Reconstruction of the O_2 Uptake Rate and CO_2 Evolution Rate on a Time Scale of Seconds

H. H. J. Bloemen
TNO TPD, Instrumentation and Information Systems, 2600 AD Delft, The Netherlands

L. Wu, W. M. van Gulik, and J. J. Heijnen
Kluyver Laboratory for Biotechnology, Faculty of Applied Sciences, Delft University of Technology, 2628 BC Delft, The Netherlands

M. H. G. Verhaegen
Control Systems Engineering, Faculty of Information Technology and Systems, Delft University of Technology, 2600 GA Delft, The Netherlands

The reconstruction is addressed of the dynamics of the microbial oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER) from off-gas concentration measurements and dissolved oxygen measurements during highly dynamic conditions, encountered, for example, during a substrate-pulse experiment. The reconstruction of the OUR and CER is done by fitting a model for the mass transport to the measured data using smoothing techniques. Most of the parameters of the mass-transport model, including the sensor dynamics, are derived from physical knowledge. A few remaining parameters that cannot be derived from physical knowledge are computed by fitting the model to an identification data set. It is demonstrated that the net gas production or consumption can have a significant impact on the estimated OUR, and, therefore, is included in the model. Besides the reconstruction of the OUR and the CER, the reconstruction algorithm also incorporates the monitoring of the mass-transfer coefficient $k_l a$.

Introduction

The development of recombinant-DNA techniques, the unraveling of complete genomes, and genome wide information measurement (DNA chips) have recently opened the possibility of precise modifications in a microbial metabolism. The goals of such rational metabolic engineering are de novo or improved production of desirable chemical compounds. However, often product biosynthesis pathways are interconnected at various points with the primary metabolic pathways, such as for the supply of carbon precursors, various cofactors, and metabolic energy. Due to the complexity of the resulting metabolic network, recombinant-DNA based improvement of product formation and/or introduction of novel pathways driven by intuition alone is not feasible. A detailed insight in the stoichiometry and *in vivo* enzyme kinetics of the metabolic system, in the form of a mathematical model, is therefore indispensable (Bailey, 1998; Gombert and Nielsen, 2000). Estimation of *in vivo* kinetic parameters of such models requires pulse-response measurements, that is, disturbance of a metabolic steady state followed by the measurement of the responses (concentration profiles) of all relevant intracellular and extracellular metabolites within a short time window (typically 300 s) (Rizzi et al., 1997; Visser et al., 2000). In these models it is usually assumed that the concentrations of the enzymes, which catalyze the reactions of the metabolism, remain constant. However, the cells possess regulatory pathways that adapt the enzyme concentrations when the metabolite concentrations change. Fortunately, the time constant of adapting the enzyme concentrations is much

Correspondence concerning this article should be addressed to H. H. J. Bloemen.

larger than those of the reactions of the (primary) metabolism, and the enzyme concentrations may be assumed to remain constant over a time window of about 300 s after addition of the substrate pulse (Lange et al., 2001). In order to observe the fast dynamics of the different intra- and extra-cellular metabolites, and to be able to collect a sufficient amount of data within this time window, the development of rapid sampling techniques has received a lot of attention recently (for example, Lange et al., 2001; Schaefer et al., 1999; Theobald et al., 1997; Visser et al., 2002).

However, these methods allow only the measurement of metabolite concentrations in the liquid phase. In previous pulse-response studies the responses of gaseous compounds (such as oxygen and carbon dioxide) have not been measured (Buchholz et al., 2002; Buziol et al., 2002; Rizzi et al., 1997; Schmitz et al., 2002; Theobald et al., 1997; Vaseghi et al., 1999; Visser et al., 2000). Nevertheless, the dynamic response of the oxygen uptake rate (OUR) and the carbon dioxide evolution rate (CER) are closely related to the primary metabolism. Because the OUR is proportional to the aerobic energy (ATP) generation, its dynamics reflect the dynamics of the flux through the citric acid cycle and of the energy generation in oxidative phosphorylation. The CER is related to reaction rates of various parts in the primary metabolism, such as the citric acid cycle, the pyruvate branch and the pentose phosphate pathway. Moreover, knowledge of the OUR and CER gives insight in the carbon and redox balance during the pulse-response experiment. Therefore, the dynamics of the OUR and CER offer additional and valuable information of the cellular metabolism, next to the intra- and extracellular metabolite concentration measurements.

In this article we focus on the measurement of the (fast) dynamics of the OUR and the CER after the addition of a substrate pulse. However, these quantities cannot be measured directly since:

- In order to accurately measure the fast dynamics of both the OUR and the CER, the dynamics of the sensors have to be accounted for.
- The dynamics of the dissolved oxygen (DO) (and carbon dioxide) concentration are the result of both the OUR (CER) and the mass transfer between the liquid phase and the gas phase in the fermenter. Although the concentrations of oxygen and carbon dioxide in the off-gas of the fermenter are often measured, these concentrations do not reflect the actual concentrations in the gas phase of the fermenter under dynamic conditions, like in a substrate-pulse experiment. This is due to the presence of a head space and tubing, which cause additional dynamics in the off-gas system (transport delays and mixing effects).

Often, the estimation of the OUR from off-gas measurements is based on a steady-state assumption (for example, Lubenova, 1999). Under dynamical conditions, such as after the addition of a substrate pulse, the assumption of a steady state is violated, which causes significant errors in the estimated OUR (Dekkers, 1982; Zhong and Bellgardt, 1990). Under such conditions, the dynamics of the mass-transfer phenomena, including the dynamics of the sensors, need to be accounted for, which requires accurate modeling of these processes.

Although the reconstruction of the OUR and/or of the CER from either off-gas measurements or from DO measurements has been addressed before in the literature (for example, Dekkers, 1982; Lindberg and Carlsson, 1996; Wu et al., 2003; Zhong and Bellgardt, 1990), the approach in this article improves those results on a number of points as described below.

In Dekkers (1982) it is assumed that the mass transport of CO_2 between the liquid phase and the gas phase is infinitely fast, whereas, for O_2, it is assumed that the liquid-phase concentration is approximately equal to zero. Under these assumptions, only the mass balances for the gas phase need to be considered. The latter assumption certainly does not hold in well aerated fermenters, and, with respect to the former assumption, it is indicated later that the mass-transfer coefficients for O_2 and CO_2 are of the same order of magnitude. In this article also the mass balances for the liquid phase are included, presented later, and the above assumptions are not required.

In Lindberg and Carlsson (1996) and Zhong and Bellgardt (1990) only the reconstruction of the OUR is considered, either only from DO data (Lindberg and Carlsson, 1996), or only from off-gas data (Zhong and Bellgardt, 1990). In this article the objective is to obtain estimates of both the OUR and the CER. This has also been reported in Wu et al. (2003), where two independent estimators for the OUR and CER are used. However, in a later section it is demonstrated that these estimators are not independent.

Often the mixing effects in the head space and tubing to the off-gas analyzers are ignored, or inappropriately modeled. For example, in Dekkers (1982) it is assumed that the gas phase and the head space together can be modeled as one stirred tank reactor, without verifying this assumption. On the contrary, in Wu et al. (2003) the model is obtained through an identification procedure such that the model accurately describes the identification data. In this way the quality of the model can be investigated prior to using it in an estimation procedure. However, the main drawback of the approach in Wu et al. (2003) is that it follows a black-box modeling approach for the two independent models that describe, respectively, the relation between the OUR and the oxygen concentration measured by the off-gas analyzer, and the relation between the CER and the carbon dioxide concentration measured by the off-gas analyzer. For the identification of the black-box models, the problem is encountered that the OUR and CER cannot be manipulated in a known way. To solve this problem, in Wu et al. (2003) an approximation is used, which makes it possible to lump the inputs OUR and CER with the oxygen and carbon dioxide concentration, respectively, in the inlet gas. In this way the concentrations of oxygen and carbon dioxide in the inlet gas can be used to excite the system in order to collect data for the black-box modeling step in Wu et al. (2003). In this article this approximation to lump the inputs is avoided by modeling the mass transport using first principles, discussed in the next section. The parameter identification of the model parameters is discussed later. The approach followed in this article also allows us to incorporate the effects of a net gas production or consumption on the flow rate of the outlet gas. In a later section it is demonstrated that the impact of the gas production/consumption on the estimates for the OUR is significant and thus needs to be incorporated into the model. This is easy for a model based on first principles, contrary to a black-box

model. Gas production/consumption introduces a nonlinearity and a coupling between the mass-transfer phenomena for oxygen and carbon dioxide. Besides, the information provided by the DO sensor is included in the approach followed in this article. This additional information enhances the reconstruction of the OUR and gives a possibility to monitor the mass-transfer coefficient as well, as described later.

Given the dynamic model for the mass transport, the OUR and the CER can be reconstructed from the off-gas data and the DO data by fitting the response of the model to the measurements. Such an optimization problem can be solved by a smoothing algorithm, as will be discussed later. The results of this approach are demonstrated both on simulated data and on pulse-experiment data. A discussion concludes the article.

