Practical identifiability of model parameters by combined respirometric-titrimetric measurements

B. Petersen, K. Gernaey and P.A. Vanrolleghem

BIOMATH Department, Ghent University, Coupure Links 653, B-9000 Gent, Belgium

Abstract An earlier study on theoretical identifiability of parameters for a two-step nitrification model showed that a unique estimation of the yield Y_{A1} is possible with combined respirometric-titrimetric data, contrary to the case where only one type of measurement is available. Here, the practical identifiability of model parameters was investigated via evaluation of the output sensitivity functions and the corresponding Fisher Information Matrix (FIM). It appeared that the FIM was not sufficiently powerful to predict the practical identifiability of this case with combined measurements as parameters could indeed be identified despite the fact that the FIM became singular. The accuracy of parameter estimates based on respirometric and titrimetric data and combination thereof was also investigated. Estimation on titrimetric data (Hp) was very accurate and a fast convergence of the objective function towards a minimum was obtained. The latter also holds for estimation on oxygen uptake rate data (r_O), however with a lower accuracy. Parameter estimation based on oxygen concentration data (S_O) was more complex but resulted in a higher accuracy. Thus, when the highest accuracy is needed it is recommended to estimate parameters initially on Hp and/or r_O data, and to subsequently use these parameters as initial values for final, and more accurate estimation on S_O data.

Keywords Fisher Information Matrix; identifiability; parameter estimation; respirometry; titrimetry

Introduction

Theoretical parameter identifiability is based on the model structure and the available measured outputs, and gives an indication of the maximum amount of information that can be obtained from a given (theoretical) experiment. The practical identifiability on the contrary not only depends on the model structure, but is also related to the experimental conditions together with the quality and quantity of the measurements. It should be stressed that the practical parameter identifiability often does not correspond with the theoretically derived one due to poor data quality (Holmberg, 1982).

The theoretical identifiability of the parameters of a two-step nitrification model has been studied via the series expansion methods: Taylor and generating series expansions (Petersen, 2000; Petersen *et al.*, 2000). The two-step nitrification Monod model (Table 1) consists of two processes: (1) oxidation of ammonium (S_{NH}) to nitrite (S_{NO2}), and (2) oxidation of nitrite to nitrate (S_{NO3}). Both nitrification steps can be characterised by measurements of oxygen consumption, whereas only the first step can be characterised by its proton production (Hp) since the oxidation of S_{NO2} does not have a pH effect. Note that the model in Table 1 does not take biomass growth into account (X = constant). This is a reasonable assumption when the model is to be used to describe short-term experiments, as will be the case in this paper. The identifiability study considered different cases: (1) availability of only respirometric data (measurements of dissolved oxygen, S_O, or oxygen uptake rate, r_O), (2) availability of only titrimetric data (cumulative proton production, Hp), and (3) availability of a combination of both measurements. Table 2 summarises the results of the theoretical identifiability study.

An important result of the theoretical identifiability study was that the autotrophic yield for the first nitrification step (Y_{A1}) becomes uniquely identifiable when a combination of respirometric and titrimetric measurements is available. This can be concluded from the "$S_O + Hp$ or

Table 1 Two-step nitrification model, no biomass growth

Process $\downarrow$ / component $\rightarrow$	S_O	S_{NH}	S_{NO2}	S_{NO3}	Hp	Process rate
1. Nitrification step 1	$-\dfrac{3.43 - Y_{A1}}{Y_{A1}}$	$-\dfrac{1}{Y_{A1}}$	$\dfrac{1}{Y_{A1}}$		$\dfrac{1}{7Y_{A1}}$	$\mu_{\max A1} \dfrac{S_{NH}}{K_{SA1} + S_{NH}} X$
2. Nitrification step 2	$-\dfrac{1.14 - Y_{A2}}{Y_{A2}}$		$-\dfrac{1}{Y_{A2}}$	$\dfrac{1}{Y_{A2}}$		$\mu_{\max A2} \dfrac{S_{NO2}}{K_{SA2} + S_{NO2}} X$

Table 2 Overview of theoretically identifiable parameter combinations for nitrification step 1 and 2, depending on the available measurement(s) (Petersen *et al.*, 2000)

