Modeling Aerobic Carbon Source Degradation Processes Using Titrimetric Data and Combined Respirometric–Titrimetric Data:

Structural and Practical Identifiability

Krist Gernaey,[1,2*] Britta Petersen,[1,2] Denis Dochain,[3] Peter A. Vanrolleghem[1]

[1]*Biomath Department, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium*
[2]*Department of Civil, Geological and Mining Engineering, École Polytechnique de Montréal, P.O. Box 6079, Station Centre-ville, Montreal, Quebec, Canada H3C 3A7*
[3]*CESAME department, Université Catholique de Louvain, Bâtiment Euler, 4 Ave. Georges Lemaître, B-1348 Louvain-la-Neuve, Belgium*

Received 7 June 2001; accepted 7 March 2002

DOI: 10.1002/bit.10337

Abstract: The structural and practical identifiability of a model for description of respirometric–titrimetric data derived from aerobic batch substrate degradation experiments of a $C_xH_yO_z$ carbon source with activated sludge was evaluated. The model processes needed to describe titrimetric data included substrate uptake, CO_2 production, and NH_3 uptake for biomass growth. The structural identifiability was studied using the Taylor series method and a recently proposed generalization method. It showed that combining respirometric and titrimetric data allows structural identifiability of one extra parameter combination, the biomass yield, Y_H, compared to estimation on separate data sets, on condition that the nitrogen fraction in biomass (i_{XB}) is known. However, data from short-term batch substrate degradation experiments were not sufficiently informative to allow practical identification of all structurally identifiable parameters. Combining respirometry and titrimetry resulted in improvements of parameter confidence intervals compared to estimation on separate respirometric or titrimetric data sets. However, the level of the improvement seems to be substrate dependent: parameter confidence intervals improved considerably more for dextrose than for acetate degradation models. Noteworthy is the finding that the half-saturation substrate concentrations can be different depending on whether they are estimated from respirometric or titrimetric data. Moreover, this difference appears to be dependent on the carbon source considered: for dextrose, titrimetry-based K_S values are higher than respirometry-based values while for acetate the opposite was found. It was hypothesized that this can be explained by the different point in cell metabolism where the proton production or consumption takes place, leading to a corresponding difference in timing between pH effect and oxygen consumption. Finally, the biomass yield Y_H and the nitrogen content of the biomass i_{XB} could be estimated from combined respirometric–titrimetric data obtained with addition of a known amount of carbon source. Y_H can also be estimated from r_O data when the initial substrate concentration $S_S(0)$ is known. The values found correspond to values reported in literature, but, interestingly, also seem able to reflect the occurrence of storage processes when pulses of acetate and dextrose are added. *Biotechnol Bioeng* **79**: 754–767, 2002.

Keywords: aerobic; carbon source; degradation; identifiability; model; respirometry; titration; storage

Correspondence to: Krist Gernaey

**Present address:* CAPEC, Department of Chemical Engineering, Technical University of Denmark, Building 229, DK-2800 Lyngby, Denmark; telephone: +45 45 25 28 00; fax: +45 45 93 29 06; e-mail: kvg@kt.dtu.dk

Contract grant sponsors: Flemish Institute for the Promotion of Scientific-Technological Research in the Industry (IWT, Brussels); Quebec Ministry of Education; NATO

INTRODUCTION

The process studied is the aerobic degradation of a $C_xH_yO_z$-type carbon source by a mixed population of bacteria (activated sludge) during a batch substrate degradation experiment. The data that will be used further on in this paper were collected using a combined respirometric–titrimetric measurement technique (Gernaey et al., 2001, 2002). Respirometry, the measurement and interpretation of the respiration rate of activated sludge, is often applied as a tool to characterize aerobic degradation processes in activated sludge (Henze et al., 1987; Spanjers et al., 1998; Vanrolleghem et al., 1999). The titrimetric method is based on the use of a pH

control unit. Indeed, the pH value of a biological system responds to microbial reactions, and the evolution of the pH of a system thus provides an indication of the status of ongoing biological reactions. The proton consumption or production related to biological reactions can be measured by controlling the pH of the liquid medium at a constant pH setpoint through addition of acid and/or base. In that case, monitoring the acid and/or base consumption rate(s), needed to keep the pH constant, provides the rate of proton formation or consumption related to the biological reactions. This titrimetric technique has already been applied in different studies of biological nitrogen removal processes in activated sludge (Bogaert et al., 1997; Gernaey et al., 1998; 2001; Massone et al., 1996; Ramadori et al., 1980).

For anoxic degradation of a carbon source by activated sludge it has been observed that protons were consumed, and the amount of consumed protons could be linked to the amount of consumed electron acceptor (NO_3^- in this case) (Bogaert et al., 1997; Petersen et al., 2002b). For the nitrification process, i.e., the aerobic conversion of ammonium to nitrate, it has been shown that titrimetric experiments with activated sludge, where the proton production due to nitrification is measured, can yield information about biodegradation kinetics (Gernaey et al., 1998), which is similar to respirometry.

In a previous paper, a model structure to describe pH effects observed during aerobic degradation of a $C_xH_yO_z$-type carbon source was proposed (Gernaey et al., 2002). In this model, the proton consumption or production observed during short-term aerobic batch carbon source degradation experiments was related to (a) the uptake of the carbon source through the cell wall of the bacteria, (b) the release of CO_2 resulting from respiration processes in the liquid phase, and (c) the uptake of ammonium for growth.

A study of the structural identifiability of model parameters prior to practical model application, e.g., in the frame of parameter estimation or model calibration, is very important. The key question of the structural and practical identifiability analysis has, e.g., been formulated by Dochain et al. (1995) as follows: "assume that a certain number of state variables are available for measurements; on the basis of the model structure (structural identifiability) or on the basis of the type and quality of available data (practical identifiability), can we expect to obtain unique values for the model parameters?" It is important to notice the distinction between structural and practical identifiability in this statement. In the study of structural identifiability perfect noise-free and informative data is assumed whereas in practice the data may be less informative, e.g., due to noise corruption or poor excitation. As a result, parameters may be practically unidentifiable although they are structurally identifiable (Holmberg, 1982).

The purpose of this paper is to investigate the structural and practical identifiability properties of the model proposed by Gernaey et al. (2002) for interpretation of respirometric and titrimetric data collected during aerobic degradation of $C_xH_yO_z$-type carbon sources. The motivation for this detailed investigation of the parameter identifiability is that parameters derived from batch substrate degradation experiments with activated sludge are often used to verify whether the parameters used during a wastewater treatment plant model calibration have realistic values. Indeed, import efforts with respect to model calibration have already been done in the framework of activated sludge model calibrations (Kappeler and Gujer, 1992; Petersen et al., 2002a; Spérandio and Paul, 2000). The better the quality of the parameters estimated from the experimental data, the better the possibilities to compare the full-scale activated sludge model parameters with the experimental results. The structural identifiability of the model parameters is studied via the Taylor series expansion method (Pohjanpalo, 1978) and via a generalization method (Petersen et al., 2000). First, it is assumed that only respirometric or titrimetric data is available; second, the case of combined respirometric and titrimetric data is considered. The practical identifiability of the model parameters is evaluated via parameter estimations on data obtained from aerobic degradation experiments with two different carbon sources (acetate and dextrose) presented by Gernaey et al. (2002). The parameter values obtained from titrimetric and respirometric data respectively are compared. Furthermore, the objective function for the parameter estimation as a function of the estimated parameter values (contour plots) will be evaluated for one particular example. Finally, confidence intervals for parameter values estimated on the basis of a single respirometric or titrimetric data set are compared to the confidence intervals for parameter values estimated from combined respirometric–titrimetric data.