Modeling

Fermentation broth

The experimental setup is shown in Figure 1. The O_2 and CO_2 mass transfer between the liquid phase and the gas phase in the fermentation broth can be modeled by (assuming that the fermentation broth is ideally stirred)

$$V_g\frac{\partial x_{g,\mathrm{O}_2}}{\partial t} = F_{g,\mathrm{in}}x_{g,\mathrm{O}_2,\mathrm{in}} - F_{g,\mathrm{out}}x_{g,\mathrm{O}_2} - k_l a_{\mathrm{O}_2}V_l\left(\frac{x_{g,\mathrm{O}_2}}{m_{\mathrm{O}_2}} - c_{l,\mathrm{O}_2}\frac{RT}{P}\right) \quad (1)$$

$$V_l\frac{\partial c_{l,\mathrm{O}_2}}{\partial t} = F_l(c_{l,\mathrm{O}_2,\mathrm{in}} - c_{l,\mathrm{O}_2}) + k_l a_{\mathrm{O}_2}V_l\left(\frac{P}{m_{\mathrm{O}_2}RT}x_{g,\mathrm{O}_2} - c_{l,O_2}\right) - V_l\cdot\mathrm{OUR} \quad (2)$$

$$V_g\frac{\partial x_{g,\mathrm{CO}_2}}{\partial t} = F_{g,\mathrm{in}}x_{g,\mathrm{CO}_2,\mathrm{in}} - F_{g,\mathrm{out}}x_{g,\mathrm{CO}_2} - k_l a_{\mathrm{CO}_2}V_l\left(\frac{x_{g,\mathrm{CO}_2}}{m_{\mathrm{CO}_2}} - c_{l,\mathrm{CO}_2}\frac{RT}{P}\right) \quad (3)$$

$$V_l\frac{\partial c_{l,\mathrm{CO}_2}}{\partial t} = F_l(c_{l,\mathrm{CO}_2,\mathrm{in}} - c_{l,\mathrm{CO}_2}) + k_l a_{\mathrm{CO}_2}V_l\left(\frac{P}{m_{\mathrm{CO}_2}RT}x_{g,\mathrm{CO}_2} - c_{l,\mathrm{CO}_2}\right) + V_l\cdot\mathrm{CER} \quad (4)$$

where x [mole·mole^{-1}] is the mole fraction of oxygen (subscript O_2) and carbon dioxide (subscript CO_2); the subscript g denotes the gas phase; henceforth, the subscript i is used to refer either to O_2 or to CO_2 where applicable; $c_{l,i}$ [mole·m^{-3}] is the concentration in the liquid phase; V_g [m^3] and V_l [m^3] are the volumes of the gas phase and the liquid phase, respectively; F_g [m^3·s^{-1}] is the gas-flow rate; the subscripts in and out refer to the inlet and outlet gas, respectively; $k_l a_i$ [s^{-1}] is the product of the mass-transfer coefficient $k_{l,i}$ [m·s^{-1}] and the specific surface area a [m^2·m^{-3}]; m_i [—] is the partition coefficient; P [Pa] is the pressure in the fermenter; $R = 8.314$ [J·mole^{-1}·K^{-1}] is the gas constant; T [K] is the temperature in the fermenter; and OUR [mole·m^{-3}·s^{-1}] and CER [mole·m^{-3}·s^{-1}] are the oxygen uptake rate and carbon dioxide evolution rate, respectively.

Off-gas system

The gas that leaves the fermentation broth cannot directly be analyzed. First, it enters the head space (a head space is necessary to control foaming of the fermentation broth, for example). Besides, the air that leaves the fermenter needs to be dried before it can be analyzed, which results in tubing. Neither the head space, nor the tubing, can be regarded as an ideally stirred tank or as an ideal plug-flow reactor. However, the combined dynamics of the off-gas system can be approximated by a series connection of several ideally stirred tanks plus a pure time delay (back mixing from the head space to the fermentation broth is assumed to be negligible)

$$V_{s,i}\frac{\partial x_{s,i}}{\partial t} = F_{g,\mathrm{out}}(x_{s-1,i} - x_{s,i}),\ s = 1, \ldots, n_s \quad (5)$$

$$x_{a,i}[t + d_i] = x_{s,i}[t] \quad (6)$$

with $x_{0,i} = x_{g,i}$, n_s is the number of ideally stirred tanks, d is the delay, and the subscript a refers to the analyzer. Note that, since the head space is not ideally stirred, $V_{1,i}$ does not necessarily correspond to the volume of the head space. Also note that $V_{s,i}$ is chosen to be gas dependent, which is motivated by the fact that the measurement cells for oxygen and carbon dioxide in the analyzer differ. $V_{s,i}$ are nonphysical parameters that are identified by fitting the model to the identification data (see the next section). The number of ideally stirred tanks n_s, and whether $V_{s,\mathrm{O}_2} = V_{s,\mathrm{CO}_2}$ or $V_{s,\mathrm{O}_2} \neq V_{s,\mathrm{CO}_2}$ is determined from the identification data.

DO sensor

The DO sensor consists of a membrane that separates the fermentation broth from the actual electrode. In order to

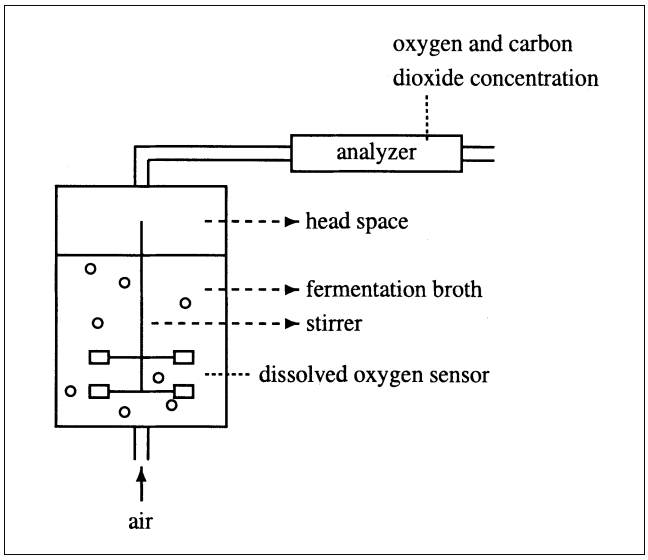

Figure 1. Off-gas system.

measure the oxygen concentration, the oxygen has to diffuse through the membrane and the stagnant layer of liquid between the membrane and the actual electrode. Physically, this mass transport can be modeled by a partial differential equation. However, often these DO dynamics can be approximated well by a first-order model (Vanrolleghem and Spanjers, 1998), plus a time delay (Lindberg and Carlsson, 1996)

$$\frac{\partial c_{s,\mathrm{O}_2}}{\partial t} = k_{DO}(c_{l,\mathrm{O}_2} - c_{s,\mathrm{O}_2}) \tag{7}$$

$$c_{DO}[t + d_{DO}] = c_{s,\mathrm{O}_2}[t] \tag{8}$$

k_{DO} is the parameter of the first-order model of the DO sensor, c_{s,O_2} denotes the oxygen concentration in the DO sensor, and c_{DO} refers to the output of the DO sensor.

Parameter Identification

The parameters of the model (Eqs. 1–8) need to be identified such that the model accurately describes the behavior of the experimental system. For this purpose, first the behavior of the experimental system is recorded in an identification experiment.

A chemostat fermentation of the yeast *Saccharomyces cerevisiae* CEN.PK 113-7D is performed in a 7×10^{-3} [m^3] fermenter with a 4×10^{-3} [m^3] working volume. The yeast is cultivated on a mineral medium containing glucose and ethanol as carbon source. Other conditions of the chemostat culture are listed in Table 1.

Ideally, one would like to record the response of the system to known variations in the OUR and the CER. However, this is not possible since the OUR and CER cannot be directly manipulated and are not known. Alternatively, the off-gas system can be excited by recording the response of the system to variations in the inlet concentrations $x_{g,\mathrm{O}_2,\mathrm{in}}$ and $x_{g,\mathrm{CO}_2,\mathrm{in}}$. One possibility is to identify all the parameters of the model (Eqs. 1–8) by fitting the model to this measured response of the system. The problem with such an approach is that the model output is insensitive to certain combinations of parameter values. Therefore, several parameter values will be obtained differently, which is indicated in Table 2.

Partition coefficients m_{O_2} and m_{CO_2} and the mass-transfer coefficients $k_l a_{O_2}$ and $k_l a_{CO_2}$

The partition coefficients m_{O_2} and m_{CO_2} follow from Henry's law, which gives (Henry constants are taken from (Janssen and Warmoeskerken, 1987))

$$m_{\mathrm{O}_2} = 34.37 \tag{9}$$

$$m_{\mathrm{CO}_2} = 1.39 \tag{10}$$

In order to be able to accurately measure $k_l a_{\mathrm{O}_2}$ and $k_l a_{\mathrm{CO}_2}$ one should adapt the concentration of the relevant component in one phase, and measure the concentration profile in the other phase, that is, change the concentrations in the gas phase and measure in the liquid phase or vice versa (Schneider and Frischknecht, 1977). Unless a special injection device for the liquid phase is available, like the one presented in Schneider and Frischknecht (1977), only the concentrations in the gas phase can be adapted easily. Since, in the experimental setup under consideration, only a dissolved-oxygen sensor is available and not a dissolved-carbon dioxide sensor, only the $k_l a_{\mathrm{O}_2}$ can be identified in this way. However, then also the dynamics of the DO sensor need to be accounted for, whereas k_{DO} and d_{DO} also need to be identified. The mass transfer from the gas phase to the DO electrode is characterized by a series connection of the mass transfer from the gas phase to the liquid phase, and of the diffusion through the membrane and the stagnant liquid layer of the DO sensor. Consequently, it is hard to identify $k_l a_{\mathrm{O}_2}$ and k_{DO} simultaneously, since an error in one of the two parameters can be compensated for easily by an error in the other parameter, that is, $k_l a_{\mathrm{O}_2}$ and k_{DO} are not accurately identifiable in this way. One possibility could be to identify the DO sensor dynamics separately by transferring the DO sensor from one solution to another with a different dissolved oxygen concentration (Lindberg and Carlsson, 1996). However, this procedure is not adopted here.