Nitrification step 1			Nitrification step 2
S_O or r_O	Hp	$S_O + Hp$ or $r_O + Hp$	S_O or r_O
$\dfrac{3.43 - Y_{A1}}{Y_{A1}} \mu_{\max A1} X$	$\dfrac{2}{14} \dfrac{\mu_{\max A1} X}{Y_{A1}}$	$\dfrac{3.43 - Y_{A1}}{Y_{A1}} \mu_{\max A1} X$	$\dfrac{1.14 - Y_{A2}}{Y_{A2}} \mu_{\max A2} X$
$(3.43 - Y_{A1}) K_{SA1}$	$\dfrac{2}{14} K_{SA1}$	$(3.43 - Y_{A1}) K_{SA1}$	$(1.14 - Y_{A2}) K_{SA2}$
$(3.43 - Y_{A1}) S_{NH}(0)$		$(3.43 - Y_{A1}) S_{NH}(0)$	$(1.14 - Y_{A2}) S_{NO2}(0)$
	$\dfrac{2}{14} S_{NH}(0)$	$\dfrac{14}{2}(3.43 - Y_{A1})$	

$r_O + Hp$" column in Table 2, where the 4. expression results in a unique identification of Y_{A1}. This means that $\mu_{\max A1} X$, K_{SA1}, Y_{A1} and $S_{NH}(0)$ are theoretically identifiable from combined respirometric-titrimetric data. On the contrary, in case only one kind of measurement is available the parameter Y_{A1} is only identifiable in a combination with other parameters, as can be seen in the "S_O or r_O" and the "Hp" column in Table 1. It is in fact not surprising that a unique identification of Y_{A1} requires two kinds of measurements, since the yield coefficient links the amount of produced biomass to the number of degraded substrate units.

The methods by which the theoretical identifiability can be studied (e.g. the series expansions) may however not be that straightforward to apply, and may result in sets of non-linear equations that are far from simple, even for models of moderate complexity. Alternatively, the practical parameter identifiability can be evaluated directly based on the sensitivity functions (Holmberg, 1982; Marsilli-Libelli, 1989) and the corresponding Fisher Information Matrix (FIM), which is a known measure for practical identifiability (Munack, 1991).

In this paper, the theoretical identifiability results (Petersen, 2000; Petersen *et al.*, 2000) are evaluated for a specific case study through interpretation of the sensitivity functions and the corresponding FIM. This evaluation was carried out as a preliminary step to apply the results of the theoretical identifiability study in the context of optimal experimental design for combined respirometric-titrimetric experiments based on optimisation of the FIM (Petersen, 2000). Furthermore, the accuracy of parameter estimates based on respirometric or titrimetric data and combined respirometric-titrimetric measurements was investigated.

Theoretical background

Parameter estimation typically aims at minimising a weighted sum of squared errors, $J(\mathbf{p})$, between model, $\mathbf{y}(t_i, \mathbf{p})$, and measured, $\mathbf{y_m}(t_i)$, outputs with the weights $\mathbf{Q_i}$ and N the number of measurements (Eq. (1)). The minimisation is obtained by optimal choice of the parameter vector $\mathbf{p}$:

$$J(\mathbf{p}) = \sum_{i=1}^{n} (\mathbf{y}(t_i, \mathbf{p}) - \mathbf{y_m}(t_i))^{\mathrm{T}} \mathbf{Q_i}(\mathbf{y}(t_i, \mathbf{p}) - \mathbf{y_m}(t_i)) \tag{1}$$

The effect of a small parameter deviation, $\partial\mathbf{p}$, on the model fit, described by $J(\mathbf{p})$, can be evaluated by introducing a linearisation of the model with respect to the parameters along the trajectory (Eq. (2)):

$$\mathbf{y}(t_i,\mathbf{p}+\partial\mathbf{p}) \approx \mathbf{y}(t_i,\mathbf{p}) + \left[\frac{\partial\mathbf{y}(t_i,\mathbf{p})}{\partial\mathbf{p}}\right]_{\mathbf{p}}\partial\mathbf{p} \approx \mathbf{y}(t_i,\mathbf{p}) + Y_{\mathrm{p}}(t_i,\mathbf{p})\partial\mathbf{p} \tag{2}$$

In Eq. (2) $Yp(t_i, \mathbf{p})$ denotes the output sensitivity functions with respect to the parameters. Thus, the expected value of the objective function $J(\mathbf{p})$ can be reformulated (Eq. (3)).

