Observability analysis and software sensor design for an animal cell culture in perfusion mode

Ines Saraiva*. Alain Vande Wouwer.*Anne-Lise Hantson***. Jaime Moreno**

*Service d’Automatique, Université de Mons, 31 Boulevard Dolez, 7000 Mons, Belgium (e-mail: Ines.Saraiva; Alain.VandeWouwer@umons.ac.be).

**Instituto de Ingenieria, Universidad Nacional Autonoma de Mexico, Cd. Universitaria, CP 04510, Mexico, D.F., Mexico (e-mail: JMorenoP@it.unam.mx)

***Service de Chimie et Biochimie appliquées, Université de Mons, 56 Rue de l’Epargne, 7000 Mons, Belgium (e-mail: Anne-Lise.Hantson@umons.ac.be).

Abstract: The cultivation of animal cells in perfusion allows the production of various biopharmaceutical products. In this work, the observability properties of a nonlinear dynamic model of these animal cell cultures is assessed using a method based on a natural dynamical interpretation of the observability/detectability concepts, leading to the description of the indistinguishable dynamics of the system. Following this analysis, a Kalman filter is designed to reconstruct on-line variables which are difficult or expensive to measure directly with a hardware sensor.

Keywords: observability, nonlinear systems, state estimation, Kalman filter, bioprocesses

1. INTRODUCTION

Some animal cell strains, such as CHO cells, can be grown in suspension in stirred tank reactors, which appear to be the most common practice in industry for large production of high value protein products. The efforts for increasing the culture productivity in these systems focus on adjusting the media composition on the one hand, and the modes of operation on the other hand. The most popular operating modes in animal cell cultivation are batch, fed-batch and perfusion modes (Fig. 1). Batch and fed-batch modes do not offer many alternatives for control, as in these cases the feed rate is either absent or limited and the growth is inhibited by the accumulation of toxic metabolites, which cannot be removed. In perfusion mode, fresh medium is fed to replenish the consumed nutrients, while an equal volume of spent medium is continuously withdrawn from it, allowing for the removal of toxic components. Cells are retained or recycled back to the reactor by some type of retention device, such as sonoperfusion filters. Higher cell concentrations and higher productivity are achieved in perfusion cultures than in conventional batch cultures.

The successful operation of cultures in perfusion mode requires tight control, and recent developments consider the use of multivariable control to manipulate the feed and bleed rates, as well as, the composition of the feed flow (Deschênes et al. 2006a,b, Sbarciog et al., 2012). However, the development of control strategies usually requires the availability of a number of on-line measurements, which can be difficult to achieve in practice (availability of the probe, costs, processing time, etc). The development of software sensors is therefore of paramount importance for the implementation of these control strategies.

In this paper, we consider a nonlinear dynamic model of Hybridoma cell cultures producing monoclonal antibodies, initially proposed in (De Tremblay al et al, 1992, 1993), and since then considered in several further works: (Pörtner and Schäfer, 1996), (Roubos et al, 1997, 1999), (Nguang et al, 2001), (Chen et al, 2002), (Sarkar and Modak, 2004), (Franco-Lara and Weuster-Botz, 2005). We first analyse the observability/detectability properties of this model using a method based on a natural dynamical interpretation of the observability/detectability concepts (Moreno et al, 2012), leading to the description of the indistinguishable dynamics of the system. Following this analysis, a Kalman filter is designed to reconstruct on-line variables which are difficult or expensive to measure directly with a hardware sensor. Various sensor configurations are considered, showing promising results in simulation.

The paper is organized as follows. The next section introduces the observability/detectability analysis method, and illustrates it with an oversimplified fictitious model of an animal cell culture. Section 3 deals with a real Hybridoma

---

Fig. 1. Bioreactor layout according to operation regime.
cell cultivation model, and the observability analysis of this more complex system. A Kalman filter is then designed and tested in simulation in section 4. Section 5 is devoted to some conclusions.

2. OBSERVABILITY/DETECTABILITY ANALYSIS

As it is well-known, the possibility of constructing an observer is tied to the observability/detectability properties of the system’s model. When only the initial conditions are unknown, observability corresponds to the (theoretical) possibility of estimating the state in a finite time-horizon, whereas if the system is only detectable the state estimation can only be attained asymptotically.