Alternatively, $k_l a_{\mathrm{O}_2}$ can be identified independently of the dynamics of the DO sensor and those of the off-gas analyzer when the fermenter is in a steady state, that is, during chemostat cultivation. Under these conditions, $k_l a_{\mathrm{O}_2}$ can be computed from

$$k_l a_{\mathrm{O}_2} = \frac{1}{V_l} \cdot \frac{1}{\dfrac{x_{g,\mathrm{O}_2}}{m_{\mathrm{O}_2}} - \dfrac{c_{l,\mathrm{O}_2} RT}{P}} \cdot \left(F_{g,\mathrm{in}} x_{g,\mathrm{O}_2,\mathrm{in}} - F_{g,\mathrm{out}} x_{g,\mathrm{O}_2} \right) \tag{11}$$

which can be derived from setting $(\partial x_{g,\mathrm{O}_2})/\partial t = 0$ in Eq. 1. Since the steady-state assumption is satisfied it follows that: $x_{g,\mathrm{O}_2} = x_{a,\mathrm{O}_2}, c_{l,\mathrm{O}_2} = c_{DO}$ and $F_{g,\mathrm{in}} = F_{g,\mathrm{out}}$ if the respiration quotient $RQ = \mathrm{CER}/\mathrm{OUR}$ equals one. Thus, if the measurements of the DO sensor and those of the gas analyzer are

Table 1. Chemostat Conditions

Quantity	Value	Unit
V_l	4×10^{-3}	(m^3)
V_g	0.2×10^{-3}	(m^3)
P	1.3×10^{5}	(Pa)
$F_{g,\mathrm{in}}$	4.3×10^{-5}	($m^3 \cdot s^{-1}$) at pressure P
F_l	5.56×10^{-8}	($m^3 \cdot s^{-1}$)
T	303	(K)
pH	5.0	
Stirrer speed	600	(rpm)

Table 2. Overview of the Identification of the Model Parameters

Parameters	Values
V_g, V_l, $F_{g,\mathrm{in}}$, F_l, T, P	Table 1
$x_{g,\mathrm{O}_2,\mathrm{in}}$	0.2095 for air
$x_{g,\mathrm{CO}_2,\mathrm{in}}$	0.00034 for air
$c_{l,\mathrm{O}_2,\mathrm{in}}$	0
$c_{l,\mathrm{CO}_2,\mathrm{in}}$	0
m_{O_2}, m_{CO_2}, $k_l a_{\mathrm{O}_2}$, $k_l a_{\mathrm{CO}_2}$	Discussed in a separate section
$V_{s,i}$, k_{DO}, d_i, d_{DO}	Discussed in a separate section

accurate, then $k_l a_{O_2}$ can accurately be determined from Eq. 11 if the assumption of a steady state is satisfied. From several steady-state cultures, the following average value of $k_l a_{O_2}$ has been obtained

$$k_l a_{O_2} = 5.9 \cdot 10^{-2} \ [s^{-1}] \tag{12}$$

$k_l a_{CO_2}$ cannot be identified in this way since a measurement of the dissolved carbon dioxide concentration is not available. However, it can be related to the value of $k_l a_{O_2}$. Note first that the specific surface area a for both parameters is the same. Besides, since the fermentation broth is well stirred, the penetration film theory may be used to describe the mass transfer. According to this theory the mass-transfer coefficient $k_{l,i}$ is a function of the square root of the diffusion coefficient of component i in liquid and of the properties of the bubbles (van't Riet and Tramper, 1991). Now, since the bubbles are the same both for oxygen transport and carbon dioxide transport, this means that $k_l a_{CO_2}$ follows from

$$\frac{k_l a_{CO_2}}{k_l a_{O_2}} = \left(\frac{D_{l,CO_2}}{D_{l,O_2}} \right)^{1/2} \approx 0.956 \tag{13}$$

where the diffusion coefficients $D_{l,CO_2} = 1.92 \times 10^{-9}$ [$m^2 \cdot s^{-1}$] and $D_{l,O_2} = 2.10 \times 10^{-9}$ [$m^2 \cdot s^{-1}$] are taken from Weast et al. (1989).

Identification of $V_{s,i}$, k_{DO} and the delays d_i and d_{DO}

The parameters $V_{s,i}$, k_{DO} and the delays d_i, d_{DO} correspond to parts of the model that are used to approximate the dynamics of a certain part of the system, and are not physical parameters. Therefore, they are not easily derived from physical knowledge. These parameters are identified by fitting the response of the model to the measured response of the system to a changing $x_{g,O_2,in}$ and $x_{g,CO_2,in}$. Pseudo random binary signals are used as excitation signals for $x_{g,O_2,in}$ and $x_{g,CO_2,in}$, that is, the concentration of either oxygen or carbon dioxide in the inflowing gas is varied between two given levels, where the switching between the low and the high level is determined in a random fashion. The levels for $x_{g,O_2,in}$ and $x_{g,CO_2,in}$ are chosen in the range of the concentrations that are observed during the steady-state and dynamic conditions of the fermenter. The maximum switching time is chosen such that the system is allowed to reach a steady state at some instances, which enables one to identify/verify the delays visually.

An important assumption during these identification experiments is that the OUR and CER remain constant. Since the cells react on changing concentrations of carbon dioxide in the inflowing gas, the identification experiment is conducted in a resting cell culture (the medium feed F_l is stopped such that the cells get into a resting state) for which holds that the OUR and CER are approximately equal to zero.

The oxygen and carbon dioxide concentration measurements of the off-gas analyzer and the DO measurements are logged every second. Since the part of the model described by Eqs. 1–4 is known, the model predictions for x_{g,O_2} and x_{g,CO_2} can be computed. These signals are the inputs to the off-gas system, and, using these signals and the measured data of the off-gas analyzer, the number of ideally stirred tanks n_s to approximate the off-gas system can be estimated using the order detection in subspace identification algorithms, for example, Verhaegen (1993). Using this algorithm, the estimated order then follows from the number of dominant singular values, which are computed from a singular value decomposition of data matrices (Verhaegen, 1993). The results of the order detection are plotted in Figure 2. From Figure 2, it follows that $n_s = 2$, both for oxygen and carbon dioxide. The next step could be to identify two second-order black-box models for the off-gas system. However, in Eq. 5 more structure has been assumed. The reason for this is that $F_{g,out}$ depends on the states during a substrate pulse experiment, discussed in the next section. This knowledge is easily incorporated into the structure in Eq. 5. Thus, the parameters are identified by fitting the response of the model to the identification data. It was observed that this yields almost equal values for V_{1,O_2} and V_{1,CO_2}; therefore, the restriction

$$V_{1,O_2} = V_{1,CO_2} = V_1 \tag{14}$$

was introduced. This restriction is also motivated by the fact that V_1 is related to the volume of the head space, which is the same both for oxygen and for carbon dioxide. The fact that $V_{2,i}$ is gas dependent is motivated from the difference in

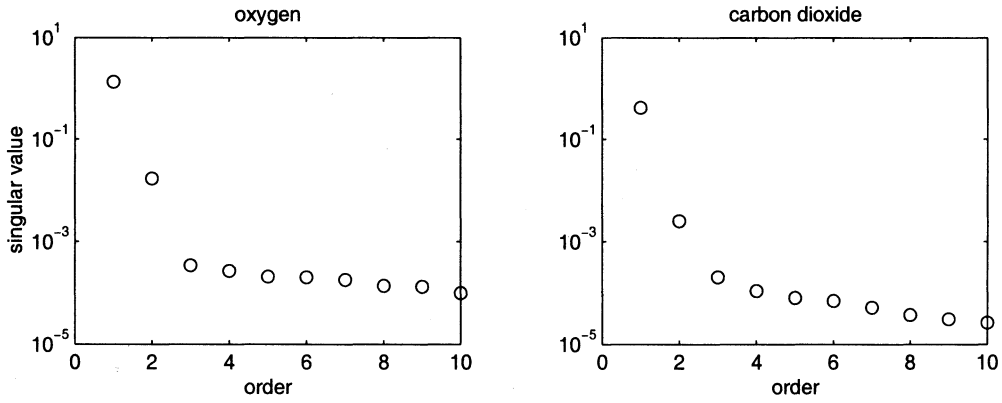

Figure 2. Order selection to estimate n_s.

Table 3. Optimized Model Parameters

Parameter	Value	Unit
d_{O_2}	47	(s)
d_{DO}	5	(s)
d_{CO_2}	36	(s)
V_1	1.92×10^{-3}	(m^3)
V_{2,O_2}	7.75×10^{-4}	(m^3)
V_{2,CO_2}	2.34×10^{-4}	(m^3)
k_{DO}	7.86×10^{-2}	(s^{-1})

measurement cells for oxygen and carbon dioxide. The fit of the response of the optimized model to the identification and validation data is presented in Figure 3. From this figure, it can be concluded that the model describes the data well. The optimized model parameters are given in Table 3.

