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

PARAMETER IDENTIFICATION TO ENFORCE PRACTICAL OBSERVABILITY OF NONLINEAR SYSTEMS

G. Goffaux [*] **L. Bodizs** [***] **A. Vande Wouwer** [*,1]
P. Bogaerts [**] **D. Bonvin** [***]

[*] *Service d'Automatique, Faculté Polytechnique de Mons, Boulevard Dolez 31, B-7000 Mons, Belgium*
[**] *Service de Chimie Générale et Biosystèmes, Université Libre de Bruxelles, B-1050 Brussels, Belgium*
[***] *Laboratoire d'Automatique, Ecole Polytechnique Fédérale de Lausanne, CH-1015, Lausanne, Switzerland*

Abstract: The sensitivity of measurements to unmeasured state variables strongly affects the rate of convergence of a state estimator. To overcome potential observability problems, the approach has been to identify the model parameters so as to reach a compromise between model accuracy and system observability. An objective function that weighs the relative importance of these two objectives has been proposed in the literature. However, this scheme relies on an extensive heuristic search to select the weighting coefficients. This paper proposes an objective function that is the product of measures of these two objectives, thus alleviating the need for the trial-and-error selection of the weighting coefficient. The proposed identification procedure is evaluated using both simulated and experimental data, and with different observer structures.

Keywords: Parameter identification, Observability, State estimation, Identification for estimation, Kalman filter, Particle filter

1. INTRODUCTION

Observability tests typically provide a binary yes/no answer and, thus, do not help assess whether practical observability problems such as slow convergence of the state estimates will occur. A study has shown that even an accurate bioprocess model can lead to poor state estimates when the measurements have a low sensitivity with respect to the unmeasured states (Bogaerts and Vande Wouwer, 2004). To alleviate this problem, the same authors have suggested a model "falsification" procedure, in which the model parameters are identified so as to achieve a compromise between model accuracy (via minimization of a criterion expressing the deviation between the model and plant states) and system observability (via a measure of observability based on sensitivity matrices). Unfortunately, the proposed objective function contains a weighting coefficient that is best determined via a trial-and-error procedure involving repeated optimization.

The contribution of this paper is to propose an objective function that (i) achieves the aforementioned compromise between model accuracy and system observability, and (ii) can be determined without trial-and-error procedure. It turns out that the objective function can be formulated as the product of two measures that are related to the sought objectives. This study also compares the classical extended Kalman filter (Maybeck, 1982) with a less classical (at least

[1] Corresponding author: alain.vandewouwer@fpms.ac.be
phone: +32 (0)65374141 fax: +32 (0)65374136

in bioprocess monitoring) particle filter (Doucet *et al.*, 2001) on two case studies, one in simulation and the other using experimental data.

The paper is organized as follows. Section 2 sets the notations for parameter identification and briefly reviews the concept of nonlinear system observability. Section 3 describes the parameter identification procedure for state estimation, while Section 4 details the results obtained with a simulated example and a real-life application. Finally, conclusions are provided in Section 5.

2. PRELIMINARIES

2.1 Parameter identification

We consider continuous-time nonlinear models associated with discrete-time measurements:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}) \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{1}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}(t_k), \boldsymbol{\theta}) \tag{2}$$

where $\mathbf{x}(t) \in \Re^{n_x}$ is the state vector, $\mathbf{u}(t) \in \Re^{n_u}$ the input vector and $\mathbf{y}_k \in \Re^{n_y}$ the output vector at the discrete time t_k. $\boldsymbol{\theta}$ is the vector of parameters to be identified. $\mathbf{f}$ and $\mathbf{h}$ are, in general, nonlinear vector functions. For simplicity of notation, the time dependency of the signals $\mathbf{x}(t)$ and $\mathbf{u}(t)$ will be dropped in the sequel.

