A FAST COMPUTATIONAL PROCEDURE FOR THE PREDETERMINATION OF PARAMETERS IN NON-LINEAR BIOPROCESS MODELS

Jens E. Haag * **Alain Vande Wouwer** *,1

* *Service d'Automatique, Faculté Polytechnique de Mons, Boulevard Dolez 31, B-7000 Mons, Belgium*

Abstract: In this study, a semi-analytical computational procedure, which allows bioprocess model parameters to be quickly evaluated from experimental data, is developed and illustrated with an application example. The approximate model can be used to investigate the qualitative behavior of key components of interest and to check modeling assumptions. The approximate model parameter values can also be used as starting points for more rigorous identification methods.

Keywords: modeling, parameter estimation, biotechnology

1. INTRODUCTION

Modeling and parameter identification of bioprocesses are notoriously difficult tasks, due to the following issues:

- the lack of experimental data, i.e. the number of samples and measurements is often restricted by time- and money-consuming laboratory analysis;
- the difficulty to perform experiments in a reproductible way, due to the large number of factors influencing cell growth;
- the lack of a priori knowledge on the biological reaction scheme, in terms of yield coefficients and kinetics;
- the potential model complexity, i.e. the potentially large number of involved components and number of reactions, and nonlinearity (mostly involved in the kinetics).

Once a model structure and parametrisation has been selected, it is required to infer the numerical values of the model parameters from experimental data. Recently, a systematic identification procedure has been proposed in (Bogaerts and Hanus, 2000), which is based on the state transformation originally introduced in (Bastin and Dochain, 1990), and which allows the yield coefficients to be estimated independently from the kinetic parameters. This procedure takes all the measurement errors into account through the formulation of maximum-likelihood criteria. Even though very efficient, this procedure can sometimes be time-consuming, and there is still a need for fast data evaluation and modeling procedure.

In this connection, this study aims at developing a semi-analytical procedure which allows a fast predetermination of model parameters. To allow analytical developments to be performed, some simplifying assumptions are required, which makes the problem solution approximate. However, a first evaluation of the model parameters can be very useful to check basic modeling assumptions, and to investigate the qualitative behavior of the key components of interest. In addition, the estimated parameter values can serve as starting points for nonlinear parameter identification approaches as those proposed in (Bogaerts and Hanus, 2000).

This paper is organized as follows. The next section deals with the devolopment of a computational procedure for the predetermination of parameters in

[1] Corresponding author, fax: +32 65 37 4136, e-mail: Alain.VandeWouwer@FPMs.ac.be

bioprocess models whereas section 3 summarizes the results. Section 4 presents a simple numerical application. Finally, section 5 is devoted to some concluding remarks.

2. PREDETERMINATION OF MODEL PARAMETERS

In order to develop an analytical procedure, we will make the assumption that the specific production/consumption, growth and death rates are time-invariant in a finite time interval, e.g. the first sampling interval $[t_0, t_1]$ (in bioprocess applications, this time interval can be relatively long, e.g. a few hours). This assumption allows the following dynamic model to be written

$$\dot{c} = \bar{\nu} x_\mathrm{v} + D(c_\mathrm{f} - c); \qquad c(t_0) = c_0 \tag{1}$$

$$\dot{x}_\mathrm{v} = \left(\bar{\mu} - \bar{k}_\mathrm{d}\right) x_\mathrm{v} - D_x x_\mathrm{v}; \qquad x_\mathrm{v}(t_0) = x_{\mathrm{v}0} \tag{2}$$

$$\dot{x}_\mathrm{d} = \bar{k}_\mathrm{d} x_\mathrm{v} - D_x x_\mathrm{d}; \qquad x_\mathrm{d}(t_0) = x_{\mathrm{d}0} \tag{3}$$

In these equations, the components c are produced or consumed with an average specific rate $\bar{\nu}$ by the viable cells x_v. The latter grow at an average specific growth rate $\bar{\mu}$ and viable cells turn into dead cells x_d with a mean mortality rate $\bar{k}_\mathrm{d}$.

D and D_x are dilution rates of the dissolved components c and the biomass x, respectively. In order to be able to describe various operating modes – batch, fed-batch, continuous and perfusion – we assume also that the reactor can be fed with fresh growth medium and that material is withdrawn from the reactor in several ways. Particularly, in a perfusion mode bioreactor, the biomass can be retained inside the bioreactor thanks to a filtering device, while the growth medium can be continuously replaced. The feed flow rate and the withdrawal rates of filtered and unfiltered culture medium are denoted Q_in, $Q_\mathrm{out,c}$ and $Q_\mathrm{out,x}$, respectively. The dilution rates are then defined as

$$D(t) = \frac{Q_\mathrm{in}(t)}{V(t)} \tag{4}$$

$$D_x(t) = \frac{Q_\mathrm{in}(t) - Q_\mathrm{out,c}}{V(t)}. \tag{5}$$

The general solution of the system (1)–(3) is written:

