concerning to the biomass concentration, the biomass growth can be considered negligible in short-term batch experiments. Thus, the real value of biomass concentration is lumped in the measurement of VSS and the estimation of the maximum process rate constants. On the other hand, the case of the storage product is not that simple. The storage product concentration fluctuates strongly along the experiment and its description is important since some of the process kinetics depend on its value. Hence, if the storage product concentration is not measured, it may be complicated to reliably identify the model parameters. In this sense, most of the literature models need to use simplifications and assumptions for a reliable estimation of the parameters. These simplifications are not always acceptable and are only useful for their particular experiments presented.

Acetate is used as substrate in the experiments conducted in this chapter, since it is the most important volatile fatty acid (VFA) present in WWTP. Moreover, it is assumed that most of acetate is stored as poly-β-hydroxybutyrate (PHB). PHB is a well-known storage polymer because it is a very important cell component and its measurement is common in many studies. In addition, many substrates are degraded in the cell with acetyl-COA as intermediate which is also the precursor of PHB formation (van Loosdrecht *et al.*, 1997). PHB also acts a reducing equivalent pool because it is more reduced than acetate (van Aalst-van Leeuwen *et al.*, 1997).

A new for the biological COD removal including simultaneous growth and storage is developed based on a critical review on previous works. This model is focused describing activated sludge processes with slowly growing biomass (i.e. high SRT) under low and alternating F/M ratio in view of modelling full-scale WWTPs. A particular emphasis was given to the kinetic description of the degradation of storage polymers under famine conditions for biomass with low PHB content, as typically found in full-scale WWTPs. To facilitate full-scale application of the model, it can be simply calibrated with batch OUR data (Chapter V.C) and titrimetric data (see Chapter V.D).

V.B.2 Review on previous biological COD removal models

V.B.2.1 THE RESEARCH LINE OF THE DELFT UNIVERSITY OF TECHNOLOGY (TUD)

The group of the TUD is probably one of the top research groups in the field of storage polymers in bioprocesses and they have been working on this topic since the 1990s. In one of their first works about the storage process, van Loosdrecht *et al.* (1997) indicated that microorganisms capable to store substrate for a posterior use have a strong competitive advantage over microorganisms without this capacity. This advantage is emphasized when they are subjected to alternating feast and famine conditions. The reason is that the formation of storage polymers allows the organism to maintain a balanced metabolism in front of sudden changes in the external substrate concentration. When subjected to limited substrate conditions, the substrate uptake machinery is working at its maximum rate. If the external substrate concentration suddenly increases, a high amount of substrate is taken up which cannot be directly converted in cells (i.e. direct growth) and part of this substrate is stored. Hence, bacteria capable of balancing their growth rate independently of the external substrate concentration will be favoured under feast-famine conditions. In addition, they asserted that the relation between oxygen and substrate consumed was very low when compared to pure cultures indicating the presence of the storage process. From their experience, a value of 0.5 g COD/g COD for aerobic growth yield could be used for a wide range of bacteria and substrates.

Van Aalst-van Leeuven *et al.* (1997) developed a metabolic model for the PHB production and consumption for *Paracoccus pantotrophus* under alternating substrate conditions. This model developed for a pure culture pretended to be also appropriate to mixed culture systems, where a similar behaviour was predicted. They proposed a model where simultaneous growth and storage was considered since acetate was degraded to acetyl-COA which could be either directly used for growth or stored as PHB. In absence of substrate the microorganisms utilised the internal stored PHB for the acetyl-COA synthesis necessary for growth. One of the big accomplishments of this model was to solve the identifiability issue of the three necessary yields by linking their value to a unique parameter: the P/O ratio (δ), i.e. the mol of ATP for mol $NADH_2$ used in the oxidative phosphorylation [eq. V.B1a-c].

$$Y_{H,S} = \frac{4\cdot\delta - 2}{4.2\cdot\delta + 4.32} \qquad Y_{STO} = \frac{4\cdot\delta - 2}{4.5\cdot\delta} \qquad Y_{H,STO} = \frac{4.5\cdot\delta - 0.5}{4.2\cdot\delta + 4.32} \qquad \text{(V.B1a-c)}$$

They found a value of δ = 1.84 mol ATP/mol $NADH_2$ for this pure culture which implied the next yield values (transformed in COD units): $Y_{H,S}$= 0.47 g COD_X/g COD_S, Y_{STO} = 0.73 g COD_{STO}/g COD_S and $Y_{H,STO}$= 0.60 g COD_X/g COD_{STO}. The storage/growth pathway ($Y_{STO}\cdot Y_{H,STO}$) produced only 7 % less biomass than the direct growth process, so storing for posterior growth seemed an efficient pathway. From a kinetic point of view, van Aalst-van Leeuven *et al.* (1997) linked the process rates with one parameter, m_{ATP}, which stands for the maintenance coefficient. The substrate conversion rate (r_S) in the feast phase was calculated as equation V.B2, whereas the rate in biomass growth (r_X) in the famine phase was calculated as equation V.B3.

$$r_S = \frac{1}{Y_{STO}}\cdot r_{STO} + \frac{1}{Y_{H,S}}\cdot r_X + m_S\cdot X \qquad \text{(V.B2)}$$

$$r_X = Y_{H,STO}\cdot(r_{STO} - m_P\cdot X_H) \qquad \text{(V.B3)}$$

where m_S and m_P are maintenance coefficients. These can also be calculated as function of δ and m_{ATP} [eqs. V.B4a,b]:

$$m_S = \frac{m_{ATP}}{2\delta - 1} \;,\; m_P = \frac{m_{ATP}}{2.25\delta - 0.25} \qquad \text{(V.B4a,b)}$$

They observed that the acetate uptake and PHB degradation rates in both the feast and famine phases were highly influenced by the PHB content in the cell. They decided to choose the simplest form of kinetic equations which was equation V.B5a for the feast phase and equation V.B5b for the famine phase. These kinetic expressions will be discussed in the model development section.

$$r_P = r_P^{MAX} \cdot \left(1 - \frac{f_{PHB}}{f_{PHB}^{MAX}}\right);\; r_P = K_P \cdot \left(f_{PHB} - f_{PHB}^{MIN}\right) \qquad \text{(V.B5a,b)}$$

where f_{PHB} = ratio of PHB in the cell (i.e. X_{PHB}/X_H) (g COD_{STO}/g COD_X)
f_{PHB}^{MAX} = maximum ratio of PHB in the cell (i.e. X_{PHB}/X_H) (g COD_{STO}/g COD_X)
f_{PHB}^{MIN} = minimum ratio of PHB in the cell (i.e. X_{PHB}/X_H) (g COD_{STO}/g COD_X)
K_P = degradation constant (g COD_X/d/g COD_{STO})
r_P = product consumption rate (1/d)

Later, Krishna and van Loosdrecht (1999) studied the simultaneous acetate storage and growth in activated sludge cultures. They utilised OUR, acetate and PHB measurements for the process analysis and proposed an ASM-like model considering simultaneous growth and storage. Their model was hand-calibrated and few parameters were modified with respect to the ASM3 model (Henze *et al.*, 2000), except for the yields which were assumed to be the same as the ones of van Aalst-van Leeuven *et al.* (1997) for *Paracoccus pantotrophus*. In spite of using the δ value from the pure culture, the modelled profiles described reasonably well the trends of the measured compounds. Particularly, the discontinuity observed in the evolution of the growth rate value was successfully described. From a kinetic point of view, they substituted the kinetic expressions of van Aalst-van Leeuven *et al.* (1997) shown in equations V.B5 for the Monod factors typical of ASM-like models and added an inhibition factor of acetate on the growth on stored product process. They assumed that when acetate was present in the media, the biomass did not grow from stored product.

Beun *et al.* (2000, 2001) extended the model of van Aalst-van Leeuven *et al.* (1997) to activated sludge cultures. They used the stoichiometric relations of V.B1 and found a value of δ between 1.6 and 2 mol ATP/mol $NADH_2$. However, it is not clear how they calculate these values for a mixed activated sludge culture. Beavis and Lehninger (1986) stated that the determination of the intrinsic or mechanistic P/O ratio of oxidative phosphorylation is difficult because of the unknown magnitude of leak fluxes. They suggested that the mechanistic P/O ratio is 1.75 for succinate oxidation and 2.75 for NADH oxidation. In any case, δ should not exceed the value of 3 due to thermodynamic considerations (Lehninger *et al.*, 1993).

Beun *et al.* (2000, 2001) performed experiments at very different SRT (from 3.8 and 19.8 d). In the feast phase, they found that the substrate uptake rate decreased 40 % with this increase in the SRT whereas PHB was produced in a more or less constant rate. With respect to the famine phase, they observed that the degradation rate of PHB was more dependent on the PHB content in the cell than on the SRT of the system. When modelling the PHB degradation rate, they used neither equation V.B5 nor classical Monod kinetics, but applied the multiple-order kinetics equation [eq. V.B6] as done in Grau *et al.* (1975), Murnleitner *et al.* (1997) or Dircks *et al.*, (2001). They found different order values (n) and maximum rate (k) for each of their experiments.

$$\frac{df_{PHB}}{dt} = -k \cdot (f_{PHB})^n \qquad \text{(V.B6)}$$

Finally, van Loosdrecht and Heijnen (2002) summarised all the prior knowledge of their research line in a very interesting paper. Once they accepted that the substrate was used for both growth and storage purposes they meditated which was the preferential use of the substrate and concluded that the effect of SRT on the growth rate is the key parameter for understanding process.

Studies of the storage phenomenon with pure cultures at low SRTs (i.e. high growth rate) showed that the accumulation rate of storage products was the difference between the maximum substrate uptake rate and the substrate uptake rate required for growth. Hence negligible storage was observed in a culture operated at a growth rate close to its maximum substrate uptake rate (i.e chemostat).