The parameter identification problem can be formulated as follows:

$$(\hat{\boldsymbol{\theta}}, \hat{\mathbf{x}}_0) = \arg\min_{\boldsymbol{\theta}, \mathbf{x}_0} J_{id}(\boldsymbol{\theta}, \mathbf{x}_0) \tag{3}$$

$$\text{given} \quad \textit{model} \;\; (1)-(2); \;\; \mathbf{y}_{meas}$$

with

$$J_{id}(\boldsymbol{\theta}, \mathbf{x}_0) = \frac{1}{2\mathrm{N}} \sum_{k=1}^{N} \left(\mathbf{y}_{meas,k} - \mathbf{y}_k(\boldsymbol{\theta}, \mathbf{x}_0)\right)^T Q_k^{-1} \left(\mathbf{y}_{meas,k} - \mathbf{y}_k(\boldsymbol{\theta}, \mathbf{x}_0)\right) \tag{4}$$

where $\mathbf{y}_{meas,k}$ represents the measured outputs at time t_k, Q_k the covariance matrix of the measurement noise, and N the data length. Note that, since the initial conditions are rarely known in practice, they can be considered as decision variables as well. This is similar to the approach taken in moving-horizon estimation (Haseltine and Rawlings, 2005).

The properties of the resulting model can be analyzed. In the context of the design of a state observer, system observability is of paramount importance.

2.2 Observability of nonlinear systems

A system is said to be completely observable if it is possible to reconstruct the state vector from a finite number of output measurements. Global observability analysis of nonlinear systems is a delicate task since observability generally depends on the system inputs. The analysis is made simpler through the introduction of canonical forms (Zeitz, 1984; Zeitz, 1989). A system is said to be globally observable if the nonlinear model can be expressed in the following canonical form (Gauthier and Kupka, 1994):

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{\mathbf{x}}^1 \\ \dots \\ \dot{\mathbf{x}}^i \\ \dots \\ \dot{\mathbf{x}}^{q-1} \\ \dot{\mathbf{x}}^q \end{bmatrix} = \begin{bmatrix} \mathbf{f}^1(\mathbf{x}^1, \mathbf{x}^2, \mathbf{u}) \\ \dots \\ \mathbf{f}^i(\mathbf{x}^1, \dots, \mathbf{x}^{i+1}, \mathbf{u}) \\ \dots \\ \mathbf{f}^{q-1}(\mathbf{x}^1, \dots, \mathbf{x}^q, \mathbf{u}) \\ \mathbf{f}^q(\mathbf{x}^1, \dots, \mathbf{x}^q, \mathbf{u}) \end{bmatrix}, \tag{5}$$

$$\mathbf{y} = \begin{bmatrix} h_1(x_1^1) \\ h_2(x_1^1, x_2^1) \\ \dots \\ h_{n_1}(x_1^1, \dots, x_{n_1}^1) \end{bmatrix} \tag{6}$$

$$\text{with} \quad \forall i \in \{1, \dots, q\} : \mathbf{x}^i = \left[x_1^i, \dots, x_{n_i}^i\right]^T, \quad n_1 \geq n_2 \geq \dots \geq n_q, \; \sum_{1 \leq i \leq q} n_i = n_x$$

and if the following conditions are satisfied:

- $\forall j \in \{1, \dots, n_1\} : \dfrac{\partial h_j}{\partial x_j^1} \neq 0$ (7)
- $\forall i \in \{1, \dots, q-1\}, \quad \forall (\mathbf{x}, \mathbf{u}) \in \Re^{n_x} \times \Re^{n_u}$:

$$\operatorname{rank}\left(\frac{\partial \mathbf{f}^i(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}^{i+1}}\right) = \mathrm{n}_{i+1} \tag{8}$$

This canonical form assumes that only the first state subvector $\mathbf{x}^1$ is measured, i.e. $n_y = n_1$. Condition (7) states that $\mathbf{x}^1$ can be inferred directly from the measurements, whereas condition (8) implies a pyramidal influence of the state subvector $\mathbf{x}^{i+1}$ on $\mathbf{x}^i$, so that any differences in the state trajectory can be detected in the measurements.

A convenient way to check condition (8) is to compute the $(n_{i+1}) \times (n_{i+1})$ matrix

$$M_i(\mathbf{x}, \mathbf{u}) = \left(\frac{\partial \mathbf{f}^i(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}^{i+1}}\right)^T \left(\frac{\partial \mathbf{f}^i(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}^{i+1}}\right) \tag{9}$$

and check the rank condition:

$$\operatorname{rank}\left[M_i(\mathbf{x}, \mathbf{u})\right] = n_{i+1} \tag{10}$$

It is shown in (Bogaerts and Vande Wouwer, 2004) that an accurate process model can lead to poor estimates when the matrices $M_i(\mathbf{x}, \mathbf{u})$ are ill-conditioned, i.e. when the internal connections between state variables are somewhat "loose", at least in some time intervals. Upon analysis, this lack of connectivity is generally related to the model structure and the selection of operating conditions.