$$c(t) = \left(c_0 + \int_{t_0}^{t} \left(\bar{\nu} x_\mathrm{v}(\tau) + D(\tau) c_\mathrm{f}(\tau)\right) e^{\int_{t_0}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau}} \mathrm{d}\tau\right) \cdot e^{-\int_{t_0}^{t} D(\tau)\mathrm{d}\tau} \tag{6}$$

$$x_\mathrm{v}(t) = x_{\mathrm{v},0} e^{\left(\bar{\mu} - \bar{k}_\mathrm{d}\right)(t - t_0) - \int_{t_0}^{t} D_x(\tau)\mathrm{d}\tau} \tag{7}$$

$$x_\mathrm{d}(t) = \left(x_{\mathrm{d},0} + \bar{k}_\mathrm{d} \int_{t_0}^{t} x_\mathrm{v}(\tau) e^{\int_{t_0}^{\tau} D_x(\tilde{\tau})\mathrm{d}\tilde{\tau}} \mathrm{d}\tau\right) \cdot e^{-\int_{t_0}^{t} D_x(\tau)\mathrm{d}\tau}. \tag{8}$$

Fitting the data triplet $\{c_1, x_{\mathrm{v},1}, x_{\mathrm{d},1}\}$ at the instant t_1 yields the following specific mean rates:

$$\bar{\mu}^* = \frac{\ln\left\{\frac{x_{\mathrm{v},1}}{x_{\mathrm{v},0}}\right\} + \int_{t_0}^{t_1} D_x(\tau)\mathrm{d}\tau}{t_1 - t_0} \tag{9}$$

$$\bar{k}_\mathrm{d} = \frac{x_{\mathrm{d},1} \exp\left\{\int_{t_0}^{t_1} D_x(\tau)\mathrm{d}\tau\right\} - x_{\mathrm{d},0}}{\int_{t_0}^{t_1} x_\mathrm{v}(\tau) \exp\left\{\int_{t_0}^{\tau} D_x(\tilde{\tau})\mathrm{d}\tilde{\tau}\right\}\mathrm{d}\tau} = \frac{x_{\mathrm{d},1} \exp\left\{\int_{t_0}^{t_1} D_x(\tau)\mathrm{d}\tau\right\} - x_{\mathrm{d},0}}{x_{\mathrm{v},1} - x_{\mathrm{v},0}} \bar{\mu}^* \tag{10}$$

$$\bar{\mu} = \bar{\mu}^* + \bar{k}_\mathrm{d} \tag{11}$$

$$\bar{\nu} = \frac{c_1 \exp\left\{\int_{t_0}^{t_1} D(\tau)\mathrm{d}\tau\right\} - c_0}{\int_{t_0}^{t_1} x_\mathrm{v}(\tau) \exp\left\{\int_{t_0}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau}\right\}\mathrm{d}\tau} - \frac{\int_{t_0}^{t_1} D(\tau) c_\mathrm{f}(\tau) \exp\left\{\int_{t_0}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau}\right\}\mathrm{d}\tau}{\int_{t_0}^{t_1} x_\mathrm{v}(\tau) \exp\left\{\int_{t_0}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau}\right\}\mathrm{d}\tau} \tag{12}$$

These calculations can be applied to each pair of consecutive experimental data points, for each time interval $[t_{i-1}, t_i]$, $i = 1, \ldots n_t$. Therefore, an approximate time evolution of the consumption and production rates can be determined.

The problem formulation becomes more complicated, if the reactor medium is diluted at discrete times with fresh medium. This corresponds however to common practice, where a certain volume $V_{\mathrm{out},i}$ is withdrawn from the reactor at the time instant t_i, and a volume $V_{\mathrm{in},i}$ of fresh medium with the concentration $c_{\mathrm{in},i}$ is added to the remaining volume in the reactor.

The algebraic relationships between the volume and the concentrations before and after the dilution are written:

$$V_r(t_i) = V_{\mathrm{r},i} - V_{\mathrm{out},i} + V_{\mathrm{in},i} \tag{13}$$

$$c(t_i) = \frac{(V_{\mathrm{r},i} - V_{\mathrm{out},i})\, c_i + V_{\mathrm{in},i} c_{\mathrm{in},i}}{V_r(t_i)} = \alpha_i c(t_i) + (1 - \alpha_i c_{\mathrm{in},i}) \tag{14}$$

$$x_\mathrm{v}(t_i) = \frac{(V_{\mathrm{r},i} - V_{\mathrm{out},i})\, x_{\mathrm{v},i}}{V_r(t_i)} = \alpha_i x_{\mathrm{v},i} \tag{15}$$

$$x_\mathrm{d}(t_i) = \frac{(V_{\mathrm{r},i} - V_{\mathrm{out},i})\, x_{\mathrm{d},i}}{V_r(t_i)} = \alpha_i x_{\mathrm{d},i}, \tag{16}$$

where $V_{\mathrm{r},i}$, c_i, $x_{\mathrm{v},i}$ and $x_{\mathrm{d},i}$ are the reactor volume and concentrations just before the dilution.

