ESTIMATION OF BIOMASS CONCENTRATION USING INTERVAL OBSERVERS IN AN *E. COLI* FED-BATCH FERMENTATION

Ana C. A. Veloso¹,², I. Rocha¹ and E. C. Ferreira¹

¹IBB - Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, Universidade do Minho 4710-057 Braga, PORTUGAL
²Escola Superior Agrária de Bragança, Campus de Santa Apolónia, Apartado 1172 5301-854 Bragança, PORTUGAL

e-mail: {anaveloso, irocha, ecferreira}@deb.uminho.pt

Abstract: In bioreactors, the measurement of variables that play a key role in the quality and productivity of fermentations, is of major importance. However, their direct measurement is often expensive or even impossible considering the current sensor technology. Therefore, on-line estimation of unmeasured variables in bioreactors can be an interesting approach.

The objective of this work is to introduce an alternative solution for the observation of biomass concentration in *E. coli* fed-batch fermentations, in cases where the kinetic model is unclear and several variables, like the concentration of the influent substrates and the initial values of the state variables are badly known, a situation that is common in many practical applications.

The simple interval observer is designed on the basis of the cooperativity properties of the observer error dynamics (Rapaport and Dochain, 2005).

The performance of the interval observer is illustrated through numerical simulation and it was found that the observer deal well with uncertainties up to 50% and with white noise in the variables measured on-line. The interval obtained for the biomass estimation is also quite narrow, indicating that it is possible to accurately predict biomass concentration under the presence of uncertainties. Copyright © 2007 IFAC

Key-words: Biomass Estimation, Interval Observers, Uncertainty, Fed-batch Fermentation.

1. INTRODUCTION

It is well known that industries are interested in decreasing the production costs and increasing the process yield, keeping the quality of the metabolic products. Thus, the ability to accurately and automatically control bioprocesses at their optimal state is of great importance, since it can contribute to achieve that goal. However, the lack of on-line instruments has limited the application of control theory to these processes. Therefore, the development of state observers, also called software sensors (Dochain, 2003) can be an attractive alternative since a large amount of additional information can be obtained, using a model together with a limited set of state variable measurements (Bernard and Gouzé, 2004, Bogaerts and Vande Wouwer, 2004).

In the literature, two classes of state observers are usually found. The first class includes the classical observers, such as the Luenberger, the Kalman, and the non-linear observers, which are based on the perfect knowledge of both model structure and parameters. However, the uncertainty in the model parameters can generate a large bias in the estimation of unmeasured state(s). The asymptotic observers
(Bastin and Dochain, 1990), which constitute the second class of observers, do not require the knowledge of the process kinetics. Nevertheless, a potential problem concerning these observers is the dependence of the estimation convergence rate on the operating conditions (Dochain, 2003). Additionally, both approaches neglect the fact that there is often additional uncertainty associated with other process parameters, like the ones associated with influent substrate.

This work aims therefore to study an alternative solution for the state observation of a high-cell density fed-batch fermentation of E. coli, assuming that the kinetic model structure is unknown and the concentration of the influent substrates and the initial values of the state variables are badly known.

The approach used is based on an interval analysis (Rapaport and Dochain, 2005; Alcaraz-Gonzalez et al., 2002; Hadj-Sadok and Gouzé, 2001; Gouzé et al., 2000). The objective is to reconstruct intervals for the biomass concentration, for which the state is certain to lie, based on a given interval of variation of the uncertain variable(s). Biomass is an important process variable but nevertheless still very difficult to measure on-line in this fermentation process. Its estimation can be used for model predictive control, estimation of specific growth rates, and optimization of the production of recombinant proteins (regarding both productivity and moment of induction).

In this study, the design of the interval observers is based on the assumption that measurements of acetate, dissolved oxygen and carbon dioxide concentrations are available. This choice is due to the fact that, nowadays, the sensors for these state variables are more developed and thus, more reliable.

