ORIGINAL PAPER

Monitoring of fed-batch *E. coli* fermentations with software sensors

A. C. A. Veloso · I. Rocha · E. C. Ferreira

Abstract Accurate monitoring and control of industrial bioprocess requires the knowledge of a great number of variables, being some of them not measurable with standard devices. To overcome this difficulty, software sensors can be used for on-line estimation of those variables and, therefore, its development is of paramount importance. An Asymptotic Observer was used for monitoring *Escherichia coli* fed-batch fermentations. Its performance was evaluated using simulated and experimental data. The results obtained showed that the observer was able to predict the biomass concentration profiles showing, however, less satisfactory results regarding the estimation of glucose and acetate concentrations. In comparison with the results obtained with an Extended Kalman Observer, the performance of the Asymptotic Observer in the fermentation monitoring was slightly better.

Keywords Software sensors · State estimation · Nonlinear systems · Fed-batch fermentation · Biomass

Introduction

The ability to measure primary process variables, such as biomass, substrate and product concentrations, is of major academic and industrial relevance in order to guarantee the successful operation and automatic control of bioprocesses at their optimal state. Nevertheless, direct on-line measurements of these biological state variables are frequently not possible due to the lack of cheap and reliable measuring devices or probes. In fact, in many practical applications, only some of the state variables involved are available for on-line measurement. Therefore, the development of methodologies, namely software sensors [1, 2], which can provide accurate estimation of process variables that are not measurable in real time, based on on-line available data while overcoming the significant model uncertainty and the non-linear and time-varying nature of the system, is of great interest [3–11].

Two principal classes of state observers are described in the literature. The first class includes the classical observers, which are based on the perfect knowledge of both model structure and parameters, such as the Luenberger and the Kalman observers, as well as the non-linear observers. Several works have been published concerning the application of such observers, mainly the Extended Kalman Observer, to biological processes [12–17]. Nevertheless, in spite of the satisfactory results reported, an uncertainty in the model parameters can generate a large bias in the estimation of unmeasured state(s) with these methodologies [2, 10]. The second class of observers, the Asymptotic Observers [1], do not require the knowledge of the process kinetics showing, however, a potential problem concerning the dependence of the estimation convergence rate on the operating conditions [1, 13, 18].

Several applications of state observers to bioprocesses can be found in the literature [9–11, 13, 19–25] although only few examples refer to their implementation in complex bioprocesses, described by dynamic models containing several balance equations and with complex kinetics.

A. C. A. Veloso · I. Rocha · E. C. Ferreira (✉)
IBB, Institute for Biotechnology and Bioengineering,
Centre of Biological Engineering, University of Minho,
Campus de Gualtar, 4710-057 Braga, Portugal
e-mail: ecferreira@deb.uminho.pt

A. C. A. Veloso
Escola Superior Agrária de Bragança, Campus de Santa
Apolónia, Apartado 1172, 5301-855 Braganca, Portugal

The main objective of this work was the design of an Asymptotic Observer (AO) for the estimation of biomass, glucose and acetate concentrations in high-cell density fed-batch fermentation of *Escherichia coli* and to compare its performance with that of a classical observer (Extended Kalman Observer, EKO). The importance of this process for the biopharmaceutical industry is widely recognized, as *E. coli* is still the most important host microorganism used to produce many recombinant proteins. However, the on-line measurement of some state variables, namely the biomass concentration, is not easily accomplished during this process, posing additional difficulties for the implementation of control algorithms [11, 26].

The application of the above-mentioned state observers in our case required the on-line measurement of a subset of state variables (dissolved oxygen and carbon dioxide concentrations) together with broth weight and gaseous mass transfer rates. Moreover, a dynamic mathematical model of the process was used, which includes balance equations for the main state variables (biomass, glucose, acetate, dissolved oxygen and carbon dioxide concentrations).

In the coming section of this work, the dynamic model of *E. coli* fed-batch fermentation is briefly presented. In the following section, the AO for state estimation is described. In the next section, the materials and methods are described, followed by the presentation and discussion of the main results achieved with the AO concerning simulated and experimental data and a comparison with the results obtained using the EKO. Finally, a last section is devoted to some general conclusions.

Process modeling

The dynamics of a reaction network in a stirred tank bioreactor can be described by the following mass balance equations written in a matrix form as [1]:

$$\frac{d\xi}{dt} = \mathbf{K}r(\xi, t) - D\xi + \mathbf{F} \tag{1}$$

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

As previously presented [27], 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 dynamic model for fed-batch fermentation can be represented as follows:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ A \\ O \\ C \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ A \\ O \\ C \end{bmatrix} + \begin{bmatrix} 0 \\ \left(\frac{F_{in}}{W}\right)S_{in} \\ 0 \\ OTR \\ -CTR \end{bmatrix} \tag{2}$$

where X, S, A, O, and C represent biomass, glucose, acetate, dissolved oxygen, and dissolved carbon dioxide concentrations, respectively; μ_1, μ_2, and μ_3 are the specific growth rates; k_i are the yield (stoichiometric) coefficients; F_{in} and S_{in} are the substrate feed rate and the glucose concentration in the feeding solution, 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.