$$E\left[J(\mathbf{p}+\partial\mathbf{p})\right] = J(\mathbf{p}) + \partial\mathbf{p}^{\mathrm{T}}\left(\sum_{i=1}^{N} Y_{\mathrm{p}}(t_i,\mathbf{p})^{\mathrm{T}}\mathbf{Q}_i Y_{\mathrm{p}}(t_i,\mathbf{p})\right)\partial\mathbf{p} \tag{3}$$

Consequently, to obtain a reliable minimum for $J(\mathbf{p})$ the difference between $J(\mathbf{p})$ and $J(\mathbf{p}+\partial\mathbf{p})$ should be maximised, i.e. a minimum is sought where $J(\mathbf{p})$ is sensitive towards changes in $\mathbf{p}$. The reliability of the minimum can be increased by maximising the term between brackets in Eq. (3). If the weighting matrix $\mathbf{Q}_i$ in Eq. (3) is chosen as the inverse measurement error covariance matrix, assuming that the measurement noise is white (i.e. independent and normally distributed with zero mean), and uncorrelated (i.e. the measurement error covariance matrix is a diagonal matrix), the term between brackets in Eq. (3) is defined as the Fisher Information Matrix (FIM). The more a measurement is noise corrupted the less it will count in the FIM. The FIM is given in Eq. (4) (Munack, 1991):

$$\mathrm{FIM} = \sum_{i=1}^{N} Y_{\mathrm{p}}(t_i,\mathbf{p})^{\mathrm{T}}\mathbf{Q}_i Y_{\mathrm{p}}(t_i,\mathbf{p}) \tag{4}$$

Moreover, the FIM is the inverse of the parameter estimation error covariance matrix, COV, and provides the Cramer-Rao lower bound on the parameter estimation errors, Eq. (5) (Ljung, 1999):

$$\mathrm{COV}(\mathbf{p}) \geq \mathrm{FIM}^{-1}(\mathbf{p}) \tag{5}$$

Thus, the FIM can be regarded as a summary of the output sensitivity functions and the measurement accuracy, thereby summarising the information concerning the model parameters gained from an experiment. Local parameter identifiability requires that the rank of the FIM is full. In case FIM is singular some of the sensitivity functions are proportional, and it will be impossible in that situation to obtain reliable parameter estimates from the data. Finally, confidence intervals can be derived since the standard deviation of the ith estimated parameter p^* can be obtained from the square root σ_i of the ith diagonal element of FIM^{-1}. An approximate confidence interval at level α is then given by Eq. (6), where t indicates the t distribution:

$$[p_i^* - t_{\alpha,(N-p)}\sigma_i, p_i^* + t_{\alpha,(N-p)}\sigma_i] \tag{6}$$

Case study
Data was obtained with a hybrid respirometer combined with a titrimetric measurement (Gernaey *et al.*, 2001). The set-up (Figure 1) consists of an open continuously aerated vessel and a closed non-aerated respiration chamber. It is equipped with two dissolved oxygen electrodes. Mixed liquor is continuously pumped between the aeration vessel and the respiration chamber. The oxygen mass balances are given in Eqs. (7) and (8), where the suffixes

1 and 2 refer to the aeration vessel and the respiration chamber respectively:

$$\frac{dS_{O,1}}{dt} = \frac{Q_{in}}{V_1}(S_{O,2} - S_{O,1}) + K_L a(S_O^o - S_{O,1}) - r_{O,1} \tag{7}$$

$$\frac{dS_{O,2}}{dt} = \frac{Q_{in}}{V_2}(S_{O,1} - S_{O,2}) - r_{O,2} \tag{8}$$

The main advantage of this respirometer is that the oxygen uptake rate r_O (i.e. $r_{O,2}$) can be calculated by a simple S_O mass balance over the closed respiration vessel (Eq. (8)), thereby avoiding the need to estimate $K_L a$ values (Gernaey *et al.*, 2001). In the hybrid respirometer substrate is added in the aeration vessel at the start of an experiment. Basically two processes contribute to r_O: (1) the immediate uptake of oxygen due to consumption of readily biodegradable substrate, i.e. the exogenous oxygen uptake rate $r_{O,ex}$, and (2) the endogenous oxygen uptake rate $r_{O,end}$. Typically $r_{O,ex}$ is modelled via a Monod expression (e.g Table 1) and $r_{O,end}$ is modelled as a first order decay process $b\,X$.