On the other hand, most WWTPs are typically operated at high SRTs (i.e. low growth rate) to achieve complete biological nutrient removal resulting in biomass with a rather low average growth rate (~1/SRT). In addition, biomass tends to be subjected to alternating feast and famine conditions. They argued that under these conditions the competition on substrate uptake rate is much more important than the competition on growth rate. Hence, bacteria maximise its substrate uptake rate (q_{MAX}) when external substrate is present whereas they grow in a balanced way. Then, these bacteria are able to grow on the storage product when external substrate is depleted. These bacteria do not optimise the growth rate when substrate is present. Consequently, the maximum substrate flux into the cell exceeds the amount used for the maximum growth of the biomass and the difference is diverted to formation of the storage polymers.

In this range of high SRT, q_{MAX} is slightly changing with SRT (Beun e*t al.*, 2001) while μ_{MAX} is strongly affected by SRT variation. Hence, the ratio of PHB produced per acetate taken up can be considered constant (around 0.67 g COD_{PHB}/ g COD_{AC}) since the storage becomes the dominant process under these conditions.

At this point, van Loosdrecht and Heijnen (2002) proposed a model to describe the variation of the growth rate on SRT based on the level of protein synthesising system (RNAs and anabolic enzymes) in the cells of the organisms. This model included the variation of the growth rate on the enzyme content, though this enzyme content could not be modelled and served only for calculation purposes. The conceptual background of this model is discussed in model development section.

V.B.2.2 OTHER RELATED WORKS

Dircks *et al.* (2001) modelled the PHB accumulation and production using biomass from pilot-plant and full-scale WWTPs. As the sludge used was withdrawn from an UCT-EBPR plant, the role of glycogen was also considered. From a COD balance, they observed that a small amount of biomass was produced in the feast period (see Table V.B1). They also observed that the amount of PHB produced was similar to the value 0.67 value predicted by van Loosdrecht and Heijnen (2002).

They measured the dependence of the PHB degradation rate on the PHB content in the cell for a wide range of PHB content values in order to obtain a reliable model. They also used the multiple-order kinetics model [eq. V.B6] and observed that optimum the reaction order (n) increased (from 0.7 to 2.2) as the PHB content in the cell decreased (i.e. SRT decreased) in a range from almost 0 to 0.12 g COD_{PHB}/g COD_X. The results obtained in this work did not allow formulating a general common kinetic factor for the dependence of the famine process rate on the PHB fraction because of the variety observed on the estimated k and n values.

Table V.B1 Yield values for the feast phase calculated by Dircks *et al.* (2001)

SRT	$Y_{H,S}$ (g COD_X/g COD_{AC})	$Y_{H,STO}$ (g COD_X/g COD_{STO})
4 d	0.07	0.67
21 d	0.05	0.65

Third *et al.* (2003) studied the effect of oxygen concentration on the feast phase processes and performed several batch experiments measuring PHB, acetate and OUR at different oxygen setpoints (0.1 to 2 mg O_2/L). They observed that both the substrate and the storage rate were strongly affected by oxygen limitation. In addition, they observed that the ratio of PHB produced to acetate uptake increased to 0.67 g COD_{PHB}/ g COD_{AC} as the DO set point decreased. Moreover, they observed that the biomass growth on low DO values was practically negligible since all the available ATP produced was used for substrate uptake. They modelled their experimental results with a metabolic model based on van Aalst-van Leeuven *et al.* (1997) and concluded that a combination of low DO and long SRT ensured maximum PHB storage with the minimum growth.

Beccari *et al.* (2002) and Karahan-Gül *et al.* (2003) did a parallel and similar work. They extended the classical ASM3 model (Henze *et al.*, 2000) so that it could describe the simultaneous growth and storage process. They both added a reaction where biomass could grow directly on external substrate (i.e. the growth process of ASM1). Karahan-Gül *et al.* (2003) also added an inhibition factor on the growth on stored product to avoid simultaneous growth on external substrate and stored product. They both calibrated the model with OUR and substrate values. However, they did not take into account identifiability issues and worked with three independent yield values. Neither acetate nor PHB was measured and only OUR was used for parameter estimation. Moreover, their respirometric experimental profiles did not show enough storage for a reliable parameter estimation.

On the other hand, Beccari *et al.* (2002) found that ASM3 could favourably describe only their OUR measurements. They had to include the growth on external substrate to describe their PHB measurements. However, this inclusion did not agree yet with their ammonia measurements and they included a new process called “internal accumulation” or “biosorption”. Although they had 4 measurements (acetate, OUR, PHB and ammonia) the model complexity they introduced (4 rates with four yields) and their parameter estimation procedure had probably some identifiability problems.

Finally, Hanada *et al.* (2002) adopted the idea that some microorganisms maximise their storage rate and some maximise their growth rate and hypothesised that there existed two different kinds of biomass in their dual biomass model. They tried to quantify the amount of each biomass present by means of PHA and biomass staining with Nile Blue A and DAPI, respectively. They considered that the fraction of “ASM3 bacteria” in each WWTP is strongly influenced by the operational plant conditions. They tested 5 different WWTP and found different PHB/Ac values (from 0.4-0.5 in C-mol basis).

V.B.2.3 UPGRADING CARBON REMOVAL MODELS WITH TITRIMETRY: THE WORK OF GERNAEY *et al.*

One of the objectives of this chapter is to upgrade the simultaneous storage and growth model with titrimetric measurements. Hence, new compounds such as protons or carbon dioxide should be included in the model. Titrimetric measurements are very helpful in monitoring biological processes since they provide high quality information about the process with simple equipment (pH control loop). These measurements have already been used to monitor biological aerobic and anoxic carbon removal processes (Gernaey *et al.*, 2002a,b; Pratt *et al.*, 2004; Sin, 2004). Based on these works, each of the processes included in the model was extended to integrate titrimetric measurements (see below).

Gernaey *et al.*, (2002a, b), from the BIOMATH research group (Ghent University), laid the foundations of the link between titrimetric measurements and carbon removal processes in view of modelling. They identified which were the processes which mostly influenced the pH (i.e. substrate and ammonia uptake and CO_2 stripping) and studied the practical identifiability of the model with combined respirometric-titrimetric measurements. The inclusion of titrimetry enabled the identification of one extra parameter combination and, hence, a more reliable biomass growth yield estimation than with only respirometric measurements. They used a simple model where biomass directly grew on external substrate, without considering the storage compounds, which certainly play an important role on the carbon removal process. Actually, their experiments were conducted with biomass which showed low storing capacity.

If substrate is considered as a weak acid (e.g. acetate or propionate), it is mostly in the dissociated form (A^-) in a pH range close to neutrality. However, it is accepted that substrate is uptaken in a neutral way, so each mole of dissociated acid taken up needs a mole of proton from the medium. The dissociation factor (ϖ) indicates the amount of acid dissociated in the medium This value can be calculated from the acidity constant and the pH [eq. V.B7]:

$$K_a = \frac{[A^-].[H^+]}{[HA]} \qquad [HA] = 10^{pKa-pH}.[A^-]$$
$$\varpi = \frac{[A^-]}{[A^-]+[HA]} = \frac{[A^-]}{[A^-]+10^{pKa-pH}.[A^-]} = \frac{1}{1+10^{pKa-pH}} = \frac{10^{-pKa}}{10^{-pKa}+10^{-pH}} \qquad \text{(V.B7)}$$

Hence, equation V.B8 shows that ϖ moles of protons would be consumed for each mol of substrate taken up (r_{VFA} = VFA uptake rate) and acid dosage would be required to maintain the pH at a certain setpoint.

$$HPR = -\varpi \cdot r_{VFA} \qquad \text{(V.B8)}$$

By the same token, ammonia is also taken up in a neutral way (NH_3). The pK_{NH4} is 9.25 and hence, in a pH range close to neutrality most of the ammonia is in NH_4^+ form. Likewise the previous equation, the dissociation (α) factor of NH_4^+ has to be taken into account. It can be deduced with pH and pK_{NH4} [eq. V.B9].

$$\alpha = \frac{[NH_3]}{[NH_3]+[NH_4^+]} = \frac{10^{-pK_{NH4}}}{10^{-pH}+10^{-pK_{NH4}}} \qquad \text{(V.B9)}$$

Hence, equation V.B10 shows that α moles of protons would be released for each mol of ammonia taken up (r_{NH4} = NH_4 uptake rate) and acid dosage would be required to maintain the pH at a certain setpoint.

$$HPR = -\alpha \cdot r_{NH4} \qquad \text{(V.B10)}$$

Finally, the carbon stripping effect on pH had to be taken into account. Acid dosage is required to balance the effect of the stripped carbon dioxide. As they used short-term experiments to calibrate their model, they assumed that under these conditions the CO_2 stripping rate could be considered constant. Hence, they predicted a constant acid addition which they called background proton production rate (BPPR). This may be the most controversial issue in their model, since this assumption is only valid under certain conditions, i.e. low oxygen transfer efficiency, high pH and short-term experiments (Pratt *et al.*, 2004; Sin, 2004).

