Preprints of the
9th International Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
June 7-10, 2015, Whistler, British Columbia, Canada

Extended and Unscented Kalman Filter design for hybridoma cell fed-batch and continuous cultures.

Sofia Fernandes * **Anne Richelle** ** **Zakaria Amribt** ** **Laurent Dewasme** *
Philippe Bogaerts ** **Alain Vande Wouwer** *

* *Service d'Automatique, University of Mons, 31 Boulevard Dolez, 7000 Mons, Belgium (e-mail: sofia.afonsofernandes, laurent.dewasme and alain.vandewouwer@umons.ac.be)*
** *3BIO-BioControl, Brussels School of Engineering, Universit Libre de Bruxelles, Av. F.-D. Roosevelt 50 C.P. 165/61, 1050 Brussels, Belgium (e-mail: arichell and Philippe.Bogaerts@ulb.ac.be)*

Abstract: In order to mantain hybridoma cell cultures in optimal operating conditions, on-line measurements of glutamine and glucose concentrations are required, implying the availability of probes, which are expensive and with poor durability. A way to overcome this problem is to design software sensors. In this work, both Extended and Unscented Kalman Filters are developed in order to estimate glucose and glutamine concentrations, based on biomass, lactate and ammonia on-line measurements. System observability conditions are first examined. The performances of both software sensors are analyzed with simulations of hybridoma cell cultures in fed-batch and continuous bioreactor operating modes. Three different tests are conducted in order to compare the performance of both observers: continuous culture with constant feeding profile, fed-batch culture with both constant and exponential feeding profiles. Also, two different sets of parameters are investigated: the ones obtained by using the least-squares method in order to minimize the error between model predictions and experimental measurements, and the ones which are modified by minimizing a cost function combining the usual least-squares criterion with a state estimation sensitivity criterion.

Keywords: Hybridoma cultures, state estimation, software sensors, observability, Extended Kalman Filter, Unscented Kalman Filter.

1. INTRODUCTION

Hybridoma cell cultures are widely used for the production of monoclonal antibodies with therapeutic purposes. In recent years, many improvements have been developed so as to ensure the growing demand for these kinds of recombinant proteins. However, large-scale production is often limited by problems related to the availability of adequate process control strategy ensuring optimal culture conditions. Moreover, the development of such control tools requires reliable monitoring systems on the production system in order to maintain the cell in a specific metabolic state (Rodrigues et al., 2010; Wurm, 2004).

Indeed, hybridoma cells exhibit an overflow metabolism phenomenon in presence of an excessive substrate concentration (glucose and/or glutamine) in the medium leading to a by-product formation (lactate and ammonia) and the inhibition of cell growth (Amribt et al., 2013). Hence, the substrate concentrations need to be maintained at a critical concentration value to avoid this undesirable effect (Amribt et al., 2014). This kind of control requires the presence of reliable probes for the on-line measurement of these variables. As these ones are expensive and present poor durability (about 1-3 months), the design of software sensors for glucose and glutamine estimation is a useful alternative strategy widely recognized in bioprocess monitoring and control (Dewasme et al., 2013; Hitzmann et al., 2000; Veloso et al., 2009).

The Extended Kalman Filter (EKF) appears as the most used state estimator when the system is nonlinear (Simon, 2006). Although the EKF is a widely used filtering strategy, it presents some disadvantages. It is reliable for systems which are almost linear on the time scale of the update intervals; it requires the calculation of Jacobians at each time step, which may be difficult to obtain for higher order systems; it does linear approximations of the system at a given time instant, which may introduce errors in the state, leading then the state to diverge over time (Julier and Uhlmann, 1997; Wan and Van Der Merwe, 2000; Zhu and Feng, 2012). In order to handle these problems, the unscented Kalman filter (UKF) was proposed by Julier and Uhlmann (2004). The UKF uses the unscented transformation (UT) based on the idea that it is easier to approximate a probability distribution than a nonlinear function. The advantage of the UKF with respect to the EKF is that no jacobians need to be computed, therefore no linearization errors are introduced.