The basic concept of the titrimetric measurement technique is that the pH of the activated sludge sample is kept at a constant set-point while the cumulative amount of base and/or acid needed to keep that set-point is measured (Ramadori *et al.*, 1980). Such titrimetric data
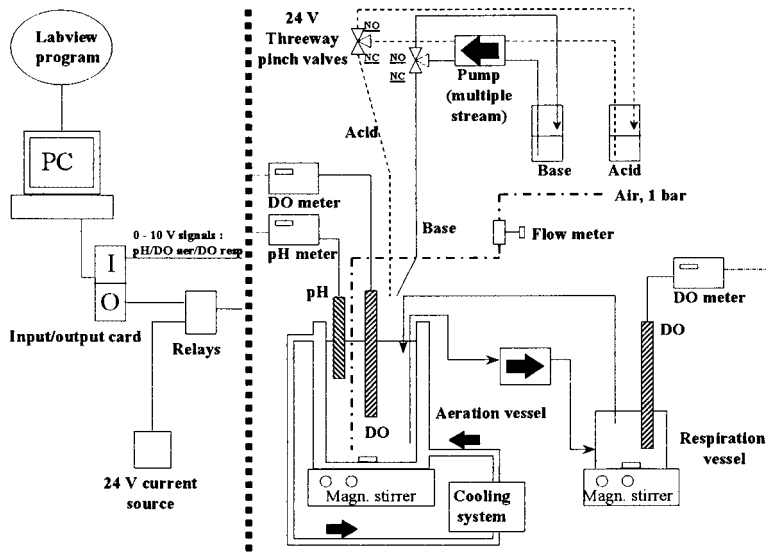
Figure 1 Experimental set-up used to collect combined respirometric-titrimetric data
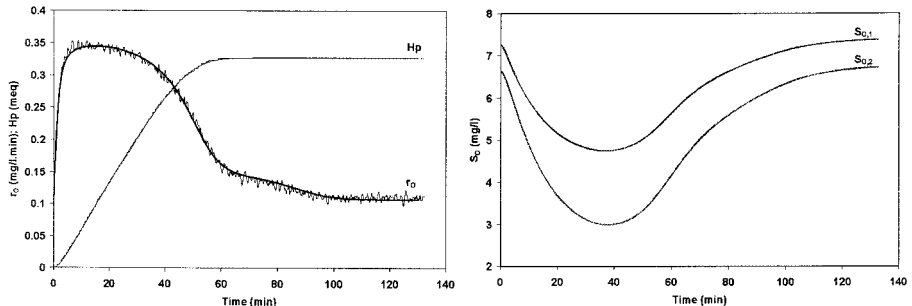

Figure 2 Data ($r_{O,2}$, Hp, $S_{O,1}$ and $S_{O,2}$) obtained after adding ammonium to activated sludge and model fit

contains kinetic information that is comparable to respirometric data in the case of nitrification, and can be modelled similarly (Gernaey *et al.*, 1998).

The experiment consisted of a S_{NH} addition to activated sludge. Data interpretation was based on the model of Table 1. Beside substrate degradation, the model included equations for the first order biomass decay and for the substrate transport in the set-up (Eqs. (7) –(8)) (Gernaey *et al.*, 2001). The data is presented in Figure 2 together with the model fits. A tail is observed in the r_O profile (from $t = 60$–100 min), indicating that the second nitrification step was slower than the first step. Here focus will only be on parameter estimation for the first nitrification step. Therefore, the parameters for the second step were fixed at known values obtained from a separate experiment with nitrite addition. Thus, according to the theoretical identifiability analysis (Table 2), the parameters μ_{maxA1}, K_{SA1}, Y_{A1} and $S_{NH}(0)$ could be estimated when considering combined $r_{O,2}$ and Hp data. If S_O data was considered $K_L a$ and S_O^o had to be estimated additionally (Eqs. (7) and (8)).

