UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Seminário do Programa de Pós-Graduação em Engenharia Química

VIII-Oktoberforum – PPGEQ

20, 21, 22 e 23 de outubro de 2009

ESTIMADORES DE ESTADO: UMA FERRAMENTA PARA A MELHORIA OPERACIONAL DE BIORREATORES

Nina Paula G. Salau[1], Fabio C. Diehl[1], Jorge O. Trierweiler[1], Argimiro R. Secchi[2]

[1] GIMSCOP
Departamento de Engenharia Química, Universidade Federal do Rio Grande do Sul (UFRGS)
R. Eng. Luis Englert, s/n. Campus Central. CEP: 90040-040 - Porto Alegre - RS - BRASIL,
E-MAIL: {ninas, fcdiehl, jorge}@enq.ufrgs.br
[2] PEQ/COPPE
Programa de Engenharia Química, Universidade Federal do Rio de Janeiro (UFRJ)
CxP. 68502, CEP: 21941-972, Rio de Janeiro, RJ, Brazil,
E-MAIL: arge@ peq.coppe.ufrj.br

Abstract: *Nowadays many advanced control tools are available to improve industrial processes (e. g. virtual analyzers, state estimators, predictive and multivariable controllers, etc.). Particularly, state estimation techniques have a long development history focused mainly to supply the lack of system measurements. Between the applications enclosed by state estimators we can highlight: more detailed process monitoring, mathematical model fitting and update, transient data reconciliation and feedback control. Although state estimators have been successfully used in many chemical processes, only few works report the application of this tool in bioprocesses. In this paper we evaluate the applications of three state estimation techniques in a biorrector: Extended Kalman Filter (EKF), Constrained Extended Kalman Filter (CEKF) and Moving Horizon Estimator (MHE). As benchmark case study we have chosen the continuous glucose fermentation with Zymomonas mobilis bacteria to produce ethanol. Our results clearly show the relevance of state estimators as a tool to improve the bioprocesses operation.*

Keyword: *ethanol, nonlinear dynamic behavior, observability analysis, state estimation, bioreactor control.*

1. Introduction

The biotechnology has an old role in the humanity history. Its first known applications were in simple empirical fermentations to produce nutritious products as wine and bread. Other applications arise during the human evolution and close to the end of XIX century end appear the industrial fermentation to synthesize ethanol and lactic acid (NAJAFPOUR, 2007). Maybe the revolutionary mark in biotechnology had been the penicillin discovered by Alexander Flemming in the 20's decade. With the penicillin was possible to produce antibiotics to combat pathogenic microorganisms. Even so bigger industrial motivation with biotechnology comes with the world wars. The wars pushed on some countries as England, United States and Germany to begin large scale production of antibiotics, glycerol, acetone, etc (STANBURY, 1995). After this, advances in chemical DNA synthesis and in genetic manipulations originate the genetic engineering that made possible the molecular hybridization of live organisms. Actually the biotechnology is in constant development, and this development depends more and more of the biological, chemical and engineering conjunction (DORAN, 1995).

According to Alford (2006), improvements to bioprocess productivity generally come from two sources: cell lines and process control. Historically, the most profit way to obtain productive increase, in a bioprocess plant, is linked to the fermentative process strain evolution (AYNSLEY *et al*., 1993). However, the last decades reveal significant advances in the control process area for to reduce production costs, conversion reaction increment, and maintains the product quality (RANI & RAO, 1999). This alternative comes to possibility that the plant be conducted to the optimum conditions for a supervisory system, and this system can be viewed as an experience operator managing the plant (BAKHTADZE, 2004). This mean furnishing the right concentration of nutrients to the culture (e.g. nitrogen, oxygen, phosphorous, sulfur), removing toxic metabolic products (e.g. CO2), and controlling important internal cellular parameters (e.g. temperature, pH) (ALFORD, 2006). To reach this goal is necessary to feedback information from the process to a control system approach in real time. Some variables as temperature, pH, pO2, agitation, foam level and flow rate are relatively easy to measure in real time. However, some essential variables (e.g. substrate, biomass and product concentration) that indicate the fermentation state are not direct on line measure due to the lack of satisfactory measuring devices (GADKAR et al., 2005). A method to solve the real time quantify difficult problem consist in estimation techniques (GONZALEZ, 1999).

Of the nonlinear state estimators, the EKF has received the most attention due to its relative simplicity and low computational effort, demonstrating effectiveness in handling some nonlinear problems. Neverthless, it can give unreliable estimates if the system presents a high nonlinearity degree and has states subject to hard constraints (e.g. nonnegative concentrations or pressures). Due to the development of effective solvers for nonlinear optimization problems, optimization-based state estimators, such as the MHE and the CEKF, simpler and computationally less demanding, has become an interesting alternative to common approaches such as EKF due to the possibility to consider states physical constraints into an optimization problem (GESTHUISEN *et al.*, 2001). However, due to the higher mathematical complexity introduced by nonlinearity and the higher computational effort, optimal nonlinear state estimation techniques are generally not used in practice. Figure 1 shows a state estimator idea, where *x*, *u* and *y* represents the states, inputs and outputs, respectively.

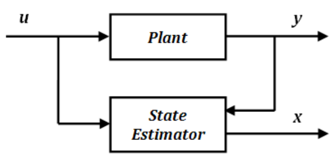

Figure 1. Generic state estimator.

In this paper we address the state estimation techniques focused on bioprocesses. The second section deals with the state estimator applications. The EKF, CEKF and MHE formulations are invetigated in Section 3. Section 4 is concerned with the computational methods required in this work. A motivating case study of a *Zymomonas mobilis* continuous fermentation is presented in Section 5. Section 6 shows the results of the state estimators applied to our example. Finally the conclusions are discussed in Section 7.