3. IDENTIFICATION FOR STATE ESTIMATION

This section discusses two ways of including the observability issue in the parameter identification: (i) a

weighted-sum cost function that penalizes the relative lack of observability (but requires an exhaustive search for determining the weighting), and (ii) a product cost function which avoids the introduction of a weighting coefficient.

3.1 Weighted-sum cost function

In (Bogaerts and Vande Wouwer, 2004), the objective function (4) has been extended to include a measure that indicates the lack of observability (*lobs*), which can be assessed from the determinant (or the condition number) of the matrix M_1:

$$J_{lobs}(\boldsymbol{\theta},\mathbf{x}_0)=\frac{1}{N}\sum_{k=1}^{N}\sum_{i=1}^{q-1}\left[\det\left(M_i\left(\mathbf{x}_k(\boldsymbol{\theta},\mathbf{x}_0),\mathbf{u}_k\right)\right)\right]^{-1/2} \tag{11}$$

A combined objective function is defined, where a compromise between model accuracy (small J_{id}) and observability (small J_{lobs}) is achieved through the weighting coefficient λ:

$$(\hat{\boldsymbol{\theta}},\hat{\mathbf{x}}_0)=\arg\min_{\boldsymbol{\theta},\mathbf{x}_0}\left\{J_{id}(\boldsymbol{\theta},\mathbf{x}_0)+\lambda J_{lobs}(\boldsymbol{\theta},\mathbf{x}_0)\right\} \tag{12}$$
$$\text{given} \quad \textit{model}\ (1)-(2);\ \mathbf{y}_{meas}$$

The optimal weighting coefficient can be chosen through extensive heuristic search, i.e. by considering different values of λ, computing the estimate (12) corresponding to each of these values, and representing the evolution of the two measures J_{id} and J_{lobs} as functions of λ. The "best" value of λ corresponds to improved system observability while keeping satisfactory model accuracy. The main disadvantage of this procedure is that it is relatively time consuming, as a series of compromises have to be tested in order to select the "best" one. The following section therefore suggests an alternative noniterative procedure.

3.2 Product cost function

The objective here is to investigate whether it is possible to transform problem (12) into a problem of the form:

$$(\hat{\boldsymbol{\theta}},\hat{\mathbf{x}}_0)=\arg\min_{\boldsymbol{\theta},\mathbf{x}_0}\left\{\tilde{J}_{id}(\boldsymbol{\theta},\mathbf{x}_0)\ \tilde{J}_{lobs}(\boldsymbol{\theta},\mathbf{x}_0)\right\} \tag{13}$$
$$\text{given} \quad \textit{model}\ (1)-(2);\ \mathbf{y}_{meas}$$

In this formulation, the weighting coefficient disappears, thus avoiding the extensive search mentioned above. However, several questions need answers:

- Is such a product objective function meaningful?
- What are appropriate choices of $\tilde{J}_{id}$ and $\tilde{J}_{lobs}$?

Formally, the goal of parameter identification for state estimation is to generate a model that is capable of describing the state evolution from the output measurements. The problem of estimating the states over a time interval can be expressed as that of estimating the initial states from measurements in that time interval (Chen, 1999). Hence, the problem of parameter identification for state estimation can be expressed as:

$$(\hat{\boldsymbol{\theta}},\hat{\mathbf{x}}_0)=\arg\min_{\boldsymbol{\theta},\mathbf{x}_0}\left\|\tilde{\mathbf{x}}_0(\boldsymbol{\theta},\mathbf{x}_0)-\mathbf{x}_0\right\| \tag{14}$$
$$\text{given} \quad \textit{model}\ (1)-(2);\ \mathbf{y}_{meas}$$

where $\tilde{\mathbf{x}}_0(\boldsymbol{\theta},\mathbf{x}_0)$ represents the initial states reconstructed from $\mathbf{y}_{meas}$ using the identified model, as explained below.