Gas Production/Consumption

The identification experiment is done on a resting cell culture, for which holds that the OUR and the CER are approximately equal to zero, in which case $F_{g,\text{out}} = F_{g,\text{in}}$. However, during substrate pulse experiments the OUR and CER are nonzero. Besides, the respiration quotient may significantly deviate from 1, resulting in a net gas production or consumption and thus $F_{g,\text{out}} \neq F_{g,\text{in}}$. Gas consumption/production affects the concentrations of both oxygen and carbon dioxide. A mass balance for the total amount of gas in the gas phase of the fermentation broth gives

$$\frac{\partial \frac{V_g P}{RT}}{\partial t} = F_{g,\text{in}}\frac{P}{RT} - F_{g,\text{out}}\frac{P}{RT} - k_l a_{O_2}\cdot V_l\left(\frac{x_{g,O_2}P}{m_{O_2}RT} - c_{l,O_2}\right) - k_l a_{CO_2}\cdot V_l\left(\frac{x_{g,CO_2}P}{m_{CO_2}RT} - c_{l,CO_2}\right) \quad (15)$$

Assuming a constant pressure and assuming a constant volume of the gas phase, the above equation gives

$$0 = F_{g,\text{in}}\frac{P}{RT} - F_{g,\text{out}}\frac{P}{RT} - k_l a_{O_2}\cdot V_l\left(\frac{x_{g,O_2}P}{m_{O_2}RT} - c_{l,O_2}\right) - k_l a_{CO_2}\cdot V_l\left(\frac{x_{g,CO_2}P}{m_{CO_2}RT} - c_{l,CO_2}\right) \quad (16)$$

that is, in Eq. 16 it is assumed that the number of moles of gas in the fermentation broth remains the same, and that uptake from and release to the gas phase immediately leads to a proportional change in gas flow, without affecting the pres-

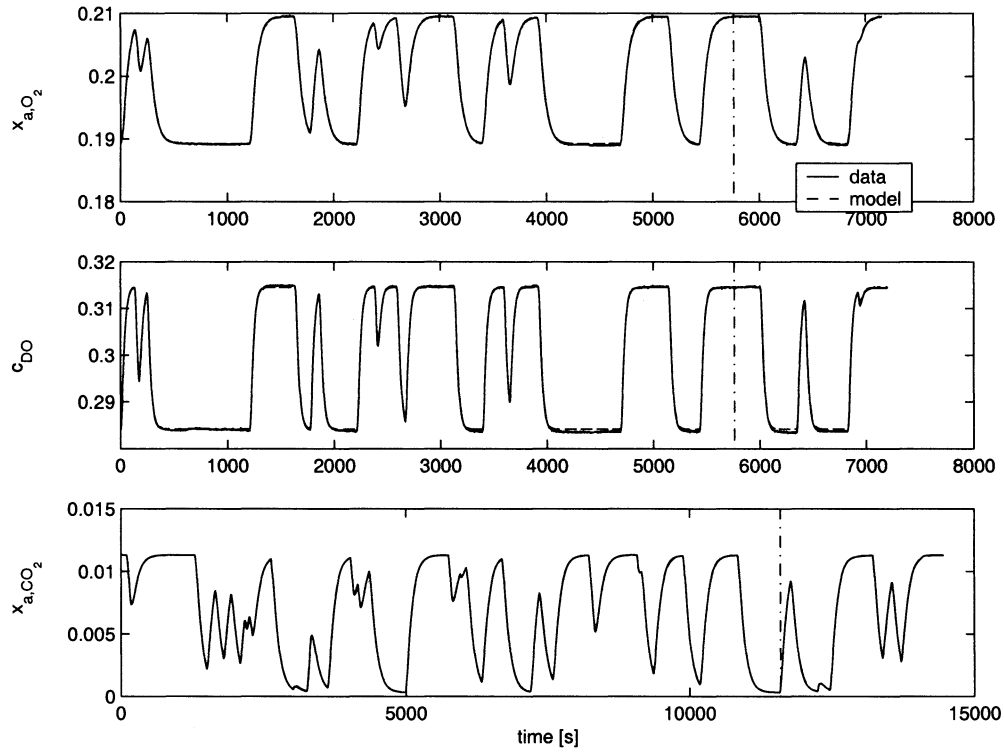

Figure 3. Fit of the optimized model to the identificaton and validation data.

The upper two plots correspond to an identification experiment during which $x_{g,O_2,\text{in}}$ was excited; the lower plot corresponds to an identification experiment during which $x_{g,CO_2,\text{in}}$ was excited. Left of the vertical dashed-dotted lines are the identification data, right of this line are the validation data.

sure and the volume. Thus

$$F_{g,\mathrm{out}} = F_{g,\mathrm{in}} - k_l a_{\mathrm{O}_2} \cdot V_l \left(\frac{x_{\mathrm{g,O}_2}}{m_{\mathrm{O}_2}} - \frac{RT}{P} c_{l,\mathrm{O}_2} \right) - k_l a_{\mathrm{CO}_2} \cdot V_l \left(\frac{x_{g,\mathrm{CO}_2}}{m_{\mathrm{CO}_2}} - \frac{RT}{P} c_{l,\mathrm{CO}_2} \right) \quad (17)$$

To investigate whether the contribution of the correction terms in Eq. 17 is significant, assume that there is a steady state in the sense that the OUR and the CER equal the volumetric mass-transfer rate of oxygen and carbon dioxide, respectively. It is assumed here that the removal via the liquid outflow can be neglected, which is justified since F_l is about a thousand times smaller than $F_{g,\mathrm{in}}$. Then

$$F_{g,\mathrm{out}} = F_{g,\mathrm{in}} - \mathrm{OUR}\frac{V_l RT}{P} + \mathrm{CER}\frac{V_l RT}{P} \quad (18)$$

If $\mathrm{CER} = 3 \cdot \mathrm{OUR}$ and $\mathrm{OUR} = 0.01$ [mol·m^{-3}·s^{-1}], then

$$F_{g,\mathrm{out}} = 4.30 \times 10^{-5} + 0.155 \times 10^{-5} = 4.455 \times 10^{-5} \quad (19)$$

The assumption $F_{g,\mathrm{out}} = F_{g,\mathrm{in}}$ then results in an error for $F_{g,\mathrm{out}}$ of about 3.5%. At first glance, this error seems to be acceptable, however, the main question is how this error affects the estimated OUR and CER. Under the assumption that $F_{g,\mathrm{out}} = F_{g,\mathrm{in}}$, the estimated steady-state OUR ($\mathrm{OUR}_{\mathrm{est}}$) follows from

$$\mathrm{OUR}_{\mathrm{est}} = F_{g,\mathrm{in}} \frac{x_{g,\mathrm{O}_2,\mathrm{in}} - x_{g,\mathrm{O}_2,\mathrm{out}}}{V_l} \frac{P}{RT} \quad (20)$$

whereas the real steady-state OUR is given by

$$\mathrm{OUR} = \frac{x_{g,\mathrm{O}_2,\mathrm{in}} F_{g,\mathrm{in}} - x_{g,\mathrm{O}_2,\mathrm{out}} F_{g,\mathrm{out}}}{V_l} \frac{P}{RT} \quad (21)$$

Now, define the relative error in the estimated OUR by

$$\mathrm{error}_{\mathrm{OUR}} = \frac{\mathrm{OUR} - \mathrm{OUR}_{\mathrm{est}}}{\mathrm{OUR}} \cdot 100\% \quad (22)$$

To proceed, first rewrite Eq. 18

$$F_{g,\mathrm{out}} = F_{g,\mathrm{in}} + \alpha \quad (23)$$

where

$$\alpha = (\mathrm{CER} - \mathrm{OUR})\frac{V_l RT}{P} = (RQ - 1)\mathrm{OUR}\frac{V_l RT}{P} \quad (24)$$

denotes the net production/consumption of gas [m^3·s^{-1}], and $RQ = \mathrm{CER}/\mathrm{OUR}$ is the respiration quotient. Substituting Eqs. 20, 21, and 23 into Eq. 22, the following is derived

$$\mathrm{error}_{\mathrm{OUR}} = \frac{\alpha \left(\mathrm{OUR} \cdot V_l - x_{g,\mathrm{O}_2,\mathrm{in}} \cdot F_{g,\mathrm{in}} \frac{P}{RT} \right)}{(F_{g,\mathrm{in}} + \alpha)\mathrm{OUR} \cdot V_l} \cdot 100\% \quad (25)$$

For the CER, a similar analysis can be made. In Figure 4 the error of the estimated values for the OUR and the CER is plotted as a function of the RQ, if it is assumed that $F_{g,\mathrm{out}} = F_{g,\mathrm{in}}$. Figure 4 demonstrates that the error can be significant, in particular for the estimate of the OUR. This indicates that the effect of gas production/consumption needs to be incorporated in the model in order to be able to accurately reconstruct the OUR. For the estimate of the CER, the error is much smaller, since $x_{g,\mathrm{CO}_2,\mathrm{in}}$ is approximately equal to zero, in which case the error in the estimate of the CER is of the same order of magnitude as the error in $F_{g,\mathrm{out}}$.

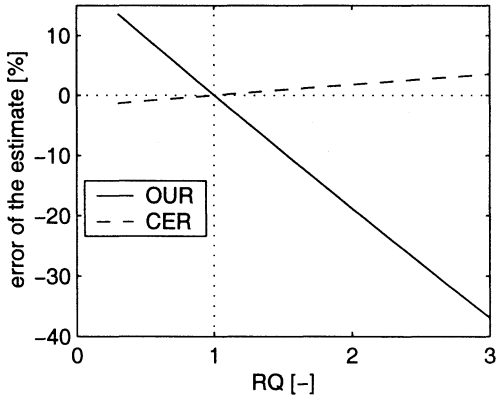

Figure 4. Relative error in the estimated OUR and CER as function of the RQ.