2. PROCESS MODELLING

The dynamics of a reaction network in a stirred tank bioreactor can be described by the following mass balance equations written in matrix form as (Bastin and Dochain, 1990):

$$\frac{d \xi}{dt} = Kr(\xi, t) - D\xi + F - Q$$  \hspace{1cm} (1)

in which $\xi$ is a vector representing the $n$ state components concentrations ($\xi \in \mathbb{R}^n$), $r$ is the growth rate vector corresponding to $m$ reactions ($r \in \mathbb{R}_+^m$), $K$ is the matrix of yield coefficients ($K \in \mathbb{R}_+^{n \times m}$), $F$ is the vector of feed rates, $Q$ is the vector of gaseous outflow rates ($F, Q \in \mathbb{R}^n$), and $D$ is the dilution rate (being $D^{-1}$ the residence time).

During the aerobic growth of E. coli with glucose as the only added substrate, the microorganism can follow three main metabolic pathways: oxidative growth on glucose, fermentative growth on glucose, and oxidative growth on acetate. The corresponding dynamical model for fed-batch fermentation can be represented as follows:

$$\frac{d X}{dt} = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_4 & -k_3 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X - \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} A + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} O \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} OTR$$

where $X, S, A, O$, and $C$ represent biomass, glucose, acetate, dissolved oxygen, and dissolved carbon dioxide concentrations, respectively; $\mu_1$, $\mu_2$, and $\mu_3$ are the specific growth rates; $k_1$ are the yield (stoichiometric) coefficients; $F_{in}$ and $S_{in}$ are the substrate feed rate and the influent glucose concentration, respectively; $W$ is the culture medium weight, $CTR$ is the carbon dioxide transfer rate from liquid to gas phase, and $OTR$ is the oxygen transfer rate from gas to liquid phase that can be approximated to the oxygen uptake rate (OUR) as follows:

$$OTR \approx OUR = (k_1 \mu_1 + k_2 \mu_2 + k_3 \mu_3) X$$  \hspace{1cm} (3)

The variation of the culture medium weight with the time is given by:

$$\frac{d W}{dt} = F_{in}$$  \hspace{1cm} (4)

where $F_{in}$ includes weight variations due to the substrate feed rate, the amount of culture removed or added during sampling, base and acid additions, evaporation and mass taken from the reactor due to gas exchanges, that can not be considered negligible in small-scale high-cell density reactors.

The observation problem is the estimation of biomass concentration from on-line measurements of acetate, dissolved oxygen and carbon dioxide concentrations. It is assumed that the kinetics are unknown, the dilution rate $D$ and the yield coefficients (matrix $K$) are known, and the influent glucose concentration $S_{in}$ is uncertain but bounded between known lower and upper bounds:

$$S_{in}(t) \leq S_{in}(t) \leq S_{in}(t)$$  \hspace{1cm} (5)

3. DESIGN OF THE INTERVAL OBSERVERS

The motivation of interval observers is to generate state estimates with bounds that are related to the uncertainty of the model or of the measurements (Alcaraz-Gonzalez et al., 2002, Gouzé et al., 2000). The design is based on the cooperativity properties of the observer error dynamics. Cooperative systems are dynamical systems for which the non-diagonal terms of the Jacobian matrix are positive (Rapaport and Dochain, 2005, Rapaport and Gouzé, 2003, Smith, 2002; Hadj-Sadok and Gouzé, 2001; Gouzé et al., 2000). The objective is to reconstruct intervals for the dynamical model for fed-batch fermentation can be represented as follows:
1996). Considering the following non-linear state space model:

\[ \frac{dx}{dt} = f(t,x) \]  

(6)

cooperation means that for any \((t,x)\):

\[ \frac{\partial f_i}{\partial x_j}(t,x) \geq 0, \quad \text{for } i \neq j \]  

(7)

Considering the cooperative system (6) and being \(f^-\) and \(f^+\) two vector fields such that:

\[ f^-(x) \leq f(t,x) \leq f^+(x), \quad \forall (t,x) \]  