Most nonlinear systems can be approximated by linear time-varying models. Thus, without loss of generality, consider the unknown LTV discrete model:

$$\begin{aligned}\mathbf{x}_{k+1}&=A_k(\boldsymbol{\theta},\mathbf{x}_0)\mathbf{x}_k+B_k(\boldsymbol{\theta},\mathbf{x}_0)\mathbf{u}_k,\quad \mathbf{x}_{k=0}=\mathbf{x}_0\\ \mathbf{y}_k&=C_k(\boldsymbol{\theta},\mathbf{x}_0)\mathbf{x}_k\end{aligned} \tag{15}$$

The initial conditions can be expressed in terms of the N output measurements (Chen, 1999):

$$\tilde{\mathbf{x}}_0=\left(\sum_{k=0}^{N-1}\Psi_k^T\Psi_k\right)^{-1}\sum_{k=0}^{N-1}\Psi_k^T\bar{\mathbf{y}}_{meas,k} \tag{16}$$

$$\begin{aligned}&\text{with}\quad \bar{\mathbf{y}}_{\mathrm{meas,k}}=\mathbf{y}_{\mathrm{meas,k}}-\mathbf{y}_{\mathrm{k}}^0\\ &\mathbf{y}_k^0=C_k(\boldsymbol{\theta},\mathbf{x}_0)\sum_{j=0}^{k-1}\left(\prod_{i=k-1}^{j+1}A_i(\boldsymbol{\theta},\mathbf{x}_0)\right)B_j(\boldsymbol{\theta},\mathbf{x}_0)\mathbf{u}_j\\ &\Psi_k(\boldsymbol{\theta},\mathbf{x}_0)=C_k(\boldsymbol{\theta},\mathbf{x}_0)\left(\prod_{j=k-1}^{0}A_j(\boldsymbol{\theta},\mathbf{x}_0)\right)\end{aligned}$$

where $\mathbf{y}_k^0$ are the output values that would result from zero initial conditions. Hence, $\bar{\mathbf{y}}_{meas,k}$ expresses the effect of the initial conditions on the outputs $\mathbf{y}_{meas,k}$.

Using (15), $\mathbf{x}_0$ in (14) can be expressed as:

$$\mathbf{x}_0=\left(\sum_{k=0}^{N-1}\Psi_k^T\Psi_k\right)^{-1}\sum_{k=0}^{N-1}\Psi_k^T\bar{\mathbf{y}}_k \tag{17}$$

with $\bar{\mathbf{y}}_k=\mathbf{y}_k-\mathbf{y}_k^0$. This way, $\mathbf{x}_0$ can be expressed in a form similar to (16), and the objective function (14) becomes:

$$\begin{aligned}&\left\|\tilde{\mathbf{x}}_0(\boldsymbol{\theta},\mathbf{x}_0)-\mathbf{x}_0\right\|=\\ &=\left\|\left(\sum_{k=0}^{N-1}\Psi_k^T\Psi_k\right)^{-1}\left(\sum_{k=0}^{N-1}\Psi_k^T\left(\mathbf{y}_{meas,k}-\mathbf{y}_k\right)\right)\right\|\\ &=\left\|\left(O^TO\right)^{-1}O^T\left(Y_{meas}-Y\right)\right\|=\left\|O^+\left(Y_{meas}-Y\right)\right\|\end{aligned} \tag{18}$$

where the $(N n_y \times n_x)$ matrix O and the $N n_y$-dimensional vector Y are given by:

$$\mathcal{O} = \begin{bmatrix} \Psi_0 \\ \Psi_1 \\ \Psi_2 \\ \vdots \\ \Psi_{N-1} \end{bmatrix} = \begin{bmatrix} C_0(\boldsymbol{\theta},\mathbf{x}_0) \\ C_1(\boldsymbol{\theta},\mathbf{x}_0)A_1(\boldsymbol{\theta},\mathbf{x}_0) \\ C_2(\boldsymbol{\theta},\mathbf{x}_0)A_2(\boldsymbol{\theta},\mathbf{x}_0)A_1(\boldsymbol{\theta},\mathbf{x}_0) \\ \vdots \\ C_{N-1}(\boldsymbol{\theta},\mathbf{x}_0)\prod_{i=N-1}^{0} A_i(\boldsymbol{\theta},\mathbf{x}_0) \end{bmatrix}$$