To take these discrete-time dilutions into account, we reformulate our time notations and consider that $[t_0, t_n]$ represents a sampling interval with the first measurement c_0 being available at t_0 and the second measurement c_n being available at t_n. In between, there are n dilutions according to (14) at the discrete times instants $t_0, \ldots, t_{n-1}$ (it is therefore assumed that the sample c_n is taken just before the dilution at t_n); see figure 1.

The evolution of the concentration on the interval $[t_{i-1}, t_i]$ is written according to (7):

$$
\begin{aligned}
c_i &= \Bigg(c(t_{i-1}) \\
&\quad + \int_{t_{i-1}}^{t_i} \left(\bar{\nu} x_\mathrm{v}(\tau) + D(\tau) c_\mathrm{f}(\tau)\right) e^{\int_{t_{i-1}}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau}} \mathrm{d}\tau \Bigg) \\
&\quad \cdot e^{-\int_{t_{i-1}}^{t_i} D(\tau)\mathrm{d}\tau} \\
&= \Bigg(\alpha_{i-1} c_{i-1} + (1-\alpha_{i-1})\, c_{\mathrm{in},i-1} \\
&\quad + \int_{t_{i-1}}^{t_i} \left(\bar{\nu} x_\mathrm{v}(\tau) + D(\tau) c_\mathrm{f}(\tau)\right) e^{\int_{t_{i-1}}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau}} \mathrm{d}\tau \Bigg) \\
&\quad \cdot e^{-\int_{t_{i-1}}^{t_i} D(\tau)\mathrm{d}\tau}
\end{aligned} \tag{17}
$$

for $i = 1, \ldots, n$, and the final concentration c_n is consequently determined recursively from the initial concentration c_0 of the sampling time interval. Using the substitutions

$$
E_D(t_i, t_{i-1}) = \exp\left\{ -\int_{t_{i-1}}^{t_i} D(\tau)\mathrm{d}\tau \right\} \tag{18}
$$

$$
I_{x,c}(t_i, t_{i-1}) = \int_{t_{i-1}}^{t_i} x_\mathrm{v}(\tau) \exp\left\{ \int_{t_{i-1}}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau} \right\} \mathrm{d}\tau \tag{19}
$$

$$
I_u(t_i, t_{i-1}) = \int_{t_{i-1}}^{t_i} D(\tau) c_\mathrm{f}(\tau) \exp\left\{ \int_{t_{i-1}}^{\tau} D(\tilde{\tau})\mathrm{d}\tilde{\tau} \right\} \mathrm{d}\tau, \tag{20}
$$

the recursion is written

$$
\begin{aligned}
c_n &= \big(\alpha_{n-1} c_{n-1} + (1-\alpha_{n-1})\, c_{\mathrm{in},n-1} \\
&\quad + \bar{\nu} I_x(t_n, t_{n-1}) + I_u(t_n, t_{n-1}) \big) E_D(t_n, t_{n-1}) \\
&= \Big(\alpha_{n-1} \big(\alpha_{n-2} c_{n-2} + (1-\alpha_{n-2} c_{\mathrm{in},n-2}) \\
&\qquad + \bar{\nu} I_x(t_{n-1}, t_{n-2}) + I_u(t_{n-1}, t_{n-2}) \big) \\
&\quad \cdot E_D(t_{n-1}, t_{n-2}) \\
&\quad + (1 - \alpha_{n-1} c_{\mathrm{in},n-1}) \\
&\quad + \bar{\nu} I_x(t_n, t_{n-1}) + I_u(t_i, t_{i-1}) \Big) \\
&\quad \cdot E_D(t_n, t_{n-1}) \\
&= \alpha_{n-1}\alpha_{n-2} c_{n-2} E_D(t_{n-1}, t_{n-2}) E_D(t_n, t_{n-1}) \\
&\quad + \alpha_{n-1} (1-\alpha_{n-2})\, c_{\mathrm{in},n-2} E_D(t_{n-1}, t_{n-2}) \\
&\quad \cdot E_D(t_n, t_{n-1}) + (1-\alpha_{n-1})\, c_{\mathrm{in},n-1} E_D(t_n, t_{n-1}) \\
&\quad + \bar{\nu} \big(\alpha_{n-1} I_x(t_{n-1}, t_{n-2}) E_D(t_{n-1}, t_{n-2}) E_D(t_n, t_{n-1}) \\
&\qquad + I_x(t_n, t_{n-1}) E_D(t_n, t_{n-1}) \big) \\
&\quad + \alpha_{n-1} I_u(t_{n-1}, t_{n-2}) E_D(t_{n-1}, t_{n-2}) E_D(t_n, t_{n-1}) \\
&\qquad + I_u(t_n, t_{n-1}) E_D(t_n, t_{n-1}) \\
&\quad \ldots
\end{aligned}
$$