V.B.2.4 CONSIDERATIONS ABOUT GERNAEY'S APPROACH TO THE CO_2 STRIPPING PROCESS

In order to check the consistency of Gernaey's approach, the evolution of CO_2 in an aerated pH controlled system was modelled. In a continuously aerated system, the evolution of S_{CO2} can be described by equation V.B11:

$$\left(\frac{dS_{CO2}}{dt}\right) = k_L a_{CO2} \cdot (S_{CO2}^{*} - S_{CO2}) + CPR_{END} \qquad \text{(V.B11)}$$

where CPR_{END} = Carbon Production Rate under endogenous conditions (mol CO_2/L/min)
$k_L a_{CO2}$ = global carbon dioxide transfer coefficient (1/min)
S_{CO2} = dissolved carbon dioxide (mol CO_2/L)
S_{CO2}^{*} = saturation concentration of carbon dioxide (mol CO_2/L)

Under steady state conditions, equation V.B11 can be simplified to equation V.B12:

$$\left(\frac{dS_{CO2}}{dt}\right) = k_L a_{CO2} \cdot (S_{CO2}^{SS} - S_{CO2}) \qquad \text{(V.B12)}$$
$$\text{where } S_{CO2}^{SS} = S_{CO2}^{*} + CPR_{END} / k_L a_{CO2}$$

When the pH of the system is controlled, there is an acid dosage which is influencing the equilibrium values. Equations V.B13-.B15 show the evolution of dissolved carbon dioxide, bicarbonate (S_{HCO3}) and protons (S_{HP}) due to the stripping, the equilibrium displacement (EQ) and the acid dosage (AD).

$$\left(\frac{dS_{CO2}}{dt}\right) = k_L a_{CO2} \cdot (S_{CO2}^{SS} - S_{CO2}) + \left(\frac{dS_{CO2}}{dt}\right)_{EQ} \qquad \text{(V.B13)}$$

$$\frac{dS_{HCO3}}{dt} = -\left(\frac{dS_{CO2}}{dt}\right)_{EQ} = k_L a_{CO2} \cdot (S_{CO2}^{SS} - S_{CO2}) - \left(\frac{dS_{CO2}}{dt}\right) \qquad \text{(V.B14)}$$

$$\frac{dS_{HP}}{dt} = \left(\frac{dS_{HP}}{dt}\right)_{AD} - \left(\frac{dS_{CO2}}{dt}\right)_{EQ} = 0 \qquad \text{(V.B15)}$$

The ratio between bicarbonate and carbon dioxide is fixed with the equilibrium constant [eqs. V.B16,b and V.B17.

$$K_1 = \frac{S_{HP} \cdot S_{HCO3}}{S_{CO2}} \rightarrow S_{HCO3} = S_{CO2} \cdot K_1 \cdot 10^{pH} \qquad \text{(V.B16 a,b)}$$

$$\frac{dS_{HCO3}}{dt} = \frac{dS_{CO2}}{dt} \cdot 10^{pH-pK1} \qquad \text{(V.B17)}$$

Hence, equation V.B18 arises from V.B14 and V.B17:

$$\left(\frac{dS_{CO2}}{dt}\right) \cdot \left(1 + 10^{pH-pK1}\right) = k_L a_{CO2} \cdot (S_{CO2}^{EQ} - S_{CO2}) \qquad \text{(V.B18)}$$

This ordinary differential equation is like $dx/(a-x)=b \cdot dt$ and has an analytical solution which corresponds to V.B19:

$$S_{CO2}(t) = S_{CO2}^{SS} + (S_{CO2}(0) - S_{CO2}^{SS}) \cdot e^{-\frac{k_L a_{CO2}}{1+10^{pH-pK1}} \cdot t} \qquad \text{(V.B19)}$$

Figures V.B2 depicts the evolution of S_{CO2} according to the previous equation with default parameters (Table V.B2).

Table V.B2 Default parameters to model S_{CO2} evolution

Parameter	Value	Parameter	Value
S_{CO2}^{SS}	0.05 mM	pH	7.5
$S_{CO2}(0)$	1 mM	pK_1	6.36
k_La_{CO2}	0.3 1/min		

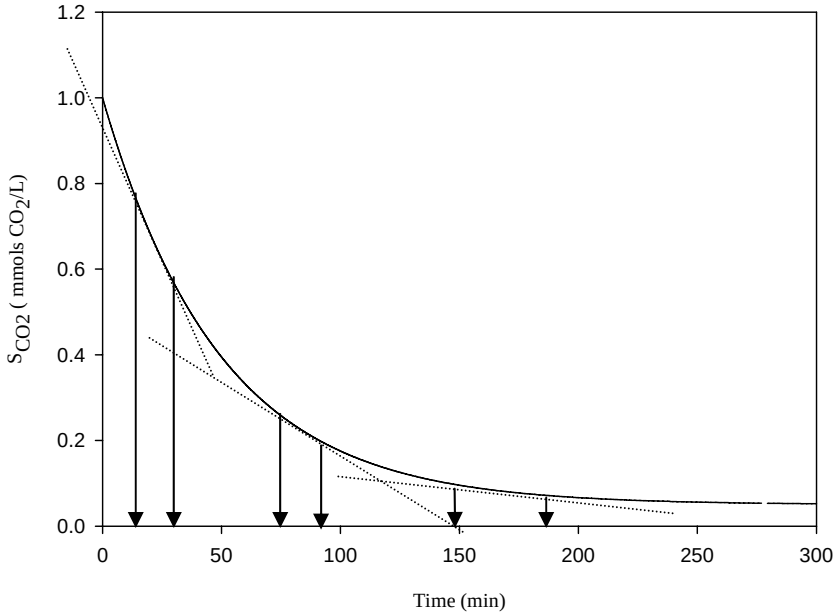


Figure V.B2 Simulation of the S_{CO2} profile in an aerated bioreactor with k_La = 0.3 1/min

As can be observed, Gernaey's approach of a constant BPPR is not appropriate if all the experiment is considered since the slope of the S_{CO2} decrease is variable. However, for short-term experiments (between arrows) it could be a fair simplification. In addition, it can be observed that the lower the S_{CO2} is the longer the time this simplification can be accepted (Figure V.B2). In addition, if the k_La_{CO2} of the system was set to a lower value (i.e. k_La_{CO2} = 0.1 1/min) the decrease would be slower and the assumption of constant BPPR could be sustained longer (Figure V.B3).

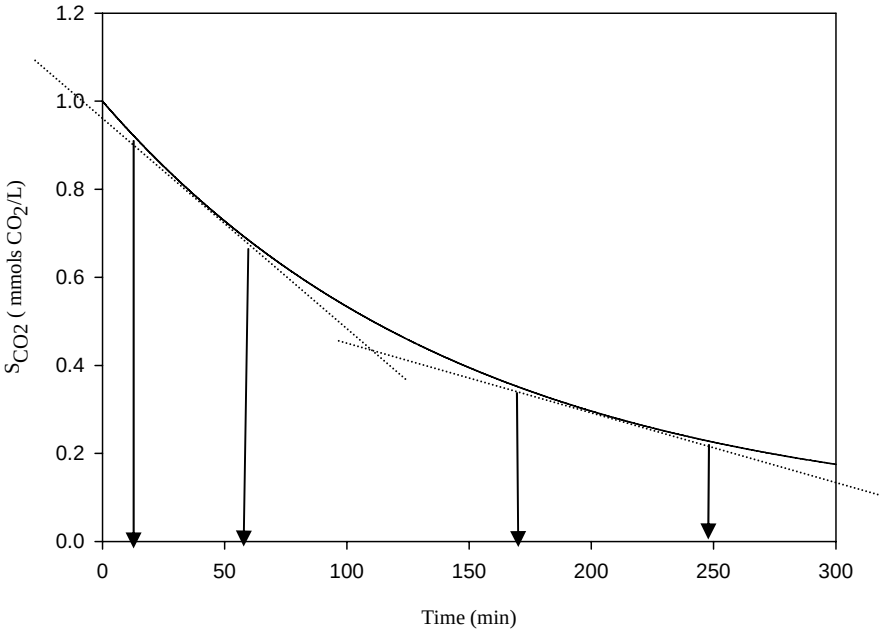


Figure V.B3 Simulation of the S_{CO2} profile in an aerated bioreactor with k_La = 0.1 1/min

V.B.2.5 EXTENDING THE CO_2 STRIPPING APPROACH

Later on, Sin (2004) from the same research group extended the Gernaey's approach by integrating the physical and chemical equilibriums involved in the dynamic CO_2 system: the carbonic acid – carbonate equilibrium, aeration, carbon stripping processes. This extension was also developed in parallel by the research group of the University of Brisbane (Pratt *et al.*, 2004). Using their sensor called TOGA with off-line gas measurements, they integrated the carbon transfer rate (CTR) measurement to the OUR and HPR measurements in view of model calibration and succeeded in closing the carbon balance. Although they accepted that simultaneous growth and storage should be included in the model, they did not adopt this assumption properly, since they model only included substrate uptake for growth and for storage. They omitted the growth on stored substrate characteristic from the famine phase though admitted that PHB was oxidised after acetate depletion. Hence, their model could not describe the classical storage tail observed in respirometric batch experiments once the substrate is depleted. Again, the fact of working with biomass with low storing capacity allowed omitting the growth on storage product process of the model.

With respect to identifiability, they used two independent yields which were not practically identifiable according to their experimental data. They linked $Y_{H,S}$ and Y_{STO} with a parameter λ, which represented the proportionality in the kinetics of both processes. However, infinite combinations of yields satisfied the values of λ found. They found that the substrate storage rate was 3-4 times faster than the growth rate on external substrate, which is in agreement with many works. It indicates that in the feast phase bacteria optimise the substrate uptake versus the growth.

Next, the whole dynamic CO_2 system description developed in the works described above is shown. By including these processes in the model, it is possible to integrate the HPR and CER as a measured variable.

A. CARBONIC ACID –CARBONATE EQUILIBRIA

The carbonic acid-carbonate stoichiometry is depicted in equation V.B20:

$$CO_2(g) + H_2O \Leftrightarrow H_2CO_3 \Leftrightarrow H^+ + HCO_3^- \Leftrightarrow 2H^+ + CO_3^{2-} \qquad \text{(V.B20)}$$

where pK_1 = 6.36 and pK_2 = 10.35 for pure water (Edwards, 1978; Spérandio and Paul, 1997 or Cai and Wang, 1998).