MATERIALS AND METHODS

Experimental Methodology

The experimental methodology to monitor aerobic batch substrate degradation processes using combined respirometric–titrimetric measurements has already been described in detail by Gernaey et al. (2001, 2002). A typical data set obtained by adding 1.563 mmol acetate (100 mg COD) to the aeration vessel of the set-up at $t = 0$ during a short-term batch experiment with activated sludge is shown in Fig. 1. Figure 1 shows the oxygen uptake rate data ($r_{O,2}$) that were obtained via respirometry. The calculated $r_{O,2}$ value is about 0.64 mg O_2/L·min during substrate degradation, and decreases to an endogenous respiration ($r_{O,end}$)

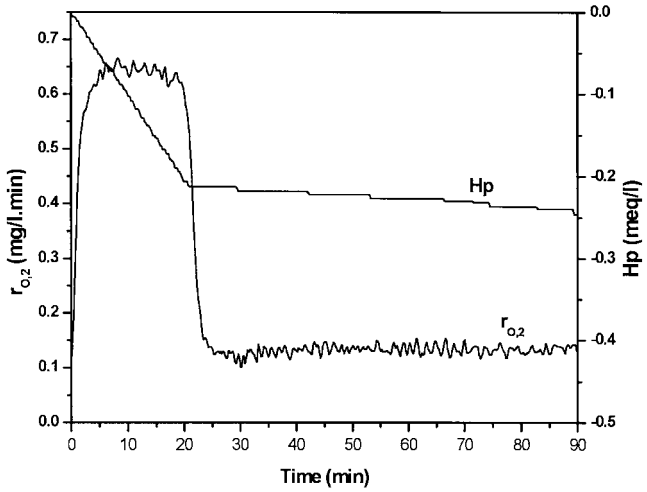

Figure 1. Typical $r_{O,2}$ and proton production data resulting from the respirometric–titrimetric measurements (Gernaey et al., 2002), in this case following the addition of 1.563 mmol acetate (100 mg COD) to 2.54 L of activated sludge at $t = 0$.

level of about 0.12 mg O_2/L·min ($t = 22$ min in Fig. 1). The titrimetric data derived from the same experiment, here expressed in terms of cumulative proton production (Gernaey et al., 2002), are also shown in Figure 1. It is clear that the substrate degradation induces proton consumption. As soon as all substrate is degraded ($t = 22$ min) the slope of the proton production curve becomes considerably less steep. This corresponds to the endogenous respiration phase in the respirometric data.

Parameter Estimation

Parameter estimation on the available data was carried out with the software package WEST (Hemmis NV, Kortrijk, Belgium), using the Simplex optimization algorithm (Nelder and Mead, 1964). Minimization of the sum of squared errors was used as the model fit criterion. Exact contour plots were created via numerous simulations over a grid of different parameter values in the neighborhood of the estimate. Finally, 95% confidence intervals have been calculated similar to the procedure outlined in Petersen et al. (2001).

MODEL AND THEORETICAL ASPECTS

Model

The model applied in this study for interpretation of respirometric, titrimetric, and combined respirometric–titrimetric data is summarized in Table I. Process 1 in this model describes the heterotrophic growth on a readily biodegradable substrate (S_S) and was derived from Gernaey et al. (2002). However, to allow interpretation of the available respirometric and titrimetric data some extra processes need to be included in the model to deal with specific properties of both data types. The calculated total respiration rate (r_O) of the biomass includes both the oxygen uptake caused by degradation of an addition of external substrate (here acetate or dextrose), called the exogenous respiration ($r_{O,ex}$), and endogenous respiration ($r_{O,end}$). The endogenous respiration is considered to be a constantly ongoing process depending on the biomass concentration and is simply described with a first order expression (process 2 in Table I), which is a simplification of the ASM1 death regeneration concept (Henze et al., 1987). The duration of the short-term batch substrate degradation experiments that provided the data sets for this paper is too short to have significant decay of biomass, in contrast to the long-term experiments of Kappeler and Gujer (1992), Keesman et al. (1998), and Spérandio and Paul (2000). It was therefore assumed that the effect of the endogenous processes could be subtracted from the raw data, resulting in H_p and r_O data that contain only information about the substrate degradation related phenomena. The parameter identifiability study will thus also be limited to the exogenous, substrate degradation related phenomena.

In addition, the aeration process can be included in the model to describe the oxygen input in the aeration vessel of the set-up that was used for the experiments, see process 3 in Table I (refer to Gernaey et al. (2001, 2002) for a detailed description of the experimental set-up). This process is necessary when biokinetic parameters are estimated from dissolved oxygen (S_O) data (Petersen et al., 2001). For some respirometers, knowledge of the oxygen transfer coefficient (K_La) is needed to obtain r_O data from S_O measurements. For the ap-

Table I. Schematic overview of the model that was used for the interpretation of respirometric and titrimetric data obtained during aerobic biodegradation of a carbon source.

	Component (i, k)					
Process (j)	1. X_B	2. S_S	3. S_O	4. S_{NH}	5. H_p	Process rate
1. Heterotrophic growth with S_S as substrate	1	$-\frac{1}{Y_H}$	$-\frac{1-Y_H}{Y_H}$	$-i_{XB}$	$-\frac{m}{C\cdot Y_H}+\frac{n\cdot(1-Y_H)\cdot x}{C\cdot Y_H}+\frac{p\cdot i_{XB}}{14}$	$\mu_{max\,H}\frac{S_S}{K_S+S_S}X_B$
2. Endogenous respiration	−1		−1		$\frac{n}{32}$	$b_H\cdot X_B$
3. Aeration			1			$K_La(S_O^0-S_O)$
4. CO_2 stripping					1	b_C

plication described here, the parameters were estimated from oxygen uptake rate (r_O) data only (Gernaey et al., 2001). With the applied respirometric measurement technique, the r_O can readily be obtained via a simple oxygen mass balance. Moreover, it has been shown that the convergence of the objective function for parameter estimation is much faster when r_O data is considered for parameter estimation compared to estimation on S_O data (Petersen et al., 2001).

The aeration induces CO_2 stripping that can, within the current experimental limitations, be assumed to be constant. Thus a constant background addition of acid or base (depending on the choice of the pH setpoint) is needed to keep the pH of the mixed liquor sample at the pH setpoint and is described in the model by a zero-order process with a rate constant b_C (process 4 in Table I). However, the endogenous respiration will also have an influence on the required background addition of either acid or base through production of CO_2. Therefore, it is assumed in the model that one CO_2 molecule is produced per molecule of oxygen that is consumed due to endogenous respiration (respiratory quotient = 1), which explains the presence of the factor $n/32$ in the H_p column of Table I for the endogenous respiration process. In Figure 1 the net background addition of either acid or base caused by CO_2 stripping and production from endogenous respiration corresponds with the part of the proton production curve recorded after about 22 min.

Furthermore, the model includes terms to describe (i) the first-order dissolved oxygen probe dynamics, with first-order constants typically in the order of 0.2–0.5 min (Spanjers and Olsson, 1992), and (ii) the biological start-up phenomena that are typically observed in batch experiments, before the oxygen uptake rate has reached its maximum value. This start-up phase, which typically lasts 0.5–2 min, is hypothesized to be the time needed by a cell to take up fresh substrate and oxidize it, and it can be described with a simple first-order term ($1 - e^{-t/\tau}$) multiplying the maximum specific growth rate (Vanrolleghem et al., 1998).

Structural Identifiability

The basis of the Taylor series approach for the study of structural identifiability is that the vector of measured outputs $\boldsymbol{y}$ and its derivatives with respect to time, typically developed around the initial time $t = 0$, can be assumed to be known and unique. For this method the model underlying the output $\boldsymbol{y}$ can originate from many different types of models (e.g., sets of algebraic equations, state space models, time series models, neural nets, etc.). The Taylor series approximation of the output around a given time t is given in Eq. (1). The successive Taylor derivatives, $\boldsymbol{a}_k(\boldsymbol{p})$, are defined as described in Eq. (2) with $\boldsymbol{p}$ as the parameter vector,

$$\boldsymbol{y}(t,\boldsymbol{p}) = \boldsymbol{y}(t,\boldsymbol{p}) + t \cdot \frac{d}{dt}(\boldsymbol{y}[t,\boldsymbol{p}]) + \frac{t^2}{2!} \cdot \frac{d^2}{dt^2}(\boldsymbol{y}[t,\boldsymbol{p}]) + \cdots + \frac{t^k}{k!} \cdot \frac{d^k}{dt^k}(\mathbf{y}[t,\mathbf{p}]) \quad [k = 0, 1, \ldots, \infty] \tag{1}$$

$$\boldsymbol{a}_k(\boldsymbol{p}) = \lim_{t \to 0} \frac{d^k}{dt^k}(\boldsymbol{y}[t,\boldsymbol{p}]) \quad [k = 0, 1, \ldots, \infty] \tag{2}$$

Thus, the Taylor derivatives, $\boldsymbol{a}_k(\boldsymbol{p})$, are functions of the parameter vector $\boldsymbol{p}$ and the method simply consists of solving a set of algebraic equations (consisting of the Taylor derivatives) with respect to the parameters or combinations thereof. A sufficient condition for the model to be structurally identifiable is that there exists a unique solution for $\boldsymbol{p}$ (Pohjanpalo, 1978). Symbolic manipulations for the study of structural parameter identifiability were carried out with the MAPLE V software package (Waterloo Maple Software).