and after some further manipulation and recursion

$$
\begin{aligned}
c_n &= \prod_{j=0}^{n-1} \alpha_j E_D(t_{j+1}, t_j) c_0 \\
&\quad + \sum_{j=1}^{n-1} \prod_{k=j}^{n-1} \alpha_k (1-\alpha_{j-1})\, c_{\mathrm{in},j-1} \prod_{k=j-1}^{n-1} E_D(t_{k+1}, t_k) \\
&\quad + (1-\alpha_{n-1})\, c_{\mathrm{in},n-1} E_D(t_n, t_{n-1}) \\
&\quad + \bar{\nu} \Bigg(\sum_{j=1}^{n-1} \prod_{k=j}^{n-1} \alpha_k I_{x,c}(t_j, t_{j-1}) \prod_{k=j-1}^{n-1} E_D(t_{k+1}, t_k) \\
&\qquad + I_{x,c}(t_n, t_{n-1}) E_D(t_n, t_{n-1}) \Bigg) \\
&\quad + \sum_{j=1}^{n-1} \prod_{k=j}^{n-1} \alpha_k I_u(t_j, t_{j-1}) \prod_{k=j-1}^{n-1} E_D(t_{k+1}, t_k) \\
&\quad + I_u(t_n, t_{n-1}) E_D(t_n, t_{n-1})
\end{aligned} \tag{21}
$$

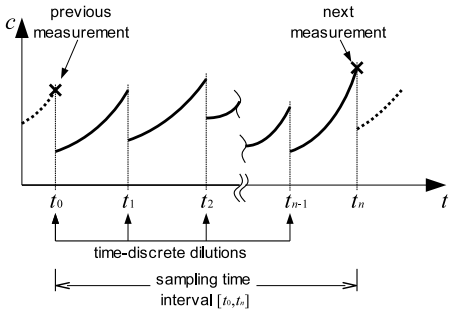

Fig. 1. Several time-discontinuous dilutions in one sampling interval $[t_0, t_n]$.

Analogously, the evolution of the biomass is given by

$$
\begin{aligned}
x_{\mathrm{v},n} &= \alpha_{n-1} x_{\mathrm{v},n-1} e^{\bar{\mu}(t_n - t_{n-1}) - \int_{t_{n-1}}^{t_n} D_x(\tau)\mathrm{d}\tau} \\
&= \alpha_{n-1}\alpha_{n-2} x_{\mathrm{v},n-2} e^{\bar{\mu}(t_{n-1} - t_{n-2}) - \int_{t_{n-2}}^{t_{n-1}} D_x(\tau)\mathrm{d}\tau} \\
&\qquad \cdot e^{\bar{\mu}(t_n - t_{n-1}) - \int_{t_{n-1}}^{t_n} D_x(\tau)\mathrm{d}\tau} \\
&= \alpha_{n-1}\alpha_{n-2} x_{\mathrm{v},n-2} e^{\bar{\mu}(t_n - t_{n-2}) - \int_{t_{n-2}}^{t_n} D_x(\tau)\mathrm{d}\tau} \\
&= \ldots .
\end{aligned}
$$

Further manipulation leads to the final equation

$$
x_{\mathrm{v},n} = \prod_{j=0}^{n-1} \alpha_j x_{\mathrm{v},0} \exp\left\{ \bar{\mu}(t_n - t_0) - \int_{t_0}^{t_n} D_x(\tau)\mathrm{d}\tau \right\}. \tag{22}
$$

The evolution equation for the dead biomass x_d is structurally identical to the evolution equation (21) for c with $D = D_x$, $\bar{\nu} = \bar{k}_\mathrm{d}$ and $c_\mathrm{f} = c_\mathrm{in} = 0$ instead.

In the case of discrete-time dilution, the more general set of equations for the determination of the mean specific rates in a time interval $[t_0, t_n]$ are therefore written as follows

$$
\bar{\mu} = \frac{\ln x_{\mathrm{v},n} - \ln x_{\mathrm{v},0} - \sum_{j=0}^{n-1} \ln \alpha_j + \int_{t_0}^{t_n} D_x(\tau)\mathrm{d}\tau}{t_n - t_0} \tag{23}
$$

$$
\bar{\nu} = \frac{c_n - P_\alpha c_0 - S_u}{S_{x,c}} \tag{24}
$$

$$
\bar{k}_\mathrm{d} = \frac{x_{\mathrm{d},n} - P_\alpha x_{\mathrm{d},0}}{S_{x,x}} \tag{25}
$$