The acid carbonic form is rather unstable and most of the carbon is on CO_2 form. The ratio of CO_2 to H_2CO_3 is 99.76% a 0.24% at 25 ºC (Musvoto *et al.,* 2000a). The second deprotonation of the carbonic acid was neglected because the common working pHs were close to neutrality. Hence, at these pH levels, most of the total inorganic carbon is on bicarbonate form. Two different equilibriums were considered taking into account the reaction between CO_2 and hydroxyl ions (Sperándio and Paul, 1997).

$$\begin{aligned} CO_2 + H_2O &\Leftrightarrow H^+ + HCO_3^- \\ CO_2 + OH^- &\Leftrightarrow HCO_3^- \end{aligned} \qquad \text{(V.B21)}$$

The evolution of these compounds can be predicted using the equilibrium kinetics as described with the set of equations V.B22-V.B26:

$$\frac{dS_{CO_2}}{dt} = -r_1 - r_2 = -\left(kS_{CO_2} - k_{-1}S_{HP}\cdot S_{HCO3}\right) - \left(k_2 S_{CO_2} S_{HP} - k_{-2} S_{HCO3}\right) \qquad \text{(V.B22)}$$

$$K_1 = \frac{S_{HP} \cdot S_{HCO3}}{S_{CO_2}} \quad K_2 = \frac{S_{HCO3}}{S_{OH} \cdot S_{CO_2}} \qquad \text{(V.B23a,b)}$$

$$K_1/K_2 = Kw; \quad k_{-1} = k_1/K_1 \quad k_{-2} = k_2/K_2 \qquad \text{(V.B24a,b,c)}$$

$$\frac{dS_{CO2}}{dt} = -\left(k_1 S_{CO2} - \frac{k_1}{10^{-pK1}} 10^{-pH} S_{HCO3}\right) - \left(k_2 S_{CO2} 10^{pH-14} - \frac{k_2}{10^{-pK2}} S_{HCO3}\right) \qquad \text{(V.B25)}$$

$$\frac{dS_{CO2}}{dt} = \left(\frac{k_1}{10^{-pK1}} 10^{-pH} + \frac{k_2}{10^{-pK2}}\right) S_{HCO3} - \left(k_1 + k_2 10^{pH-14}\right) S_{CO2} \qquad \text{(V.B26)}$$

B.CARBON DIOXIDE STRIPPING AND AERATION

These processes describe the carbon dioxide and oxygen transfer from the liquid to the gas phase due to the physical equilibrium between the two phases. These gas/liquid mass transfer equilibriums are considered to be limited by the liquid side transfer and their kinetics are generally based on a driving force and one efficiency coefficient [eq. V.B27 and V.B28]. In the experiments developed in this thesis, S_{CO2} was generally higher than S_{CO2}*, hence, CTR was negative and CO_2 was stripped. On the contrary, oxygen was transferred to the gas phase to the liquid phase.

$$r_{CTR} = k_L a_{CO2} \cdot (S_{CO2}^* - S_{CO2}) \qquad \text{(V.B27)}$$

$$r_{OTR} = k_L a_{O2} \cdot (S_{O2}^* - S_{O2}) \qquad \text{(V.B28)}$$

V.B.3 Model development: Feast phase

As discussed above, modelling simultaneous storage and growth involves two distinct but complementary phases: feast and famine. Under feast conditions, two modelling approaches have been employed: traditional ASM and metabolic approaches. The first approach, based on a traditional ASM-type model structure, requires three distinctive yield coefficients independent from each other (i.e. storage, direct growth on external substrate and growth on internal storage products). Examples of this approach are Beccari *et al*. (2002), Karahan-Gul *et al*. (2003), Carucci *et al*. (2001) or Pratt *et al.* (2004).

On the other hand, a second approach is based on the metabolic model of van Aalst-van Leeuwen *et al.* (1997) for pure cultures described above (equations V.B1).Examples of this approach are the works of Beun *et al.* (2000), Beun *et al*. (2001) and van Loosdrecht and Heijnen (2002). This approach has been used in this study since the correlation of the three yields appeared to be closer to the reality than three independent yields. Moreover, this approach makes it possible to restrict the calibration to the estimation of only one parameter (δ) instead of three yield coefficients. The latter is a very important aspect of modelling in view of reliability and full-scale applicability of models for WWTPs (Henze *et al.*, 2000).

V.B.3.1 AEROBIC GROWTH OF HETEROTROPHS ($C_iH_jO_kN_l$) ON SUBSTRATE ($C_xH_yO_z$)

A. STOICHIOMETRY

The process stoichiometry (in molar basis) is described in equation V.B29. The stoichiometry of the direct aerobic growth on substrate in view of modelling with titrimetric measurements has already been developed in the literature (Pratt *et al.*, 2003, Sin *et al.*, 2004).

$$a\ C_xH_yO_z + b\ O_2 + c\ NH_3 \rightarrow C_iH_jO_kN_l + d\ CO_2 + e\ H_2O \qquad \text{(V.B29)}$$

As this expression was developed in C-mol basis, it can be assumed x and i = 1.The mass and degree of reduction balances of the process corresponded to the set of equations V.B30 (a-d):

$$C)\ a = 1 + d \qquad \text{(V.B30a)}$$
$$H)\ a\cdot y + 3c = j + 2\cdot e \qquad \text{(V.B30b)}$$
$$N)\ c = l \qquad \text{(V.B30c)}$$
$$\text{degree of reduction})\ a\cdot(\gamma_S) + b\cdot(-4) + c\cdot(0) = \gamma_X + d\cdot(0) + e\cdot(0) \qquad \text{(V.B30d)}$$

where γ_S is the degree of reduction of the substrate and γ_X is the degree of reduction of the biomass.

The balance of the degree of reduction was developed with NH_3 as N-compound reference (Roels 1983, Heijnen, 1999). Once the substrate and the biomass compositions were known, there were 5 unknown variables (a, b, c, d and e) and only 4 equations, which made the system unsolvable, unless an extra restriction or equation was obtained. This degree of freedom was the biomass growth yield on external substrate ($Y_{H,S,C}$), where the sub index C stands for molar basis.

Hence, $a = 1/Y_{H,S,C}$ where $Y_{H,S,C}$ (moles C_X/moles C_S).

Solving the balances [eqs. V.B30], equation V.B29 became V.B31:

$$\frac{1}{Y_{H,S,C}}CH_yO_z + \frac{\frac{1}{Y_{H,S,C}}\gamma_S - \gamma_X}{4}O_2 + l\cdot NH_3 \rightarrow CH_jO_kN_l + \left(\frac{1}{Y_{H,S,C}} - 1\right)CO_2 + \frac{\frac{1}{Y_{H,S,C}}\cdot y + 3\cdot l - j}{2}H_2O \qquad \text{(V.B31)}$$

The stoichiometric coefficients of substrate, oxygen, ammonia and biomass could be converted to COD and N weight units [eq. V.B32] assuming that:

1. one mole of electron corresponded to 8g of COD (likewise to an oxygen mol with 32g and 4 electron moles).
2. the degree of reduction of one compound is the number of electron moles present in a C mol of compound. Hence, 1 mol of $C_S = 8\cdot\gamma_S$ g COD.

$$\frac{8\cdot\gamma_S}{Y_{H,S,C}}CH_yO_z + 32\frac{\frac{1}{Y_{H,S,C}}\gamma_S - \gamma_X}{4}O_2 + 14\cdot l\,NH_3 \rightarrow$$
$$8\cdot\gamma_X\,CH_jO_kN_l + \left(\frac{1}{Y_{H,S,C}} - 1\right)CO_2 + \frac{\frac{1}{Y_{H,S,C}}\cdot y + 3\cdot l - j}{2}H_2O \qquad \text{(V.B32)}$$

Equation V.B33 arises dividing all the stoichiometric coefficients by $8\cdot\gamma_X$:

$$\frac{\gamma_S}{\gamma_X\cdot Y_{H,S,C}}CH_yO_z + \left(\frac{\gamma_S}{\gamma_X\cdot Y_{H,S,C}} - 1\right)O_2 + \frac{14\cdot l}{8\gamma_X}NH_3 \rightarrow$$
$$CH_jO_kN_l + \frac{1}{8\gamma_X}\left(\frac{1}{Y_{H,S,C}} - 1\right)CO_2 + \frac{1}{8\gamma_X}\left(\frac{\frac{1}{Y_{H,S,C}}\cdot y + 3\cdot l - j}{2}\right)H_2O \qquad \text{(V.B33)}$$

Equation V.B34 shows the conversion of the biomass growth yield in molar units ($Y_{H,S,C}$) to the yield in weight units ($Y_{H,S}$). $Y_{H,S,C}$ was substituted for $Y_{H,S}$ in equation V.B34. The percentage of nitrogen in biomass in weight basis (i_{NB}) could be calculated as $14 \cdot l / 8\gamma x$.

$$Y_{H,S,C} = \frac{8\gamma_S}{8\gamma_X} Y_{H,S} \text{ where } Y_{H,S} \text{ was g } COD_X/\text{g } COD_S \qquad \text{(V.B34)}$$

$$\frac{1}{Y_{H,S}} CH_yO_z + \left(\frac{1}{Y_{H,S}} - 1\right) O_2 + i_{NB} NH_3 \rightarrow CH_jO_kN_l + \frac{1}{8\gamma_X}\left(\frac{\gamma_X}{\gamma_S \cdot Y_{H,S}} - 1\right) CO_2 + \frac{\frac{\gamma_X}{\gamma_S \cdot Y_{H,S}} \cdot y + 3 \cdot l - j}{16\gamma_X} H_2O \qquad \text{(V.B35)}$$