2. Applicability of State Estimators

2.1. Virtual Analyser & Feedback Control

Process monitoring and control require real-time information on the state variables of a process to ensure proper operation of the plant (ENGELL, 2006). To calculate control actions the controller require to feedback key process variables, such as product and biomass concentration or indices as growing rate. Measurements of these indices are rarely available on-line and are usually obtained by laboratory sample analyses. These laboratory measurements are available infrequently with substantial time delays between sampling times. Thus, these measurements must be estimated or inferred to be valid to control design. State estimation techniques can be used to implement a state-feedback control. This functionality enables the robust operation of the control system at various sample rates of plant-product quality measurements.

2.2. Parameters fitting and update

A suitable design of state estimators requires a representative model for capturing the plant behavior. State estimation theory can be use also to estimate unknown parameters of a model to improve its fidelity. In this case, the state vector must be augmented with the constant parameter vector (SIMON, 2006). Differential equations for the estimated parameters must be added to the model. As it is assumed that parameters are constants, the right side of corresponding differential equations is zero. Another way to estimate parameters via state estimation requires a dynamic model for each of unknown parameters to be estimated such as "*random walk*" and "*random ramp*". Once appropriate parameter models are chosen, a state estimator is used to estimate the process parameters that appear as a subset of the state variables of the combined process and parameter models. This method has been used widely in chemical and biological engineering (SOROUSH, 1998).

2.3. Transient Data Reconciliation

Poor measurements can lead to estimates that violate the conservation laws used to model the system. However, states estimators (e.g. EKF) can also be applied as a tool for data pre-processing and transient data reconciliation in order to reduce the inaccuracy of process data due to measurement errors (NARASIMHAN & JORDACHE, 2000).

In Rao & Rawlings (2002) we can find a comparison between MHE and EKF in the problem of detecting the location and magnitude of a leak (gross error) in the wastewater treatment process. While MHE is able to provide a fairly accurate estimate of the total losses, the Kalman filter underestimates the total losses.

The application of state estimation techniques are summarized in Figure 2.

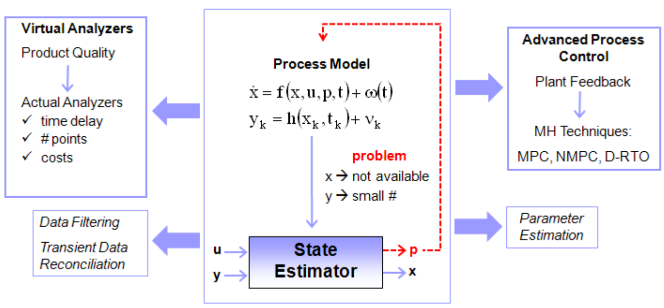

Figure 2. Application of state estimation techniques.

3. EKF, CEKF and MHE

Consider the following continuous-time system with discrete-time measurements:

$$\dot{x} = f(x,u,t) + \omega(t) \tag{1}$$

$$y_k = h(x_k, t_k) + v_k \tag{2}$$

where $x \in \mathbb{R}$ denotes the state vector, $u \in \mathbb{R}$ the known system input vector and $y \in \mathbb{R}$ the measured output. The process-noise, ω, and the measurement-noise, v_k, are assumed to be white Gaussian random process with zero mean and covariance Q and R_k, respectively.

The system is linearized at each time step to obtain the local state-space matrices as below:

$$F = \left(\frac{\partial f}{\partial x} \right)_{x,u,t,p} \quad (3)$$

$$H = \left(\frac{\partial h}{\partial x} \right)_{x,u,t,p} \quad (4)$$

The discrete version of Eq. 1 can be written as follows:

$$x_k = \varphi_{k-1} x_{k-1} + \omega_{k-1} \quad (5)$$

$$y_k = h_k \left(x_k, t_k \right) + \nu_k \quad (6)$$

where k denotes the discrete-time index and φ_{k-1} is the *state transition matrix* for the state at t_k given as

$$\varphi_{k-1} = e^{F(t_k - t_{k-1})} \quad (7)$$

ω_{k-1} and ν_k are assumed to be constant over the sampling period and both are modeled as random variables solely characterized by their PDFs.

The filter algorithm is initialized by the initially expected state and state covariance:

$$\hat{x}_{0|0} = E\left[x(0) \right] \quad (8)$$

$$P_{0|0} = E\left[\left(x(0) - \hat{x}_{0|0} \right) \left(x(0) - \hat{x}_{0|0} \right)^T \right] \quad (9)$$

3.1. Extended Kalman Filter

We will focus on the discrete-time version of the EKF since it is used in most applications. The equations that compose the different steps in the EKF are given below.

State Transition Equation:

$$\hat{x}_{k|k-1} = \hat{x}_{k-1|k-1} + \int_{k-1}^{k} f\left(\hat{x}, u, \tau, p \right) d\tau \quad (10)$$

Kalman Gain Equation:

$$K_k = P_{k|k-1} H_k^T \left[H_k P_{k|k-1} H_k^T + R_k \right]^{-1} \quad (11)$$

State Covariance Transition Equation:

$$P_{k|k-1} = \varphi_{k-1} P_{k-1|k-1} \varphi_{k-1}^T + Q_{k-1} \quad (12)$$

State Update Equation:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \left[y_k - h_k \left(\hat{x}_{k|k-1}, t_k \right) \right] \quad (13)$$

State Covariance Update Equation:

$$P_{k|k} = \left[I_n - K_k H_k \right] P_{k|k-1} \left[I_n - K_k H_k \right]^T + K_k R_k K_k^T \quad (14)$$

where n is the number of states.