(8)

and the initial conditions \(x_0^-, x_0^+, x_0^+\) (such that \(x_0^- \leq x_0 \leq x_0^+\)) the solution of the dynamical system is:

\[ \frac{dx}{dt} = \begin{bmatrix} f^-(x) \\ f(t,x) \\ f^+(x) \end{bmatrix}, \quad x(0) = \begin{bmatrix} x^- \\ x_0 \\ x^+ \end{bmatrix} \]  

(9)

which fulfills the following property:

\[ x^-(t) \leq x(t) \leq x^+(t), \quad t \geq 0 \]  

(10)

Therefore, two estimates can be computed, an upper one and a lower one, that bound the unmeasured variables. Since the best final estimate is aimed, the interval \([x(t), x^+(t)]\) should become smaller (or ideally tends to \(\{x(t)\}\)) when the time \(t\) increases (Raissi et al., 2005; Gouzé et al., 2000).

The observer (being \(g_1\) to \(g_5\) the observer gains), designed directly from the dynamical model given by eq.(2), can be written as:

\[ \frac{dx}{dt} = (\mu_1 + \mu_2)X - DX + g_1(x - \bar{A}) + g_2(\bar{O} - \bar{C}) + g_3(C - \bar{C}) \]  

(11a)

\[ \frac{dx}{dt} = (-k_1 - k_2)X - DS + g_4(x - \bar{A}) + g_5(\bar{O} - \bar{C}) + g_6(C - \bar{C}) \]  

(11b)

\[ \frac{dx}{dt} = (k_1 - k_2)X - D\bar{A} + g_7(x - \bar{A}) + g_8(\bar{O} - \bar{C}) + g_9(C - \bar{C}) \]  

(11c)

\[ \frac{dx}{dt} = (-k_1 - k_2)X - D\bar{A} + g_7(x - \bar{A}) + g_8(\bar{O} - \bar{C}) + g_9(C - \bar{C}) \]  

(11d)

\[ \frac{dx}{dt} = (k_1 - k_2)X - D\bar{A} - D\bar{C} + r(x - \bar{A}) + g_7(x - \bar{A}) + g_8(\bar{O} - \bar{C}) + g_9(C - \bar{C}) \]  

(11e)

It should be noticed that the interval observer cannot be designed directly from the observer given by eq. (11). In fact, the off-diagonal term of the Jacobian matrix (eq. (12)) of the observer does not fulfill the condition of eq. (7), as the off-diagonal term \((-K\mu)\) is negative and \((-K\mu^+)\) can also be negative in some cases:

\[ \frac{d\xi}{dt} = -D\bar{Z} - K_1K_1^{-1}\xi + (F - Q) \]  

(15)

If the measured variables are \(A, O\) and \(C\) the matrix used in the state transformation of eq. (14) will be:

\[ K, K_1^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -k_1 & -k_2 \\ k_1 & k_2 & k_3 \end{bmatrix} \]  

(16)

Using this approach, several interval observers can be defined from which two examples are presented in the next sections.

3.1 Simple Interval Observers

The simplest formulation is obtained directly from eqs. (14) and (15):

\[ \frac{d\hat{Z}}{dt} = -D\hat{Z} - K_1K_1^{-1}\xi + (F - Q) \]  

(17a)

\[ \hat{\xi} = \hat{Z} + K_1K_1^{-1}\xi \]  

(17b)

The observer, for the case studied, is given by the following equations:
The following observer is therefore designed with the incorporation of gains that are associated with OTR measurements:

\[
\frac{d [\hat{Z}]_i}{dt} = -D [\hat{Z}]_i - [a_i, a_i, a_i] \begin{bmatrix} 0 & 0 \\ O & -CTR \end{bmatrix} + \frac{F_i}{W} S_i (25)
\]

being,

\[\hat{O}TR = (k_i, \hat{\mu_i} + k_i, \hat{\mu_i} + k_i, \hat{\mu_i})\hat{X}\]