The observability/detectability analysis proposed in (Moreno et al., 2012) is based on a natural dynamical interpretation of the observability/detectability concepts, and is introduced using a simple example for the sake of illustration. A continuously perfused bioreactor is considered where a culture of suspended cells takes place. Besides the feeding of substrates, a perfusion filter allows cell retention, whilst a small bleed stream lets all components out.

Only two biological macroreactions comprising substrates glucose (Glc) and glutamine (Gln), living (Xv) and dead (Xd) biomass are considered:

\[ (-v_{21}) \text{Glc} + (-v_{31}) \text{Gln} \xrightarrow{\text{in}} \text{Xv} \]
\[ (-v_{12}) \text{Xv} \xrightarrow{\text{out}} v_{12} \text{Xd} \]

This leads to a system of mass balance equations where the following symbols represent states, kinetic parameters and inputs:

\[ x_1 = \text{Xv} \; ; \; x_2 = \text{Glc} \; ; \; x_3 = \text{Lac} \; ; \; u_1 = D = F / V \; ; \; \]
\[ u_2 = D_{\text{perf}} = F_{\text{perf}} / V \; ; \; k_1 = v_{21} = 1 / Y_{\text{Xv/Glc}} = v_{31} = 1 / Y_{\text{Xv/Gln}} \; ; \; \]
\[ k_2 = \text{Gln}^\text{N} \; ; \; k_3 = \text{Gln} \; ; \; k_4 = \mu_{\text{max}} \; ; \; k_5 = \mu_{\text{Glc}} \; ; \; k_6 = \mu_{\text{Gln}} \; ; \; k_7 = \mu_{\text{dead}}. \]

\[
\begin{align*}
\frac{dx_1}{dt} &= \left( k_1 \frac{x_2}{k_3 + x_2} \frac{x_3}{k_4 + x_4} \right) x_1 - k_1 x_1 - u_1 x_1 + u_2 x_1 \quad (1) \\
\frac{dx_2}{dt} &= -k_1 \left( k_2 \frac{x_2}{k_3 + x_2} \frac{x_3}{k_4 + x_4} \right) x_2 - u_1 x_2 + k_2 u_1 \quad (2) \\
\frac{dx_3}{dt} &= -k_1 \left( k_2 \frac{x_2}{k_3 + x_2} \frac{x_3}{k_4 + x_4} \right) x_3 - u_1 x_3 + k_3 u_1 \quad (3)
\end{align*}
\]

A dynamical interpretation of the concepts of observability/detectability can be obtained considering a copy of the system with variables \( z_i, i = 1, 2, 3 \).

\[
\begin{align*}
\frac{dz_1}{dt} &= \left( k_4 \frac{z_2}{k_5 + z_2} \frac{z_3}{k_6 + z_3} \right) z_1 - k_1 z_1 - u_1 z_1 + u_2 z_1 \quad (4) \\
\frac{dz_2}{dt} &= -k_1 \left( k_2 \frac{z_2}{k_5 + z_2} \frac{z_3}{k_6 + z_3} \right) z_2 - u_1 z_2 + k_2 u_1 \quad (5) \\
\frac{dz_3}{dt} &= -k_1 \left( k_2 \frac{z_2}{k_5 + z_2} \frac{z_3}{k_6 + z_3} \right) z_3 - u_1 z_3 + k_3 u_1 \quad (6)
\end{align*}
\]

and defining deviations \( \epsilon_i \) between the system states \( x_i \) and their homologs \( z_i \):

\[
\epsilon_i = x_i - z_i \quad (7)
\]

Subtracting (4-6) from (1-3) leads to

\[
\begin{align*}
\frac{d\epsilon_1}{dt} &= \mu_1 x_1 - \mu'_1 \left( x_1 - \epsilon_1 \right) - k_1 \epsilon_1 - \epsilon_2 u_1 + \epsilon_2 u_2 \quad (8) \\
\frac{d\epsilon_2}{dt} &= -k_1 \mu_1 x_1 + k_1 \mu'_1 \left( x_1 - \epsilon_1 \right) - \epsilon_2 u_1 \quad (9) \\
\frac{d\epsilon_3}{dt} &= -k_1 \mu_1 x_1 + k_1 \mu'_1 \left( x_1 - \epsilon_1 \right) - \epsilon_2 u_1 \quad (10)
\end{align*}
\]

with

\[
\begin{align*}
\mu_1 &= k_4 \frac{x_2}{k_5 + x_2} \frac{x_3}{k_6 + x_3} \\
\mu'_1 &= k_7 \frac{(x_2 - \epsilon_2)}{(x_2 - \epsilon_2)} \frac{(x_3 - \epsilon_3)}{(x_3 - \epsilon_3)} \quad (12)
\end{align*}
\]