The identifiable parameter combinations were also evaluated via the generalization method that was proposed by Petersen et al. (2000). The generalization method has been tested on Monod kinetics based models that have been written down in an ASM1-like matrix notation (Table I). The generalization method has not been checked further for any other processes, something that may be done in the future. A set of simple rules was defined for Monod kinetics based models. From these rules the structurally identifiable parameter combinations can be predicted directly based on the knowledge of the model under study (in a matrix representation), on the measured variables and on the substrate. The rules of this generalization are illustrated in Table II. With reference to Table I, v denotes the stoichiometric coefficient, j the process, and i the measured component, while the substrate under study is denoted k. Considering that some components are consumed (e.g., S_O, S_S) whereas others are produced (e.g., X_B, H_p), the absolute values of the stoichiometric coefficients v should be taken. In case two variables are measured for the same process, the parameter combinations for a single measurement listed in Table II still

Table II. Parameter combinations for Monod models considering measurements of one or two components; see text for a detailed explanation of the generalization rules (Petersen et al., 2000).

Measurement of one component	Measurement of two components
$\mu_{max,j}$	$\mu_{max,j}$
$\lvert v_{i,j}\rvert \cdot X_B(0)$	$\lvert v_{i,j}\rvert \cdot X_B(0)$
$\frac{\lvert v_{i,j}\rvert}{\lvert v_{k,j}\rvert} \cdot K_j$	$\frac{\lvert v_{i,j}\rvert}{\lvert v_{k,j}\rvert} \cdot K_j$
$\frac{\lvert v_{i,j}\rvert}{\lvert v_{k,j}\rvert} \cdot S_k(0)$	$\frac{\lvert v_{i,j}\rvert}{\lvert v_{k,j}\rvert} \cdot S_k(0)$
	$\frac{\lvert v_{i(1),j}\rvert}{\lvert v_{i(2),j}\rvert}$

hold, but with the additional identifiable parameter combination $v_{i(1),j}/v_{i(2),j}$, where (1) and (2) indicate the two measured variables, respectively. Thus, the structurally identifiable parameters can be obtained directly without considering the mathematical aspects of, e.g., the Taylor series expansion, thereby significantly reducing the time needed for a structural identifiability study. The generalization of Table II was confirmed with several examples in Petersen (2000) and will be applied and evaluated for the model in this study as well.

RESULTS

Structural Identifiability

The structural identifiability was only investigated for the heterotrophic growth process with S_S as substrate (process 1 in Table I), i.e., the paper focuses only on substrate degradation related phenomena, as explained before. It is clear that if r_O data are considered, the decay rate, b_H, in the first-order expression of the endogenous respiration is structurally identifiable directly from the part of the data either before substrate addition or after the substrate is fully degraded and the rate of process 1 is zero (e.g., the r_O data points for $t > 22$ min in Fig. 1). For a short-term experiment as was used here for data collection the biomass concentration, $X_B(t)$, must then be known to allow practical identifiability of b_H (Petersen, 2000). Also, the net background addition of either acid or base caused by CO_2 stripping and production from endogenous respiration, as explained above, will be identifiable directly from the H_p data set, again using the data points collected at $t > 22$ min for the example given in Figure 1. Thus, the parameter b_C in Table I can be readily identified since the parameter b_H is obtained from the r_O data, the rate of process 1 is zero and the value of the pH depending factor n is known since the experimental pH is known (and constant) (Gernaey et al., 2002).

To study the identifiability of the heterotrophic growth process with S_S as substrate (process 1 in Table I), only the model describing respirometric data was considered (Eqs. (3)–(5)). Note that it is only the exogenous (= substrate degradation related) responses that are considered in Eqs. (3)–(5). It is assumed for short-term batch experiments that the endogenous (= biomass decay related responses) can be removed from the data collected during a short-term batch experiment, as explained before.

$$rX = \mu_{max\,H}\, X_B(t) \frac{S_S(t)}{K_S + S_S(t)} \tag{3}$$

$$rS_S = -\frac{1}{Y_H}\mu_{max\,H}\, X_B(t) \frac{S_S(t)}{K_S + S_S(t)} \tag{4}$$

$$r_O = \frac{(1 - Y_H)}{Y_H}\mu_{max\,H}\, X_B(t) \frac{S_S(t)}{K_S + S_S(t)} \tag{5}$$

This study is very comparable to the one of Dochain et al. (1995), however, with the difference that biomass growth is explicitly considered. Thus, the biomass concentration X_B is not treated as a constant in the development of the Taylor series but is time-varying. Eq. (5) is the first term in the identifiability analysis, comparable to the term $\boldsymbol{y}(t,\boldsymbol{p})$ on the right-hand side of Eq. (1). Eq. (6) is the first derivative of Eq. (5) with respect to time,

$$\frac{dr_O(0)}{dt} = -\frac{\mu_{max\,H}^2 \cdot X_B(0) \cdot S_S(0) \cdot (1 - Y_H) \cdot (X_B(0) \cdot K_S - S_S(0) \cdot Y_H \cdot (S_S(0) + K_S))}{Y_H^2 \cdot (K_S + S_S(0))^3} \tag{6}$$

similar to the derivatives shown in Eq. (2). Only the first Taylor derivative, derived around $t = 0$, is given as an illustration in Eq. (6). The second and third derivatives were also needed for the structural identifiability. Higher derivatives are not shown due to the high complexity of higher-order derivatives.

The result of the study is listed in Table III, column 1, where it can be seen that four parameter combinations are identifiable. In the study of Dochain et al. (1995) the maximum specific growth rate ($\mu_{max\,H}$), the heterotrophic yield coefficient (Y_H) and the biomass concentration present at the beginning of the experiment ($X_B[0]$) are combined in one identifiable parameter combination. Here $\mu_{max\,H}$ can be separated from the parameter combination including the initial biomass concentration since the biomass concentration is now considered to vary with time. The parameter K_S and the initial conditions, $S_S(0)$ and $X_B(0)$, can, however, only be identified in combination with Y_H. Hence, in case the value of Y_H is known, all model parameters can be identified uniquely, i.e., unique values could be given to $\mu_{max\,H}$, K_S, $S_S(0)$, and $X_B(0)$.

Next, the model for titrimetric data is evaluated (Eq. [7]). Again, biomass growth is considered and X_B is therefore a function of time.

$$rHp = \left(-\frac{m}{C \cdot Y_H} + \frac{n \cdot (1 - Y_H) \cdot x}{C \cdot Y_H} + \frac{p \cdot i_{XB}}{14}\right) \times \mu_{max\,H}\, X_B(t) \frac{S_S(t)}{K_S + S_S(t).} \tag{7}$$

Table III. Schematic overview of the structurally identifiable parameter combinations for an aerobic degradation process (assuming biomass growth) with S_S as substrate, depending on the available measurement (r_O, H_p, or a combination of both).

Process	Measurement		
	r_O	H_p	$r_O + H_p$
Heterotrophic growth with S_S as substrate	$\mu_{max\,H}$	$\mu_{max\,H}$	$\mu_{max\,H}$
	$\frac{1-Y_H}{Y_H} X_B(0)$	$\left(-\frac{m}{C \cdot Y_H} + \frac{n(1-Y_H)x}{C \cdot Y_H} + \frac{p \cdot i_{XB}}{14}\right) X_B(0)$	$\frac{1-Y_H}{Y_H} X_B(0)$
	$(1-Y_H)K_S$	$\left(-\frac{m}{C} + \frac{n(1-Y_H)x}{C} + \frac{p \cdot i_{XB} \cdot Y_H}{14}\right) K_S$	$(1-Y_H)K_S$
	$(1-Y_H)S_S(0)$	$\left(-\frac{m}{C} + \frac{n(1-Y_H)x}{C} + \frac{p \cdot i_{XB} \cdot Y_H}{14}\right) S_S$	$(1-Y_H)S_S(0)$
			$\frac{1}{1-Y_H}\left(-\frac{m}{C} + \frac{n(1-Y_H)x}{C} + \frac{p \cdot i_{XB} \cdot Y_H}{14}\right)$

Note that the parameters m, n, and p are all pH-dependent constants linked to chemical equilibria (Gernaey et al., 2002), and the parameters C and x are factors related to the specific known carbon source under study. For acetate ($C_2H_4O_2$), for instance, C is 64 and x is 2, and for dextrose ($C_6H_{12}O_6$), C is 192 and x is 6. The values of these five parameters can therefore be considered known and constant for a given pH value and carbon source (Gernaey et al., 2002).