In order to compare the EKF and the optimization-based estimators (MHE and CEKF), we first need to derive yet another form for the EKF. After combining and rearranging Eqs. (11), (12) and (14), one obtains an alternate form for the one-step state covariance equation as follows

$$\begin{aligned} P_{k|k} &= \varphi_k P_{k-1|k-1} \varphi_k^T \\ &- \varphi_k P_{k-1|k-1} H_k^T \left[H_k P_{k-1|k-1} H_k^T + R_k \right]^{-1} H_k P_{k-1|k-1} \varphi_k^T \\ &+ Q_k \end{aligned} \quad (15)$$

The equation above is the discrete Riccati equation. Details on the derivation of this alternate EKF form are found in Simon (2008).

3.2. Moving Horizon Estimator

Before explaining the CEKF formulation, the basic aspects about the MHE (MUSKE & RAWLINGS, 1994; ROBERTSON *et al*., 1996; RAO *et al*., 2003) is presented. The basic idea of MHE is to proceed with state estimation by using only the most recent N+1 measurements, where N is the time horizon size.

The moving horizon approximation of the objective function is given by

The moving horizon approximation of the objective function is given by

$$\min_{\substack{\omega_{k-N-1|k}, \cdots, \omega_{k-1|k} \\ \nu_{k-N|k}, \cdots, \nu_{k|k}}} \Psi_k^N = \begin{bmatrix} \hat{\omega}_{k-N-1|k}^T \left(P_{k-N-1|k-1} \right)^{-1} \hat{\omega}_{k-N-1|k} \\ + \sum_{j=k-N}^{k-1} \hat{\omega}_{j|k}^T \left(Q_{k-1} \right)^{-1} \hat{\omega}_{j|k} \\ + \sum_{j=k-N}^{k-1} + \sum_{j=k-N}^{k} \hat{\nu}_{j|k}^T \left(R_k \right)^{-1} \hat{\nu}_{j|k} \end{bmatrix} \quad (16)$$

subject to the equality constraints

$$\hat{x}_{k-N|k} = \hat{x}_{k-N|k-1} + \hat{\omega}_{k-N-1|k}$$

$$\hat{x}_{j+1|k} = \int_j^{j+1} f\left(\hat{x}, u, \tau \right) d\tau + \hat{\omega}_{j|k}, \quad j = k-N, \cdots, k-1 \quad (17)$$

$$y_j = h\left(\hat{x}_{j|k} \right) + \hat{\nu}_{j|k}, \quad j = k-N, \cdots, k$$

and the inequality constraints

$$\begin{aligned} & x_{\min} \le \hat{x}_{j|k} \le x_{\max} \\ & \hat{\omega}_{\min} \le \hat{\omega}_{j-1|k} \le \hat{\omega}_{\max}, \quad j = k-N, \cdots, k \\ & \hat{\nu}_{\min} \le \hat{\nu}_{j|k} \le \hat{\nu}_{\max}, \quad j = k-N, \cdots, k \end{aligned} \quad (18)$$

Rao *et al*.(2003) suggests computing the state covariance matrix equation $P_{k|k}^N$ (Eq. 31) recursively using the discrete Riccati equation.

$$\begin{aligned} P_{k|k}^N &= \varphi_k P_{k-1|k-1}^N \varphi_k^T \\ &- \varphi_k P_{k-1|k-1}^N H_k^T \left[H_k P_{k-1|k-1}^N H_k^T + R_k \right]^{-1} H_k P_{k-1|k-1}^N \varphi_k^T \\ &+ Q_k \end{aligned} \quad (19)$$

Note that the equation above is the same as for Eq. 15 applied to each horizon step.

3.3. Constrained Extended Kalman Filter

CEKF follows from the MHE when the horizon length is set to zero (GESTHUISEN *ET AL*., 2001). Zero length implies that ODEs are not considered into the optimization problem, which simplifies the complexity of solving a nonlinear dynamic optimization problem.

Setting (N=0) into the MHE optimization problem (Eq. 16), the resulting formulation is exactly the CEKF formulation problem

$$\min_{\omega_{k-1|k}, \nu_{k|k}} \Psi_k = \begin{bmatrix} \hat{\omega}_{k-1|k}^T \left(P_{k-1|k-1} \right)^{-1} \hat{\omega}_{k-1|k} \\ + \hat{\nu}_{k|k}^T \left(R_k \right)^{-1} \hat{\nu}_{k|k} \end{bmatrix} \quad (20)$$

subject to the equality constraints

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \hat{\omega}_{k-1|k}$$
$$y_k = h\left(\hat{x}_{k|k}\right) + \hat{v}_{k|k} \quad (21)$$

and inequality constraints

$$x_{\min} \le \hat{x}_{k|k} \le x_{\max}$$
$$\hat{\omega}_{\min} \le \hat{\omega}_{k-1|k} \le \hat{\omega}_{\max} \quad (22)$$
$$\hat{v}_{\min} \le \hat{v}_{k|k} \le \hat{v}_{\max}$$

4. Computational Methods

During the last decades many tools to bifurcation analysis were developed. The main tools are: AUTO, CONTENT, LOCBIF, PITCONT and BIFPACK. In this work was used a MatLab toolbox called MATCONT, developed based upon CONTENT (DHOOGE *et al*. 2006).