with the respective substitutions

$$P_\alpha = \prod_{j=0}^{n-1} \alpha_j E_D(t_n, t_0) \tag{26}$$

$$\begin{aligned} S_u = &\sum_{j=1}^{n-1} \prod_{k=j}^{n-1} \alpha_k \left(I_u(t_j, t_{j-1}) + (1-\alpha_{j-1})\, c_{\mathrm{in},j-1} \right) \\ &\cdot E_D(t_n, t_{j-1}) \\ &+ \left(I_u(t_n, t_{n-1}) + (1-\alpha_n)\, c_{\mathrm{in},n-1} \right) \\ &\cdot E_D(t_n, t_{n-1}) \end{aligned} \tag{27}$$

$$\begin{aligned} S_{x,c} = &\sum_{j=1}^{n-1} \prod_{k=j}^{n-1} \alpha_k I_{x,c}(t_j, t_{j-1}) E_D(t_n, t_{j-1}) \\ &+ I_{x,c}(t_n, t_{n-1}) E_D(t_n, t_{n-1}) \end{aligned} \tag{28}$$

$$\begin{aligned} S_{x,x} = &\sum_{j=1}^{n-1} \prod_{k=j}^{n-1} \alpha_k I_{x,x}(t_j, t_{j-1}) E_{D_x}(t_n, t_{j-1}) \\ &+ I_{x,x}(t_n, t_{n-1}) E_{D_x}(t_n, t_{n-1}). \end{aligned} \tag{29}$$

The equations developed so far are general and can be applied to bioreactors in various operating modes. In perfusion mode, with constant feed concentration c_f, constant perfusion rate D, constant reactor volume V and complete cell retention ($D_x \equiv 0$), the following closed analytical expressions are obtained:

$$\bar{\mu}^* = \frac{\ln x_{\mathrm{v}1} - \ln x_{\mathrm{v}0}}{t_1 - t_0} \tag{30}$$

$$\bar{k}_\mathrm{d} = \frac{x_{\mathrm{d}1} - x_{\mathrm{d}0}}{x_{\mathrm{v}1} - x_{\mathrm{v}0}} - \bar{\mu}^* \tag{31}$$

$$\bar{\mu} = \bar{\mu}^* + \bar{k}_\mathrm{d} \tag{32}$$

$$\bar{\nu} = \frac{(c_1 - c_\mathrm{f})\, e^{Dt_1} - (c_0 - c_\mathrm{f})\, e^{Dt_0}}{x_{\mathrm{v}1} e^{Dt_1} - x_{\mathrm{v}0} e^{Dt_0}} \left(\bar{\mu} + D \right). \tag{33}$$

These mean consumption/production, growth and mortality rates can now be used to make a preliminary estimation of the complete stoichiometry of the reaction system and the respective reaction rates. Thus far, the specific rates for each component are determined independently of each other. However, macroscopic reaction schemes relate the consumption of some reactants with the formation of some products according to the stoichiometry. In other words, the stoichiometric matrix $\Upsilon \in \mathbb{R}^{n_q \times n_r}$ maps the reaction rate vector $\rho \in \mathbb{R}^{n_r}$ into a consumption/production rate vector $q \in \mathbb{R}^{n_q}$ of each component by linear combination of the individual reaction rates of the reaction scheme

$$q = \Upsilon \rho \tag{34}$$

with q and ρ respectively defined as

$$q = \begin{bmatrix} \nu_1 \\ \vdots \\ \nu_{n_q-2} \\ \mu^* \\ k_\mathrm{d} \end{bmatrix}; \qquad \rho = \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_{n_r-1} \\ k_\mathrm{d} \end{bmatrix}. \tag{35}$$

The consumption/production rates have been determined approximately for each sampling time interval, while the constant stoichiometric matrix and the reaction rates in each sampling time interval are unknown.

To determine these unknown values, a necessary condition is therefore $n_r < n_q$, i.e. the number of reactions in the macroscopic reaction scheme must be lower than the number of components. More precisely, if n_c different concentrations are measured, then $n_t n_c$ data points are available for the determination of n_r reaction rates at each of the n_t sampling intervals and the stoichiometric coefficients.

If the stoichiometric matrix Υ is constructed as follows:

$$\Upsilon = \left[\begin{array}{c|c} Y & \begin{matrix} 0 \\ -1 \end{matrix} \\ \hline \begin{matrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 \end{matrix} & \begin{matrix} -1 \\ 1 \end{matrix} \end{array}\right] \tag{36}$$

with the yield matrix $Y \in \mathbb{R}^{(n_q-2)(n_r-1)}$, the total number of unknowns is $n_r n_c + (n_q - 2)(n_r - 1)$. In this expression, it is assumed that the dead biomass is exclusively formed through the mortality reaction, and the stoeciometry is normalized with respect to the viable biomass.

3. A PRACTICAL RECIPE

In the following, a recipe summarizing the results obtained so far is given.