$$Y_{meas} = \begin{bmatrix} \mathbf{y}_{meas,0} \\ \mathbf{y}_{meas,1} \\ \mathbf{y}_{meas,2} \\ \vdots \\ \mathbf{y}_{meas,N-1} \end{bmatrix} ; Y = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{N-1} \end{bmatrix}$$

and

$$\mathcal{O}^+ = \left(\mathcal{O}^T\mathcal{O}\right)^{-1}\mathcal{O}^T$$

Remarks

- The term $\mathcal{O}$ is an observability measure for LTV systems (note the resemblance with the Kalman observability matrix for LTI systems), while the term $(Y_{meas} - Y)$ expresses the output error. Hence, the objective function in (14) involves the *product* of terms expressing these two objectives.
- The objective function (18) can be bounded as follows:

$$\left\|\mathcal{O}^+(Y_{meas} - Y)\right\| \leq \left\|\mathcal{O}^+\right\| \|Y_{meas} - Y\| \quad (19)$$

which allows writing:

$$\tilde{J}_{lobs} = \left\|\mathcal{O}^+\right\| \qquad \tilde{J}_{id} = \|Y_{meas} - Y\| \quad (20)$$

The output error norm can be computed as in (4), while the $\mathcal{L}_2$ norm of $\mathcal{O}^+$ can be chosen, i.e. its largest singular value:

$$\left\|\mathcal{O}^+\right\|_2 = \sigma_{max}(\mathcal{O}^+) = \sigma^{-1}_{min}(\mathcal{O})$$

- In the final cost functions (18-19), an actual model linearization is only needed to derive the expression of $\mathcal{O}^+$, whereas the output Y is obtained through simulation of the nonlinear model equations using the parameter values θ and starting from the initial conditions x_0. As such, Y can be written in a more explicit way as $Y(\theta, x_0)$.

3.3 Case of full state state measurement

In (Bogaerts and Vande Wouwer, 2004), the parameter identification problem (3) has assumed that the full state vector can be measured off-line. This way, maximal experimental information is used for system identification. To mimic this situation here, $\mathbf{x}_{meas}$ should replace $\mathbf{y}_{meas}$ in the computation of J_{id} and $\tilde{J}_{id}$ in (12) and (20), respectively. However, since the on-line estimator will use the measurements of only a subset of the state variables, practical observability problems might result. This structural "lack of observability" is expressed by J_{lobs} and $\tilde{J}_{lobs}$, i.e., these quantities need to be computed from the model using the output $\mathbf{y}$ and not $\mathbf{y} = \mathbf{x}$.

4. CASE STUDIES

Two case studies are considered to illustrate the concept of parameter identification for state estimation. To show that the results do not depend on a particular state estimation algorithm, two different observers are tested, i.e. an extended Kalman filter (Maybeck, 1982) and a particle observer (Doucet *et al.*, 2001).

4.1 Simulation case study

A fed-batch bioreactor is considered. The reaction scheme describes the growth of biomass on glutamine and a maintenance process based on glucose that leads to the production of lactate:

Growth:

$$\nu_{Gln}\,Gln \xrightarrow{\varphi_g}{}^{\wedge}\; X \quad (21)$$

Maintenance:

$$G + \nu_X X \xrightarrow{\varphi_m} \nu_X X + \nu_L L \quad (22)$$

where G, Gln, X and L denote glucose, glutamine, biomass and lactate, respectively. ν_{Gln} , ν_X et ν_L are yield coefficients. The symbol "$\longrightarrow^{\wedge}$" means that the growth reaction is auto-catalysed by X, and the presence of biomass on both sides of the maintenance reaction indicates that X catalyses this reaction.