If biomass is measured along time, then it can be assumed that \( \epsilon_1 = 0 \) and \( \frac{d\epsilon_j}{dt} = 0 \). Thus equations (8)-(10) become:

\[
\begin{align*}
\frac{d\epsilon_1}{dt} &= \mu_1 x_1 - \mu'_1 \left( x_1 - \epsilon_1 \right) - k_1 \epsilon_1 - \epsilon_2 u_1 + \epsilon_2 u_2 \quad (13) \\
\frac{d\epsilon_2}{dt} &= -k_1 \mu_1 x_1 + k_1 \mu'_1 \left( x_1 - \epsilon_1 \right) - \epsilon_2 u_1 \quad (14) \\
\frac{d\epsilon_3}{dt} &= -k_1 \mu_1 x_1 + k_1 \mu'_1 \left( x_1 - \epsilon_1 \right) - \epsilon_2 u_1 \quad (15)
\end{align*}
\]

This method allows us to conclude that asymptotic convergence occurs for a dilution ratio \( u_1 > 0 \). The system is detectable (asymptotic convergence) because:

\[
\lim_{t \to +\infty} \epsilon_j(t) = 0 \quad (16) \\
\lim_{t \to +\infty} \epsilon_j(t) = 0 \quad (17)
\]

Thus, except for the batch mode, system states \( x_2 \) and \( x_3 \) can be distinguished if \( x_1 \) is measured.

3. PROCESS MODEL AND ANALYSIS

A real dynamic model of hybridoma cells producing monoclonal antibody is considered (De Tremblay et al., 1992). It comprises 7 states (biomass, glucose, lactate, glutamine, ammonium, monoclonal antibodies, volume), 16 parameters and describes typical animal cell culture phenomena.

\[
\begin{align*}
(-v_{21}) \text{Glc} + (-v_{31}) \text{Gln} \xrightarrow{\text{in}} \text{Xv} + v_{13} \text{Lac} + v_{31} \text{Amm} \\
(-1) \text{Xv} \xrightarrow{\text{out}} \text{Xd} \\
(-1) \text{Glc} + (-v_{13}) \text{Xv} \xrightarrow{\text{out}} v_{13} \text{Xv} \\
(-v_{12}) \text{Xv} \xrightarrow{\text{out}} v_{14} \text{Xv} + \text{MAb}
\end{align*}
\]

Mass balance equations are given by
with the following notation:

\[
\begin{align*}
    x_1 &= X_v; \quad x_2 = \text{Glc}; \quad x_3 = \text{Lac}; \\
    x_4 &= \text{Gln}; \quad x_5 = \text{Amm}; \quad x_6 = \text{MAb}; \quad k_1 = \alpha; \quad k_2 = \beta; \quad k_3 = \mu_\text{p}; \\
    k_4 &= k_\text{d.Amm}; \quad k_5 = k_\text{d.Glc}; \quad k_6 = k_\text{d.Lac}; \quad k_7 = k_\text{d.Gln}; \quad k_8 = k_\text{d}; \\
    k_9 &= k_\text{p.Gln}; \quad k_{10} = m_\text{Gl}
\end{align*}
\]

The reaction rates are given by:

\[
\begin{align*}
    \phi_i &= \mu_i x_i, \\
    \mu_i &= k_{12} \frac{x_2}{k_7 + x_2} - \frac{x_3}{k_5 + x_3}, \\
    \mu_2 &= k_{11} \frac{1}{k_{12} - k_6 x_3} - \frac{1}{k_{12} - k_{14} x_3} \frac{k_4}{k_5 + x_3}, \\
    \mu_3 &= k_{10} \frac{x_2}{k_7 + x_2}, \\
    \mu_4 &= k_2 + k_1 \frac{\mu_1}{k_5 + \mu_i}.
\end{align*}
\]

Since \( x_4 \) is has no influence on the remaining states, it is not considered any longer in the following.