The estimation of the state variables is then given by:

\[
\begin{bmatrix} \hat{X} \\ \hat{S} \end{bmatrix} = \begin{bmatrix} a_i & a_i & a_i \\ \hat{Z}_i & a_i & a_i \end{bmatrix} \begin{bmatrix} A \\ O \\ C \end{bmatrix} (27)
\]

In this case the error dynamics is given by:

\[
\frac{d [e_i]}{dt} = -D [e_i] - \begin{bmatrix} g_i(k_i, \hat{\mu_i} + k_i, \hat{\mu_i} + k_i, \hat{\mu_i}) e_i \\ g_i(k_i, \hat{\mu_i} + k_i, \hat{\mu_i} + k_i, \hat{\mu_i}) e_i \end{bmatrix} (28)
\]

with the following Jacobian matrix:

\[
J(e_i, e_i) = \begin{bmatrix} -D & 0 \\ 0 & -D \end{bmatrix} (29)
\]

The system is cooperative if \(g_2\) is negative or equal to zero. In the last case, this interval observer becomes identical to the one deduced in the previous section (eq. 24). When \(g_2\) is negative, the following observer will be obtained:

\[
\frac{d [\hat{Z}]_i}{dt} = -D [\hat{Z}]_i - \begin{bmatrix} 0 & 0 \\ O & -CTR \end{bmatrix} + \frac{F_i}{W} S_i (25)
\]

\[
\begin{bmatrix} \hat{X} \\ \hat{S} \end{bmatrix} = \begin{bmatrix} a_i & a_i & a_i \\ \hat{Z}_i & a_i & a_i \end{bmatrix} \begin{bmatrix} A \\ O \\ C \end{bmatrix} (30a)
\]

For obtaining the specific growth rates \(\mu_i\) to \(\mu_k\), used in eq. (30a), the estimator deduced in Rocha et al. (2006), based on the formulation proposed by Bastin and Dochain (1990) and reformulated by Pomerleau and Perrier (1990), was used. However, for this application, this estimator takes into account both the upper and lower bounds of the variables calculated as expressed below:

\[
\frac{d\psi_i^+}{dt} = \dot{\psi}_i^+ - D\dot{\psi}_i^+ + a_{i, OTR} - a_{i, CTR} + a_{i, OTR} (\psi_i^+, \psi_i^-) (31a)
\]

\[
\frac{d\psi_i^-}{dt} = \dot{\psi}_i^- - D\dot{\psi}_i^- + a_{i, OTR} - a_{i, CTR} + a_{i, OTR} (\psi_i^-, \psi_i^-) (31b)
\]

\[
\frac{d\psi_i}{dt} = \dot{\psi}_i - D\dot{\psi}_i + a_{i, OTR} - a_{i, CTR} + a_{i, OTR} (\psi_i^-) (31c)
\]

\[
\frac{d\psi_i}{dt} = a_{i, OTR} (\psi_i^-, \psi_i^-) (31d)
\]

\[
\frac{d\psi_i}{dt} = a_{i, OTR} (\psi_i^-) (31e)
\]

\[
\frac{d\psi_i}{dt} = a_{i, OTR} (\psi_i^-) (31f)
\]

being \(\psi\) obtained from the transformation \(\psi = K_i^{-1} \xi_i\) which allows a decoupled estimation of each specific growth parameter. The calculation of the estimator’s
gains \( \omega \) is made such that a second order dynamics is obtained (Oliveira et al., 1996):

\[
\omega^i_{x_0} = \frac{2s_\omega}{r_s} \frac{\hat{X}_i - X_s}{X_s} \\
\omega^u_{x_0} = \frac{2s_\omega}{r_s} \frac{\hat{X}_u - X_s}{X_s} \\
\omega^0_{x_0} = \frac{1}{r_sX_s} \tag{32a}
\]

\[
\omega^i_{x_0} = \frac{1}{r_sX_s} \\
\omega^u_{x_0} = \frac{1}{r_sX_s} \\
\omega^0_{x_0} = \frac{1}{r_sX_s} \tag{33a}
\]

where \( T \) is the integration step.