The present study uses a macroscopic model taking account of an overflow metabolism within glycolysis and glutaminolysis (Amribt et al., 2013) in order to compare the performance between both Extended and Unscented Kalman Filters.

In many study cases, sensitivity of measured model states with respect to the unmeasured ones is poor, leading to poor estimation quality. This is due to the fact that when using least-squares method to identify model parameters, no guarantee is given about sensitivity of measured with respect to unmeasured states. To overcome this problem, a parameter identification

procedure is proposed in Bogaerts and Vande Wouwer (2004), which is based on a cost function combining the usual least-squares criterion with a state estimation sensitivity criterion.

The motivation of this work is twofold: on the one hand to compare the performance of both EKF and UKF observers, to estimate glucose and glutamine concentrations, based on biomass, lactate and ammonia on-line measurements and, on the other hand, to show the effectiveness of the parameter identification for state estimation procedure.

This paper is organized as follows. The dynamic model is presented in the next section. In section 3, the system observability condition is analyzed and both methods for parameter identification are introduced: the least-squares method and the cost function combining the usual least-squares criterion with a state estimation sensitivity criterion. Sections 4 and 5 describe the EKF and UKF, respectively. Results are presented in section 6 and conclusions are pointed out in section 7.

2. OVERFLOW MODEL

The present study uses a macroscopic model developed in Amribt et al. (2013) which takes into account the overflow metabolism of glycolysis and glutaminolysis of hybridoma cells. Macroscopic reactions representing the respiratory and the overflow metabolism are assumed as follows:

(1) Respiratory metabolism

$$Glc \xrightarrow{\varphi_{Glc}} a \times X + b \times L \tag{1}$$

$$Gln \xrightarrow{\varphi_{Gln}} c \times X + d \times N \tag{2}$$

(2) Overflow metabolism

$$Glc \xrightarrow{\varphi_{Glc\_over}} 2 \times L \tag{3}$$

$$Gln \xrightarrow{\varphi_{Gln\_over}} N + \frac{1}{2} \times L \tag{4}$$

where X, Glc, Gln, L and N are the concentrations of biomass, glucose, glutamine, lactate and ammonia, respectively. Note that a, b, c and d are the stoichiometric coefficients. φ_{Glc}, φ_{Gln}, φ_{Glc_over} and φ_{Gln_over} are the reaction rates used to describe respiratory and overflow metabolisms. Reaction rates are given by:

$$\varphi_{Glc} = min(\varphi_{Glc1}, \varphi_{Glc\_max}) \tag{5}$$

$$\varphi_{Gln} = min(\varphi_{Gln1}, \varphi_{Gln\_max}) \tag{6}$$

$$\varphi_{Glc\_over} = max(0, \varphi_{Glc1} - \varphi_{Glc\_max}) \tag{7}$$

$$\varphi_{Gln\_over} = max(0, \varphi_{Gln1} - \varphi_{Gln\_max}) \tag{8}$$

where glucose and glutamine consumption rates (φ_{Glc1} and φ_{Gln1}) are represented as extended Monod-kinetics. Maximum growth capacity for glucose and glutamine (φ_{Glc_max} and φ_{Gln_max}) are expressed as first order kinetics with respect to the biomass concentration:

$$\varphi_{Glc1} = \mu_{Glc\_max1} \frac{Glc}{K_{Glc} + Glc} \frac{Gln}{K_{Gln1} + Gln} X_v \tag{9}$$

$$\varphi_{Gln1} = \mu_{Gln\_max1} \frac{Gln}{K_{Gln} + Gln} \frac{K_N}{K_N + N} X_v \tag{10}$$

$$\varphi_{Glc\_max} = \mu_{Glc\_max2} X_v \tag{11}$$

$$\varphi_{Gln\_max} = \mu_{Gln\_max2} X_v \tag{12}$$

where $\mu_{i_{max\,j}}$ ($i = Glc, Gln$, $j = 1, 2$) are the maximum values of the specific rates and K_{Glc}, K_{Gln1} and K_{Gln} are the saturation coefficients. K_N is the ammonia inhibition constant over the oxidation of glutamine.