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

$$\frac{dW}{dt} = \mathbf{F}_{tot} \tag{3}$$

where $\mathbf{F}_{tot}$ includes weight variations due to the substrate feed rate, the amount of culture removed during sampling, base and acid additions, evaporation and mass loss from the reactor due to gas exchanges, that cannot be considered negligible in small-scale high-cell density reactors.

It should be remarked that the three above mentioned metabolic pathways represented in the mathematical model do not occur simultaneously in the cell, originating four partial models corresponding to different metabolic regimens:

- regimen A: simultaneous oxidative and fermentative growth on glucose ($\mu_1, \mu_2, > 0$; $\mu_3 = 0$)
- regimen B: oxidative growth on glucose ($\mu_1 > 0$; $\mu_2 = \mu_3 = 0$)
- regimen C: simultaneous oxidative growth on acetate and glucose ($\mu_1, \mu_3, > 0$; $\mu_2 = 0$)
- regimen D: oxidative growth on acetate ($\mu_3 > 0$; $\mu_1 = \mu_2 = 0$)

Observability of *E. coli* fed-batch model

The AO allows the estimation of the missing state variables even when the process is not exponentially observable [1], being required that the number of the measured state variables is equal or greater than the rank of the matrix $\mathbf{K}$.

Based on these facts, the possibility of applying AO was studied for nine different combinations of measured and estimated variables, among the five state variables (X, S, A,

Table 1 Observability of the model for different combinations of the measured and estimated variables using the AO

Case	Measured variables	Estimated variables	Observability
1	A, O, C	X, S	FM
2	S, A, O	X, C	FM
3	S, O, C	X, A	FM
4	O, C	X, S, A	PM
5	A, O	X, S, C	PM
6	S, O	X, A, C	PM
7	S, C	X, A, O	PM
8	S, A	X, O, C	PM
9	A, C	X, S, O	PM

OTR, CTR, W and *F* are measured on-line for all cases

FM full model, *PM* partial model

O and C). Each of those cases was classified as observable in the situation when the full model (FM) described by Eq. 2 is used, observable only under some of the regimens described previously (described by a partial model—PM), or as not observable. Those results are illustrated in Table 1, showing that if three reactions are to be considered, the number of measured variables must be equal or greater than 3. Therefore, all the combinations of three measured variables are theoretically possible. The same occurs when a partial model is considered together with the measurement of two state variables, as illustrated in Table 1.

Asymptotic Observer (AO)

The idea behind the AO consists in using the structure of the dynamic model to obtain a part of it in a form that is independent of the kinetics. These observers allow reconstructing the missing state variables even the process is not exponentially observable and the kinetics are unknown. The following additional assumptions should also be made for the design of these observers: (1) the yield coefficients (matrix $\mathbf{K}$) are known; and (2) the number q of measured state variables is equal to or greater than the rank of the matrix $\mathbf{K}$: $q \geq p = \text{rank } (\mathbf{K})$ [1].

Considering a partition in the state variables vector ξ induced by the measured (ξ_1) and unmeasured (ξ_2) variables the dynamic model can be re-written as follows:

$$\frac{d\xi_1}{dt} = \mathbf{K}_1 \mathbf{r}(\xi, t) - D\xi_1 + \mathbf{F}_1 \tag{4a}$$

$$\frac{d\xi_2}{dt} = \mathbf{K}_2 \mathbf{r}(\xi, t) - D\xi_2 + \mathbf{F}_2 \tag{4b}$$

The following state transformation can then be defined:

$$\mathbf{Z} \equiv \xi_2 - \mathbf{K}_2 \mathbf{K}_1^{-1} \xi_1 \tag{5}$$

where $\mathbf{K}_1^{-1}$ is the pseudo-inverse of the matrix $\mathbf{K}_1$, considering that $\mathbf{K}_1$ has full rank. $\mathbf{K}_1$ and $\mathbf{K}_2$ are obtained from the matrix $\mathbf{K}$ applying the induced partition.