(1) Set up a time vector $t = \begin{bmatrix} t_0 & \cdots & t_{n_t} \end{bmatrix}^\mathrm{T}$ containing all discrete time instants corresponding to sampling and/or dilutions.
(2) Determine the specific mean growth rate $\bar{\mu}^*$ for each time interval $[t_i, t_{i+1}]$, $i = 0, \ldots, n_t - 1$, according to equation (23).
(3) Determine the unmeasured viable cell concentrations at all time instants of t according to

$$\begin{aligned} &x_{\mathrm{v},i+1} = \\ &\quad x_{\mathrm{v},i} \alpha_i \exp\left\{ \bar{\mu}^* (t_{i+1} - t_i) - \int_{t_i}^{t_{i+1}} D_x(\tau) \mathrm{d}\tau \right\} \\ &\qquad\qquad i = 1, \ldots, n_t - 1, \end{aligned} \tag{37}$$

which is easily deduced from (22).
(4) Determine the exponential $E_D(t_{i+1}, t_i)$by equation (18) for all time intervals.
(5) Determine the integrals $I_u(t_{i+1}, t_i)$ and $I_{x,c}(t_{i+1}, t_i)$ by equations (20) and (19) for all time intervals.
(6) Determine the dilution factor $\alpha_i = \frac{V_{\mathrm{r},i} - V_{\mathrm{out},i}}{V_{\mathrm{r},i} - V_{\mathrm{out},i} + V_{\mathrm{in},i}}$ according to its definition in equation (14).
(7) Calculate the specific mean rates $\bar{\nu}_i$ and $\bar{k}_{\mathrm{d},i}$ by equations (24) and (25) and $\bar{\mu}$ by equation (11).
(8) The mean production/consumption, growth and mortality rates can serve for the preliminary estimation of the yield matrix $\hat{Y}$ and the specific reaction rates $\hat{\rho}_i$ subject to equation (34), e.g. by optimization of a weighted least-squares criterion

$$[\hat{Y},\hat{\rho}] = \arg\min_{Y,\rho} \sum_{i=0}^{n_t-1} (\bar{q}_i - \Upsilon(Y)\rho_i)^{\mathrm{T}} \Omega^{-1} \cdot (\bar{q}_i - \Upsilon(Y)\rho_i) \quad (38)$$

with a weighting matrix Ω^{-1}, e.g. chosen according to the order of magnitude, $\Omega = \mathrm{diag}\left\{\max_i q_{ji}^2\right\}_j$.

In a second step, the estimated specific reaction rates $\hat{\rho}_i$ for each time interval $[t_i, t_{i+1}]$ can be used to estimate the kinetic parameters. At this stage, the states have to be known, whereas they are available at some discrete time instants only. In order to get a representative value for each interval, a mean value of the initial and final values is evaluated:

$$\bar{c}_i = \frac{c_{i+1} + \alpha_i c_i + (1-\alpha_i)\, c_{\mathrm{in},i}}{2} \quad (39)$$

$$\bar{x}_{\mathrm{v},i} = \frac{x_{\mathrm{v},i+1} + \alpha_i x_{\mathrm{v},i}}{2} \quad (40)$$

$$\bar{x}_{\mathrm{d},i} = \frac{x_{\mathrm{d},i+1} + \alpha_i x_{\mathrm{d},i}}{2}. \quad (41)$$

In this manner, approximate points representing the dependence of the kinetics on the state vector can be collected for all time intervals $[t_{i-1}, t_i]$, $i = 1, \ldots, n_t$, and fit to an assumed structure of the kinetics. This structure is generally written in the following form:

$$\rho(c, x_{\mathrm{v}}, x_{\mathrm{d}}, p_{\mathrm{kin}}) = \begin{bmatrix} \nu(c, x_{\mathrm{v}}, x_{\mathrm{d}}, p_{\mathrm{kin}}) \\ \mu(c, x_{\mathrm{v}}, x_{\mathrm{d}}, p_{\mathrm{kin}}) \\ k_{\mathrm{d}}(c, x_{\mathrm{v}}, x_{\mathrm{d}}, p_{\mathrm{kin}}) \end{bmatrix} \quad (42)$$

Based on this reaction vector, the kinetic parameter vector p_{kin} can be estimated using the corresponding $\hat{\rho}_i$ that has been previously determined for all intervals $i = 1, \ldots, n_t - 1$. For instance, the weighted least-squares criterion

$$\hat{p}_{\mathrm{kin}} = \arg\min_{p_{\mathrm{kin}}} \left\{ \sum_{i=0}^{n_t-1} (\hat{\rho}_i - \rho(\bar{c}_i, \bar{x}_{\mathrm{v},i}, \bar{x}_{\mathrm{d},i}, p_{\mathrm{kin}}))^{\mathrm{T}} \Omega^{-1} \cdot (\hat{\rho}_i - \rho(\bar{c}_i, \bar{x}_{\mathrm{v},i}, \bar{x}_{\mathrm{d},i}, p_{\mathrm{kin}})) \right\} \quad (43)$$

could be used.