The growth rate φ_g and the maintenance rate φ_m are described by Monod laws with inhibition factors:

$$\varphi_g(X,G,Gln) = \mu^g_{\max}\frac{Gln}{K^g_M + Gln}\frac{K^g_i}{K^g_i + G}X \quad (23)$$

$$\varphi_m(X,G) = \mu^m_{\max}\frac{G}{K^m_M + G}\frac{K^m_i}{K^m_i + X}X \quad (24)$$

The mathematical model is obtained from mass balances:

$$\frac{dX}{dt} = \varphi_g - DX \quad (25)$$

$$\frac{dG}{dt} = -\varphi_m + D(G_{in} - G) \quad (26)$$

$$\frac{dGln}{dt} = -\nu_{Gln}\varphi_g + D(Gln_{in} - Gln) \quad (27)$$

$$\frac{dL}{dt} = \nu_L\varphi_m - DL \quad (28)$$

$$\frac{dV}{dt} = Q_{in} \quad (29)$$

where Q_{in} is the feed rate, V is the reactor volume, D is the dilution rate ($D = \frac{Q_{in}}{V}$), and G_{in} and Gln_{in} are the feed concentrations. Numerical values of the model parameters are listed in Table 1. Three experiments

ν_L	1.7	ν_{Gln}(mM/(10^5 cell/ml))	0.2
$\mu^g_{\max}$ (h^{-1})	0.05	$\mu^m_{\max}$ (h^{-1})	0.1
K^g_M (mM)	0.1	K^m_M (mM)	0.2
K^g_i (mM)	70	K^m_i(10^5cell/ml)	3

Table 1. *Model parameters*

are performed that differ in their initial conditions,

inlet substrate concentrations and flow rate (Table 2). The inlet flow rate is chosen as $Q_{in} = 0$ for $t < t_b$ and $Q_{in} = \alpha(t - t_b)$ for $t_b \leq t \leq t_f$, with $\alpha = 5 \times 10^{-4}\ l/h^2$. Samples are taken every 8 h and are used to evaluate the performance of the various state estimators. The measurements are corrupted by noises, corresponding to the assumption of constant relative errors, e.g. $\varepsilon_X^{rel} = 0.1$, $\varepsilon_G^{rel} = 0.05$, $\varepsilon_{Gln}^{rel} = 0.05$ and $\varepsilon_L^{rel} = 0.05$.

	Exp 1	Exp 2	Exp 3
X(0) (10^5 cell/ml)	1	1	1
G(0) (mM)	10	10	8
Gln(0) (mM)	8	1	5
L(0) (mM)	1	1	1
V(0) (l)	0.5	0.5	0.5
G_{in} (mM)	5	5	10
Gln_{in} (mM)	1	1	3
t_b (h)	30	35	50
t_f (h)	80	80	240

Table 2. *Experimental conditions*

Based on full-state measurements from the first two experiments, the identification procedure provides accurate model parameters, thus demonstrating that no identifiability problem occurs.

Figure 1 illustrates state estimation with an EKF based on a model identified using one of three objective functions, i.e. with either the objective function "output error" (3), "weighted sum" (12) or "product" (19). Experiment 3 is used, and the various estimates are compared to the off-line measurements with a 99% confidence interval. The EKF uses on-line glutamine and lactate measurements, and the initial conditions for the unmeasured state are taken as $\hat{X}(0) = 5 \times 10^5$ cell/ml, $\hat{G}(0) = 15$ mM with the standard deviations $\sigma_X(0) = 10^6$ cell/ml, $\sigma_G(0) = 10$ mM. Clearly, the convergence of the state estimates, in particular for glucose, improves when using an objective function that considers system observability. Figure 2 shows the same trend with a particle-filter observer. Interestingly, the particle filter (PF) appears more robust, particularly when the number of particles is sufficiently large so as to have a good exploration of the state space. PF has certainly the advantage that it does not rely on a linearization of the process model.