Mass balances for viable biomass (X_v), glucose, glutamine, lactate, ammonia are given by the following differential equations:

$$\frac{dX_v}{dt} = a\varphi_{Glc} + c\varphi_{Gln} - \mu_d X_v - DX_v \tag{13}$$

$$\frac{dGlc}{dt} = -\varphi_{Glc} - m_G X_v - \varphi_{Glc\_over} + D(Glc_{in} - Glc) \tag{14}$$

$$\frac{dGln}{dt} = -\varphi_{Gln} - \varphi_{Gln\_over} + D(Gln_{in} - Gln) \tag{15}$$

$$\frac{dL}{dt} = b\varphi_{Gln} + 2\varphi_{Glc\_over} + \frac{1}{2}\varphi_{Gln\_over} - DL \tag{16}$$

$$\frac{dN}{dt} = d\varphi_{Gln} + \varphi_{Gln\_over} - DN \tag{17}$$

$$\frac{dV}{dt} = F_{in} \tag{18}$$

with dilution rate $D = \frac{F_{in}}{V}$, where F_{in} is the inlet feed rate and V the broth volume. Glc_{in} and Gln_{in} are the glucose and glutamine concentrations in the feed medium, respectively. m_G represents the maintenance coefficient of glucose and the specific death rate μ_d is assumed to be constant. Note that in the case of this study, the authors have chosen not to consider the dynamics associated with dead biomass which is normally included as a state variable in the model of Amribt et al. (2013). Indeed, this state variable can be easily excluded from the model only taking into account the viable biomass concentration.

3. SYSTEM OBSERVABILITY

Before proceeding to state estimation, system observability has to be checked. Observability is a system property, which depends on the input signal in nonlinear systems. The global observability analysis of nonlinear models can be simplified through the introduction of a canonical form (Zeitz, 1984; Gauthier and Kupka, 1994) given by:

$$\dot{\xi} = \begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \vdots \\ \dot{\xi}_{q-1} \\ \dot{\xi}_q \end{bmatrix} = \begin{bmatrix} f_1(\xi_1, \xi_2) \\ f_2(\xi_1, \xi_2, \xi_3) \\ \vdots \\ f_{q-1}(\xi_1, \ldots, \xi_{q-1}, \xi_q) \\ f_q(\xi_1, \ldots, \xi_{q-1}, \xi_q) \end{bmatrix} \tag{19}$$

$$y = \begin{bmatrix} h_1(\xi_1^1) \\ h_2(\xi_1^1, \xi_1^2) \\ \vdots \\ h_{n_1}(\xi_1^1, \ldots, \xi_1^{n_1}) \end{bmatrix} \tag{20}$$

where $\forall i \in 1, \ldots, q$, $\xi_i \in \Re^{n_i}$, $n_1 \geq n_2 \geq \cdots \geq n_q$, $\sum_{1 \leq i \leq q} n_i = N = dim\xi$. Then, a system is said globally observable if the following conditions are satisfied:

$$\frac{\partial h_j}{\partial \xi_1^j} \neq 0, \quad \forall j \in 1, \ldots, n_1 \tag{21}$$

$$rank \frac{\partial f_i}{\partial \xi_{i+1}} = n_{i+1}, \quad \forall i \in 1, \ldots, q-1 \tag{22}$$

The first condition implies that the first n_1 state variables can be inferred from the measurements, while the second ensures that any differences in the state trajectory can be detected in the measurements thanks to a pyramidal influence of the state subvector ξ_{i+1} on the evolution equations $\dot{\xi}_i$.