The dynamics of $\mathbf{Z}$ are independent of the reaction rates $\mathbf{r}(\xi, t)$:

$$\frac{d\mathbf{Z}}{dt} = -D\mathbf{Z} - \mathbf{K}_2 \mathbf{K}_1^{-1} \mathbf{F}_1 + \mathbf{F}_2 \tag{6}$$

Finally, the equation of the AO is given by:

$$\frac{d\hat{\mathbf{Z}}}{dt} = -D\hat{\mathbf{Z}} - \mathbf{K}_2 \mathbf{K}_1^{-1} \mathbf{F}_1 + \mathbf{F}_2 \tag{7a}$$

$$\hat{\xi}_2 = \hat{\mathbf{Z}} + \mathbf{K}_2 \mathbf{K}_1^{-1} \xi_1 \tag{7b}$$

where, $\hat{\mathbf{Z}}$ and $\hat{\xi}_2$ are the on-line estimations of $\mathbf{Z}$ and ξ_2, respectively.

The most important advantage of this observer is its completely independence of the process kinetics. However, the speed of convergence of the estimation is completely determined by the experimental condition through the value of the dilution rate, implying that $D(t)$ can not remain equal to zero for an excessively long period of time [1, 2].

Taking into account the results obtained from the observability study carried out and since the dissolved oxygen and carbon dioxide concentrations are variables on-line accessible during industrial fermentation runs using standard on-line instruments, case 4 is given as an example for the design of this observer. Therefore, the following state partition can be chosen: $\xi_1^T = [O \quad C]$ and $\xi_2^T = [X \quad S \quad A]$.

In this case, the application of the AO requires the use of partial models. Therefore, the matrix used in the state transformation of Eq. 5, regimen A, will be:

$$\mathbf{K}_2 \mathbf{K}_1^{-1} = \begin{bmatrix} 1 & 1 \\ -k_1 & -k_2 \\ 0 & k_3 \end{bmatrix} \begin{bmatrix} -k_5 & -k_6 \\ k_8 & k_9 \end{bmatrix}^{-1} = \begin{bmatrix} \alpha_1 & \alpha_4 \\ \alpha_2 & \alpha_5 \\ \alpha_3 & \alpha_6 \end{bmatrix} \tag{8}$$

and for regimen C:

$$\mathbf{K}_2 \mathbf{K}_1^{-1} = \begin{bmatrix} 1 & 1 \\ -k_1 & 0 \\ 0 & -k_4 \end{bmatrix} \begin{bmatrix} -k_5 & -k_7 \\ k_8 & k_{10} \end{bmatrix}^{-1} = \begin{bmatrix} \alpha_1 & \alpha_4 \\ \alpha_2 & \alpha_5 \\ \alpha_3 & \alpha_6 \end{bmatrix} \tag{9}$$

The observer is given by the following equations:

$$\frac{d}{dt} \begin{bmatrix} \hat{Z}_1 \\ \hat{Z}_2 \\ \hat{Z}_3 \end{bmatrix} = -D \begin{bmatrix} \hat{Z}_1 \\ \hat{Z}_2 \\ \hat{Z}_3 \end{bmatrix} - \begin{bmatrix} \alpha_1 & \alpha_4 \\ \alpha_2 & \alpha_5 \\ \alpha_3 & \alpha_6 \end{bmatrix} \begin{bmatrix} OTR \\ -CTR \end{bmatrix}$$
$$+ \begin{bmatrix} 0 \\ \left(\frac{F_{in}}{W}\right) S_{in} \\ 0 \end{bmatrix} \tag{10a}$$

$$\begin{bmatrix} \hat{X} \\ \hat{S} \\ \hat{A} \end{bmatrix} = \begin{bmatrix} \hat{Z}_1 \\ \hat{Z}_2 \\ \hat{Z}_3 \end{bmatrix} + \begin{bmatrix} \alpha_1 & \alpha_4 \\ \alpha_2 & \alpha_5 \\ \alpha_3 & \alpha_6 \end{bmatrix} \begin{bmatrix} O \\ C \end{bmatrix} \qquad (10b)$$

Another advantage of this kind of observers is that there are no tuning parameters since the initial values of the **Z** variables can be obtained directly from Eqs. 7b or 10b, by replacing the values of the estimated variables by their experimental initial values, usually known. However, in the cases where partial models are used, the different regimens are described using different mathematical equations, being necessary to alternate between algorithms. Therefore, a detection mechanism to identify which regimen the cells exhibit in a given moment is needed that can be based on the values of the specific growth rates or on changes of acetate concentration. In this study, a detection mechanism based on the specific growth rate was used. For obtaining the specific growth rates μ_1 to μ_3, the estimator deduced in Rocha et al. [29], based on the formulation proposed by Bastin and Dochain [1] and reformulated by Pomerleau and Perrier [30], was used.