At this point the Respiratory Quotient (RQ) of this process can be obtained [eq. V.B36].

$$RQ = \frac{CPR[molCO_2]}{OUR[molO_2]} = \frac{\frac{1}{8\gamma_X}\left(\frac{\gamma_X}{\gamma_S \cdot Y_{H,S}} - 1\right)}{\frac{1}{32}\left(\frac{1}{Y_{H,S}} - 1\right)} = 4 \cdot \frac{\left(\frac{1}{\gamma_S} - \frac{Y_{H,S}}{\gamma_X}\right)}{(1 - Y_{H,s})} \qquad \text{(V.B36)}$$

For example: for acetate as substrate (γ_S=4) and biomass as $CH_{1.4}N_{0.2}O_{0.4}$ (γ_X=4) this process in molar basis would become (V.B37).

$$\frac{1}{Y_{H,S,C}} CH_2O + \left(\frac{1}{Y_{H,S,C}} - 1\right) O_2 + 0.2NH_3 \rightarrow CH_{1.4}O_{0.4}N_{0.2} + \left(\frac{1}{Y_{H,S,C}} - 1\right) CO_2 + \left(\frac{2}{Y_{H,S,C}} - 0.1\right) H_2O \qquad \text{(V.B37)}$$

Then RQ = 1 mol CO_2/ mol O_2

Equation V.B38 shows that the amount of proton consumed in this process should be calculated taking into account two different effects: the substrate and the ammonia uptake [eqs. V.B8 and V.B10].

$$HPR = \left(\frac{\alpha \cdot i_{NB}}{14} - \frac{\varpi}{i_{CSS} \cdot 8 \cdot \gamma_S \cdot Y_{H,S}}\right) r_X \qquad \text{(V.B38)}$$

where i_{CSS} represents the number of carbon atoms in the substrate.
r_X = biomass growth rate (mg COD_X /L/min)

B. KINETICS

The process kinetics are shown on equation V.B39. As can be observed, the maximum growth rate on external substrate ($\mu_{MAX,S}$) is corrected with the corresponding substrate limitations factors (Monod kinetics). Moreover, a first order delay is added to describe the start-up phenomenon as described in Chapter I.2.

$$r_1 = \left(1 - e^{t/\tau}\right) \mu_{MAX,S} \cdot \frac{S_S}{K_S + S_S} \cdot \frac{S_o}{K_O + S_O} \cdot \frac{S_{NH}}{K_{NH} + S_{NH}} \cdot X_H \qquad \text{(V.B39)}$$

where K_{NH} = ammonia affinity constant (mg N/L)
K_O= oxygen affinity constant (mg O_2/L)
K_S = external substrate affinity constant (mg COD_S /L)
S_{NH} = ammonia concentration (mg N/L)
S_O = oxygen concentration (mg O_2/L)
S_S = external substrate concentration(mg COD_S/L)
τ = time constant of the start-up phase (min)

V.B.3.2 FORMATION OF STORAGE PRODUCTS ($C_oH_pO_q$) FROM SUBSTRATE ($C_xH_yO_z$)

A. STOICHIOMETRY

The process stoichiometry (in molar basis) is described in equation V.B40:

$$a\ C_xH_yO_z + b\ O_2 \rightarrow C_oH_pO_q + c\ CO_2 + d\ H_2O \qquad \text{(V.B40)}$$

As this expression was developed in C-mol basis, it can be assumed x and o = 1.The mass and degree of reduction balances of the process corresponded to the set of equations V.B41 (a-c):

$$\text{C)}\ a = 1+c \qquad \text{(V.B41a)}$$
$$\text{H)}\ a\cdot y = p+2\cdot d \qquad \text{(V.B41b)}$$
$$\text{degree of reduction)}\ a\cdot(\gamma_S) + b\cdot(-4) = \gamma_{STO} + c\cdot(0) + d\cdot(0) \qquad \text{(V.B41c)}$$

Once the substrate and storage product compositions were known, there were 4 unknown variables (a, b, c and d) and only 3 equations, which made the system unsolvable, unless an extra restriction or equation was obtained. This degree of freedom was the storage yield ($Y_{STO,C}$), where the sub index C stands for molar basis.

Hence, $a = 1/Y_{STO,C}$ where $Y_{STO,C}$ (moles C_{STO}/moles C_S).

Solving the balances [eqs. V.B41], equation V.B40 became V.B42:

$$\frac{1}{Y_{STO,C}} CH_yO_z + \frac{\frac{1}{Y_{STO,C}}\gamma_S - \gamma_{STO}}{4} O_2 \rightarrow CH_pO_q + \left(\frac{1}{Y_{STO,C}} - 1\right) CO_2 + \left(\frac{\frac{1}{Y_{STO,C}} y - p}{2}\right) H_2O \qquad \text{(V.B42)}$$

The stoichiometric coefficients of substrate, oxygen, and stored product could be converted to COD weight units [eq. V.B43] likewise equation V.B32:

$$\frac{8\cdot\gamma_S}{Y_{STO,C}} CH_yO_z - 32\frac{\frac{1}{Y_{STO,C}}\gamma_S - \gamma_{STO}}{4} O_2 \rightarrow 8\cdot\gamma_{STO}\, CH_pO_q + \left(\frac{1}{Y_{STO,C}} - 1\right) CO_2 + \left(\frac{\frac{1}{Y_{STO,C}} y - p}{2}\right) H_2O \qquad \text{(V.B43)}$$

Equation V.B44 arises dividing all the stoichiometric coefficients by $8\cdot\gamma_{STO}$:

$$\frac{\gamma_S}{\gamma_{STO}\cdot Y_{STO,C}} CH_yO_z - \frac{\frac{1}{Y_{STO,C}}\gamma_S - \gamma_{STO}}{\gamma_{STO}} O_2 \rightarrow$$
$$CH_pO_q + \frac{1}{8\gamma_{STO}}\left(\frac{1}{Y_{STO,C}} - 1\right) CO_2 + \frac{1}{8\gamma_{STO}}\left(\frac{\frac{1}{Y_{STO,C}} y - p}{2}\right) H_2O \qquad \text{(V.B44)}$$

Equation V.B45 shows the conversion of the storage yield in molar units ($Y_{STO,C}$) to the yield in weight units (Y_{STO}). $Y_{STO,C}$ was substituted for Y_{STO} in equation V.B46.

$$Y_{STO,C} = \frac{8\gamma_S}{8\gamma_{STO}} Y_{STO} \text{ where } Y_{STO} \text{ was g } COD_{STO}/\text{g } COD_S \quad \text{(V.B45)}$$

$$\frac{1}{Y_{STO}} CH_yO_z + \left(1 - \frac{1}{Y_{STO}}\right) O_2 \rightarrow CH_pO_q + \frac{1}{8\gamma_{STO}} \left(\frac{\gamma_{STO}}{Y_{STO} \cdot \gamma_S} - 1\right) CO_2 + \left(\frac{\frac{y}{\gamma_S Y_{STO,}} - \frac{p}{\gamma_{STO}}}{16}\right) H_2O \quad \text{(V.B46)}$$

At this point the Respiratory Quotient (RQ) of this reaction can be obtained [eq. V.B47].

$$RQ = \frac{CPR}{OUR} = 32 \frac{\frac{1}{8\gamma_{STO}} \left(\frac{\gamma_{STO}}{Y_{STO} \cdot \gamma_S} - 1\right)}{\left(1 - \frac{1}{Y_{STO}}\right)} = 4 \cdot \frac{\left(\frac{1}{\gamma_S} - \frac{Y_{STO}}{\gamma_{STO}}\right)}{(1 - Y_{STO})} \quad \text{(V.B47)}$$

For example: for acetate as substrate (γ_S=4) and its storage as PHB (γ_{STO}=4.5) this process in molar basis would become:

$$\frac{1}{Y_{STO,C}} CH_2O + \left(\frac{1}{Y_{STO,C}} - 1\right) O_2 \rightarrow CH_{1.5}O_{0.5} + \left(\frac{1}{Y_{STO,C}} - 1.125\right) CO_2 + \left(\frac{1}{Y_{STO,C}} - 0.75\right) H_2O \quad \text{(V.B48)}$$

If Y_{STO} = 0.85 g COD_{STO}/g COD_S (from ASM3; Henze *et al.*, 2000) then RQ = 1.63 .

Equation V.B49 shows the amount of proton consumed in this process, which is calculated with the electron balance of the substrate [eq. V.B8].

$$HPR = -\frac{\varpi}{i_{CSS} \cdot 8 \cdot \gamma_S \cdot Y_{STO}} \cdot r_{STO} \quad \text{(V.B49)}$$

where r_{STO} stands for the storage rate (mg COD_{STO}/L/min)

B. KINETICS

The process kinetics are shown on equation V.B50. As can be observed, the maximum storage rate (k_{STO}) is corrected with the corresponding substrate limitations factors (Monod kinetics). Moreover, a first order delay is added to describe the start-up phenomenon as described in Chapter I.2.

$$r_2 = \left(1 - e^{t/\tau}\right) k_{STO} \cdot \frac{S_O}{K_O + K_O} \cdot \frac{S_S}{K_S + K_S} \cdot X_H \quad \text{(V.B50)}$$

V.B.3.3 FEAST PHASE: PROCESS RATES

According to the research line of TUD, the kinetics of the feast phase strongly depend on the evolution of the growth rate with SRT. At low SRTs (i.e. high growth rate) negligible storage is observed since the growth rate is close to its maximum substrate uptake rate. On the other hand, at high SRTs (i.e. low growth rate) the substrate uptake rate of the biomass is higher than the the average growth rate. Consequently, the difference between substrate uptake (q_{MAX}) and substrate used for growth (μ_{MAX}) is diverted to formation of the storage polymers. Under these conditions, van Loosdrecht and Heijnen

(2002) asserted that the q_{MAX} was slightly changing with SRT while μ_{MAX} is strongly affected by SRT variation. Hence, the ratio of PHB produced per acetate taken up can be considered constant since the storage becomes the dominant process under these conditions (around 0.67 g COD_{STO}/g COD_S). This premise has been experimentally confirmed in several works such as Beun *et al.*, (2000); Beun *et al.*, (2001); Dircks *et al.*, (2001) or Third *et al.*, (2003).