The four identifiable parameter combinations are listed in Table III, column 2. The structure of these parameter combinations is comparable to the ones for the model of respirometric data, only the stoichiometric coefficients differ (see also Table I).

When combined respirometric–titrimetric measurements are considered in the structural identifiability study, it appeared that a fifth parameter combination was identifiable (column 3, last combination). This parameter combination appears to be just the ratio between the stoichiometric coefficients (see Table I) that link the oxygen uptake and the proton production to the substrate degradation process. From this parameter combination the heterotrophic biomass yield coefficient Y_H can be structurally identified assuming that the parameter expressing the amount of nitrogen incorporated into biomass, i_{XB}, is known. The parameter i_{XB} is a biomass composition related parameter that is assumed to be constant for a bacterial species. Inserting the yield coefficient in the remaining four parameter combinations will thereby result in structural identifiability of $\mu_{max\,H}$, K_S, $X_B(0)$, and $S_S(0)$. Alternatively, if $S_S(0)$ is known (e.g., for a given addition of carbon source), i_{XB} becomes structurally identifiable.

The generalization method of Petersen et al. (2000) was also applied as an alternative to the Taylor series approach. The method readily gave exactly the same structurally identifiable parameter combinations (Table II).

Practical Identifiability

The practical identifiability study also focuses on the substrate degradation related phenomena. In parameter estimation on separate respirometric or titrimetric data sets obtained from short-term batch substrate degradation experiments, it appeared that only three parameter combinations could be identified from the available data as $X_B(0)$ could only be practically identified in a combination with Y_H and $\mu_{max\,H}$. The same was observed when combined respirometric–titrimetric data was considered. The explanation for this discrepancy between structural and practical parameter identifiability is that the biomass grows insufficiently during the short-term batch experiments from which the data was obtained. The initial biomass concentration $X_B(0)$ in combination with Y_H and $\mu_{max\,H}$ would only be practically identifiable as separate parameters when considerable growth can be observed in the data. Thus, the current data (see Fig. 1 for an example) derived from short-term batch experiments are not informative enough to practically identify all the structurally identifiable parameters. Therefore, for the practical parameter identifiability in this study $\mu_{max\,H}$ and $X_B(0)$ had to be considered as one parameter.

Examples of the model fit to combined respirometric–titrimetric data of acetate and dextrose experiments are shown in Figures 2 and 3, respectively. For both substrates the model is able to describe the data well. Note that acetate degradation resulted in proton consumption, while dextrose degradation resulted in proton production (see also Gernaey et al., 2002).

The parameter values obtained for the acetate and dextrose degradation experiments are given in Tables IV and V, and the 95% confidence intervals for the parameters $\mu_{max\,H}\,X_B(0)$ and K_S are listed in Tables VI and VII.

Table IVA gives the results of the estimated parameters based on r_O data for five experiments with addition of acetate. The initial amount of added substrate $S_S(0)$ was known. Consequently, Y_H was structurally identi-

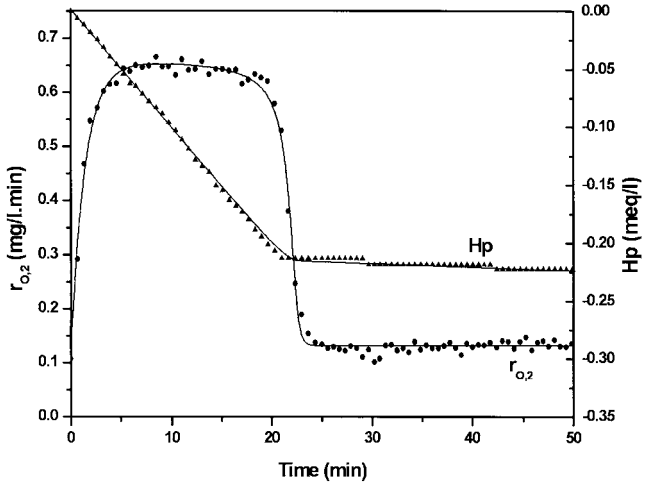

Figure 2. Model fit on combined r_O and H_p data for an example with acetate addition ($\mu_{max\ H}\ X_B$ = 1.50 mg COD/L·min; K_S = 0.62 mg COD/L; Y_H = 0.74; $S_{S,1}(0)$ = 49.72 mg COD/L, i_{XB} = 0.051 mg N/mg biomass COD). Note that only each 3rd data point is shown, to keep the figure from becoming too overloaded.

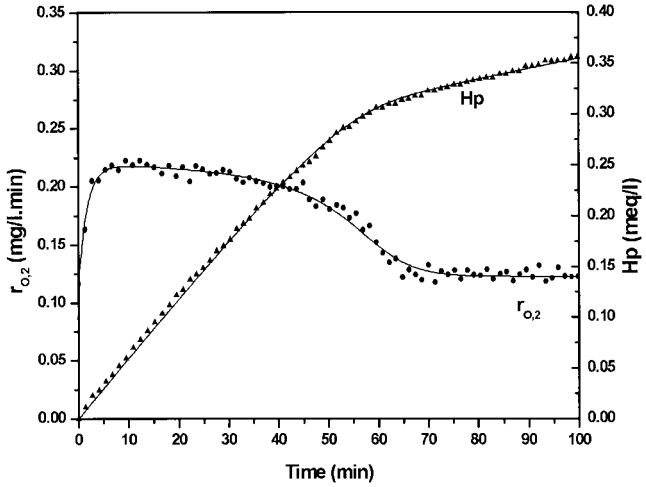

Figure 3. Model fit on combined r_O and H_p data for an example with dextrose addition ($\mu_{max\ H}\ X_B$ = 0.74 mg COD/L·min; K_S = 3.80 mg COD/L; Y_H =0.87; $S_{S,1}(0)$ = 47.62 mg COD/L, i_{XB} = 0.043 mg N/mg biomass COD). Note that only each 6th data point is shown, to prevent the figure from becoming too overloaded.

fiable together with $\mu_{max\ H}\ X_B(0)$ and K_S (see Table III). It is clear that the variation in estimated parameter values between the five experiments is minor, especially for $\mu_{max\ H}\ X_B(0)$ and Y_H, whereas the coefficient of variance (CV, equal to the ratio between the standard deviation and the mean) (= 9.3%) is higher for the K_S estimates. The results of the parameter estimation based on H_p data for the same five experiments are listed in Table IVB. Here, too, the initial substrate addition $S_S(0)$ was considered to be known. Furthermore, it was de-

Table IV. Results of parameter estimations for 5 experiments in which different amounts of acetate were added to activated sludge.*

Exp. No.	$S_S(0)$ (mg COD/L)	Y_H	$\mu_{max\ H}\ X_B(0)$ (mg COD/L · min)	K_S (mg COD/L · min)	i_{XB} (mg N/mg biomass COD)
A. Respirometric data					
1	*25.22*	0.75	1.50	0.72	
2	*37.61*	0.75	1.48	0.63	
3	*49.72*	0.74	1.50	0.62	
4	*61.39*	0.73	1.52	0.58	
5	*72.49*	0.73	1.59	0.71	
Avg.		0.74	1.52	0.65	
CV (%)		1.4	2.8	9.3	
B. Titrimetric data					
1	*25.22*	*0.75*	1.54	0.30	0.078
2	*37.67*	*0.75*	1.48	0.10	0.081
3	*49.72*	*0.74*	1.51	0.10	0.055
4	*61.39*	*0.73*	1.60	0.60	0.052
5	*72.49*	*0.73*	1.61	0.17	0.038
Avg.			1.55	0.25	0.061
CV (%)			3.6	82.7	30.1
C. Combined respirometric–titrimetric data					
1	*25.22*	0.75	1.50	0.73	0.080
2	*37.61*	0.75	1.48	0.65	0.078
3	*49.72*	0.74	1.50	0.62	0.051
4	*61.39*	0.73	1.52	0.59	0.049
5	*72.49*	0.74	1.60	0.69	0.038
Avg.		0.74	1.52	0.66	0.059
CV (%)		1.1	3.1	8.5	31.7

*A: Respirometric data (values for $S_S(0)$, in italics, were assumed to be known. Y_H, $\mu_{max\ H}\ X_B(0)$, and K_S were estimated). B: Titrimetric data (values for $S_S(0)$ and Y_H (from Table IVA), in italics, were assumed to be known. $\mu_{max\ H}\ X_B(0)$, K_S, and i_{XB} were estimated). C: Combined respirometric–titrimetric data (values for $S_S(0)$, in italics, were assumed to be known. Y_H, $\mu_{max\ H}\ X_B(0)$, K_S, and i_{XB} were estimated).