Results and discussion

Sensitivity functions and practical identifiability

The output sensitivity functions of S_O, $r_{O,2}$ and Hp with respect to the different parameters (Figure 3) were analytically derived in MAPLE V (Waterloo Maple Software). The well-known almost linear dependence between the sensitivities of the outputs with respect to μ_{maxA1} and K_{SA1} was observed (data not shown, see Petersen, 2000). This indicates that μ_{maxA1} and K_{SA1} will be correlated, and they may be difficult to identify in practice (Holmberg, 1982). Furthermore, Figure 3(a) and 3(b) indicate that the sensitivity functions of the outputs $r_{O,2}$ and Hp with respect to μ_{maxA1} and Y_{A1} appear proportional. This is confirmed by the sensitivity function equations that are proportional (Petersen, 2000). Consequently, based on these results it will not be possible to obtain reliable estimates for μ_{maxA1} and Y_{A1}.

As mentioned above, the sensitivity functions can pinpoint the experimental conditions under which the dependency of the outputs on the parameters is largest, and thereby under which conditions most information can be obtained on the parameters. For example, the sensitivities of $r_{O,2}$ with respect to the parameters exhibit rather sharp peaks indicating that at certain points the sensitivity of the outputs towards the parameters is rather significant (see Figure 3(a)). This indicates that the full information contained in the data is only
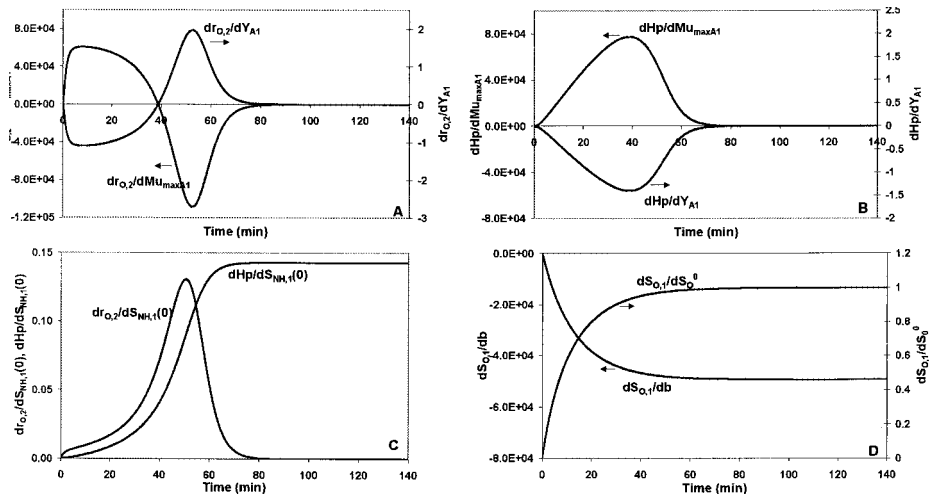

Figure 3 Sensitivity functions: A: $dr_{O,2}/d\mu_{maxA1}$ and $dr_{O,2}/dY_{A1}$; B: $dHp/d\mu_{maxA1}$ and dHp/dY_{A1}; C: $dr_{O,2}/dS_{NH,1}(0)$ and $dHp/dS_{NH,1}(0)$; D: $dS_{O,1}/d$ and $dS_{O,1}/db$

available during a rather short part of a typical batch experiment and, further, that large estimation errors may be generated if one does not collect sufficient data during this "sensitive" time interval (Marsilli-Libelli, 1989). On the other hand the profiles of the sensitivity functions of Hp do not exhibit as sharp peaks (Figure 3(b)). The information is not only concentrated in a limited time interval. In case of equidistant measurements, this may indicate that with Hp data better advantage is taken of the information provided by the entire data set. The sensitivity functions with respect to $S_{NH}(0)$ seemed to be clearly distinguishable from the other sensitivities indicating that a unique estimate can be expected (Figure 3(c)). The rather different shapes of the sensitivity functions of $r_{O,2}$ compared to Hp may illustrate the difference between the sensitivity of concentration (given by Hp) versus rate data (given by $r_{O,2}$). The sensitivity functions of S_O with respect to S_O^o and b in Figure 3(d) show that these are proportional, whereas it was found that the shape of the sensitivity function of $K_L a$ was distinguishable, indicating a possible reliable $K_L a$ estimation (Petersen, 2000). Summarising, it can be expected from the detailed sensitivity analysis study that estimation of μ_{maxA1} and K_{SA1} may cause practical problems. Y_{A1} is not identifiable when considering separate measurements of $r_{O,2}$, S_O or Hp. Furthermore, it was indicated that a separation of the parameters and b may be problematic.