The computations for validating the estimation and filtering strategies were performed on MatLab. We solve Eq. (16) using sequential quadratic programming (SQP) as implemented in the medium-scale algorithm of *fmincon* function. For the successive integration of Eq. (15) we use the algorithm *ode45* that is based on an explicit Runge-Kutta formula (SHAMPINE & REICHELT, 1997).

Because the measurement equation of our example is linear, the optimization problem of Eq. 20 was solved using quadratic programming (QP) as implemented in the algorithm of *quadprog* function. The CEKF algorithm implemented in MatLab is shown in the Appendix.

5. Case Study

The computational case study is based on the ethanol production from glucose fermentation by *Zymomonas mobilis* bacteria. This system shows an interesting dynamic (continuous oscillations, equilibrium multiplicities, unstable regions). *Zymomonas mobilis* has attracted considerable interest over the past decades as a result of its unique metabolism and ability to rapidly and efficiently produce ethanol from simple sugars. However, despite its apparent advantages of higher yields and faster specific rates when compared to yeasts, little attention has been focused in *Zymomonas mobilis* for the manufacture of fuel ethanol. In addition to ethanol depending on the substrate other fermentation products can occur, such as lactic acid, acetic acid, formic acid, acetone, and sorbitol. See Rogers *et al*. (2007) for a detailed review.

Various models have been proposed to describe the oscillatory dynamics of continuous Zymomonas mobilis cultures (Daugulis *et al*., 1997; Jarsebski, 1992; Jöbses *et al*., 1986). Since the Jöbses's model can predict a branch with higher ethanol production, which has been experimentally confirmed (at least for low dilution rates) by Elnashaie *et al*. (2006), we have decided to use this model as our case study.

Jöbses's model is given by the following equations:

$$\frac{dC_S}{dt} = -\left(\frac{\mu_{\max} C_S C_e}{Y_{Sx}\left(K_S + C_S\right)}\right) - m_S C_x + D\left(C_{So} - C_S\right) \quad (23)$$

$$\frac{dC_x}{dt} = \left(\frac{\mu_{\max} C_S C_e}{K_S + C_S}\right) + D\left(C_{xo} - C_x\right) \quad (24)$$

$$\frac{dC_e}{dt} = K_E\left(C_P - c_1\right)\left(C_P - c_2\right)\left(\frac{C_S C_e}{K_S + C_S}\right) + D\left(C_{eo} - C_e\right) \quad (25)$$

$$\frac{dC_P}{dt} = \left(\frac{\mu_{\max} C_S C_e}{Y_{Px}\left(K_S + C_S\right)}\right) - m_P C_x + D\left(C_{Po} - C_P\right) \quad (26)$$

where C_S is the substrate (glucose) concentration, C_x is the biomass (*Zymomonas mobilis*), C_P is the product (ethanol) concentration, and C_e is an auxiliary variable used to lag the effect of the ethanol concentration in the kinetic model. The polynomial K_E (C_P-c_1) (C_P-c_2), experimentally adjusted in C_e description, makes possible the model to depict the oscillatory behaviors and have output multiplicity. The variables C_{S0}, C_{x0}, C_{e0} and C_{P0} complete the mass balance representing the states inputs in the reactor, normally only C_{S0} (substrate inlet) is different to zero. The rate dilution (D) meaning the same as the residence time inverse. The other variables are listed in Table 1 and more detailed description can be found in the Jöbses et al. (1986).

Tabela 1. Jöbses's model parameters.

Parameters	Values	Parameters	Values
$k_3\left[\frac{m^6}{kg^2.h}\right]$	0.00383	$m_S\left[\frac{kg}{kg.h}\right]$	2.160
$c_1\left[\frac{kg}{m^3}\right]$	59.2085	$m_P\left[\frac{kg}{kg.h}\right]$	1.100
$c_2\left[\frac{kg}{m^3}\right]$	70.5565	$Y_{SX}, Y_{PX}\left[\frac{kg}{kg}\right]$	(0.02445, 0.05263)
$K_S\left[\frac{kg}{m^3}\right]$	0.500	$\mu_{\max}\left[\frac{1}{h}\right]$	1.0

The Figure 3 shows the dynamic simulation (with C_{S0} = 200 kg/m³ and D = 2.0 h^{-1}) of the product to different initial conditions: IC1 (C_S = 10 kg/m³, C_e = 3 kg/m³, C_x = 0.1 kg/m³ e C_P = 20 kg/m³) e IC2 (C_S = 10 kg/m³, C_e = 3 kg/m³, C_x = 0.1 kg/m³ e C_P = 100 kg/m³). A perturbation in the initial condition of C_P leads the system to converge to a different steady state, indicating the occurrence of bistability. Such behavior characterizes the existence of multiple solutions.

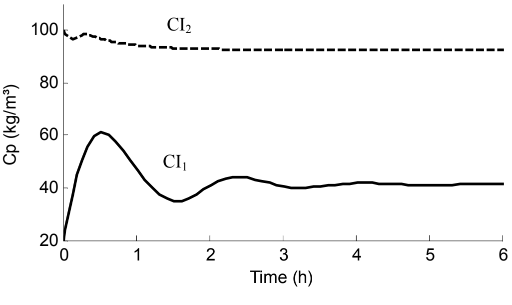

Figure 3. Jöbses's model dynamic simulation (Jöbses et al., 1986).