To give an idea of the computational load, an optimization run takes about 5 to 20 minutes using a Pentium 4 computer and the MATLAB optimization toolbox. This load depends on the initial starting point for the optimization problem. It is nearly the same for all objective function types. However, with the objective function (12), it is necessary to repeat the identification procedure for several values of λ. In this simulation example, the λ values were chosen in the interval $[0, \ldots, 1]$ as follows: $0, 10^{-10}, \ldots, 10^{-1}, 1$. This range can be refined if necessary to improve precision. Consequently, the computation time using the objective function (12) can be 10 to 20 times longer than with the objective function (13). However, the latter objective function can be more prone to local minima, and some kind of multi-start strategy might

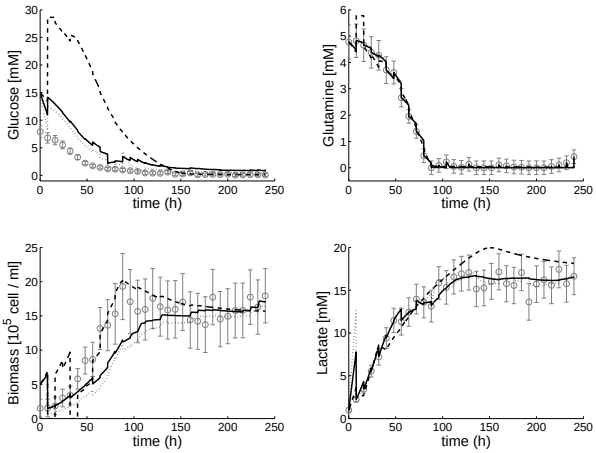

Fig. 1. State estimation with EKF using the objective function "output error" (3) (- -), "weighted sum" (12) ($\cdots$), and "product" (19) (–).

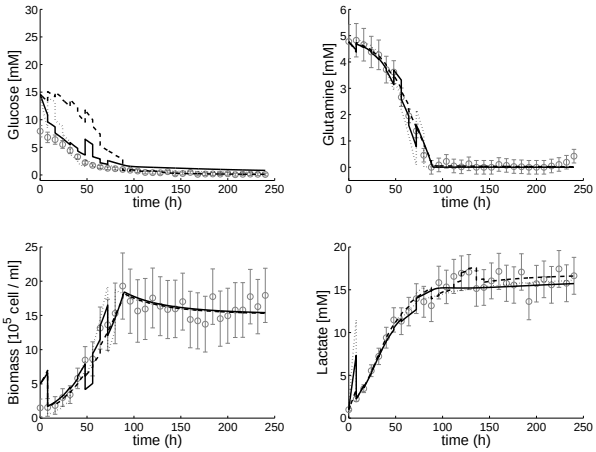

Fig. 2. State estimation with a particle filter ($N = 500$) using the objective function "output error" (3) (- -), "weighted sum" (12) ($\cdots$), and "product" (19) (–).

be necessary, which slightly reduces the benefits of the procedure. These local minima are essentially of two types: (a) minima that ensure a good compromise between J_{id} and J_{lobs} and (b) minima that correspond to the particularly damaging situation where one of the two measures can be reduced to a very low level, without taking care of the value of the other measure. In this latter case, the improvement in observability can be detrimental to accuracy, or vice versa.

4.2 Experimental application

The experimental study concerns batch animal cell (CHO-K1) cultures. The model equations are those of the simulated example with $Q_{in} = D = 0$. The kinetics are described by

$$\begin{aligned} \varphi_g = {} & \alpha_g X^{\gamma_{g,X}} G^{\gamma_{g,G}} Gln^{\gamma_{g,Gln}} \\ & e^{-\beta_{g,X} X} e^{-\beta_{g,G} G} e^{-\beta_{g,Gln} Gln} e^{-\beta_{g,L} L} \\ \varphi_m = {} & \alpha_m X^{\gamma_{m,X}} G^{\gamma_{m,G}} Gln^{\gamma_{m,Gln}} \\ & e^{-\beta_{m,X} X} e^{-\beta_{m,G} G} e^{-\beta_{m,Gln} Gln} e^{-\beta_{m,L} L} \end{aligned} \tag{30}$$

Rare and asynchronous measurements of biomass X, glutamine Gln, glucose G and lactate L concentrations are available for model identification, whereas glutamine and lactate measurements only are used for state estimation. Again, the three objective functions are used to provide model parameters. First, a standard parameter identification, i.e. without consideration of system observability, is performed, and the model cross-validated is shown in Fig. 3. With the objective function (12), a series of repeated optimization runs suggests the value $\lambda = 10^{-6}$.