According to the findings of the theoretical identifiability study (Table 2) Y_{A1} should become uniquely identifiable when measurements are combined. Surprisingly, however, when summarising the information of the sensitivity functions in the FIM for combined respirometric (S_O or $r_{O,2}$) and titrimetric (Hp) measurements, the FIM becomes singular indicating an unidentifiable situation. The inclusion of the sensitivity function with respect to Y_{A1} causes the singularity, since the sensitivity functions of μ_{maxA1} and Y_{A1} were proportional. Thus, there seems to appear a conflict between the application of FIM as a measure for local parameter identifiability and the results derived from the theoretical identifiability studies (Table 2). The FIM does not seem to reflect the improved theoretical identifiability achieved by combining measurements. Something similar was observed with the parameters S_O^o and b. The parameter b is theoretically identifiable when combining two oxygen measurements in the hybrid respirometer (Petersen, 2000). However, the inclusion of the sensitivity function of b also causes singularity of the FIM.

The reason for this discrepancy is not clear. It may be hypothesised that information on parameter identifiability obtained from the combination of measured outputs may be lost due to the local first order linearisation of the model with respect to the parameters on which FIM is based. To further investigate the practical identifiability of Y_{A1}, simultaneous estimation of μ_{maxA1}, K_{A1}, $S_{NH}(0)$ and Y_{A1} was carried out. A contour plot of the objective function as a function of the possibly correlated parameters μ_{maxA1} and Y_{A1} is given in Figure 4. As would be expected from the FIM results, it appears that the parameters μ_{maxA1}

Table 3 95% confidence intervals expressed as percentage of the parameter values, Q, via measurement errors or via objective function

Evaluation method Q	Q via measurement error			Q via objective function		
Data (↓) / Parameter (→)	μ_{maxA1}	K_{SA1}	$S_{NH}(0)$	μ_{maxA1}	K_{SA1}	$S_{NH}(0)$
$r_{O,2}$	1.385	6.328	0.510	1.884	8.805	2.433
Hp	0.905	5.780	0.074	0.911	5.820	0.074
$r_{O,2} + Hp$	0.696	3.942	0.070	0.826	4.740	0.077
$S_{O,1}$	0.231	0.943	0.241	0.308	1.255	0.321
$S_{O,1} + S_{O,2}$	0.093	0.506	0.028	0.335	1.821	0.104
$S_{O,1} + S_{O,2} + Hp$	0.091	0.511	0.026	0.264	1.503	0.066
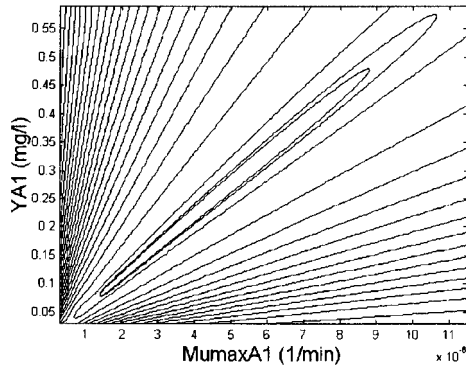
Figure 4 Contour plot of the objective function for the parameters μ_{maxA1} and Y_{A1}

and Y_{A1} are highly correlated in practice since a long valley is observed, indicating a severe practical identifiability problem. However, it is also obvious that the contour is closed. Hence, the result that Y_{A1} is theoretically identifiable is confirmed if one evaluates the non-linear objective function and not its linear approximation close to the minimum as done when applying a FIM-based analysis. Obviously this conflict between methods for identifiability analysis deserves further theoretical analysis.

Evaluation of parameter estimation accuracy with respirometric-titrimetric data

The confidence intervals of the estimated parameters are evaluated based on Eq. (6). It will be discussed whether respirometric data may be more powerful than titrimetric data for accurate parameter estimations.