In doing so, the model presented above can be put in the canonical form as follows:

$$\dot{\xi} = \begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} f_1(\xi_1, \xi_2) \\ f_2(\xi_1, \xi_2) \end{bmatrix}, \quad y = \xi_1 \tag{23}$$

where $\xi_1 = [X_v \;\; L \;\; N]^T$ and $\xi_2 = [Glc \;\; Gln]^T$. The observability test is given by:

$$rank \frac{\partial f_1}{\partial \xi_2} = \begin{bmatrix} \frac{\partial \dot{X}_v}{\partial Glc} & \frac{\partial \dot{X}_v}{\partial Gln} \\ \frac{\partial \dot{L}}{\partial Glc} & \frac{\partial \dot{L}}{\partial Gln} \\ \frac{\partial \dot{N}}{\partial Glc} & \frac{\partial \dot{N}}{\partial Gln} \end{bmatrix} = n_2 = 2 \tag{24}$$

and is verified if Glc, Gln and N concentrations do not vanish.

3.1 Parameter identification for state estimation

One of the requirements to build state observers is to investigate the observability of the system. What happens sometimes is that even if the system is observable, the ability to detect, in the output trajectories, a difference in the initial states is difficult, leading to poor estimation quality. This fact was a motivation to define a new cost function $F(\theta)$ combining the identification criterion $J(\theta)$ with the observability measure $F_{obs}(\theta)$ in Bogaerts and Vande Wouwer (2004), yielding, therefore, a model dedicated to state estimation purposes.

$$F(\theta) = J(\theta) + \lambda F_{obs}(\theta) \tag{25}$$

where λ is the weighting factor and F_{obs} is given by:

$$F_{obs}(\theta) = \sum_{j=1}^{q} \sum_{i=1}^{p} \sqrt{cond \left[\left(\frac{\partial f_1}{\partial \xi_2} \right)_{ij}^T \left(\frac{\partial f_1}{\partial \xi_2} \right)_{ij} \right]} \tag{26}$$

"cond" being the condition number of the matrix (the ratio of its largest to its smallest eigenvalue), q the number of experiments and p the number of measurements.

Both nominal and modified sets of parameters are used in the present study in order to design and compare the performances of both EKF and UKF software sensors.

4. EXTENDED KALMAN FILTER

The Extended Kalman Filter (EKF) is widely used to estimate the states when the system is nonlinear (Simon, 2006). The EKF algorithm involves a linearization of the nonlinear state equations around the current state estimate $\hat{x}$. As a result, a group of iterative equations is obtained, which are similar to the Kalman Filter (Kalman, 1960) equations for linear systems.

Given the following nonlinear system:

$$x_k = f(x_{k-1}, \upsilon_{k-1}, u_{k-1}), \tag{27}$$

$$y_k = h(x_k, n_k, u_k), \tag{28}$$

where $x \in \mathbb{R}^{n_x}$ is the system state, $\upsilon \in \mathbb{R}^{n_\upsilon}$ the process noise, $n \in \mathbb{R}^{n_n}$ the observation noise, u the input and y the noisy observation of the system. The EKF algorithm proceeds in two steps: a prediction step (corresponding to the time period between two measurement times) and a correction step occurring each time a new measurement is available.

Step 1: Prediction step between t_k and t_{k+1}

Prediction step linearizes the system dynamics, yielding the state estimate $\hat{x}_k^-$ as well as the covariance $P_{x_k}^-$ *a priori* by using the Jacobian function, as follows:

$$\begin{aligned} \hat{x}_k^- &= f(\hat{x}_{k-1}, u_k), \\ P_{x_k}^- &= F_{x_k} P_{x_k} F_{x_k}^T + Q_{k-1} \end{aligned} \tag{29}$$

where Q_{k-1} is the process noise covariance matrix and F_{x_k} is the Jacobian function.

$$F_{x_k} = \left. \frac{\partial f}{\partial x} \right|_{x=\hat{x}(k)} \tag{30}$$

Step 2: Correction step at time t_{k+1}

The observation dynamics y_k are linearized around the *a priori* state estimate $\hat{x}_k^-$.