Simulation and experimental validation

Experimental conditions and equipment

The *E. coli* fed-batch fermentations were performed in a 5 L Biostat MD fermenter from B. Braun Biotech (Germany) connected to a Digital Control Unit, with an initial medium weight of 3 kg with an initial concentration of glucose of 5 g kg^{-1}. A pre-programmed profile of feeding rates was used for the addition of 2 L of feeding solution with a concentration of 250 g kg^{-1} of glucose. Aqueous solutions of phosphoric acid (85%) and ammonia (25%) were added during the fermentations for pH control. The set-points for the environmental variables were 37 °C for temperature, pH 7, and 30% of saturation for dissolved oxygen. The dissolved oxygen and carbon dioxide concentrations were monitored with commercial sensors (Metter Toledo, Switzerland). For offgas measurements, a mass spectrometer Dymaxion (DM 200 M) from Ametek (USA) connected to the exhaust gas line of the fermenter and to the inlet, aeration line was used. The gas transfer rates were calculated from these gas analysis data. The total weights of the fermenter and of the feeding solution reservoir were measured using two analytical balances (Metter Toledo, Switzerland). Biomass concentration was monitored by measuring OD at 600 nm. A modular liquid chromatograph (Jasco, Japan) with refractive index and UV/Vis detectors was used for measuring glucose and acetic acid concentrations, respectively, with an organic acids column (Chrompack 257001).

Hardware and software

A program developed in LabVIEW (version 7.0) was used to acquire and store the data obtained from the fermenter Digital Control Unit, concerning the F_{in}, W, O and C.

The model simulations were performed by solving the differential equations of Eq. 2 using the MATLAB version 7.1 (The MathWorks, Inc., USA) subroutine ODE23s. The implementation of the observers using both experimental and simulated data was conducted using the Euler integration method. The calculations regarding the observability of the model, together with most of the mathematical operations behind the design of the state observers were performed using the Symbolic Math toolbox running in MATLAB 7.1.

AO performance

In order to study the robustness of the AO, cases 1 and 4 (Table 1) were used. For these cases, simulated "pseudo-real" values of the state variables, obtained by numerical integration of the differential equation of Eq. 2, were used. These "pseudo-real" values were then corrupted with white noise, according to the standard deviations typically found in this process at the authors' lab, originating "pseudo-experimental" values. Then, the observer algorithm was used to obtain the "estimated" variables from the "pseudo-experimental" data corresponding to the measured variables. Experimental validation was also carried out for case 4, which was selected since it required the measurement of few state variables, which are commonly measured in industry as mentioned before.

The AO performance was evaluated by calculating the quadratic error between experimental and estimated data, according to the following equation:

$$\mathrm{dif}_\xi = \sum_{j=1}^{np} \left(\frac{\xi_{\exp,j} - \xi_{\mathrm{est},j}}{\bar{\xi}_{\exp,j}} \right)^2 \qquad (11)$$

where *np* is the number of experimental points and $\xi_{\exp}$ and ξ_{est} are experimental and estimated values of the state variable ξ and $\bar{\xi}_{\exp,j}$ is the average of the $\xi_{\exp}$ values.

In order to illustrate the robustness of the AO a comparison with the performance of a classical observer, such as the EKO (see "Appendix" for detailed information), was made. EKO was selected since it still is often used, mostly because the algorithm can be derived directly from the state space model [31].

The AO and EKO performances, regarding the simulation results obtained in the estimation of biomass, glucose and acetate concentrations, calculated using Eq. 11, for cases 1 and 4, are shown in Table 2. The analysis of the results obtained show that the state variables are well

Table 2 Performance indexes obtained for AO for cases 1 and 4 of Table 1, using simulated data

Observer	Case	dif_X	dif_S	dif_A
AO	1	0.196	1.42	–
	4	1.09×10^{-2}	0.311	0.389
EKO	1	4.55×10^{-3}	4.80×10^{-2}	–
	4	4.81×10^{-3}	0.101	0.165

Comparison with the results obtained with EKO

Table 3 Performance indexes for AO for two fermentations (Exp_1 and Exp_2)

Observer	Case	dif_X	dif_S	dif_A
AO	Exp_1	8.02	656	370
	Exp_2	4.42	734	738
EKO	Exp_1	8.34	442	369
	Exp_2	17.7	2,069	738

Comparison with the results obtained with EKO

estimated, in spite of the introduction of noise, showing the robustness of both AO and EKO, being the AO more susceptible to noise effects.

The experimental validation of both observers was performed using two experiments (Exp_1 and Exp_2). Table 3 shows the performance indexes obtained with AO and EKO. The results obtained show that the observers studied presented a satisfactory performance, for both experiments, allowing the adequate estimation of biomass concentration. However, the performance of both observers is less satisfactory for the estimation of glucose and acetate concentrations. In Fig. 1, the on-line data acquired during Exp_2 are shown. The variables measured on-line were O, C, CTR, OTR, F_{in} and W. Figure 2 shows the good agreement obtained for both observers, between the estimated biomass concentration values and the off-line measurements, although some divergences can be observed at the end of the fermentation.