Table V. Results of parameter estimations for 3 experiments in which different amounts of dextrose were added to activated sludge.*

Exp. No.	$S_S(0)$(mg COD/L)	Y_H	$\mu_{max\ H}\ X_B(0)$ (mg COD/L · min)	K_S (mg COD/L · min)	i_{XB} (mg N/mg biomass COD)
A. Respirometric data					
1	*23.92*	0.88	0.61	1.59	
2	*47.62*	0.87	0.71	2.95	
3	*70.42*	0.88	0.87	5.60	
Avg.		0.88	0.73	3.38	
CV		0.7	18.0	60.4	
B. Titrimetric data					
1	*23.92*	*0.88*	0.72	3.75	0.042
2	*47.62*	*0.87*	0.81	6.25	0.046
3	*70.42*	*0.88*	0.95	8.13	0.046
Avg.			0.83	6.04	0.045
CV			14.0	36.4	5.2
C. Combined respirometric–titrimetric data					
1	*23.92*	0.89	0.62	1.80	0.041
2	*47.62*	0.87	0.74	3.80	0.043
3	*70.42*	0.88	0.93	7.63	0.047
Avg.		0.88	0.76	4.41	0.044
CV		1.1	20.5	67.2	7.0

*A: Respirometric data (values for $S_S[0]$, in italics, were assumed to be known. Y_H, $\mu_{max\ H}X_B[0]$, and K_S were estimated). B: Titrimetric data (values for $S_S(0)$ and Y_H (from Table VA), in italics, were assumed to be known. $\mu_{max\ H}X_B(0)$, K_S, and i_{XB} were estimated). C: Combined respirometric–titrimetric data (values for $S_S[0]$, in italics, were assumed to be known. Y_H, $\mu_{max\ H}X_B[0]$, K_S, and i_{XB} were estimated).

cided to fix the value of Y_H based on the knowledge derived from the respirometric data, allowing estimation of the combination $\mu_{max\ H}\ X_B(0)$, K_S, and i_{XB} (see Table III). It is again observed that the variation in the values of the parameter combination $\mu_{max\ H}\ X_B(0)$ is minor between the five experiments, whereas a higher variation is observed for K_S resulting in a significantly higher CV (=82.7%) compared to the estimation based on respirometric data. Finally, Table IVC gives the results of the estimation on combined respirometric–titrimetric data sets where Y_H, $\mu_{max\ H}\ X_B(0)$, K_S, and i_{XB} were estimated, with $S_S(0)$ known. Both data sets were given equal weights during the estimation procedure. Here the CV of the K_S estimates is in the same order of magnitude as for the estimation based on respirometric data alone.

The same parameter estimation approach was carried out for three experiments with additions of dextrose (see Table V). The variation of the parameters $\mu_{max\ H}\ X_B(0)$ and K_S (except for separate titrimetric data) is larger, i.e., higher CV values, between the three experiments with dextrose compared to the acetate experiments. On the contrary, the CV for the i_{XB} estimates for dextrose as substrate is lower than those found for acetate.

The practical parameter identifiability was evaluated in more detail for one particular example with acetate

Table VI. 95% confidence intervals for the parameters $\mu_{max\ H}\ X_B(0)$ and K_S (as absolute values, and expressed as percentage of parameter values between brackets) for 5 experiments with acetate.

Exp. No.	$S_S(0)$	Measurement	$\mu_{max\ H}\ X_B(0)$	K_S
1	25.22	r_O	1.50 ± 0.019 (1.2)	0.72 ± 0.087 (12.1)
		H_p	1.54 ± 0.419 (3.8)	0.30 ± 0.419 (141.7)
		$r_O + H_p$	1.50 ± 0.083 (1.1)	0.73 ± 0.083 (11.4)
2	37.61	r_O	1.48 ± 0.012 (0.8)	0.63 ± 0.077 (12.3)
		H_p	1.48 ± 0.044 (3.0)	0.10 ± 0.426 (426.2)
		$r_O + H_p$	1.48 ± 0.012 (0.8)	0.65 ± 0.076 (11.7)
3	49.72	r_O	1.50 ± 0.010 (0.7)	0.62 ± 0.080 (12.9)
		H_p	1.51 ± 0.018 (1.2)	0.10 ± 0.223 (222.4)
		$r_O + H_p$	1.50 ± 0.008 (0.5)	0.62 ± 0.071 (11.4)
4	61.39	r_O	1.52 ± 0.007 (0.5)	0.58 ± 0.063 (10.8)
		H_p	1.60 ± 0.020 (1.3)	0.60 ± 0.306 (50.9)
		$r_O + H_p$	1.52 ± 0.006 (0.4)	0.59 ± 0.058 (9.9)
5	72.49	r_O	1.59 ± 0.005 (0.3)	0.71 ± 0.052 (7.4)
		H_p	1.61 ± 0.014 (0.9)	0.17 ± 0.220 (129.9)
		$r_O + H_p$	1.60 ± 0.004 (0.3)	0.69 ± 0.048 (6.9)

Table VII. 95% confidence intervals for the parameters $\mu_{max\ H}\ X_B(0)$ and K_S (as absolute values and expressed as percentage of parameter values between brackets) for 3 experiments with dextrose.

Exp. No.	$S_S(0)$	Measurement	$\mu_{max\ H}\ X_B(0)$	K_S
1	23.92	r_O	0.61 ± 0.024 (3.9)	1.59 ± 0.378 (23.8)
		H_p	0.72 ± 0.029 (4.0)	3.75 ± 0.570 (15.2)
		$r_O + H_p$	0.62 ± 0.015 (2.4)	1.80 ± 0.277 (15.3)
2	47.62	r_O	0.71 ± 0.014 (2.0)	2.95 ± 0.373 (12.6)
		H_p	0.81 ± 0.009 (1.2)	6.25 ± 0.305 (4.9)
		$r_O + H_p$	0.74 ± 0.007 (0.9)	3.80 ± 0.206 (5.4)
3	70.42	r_O	0.87 ± 0.008 (1.0)	5.60 ± 0.318 (5.7)
		H_p	0.95 ± 0.007 (0.7)	8.13 ± 0.275 (3.4)
		$r_O + H_p$	0.93 ± 0.006 (0.6)	7.63 ± 0.239 (3.1)

considering combined respirometric–titrimetric data (experiment 2 in Table IV). For this evaluation the parameter i_{XB} was assumed to be known, and the remaining unknown parameters $\mu_{max\ H}\ X_B(0)$, K_S, Y_H, and $S_S(0)$ were estimated. The evolution of the objective function as a function of the values of the four estimated parameters ($\mu_{max\ H}\ X_B[0]$, K_S, Y_H, and $S_S[0]$) was evaluated via contour plots (see Fig. 4). In addition, 95% confidence intervals were estimated here based on the procedure of Beale (1960). In this procedure the $(1 - \alpha)$ confidence region corresponds with the region for which the objective function MSE (mean square error, $J_{min}/[N - f]$) is below the threshold given by $MSE \cdot (1 + f/[N - f]^* F_{\alpha,f,N-f})$, where N = number of data points, f = number of estimated parameters, and $F_{\alpha,f,N-f}$ = value of the F distribution with f and $N - f$ degrees of freedom at a confidence level α. In Figure 4, the inner bold line indicates the 95% confidence limits according to Beale (1960) while the second bold line indicates the points where the objective function reaches a value corresponding to the double of its minimum value. From Figure 4 it is clear that all four parameters were practically identifiable in this case since the contour lines are closed.

The 95% confidence intervals of the parameters $\mu_{max\ H}\ X_B(0)$ and K_S were evaluated for each of the acetate and dextrose experiments according to the method described in Petersen et al. (2001), and results are given in Tables VI and VII, respectively.

Tables VI and VII both show that the confidence intervals narrow as more initial substrate is added. The explanation for this is that a higher initial substrate concentration results in a larger number of available data points, and, as a result, more information is available for the parameter estimation procedure. Exceptions to this observation are the titrimetric experiments with acetate, as will be discussed below.