At the end, the state estimate $\hat{x}_k$ and the covariance P_{x_k} *a posteriori* are obtained thanks to the correction term K_k.

$$\begin{aligned} K_k &= P_{x_k}^- H_{x_k}^T (H_{x_k} P_{x_k}^- H_{x_k} + R_k)^{-1}, \\ \hat{x}_k &= \hat{x}_k^- + K_k [y_k - H(\hat{x}_k^-)], \\ P_{x_k} &= (I - K_k H_{x_k}) P_{x_k}^-. \end{aligned} \tag{31}$$

where R_k is the measurement noise covariance matrix and H_{x_k} is the Jacobian function given by:

$$H_{x_k} = \left. \frac{\partial h}{\partial x} \right|_{x=\hat{x}(k)} \tag{32}$$

5. UNSCENTED KALMAN FILTER

UKF addresses the mentioned problems of EKF by using a deterministic sampling approach. The state distribution is represented by a GRV (Gaussian Random Variable), but is now specified using a minimal set of carefully chosen sample points (called sigma points). These sample points completely capture the true mean and covariance of the GRV, and when propagated through the true non-linear system, captures the posterior mean and covariance accurately.

5.1 UKF Algorithm

UKF computes a set of $2n+1$ sigma points χ_i (with corresponding weight w_i) of the random variable x of dimension n, mean $\bar{x}$, and covariance P_x. The statistics of the random variable x is calculated using the unscented transforms (UT) as follows:

$$\begin{aligned}
&\chi_0 = \bar{x} \\
&\chi_i = \bar{x} + (\sqrt{(n+\lambda)P_x})_i && i = 1, ..., n \\
&\chi_i = \bar{x} - (\sqrt{(n+\lambda)P_x})_{i-n} && i = n+1, ..., 2n \\
&w_0^m = \lambda/(n+\lambda) \\
&w_0^c = \lambda/(n+\lambda) + (1-\alpha^2+\beta) \\
&w_i^m = w_i^c = \lambda/(2(n+\lambda)) && i = 1, ..., 2n
\end{aligned}$$

where $\lambda = \alpha^2(n+\kappa) - n$ is a scaling parameter, α determines the spread of the sigma points around $\bar{x}$ and is usually set to a small positive value ($0 < \alpha \leq 1$). $\kappa \geq 0$ must be chosen to guarantee the semi-positive definiteness of the covariance matrix. A good default choise is $\kappa = 0$. β is a tuning parameter, which can be used to incorporate knowledge of the distribution. For a Gaussian distribution the optimal choice is $\beta = 2$ (Kandepu et al., 2008).

Step 1 - During the prediction step, the UKF algorithm exhibits as follows:

The transformed sigma points are given by instantiating each point through the process model:

$$\chi^-_{i(k)} = f[x_k, \chi^-_{i(k-1)}] \tag{33}$$

The predicted mean $\hat{x}_k^-$ and covariance P_k^- are computed as follow:

$$\begin{aligned}
&\chi^-_{i(k)} = f[x_k, \chi^-_{i(k-1)}] \\
&\hat{x}_k^- = \sum_{i=0}^{2n} w_i \chi^-_{i(k)} \\
&P_k^- = \sum_{i=0}^{2n} w_i [\chi^-_{i(k)} - \hat{x}_k^-][\chi^-_{i(k)} - \hat{x}_k^-]^T + Q_{k-1}
\end{aligned} \tag{34}$$

Step 2 - During the correction step of UKF, the mean and covariance calculated during the prediction are used together with the measurements at time k to correct the new values.