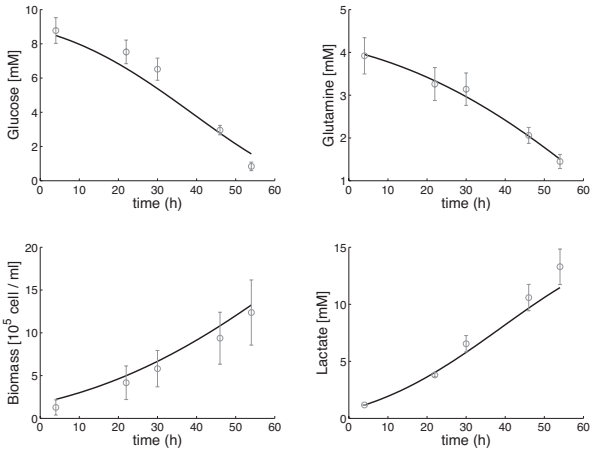

Fig. 3. Validation of models obtained via standard parameter identification, i.e. without consideration of system observability.

Figure 4 illustrates the performance of a PF, in particular its good convergence properties.

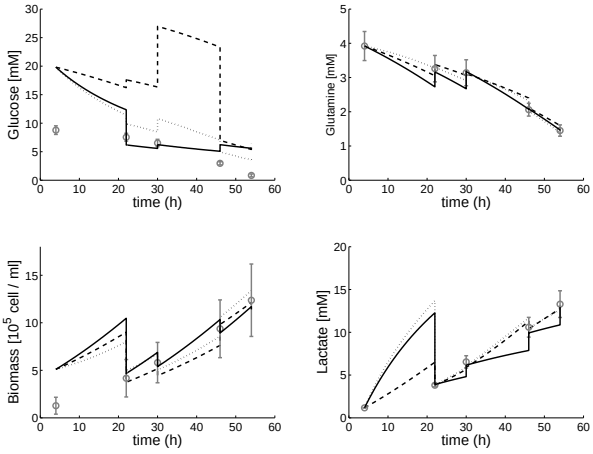

Fig. 4. State estimation with a particle filter ($N = 500$) using the objective function "output error" (3) (--), "weighted sum" (12) ($\cdots$), and "product" (19) (–).

5. CONCLUSIONS

A variation to the objective function given in (Bogaerts and Vande Wouwer, 2004) has been proposed. It uses the product of a measure of model accuracy and a measure of system observability rather than a weighted sum of the two. Consideration of system observability improves the rate of convergence by increasing the sensitivity to the measurements. The advantage of the proposed objective function is twofold: (i) The identification process is no longer repetitive, and (ii) the objective function has a sound theoretical (less heuristic) justification. The disadvantage of the product objective function is however the existence of local minima corresponding to the sought compromise and to unbalanced situations where one of the two model measures (accuracy or observability) is pushed to the extreme at the expense of the other measure. A use of a multi-start strategy is therefore advised.

One could also argue that a "falsified" model will lead to steady-state offsets, and this might indeed be observed in practice. However, the proposed procedure is mostly intended for speeding up the convergence of state estimates in batch or fed-batch processes, i.e. where time of operation is limited and monitoring is important from the beginning.

6. ACKNOWLEDGEMENT

This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its author(s).

REFERENCES

Bogaerts, P. and A. Vande Wouwer (2004). Parameter identification for state estimation application to bioprocess software sensors. *Elsevier, Chemical Engineering Science* **59**(12), 2465–2476.

Chen, C. T. (1999). *Linear System Theory and Design*. Oxford Unviersity Press.

Doucet, A., N. de Freitas and N. Gordon (2001). *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New-York.

Gauthier, J.-P. and I. Kupka (1994). Observability and observers for nonlinear systems. *SIAM Journal Control and Optimization* **32**(4), 975–994.

Haseltine, E.L. and J.B. Rawlings (2005). Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Industrial and Engineering Chemistry Research* **44**(8), 2451–2460.

Maybeck, P.S. (1982). *Stochastic Models, Estimation, and Control, Volume 2*. Academic Press, New York.

Zeitz, M. (1984). Observability canonical (phase-variable) form for nonlinear time-variable systems. *International Journal of System Science* **15**(9), 949–958.

Zeitz, M. (1989). *Nonlinear Control Systems Design*. Chap. Canonical forms for nonlinear systems, pp. 33–38. Pergamon Press.