Reconstruction of the OUR and CER

During operation of the fermenter, the following data are logged: the oxygen and carbon dioxide concentrations in the off-gas and the dissolved oxygen concentration. If the fermenter is in a steady state, then the steady state OUR and CER can be calculated from (the subscript *ss* denotes steady state)

$$\mathrm{OUR}_{ss} = \frac{x_{g,\mathrm{O}_2,in} F_{g,\mathrm{in}} - x_{g,\mathrm{O}_2,\mathrm{out}} F_{g,\mathrm{out}}}{V_l} \frac{P}{RT} \quad (26)$$

$$\mathrm{CER}_{ss} = \frac{x_{\mathrm{g,CO}_2,\mathrm{out}} F_{g,\mathrm{out}} - x_{g,\mathrm{CO}_2,\mathrm{in}} F_{g,\mathrm{in}}}{V_l} \frac{P}{RT} \quad (27)$$

If the fermenter is not in a steady state, such as after the addition of a substrate pulse, then Eqs. 26 and 27 are no longer valid. Thus, the problem is how to reconstruct the OUR and the CER from the logged data under nonsteady-state conditions. Note that, besides the data, the dynamic model (Eqs. 1–8) is also available, and this model describes the relation between the unknowns OUR and CER, which are not constant, but may vary with time, and the quantities

x_{a,O_2}, x_{a,CO_2}, and c_{DO} of which the measured values $x_{a,O_2,m}$, $x_{a,CO_2,m}$, and $c_{DO,m}$ are available (the subscript m refers to the measured values). Thus, the problem consists of fitting the model output x_{a,O_2}, x_{a,CO_2}, and c_{DO} to the measurements $x_{a,O_2,m}$, $x_{a,CO_2,m}$, and $c_{DO,m}$ where the OUR and CER are the degrees of freedom. This optimization problem can be stated as

$$\min_{x[0],u[0],\ldots,u[N-1]} \sum_{j=1}^{N} (y[j]-y_m[j])' R_n[j]^{-1} (y[j]-y_m[j]) + \Delta u[j-1]' Q_n[j-1]^{-1} \Delta u[j-1] + (x[0]-x_{ss})' \Pi^{-1} (x[0]-x_{ss}) \quad (28)$$

with

$$y[j] = \left[x_{a,O_2}\left[j+d_{O_2}\right] \;\; c_{DO}[j+d_{DO}] \;\; x_{a,CO_2}\left[j+d_{CO_2}\right] \right]' = \left[x_{2,O_2}[j] \;\; c_{s,O_2}[j] \;\; x_{2,CO_2}[j] \right]' \quad (29)$$

$$y_m[j] = \left[x_{a,O_2,m}\left[j+d_{O_2}\right] \;\; c_{DO,m}[j+d_{DO}] \;\; x_{a,CO_2,m}\left[j+d_{CO_2}\right] \right]' \quad (30)$$

$$u = [\mathrm{OUR} \;\; \mathrm{CER}]' \quad (31)$$

$$\Delta u[j] = u[j] - u[j-1] \quad (32)$$

$$x = [x_{g,O_2} \;\; c_{l,O_2} \;\; x_{1,O_2} \;\; x_{2,O_2} \;\; c_{s,O_2} \;\; x_{g,CO_2} \;\; c_{l,CO_2} \;\; x_{1,CO_2} \;\; x_{2,CO_2}]' \quad (33)$$

j denotes the sampling instant, N is the window size, the prime denotes the transpose and R_n, Q_n and Π are positive definite weighting matrices. The first quadratic term in Eq. 28 penalizes the error between the measurements and the model output. The second quadratic term in Eq. 28 penalizes the changes in the OUR and the CER, which ensures that the problem is well behaved and is used to prevent the model from tracking the measurement noise in y_m. Comparing the first quadratic term and the second quadratic term, a time lag of one sampling instant is observed. This is due to the fact that there is no direct feed-through term present in the model (that is, $u[j]$ does not affect $y[j]$). The third term in Eq. 28 penalizes the deviation of the initial state from the (computed) steady state prior to the addition of the substrate pulse. To get rid of initialization errors in x_{ss} and $u[-1]$, the minimization in Eq. 28 is done over a window that starts several sampling instants before the addition of the substrate pulse. The optimization in Eq. 28 is nonconvex due to the nonlinearity of the model caused by Eq. 17. Moreover, N is usually large since the sampling time is small (1 s) and the objective is to reconstruct the OUR and the CER over a horizon of about 300 s after the addition of the substrate pulse. Consequently, directly solving Eq. 28 involves a high computational demand (due to a large number of degrees of freedom).

The computational complexity of solving Eq. 28 can be alleviated by recasting Eq. 28 as a smoothing problem and exploiting the properties of the smoothing problem (see Kailath (2000)). Due to the discrete-time nature of the sampled data, the continuous-time time-invariant nonlinear mass-transport model (Eqs. 1–8) with Eq. 17 is approximated by a discrete-time time-varying linear model, which is obtained by linearizing (and discretizing) the time-invariant nonlinear model. Moreover, following Eq. 32, the OUR and the CER are modeled as state equations

$$\mathrm{OUR}[j+1] = \mathrm{OUR}[j] + w_1 \quad (34)$$

$$\mathrm{CER}[j+1] = \mathrm{CER}[j] + w_2 \quad (35)$$

The discrete-time time-varying linear model can be represented by

$$z[j+1] = A[j]z[j] + G[j]w[j] + K[j] \quad (36)$$

$$y[j] = Cz[j] + v[j] \quad (37)$$

where the augmented state z represents

$$z = [x' \;\; \mathrm{OUR} \;\; \mathrm{CER}]' \quad (38)$$

In Eq. 36 $w = [w_1 \;\; w_2]'$ is regarded as process noise, and a measurement-noise term v has been added in Eq. 37. $K[j]$ is a time-varying off-set term as a result of the linearization and also includes the effects of the inputs $x_{g,O_2,\mathrm{in}}$ and $x_{g,CO_2,\mathrm{in}}$, which are constant during the pulse experiments. For the model in this article, C is a constant matrix, and, therefore, lacks the time index. In order to get the most accurate approximation of the nonlinear model, at each time instant j the model parameters $A[j]$, $G[j]$, $K[j]$ are recomputed through a relinearization (and discretization) of the nonlinear model.

Now, the problem of solving Eq. 28, where u are the degrees of freedom, can be recast as a problem of determining the least-mean-squares estimator $\hat{z}$ for the state z, where the reconstructed OUR and CER follow from the estimated state $\hat{z}$ (see Eq. 38). This is a so-called smoothing problem for the state z: the problem of determining the least-mean-squares estimator $\hat{z}[j|N]$ of $z[j]$ given the data $\{y_m[1], \ldots, y_m[N]\}$. The second argument N in $\hat{z}[j|N]$ is used to express the estimate of $z[j]$ based on information/data up to time N. This smoothed estimator can be computed from a backwards-time recursion and knowledge of the prediction estimator $\hat{z}[j|j-1]$, which is the least-mean-squares estimator of $z[j]$ given the data $\{y_m[1], \ldots, y_m[j-1]\}$. The latter ($\hat{z}[j|j-1]$) can be computed from forwards-time recursions, also known as Extended Kalman Filtering. Thus, the smoothing problem can be solved in two steps:

(1) Forwards-time recursions to find $\hat{z}[j|j-1]$ (see the next subsection).

(2) Backwards-time recursions to find $\hat{z}[j|N]$, using knowledge of $\hat{z}[j|j-1]$ (see the second subsection).

For the theoretical background behind the formulas in the next two subsections refer to Kailath et al. (2000), for example.

Forwards-time recursions: Extended Kalman Filtering

Since v and w are nonmeasurable and are assumed to be zero-mean white noise signals, the prediction model for the

estimate $\hat{z}$ of z is given by

$$\hat{z}[j+1|j] = A[j]\hat{z}[j|j] + K[j] \quad (39)$$

$$\hat{y}[j|j] = C\hat{z}[j|j] \quad (40)$$

When at time instant j a new measurement is available, the observation error e can be computed from

$$e[j] = y_m[j] - C\hat{z}[j|j-1] \quad (41)$$

This observed error is used to update the estimated state

$$\hat{z}[j|j] = \hat{z}[j|j-1] + L[j]e[j] \quad (42)$$

(with $\hat{z}[1|0] = z_{ss}$, since the first data are from the steady state before the addition of the substrate pulse), where $L[j]$ follows from the recursion (Kailath, 2000)

$$L[j] = \Phi[j|j-1]C'\left(C\Phi[j|j-1]C' + R_n[j]\right)^{-1} \quad (43)$$

$$\Phi[j|j] = (I - L[j]C)\Phi[j|j-1] \quad (44)$$

$$\Phi[j+1|j] = A[j]\Phi[j|j]A[j]' + G[j]Q_n[j]G[j]' \quad (45)$$

where the weighting matrices R_n and Q_n are the covariance matrices of the measurement noise and process noise, respectively. The non-negative-definite symmetric covariance matrix $\Phi[1|0]$ needs to be specified to initialize the recursion. One possibility is to choose $\Phi[1|0] = \Phi^*$, where Φ^* is the solution to the algebraic Riccati equation

$$\Phi^* = A[1]\Phi^*A[1]' + G[1]Q_n[1]G[1]' - A[1]\Phi^*C'\left(C\Phi^*C' + R_n[1]\right)^{-1}C\Phi^*A[1]' \quad (46)$$

Alternative choices for $\Phi[1|0]$ are possible as well, as long as $\Phi[1|0] \geq 0$ (non-negative-definite). The motivation to choose $\Phi[1|0] = \Phi^*$ comes from the fact that in steady state, that is, when $A[j]$, $G[j]$, $R_n[j]$ and $Q_n[j]$ are constant, the recursions for Φ in Eqs. 44–45 exponentially tend to Φ^* for any non-negative-definite $\Phi[1|0]$ (Kailath, 2000). To obtain $A[j]$, $G[j]$, and $K[j]$ for the recursion in Eq. 45 and for the prediction of the estimated state in Eq. 39, the nonlinear continuous-time model is linearized (and discretized) at the linearization point $\hat{z}[j|j]$, which corresponds to the best estimate of the current state, based on the information available up to the current sampling time.