6. Results and Analyses

6.1. Virtual Analyser & Feedback Control

A multivariable control strategy for *Zymomonas mobilis* bioreactor is proposed in Diehl & Trierweiler, (2009), where C_{S0} and D are manipulated variables and C_P and C_S are controlled variables. This structure provides flexibility to the system permitting to maximize the substrate to product conversion. The frequency domain PI (proportional-integral) controller is shown in (Eq. 27), where the considerate pairing is $(D,\ C_{S0}) \rightarrow (C_P,\ C_S)$ and the control sampling time of 1 minute.

$$PI = \begin{bmatrix} -0.016\times\left(1+\frac{1}{0.565s}\right) & 0.7077\times\left(1+\frac{1}{0.015s}\right) \\ 6.133\times\left(1+\frac{1}{0.8076s}\right) & 3.2102\times\left(1+\frac{1}{0.4393s}\right) \end{bmatrix} \quad (27)$$

The design of a state estimator assumes system observability, i.e., a set of measurements must provide enough information to estimate all the system states. Thus, a system observability analysis is necessary before applying the state estimation techniques.

According to the Popov-Bevelic-Hautus (PBH) criterion**,** the system (F, H) (cf. Eqs.3 and 4) is state observable if and only if

$$M(i) = \begin{bmatrix} \lambda(i) & I-F \\ & H \end{bmatrix}; \quad rank\left[M(i)\right] = n \quad (28)$$

has rank n (full column rank) for all eigenvalues $\lambda_i\left(i=1,2,\ldots,n\right)$ of matrix F (Sontag, 1998).

As result, the *Zymomonas mobilis* bioreactor system is observable if C_S is measured.

The CEKF works as a virtual analyzer in this section, filtering C_S (assumed to be a measured variable with random noise) and estimating the other states, enclosing C_P. Afterwards, the information on C_S and C_P will be used to feedback the PI controller.

We consider state estimation with the following initial guesses and parameters:

$$x_0 = \left[1.24 \quad 4.74 \quad 13.31 \quad 92.56\right] \quad (29)$$

$$P_0 = 0.05^2 I_{4x4} \quad (30)$$

$$\Delta t = t_k - t_{k-1} = 0.0167 \quad (31)$$

$$R = 1 \quad (32)$$

$$Q = 0.01^2 I_{4x4} \quad (33)$$

$$x_{min} = \left[0.01 \quad 0.01 \quad 0.01 \quad 80\right] \quad (34)$$

$$x_{max} = \left[10 \quad 20 \quad 50 \quad 110\right] \quad (35)$$

Typically in bioprocesses the concentrations are measured by techniques with inherent time delay. For instance, it is common to use high performance liquid chromatography (HPLC) (DAVIS *et al.*, 2006), infrared (MAZAREVICA *et al.*, 2004) or Raman (SHAW & KELL, 1999) spectroscopy. However some new methods, as two-dimensional fluorescence spectroscopy (HALTELMANN *et al.*, 2006) can reduce this pure time delay considerably. In our example we consider hence online measurements available each 20 minutes and without pure time delay.

The control system with the CEKF feedback is shown in Figure 4. As the control and state estimation have different sample times (1 and 20 minutes, respectively), the CEKF simulation model is called every 1 minute and the CEKF update step (with the new measurement) is called only each 20 minutes.

According to the results of Figure 4, the state estimation with CEKF guarantees a suitable performance of the PI controller.

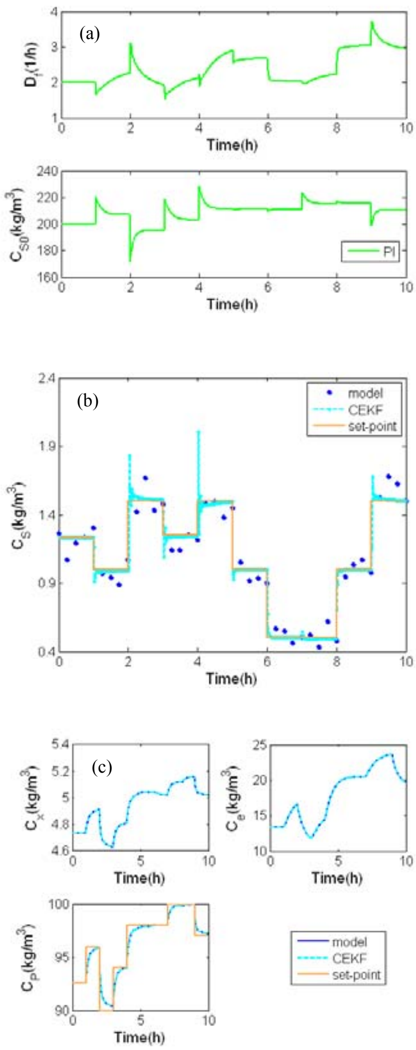

Figure 4. CEKF performance for feedback control: (a) manipulated variables, (b) filtered measured states and (c) estimated states.

6.3. Parameters fitting and update

Here, we have chosen the c_1 to be estimated with the states and, thereby, the state vector x is augmented with this parameter to obtain an augmented state vector x', i.e.

$$x' = \begin{bmatrix} x \\ c_1 \end{bmatrix} \quad (36)$$

As the parameter is constant then the augmented system model of Eqs. (1) and (2) can be rewritten as

$$\dot{x}' = \begin{bmatrix} f\left(x,u,t,c_1\right) \\ 0 \end{bmatrix} + \omega(t) \quad (37)$$

$$y_k = \left[h\left(x_k,t_k\right) \quad 0\right] \begin{bmatrix} x_k \\ c_l \end{bmatrix} + \nu_k \quad (38)$$

The common praxis of Eq. 36 to augment the state vector definition to make possible a simultaneous state and parameter estimation is in general not recommended, since the additional differential equation produces a local-state matrix F (cf. Eq. 3) with corresponding row elements equal to zero. It makes more difficult to find a set of measurements for system observability, being usually necessary to include additional measurements.