In general, the performance of the experimental implementation of AO was quite satisfactory and superior to that of the EKO. In fact, and although EKO are still often used [31], they must be developed based on the perfect knowledge of the system structure and parameters, especially those related with the process kinetics. Therefore, their implementation will be successful if the process dynamics are (almost) exactly described by the kinetic model equations and if the observers are initialized, close enough to

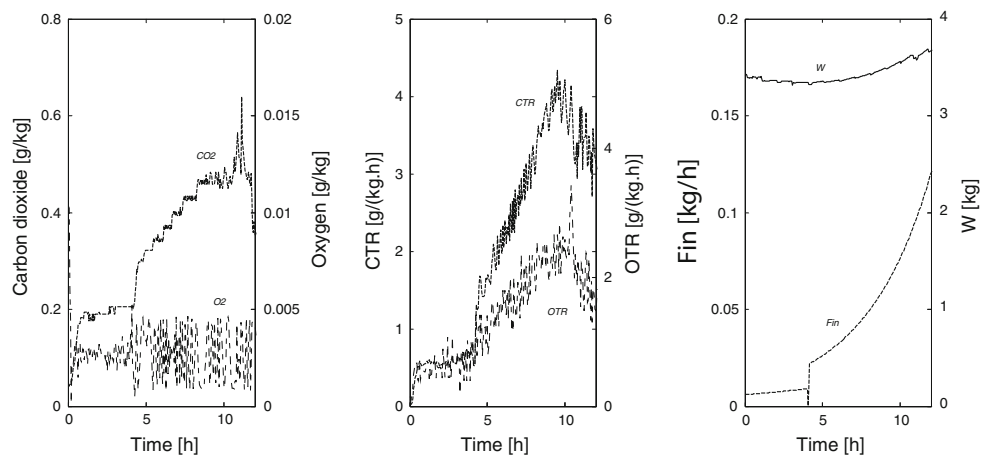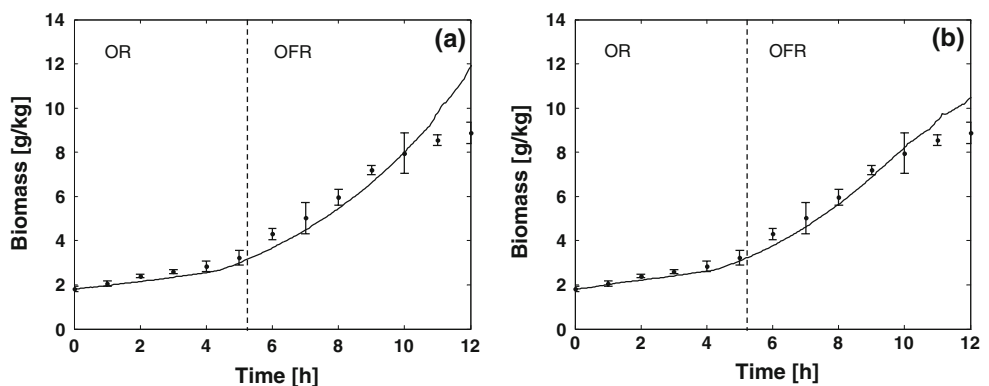**Fig. 1** State observation in a fed-batch fermentation of *E. coli*: on-line data

Fig. 2 State observation for Exp_2. Estimated variable (biomass) obtained with: **a** EKO and **b** AO. *Symbols* represent real values and *lines* correspond to the estimated values. (*OR* oxidative regimen—regimen B, *OFR* oxidative-fermentative regimen—regimen A)

the true state values [2]. Since the implementation of the AO does not require the knowledge of the kinetics structure and parameters, where most of the model uncertainties usually rely, it is not surprising that these Observers behave better with real data [32].

Moreover, since a model linearization step is used in the EKO design, the stability and convergence properties are essentially local and valid around an equilibrium point, and therefore it is difficult to assure their stability over a wide range of operation conditions [2, 31]. Even though the AO does not contain any estimation error driving term its asymptotic stability is guaranteed even if the dynamical model is unstable [2, 32]. However, a note of caution should be drawn for the cases, not contemplated in this paper, when the dilution rate [$D(t)$] is kept close to zero for long periods, since the rate of convergence of the estimation fully depends on the values of that variable[1, 2]. In those cases, the performance of AO is expected to be much poorer.

Furthermore, although the kinetics of the process may be considered unknown in these observers, the state variables estimation requires the knowledge of the correct reaction scheme, with known stoichiometric coefficients (matrix $\mathbf{K}$). Consequently, uncertainties on these model parameters together with the noise eventually associated with on-line measurements can generate a large bias in the estimation of the unmeasured state variables [2]. These facts can help to explain the lower performance on the estimation of biomass at the end of the fermentation. In order to improve the accuracy of the stoichiometric coefficients, which is indispensable for the efficient implementation of such observers, an *ad-hoc* experimental design can be carried out.