Backwards-time recursions: the Bryson-Frazier formulas

The prediction estimator $\hat{z}[j|j-1]$ only uses information up to time $j-1$, whereas information over the entire window is available. To incorporate the knowledge of the future time instants from j until N, the smoothed estimate $\hat{z}[j|N]$ can easily be computed using knowledge of $\hat{z}[j|j-1]$ (Kailath, 2000)

$$\hat{z}[j|N] = \hat{z}[j|j-1] + \Phi[j|j]\lambda[j|N] \quad (47)$$

where $\lambda[j|N]$ follows from the backwards-time recursion

$$\lambda[j|N] = (A[j] - L[j]C)'\lambda[j+1|N] + C'R_n[j]^{-1}e[j] \quad (48)$$

with $\lambda[N+1|N] = 0$. An important motivation for computing these backwards-time recursions is that the forwards-time recursions of the Extended Kalman Filter discussed earlier in this section would introduce a time lag in the reconstructed OUR and CER, if these were to be computed from the estimates $\hat{z}[j|j-1]$. This is because these forwards-time recursions consist of a pure feedback mechanism: the estimated state is only updated after the error has occurred. The backwards-time recursions by means of the Bryson-Frazier formulas (Eqs. 47–48) remove this time lag in the estimates $\hat{z}[j|N]$, and, consequently, in the reconstructed OUR and CER, which follow from $\hat{z}[j|N]$.

Results

After the data $y_m[j]$ for $j = 1, \ldots, N$ have been collected, the OUR and CER can be reconstructed using Eqs. 39–48 once $Q_n[j]$ and $R_n[j]$ have been specified. $R_n[j]$ is the covariance matrix of the measurement noise v and can be computed from the measurements y_m over a certain horizon prior to the addition of the substrate pulse, that is, when the fermenter is in a steady state. $Q_n[j]$ is the assumed covariance matrix of the process noise w, which is used as a tuning parameter to tune the trade-off between:

- fitting the measurements y_m with the output $C\hat{z}[j|N]$ of the smoothing algorithm; and
- obtaining "smooth" estimates for the OUR and CER.

That is, for $Q_n[j]$ small, the smoothing algorithm will tend to small perturbations of the estimated OUR and CER since then it is assumed that the covariance of the process noise w is small. This may result in a bad fit of the measurements if the real OUR and CER are prone to heavier changes. On the contrary, for $Q_n[j]$ large, the smoothing algorithm will tend to more severe perturbations of the estimated OUR and CER. This may cause the smoothing algorithm to try to fit the measurement noise by means of the process noise, which will result in noisy estimates for the OUR and the CER. A proper tuning is found by trial and error, where the quality of Q_n is evaluated by means of a visual inspection of the fit of the measurements and by means of a visual inspection of the reconstructed OUR and CER.

The time indices of $R_n[j]$ and $Q_n[j]$ indicate that these covariance matrices may be time-varying. The measurement noise is assumed to have a constant covariance during the experiment, so

$$R_n[j] = R_{n,ss} \qquad \forall j = 1, \ldots, N \quad (49)$$

where $R_{n,ss}$ is the covariance matrix of the measurements taken from the steady state before the addition of the substrate pulse. For $Q_n[j]$, the story is different. Before the addition of the substrate pulse, when the fermenter is in a steady state, accurate steady-state values for the OUR and CER can be determined. Thus, before the addition of the substrate pulse, $Q_n[j]$ is preferably taken small to express a high confi-

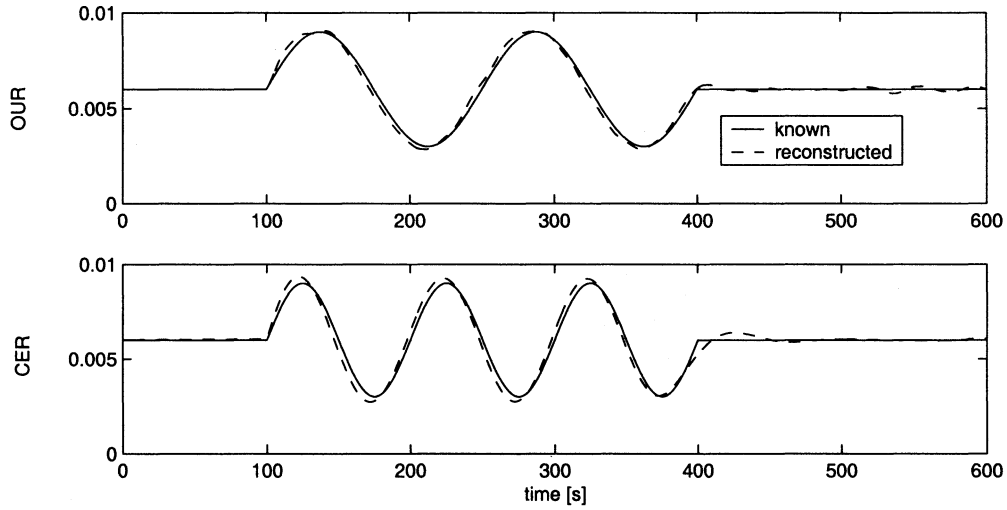

Figure 5. Known and reconstructed OUR and CER for the simulated data.

dence in the estimated steady-state values for the OUR and CER. On the other hand, after the addition of the substrate pulse, the OUR and CER are likely to change, and, thus, one should use larger values for $Q_n[j]$. In this case the following is chosen

$$Q_n[j] = \frac{1}{1000} Q_{n,p} \qquad \forall j = 1, \ldots, j_p$$

$$Q_n[j] = Q_{n,p} \qquad \forall j = j_p, \ldots, N \tag{50}$$

where $Q_{n,p}$ is the assumed covariance matrix of the process noise during the pulse experiment, which is used as a tuning parameter, and j_p is the time index at which the substrate pulse is added to the fermenter.

First, in the next subsection the performance of the smoothing algorithm is investigated on simulated data. This provides an opportunity to check whether the reconstructed OUR and CER correspond to those used for generating the data. Secondly, the results of the smoothing algorithm to reconstruct the OUR and CER from data obtained from real-life pulse experiments is presented.

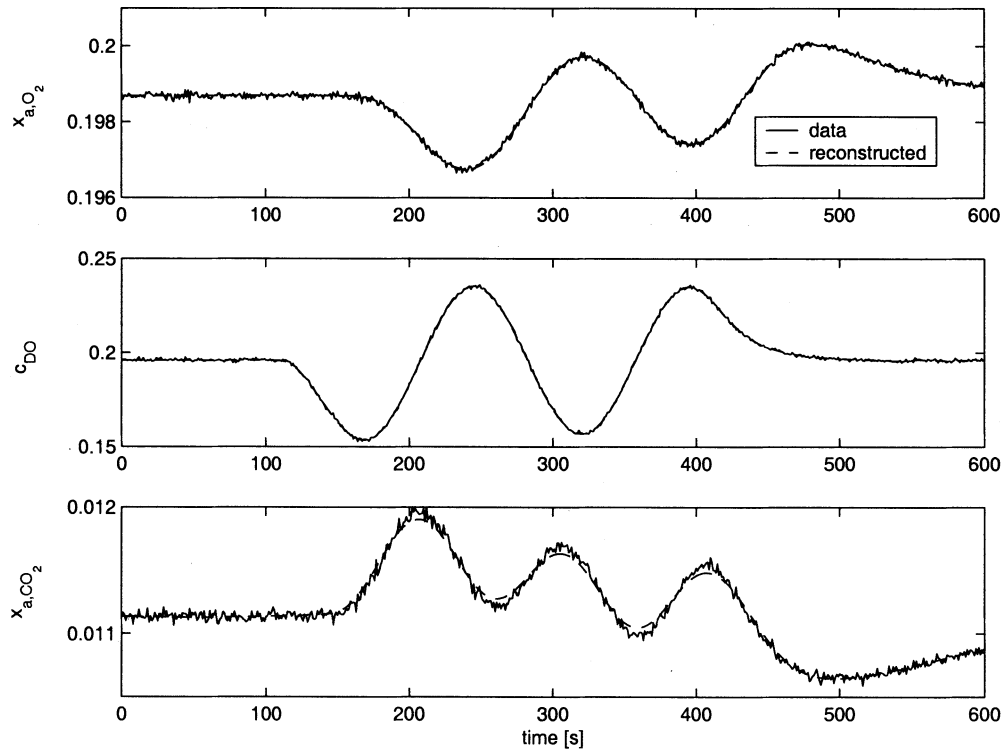

Figure 6. Simulated data and the reconstructed outputs of the smoothing algorithm.

Simulation study

To investigate the performance of the smoothing algorithm in the nominal case, that is, when the model is perfect, the algorithm is used for data that are generated by simulating the model with a known OUR and CER. The input signals OUR and CER used to simulate the model consist of constant signals superimposed by sinusoids from time instant $j_p = 100$ during a window of 300 s, respectively. These signals are given in Figure 5 (note that these signals are not motivated from a physical background, but only serve to validate the smoothing algorithm). Normally distributed measurement noise with a covariance matrix of

$$R_{n,ss} = \begin{bmatrix} 0.036 & 0 & 0 \\ 0 & 3.6 & 0 \\ 0 & 0 & 0.009 \end{bmatrix} \cdot 10^{-5} \tag{51}$$

was added to the simulated outputs, respectively. This covariance matrix is derived from measurements under steady-state conditions of the fermenter. In Figure 6 the simulated outputs are given. These outputs are used for the smoothing algorithm to reconstruct the OUR and CER. $Q_{n,p}$ was set to

$$Q_{n,p} = \begin{bmatrix} 3.6 & 0 \\ 0 & 1.8 \end{bmatrix} \cdot 10^{-7} \tag{52}$$

The output of the smoothed estimator, that is, $C\hat{z}[j|N]$ shifted with the respective delays, is plotted in Figure 6. From Figure 6, it is clear that the output of the smoothed estimator closely fits the simulated data. The reconstructed signals OUR and CER are plotted in Figure 5, from which it is clear that the smoothing algorithm is able to accurately reconstruct the known OUR and CER.