As result of the new observability analisys (Eq. 28), not only C_s but also C_p must be measured to make the augmented system (Eqs. 37 and 38) observable.

For our analyses, the dilution rate is increased from $2h^{-1}$ to $2.5h^{-1}$ at $5h$ to lead the system to converge to a different steady-state.

Now we consider state estimation with the following initial guesses and parameters:

$$x_0 = [8.78 \quad 4.55 \quad 9.63 \quad 89.05 \quad 56.25] \tag{39}$$

$$P_0 = 0.05^2 I_{5x5} \tag{40}$$

$$\Delta t = t_k - t_{k-1} = 0.25 \tag{41}$$

$$R = 0.1^2 I_{2x2} \tag{42}$$

$$Q = 0.5^2 I_{5x5} \tag{43}$$

$$x_{\min} = [0.15 \quad 1.2 \quad 1.8 \quad 30 \quad 53.28] \tag{44}$$

$$x_{\max} = [150 \quad 5 \quad 41 \quad 121 \quad 65.13] \tag{45}$$

In this section, we have enlarged the MHE estimation window to N=2 and compared its performance with the CEKF and EKF in Figure 5.

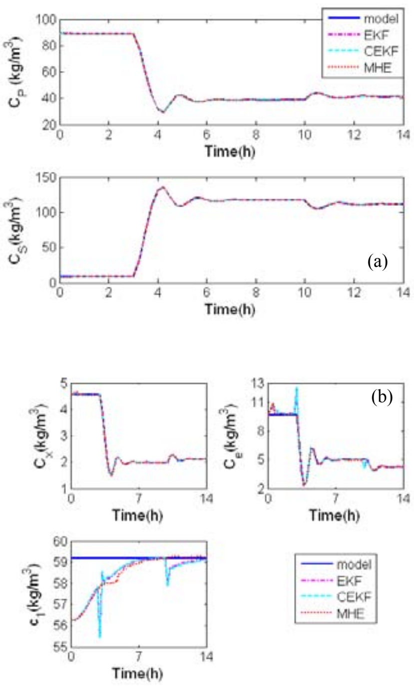

Figure 5. Comparison between EKF, CEKF and MHE (N=2) performances for parameter estimation: (a) filtered measured states and (b) estimated states.

As in the EKF constraints are not taken into account, its performance for parameter estimation is not as good as the optimization based approaches performance, as shown in Figure 5b.

6.4. Transient Data Reconciliation

We can also augment the state vector x with new artificial parameters corresponding to the error in the variables to be reconciliated. Afterwards, one can apply the same procedure as for the parameter estimation of the previous secction, taking appropriate noise variances into account.

Supposing a leak in the process, it was considered an error of $2h^{-1}$ in the manipulated dilution rate (ΔD) and no error in the manipulated inlet substrate concentration (ΔC_{S0}).

The augmented state vector is now given by:

$$x' = \begin{bmatrix} x \\ \Delta D \\ \Delta C_{S0} \end{bmatrix} \tag{46}$$

The augmented system model originated by the augmented state vector above is also observable with a set of measurements composed by C_s and C_p.

Again the dilution rate is increased from $2h^{-1}$ to $2.5h^{1}$ at $5h$ to lead the system to converge to a different steady-state.

Here we consider state estimation with the following poor state initial guesses $\bar{x}_0$ and parameters:

$$\bar{x}_0 = [111.34 \quad 2.11 \quad 4.24 \quad 41.29 \quad 0 \quad 0] \tag{47}$$

$$P_0 = 0.75^2 I_{6x6} \tag{48}$$

$$\Delta t = t_k - t_{k-1} = 0.25 \tag{49}$$

$$R = 0.1^2 I_{2x2} \tag{50}$$

$$Q = diag\left(0.05^2 \quad 0.05^2 \quad 0.05^2 \quad 0.05^2 \quad 0.1^2 \quad 0.1^2\right) \tag{51}$$

$$x_{\min} = [0.15 \quad 1.2 \quad 1.8 \quad 30 \quad -5 \quad 0] \tag{52}$$

$$x_{\max} = [150 \quad 5 \quad 41 \quad 121 \quad 0 \quad 0] \tag{53}$$

As the location of the leak is unknown to the estimator, we design the estimator with a process-noise covariance for ΔD and ΔC_{S0} equal to the measurement-noise covariance. We have also enlarged the MHE estimation window to N=2 and compared its performance with the CEKF and EKF in Figure 6.

According to the results of Figure 6b, MHE and CEKF provide a good estimate of the total losses for the leak, can identify the error in ΔD, and that there is no error in ΔC_{S0}. Otherwise, EKF cannot provide such good estimates.

MHE presents the best performance and swifter converge to actual states because over the horizon length no information about the nonlinear system is lost. Further, MHE provides improved state estimation and greater robustness to poor guesses of the initial state.

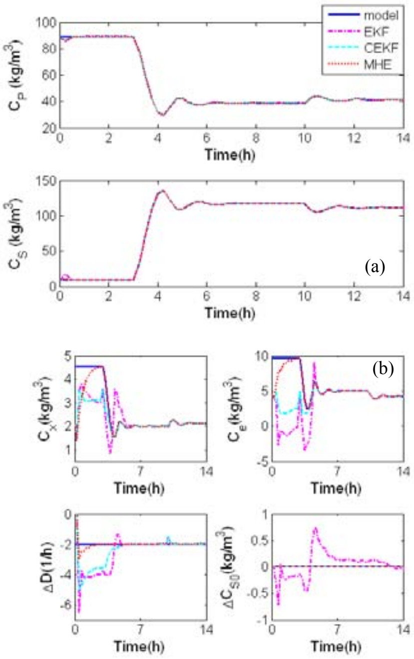

Figure 6. Comparison between EKF, CEKF and MHE (N=2) performances for transient data reconciliation: (a) filtered measured states and (b) estimated states.