A copy of the original system is built.

\[
\begin{align*}
    \frac{dz_1}{dt} &= k_{12} \frac{z_2}{k_7 + z_2} \frac{z_3}{k_5 + z_3} z_1, \\
    \frac{dz_2}{dt} &= -k_{14} \left( k_{12} \frac{z_2}{k_7 + z_2} \frac{z_3}{k_5 + z_3} + k_{10} \right) z_1 - z_1 u_1 + k_{10} u_1, \\
    \frac{dz_3}{dt} &= k_{16} \left( k_{12} \frac{z_2}{k_7 + z_2} \frac{z_3}{k_5 + z_3} + k_9 \right) z_1 + k_{10} \frac{z_2}{k_7 + z_2} z_1 - z_2 u_1 + k_{10} u_1, \\
    \frac{dz_4}{dt} &= -k_{11} \left( k_{12} \frac{z_2}{k_7 + z_2} \frac{z_3}{k_5 + z_3} + k_6 \right) z_1 - z_4 u_1 + k_{10} u_1.
\end{align*}
\]

Following the procedure introduced in the previous section, the error dynamics can be expressed as:

\[
\frac{dx_4}{dt} = \mu_1^* (x_1 - e_1) - \mu_2 x_1 + \mu_3^* (x_1 - e_1) - e_1 u_1 + e_2 u_2
\]

Two practical measurement configurations are considered:

- Living biomass and extracellular glucose concentrations are measured on-line;
- Living biomass only is measured (this can be achieved with an impedance probe);

### 3.1 Case A – Biomass and glucose measurements

If we consider that biomass and glucose are being measured then \( e_1 = 0, e_2 = 0 \) and it follows that \( \frac{dx_4}{dt} = 0, \frac{de_2}{dt} = 0 \). Equations (32)-(36) allow us to conclude that

\[
\frac{dx_4}{dt} = \frac{(x_4 - e_1)}{k_4 + (x_4 - e_2)} \Rightarrow e_4 = 0 \quad (x_4 \neq 0)
\]

\[
\frac{de_{21}}{dt} = -u_1 e_1 \rightarrow \text{since } u_1 > 0 \text{ then } \lim_{t \to \infty} e_1(t) = 0
\]

\[
\frac{de_{22}}{dt} = -u_1 e_1 \rightarrow \text{since } u_1 > 0 \text{ then } \lim_{t \to \infty} e_1(t) = 0
\]

We conclude that the measurements of \( x_1 \) and \( x_2 \) provide sufficient information to ensure detectability (except in batch mode).
3.2 Case B – Biomass Measurements

If we consider that only biomass is being measured then \( \epsilon_1 = 0, \ \frac{d\epsilon}{dt} = 0 \). Equations (32)-(36) lead to

\[
\begin{align*}
\bar{\mu} - \bar{\mu}_2 &= 0, \\
\frac{d\epsilon}{dt} &= (-\bar{\mu}_3 - k_i \bar{\mu}_2) x_1 - u_i \epsilon_2, \\
\frac{d\epsilon}{dt} &= (k_i \bar{\mu}_3 + k_i \bar{\mu}_2) x_1 - u_i \epsilon_3, \\
\frac{d\epsilon}{dt} &= (-k_i \bar{\mu}_3) x_1 - u_i \epsilon_4, \\
\frac{d\epsilon}{dt} &= (k_i \bar{\mu}_2) x_1 - u_i \epsilon_5,
\end{align*}
\]

where

\[
\bar{\mu}_i = \mu_i - \mu_i^e.
\]

Unfortunately, it is not easy to conclude in a straightforward manner about the observability/detectability conditions.

Computation of the observability map and linearization suggests that the system would be locally observable (Fig. 2) for illustrative normal operating conditions. The evolution of normalized values attained by the determinant is listed on the left column and plotted on the right (top) along with the states (middle and bottom).

![Graph of determinant and states](image)

**Fig. 2.** Evolution of the determinant of the observability matrix and concentrations during a normal culture.

4. KALMAN FILTER DESIGN

Following the previous analysis, a Kalman filter is designed considering the two measurement configurations. The implemented algorithm extends the use of the filter to nonlinear systems by use of a linearization along the state trajectory.

System:

\[
\begin{align*}
\dot{x}(t) &= f(\hat{x}(t), u(t)) + \eta(t); \\
y(t_k) &= C(t_k) x(t_k) + \epsilon(t_k) \\
\end{align*}
\]

Gaussian white noises:

\[
\begin{align*}
\eta(t) &\sim N(0, R_\eta(t)) \\
\epsilon(t) &\sim N(0, R_\epsilon(t))
\end{align*}
\]

A continuous-discrete version of the extended Kalman filter (EKF) is considered.