Pulse experiments with Saccharomyces cerevisiae

A pulse experiment consists of perturbing the chemostat culture (described in the section about parameter identification) with a sudden increase of the growth limiting substrate concentration. In the case of a glucose pulse, a small volume of a concentrated glucose solution is added to the fermenter such that the glucose concentration is abruptly raised to 1 [kg·m^{-3}] in the fermentation broth (the steady-state glucose concentration of the chemostat is about 50 times smaller); in the case of an ethanol pulse, the ethanol concentration is abruptly raised to 0.5 [kg·m^{-3}] in the fermentation broth. The off-gas data and DO data are logged every second and are used to reconstruct the OUR and the CER.

In the ideal case, the only mismatch between the model and the true process consists of the lack of knowledge of the OUR and the CER for the model. Then, the smoothing algorithm can accurately reconstruct these signals, as demonstrated in the previous subsection. Although the model closely describes the behavior of the true process (see Figure 3), additional sources of model mismatch may be present. Since the process noise w_1 and w_2 are assumed to only affect the states corresponding to the OUR and the CER (Eqs. 34–35), the smoothing algorithm will try to explain any additional model mismatch in terms of the OUR and the CER. Unfortunately, the "real" OUR and CER are not available, like in a simulation environment, so the performance of the smoothing algorithm must be evaluated differently. This can be done by a visual inspection of the fit of the measured data, and, by analyzing the residuals e, that is, the error between the measurements and the model predictions, as defined in Eq. 41. If the data cannot be explained in terms of the OUR and the CER, this will result in a bad fit of the model, which affects the residuals e. The basic assumption of the smoothing algorithm is that e is a zero-mean white-noise signal (Kailath, 2000), whereas a bad fit of the model may result, for example, in a residual with an offset. Using the smoothing algorithm to reconstruct the OUR and CER from off-gas data measured during a pulse experiment, under the assumption that only the states OUR and CER are corrupted with process noise, resulted in a significant error between the

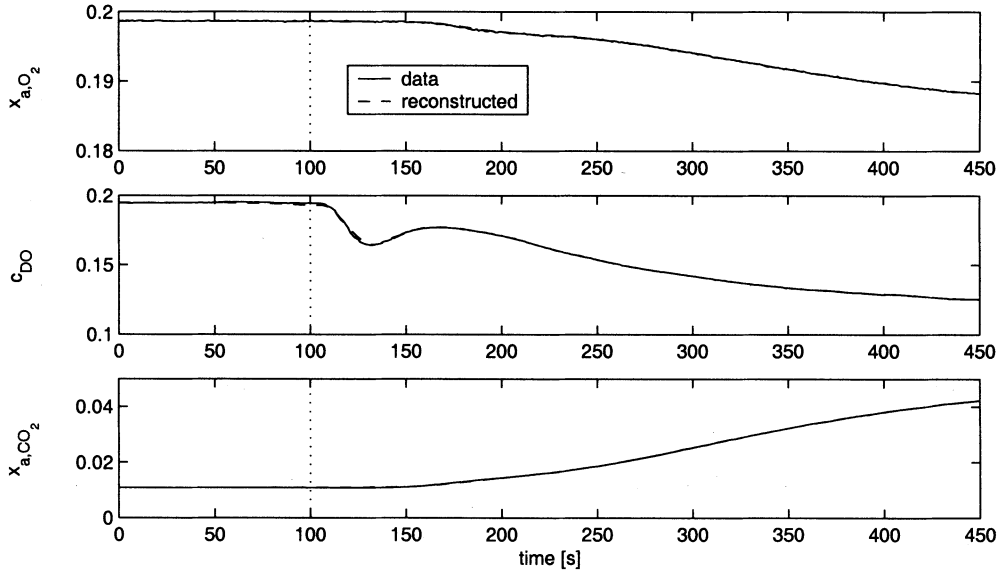

Figure 7. Measured data and the reconstructed output of the smoothing algorithm in case of a glucose pulse.

At $j_p = 100$ (dotted line), the substrate pulse is added.

smoothed estimator and the data; it was observed that it was hard to fit both the off-gas oxygen data and the dissolved oxygen data simultaneously (data not shown). A physical explanation for the fact that it is not possible to fit both the oxygen concentration of the off-gas and that of the fermentation broth (DO) may be due to an error in the $k_l a_{O_2}$. To understand this, consider a steady state for simplicity. Then, the OUR can be computed from Eq. 21. In a steady state also Eq. 2 should be equal to zero, which gives

$$\text{OUR} = -\frac{F_l}{V_l} \cdot c_{l,O_2} + k_l a_{O_2} \left(\frac{P}{m_{O_2} RT} x_{g,O_2} - c_{l,O_2} \right) \quad (53)$$

Equation 21 gives a value for the OUR, but, if the $k_l a_{O_2}$ is incorrect, then Eq. 53 cannot hold simultaneously. Depending on the weighting matrix R_n, the smoothing algorithm will balance the error for x_{g,O_2} and c_{l,O_2}, which, in steady state, are equal to x_{a,O_2} and c_{DO}, respectively. To prevent this, $k_l a_{O_2}$ may be adapted such that Eqs. 21 and 53 will hold simultaneously. To compensate for an error in $k_l a_{O_2}$, its value may be estimated through the use of the smoothing algorithm. Similar to the OUR and CER (Eqs. 34 and 35), $k_l a_{O_2}$ can be modeled by

$$k_l a_{O_2}[j+1] = k_l a_{O_2}[j] + w_3 \quad (54)$$

that is, $k_l a_{O_2}$ is now considered as a state of the model. Using this difference equation for $k_l a_{O_2}$, the smoothing algorithm is now able to reconstruct the OUR, CER and $k_l a_{O_2}$, where the covariance matrix of the process noise was set to

$$Q_n[j] = \begin{bmatrix} \frac{1}{1{,}000} Q_{n,p} & 0 \\ 0 & 1 \times 10^{-7} \end{bmatrix} \quad \forall j = 1, \ldots, j_p$$

$$Q_n[j] = \begin{bmatrix} Q_{n,p} & 0 \\ 0 & 1 \times 10^{-7} \end{bmatrix} \quad \forall j = j_p + 1, \ldots, N \quad (55)$$

which gave good results. Note that the covariance of w_3 is assumed to be constant over the entire horizon. The reason for a relatively large covariance of w_3 prior to the addition of the substrate pulse is that this provides a way to allow for a steady state $k_l a_{O_2}$ that is slightly different from the one in Eq. 12, which is an average value determined from several steady states. In Figure 7 an example of the fit of the data by the smoothed estimator is shown. Clearly, the estimator is able to accurately track the measured responses. The reconstructed OUR, CER, and $k_l a_{O_2}$ for both a glucose pulse experiment and an ethanol pulse experiment are plotted in Figure 8. The differences between the responses due to the glucose pulse and those due to the ethanol pulse in Figure 8 are caused by the fact that these two substrates enter the metabolism of *Saccharomyces cerevisiae* at different points. Thus, as explained in the introduction, these different responses for different substrate pulses give insight in the stoichiometry and kinetics of the metabolism of the organism under investigation, see also Visser (2002).

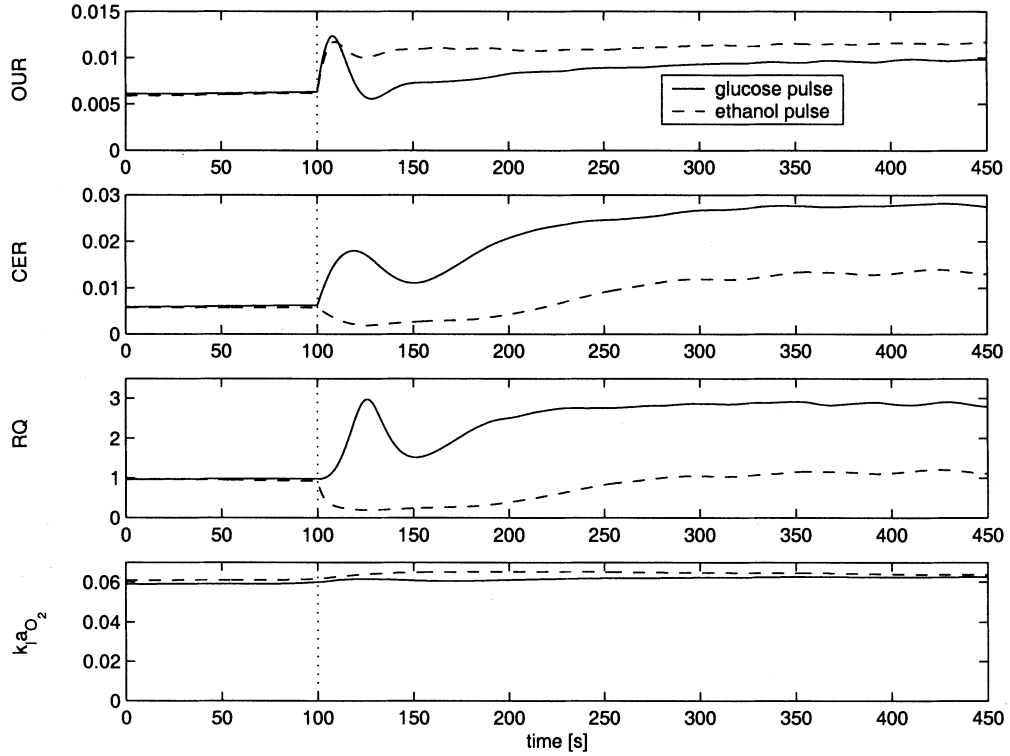

Figure 8. Reconstructed OUR, CER, RQ and $k_l a_{O_2}$ from a glucose pulse experiment and an ethanol pulse experiment.