The prediction sigma points are propagated through the observation model:

$$\hat{Y}_{i(k)} = h[x_k, \chi^-_{i(k)}] \tag{35}$$

$$\begin{aligned}
&\hat{y}_k = \sum_{i=0}^{2n} w_i^m \hat{Y}_{i(k)} \\
&P_{\bar{y}_k \bar{y}_k} = \sum_{i=0}^{2n} w_i^c [\hat{Y}_{i(k)} - \hat{y}_k][\hat{Y}_{i(k)} - \hat{y}_k]^T \\
&P_{x_k y_k} = \sum_{i=0}^{2n} w_i^c [\chi^-_{i(k)} - \hat{x}_k^-][\hat{Y}_{i(k)} - \hat{y}_k]^T \\
&S_k = P_{\bar{y}_k \bar{y}_k} + R_{k-1} \\
&K_k = P_{x_k y_k} S_k^{-1} \\
&\hat{x}_k = \hat{x}_k^- + K_k(y_k - \hat{y}_k) \\
&P_k = P_k^- - k_k S_k k_k^T
\end{aligned} \tag{36}$$

where $\hat{y}_k^-$, $P_{\bar{y}_k}^-$, $P_{x_k y_k}$, K_k, $\hat{x}_k$ and P_k are the mean and covariance of the measurement vector, the cross covariance, the Kalman gain, estimated state and covariance, respectively.

6. RESULTS AND DISCUSSION

EKF and UKF observers are applied in order to estimate glucose and glutamine concentrations from biomass, lactate and ammonia measurements. Biomass can be measured off-line using microscopic cell counting and on-line using a Fogale Nanotech probe. Lactate and ammonium can be measured off-line by ultra performance liquid chromatography (UPLC) and in-line using a calibrated Near Infra-Red (NIR) probe.

Due to the poor estimation of glutamine concentration, a set of modified parameters based on a cost function (25) combining the usual least-squares criterion with a state estimation sensitivity criterion, when applied to a structurally comparable model defined by a set of nominal parameters identified by a classical least-squares method is also applied. This procedure was also applied in Amribt et al. (2014), from where the parameters values (see table 1) are taken.

Table 1. Identified parameter values.

Parameters	Nominal values	Modified values	Units
μ_{Glc_max1}	1.0006	1.5265	$mmol/(10^9 cellsh)$
μ_{Glc_max2}	0.0283	0.0371	$mmol/(10^9 cellsh)$
μ_{Gln_max1}	0.1992	0.1447	$mmol/(10^9 cellsh)$
μ_{Gln_max2}	0.0203	0.0222	$mmol/(10^9 cellsh)$
μ_{dmax}	0.0111	0.4753	h^-1
K_{Glc}	23.2350	42.7822	mM
K_{Gln}	0.0004	0.2770	mM
K_N	0.9931	3.5332	mM
K_{Gln1}	0.0005	0.2006	mM
K_{Gd}	2.1862	1.7429	mM
a	1.1462	0.8757	$10^9 cells/mmol$
b	1.2939	1.1806	$mmol/mmol$
c	0.1186	0.0805	$10^9 cells/mmol$
d	0.3000	0.4099	$mmol/mmol$
m_G	0.0367	0.0352	$mmol/(10^9 cellsh)$
K_{Gnd}	0.002	0.002	mM

Three different tests are conducted in order to compare the performance of both observers: continuous culture with constant feeding profile, fed-batch culture with both constant and exponential feeding profiles. Continuous culture conditions are the following: initial volume $V = 0.35$L, dilution rate $D = 0.0197h^{-1}$. The feeding starts at time $t = 35$h. Glucose and glutamine feeding are $Glc_{in} = 15$mM and $Gln_{in} = 4$mM, respectively and final culture time is $t_f = 200$h. Concerning the fed-batch culture with constant feeding, the same conditions as the ones used in continuous culture are considered, except the dilution rate. From $D = \frac{F_{in}}{V}$, we assume an input feeding of

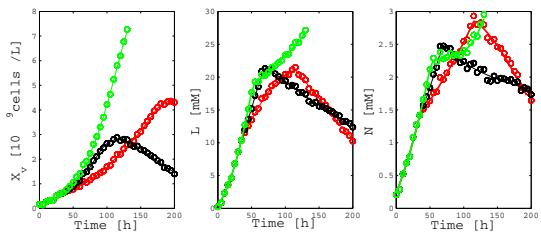

Fig. 1. Biomass, lactate and ammonia on-line measurements simulated with nominal parameter values. Red: Continuous culture. Black: Fed-batch culture with constant feeding. Green: Fed-batch culture with exponential feeding.