Finally, a complete optimization can be performed with respect to the yield matrix $Y(p_{\mathrm{st}})$ containing the stoichiometric parameters p_{st} as unknown elements and the kinetic parameters p_{kin}, given the specific mean rates $\bar{q}_i$ and subject to

$$q = \Upsilon(p_{\mathrm{st}})\rho(c, x_{\mathrm{v}}, x_{\mathrm{d}}, p_{\mathrm{kin}}), \quad (44)$$

e.g. using again the weighted least-squares criterion

$$[\hat{p}_{\mathrm{st}}, \hat{p}_{\mathrm{kin}}] = \arg\min_{p_{\mathrm{st}}, p_{\mathrm{kin}}} \left\{ \sum_{i=0}^{n_t-1} (\bar{q}_i - \Upsilon(p_{\mathrm{st}})\rho(\bar{c}_i, \bar{x}_{\mathrm{v},i}, \bar{x}_{\mathrm{d},i}, p_{\mathrm{kin}}))^{\mathrm{T}} \cdot \Omega^{-1} (\bar{q}_i - \Upsilon(p_{\mathrm{st}})\rho(\bar{c}_i, \bar{x}_{\mathrm{v},i}, \bar{x}_{\mathrm{d},i}, p_{\mathrm{kin}})) \right\} \quad (45)$$

Note that relatively small measurement errors directly influence the results of the preliminary data evaluation, and that sufficiently small sampling time intervals are needed in order to avoid a too strong deviation between the real reaction rates and their mean on each time interval (which is basic assumption of the parameter redetermination).

4. A SIMPLE APPLICATION

Consider the following reaction system consisting of two metabolic reactions and one mortality reaction

$$\mathrm{S}_1 \xrightarrow{\rho_1} \mathrm{X_v} + \mathrm{P} \quad (46)$$

$$\mathrm{S}_2 \xrightarrow{\rho_2} \mathrm{X_v} \quad (47)$$

$$\mathrm{X_v} \xrightarrow{\rho_3} \mathrm{X_d} \quad (48)$$

involving five components: the substrates S_1 and S_2, the product P and the viable and non-viable biomasses, $\mathrm{X_v}$ and $\mathrm{X_d}$, respectively. The specific growth rates are defined by Monod kinetics

$$\rho_1(s_1) = \mu_{1,\max} \frac{s_1}{s_1 + K_1} \quad (49)$$

$$\rho_2(s_2) = \mu_{2,\max} \frac{s_2}{s_2 + K_2} \quad (50)$$

with $\mu_{1,\max} = 0.3$, $K_1 = 1$, $\mu_{2,\max} = 0.2$ and $K_2 = 2$, while the specific mortality rate is constant, $\rho_3 = 0.1$.

The culture is performed in a stirred-tank reactor, first operated in batch mode, $D = 0$, for $0 \le t < 5$, with the initial condition

$$\begin{bmatrix} s_{1,0} & s_{2,0} & p_0 & x_{\mathrm{v},0} & x_{\mathrm{d},0} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 2 & 3 & 0 & 1 & 0 \end{bmatrix}, \quad (51)$$

and then operated in continuous mode with a perfusion rate, $D = 1$, for $5 \le t \le 10$, with the following feed concentrations:

$$\begin{bmatrix} s_{1,\mathrm{f}} & s_{2,\mathrm{f}} & p_{\mathrm{f}} & x_{\mathrm{v,f}} & x_{\mathrm{d,f}} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 2 & 2 & 0 & 0 & 0 \end{bmatrix} \quad (52)$$

Samples are taken every $\Delta t = 1$ time units, i.e. 11 samples are taken during the experiment and a total of 55 measurements are available for identification purposes (see Figure 2). They serve for a preliminary calculation of the consumption and production rates in each of the 10 sampling time intervals (Table 1).

Table 1. Estimated average production and consumption rates in each time interval.

	[0,1]	[1,2]	[2,3]	[3,4]	[4,5]
S_1	−0.196	−0.187	−0.174	−0.155	−0.130
S_2	−0.119	−0.117	−0.113	−0.109	−0.105
P	0.393	0.374	0.347	0.311	0.260
$\mathrm{X_v}$	0.415	0.404	0.387	0.365	0.335
$\mathrm{X_d}$	0.100	0.100	0.100	0.100	0.100
	[5,6]	[6,7]	[7,8]	[8,9]	[9,10]
S_1	−0.153	−0.172	−0.174	−0.171	−0.166
S_2	−0.098	−0.095	−0.093	−0.091	−0.089
P	0.305	0.343	0.347	0.342	0.332
$\mathrm{X_v}$	0.346	0.366	0.366	0.362	0.356
$\mathrm{X_d}$	0.100	0.100	0.100	0.100	0.100

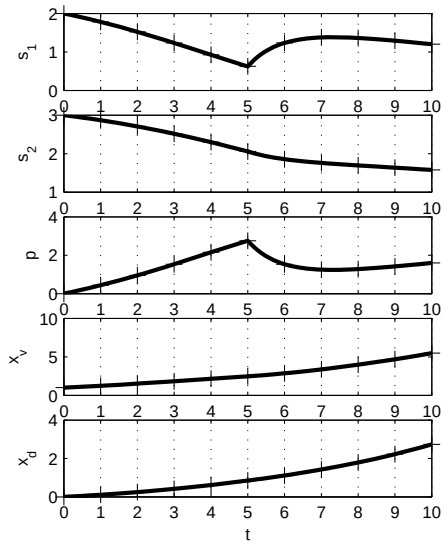

Fig. 2. Measurements in simulated experiment.