At $j_p = 100$ (dotted line), the substrate pulse is added.

Note that the $k_l a_{O_2}$ remains fairly constant in Figure 8, and only shows a small increase after the addition of the substrate pulse. This effect may be attributed to the increased ethanol concentration, where it is noted that after the glucose pulse, the ethanol concentration increases due to ethanol production by the yeast. Production of ethanol can be derived from the increased *RQ* in Figure 8 for the glucose pulse. The surface tension decreases with an increasing ethanol concentration, which consequently enhances the mass transport (Wong and Shiuan, 1986).

Discussion

In this article an algorithm is presented that is used to reconstruct the dynamics of the OUR and CER during highly dynamic conditions (like during a pulse experiment), using the dissolved oxygen measurements and the oxygen and carbon dioxide concentration measurements of the off-gas. The dynamics of the OUR and CER is reconstructed by fitting a model for the mass transport to these measured data, using smoothing techniques. In a simulation study it is demonstrated that the algorithm performs well in the nominal case. The reconstructed OUR and CER during a pulse experiment, along with the measurements of intra- and extra-cellular metabolites obtained through rapid sampling techniques (Lange et al., 2001; Schaefer et al., 1999; Theobald et al., 1997; Visser et al., 2002), can be used to model the *in-vivo* kinetics of the primary metabolism (Buchholz et al., 2002; Buziol et al., 2002; Rizzi et al., 1997; Schmitz et al., 2002; Theobald et al., 1997; Vaseghi et al., 1999; Visser, 2002; Visser et al., 2000). The use of different substrates in the pulse experiments provides a possibility to excite the metabolism at different points. As demonstrated, this causes remarkable differences in the responses of the OUR and CER, and this knowledge is valuable for the purpose of metabolic modeling.

The presented algorithm takes the effect of gas production/consumption into account by modeling the flow rate of the outlet gas as a function of the states of the model. This connects the mass transport phenomena of oxygen and carbon dioxide in the model. It has been demonstrated that the net gas production/consumption has a significant impact on the estimate for the OUR and, therefore, needs to be included into the model.

The algorithm presented in this article uses both off-gas data and DO data. The main advantage of the off-gas analyzer over the DO sensor is that it can be re-calibrated at any time during the course of a fermentation, whereas the DO sensor is only calibrated prior to sterilization of the fermenter. In fact, it sometimes may happen that the DO sensor becomes useless after sterilization or due to fouling of the membrane during a fermentation. If, somehow, the DO sensor appears to be unreliable or useless during a pulse experiment, the OUR and CER can still be reconstructed, using only the off-gas data (the part of the model that describes the DO sensor dynamics (Eqs. 7–8) can then be discarded). Of course, if the DO sensor is working properly, the information of this sensor should be used as well. The advantage of the DO sensor over the off-gas analyzer is that it is located "closer" to the OUR: the fast dynamics in the OUR are observed easier from the DO sensor data, whereas the mixing effects in the head space and tubing to the analyzer tend to filter out the dynamics of the OUR. Besides, it is noted that the algorithm that is presented in this article may also be used for other applications, such as, if one is interested in the carbon dioxide concentration and no dissolved carbon dioxide sensor is available (Royce and Thornhill, 1991), or if one is interested in monitoring the $k_l a$.

Literature Cited

Bailey, J. E., "Mathematical Modelling and Analysis in Biochemical Engineering: Past Accomplishments and Future Opportunities," *Biotechnol. Prog.*, **14**, 8 (1998).

Buchholz, A., J. Hurlebaus, C. Wandrey, and R. Takors, "Metabolomics: Quantification of Intracellular Metabolite Dynamics," *Biomolecular Eng.*, **19**, 5 (2002).

Buziol, S., I. Bashir, A. Baumeister, W. Claassen, N. Noisommit-Rizzi, W. Mailinger, and M. Reuss, "New Bioreactor-Coupled Rapid Stopped-Flow Sampling Technique for Measurements of Metabolite Dynamics on a Subsecond Time Scale," *Biotechnol. and Bioeng.*, **80**, 636 (2002).

Dekkers, R. M., "State Estimation of a Fed-Batch Baker's Yeast Fermentation," *Modelling and Control of Biotechnical Processes*, Helsinki, Finland, published by Pergamon, Oxford, U.K., pp. 201–211 (1982).

Gombert, A. K., and J. Nielsen, "Mathematical Modelling of Metabolism," *Current Opinion in Biotechnology*, **11**, 180 (2000).

Janssen, L. P. B. M., and M. M. C. G. Warmoeskerken, *Transport Phenomena Data Companion*, Arnold, London (1987).

Kailath, T., A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, NJ (2000).

Lange, H. C., M. Eman, G. van Zuijlen, D. Visser, J. C. van Dam, J. Frank, M. J. Teixeira de Mattos, and J. J. Heijnen, "Improved Rapid Sampling for in vivo Kinetics of Intracellular Metabolites in Saccharomyces cerevisiae," *Biotechnol. and Bioeng.*, **75**(4), 406 (2001).

Lindberg, C.-F., and B. Carlsson, "Estimation of the Respiration Rate and Oxygen Transfer Function Utilizing a Slow DO Sensor," *Water Sci. and Technol.*, **33**, 325 (1996).

Lubenova, V. N., "Stable Adaptive Algorithm for Simultaneous Estimation of Time-Varying Parameters and State Variables in Aerobic Bioprocesses," *Bioprocess Eng.*, **21**, 219 (1999).

Rizzi, M., M. Baltes, U. Theobald, and M. Reuss, "*In vivo* Analysis of Metabolic Dynamics in Saccharomyces cerevisiae: II. Mathematical Model," *Biotechnol. and Bioeng.*, **55**(4), 592 (1997).

Royce, P. N. C., and N. F. Thornhill, "Estimation of Dissolved Carbon Dioxide Concentrations in Aerobic Fermentions," *AIChE J.*, **37**, 1680 (1991).

Schaefer, U., W. Boos, R. Takors, and D. Weuster-Botz, "Automated Sampling Device for Monitoring Intracellular Metabolite Dynamics," *Analytical Biochemistry*, **270**, 88 (1999).

Schmitz, M., E. Hirsch, J. Bongaerts, and R. Takors, "Pulse Experiments as a Prerequisite for the Quantification of in vivo Enzyme Kinetics in Aromatic Amino Acid Pathway of Escherichia coli," *Biotechnol. Prog.*, **18**(5), 935 (2002).

Schneider, K., and K. Frischknecht, "Determination of the O_2 and CO_2 $K_L a$ Values in Fermenters with the Dynamic Method Measuring the Step Responses in the Gas Phase," *J. of Appl. Chemistry and Biotechnol.*, **27**, 631 (1977).

Theobald, U., W. Mailinger, M. Baltes, M. Rizzi, and M. Reuss, "In vivo Analysis of Metabolic Dynamics in Saccharomyces cerevisiae: I. Experimental Observations," *Biotechnol. and Bioeng.*, **55**(2), 305 (1997).

van't Riet, K., and J. Tramper, *Basic Bioreactor Design*, Marcel Dekker, New York (1991).

Vanrolleghem, P. A., and H. Spanjers, "A Hybrid Respirometric Method for More Reliable Assessment of Activated Sludge Model Parameter," *Water Sci. and Technol.*, **37**, 237 (1998).

Vaseghi, S., A. Baumeister, M. Rizzi, and M. Reuss, "*In vivo* Dynamics of the Pentose Phosphate Pathway in Saccharomyces cerevisiae," *Metabolic Eng.*, **1**, 128 (1999).

Verhaegen, M., "Subspace Model Identification. Part 3. Analysis of the Ordinary Output-Error State-Space Model Identification Algorithm," *Int. J. of Control*, **58**(3), 555 (1993).

Visser, D., "Measuring and Modeling in vivo Kinetics of Primary Metabolism," PhD Thesis, Delft University of Technology (2002).
Visser, D., R. van der Heijden, K. Mauch, M. Reuss, and S. Heijnen, "Tendency Modeling: A New Approach to Obtain Simplified Kinetic Models of Metabolism Applied to Saccharomyces cerevisiae," *Metabolic Eng.*, **2**, 252 (2000).
Visser, D., G. A. van Zuylen, J. C. van Dam, A. Oudshoorn, M. R. Eman, C. Ras, W. M. van Gulik, J. Frank, G. W. van Dedem, and J. J. Heijnen, "Rapid Sampling for Analysis of in vivo Kinetics using the BioScope: A System for Continuous-Pulse Experiments," *Biotechnol. and Bioeng.*, **79**, 674 (2002).
Weast, R. C., D. R. Lide, M. J. Astle, and W. H. Beyer, *CRC Handbook of Chemistry and Physics: A Ready-Reference Book of Chemical and Physical Data 1989–1990*, 70th ed., CRC Press, Boca Raton (1989).
Wong, C. W., and J. H. Shiuan, "Effect of Additives on Mass Transfer in an Aerated Mixing Vessel," *Chem. Eng. Commun.*, **43**, 133 (1986).
Wu, L., H. C. Lange, W. M. van Gulik, and J. J. Heijnen, "Determination of in vivo Oxygen Uptake and Carbon Dioxide Evolution Rates from Off-Gas Measurements under Highly Dynamic Conditions," *Biotechnol. and Bioeng.*, **81**, 448 (2003).
Zhong, Q.-Q., and K. G. Bellgardt, "Dynamic Estimation and Control of the Specific Oxygen Uptake Rate in Processes with Oxygen Limitation," *DECHEMA Biotechnol. Conf.*, Frankfurt am Main, Germany, p. 709 (1990).

Manuscript received Jan. 15, 2003, and final revision received Mar. 19, 2003.