7. Conclusions

In this paper we have shown the advantages of using state estimators as a tool for better operation of a *Zymomonas mobilis* bioreactor. Based on this illustrative example and on many other case studies presented elsewhere, the following conclusions can be formulate:

- State estimators can be successfully applied for process monitoring and feedback control because they provide information on unmeasured state which is essencial to ensure proper operation of any plant.
- This tool can also be used to fit and update parameters and to reconciliate the transient data. However, the strategy of augmenting the state vector with new variables may become more difficult to find a set of measurements to garantee the system observability, being usually necessary to include additional measurements.
- In general the EKF formulation has worse performance when compared with CEKF and MHE strategies;
- MHE can improve state estimation and provide better robustness to poor guesses of the initial state. However, afterwards converge towards the actual states, both MHE and CEKF perform equally accurately and, therefore, the use of MHE becomes needless.
- The better relationship between performance and practical application is obtained with the CEKF formulation, because it requires small computational effort than MHE with comparable performance. Therefore, we recommend using CEKF as first choice. The Appendix illustrates how a CEKF can be easily implemented.

8. Appendix: CEKF algorithm implemented in a MatLab –like language.

The CEKF algorithm for state estimation can be carried out by the steps below:

1. Simulate the model from $t_{k\text{-}1}$ to t_k (Δt) to obtain $\hat{x}_{k|\,k-1}$ (Eq. 1).

```
[tk,xkk_1]=ode45(Model,[0 delta_t], xk_1k_1);
```

2. Compute the linear models F (Eq.3) and H (Eq.4) at t_k.

```
n= length(xkk_1);
ny=length(y);
for i=1:n,
   for j=1:n,
      if i~=j,
         F(i,j)= diff(dxdt(i), 'x(j)')
            else
               F(i,i)= diff(dxdt(i), 'x(i)')
end

for i=1:ny,
   for j=1:n,
      H(i,j)= diff(y(i), 'x(j)');
end
```

3. Calculate the state transition matrix $\varphi_{k\text{-}1}$ for the state at t_k (Eq. 7).

```
phi=expm(F*delta_t);
```

4. Compute the state covariance matrix $P_{k|k}$ (Eq. 13).

```
Pkk= phi* Pk_1k_1*(phi')-phi*Pk_1k_1*(H')*...
      inv(H*Pk_1k_1*(H')+R)*H*Pk_1k_1*(phi')...
      +Q;
```

6. In case of linear measurement equation, solve the optimization problem of Eq. 20 using a quadratic programming to obtain $\omega_{k-1|k}$ and $v_{k|k}$.

$$\min_{\Theta_{k|k}} \Psi_k = \hat{\Theta}^T_{k|k} S^{-1}_{k|k} \hat{\Theta}_{k|k} + d^T \hat{\Theta}_{k|k}$$

where

$$\hat{\Theta}_{k|k} = \begin{bmatrix} \omega_{k-1|k} \\ v_{k|k} \end{bmatrix} \quad S = \begin{bmatrix} P_{k-1|k-1} & 0 \\ 0 & R_k \end{bmatrix} \quad d = 0$$

subject to the equality constraints (cf. Eq. 21)

$$\begin{bmatrix} H & I_{ny} \end{bmatrix} \hat{\Theta}_{k|k} = y_k - h\left(\hat{x}_{k|k-1}\right)$$

and to the inequality contraints of Eq. 22.

```
T=[inv(Pk_1k_1) zeros(n,ny);zeros(ny,n) inv(R)];
T=T+T'/2;
wmin=xmin-xkk_1;
wmax=xmax- xkk_1;
vmin=H*xmin-H*xkk_1;
vmax= H*xmax-H*xkk_1;
opt=optimset('LargeScale', 'off');
w_v=quadprog(T,zeros(n+ny,1),[],[],[H    eye(ny)],...
      [y(:)-H*xkk_1],[wmin vmin],...
      [wmax vmax], w_v0,opt);
```

7. Finally, compute the update state $\hat{x}_{k|k}$ and measurement y_k (Eq. 19).

```
w=w_v (1:n);
v=w_v (n+1:n+ny);
xkk= xkk_1+w(:);
yk= H*xkk_1+v(:);
```

8. Repeat this procedure until the final time.

9. Referências

NAJAFPOUR G.D. *Biochemical Engineering and Biotechnology*. 1ª ed., Elsevier B. V., Amsterdam, 2007.

STANBURY, P. F.; WHITAKER, A.; HALL, S. *Principles of fermentation technology*. 2ª ed., Elsevier Science Ltd., Burlington, 1995.

DORAN P.M. *Bioprocess Engineering Principles*. Elsevier Science & Technology Books, 1995.

ALFORD J.S. Bioprocess control: Advances and challenges. *Computers and Chemical Engineering*, v. 30, p. 1464-1475, 2006.

AYNSLEY, M.; HOLFLAND, A.; MORRIS, A.J., MONTAGUE G.A.; DI MASSIMO C. *Artificial Intelligence and the Supervision of Bioprocesses (Real-Time Knowledge-Based Systems and Neural Networks). Bioprocess Design and Control*. Springer-Verlag Berlin & Heidelberg, Berlin, 1993.