From a mathematical point of view, the biological control of substrate flux into the cell can be illustrated using a branch-pipe analogy (see Figure V.B4). In this branch-pipe analogy, the flow F1 stands for the substrate influx into the cell, F3 for the substrate flux diverted to growth and F2 for the substrate flux diverted to storage. Experimental observations with slowly growing systems showed that the ratio of storage products to substrate taken up (F2/F1) could be considered constant. The remaining substrate flux is diverted to growth (F3). This experimental observation can be modelled by considering a ratio controller on the flow F2 that is indicated by valve A. In this way, the flow of F2 can be controlled by fixing its value to a certain fraction of F1, f_{STO} [eq. V.B51] which means that the substrate flow to F3 is also controlled [eq. V.B52].

$$F2 = f_{STO} * F1 \qquad \text{(V.B51)}$$

$$F1\text{-}F2 = F3;\ F3 = (1\text{-}f_{STO}) * F1 \qquad \text{(V.B52)}$$

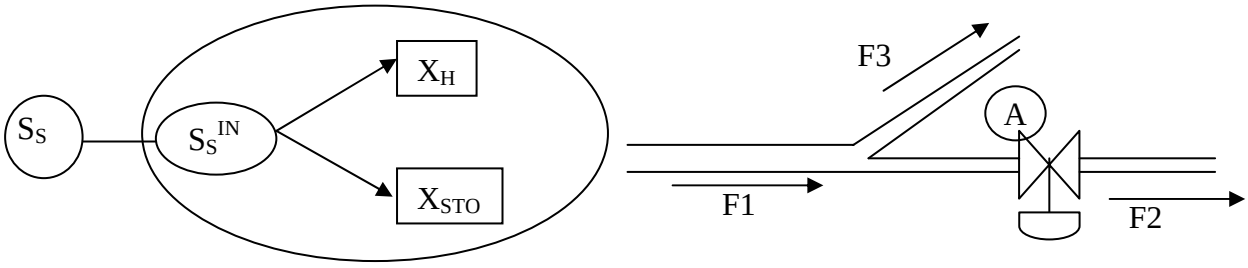


Figure V.B4. Illustration of substrate flux into the cell (LEFT) and branch-pipe analogy for control of substrate flux under feast conditions (RIGHT)

From a *strict* mathematical point of view, it is not important where the control valve is allocated to allow for a good description of the experimental observations. However, biochemically, it has two different meanings depending whether the limiting step is the growth rate or the storage rate. Most probably the reality consists of a mixture of biomass with different strategies, but this is beyond the scope of this study. Again, the mathematical representation of both strategies remains largely the same.

One of the most difficult issues to model is the internal substrate (S_S^{IN}). According to previous metabolic models (van Aalst-van Leeuven *et al.,* 1997), this compound would correspond to acetyl-COA. Including this compound would increase too much the model complexity with new stoichiometric and kinetic parameters. In addition, it is not a compound normally measured. On the other hand, this compound becomes necessary since a unique substrate uptake process must be included. Otherwise, biomass would uptake substrate for growth and for storage directly for the exterior.

The substrate flux under feast conditions can be modelled as follows [eq. V.B53].

$$r_S^{IN} = r_S - \frac{1}{Y_{HS}} r_X - \frac{1}{Y_{STO}} r_{STO} \qquad \text{(V.B53)}$$

where r_S^{IN} is the rate of substrate conversion in terms of mg COD_S/L/min

The internal substrate concentration, S_S^{IN}, can be assumed at steady state (i.e. $r_S^{IN} = 0$).

The steady state assumption of S_S^{IN} is mostly correct during the feast phase in a pulse experiment, except for two short unsteady-state phases:

- at the time of S_S pulse addition The description of the first unsteady-state part can be lumped into the description of the transient response usually observed in batch experiments (Vanrolleghem *et al.,* 2004)

- just after depletion of S_S respectively. This second unsteady-state phase will be captured by a small change in the substrate affinity constant (K_S). Therefore, for most of the time equation V.B54 is accomplished.

$$r_S = \frac{1}{Y_{H,S}} r_{XH} + \frac{1}{Y_{STO}} r_{STO} \qquad \text{(V.B54)}$$

Equation V.B55 can be translated into the following equality assuming no substrate limitations:

$$q_{MAX} = \frac{\mu_{MAX,S}}{Y_{H,S}} + \frac{k_{STO}}{Y_{STO}} \qquad \text{(V.B55)}$$

Based on experimental observations of the constant ratio of substrate uptake/storage as discussed above in detail, the control is assumed on F2 which means that biomass is maintaining storage rate in constant value [eq. V.B56]:

$$k_{STO} = f_{STO} \cdot q_{MAX} \cdot Y_{STO} \text{ and } \mu_{MAX,S} = (1 - f_{STO}) \cdot q_{MAX} \cdot Y_{H,S} \qquad \text{(V.B56)}$$

where f_{STO} is the fraction of the substrate flux diverted to the storage products (g COD_S/g COD_S). As discussed below, $f_{STO}.Y_{STO}$ should be around 0.6-0.7 for WWTP with high SRT. However this value should not be taken as universal, since it is influenced by many factors such as the alternating feed pattern, the plant SRT… In this way, modelling the rates of simultaneous storage and growth reduces to estimation of three parameters, i.e. f_{STO}, δ and q_{MAX}.

It is important to note that from a metabolic point of view, equation V.B55 should also include a fraction of substrate used for maintenance such as the Herbert-Pirt equation (Beun *et al*., 2000; van Loosdrecht and Heijnen, 2002). However, in traditional ASM models (Henze *et al*., 2000) the maintenance concept of the biomass is already lumped into the endogenous decay coefficient describing many other processes such as death, predation, lysis etc. According to the classical approach, the maintenance of biomass was implicitly included in the endogenous decay coefficient in order to keep the model at a reasonable complexity.

Although based on the same conceptual background, the major difference between the proposed model and the model of van Loosdrecht and Heijnen (2002) is the approach to this internal substrate issue (S_S^{IN}). They did not include this compound in their model and all the substrate taken up was directly stored as PHB. Then, part of this PHB is used for growth. The process kinetics and stoichiometry are developed so that when model was simulated biomass is observed to grow on external substrate.

The model of van Loosdrecht and Heijnen (2002) uses the X_{STO} component to divert the substrate flux to growth and to storage in order to keep the modelling simple (see Figure V.B5). Hence, the direct growth on substrate is modelled with X_{STO} as intermediate although the kinetics of both process result in a fast conversion of substrate to biomass. However, in this way, the interpretation of the mathematical model becomes dangerous since it implies that the biomass grows both in the feast phase and the famine phase using X_{STO} but then in two different ways, with two different yields and two different

growth rates. One of the two processes is less efficient than the other (in energy terms). Hence, the biomass would also choose the most efficient pathway under the same conditions.

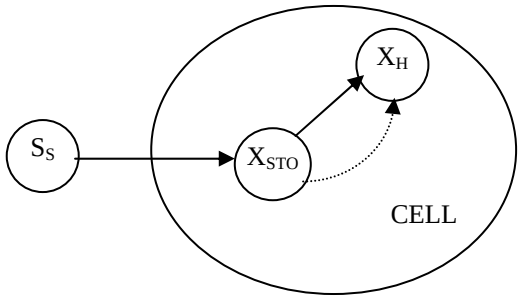


Figure V.B5 Mathematical formulation used in the model of van Loosdrecht and Heijnen (2002) (—— feast phase, ------- famine phase)

V.B.4 Model development: Famine phase

V.B.4.1 AEROBIC GROWTH OF HETEROTROPHS ($C_iH_jO_kN_l$) ON STORAGE PRODUCT ($C_oH_pO_q$)

A. STOICHIOMETRY

The process stoichiometry (in molar basis) is described in equation V.B57:

$$a\ C_oH_pO_q + b\ O_2 + c\ NH_3 \rightarrow C_iH_jO_kN_l + d\ CO_2 + e\ H_2O \qquad \text{(V.B57)}$$

As this expression was developed in C-mol basis, it can be assumed i and o = 1The mass and degree of reduction balances of the process corresponded to the set of equations V.B58 (a-d):

$$C)\ a = 1+d \qquad \text{(V.B58a)}$$
$$H)\ a\cdot p + 3c = j+2\cdot e \qquad \text{(V.B58b)}$$
$$N)\ c = l \qquad \text{(V.B58c)}$$
$$\text{degree of reduction})\ a\cdot(\gamma_{STO}) + b\cdot(-4) + c\ (0) = \gamma_X + d\cdot(0) + e\cdot(0) \qquad \text{(V.B58d)}$$

The balance of the degree of reduction was developed with NH_3 as N-compound reference (Roels 1983, Heijnen, 1999). Once the biomass and storage product compositions were known, there were 5 unknown variables (a, b, c, d and e) and only 4 equations, which made the system unsolvable, unless an extra restriction or equation was obtained. This degree of freedom was the biomass growth yield on stored product ($Y_{H,STO,C}$), where the sub index C stands for molar basis.

Hence, $a = 1/Y_{H,STO,C}$ where $Y_{H,STO,C}$ (moles C_X/moles C_{STO}).