In general the variance on the $\mu_{max\ H}\ X_B(0)$ estimates is lower than for K_S, as found in many similar studies (Holmberg, 1982; Vanrolleghem et al., 1995). In fact this is also indicated in the contour plots with K_S (Fig. 4). Indeed, for the contour plots involving K_S the contour lines actually indicate that the objective function is rather flat in the direction of K_S. The latter becomes obvious when considering the relative parameter value axes (top and right axis for each figure). The relative parameter value axes were obtained via a normalization, by division of the actual parameter values by the optimal parameter values. For the $\mu_{max\ H}\ X_B(0)$ and K_S depending contour lines for example (Fig. 4C), applying the same relative scale to the X and the Y axes (e.g., scaling both relative parameter values from 0.6 to 1.6) would result in a rather long elliptic contour that indicates a much higher variance on K_S than on $\mu_{max\ H}\ X_B(0)$. This observation corresponds with the results of Table VI, where the 95% confidence intervals expressed as percent of the actual parameter values are 10 to 20 times higher for K_S compared to $\mu_{max\ H}\ X_B(0)$ when combined respirometric–titrimetric data are considered.

Table VI indicates that the parameter variance is especially high for the K_S values obtained for the acetate experiments when H_p measurements are considered. In fact the 95% confidence intervals in percentage of the actual parameter values are in the order of 50%–400%, which means that the K_S values can even become negative within these intervals. As mentioned above the actual estimated K_S values obtained for the acetate experiments based on H_p measurements are rather low. Practically, this means that information on this parameter is only available from very few data points (the titrimetric profile bends very sharply when the substrate is consumed), thereby obviously decreasing the accuracy of its estimation. On the contrary, for the dextrose experiments the actual K_S values estimated from titrimetric data were higher than the respirometry-based values (see Table V). As a result, for dextrose, the confidence intervals for the titrimetry-based K_S values were smaller compared to the intervals obtained from respirometric data (see Table VI). In general the higher the K_S values, the more data points are available for the K_S parameter estimation and the more reliable that estimation will be.

Considering both acetate and dextrose experiments, an improvement in parameter variance is in most cases observed when combined data are considered for parameter estimation, compared to the situation where

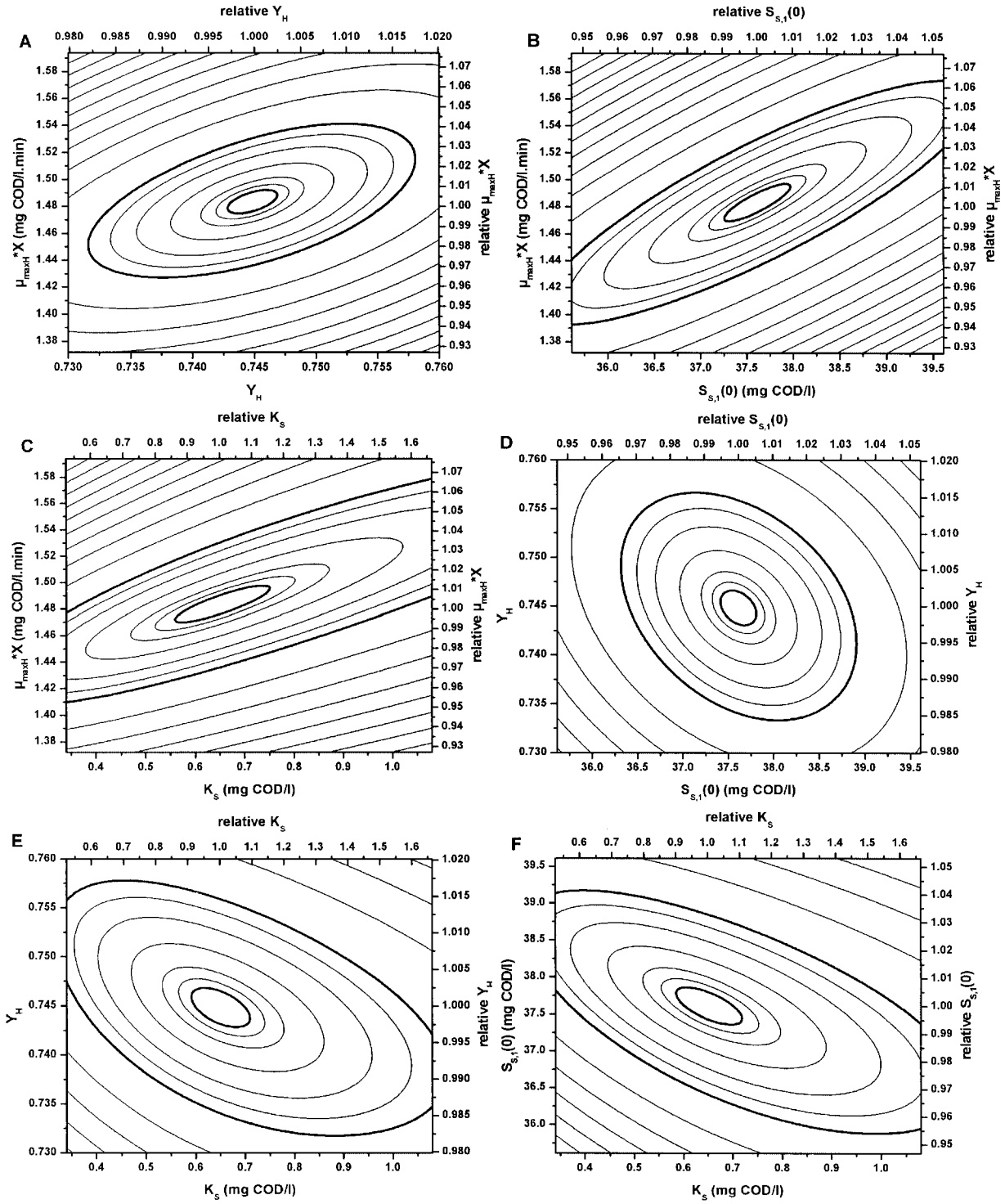

Figure 4. Contour plots obtained from parameter estimation on combined respirometric–titrimetric data. Estimated parameters (optimal values between parentheses) $\mu_{max\ H}\ X_B$ (1.48 mg COD/L · min), K_S (0.65 mg COD/L), Y_H (0.75), $S_{S,1}(0)$ (37.61 mg COD/L). The factor i_{XB} was fixed at 0.078 mg N/mg biomass COD.

only one data set (either respirometric or titrimetric data) is used for the estimation. This observation is more clear for the dextrose data than for the acetate data, where the improvement in the estimation accuracy of both $\mu_{max\ H}\ X_B(0)$ and K_S is in the order of 35%–55% when combined data are considered.

DISCUSSION

This study dealt with the structural and practical parameter identifiability of a model proposed to describe short-term aerobic batch carbon source degradation experiments using respirometric and/or titrimetric data.

Processes included in the model to describe titrimetric data were substrate uptake, CO_2 production, and NH_3 uptake for biomass growth, as discussed in Gernaey et al. (2002).

In the first phase the structural identifiability properties of the model were examined using both the Taylor series expansion method (Pohjanpalo, 1978) and the generalization method of Petersen et al. (2000). As expected, both methods yielded identical results. However, the generalization method is considerably less time consuming compared to the Taylor series expansion method because one does not have to spend time on developing rather complicated mathematical equations and trying to solve a relatively large set of non-linear equations for the unknown parameter combinations. As such, based on the model matrix (see Table I for an example) the generalization method allows to perform a quick structural identifiability analysis for Monod kinetics based models by applying a set of simple rules.

In the study of Dochain et al. (1995), only respirometric data were considered and the biomass concentration was assumed to be constant. This resulted in three identifiable parameter combinations ($[(1-Y_H)/Y_H]\mu_{max\ H}\ X_B$, $(1-Y_H)K_S$, $(1-Y_H)S_S(0)$). In this study growth was included in the model and four parameter combinations became identifiable from respirometric data ($\mu_{max\ H}$, $[(1-Y_H)/Y_H]X_B(0)$, $(1-Y_H)K_S$, $(1-Y_H)S_S(0)$) since the combination including $\mu_{max\ H}$ and X_B (Dochain et al., 1995) could be split up further. This finding is similar to the results of Spérandio and Paul (2000). The structure of the parameter combinations that are structurally identifiable on the basis of titrimetric data was similar to the ones with respirometric data, only with different stoichiometric factors.

