10th International IFAC Symposium on
Computer Applications in Biotechnology
Proceedings Vol.1, June 4-6, 2007, Cancún, Mexico

BIOMASS GROWTH AND k_La ESTIMATION USING ONLINE AND OFFLINE MEASUREMENTS

Z. I. T. A. Soons[a,b], J. Shi[a], L. A. van der Pol[b], G. van Straten[a] and A. J. B. van Boxtel[a]

[a]Wageningen University, Systems and Control Group
P.O. box 17, 6700 AA Wageningen, The Netherlands, Zita.Soons@wur.nl
[b]Netherlands Vaccine Institute
Antonie van Leeuwenhoeklaan 9-11, P.O. box 457, 3720 AL Bilthoven, The Netherlands

Abstract: Measurement of the key process variables is essential during biopharmaceutical production. These measurements are often not available online. This work combines frequent online measurements with infrequent offline measurements to estimate the specific growth rate, biomass, and the oxygen mass transfer coefficient during continuous and fed-batch cultivations of *Bordetella pertussis* online using an Extended Kalman filter, parameter adaptation, and learning.

Keywords: Extended Kalman filter, online measurements, delayed offline measurements, oxygen transfer, specific growth rate, biomass, bioreactor, vaccine production, *Bordetella pertussis*.

1. INTRODUCTION

Most biopharmaceuticals are produced in a batch or fed-batch cultivation. The quality of the product is formed in this step and is the result of the metabolic state of the micro-organisms. It is therefore essential to measure the physiological state of the process. Metabolic activity is difficult to measure directly due to the lack of sensors, but respiration can be monitored by the oxygen mass balance. The oxygen uptake rate can in turn be used to estimate the specific growth rate and biomass. The specific growth rate and biomass concentration are key parameters that define the metabolic state of micro-organisms.

In this application only dissolved oxygen measurements were available. Therefore, a software sensor based on an Extended Kalman filter (EKF) was developed to estimate specific growth rate and biomass every minute using the oxygen uptake rate (*OUR*) as input. The choice for an EKF is in line with the process, in which the instruments and the measurement noise are known.

In bioreactors, aerated by a high air flow entering the headspace, the difference between the inlet and exhaust oxygen fraction is small and can therefore not be measured accurately. Hence *OUR* must be calculated using the oxygen balance in the liquid phase (Neeleman, 2002, Soons *et al*., 2006, 2006a):

$$OUR \approx k_L a \cdot (C_O^{in} - C_O^L) \qquad (1)$$

Where k_La is the oxygen transfer coefficient. The dissolved oxygen concentration (C_O^L) is assumed to be at pseudo-steady state i.e. accumulation of oxygen in the bioreactor is negligible. The oxygen concentration entering the bioreactor (C_O^{in}) is

assumed to be equal to the liquid phase oxygen concentration in equilibrium with the gas phase in the headspace.

In most applications, the oxygen transfer rate k_La is measured in advance of the cultivation in medium and is assumed to depend on stirrer speed and volume only. However, the formation of cells, proteins, and other molecules, which absorb at gas-liquid interfaces, cause interfacial blanketing and reduce the oxygen transfer rate (Doran, 1995). Because concentrations of cells, substrates, and products change during (fed-) batch cultivation, the value of k_La also changes. An example of change in k_La due to these factors is given in Sabra *et al* (2002). Changing k_La causes errors in the *OUR* calculation and the estimation of the specific growth rate and biomass. It is therefore essential to deal with time-varying k_La.

Offline measurements are mostly considered as not suited for control and estimation purposes, because they become available with a delay and at infrequent and irregular times. These measurements however contain valuable information about the states of the system and can make the estimator more robust (Dondo and Marqués, 2003). Amongst the literature on bioprocess monitoring (e.g. Bastin and Dochain, 1990) the use of offline information for online estimation is relatively small. Myers *et al.* (1995) and Tatiraju *et al.* (1997) use offline and online measurements for state estimation, but do not estimate parameters; Lubenova *et al*. (2003); Gudi *et al.* (1995); Dondo and Marqués (2003); and Ignatova *et al.* (2003) estimate parameters in addition. In these approaches the parameters are part of the input-output equations. In this work, however, better results are obtained if the parameter k_La is considered as a part of the *OUR* calculations, which is done separately from the input-output equations (see figure 1).