It should also be remarked that the AO requires, in order to achieve convergence, relatively high sampling intervals (3 min), compatible with on-line practical implementation requirements of some instruments, like FIA [28].

As known, state estimation techniques are of major importance in several areas of science and engineering, especially in biotechnology, since hardware sensors are extremely expensive and have to deal with stringent operating conditions, like sterilization and long processing times. The present work shows that the AO can be successfully applied to the on-line monitoring of a complex *E.coli* fermentation, being an effective alternative to the use of hardware sensors.

Conclusions

The design of monitoring and control algorithms to improve the performance of bioprocesses is of major importance. However, it is difficult to find inexpensive and robust commercial sensors that allow real-time monitoring of important process variables, namely the biomass

concentration, required for the implementation of control and optimization strategies. In fact, nowadays the biomass concentration is still mostly measured using off-line laboratory analysis, making it of limited use for control purposes. Therefore, the development of software sensors is of paramount importance.

In this work, an AO was applied for the monitoring of fed-batch *E. coli* fermentation and its performance and flexibility was evaluated and compared with a classical observer. In addition, the robustness of the AO algorithm concerning the effect of experimental noise in the measured variables was checked with numerical simulations and experimental data.

The observer algorithm studied only requires on-line measurements of dissolved oxygen and carbon dioxide concentrations, together with the gaseous transfer rates, which represent common measurements both in industrial and academic facilities. The sampling frequency required is also compatible with most existing data acquisition systems.

The results obtained showed that the Asymptotic Observer is rather flexible regarding the choice of the measured variables, exhibiting a satisfactory performance in estimating variables from experimental data.

Appendix: Extended Kalman Observer

Bastin and Dochain [1] proposed the following general equation for a state observer of non-linear systems described by Eq. 1, when a sub-set of the state vector, ξ_1 is obtained by on-line measurement:

$$\frac{d\hat{\xi}}{dt} = \mathbf{K}r\left(\hat{\xi}, t\right) - D\hat{\xi} + \mathbf{F} + \mathbf{\Omega}\left(\hat{\xi}\right)\left[\xi_1 - \hat{\xi}_1\right] \quad (12)$$

where $\hat{\xi}$ and $\hat{\xi}_1$ represent the on-line estimation of ξ and ξ_1, respectively. $\mathbf{\Omega}\left(\hat{\xi}\right)$ is the gain matrix of dimension $n \times q$, function of $\hat{\xi}$, and q is the number of on-line measured variables.

This last equation can be seen as an extension of the model given by Eq. 1 with an additional driving term which is proportional to the observation error of the measured part of the state. The state observer design problem consists then in choosing the more adequate gain matrix, $\mathbf{\Omega}\left(\hat{\xi}\right)$.

In order to obtain the Extended Kalman Observer (EKO), the following assumptions are made: (1) a full knowledge of the model is available, i.e. the structure of the reaction kinetics $r(\xi,t)$ is completely known and the

numerical values of all the coefficients involved in the model (yield and kinetic coefficients) are given; and (2) D and $\mathbf{F}$ are known on-line, together with a q subset of state variables. This vector of state variables measured ξ_1 is related to the state of the system as follows:

$$\xi_1 = \mathbf{L}\xi \qquad (13)$$

where the $q \times n$ matrix $\mathbf{L}$ is an elementary matrix which selects the measured components of ξ. On the other hand, the vector of unmeasured states is denoted ξ_2, so that (ξ_1, ξ_2) constitutes a partition of ξ.

If the measured variables are O and C (as discused for the AO), the following state partition can be chosen: $\xi_1^T = [O \quad C]$ and $\xi_2^T = [X \quad S \quad A]$. The matrix $\mathbf{L}$ is as follows:

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (14)$$

To solve the problem of selecting an appropriate matrix $\mathbf{\Omega}\left(\hat{\xi}, t\right)$ (Eq. 12), the observation error, $e = \xi - \hat{\xi}$, and its dynamics can be used [1]. Considering a linearized tangent approximation of the dynamic model of the observation error around $e = 0$, the following equation can be obtained [1]:

$$\frac{de}{dt} = \left[\mathbf{M}\left(\hat{\xi}\right) - \mathbf{\Omega}\left(\hat{\xi}\right)\mathbf{L}\right]e \qquad (15)$$

with:

$$\mathbf{M}\left(\hat{\xi}\right) \equiv \mathbf{K}\left[\frac{\partial \mathbf{r}(\xi, t)}{\partial \xi}\right]_{\xi=\hat{\xi}} - D\mathbf{I}_N \qquad (16)$$