Furthermore, from the structural identifiability study it became clear that an additional parameter combination is identifiable when combined data are considered, compared to a situation where only one measured variable is available. This additional parameter combination appears to be nothing else than the ratio of the two stoichiometric factors that relate the respective measured variables to substrate degradation, and allows the identifiability of the biomass yield, Y_H. For biological models the yield coefficient is an important factor because it will determine the distribution of substrate between biomass production and electron acceptor consumption. It is therefore of major importance to have a unique value for this parameter. As such it is not surprising that the identifiability of the biomass yield requires two kinds of measurements, since the yield coefficient relates two measures that link how much biomass is produced per unit of substrate degraded. Holmberg (1982), who identified the yield coefficient uniquely by assuming combined measurements of biomass and substrate, already proved this.

However, to practically identify all structurally identifiable parameter combinations, especially the $\mu_{max\ H}$ and $X_B(0)$ parameters, the data must exhibit a significant increase in biomass. Such an increase in biomass concentration is typically only visible in a long-term experiment with higher initial substrate concentration. A good example for a respirometric experiment that does show significant biomass growth is, e.g., presented in the work of Kappeler and Gujer (1992) and was also applied by Spérandio and Paul (2000). Care should be taken with such high-substrate experiments since the resulting significant biomass growth may at the same time induce a shift in the microbial composition of the biomass (fast growers overtaking slow growing organisms), leading to parameter estimates that are no longer representative for the original system. Under the experimental conditions applied in this study (lower substrate concentrations), the evaluation of the contour plots for one example (see Fig. 4) showed that it was practically possible to identify only four of the possibly five parameter combinations with combined respirometric–titrimetric data, i.e., $S_S(0)$, Y_H, K_S, and the combination $\mu_{max\ H}\ X_B$, assuming that i_{XB} was known.

In the model for description of the titrimetric data, one additional biomass-related parameter, i.e., i_{XB}, must be introduced compared to a model developed to describe respirometric data. Indeed, the uptake of nitrogen in the form of ammonia has a significant pH effect resulting from its proton production (Gernaey et al., 2002). This contribution to the overall proton balance was observable in the titrimetric data even though biomass growth was not. This explains why it was decided to include biomass growth in the model despite the fact that the actual data appeared not to be informative enough to enable separate practical identifiability of $\mu_{max\ H}$ and $X_B(0)$. The structural identifiability of this parameter i_{XB} was also assessed. Considering the parameter combination involving i_{XB} (Table III) and the fact that the yield coefficient Y_H is structurally identifiable when respirometric and titrimetric data are combined and i_{XB} is known, it is easy to see that i_{XB} is structurally identifiable when $S_S(0)$ is known.

When comparing the results of Tables IV and V it is striking that the estimated values of $\mu_{max\ H}\ X_B(0)$ on separate respirometric and titrimetric data are well comparable, both for acetate and dextrose, whereas the differences in estimated K_S values are larger. For dextrose, titrimetry-based K_S values are higher than respirometry-based ones (6.04 compared to 3.38 mg COD/L). One hypothesis could be that the CO_2 produced during dextrose oxidation, which is the main phenomenon observed via the titrimetric data, is only slowly released into the mixed liquor, resulting in estimated K_S values that are higher than those obtained from respirometric data that reflect the earlier uptake of oxygen from the mixed liquor. For acetate, on the other hand, titrimetry-based K_S values are lower than the ones obtained from respirometric data (0.25 compared to 0.65 mg COD/L). In this case the observed proton produc-

tion is mainly due to the acetate uptake and, thus, takes place before substrate oxidation. The affinity of the bacteria for substrate uptake is possibly higher than for substrate oxidation. Further research is, however, needed to verify and include such phenomena in a model.

Assuming that either $S_S(0)$ or i_{XB} is known, combined respirometric–titrimetric data sets derived from short-term batch experiments allowed practical identifiability of an extra parameter, the heterotrophic yield coefficient Y_H, compared to separate respirometric or titrimetric data (see both Tables IV and V), which confirms the results of the structural identifiability study. Note here that Y_H can also be estimated from r_O data assuming that $S_S(0)$ is known. However, it is also interesting to have a closer look at the estimates for the yield coefficients themselves, especially in relation to the estimated biomass nitrogen contents i_{XB} that could be obtained because the initial substrate concentrations $S_S(0)$ were known. The yield coefficients obtained are considerably higher than the values that are typically observed for heterotrophic growth (0.6–0.7; Henze et al., 1987), which is indicative of storage of carbon source in the cell (van Loosdrecht et al., 1997). It can be noticed that even more storage occurs for dextrose uptake compared to acetate uptake, which was also reported by Dircks et al. (1999). Turning to the estimated biomass nitrogen contents, we observe that the values found are lower than the values of the activated sludge models (i_{XB} = 0.086 g_N/g_{COD}; Henze et al., 1987) that are based on typical microbial nitrogen contents. However, when taking into account that the conversion process considered may include considerable storage activity, it becomes quite acceptable to find lower i_{XB} values because not all COD consumed gives rise to the formation of new biomass. This will therefore result in an apparent lower i_{XB}, ammonium uptake and pH effect. The fact that the i_{XB} values for dextrose are lower than the ones of acetate corroborates this interpretation.

The practical identifiability study also showed that the parameters $\mu_{max\ H}$ $X_B(0)$ and K_S were rather correlated, which appeared clearly from the contour plots (Fig. 4C). This observation has been recorded earlier in literature, e.g., Holmberg (1982). Especially the estimation of K_S is difficult and rather large confidence intervals are often obtained. Indeed, in this study with acetate as substrate and titrimetric measurements, the confidence intervals on K_S were sometimes that large that the zero value was contained in the intervals, similar to observations recorded in the study of Holmberg (1982). Thus, especially in the case of acetate the data were not informative enough to allow reliable estimation of K_S. Obviously a better experimental design could improve the practical identifiability of the K_S parameter, for instance, as shown by Vanrolleghem et al. (1995).

Although leading to more complex estimation problems, it is obviously more optimal to estimate parameters on the combined data sets since the resulting parameter values will reflect the information contained in both data sets. Furthermore, one may expect that the parameter variance decreases when two data sets are available for parameter estimation. Petersen et al. (2001) studied the nitrification process and reported an improvement in parameter estimation accuracy for $\mu_{max\ A}$ and K_{NH} with about 50% when comparing parameter estimation on respirometric data to parameter estimation on combined respirometric–titrimetric data. In that particular example adding titrimetry as an extra measurement to the experimental set-up yielded a clear advantage when it came to estimating kinetic parameters. For the aerobic carbon source degradation model studied in this paper, the improvements in confidence interval vary considerably depending on the carbon source. Especially for acetate (see Table VI) estimating parameters on combined data sets does not always result in a significant improvement of the confidence intervals compared to parameter estimation on respirometric data alone (0%–29% and 6%–12% improvement in confidence intervals for $\mu_{max\ H}$ $X_B[0]$ and K_S, respectively). For dextrose (see Table VII), the use of titrimetric data as a second information source for parameter estimations next to respirometric data resulted in confidence intervals that improved with 38%–55% and 36%–57% for $\mu_{max\ H}$ $X_B(0)$ and K_S, respectively. The results with dextrose are thus more or less comparable with the results reported by Petersen et al. (2001), whereas the results obtained for acetate are not. An explanation for this difference in improvement between the two carbon sources might be found in the observation that the K_S values obtained with titrimetric and respirometric data, respectively, were different. In the case of the acetate degradation, these differences may have hampered the potential improvement that could be achieved by combining titrimetric and respirometric data sets. The model mismatch that is the result of trying to fit the model to both data sets despite their clearly different K_S values, is reflected in larger parameter confidence regions.

CONCLUSIONS

The structural and practical identifiability of a model for description of respirometric–titrimetric data derived from aerobic batch carbon source degradation experiments with activated sludge was evaluated. Only substrate degradation related responses were considered. The structural identifiability study showed that combined respirometric–titrimetric data allow identification of one extra parameter combination compared to estimation on separate respirometric or titrimetric data. This extra parameter combination results in structural identifiability of the biomass yield Y_H on condition that the parameter i_{XB} (nitrogen content of the biomass) is

known. The model described the data collected for aerobic degradation of acetate and dextrose well. Parameter estimation on combined respirometric–titrimetric data results in significant improvements of parameter confidence intervals. However, these improvements seem to be substrate depending. Typically 95% confidence intervals lie in the range of 1% of the estimated growth rate and 10% of the estimated half-saturation substrate concentration. The biomass yield Y_H and the biomass nitrogen content i_{XB} could be estimated reliably from combined respirometric–titrimetric data when the initial substrate concentration $S_S(0)$ was known. Y_H can also be estimated from r_O data when $S_S(0)$ is known. The values obtained are close to reported values and, most interestingly, seem to reflect well the substrate storage processes that can be expected when pulse feeding acetate and dextrose.