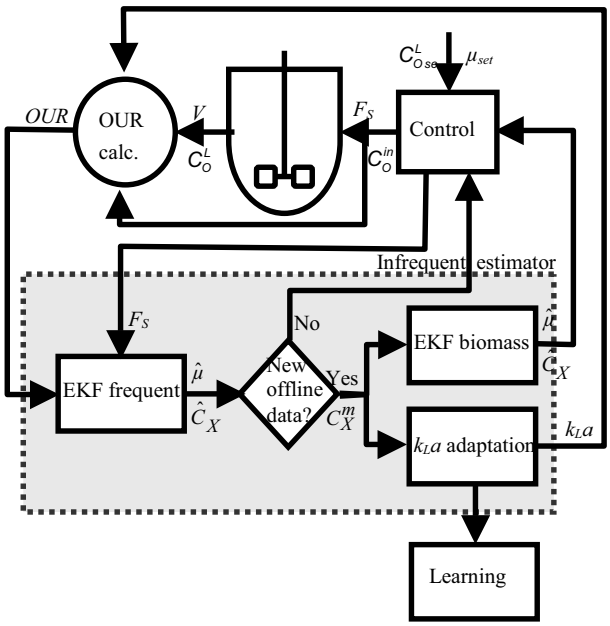

Fig. 1. System configuration (symbolic: see text or nomenclature).

This work combines frequent online measurements (oxygen uptake rate) with infrequent offline measurements (biomass) to estimate the specific growth rate and biomass accurately. Figure 1 shows an overview of the system, in which two types of estimators an learning are involved: a *frequent* estimator using the online data and an *infrequent* estimator activated by sampled offline data. The offline measurements are also used to adapt the time-varying k_La. After each run improved values for k_La are obtained and can be used in a learning process to acquire the appropriate k_La time pattern.

2. EXTENDED KALMAN FILTER

The estimator is based on a nonlinear continuous-time model:

$$\begin{aligned} \frac{dx}{dt} &= f(x,u) \\ y &= h(x) \end{aligned} \qquad (2)$$

With f a nonlinear function of the states x and inputs u, and y the output.

Lewis (1986) gives a good explanation of an Extended Kalman filter. The application in biotechnological applications is amongst others discussed by Gudi *et al.* (1995), Neeleman (2002) and Keesman (2002). The structure of a discrete time EKF is shown in Fig. 2, and is based on the following equations following Lewis (1986):

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + w_k \\ y_k &= C_k x_k + v_k \end{aligned} \qquad (3)$$

Where A_k and C_k and B_k at each time instant follow from discretization and linearization of Eq. 2 for a time step τ:

$$\begin{aligned} F &= \left[\frac{\partial f}{\partial x}\right]_{\hat{x},u}, \quad C_k = \left[\frac{\partial h}{\partial x}\right]_{\hat{x},u} \\ A_k &= I + F\tau \quad , B_k = \left[\frac{\partial f}{\partial u}\right]_{\hat{x},u} \tau \end{aligned} \qquad (4)$$

and u_k is the input vector. The initial states x_0 are stochastic variables with average $\bar{x}_0$ and variance P_0: $x_0 \sim (\bar{x}_0, P_0)$; $w_k \sim (0, Q_k)$ is system noise and consists of model errors an unknown inputs; and $v_k \sim (0, R_k)$ is measurement noise. The algorithm has two steps. The time update and the measurement update.

2.1 Time update

When a sample comes available at time k, first the time update $k+1$ is calculated using the original nonlinear model.