The measurement error covariance matrix Q in Eq. (4) is practically estimated based on the actual S_{O}, $r_{\text{O,2}}$ and Hp data (Figure 2) (Petersen, 2000). For S_{O} and r_{O} data the measurement error (s^2) is estimated from a data series obtained during endogenous respiration (typically before or after the substrate addition). For Hp data it would, however, give an unrealistic, optimistic picture to estimate the measurement error from data after the point where substrate degradation is terminated, since the Hp profile in this case is a perfect horizontal line. As a consequence s^2 is estimated based on the data from $t = 15$–35 min where the slope is constant. It was found that the actual measurement error on S_{O} and $r_{\text{O,2}}$ is in the same order of magnitude, however relatively the measurement error on $r_{\text{O,2}}$ data is about 100 times larger than the relative error on S_{O} data. This is simply caused by the noise-introducing derivation of S_{O} data needed to obtain r_{O}. The 95% confidence intervals (expressed as percentage of the parameter values) are now retrieved by calculation of the inverse FIM, and by inserting the diagonal values into Eq. (6) (Table 3). It should be noted that due to the singularity problems with the FIM, the sensitivity functions of Y_{A1} and b were excluded from the FIM calculation. Thus, confidence intervals for these 2 parameters are not calculated.

First, these results indicate that it is more accurate to estimate the parameters based on Hp data than on $r_{\text{O,2}}$ data. This is especially the case for $S_{\text{NH}}(0)$ where the 95% confidence interval is as low as $\pm 0.074\%$ with Hp data. If one considers the Hp profile in Figure 2 this is not surprising since $S_{\text{NH}}(0)$ is in fact determined by the location of the constant horizontal plateau and many data points contain this information. Furthermore, it becomes obvious from Table 3 that the confidence intervals improve by the application of combined $r_{\text{O,2}}$ and Hp data. This is especially the case for the parameters μ_{maxA1} and K_{SA1}, where an improvement of about 50% is observed compared to measurements of $r_{\text{O,2}}$ alone. The accuracy of the $S_{\text{NH}}(0)$ estimate does, however, not improve further by applying combined $r_{\text{O,2}}$ and Hp measurements. When the accuracy based on S_{O} data is compared with the accuracy

obtained via $r_{O,2}$, it is clearly more accurate to estimate on the S_O data. This is due to the much higher measurement noise on $r_{O,2}$ data, as mentioned above. Furthermore, a significant accuracy improvement is observed when two dissolved oxygen measurements are available ($S_{O,1}$ + $S_{O,2}$) compared to one ($S_{O,1}$). This confirms the statement of Vanrolleghem and Spanjers (1998) that this set-up provides two measures of the respiration rate, thus duplicating the information on the kinetic parameters. The added value of Hp to combined $S_{O,1}$ and $S_{O,2}$ measurements seems to be insignificant.

The accuracies based on measurement errors reported in Table 3 are quite impressive and warrant some verification. However, the measurement noise of the applied experimental method is indeed not very significant, as indicated in Figure 2. The confidence intervals were verified via simulations where the parameter values were set to the limits of the 95% confidence intervals. The resulting simulated curves indeed just laid within the edges of the measurement noise, confirming that the calculated 95% confidence intervals are reliable. A final element of discussion is that the weighting matrix Q was based on the measurement errors only and did not include the model errors. Therefore, as a second evaluation of the parameter accuracy, Q was based on the values of $J(\mathbf{p})$ obtained from the parameter estimation, thereby including both measurement noise and model errors. The resulting 95% confidence intervals are given in the second half of Table 3.

In general the differences between the 95% confidence intervals obtained with the two error calculation approaches are not very large when considering $r_{O,2}$, Hp, $r_{O,2}$+Hp and $S_{O,1}$ data. This indicates that the model has been able to describe these data adequately since the main part of the errors is included in the measurement noise. This is especially the case for Hp measurements, where the confidence intervals obtained with both approaches are almost identical. On the contrary the difference in 95% confidence intervals with $S_{O,1}$ + $S_{O,2}$ data seems significant, and when model errors are considered the accuracy does not improve compared to $S_{O,1}$ data alone. However, the inclusion of Hp data now improves the accuracy. The reason for the lack of improvement in accuracy when considering two S_O data sets is probably because the complexity of the parameter estimation procedure increases drastically in this case. Consequently, if only measurement errors are considered, one may conclude that the most accurate parameters are obtained by applying the two oxygen measurements, $S_{O,1}$ and $S_{O,2}$, from the hybrid set-up. This is however contradicted if the modelling errors are included as well. The $J(\mathbf{p})$ based approach points more in the direction of applying only one set of oxygen measurements, i.e. $S_{O,1}$ or combined $S_{O,1}$ + Hp data.