4. SIMULATION RESULTS AND DISCUSSION

The performance of the interval observers has been tested by numerical simulations. The model simulations were performed by integration of the differential equations of eq. (2), using the MATLAB version 7.1 subroutine ODE23s. The implementation of the observer and the estimators was conducted using the Euler integration method. Most of the mathematical operations behind the design of the observer and the estimators were performed using the Symbolic Math toolbox running in MATLAB 7.1.

For validating the developed algorithms “real” values of the state variables were obtained by integration of eq. (2). The “real” values of the variables that can be obtained on-line, i.e., \( A, O, C, OTR, CTR \) and \( W \) were then corrupted with white noise, according to the standard deviations typically found in this process at the authors’ lab, originating “experimental” values. Then, the observer and estimator algorithms were used to obtain the “estimated” variables from the “experimental” data corresponding to the measured variables.

It is assumed that glucose concentration in the feed \( S_0 \) is unknown, but it is assumed that its time varying bounds are known (\( S_{m_{g1}}(t) \leq S(t) \leq S_{m_{g2}}(t) \), \( \forall t \geq 0 \)). It is also assumed that a priori bounds on initial values of \( X_0 \) and \( S_0 \) are known.

Therefore, the simulation results have been performed by considering that the uncertainty is concentrated on the influent glucose concentration and on the initial state variables \( X \) and \( S \) with 50% variation around their nominal values (\( S_m = 250 \text{ g/kg}; X_m = 1.5 \text{ g/kg}; S_i = 0 \text{ g/kg} \)).

The performance of the simple observer is shown in Figure 1, while the performance of the observer with gains from OTR is shown in Figure 2. As it can be seen, the performance of both observers is satisfactory, although for the first case the “real” values of biomass are not always included in the interval given by the observer. Additionally, the observer with gains from OTR originates a slightly lower noise in the predictions, being therefore the methodology to be chosen for this application. The only disadvantage of this observer is the need to tune two gains. However, the trial-and-error approach used here gives satisfactory results after a few iterations.

Another observation that can be drawn from the results obtained is that the interval given by both observers is rather narrow, indicating that it is possible to accurately predict biomass concentration for this process in the presence of significant uncertainties. In fact, imposing a 50% and a 100% variation around the nominal value of \( X_0 \) both observers allowed to predict the final value of \( X \) with a variation, from its nominal value, of 1.4% and 3.2%, respectively.

5. CONCLUSIONS

In this work, two types of interval observers are presented based on the works of Rapaport and Dochain (2005) and Gouzé et al., (2000) in order to
handle the uncertainties on the influent glucose concentration and on the initial values of relevant state variables. A key issue associated with interval observer is the cooperativity of the observer error dynamics. An appropriate state transformation and conditions that guarantee system cooperativity have been introduced for that purpose.

The first observer is directly derived from the state transformation equation, while the second one possesses an additional gain term associated with the errors in the measurements of OTR.

Simulation experiments were conducted to validate the algorithms deduced and it was found that both observers deal well with uncertainties up to 50% and with white noise in the variables measured on-line. The interval obtained for the biomass estimation is also quite narrow, indicating that it is possible to accurately predict biomass concentration under the presence of uncertainties.

Additionally, the observer with gains from OTR exhibits a better performance regarding the noise associated with the estimations and the inclusion of the real values within the interval of observation.

Nevertheless the good results obtained, experimental validation of this work is needed and is under investigation.

Finally, and as pointed out by Rapaport and Dochain (2005) the interval observer principles used can also be applied in order to account for the uncertainties in the yield coefficients as well as for bounded noise on the outputs. Further studies taking into account the above-mentioned approaches are undergoing examination.

ACKNOWLEDGEMENTS

A. C. A. Veloso is most grateful for the financial support provided by PRODEP. This work was supported by FCT - Portuguese Science Foundation under the scope of the recSysBio Project (POCI/BIO/60139/2004).

REFERENCES