where $\mathbf{I}_N$ is the $n \times n$ identity matrix. For the *E. coli* model used matrix $\mathbf{M}\left(\hat{\xi}\right) = \mathbf{M}(\hat{X}, \hat{S}, \hat{A}, \hat{O}, \hat{C})$ and:

$$\left[\frac{\partial \mathbf{r}(\xi, t)}{\partial \xi}\right]_{\xi=\hat{\xi}} = \begin{bmatrix} \frac{\partial(\mu_1\hat{X})}{\partial\hat{X}} & \frac{\partial(\mu_1\hat{X})}{\partial\hat{S}} & \frac{\partial(\mu_1\hat{X})}{\partial\hat{A}} & \frac{\partial(\mu_1\hat{X})}{\partial\hat{O}} & \frac{\partial(\mu_1\hat{X})}{\partial\hat{C}} \\ \frac{\partial(\mu_2\hat{X})}{\partial\hat{X}} & \frac{\partial(\mu_2\hat{X})}{\partial\hat{S}} & \frac{\partial(\mu_2\hat{X})}{\partial\hat{A}} & \frac{\partial(\mu_2\hat{X})}{\partial\hat{O}} & \frac{\partial(\mu_2\hat{X})}{\partial\hat{C}} \\ \frac{\partial(\mu_3\hat{X})}{\partial\hat{X}} & \frac{\partial(\mu_3\hat{X})}{\partial\hat{S}} & \frac{\partial(\mu_3\hat{X})}{\partial\hat{A}} & \frac{\partial(\mu_3\hat{X})}{\partial\hat{O}} & \frac{\partial(\mu_3\hat{X})}{\partial\hat{C}} \end{bmatrix} \qquad (17)$$

The design of the EKO is reduced to the quadratic optimization problem of finding the matrix $\mathbf{\Omega}\left(\hat{\xi}, t\right)$ that minimizes the mean square observation error taking into account the constraint of the linear tangent error model (Eqs. 15 and 16). The solution of this optimization problem is given by [1]:

$$\mathbf{\Omega}\left(\hat{\xi}, t\right) = \mathbf{R}\left(\hat{\xi}, t\right)\mathbf{L}^T \qquad (18)$$

where the $n \times n$ square symmetric matrix $\mathbf{R}\left(\hat{\xi}, t\right)$ is generated by the Riccati equation:

$$\frac{d\mathbf{R}}{dt} = -\mathbf{R}\mathbf{L}^T\mathbf{L}\mathbf{R} + \mathbf{R}\mathbf{M}^T\left(\hat{\xi}, t\right) + \mathbf{M}\left(\hat{\xi}, t\right)\mathbf{R} \qquad (19)$$

The observer is then written from Eqs. 1 and 12 as:

$$\frac{d}{dt}\begin{bmatrix} \hat{X} \\ \hat{S} \\ \hat{A} \\ \hat{O} \\ \hat{C} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}\hat{X} - D\begin{bmatrix} \hat{X} \\ \hat{S} \\ \hat{A} \\ \hat{O} \\ \hat{C} \end{bmatrix}$$
$$+ \begin{bmatrix} 0 \\ \left(\frac{F_{in}}{W}\right)S_{in} \\ 0 \\ OTR \\ -CTR \end{bmatrix} + \begin{bmatrix} R_1 & R_2 \\ R_3 & R_4 \\ R_5 & R_6 \\ R_7 & R_8 \\ R_9 & R_{10} \end{bmatrix}\begin{bmatrix} O - \hat{O} \\ C - \hat{C} \end{bmatrix} \qquad (20)$$

where the elements of $\mathbf{R}$ are time dependent and obtained by the solution of Eq. 19. Therefore, the only tuning parameters for this observer are the initial values of the elements of that matrix.