(i) Initialization

Initially we consider a given initial condition and covariance \( P \):

\[
\begin{align*}
m_{t_0} &= x_0 \\
P_{t_0} &= P_0
\end{align*}
\]

(ii) Continuous Prediction for \( t_{k-1} < t < t_k \):

\[
\begin{align*}
\hat{x}(t) &= f(\hat{x}(t), u(t)); \\
\dot{\hat{x}}(t_{k-1}) &= m_{t_{k-1}} \eta_{t_{k-1}}, \\
P(t) &= A P(t) + P(t) A^T + R_x(t); \\
P(t_{k-1}) &= P_{t_{k-1}}
\end{align*}
\]

where \( A = \text{Jac}(\hat{x}(t), u(t)) \frac{\partial f(\hat{x}(t), u(t))}{\partial x} \). (48)

(iii) Discrete-time correction at \( t=t_k \)

\[
\begin{align*}
m_{t_k} &= \hat{x}(t_k) \\
P_{t_k} &= P(t_k)
\end{align*}
\]

\[
K(t_k) = P_{t_k} C^T(t_k) \left( C(t_k) P_{t_k} C^T(t_k) + R_x(t_k) \right)^{-1}
\]

\[
\begin{align*}
m_{t_{k+1}} &= m_{t_k} + K(t_k) \left[ y(t_k) - C(t_k) m_{t_k} \right] \\
P_{t_{k+1}} &= P_{t_k} - K(t_k) C(t_k) P_{t_k}
\end{align*}
\]

This algorithm is appropriate for bioprocesses, which can have low sampling frequencies of the available probes, whereas the process models are nonlinear mass balance equations. The state estimator can be coupled to a model predictive controller, such as the one described in (Sbarciog et al., 2012).

The convergence of the filter is illustrated in the case of biomass and glucose measurements in Fig. 3.

As for the case where only biomass measurements are available, the filter appears to perform sufficiently well since estimates still converge to the real values even with an extremely poor knowledge of initial conditions.

In reality, an animal cell culture begins with initial conditions that are relatively well known: the medium is generally prepared beforehand and its concentrations of glucose and glutamine fairly well known. On the other hand, before cells are inoculated, the concentration of metabolites lactate and ammonium are close to zero and the concentration of the synthesized bioprotein is zero.
Another issue to be addressed is the performance of the filter given measurement noise. The figures presented assume a small level of noise, which is in fact coherent with the performances of probes such as the Fogale, illustrated in Fig. 5.

\[ R_\eta = \text{diag}([0 \ 0 \ 0 \ 0 \ 0 \ 0]) \]
\[ R_\eta = \text{diag}([0.1 \ 0.01 \ 30 \ 0.1 \ 0.01 \ 1]) \]
\[ P_0 = \text{diag}([10 \ 10 \ 10 \ 10 \ 10 \ 10]) \]

Estimation (magenta) – real process variables (blue).

Fig. 3. EKF based on biomass and glucose measurements.

Lowest IC=[12; 2; Xv0; 10; 2; 40]; lower IC=[20; 3; Xv0; 1; 1; 10]; perfect IC=[25; 4; Xv0; 0; 0; 0]; higher IC=[30; 5; Xv0; 1; 1; 10]; highest IC=[38; 6; Xv0; 10; 2; 40].

Fig. 4. EKF based on biomass measurements.
5. CONCLUSIONS

This application-oriented paper addresses two questions.

The first is the assessment of observability of nonlinear dynamic models of cell cultures, in the case of minimal measurement configurations. Given the complexity of these systems, most available methods such as those presented in (Dochain, 1992) fail to be helpful. Canonical forms such as those used in bacterial cultures in (Dewasme et al, 2012) are also of difficult application. The method proposed in (Moreno et al., 2012) has however shown to be useful, even though the solution of the differential algebraic system for the error dynamics can be delicate in some cases.

The second is the design of observers for cell cultures using a few measurement probes, and in particular the relatively recent biomass probes which provide almost time-continuous evolution of the biomass with low levels of noise. This has to be contrasted with bioanalyzers that automate enzymatic kit analyses. Besides not being very frequent, these analyses have high operation costs and can present some errors.

The authors gratefully acknowledge the support of the FEDER program 2007-2013, in the framework of Hainaut-Biomed, and the research project OCPAM. This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The authors also gratefully acknowledge the support of FNRS and CONACYT. The scientific responsibility rests with the author(s).

REFERENCES