Assuming a reaction scheme with two reactions, as it is the case for the reference system, the stoichiometric matrix $\hat{\Upsilon}_0$ (Table 2) is obtained from the identification of the linear system (34) as well as the corresponding matrix of specific reaction rates ρ at all time intervals. The resulting predicted specific production and consumption rates are shown in Figure 3 together with the average rates from Table 1 and the true continuous time evolution.

Table 2. Stoichiometric matrices of the reference and approximate systems.

	reference			estimation		
rcn. #	1	2	3	1	2	3
S_1	-1	0	0	-1.260	-0.107	0
S_2	0	-1	0	0.275	-0.906	0
P	2	0	0	2.521	0.213	0
X_v	1	1	-1	1	1	-1
X_d	0	0	1	0	0	1

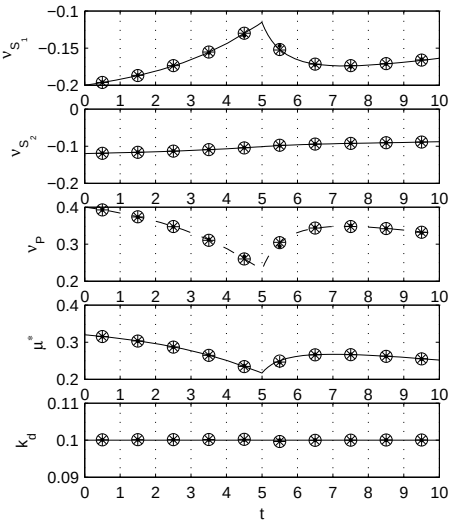

Fig. 3. Average production rates. Stars: preliminary computation, circles: first estimation step (stoichiometry), dots: last estimation step (kinetics).

The reaction rates ρ are then used for a preliminary estimation of the kinetic coefficients $\hat{\rho}_{j,\max,0}$ and $\hat{K}^*_{ij,0}$, $i = 1,\ldots,5$ and $j = 1,2$, according to the general kinetic model structure introduced by Haag *et al.* (2003). These kinetic coefficients can then be used together with the stoichiometric matrix Υ_0 as initial estimates for a final estimation of the kinetic model in combination with the stoichiometry. The result is displayed in table 3. The resulting stoichiometric matrix remains almost unchanged, $\hat{\Upsilon} \approx \hat{\Upsilon}_0$.

Table 3. Kinetic parameters ρ_{max} and K^* for the reference and identified systems.

	reference			estimation		
rcn. #	1	2	3	1	2	3
$\rho_{\max}$	0.3	0.2	0.1	0.270	0.261	0.106
$K^*_{S_1}$	1	0	0	0.944	0.299	-0.161
$K^*_{S_2}$	0	1.41	0	0.300	1.101	0.113
K^*_{P}	0	0	0	-0.169	-0.143	0.116
$K^*_{X_v}$	0	0	0	-0.141	0.091	0.063
$K^*_{X_d}$	0	0	0	0.000	0.011	0.006

As is shown in Tables 2 and 3, these preliminary results are already very close to the real model parameters, despite the strong assumption of constant reaction rates in each sampling interval.

5. CONCLUSION

In this study, a procedure for the fast prototyping of bioprocess models is proposed. In a first step, mean production/consumption, growth and mortality rates are obtained, which can serve for a preliminary evaluation of the stoichiometric matrix and specific reaction rates. In a second step, approximate values of the parameters of an assumed kinetic model structure can be evaluated. Finally, in a third step, all the stoichiometric and kinetic parameters can be re-estimated more accurately. The proposed procedure allows experimental data to be assessed and modeling assumptions to be checked. Finally, the estimated parameter values can be used as starting points for more rigorous identification methods.

REFERENCES

Bastin, G. and D. Dochain (1990). *On-Line Estimation and Adaptive Control of Bioreactors*. Vol. 1 of *Process Measurement and Control*. Elsevier. Amsterdam.

Bogaerts, Ph. and R. Hanus (2000). Macroscopic modelling of bioprocesses with a view to engineering applications. In: *Focus on Biotechnology, vol. IV (Engineering and Manufacturing for Biotechnology)* (Ph. Thonart and M. Hofman, Eds.). Kluwer Academic Publishers.

Haag, J. E., A. Vande Wouwer and M. Remy (2003). A general model of reaction kinetics in biological systems. In: *European Control Conference*. Cambridge, UK.