Solving the balances [eqs. V.B58], equation V.B57 became V.B59:

$$\frac{1}{Y_{H,STO,C}} CH_pO_q + \frac{\frac{1}{Y_{H,STO,C}}\gamma_{STO} - \gamma_X}{4} O_2 + l\cdot NH_3 \rightarrow CH_jO_kN_l + \left(\frac{1}{Y_{H,STO,C}} - 1\right) CO_2 + \frac{\frac{1}{Y_{H,STO,C}}p - 3l - j}{2} H_2O \qquad \text{(V.B59)}$$

The stoichiometric coefficients of stored product, oxygen, biomass and ammonia could be converted to COD and N weight units [eq. V.B60] likewise V.B32:

$$\frac{8\gamma_{STO}}{Y_{H,STO,C}} CH_pO_q - 32\cdot \frac{\frac{1}{Y_{H,STO,C}}\gamma_{STO} - \gamma_X}{4} O_2 + 14\cdot l\, NH_3 \rightarrow$$

$$8\cdot\gamma_X\cdot CH_jO_kN_l + \left(\frac{1}{Y_{H,STO,C}} - 1\right) CO_2 + \frac{\frac{1}{Y_{H,STO,C}} p - 3l - j}{2} H_2O \qquad \text{(V.B60)}$$

Equation V.B61 arises dividing all the stoichiometric coefficients by $8\cdot\gamma_S$:

$$\frac{8\gamma_{STO}}{Y_{H,STO,C}}\frac{1}{8\gamma_X} CH_pO_q - \frac{\frac{1}{Y_{H,STO,C}}\gamma_{STO} - \gamma_X}{\gamma_X} O_2 + \frac{14\cdot l}{8\cdot\gamma_X} NH_3 \rightarrow$$

$$CH_jO_kN_l + \frac{1}{8\cdot\gamma_X}\left(\frac{1}{Y_{H,STO,C}} - 1\right) CO_2 + \frac{1}{8\cdot\gamma_X}\left(\frac{\frac{1}{Y_{H,STO,C}} p - 3l - j}{2}\right) H_2O \qquad \text{(V.B61)}$$

Equation V.B62 shows the conversion of the storage yield in molar units ($Y_{H,STO,C}$) to the yield in weight units ($Y_{H,STO}$). $Y_{H,STO,C}$ was substituted for $Y_{H,STO}$ in equation V.B63. The percentage of nitrogen in biomass in weight basis (i_{NB}) could be calculated as $14\cdot l/8\gamma x$.

$$Y_{H,STO,C} = \frac{8\gamma_{STO}}{8\gamma_X} Y_{H,STO} \text{ where } Y_{H,STO} \text{ was g } COD_X/\text{g } COD_{STO} \qquad \text{(V.B62)}$$

$$\frac{1}{Y_{H,STO}} CH_pO_q + \left(\frac{1}{Y_{H,STO}} - 1\right) O_2 + i_{NB}\, NH_3 \rightarrow$$

$$CH_jO_kN_l + \frac{1}{8\cdot\gamma_X}\left(\frac{\gamma_X}{Y_{H,STO}\cdot\gamma_{STO}} - 1\right) CO_2 + \frac{1}{16}\left(\frac{p}{\gamma_{STO}Y_{H,STO}} - \frac{3l - j}{\gamma_X}\right) H_2O \qquad \text{(V.B63)}$$

At this point the Respiratory Quotient of this reaction can be obtained [eq. V.B64]:

$$RQ = \frac{CPR}{OUR} = 32\frac{\frac{1}{8\cdot\gamma_X}\left(\frac{\gamma_X}{Y_{H,STO}\cdot\gamma_{STO}} - 1\right)}{\left(1 - \frac{1}{Y_{H,STO}}\right)} = 4\frac{\left(\frac{1}{\gamma_{STO}} - \frac{Y_{H,STO}}{\gamma_X}\right)}{(1 - Y_{H,STO})} \qquad \text{(V.B64)}$$

For example: for PHB as storage polymer (γ_{STO}=4.5) and $CH_{1.4}O_{0.4}N_{0.2}$ as biomass (γ_X=4.2) this process in molar basis would become [eq. V.B65].

$$\frac{1}{Y_{H,STO,C}} CH_{1.5}O_{0.5} + \left(\frac{1.125}{Y_{H,STO,C}} - 1\right) O_2 + 0.2NH_3 \rightarrow CH_{1.4}O_{0.4}N_{0.2l} + \left(\frac{1}{Y_{H,STO,C}} - 1\right) CO_2 + \left(\frac{0.75}{Y_{H,STO,C}} - 1\right) H_2O \qquad \text{(V.B65)}$$

If $Y_{H,S}$ = 0.63 g COD_X/g COD_S (from ASM3; Henze *et al.*, 2000) then RQ = 0.7. Equation V.B66 shows that the amount of proton required in this process can be calculated due to the ammonia uptake for growth as described in equation V.B10.

$$HPR = \frac{\alpha\cdot i_{NB}}{14} r_X \qquad \text{(V.B66)}$$

V.B.4.2 FAMINE PHASE KINETICS

The typical process of the famine phase is the growth on the storage product with the energy coming from its oxidation. The PHB is degraded for energy obtainment and growth purposes. However, there is no commonly agreed kinetic model yet. Two approaches have been usually employed to describe the kinetics of degradation of X_{STO} under famine conditions: surface saturation-type kinetics and a first-order model.

On the one hand, the surface saturation-type kinetics [eq. V.B67] (Henze *et al*., 2000; Krishna and van Loosdrecht, 1999; Beccari *et al*., 2002; Karahan-Gul *et al*., 2003), e.g. as in ASM3, has been shown to cause severe practical identifiability problems due to its structure, resulting in unrealistic parameter estimates (see Chapter V.A).

$$f(X_{STO}/X_H) = \frac{X_{STO}/X_H}{K_{STO} + X_{STO}/X_H} \qquad \text{(V.B67)}$$

On the other hand, the multiple-order type models [eq. V.B6] (van Aalst-van Leeuwen *et al*., 1997; Beun *et al*., 2000; Beun *et al*., 2001; Dircks *et al*., 2001; van Loosdrecht and Heijnen, 2002) were developed and applied for experimental conditions leading to biomass with high internal storage products content. However, activated sludge from full-scale WWTPs has a much lower fraction of storage products due to the limited availability of external substrate sources as opposed to the studies in well-controlled lab environments. Hence, the multiple-order type kinetics may not be proper for full-scale WWTPs.

In this study several model structures including the above mentioned models have been applied to OUR data obtained with sludge from full-scale WWTPs with low PHB content (results not shown). The following kinetic expression was found to describe the degradation of storage products reasonably well:

$$f(X_{STO}/X_H) = \frac{X_{STO}/X_H}{K_{STO} + X_{STO}/X_H} \cdot \frac{X_{STO}/X_H}{f_{XSTO}^{REG}} \qquad \text{(V.B68)}$$

The first part of this mathematical expression describes the surface-saturation type degradation kinetics of X_{STO} likewise ASM3 (Henze *et al*., 2000). The second part assumes that the degradation of X_{STO} is regulated as function of the storage content of the cell, $f_{XSTO} = X_{STO}/X_H$ likewise Dircks *et al*. (2001). This means that when f_{XSTO} is high, the degradation of X_{STO} is faster, depending on the regulation constant of the cell, f_{XSTO}^{REG}. However, when f_{XSTO} is decreasing and approaching a minimum level in the biomass, the biomass starts to limit the degradation rate of X_{STO}.

This expression was chosen because:

- Dircks *et al*. (2001) explicitly showed that the degradation rate of PHB strongly depended on the PHB content of the cell.
- van Aalst-van Leeuwen *et al*. (1997) hypothesised that biomass always contains a minimum PHB content. This implies that biomass is likely to control the degradation rate of storage products such that a minimum level of storage products can be maintained.
- Experimental observations (particularly OUR from batch experiments) showed that there are at least two phenomena corresponding to a fast and a slow degradation rate of X_{STO} under famine conditions (see below). This fact cannot be described with a classical Monod kinetics.

Equation V.B67 can be rewritten as follows, resulting in a second-order type kinetic expression [eq. V.B69]:

$$f(X_{STO}/X_H) = \frac{\left(X_{STO}/X_H\right)^2}{K_2 + K_1 . X_{STO}/X_H} \qquad \text{(V.B69)}$$

where $K_2 = K_{STO} * f_{XSTO}^{REG}$ and $K_1 = f_{XSTO}^{REG}$.

In this expression, K_2 becomes the affinity of the biomass towards X_{STO}/X_H (g COD_{STO}/g COD_X) and K_1 is nothing but the regulation constant of the biomass as function of X_{STO}/X_H (g COD/g COD). Similar to previous studies (e.g. van Loosdrecht and Heijnen, 2002), the growth rate of biomass on X_{STO} is assumed to occur under strictly famine conditions, i.e. a Monod inhibition function for external substrate is added to the kinetic description of $r_{STO.}$ This is a significant assumption that will be discussed in Chapter V.C.

V.B.5 Model development: Endogenous processes

V.B.5.1 ENDOGENOUS BIOMASS DECAY

A. STOICHIOMETRY

The process stoichiometry (in molar basis) is described in equation V.B70:

$$a\ C_iH_jO_kN_l + b\ O_2 \rightarrow c\ CO_2 + d\ NH_3 + e\ H_2O \qquad \text{(V.B70)}$$

As this expression was developed in C-mol basis, it can be assumed i = 1. The mass and degree of reduction balances of the process corresponded to the set of equations V.B71 (a-d):

$$C)\ a = c \qquad \text{(V.B71a)}$$
$$H)\ a\cdot j = 3d + 2f \qquad \text{(V.B71b)}$$
$$N)\ a\cdot l = d \qquad \text{(V.B71c)}$$
$$\text{degree of reduction})\ a\cdot(\gamma_X) + b\cdot(-4) = c\cdot(0) + d\cdot(0) + e\cdot(0) \qquad \text{(V.B71d)}$$

The balance of the degree of reduction was developed with NH_3 as N-compound reference (Roels 1983, Heijnen, 1999). Once the biomass composition was known, there were 4 unknown variables (a, b, c and d) and 4 equations, which made the system solvable without any extra restriction or equation.