The research of Krist Gernaey was financed by a postdoctoral scholarship of the Flemish Institute for the Promotion of Scientific-Technological Research in the Industry (IWT, Brussels; 1998–1999), a scholarship of the Quebec Ministry of Education in the frame of the "Bourses d'Excellence" program (2000), and a NATO scholarship (2000).

NOMENCLATURE

$\boldsymbol{a}_k(\boldsymbol{p})$	Taylor derivatives
b_H	decay rate (1/min)
b_C	CO_2 stripping rate (meq/L · min)
C	conversion factor (g COD/mol)
COD	chemical oxygen demand
CV	coefficient of variance (%)
f	number of parameters
$F_{\alpha,f,N-f}$	value F distribution with f and $N - f$ degrees of freedom at a confidence level α
H_p	proton concentration in mixed liquor (meq/L)
i	index for the process under study (generalization method)
i_{XB}	fraction of N in biomass (g_N/g_{COD} biomass)
j	index for the measured component (generalization method)
k	index for the substrate under study (generalization method)
K_La	oxygen transfer coefficient (1/min)
K_{NH}	autotrophic half-saturation substrate concentration (mg N/L)
K_S	heterotrophic half-saturation substrate concentration (mg COD/L)
m	fraction of dissociated acid A^- in the liquid phase for a monoprotic acid HA
MSE	mean square error
n	number of protons produced per CO_2 molecule released
N	number of data points
p	fraction of NH_4^+ in liquid phase
$\boldsymbol{p}$	parameter vector
rH_p	proton production rate (meq/L · min)
r_O	oxygen uptake rate (mg/L · min)
$r_{O,2}$	oxygen uptake rate in the respiration chamber (mg/L · min)
$r_{O,ex}$	exogenous respiration rate (mg/L · min)
$r_{O,end}$	endogenous respiration rate (mg/L · min)
r_{S_S}	substrate consumption rate (mg COD/L · min)
r_X	biomass formation rate (mg COD/L · min)
S_{NH}	ammonium (concentration) (mg N/L)
S_O	dissolved oxygen concentration in the liquid phase (mg/L)
S_O^0	saturation dissolved oxygen concentration (mg/L)
S_S	readily biodegradable substrate concentration (mg COD/L)
x	number of carbon atoms per substrate molecule for a $C_xH_yO_z$ carbon source
X_B	biomass concentration (mg COD/L)
y	number of hydrogen atoms per substrate molecule for a $C_xH_yO_z$ carbon source
$\boldsymbol{y}$	output
Y_H	yield coefficient for heterotrophic biomass (g_{COD}/g_{COD})
z	number of oxygen atoms per substrate molecule for a $C_xH_yO_z$ carbon source
$\mu_{max\ A}$	maximum specific growth rate for autotrophic biomass (1/min)
$\mu_{max\ H}$	maximum specific growth rate for heterotrophic biomass (1/min)
v	stoichiometric coefficient (generalization method)

References

Beale EML. 1960. Confidence regions in non-linear estimation. J R Stat Soc B22:41–88.

Bogaert H, Vanderhasselt A, Gernaey K, Yuan Z, Thoeye C, Verstraete W. 1997. New sensor based on pH-effect of denitrification process. J Environ Eng 123:884–891.

Dircks K, Pind PF, Mosbaek H, Henze M. 1999. Yield determination by respirometry: the possible influence of storage under aerobic conditions in activated sludge. Water SA 25:69–74.

Dochain D, Vanrolleghem PA, Van Daele M. 1995. Structural identifiability of biokinetic models of activated sludge respiration. Water Res 29:2571–2578.

Gernaey K, Vanrolleghem PA, Verstraete W. 1998. On-line estimation of *Nitrosomonas* kinetic parameters in activated sludge samples using titration in-sensor-experiments. Water Res 32:71–80.

Gernaey K, Petersen B, Nopens I, Comeau Y, Vanrolleghem PA. 2002. Modeling aerobic carbon source degradation processes using titrimetric data and combined respirometric–titrimetric data: experimental data and model structure. Biotechnol Bioeng 79:000–000.

Gernaey K, Petersen B, Ottoy JP, Vanrolleghem PA. 2001. Activated sludge monitoring with combined respirometric–titrimetric measurements. Water Res 35:1280–1294.

Henze M, Grady CPL Jr, Gujer W, Marais GvR, Matsuo T. 1987. Activated sludge model No. 1. IAWQ scientific and technical report no. 1. London: International Association on Water Quality.

Holmberg A. 1982. On the practical identifiability of microbial growth models incorporating Michaelis–Menten type nonlinearities. Math Biosci 62:23–43.

Kappeler J, Gujer W. 1992. Estimation of kinetic parameters of heterotrophic biomass under aerobic conditions and characterization of wastewater for activated sludge modelling. Water Sci Technol 25(6):125–139.

Keesman KJ, Spanjers H, van Straten G. 1998. Analysis of endogenous behavior in activated sludge. Biotechnol Bioeng 57:155–163.

Massone A, Gernaey K, Bogaert H, Vanderhasselt A, Rozzi A, Verstraete W. 1996. Biosensors for nitrogen control in wastewaters. Water Sci Technol 34(1-2):213–220.

Nelder JA, Mead R. 1964. A simplex method for function minimization. Computer J 7:308–313.

Petersen B. 2000. Calibration, identifiability and optimal experimental design for activated sludge models. Ph.D. thesis. Faculty for Agricultural and Applied Biological Sciences, Ghent University, Belgium.

Petersen B, Gernaey K, Vanrolleghem PA. 2000. Improved theoretical identifiability of model parameters by combined respirometric–titrimetric measurements. A generalisation of results. In: Proceedings of the IMACS 3rd Symposium on Mathematical Modelling, February 2–4, 2000, Vienna University of Technology, Austria. Vol 2, p 639–642.

Petersen B, Gernaey K, Vanrolleghem PA. 2001. Practical identifiability of model parameters by combined respirometric–titrimetric measurements. Water Sci Technol 43(7):347–356.

Petersen B, Gernaey K, Henze M, Vanrolleghem PA. 2002a. Evaluation of an ASM1 model calibration procedure on a municipal–industrial wastewater treatment plant. J Hydroinf 4:15–38.

Petersen B, Gernaey K, Vanrolleghem PA. 2002b. Anoxic activated sludge monitoring with combined nitrate and titrimetric measurements. Water Sci Technol 45(4-5):787–790.

Pohjanpalo H. 1978. System identifiability based on the power series expansion of the solution. Math Biosci 41:21–33.

Ramadori R, Rozzi A, Tandoi V. 1980. An automated system for monitoring the kinetics of biological oxidation of ammonia. Water Res 14:1555–1557.

Spanjers H, Olsson G. 1992. Modelling of the dissolved oxygen probe response in the improvement of the performance of a continuous respiration meter. Water Res 26:945–954.

Spanjers H, Vanrolleghem PA, Olsson G, Dold PL. 1998. Respirometry in control of the activated sludge process: principles. IAWQ Scientific and Technical Report No. 7. London, UK: International Association on Water Quality.

Spérandio M, Paul E. 2000. Estimation of wastewater biodegradable COD fractions by combining respirometric experiments in various $S_O/X_O/$ ratios. Water Res 34:1233–1246.

van Loosdrecht MCM, Pot M, Heijnen JJ. 1997. Importance of bacterial storage polymers in bioprocesses. Water Sci Technol 35(1):41–47.

Vanrolleghem PA, Gernaey K, Coen F, Petersen B, De Clercq B, Ottoy JP. 1998. Limitations of short-term experiments designed for identification of activated sludge biodegradation models by fast dynamic phenomena. In: Proceedings of the 7th IFAC conference on computer applications in biotechnology (CAB7), Osaka, Japan, May 31–June 4, 1998, p 567–572.

Vanrolleghem PA, Van Daele M, Dochain D. 1995. Practical identifiability of a biokinetic model of activated sludge respiration. Water Res 29:2561–2570.

Vanrolleghem PA, Spanjers H, Petersen B, Ginestet P, Takacs I. 1999. Estimating (combinations of) activated sludge model No. 1 parameters and components by respirometry. Water Sci Technol 39(1):195–214.