Another important factor to consider before choosing the most adequate measured variables for a parameter estimation problem is the rate of convergence of $J(\mathbf{p})$ towards a minimum and, quite related, the sensitivity for local minima. It appeared that convergence is significantly faster when $r_{O,2}$ and/or Hp data are used rather than S_O data. This is due to the increased complexity of the estimation problem when using S_O data, since $K_L a$ and both need to be estimated simultaneously with the kinetic parameters μ_{maxA1}, K_{SA1} and $S_{NH}(0)$. Especially the parameter estimation based on Hp data alone is very fast and is, in addition, also very accurate, as indicated in Table 3. Thus, in practice one may initially choose to estimate the kinetic parameters on Hp and/or $r_{O,2}$ data, and to obtain an even higher accuracy, especially for μ_{maxA1} and K_{SA1}, apply these parameter values as initial guesses for estimations based on S_O data. The introduction of a second S_O data source, however, may seem questionable based on the case under study, due to increased complexity of the parameter estimation problem and its effects on convergence rates.

Conclusions

The practical identifiability was evaluated on a nitrification example via output sensitivity functions and the corresponding Fisher Information Matrix (FIM), considering combined

respirometric-titrimetric outputs. The FIM became singular indicating an unidentifiable situation despite the fact that a theoretical identifiability study had shown that the chosen parameter set, including a unique identifiability of the yield Y_{A1}, should be identifiable. The FIM seemed inadequate to evaluate this improved theoretical identifiability since the addition of the sensitivity of Y_{A1} in the FIM caused the singularity. Some information on the parameters may be lost when applying FIM due to the local first order linearisation of the model with respect to the parameters on which the FIM is based. Estimation of Y_{A1} was indeed possible in practice, although Y_{A1} estimates were strongly correlated with μ_{maxA1}. For this case study it thus seemed that a parameter identifiability evaluation based on FIM gave too pessimistic a picture.

The accuracy of parameter estimates based on respirometric and titrimetric data was evaluated considering (i) measurement errors, (ii) model errors and (iii) the complexity of the parameter estimation as characterised by the convergence rate of the estimation algorithm towards a minimum. Estimation of parameters on Hp data is very accurate and a fast convergence is obtained. The same holds for r_O data although the accuracy is less. An even higher accuracy can be obtained when the parameter estimates based on Hp and/or r_O are applied as initial values for the more complex parameter estimation based on S_O data.

References

Gernaey, K., Vanrolleghem, P.A. and Verstraete, W. (1998).On-line estimation of Nitrosomonas kinetic parameters in activated sludge samples using titration in-sensor-experiments. *Wat. Res*., **32**, 71–80.

Gernaey, K., Petersen, B., Ottoy, J.P. and Vanrolleghem,P.A. (2001).Activated sludge monitoring with combined respirometric – titrimetric measurements. *Wat. Res*. **35**, 1280–1294.

Holmberg, A. (1982).On the practical identifiability of microbial growth models incorporating Michaelis-Menten type nonlinearities. *Math. Biosci.*, **62**, 23–43.

Ljung (1999).Experiment design. System Identification – Theory for the User. Prentice-Hall, Englewood Cliffs, New Jersey.

Marsilli-Libelli, S. (1989).Modelling, identification and control of the activated sludge process. *Adv. Biochem. Eng. Biotechnol.*, **38**, 90–148.

Munack, A. (1991).Optimisation of sampling. *Biotechnology, a Multi-volume Comprehensive Treatise, Measuring, Modelling and Control.* K. Schügerl (ed.). VCH, Weinheim, **4**, 251–264.

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. and Vanrolleghem, P.A. (2000). Improved theoretical identifiability of model parameters by combined respirometric-titrimetric measurements. A generalisation of results. *Proceedings IMACS 3rd Symposium on Mathematical Modelling,* February 2–4, 2000, Vienna University of Technology, Austria. Vol.2, 639–642.

Ramadori, R., Rozzi, A. and Tandoi, V. (1980).An automated system for monitoring the kinetics of biological oxidation of ammonia. *Wat. Res*., **14**, 1555–1557.

Vanrolleghem, P.A. and Spanjers, H. (1998).A hybrid respirometric method for more reliable assessment of activated sludge model parameters. *Wat. Sci. Tech*., **37**(12), 237–246.