References

1. Bastin G, Dochain D (1990) On-line estimation and adaptive control of bioreactors. Elsevier, Amsterdam
2. Dochain D (2003) State and parameter estimation in chemical and biochemical processes: a tutorial. J Proc Control 13:801–818
3. Cheruy A (1997) Software sensors in bioprocess engineering. J Biotechnol 52:193–199
4. Lombardi M, Fiaty K, Laurent P (1999) Implementation of observer for on-line estimation of concentration in continuous-stirred membrane bioreactor: Application to the fermentation of lactose. Chem Eng Sci 54:2689–2696
5. Assis AJ, Filho RM (2000) Soft sensors development for on-line bioreactor state estimation. Comput Chem Eng 24:1099–1103
6. Bogaerts P, Vande Wouwer A (2004) Parameter identification for state estimation—application to bioprocess software sensors. Chem Eng Sci 59:2465–2476
7. Aguilar-Lopez R, Maya-Yescas R (2005) State estimation for nonlinear systems under model uncertainties: a class of sliding-mode observers. J Proc Control 15:363–370
8. Jenzsch M, Simutis R, Eisbrenner G, Stückrath I, Lübbert A (2006) Estimation of biomass concentrations in fermentation processes for recombinant protein production. Bioprocess Biosyst Eng 29:19–27
9. Nadri M, Trezzani I, Hammouri H, Dhurjati P, Longin R, Lieto J (2006) Modeling and observer design for recombinant *Escherichia coli* strain. Bioprocess Biosyst Eng 28:217–225
10. Chachuat B, Bernard O (2006) Probabilistic observers for a class of uncertain biological processes. Int J Robust Nonlinear Control 16:157–171
11. Sundström H, Enfors S-O (2008) Software sensors for fermentation processes. Bioprocess Biosyst Eng 31:145–152
12. Dubach A, Märkl H (1992) Application of an extended kalman filter method for monitoring high density cultivation of *Escherichia coli*. J Ferment Bioeng 73:396–402
13. Ph Bogaerts (1999) A hybrid asymptotic-Kalman observer for bioprocesses. Bioproc Eng 20:249–255
14. Neeleman R, van den End EJ, Van Boxtel AJB (2000) Estimation of the respiratory quotient in a bicarbonate buffered batch cell cultivation. J Biotechnol 80:85–95
15. Li J, Xu NS, Su WW (2003) Online estimation of stirred-tank microalgal photobioreactor cultures based on dissolved oxygen measurement. Biochem Eng J 14:51–65
16. Arndt M, Hitzmann B (2004) Kalman filter based glucose control at small set points during fed-batch cultivation of *Saccharomyces cerevisiae*. Biotechnol Progr 20:377–383

17. Arndt M, Kleist S, Miksch G, Friehs K, Flaschel E, Trierweiler J, Hitzmann B (2005) A feedforward-feedback substrate controller based on a Kalman filter for a fed-batch cultivation of *Escherichia coli* producing phytase. Comput Chem Eng 29:1113–1120

18. Dochain D (2001) State observation and adaptive linearizing control for distributed parameter (bio) chemical reactors. Int J Adapt Control Signal Process 15:633–653

19. Farza M, Hammouri H, Jallut C, Lieto J (1999) State observation of a nonlinear system: application to (bio)chemical processes. AIChE J 45:93–106

20. Zhang T, Guay M (2002) Adaptive nonlinear observers of microbial growth processes. J Proc Control 12:633–643

21. Dochain D (2003) State observers for processes with uncertain kinetics. Int J Control 76:1483–1492

22. Rapaport A, Dochain D (2005) Interval observers for biochemical processes with uncertain kinetics and inputs. Math Biosci 193:235–253

23. Raissi T, Ramdani N, Candau Y (2005) Bounded error moving horizon state estimator for non-linear continuous- time systems: application to a bioprocess system. J Proc Control 15:537–545

24. Hulhoven X, Vande Wouwer A, Bogaerts Ph (2007) State observer scheme for joint kinetic parameter and state estimation. Chem Eng Sci. doi:10.1016/j.ces.2007.11.042

25. Ko C-L, Wang F-S (2007) On-line estimation of biomass and intracellular protein for recombinant *Escherichia coli* cultivated in batch and fed-batch modes. J Chin Inst Chem Engrs 38:197–203

26. Gnoth S, Jenzsch M, Simutis R, Lübbert A (2008) Control of cultivation processes for recombinant protein production: a review. Bioprocess Biosyst Eng 31:21–39

27. Rocha I, Ferreira EC (2004) Yield and kinetic parameters estimation for model reduction in a recombinant *E. coli* fermentation. ESCAPE–14, Lisbon (Proceedings)

28. Rocha I, Ferreira EC (2002) On-line simultaneous monitoring of glucose and acetate with FIA during high cell density fermentation of recombinant *E. coli*. Anal Chim Acta 462:293–304

29. Rocha I, Veloso ACA, Ferreira EC (2006) Design of estimators for specific growth rate control in a fed-batch *E. coli* fermentation. In: I Troch, F Breitenecker (eds) 5th MATHMOD Vienna, vol 2: full papers CD, 1.1–1.9, (ISBN 3-901608-30-3)

30. Pomerleau Y, Perrier M (1990) Estimation of multiple specific growth rates in bioprocesses. AIChE J 136:207–215

31. Perrier M, Feyo de Azevedo S, Ferreira EC, Dochain D (2000) Tuning of observer-based estimators: theory and application to the on-line estimation of kinetic parameters. Control Eng Pract 8:377–388

32. Bernard O, Gouzé J-L (2004) Closed loop observers bundle for uncertain biotechnological models. J Proc Control 14:765–774