Solving the balances [eqs. V.B71], equation V.B70 became V.B72:

$$CH_jO_kN_l + \frac{\gamma_X}{4} O_2 \rightarrow CO_2 + l\cdot NH_3 + \frac{j - 3\cdot l}{2} H_2O \qquad \text{(V.B72)}$$

The stoichiometric coefficients of biomass, oxygen and ammonia could be converted to COD and N weight units [eq. V.B73] by the same token as equation V.B32.

$$8\gamma_X \cdot CH_jO_kN_l + 32\frac{\gamma_X}{4} O_2 \rightarrow CO_2 + l\cdot NH_3 + \frac{j - 3\cdot l}{2} H_2O \qquad \text{(V.B73)}$$

Equation V.B74 arises dividing all the stoichiometric coefficients into $8 \cdot \gamma_X$:

$$CH_jO_kN_l + O2 \rightarrow \frac{1}{8\gamma_X} CO_2 + \frac{l}{8\gamma_X} NH_3 + \frac{j - 3 \cdot l}{16\gamma_X} H_2O \quad \text{(V.B74)}$$

All the stoichiometric values obtained should be multiplied for $(1\text{-}f_{XI})$ since part of the biomass (f_{XI}) decays into inert fraction. Hence, the real stoichiometric equation would be equation V.B75. The percentage of nitrogen in biomass in weight basis (i_{NB}) could be calculated as $14 \cdot l / 8\gamma x$.

$$(1 - f_{XI}) \cdot CH_jO_kN_l + (1 - f_{XI})O_2 \rightarrow \frac{(1 - f_{XI})}{8\gamma_X} CO_2 + \frac{(1 - f_{XI}) \cdot i_{NB}}{14} NH_3 + (1 - f_{XI}) \frac{j - 3 \cdot l}{16\gamma_X} H_2O \quad \text{(V.B75)}$$

Finally, the amount of proton can be calculated taking into account the ammonia released coming from the biomass and the ammonia concentration in the inert fraction:

$$HPR = \frac{-\alpha \cdot (1 - f_{XI}) i_{NB}}{14} \cdot r_{ENDX} \quad \text{(V.B76)}$$

where r_{ENDX} is the endogenous biomass decay rate (mg COD_X/L/min).

B. KINETICS

The kinetics of this process were assumed from the ASM models (Henze *et al.*, 2000) and consisted of a maximum rate limited by a Monod on oxygen [eq. V.B77]:

$$r_{ENDX} = b_H \frac{S_O}{K_O + S_O} X_H \quad \text{(V.B77)}$$

V.B.5.2 AEROBIC RESPIRATION OF STORAGE PRODUCTS

A. STOICHIOMETRY

The process stoichiometry (in molar basis) is described in equation V.B78:

$$a\ C_oH_pO_q + b\ O_2 \rightarrow c\ CO_2 + d\ H_2O \quad \text{(V.B78)}$$

As this expression was developed in C-mol basis, it can be assumed o = 1.The mass and degree of reduction balances of the process corresponded to the set of equations V.B79 (a-c):

$$C\ )\ a = c \quad \text{(V.B79a)}$$
$$H\ )\ a \cdot p = 2d \quad \text{(V.B79b)}$$
$$\text{degree of reduction })\ a \cdot (\gamma_{STO}) + b \cdot (-4) = c \cdot (0) + d \cdot (0) + f \cdot (0) \quad \text{(V.B79c)}$$

Once the storage product composition was known, there were 3 unknown variables (a, b, and c) and 3 equations, which made the system solvable without any extra restriction or equation.

Solving the balances [eqs. V.B79], equation V.B78 became V.B80:

$$CH_pO_q + \frac{\gamma_{STO}}{4} O_2 \rightarrow CO_2 + \frac{p}{2} H_2O \quad \text{(V.B80)}$$

The stoichiometric coefficients of storage product and oxygen could be converted to COD weight units [eq. V.B81] by the same token as equation V.B32:

$$8\gamma_{STO}CH_pO_q + 32\frac{\gamma_{STO}}{4}O_2 \rightarrow CO_2 + \frac{p}{2}H_2O \qquad \text{(V.B81)}$$

Equation V.B76 arises dividing all the stoichiometric coefficients into $8\cdot\gamma_{STO}$:

$$CH_pO_q + O_2 \rightarrow \frac{1}{8\gamma_{STO}}CO_2 + \frac{p}{16\gamma_{STO}}H_2O \qquad \text{(V.B82)}$$

B. KINETICS

The kinetics of this process were assumed from the ASM models (Henze *et al.*, 2000) and consisted of a maximum rate limited by a Monod on oxygen [eq. V.B83]:

$$r_{ENDSTO} = b_{STO}\frac{S_O}{K_O + S_O}X_{STO} \qquad \text{(V.B83)}$$

V.B.6 Model stoichiometric and kinetic matrix

Table V.B3 Model stoichiometric and kinetic matrix

Processes	S_O	S_S	S_{NH}	S_{HP}	S_{CO2}	S_{HCO3}	X_H	X_I	X_{STO}	**Kinetics**
	g O_2	g COD	g N	mol H^+	mol CO_2	mol HCO_3^-	g COD	g COD	g COD	
Aerobic growth on external substrate	$-\left(\frac{1}{Y_{H,S}}-1\right)$	$\frac{-1}{Y_{H,S}}$	$-i_{NB}$	$-\frac{\varpi}{i_{CSS}\cdot 8\cdot\gamma_S\cdot Y_{H,S}}+\frac{\alpha\cdot i_{NB}}{14}$	$\frac{1}{8\gamma_X}\left(\frac{\gamma_X}{\gamma_S\cdot Y_{H,S}}-1\right)$		1			$\left(1-e^{t/\tau}\right)\mu_{MAX,S}\cdot M_S M_O M_{NH} X_H$
Formation of storage products	$-\left(\frac{1}{Y_{STO}}-1\right)$	$\frac{-1}{Y_{STO}}$		$-\frac{\varpi}{i_{CSS}\cdot 8\cdot\gamma_S\cdot Y_{STO}}$	$\frac{1}{8\gamma_{STO}}\left(\frac{\gamma_{STO}}{Y_{STO}\cdot\gamma_S}-1\right)$				1	$\left(1-e^{t/\tau}\right)k_{STO}\cdot M_O\cdot M_S\cdot X_H$
Aerobic growth on storage products	$-\left(\frac{1}{Y_{H,STO}}-1\right)$		$-i_{NB}$	$\frac{\alpha\cdot i_{NB}}{14}$	$\frac{1}{8\cdot\gamma_X}\left(\frac{\gamma_X}{Y_{H,STO}\cdot\gamma_{STO}}-1\right)$		1		$-\frac{1}{Y_{H,STO}}$	$\mu_{MAX,STO} M_O M_{NH}\frac{\left(X_{STO}/X_H\right)^2}{K_2+X_{STO}/X_H\cdot K_1}\cdot\frac{K_S}{S_S+K_S}X_H$
Endogenous biomass decay	$-(1-f_{XI})$		$i_{NB}\cdot(1-f_{XI})$	$\frac{-\alpha\cdot(1-f_{XI})}{14}i_{NB}$	$\frac{(1-f_{XI})}{8\gamma_X}$		-1	f_{XI}		$b_H M_O X_H$
Aerobic respiration of X_{STO}	-1				$\frac{1}{8\gamma_{STO}}$				-1	$b_{STO} M_O X_{STO}$
Aeration	1									$K_L a\cdot\left(S_O^*-S_O\right)$
CO_2 stripping					1					$K_L a_{CO2}\cdot\left(S_{cO2}^*-S_{cO2}\right)$
Carbonic acid-carbonate equilibrium					1	-1				$\left(\frac{k_1}{10^{-pK1}}10^{-pH}+\frac{k_2}{10^{-pK2}}\right)S_{HCO3}-\left(k_1+k_2 10^{pH-14}\right)S_{CO2}$

$k_{STO}=f_{STO}\cdot q_{MAX}\cdot Y_{STO}$; $\mu_{MAX,S}=(1-f_{STO})\cdot q_{MAX}\cdot Y_{H,S}$;

M stands for a Monod kinetic function (the substrate considered is indicated in the subscript) e.g. $M_S = S_S/(K_S+S_S)$.

The first-order empirical model ($1-e^{t/\tau}$) is used to model the transient response observed in OUR data obtained from batch experiments as in Vanrolleghem *et al.,* 2004 or Guisasola *et al*., 2003.

Chapter V.B Conclusions

- A new model for biological COD removal was developed using mass and reduction balances. The major improvement of this model is to consider the simultaneous growth and storage process. Moreover it includes:

    1. the three yields linked to one unique parameter
    2. a new approach to describe the substrate kinetics in the feast phase
    3. the dynamic CO_2 system to describe titrimetric data
    4. a new factor to describe the PHB degradation kinetics on biomass with low PHB content.

- The simultaneous growth and storage processes were included based on the experimental observations previous works (particularly from TUD). However, these works include models which have been improved in this chapter.

- The utilisation of the simplification of constant BPPR used in Gernaey *et al.*, 2002a, b is constrained to short-term experiments with low gas/liquid mass transfer efficiency.

- The model was upgraded with titrimetric data using the dynamic CO_2 system developed by Pratt *et al.*, (2004) and Sin (2004). Hence, the titrimetric data can be described without time constraints.