RANI, K.Y.; RAO, R.V.S. Control of fermenters – a review. *Bioprocess Engineering*, v. 21, p. 77-88, 1999.

BAKHTADZE, N.N. Virtual Analyzers: Identification Approach. *Automation and Remote Control*, v. 65(11), p. 1691-1709, 2004.

GADKAR, K.G.; MEHRA, S.; GOMES, J. On-line adaptation of neural networks for bioprocess control. *Computers and Chemical Engineering*, v. 29, p. 1047-1057, 2005.

GONZALEZ, G. D. Soft Sensors for Processing Plants. In: 2nd INTERNATIONAL CONFERENCE ON INTELLIGENT PROCESSING AND MANUFACTURING OF MATERIALS, 1999, Hawaii, Proceedings.

GESTHUISEN, R.; KLATT, K. -U.; ENGELL, S. Optimization-Based State Estimation - A Comparative Study for the Batch Polycondensation of PET. In: ECC, 2001, Porto, Proceedings.

ENGELL S. Feedback Control for Optimal Process Operation. Journal of Process Control, v. 17, 2007, p. 203-219, 2006.

SIMON, D. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*, Hoboken, Wiley-Interscience, 2006.

SOROUSH, M. State and Parameter Estimations and Their Applications in Process Control. *Computers and Chemical Engineering*, v. 23(2), p. 229-245, 1998.

NARASIMHAN, S.; JORDACHE, C. *Data Reconciliation & Gross Error Detections: An Intelligent Use of Process Data*, Gulf Professional Publishing, Houston, 2000.

RAO, C.V.; RAWLING, J.B. Constrained Process Monitoring: Moving-Horizon Approach. *AIChE Journal*, v.48(1), p. 97-109, 2002.

MUSKE, K.R.; RAWLINGS, J.B. *Nonlinear Moving Horizon State Estimation. NATO ASI Series*, v. 293, The Netherlands, Kluwer Academic, 1994.

ROBERTSON, D. G., LEE, J. H., AND RAWLINGS, J. B. A moving horizon-based approach for least-squares estimation. AIChE Journal, v. 42(8), p. 2209-2224, 1996.

RAO, C.V.; RAWLING, J.B.; MAYNE, D.Q. Constrained State Estimation for Nonlinear Discrete-Time Systems: Stability and Moving Horizon Approximations. *IEEE Trans. Autom. Control*, v. 48(2), p. 246-258, 2003.

DHOOGE, A.; GOVAERTS, W.; KUZNETSOV, Y. A.; MESTROM W.; RIET, A. M.; SOUTOIS, B. *MATCONT and CL_MATCONT: Continuation toolboxes in MATLAB*, 2006.

SHAMPINE, L.; REICHELT, M. The MATLAB ODE Suite. *SIAM Journal on Scientific Computing*, v. 18, p. 1–22, 1997.

JÖBSES, I.M.L.; EGBERTS, G.T.C; LUYBEN, K.C.A.M.; ROELS, J.A. 1986. Fermentation Kinetics of *Zymomonas Mobilis* an High Ethanol Concentrations: Oscillations in Continuous Cultures. *Biotechnology and Bioengineering*, v. 28, p. 868-877, 1986.

ROGERS, P.; JEON, Y.; LEE, K.; LAWFORD, H. *Zymomonas Mobilis for Fuel Ethanol and Higher Value Products*. *Biofuels*. Springer Berlin & Heidelberg, 2007.

DAUGULIS, A.J.; MCLELLAN, P.J.; LI, J. Experimental Investigation and Modeling of Oscillatory Behaviour in the Continuous Culture of *Zymomonas mobilis*. *Biotechnology and Bioengineering*, v. 1, p. 99-105, 1997.

JARZEBSKI, A.B. Modelling of Oscillatory Behaviour in Continuous Ethanol Fermentation. *Biotechnology Letters*, v. 2, p. 137-142. (1992).

ELNASHAINE, S. S. E. H.; CHEN, Z.; GARHYAN, P.; PRASAD, P.; MAHECA-BOTERO, A. Practical Implications of Bifurcation Chaos in Chemical and Biological Reaction Engineering. *International Journal of Chemical Reactor Engineering*, v. 4, 2006.

SONTAG, E.D. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd Edition, Springer, New York, 1998.

TRIERWEILER J.O.; DIEHL D.C. Analysis, control, and operational optimization of a Zymomonas mobilis reactor with equilibrium multiplicity. In: ADCHEM, 2009, Istanbul, Proceedings.

DAVIS, L.; ROGERS, P.; PEARCE, J.; PEIRIS, P. Evaluation of Zymomonas-Based Ethanol Production from a Hydrolyzed Waste Starch Stream. *Biomass and Bioenergy*, v. 30, p. 809-814, 2006.

MAZAREVICA, G.; DIEWOK, J.; BAENA, J.R.; ROSENBERG, E.; LENDL, B. On-Line Fermentation Monitoring by Mid-Infrared Spectroscopy. *Society for Applied Spectroscopy*, v. 58, p. 804-810, 2004.

SHAW, A.D.; KELL, J.J. Noninvasive On-Line Monitoring of the Biotransformation by Yeast of Glucose to Ethanol Using Dispersive Raman Spectroscopy and Chemometrics. *Society for Applied Spectroscopy*, v. 53, p. 1419-1428, 1999.

HALTELMANN, K.; KOLLECKER, M.; HÜLL, D.; HITZMANN, B.; SCHEPER T. Two-Dimensional Fluorescence Spectroscopy: A Novel Approach for Controlling Fed-Batch Cultivations. *Journal of Biotechnology*, v.121, p. 410-417, 2006.