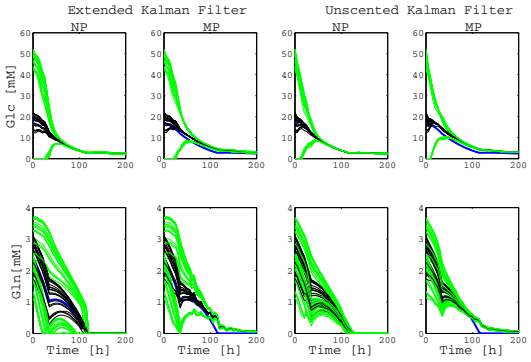

Fig. 2. Glucose and glutamine estimation with EKF and UKF using nominal and modified parameters values (20 runs by varying initial conditions - black curves) for a continuous culture with constant feeding. In blue: model evolution. In green: Confidence intervals at 95%.

F_{in} = 0.1L/day, which starts at time t = 35h. For the last test, the fed-batch culture with exponential feeding, the feeding is represented by $F(t) = 6.10^{-4}e^{0.038\times(t-52)}$ and starts at t = 52h. Initial volume V = 0.35L, glucose and glutamine feeding are Glc_{in} = 38mM and Gln_{in} = 10mM, respectively. Culture ends at time t_f = 133h.

Observer initial conditions are taken randomly at maximum 30% of the real initial conditions: $Xv = 0.185 \times 10^6$cells/mL, Glc = 17.17mM, Gln = 2.41mM, L = 0.36mM, N = 0.23mM.

The noise standard deviation is chosen as 5% for biomass and ammonia and 25% for lactate. Being x_{i_0} the initial conditions and N the number of considered states; initial covariance and model noise covariance are given by: $P_0 = (diag[0.2 \times x_{i_0}]^2)$ $Q_0 = 0.1^2 \times eye(N)$ are given by.

As dynamics are changing depending on bioreactor mode, the sigma points tuning parameters α and κ also have to change. For the first test α = 0.9 and κ = 0.6. For the second test $\alpha = 1.0^{-3}$ and $\kappa = 1.0^{-4}$. For the third test α = 0.4 and κ = 0.

In order to compare the different simulation results, Root Mean Square Error (RMSE) of the estimation of glucose and glutamine when varying initial conditions are calculated as follows:

$$RMSE_k = \sqrt{\frac{\sum_{j=1}^{N}\sum_{i=1}^{n}(Xobs_{k,ij} - Xmod_{k,ij})^2}{n \times N}} \qquad (37)$$

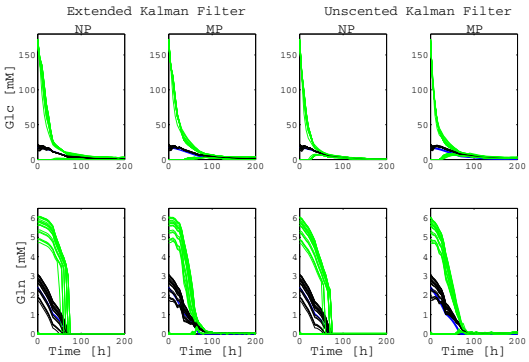

Fig. 3. Glucose and glutamine estimation with EKF and UKF using nominal and modified parameters values (20 runs by varying initial conditions - black curves) for a fed-batch culture with constant feeding. In blue: model evolution. In green: Confidence intervals at 95%.

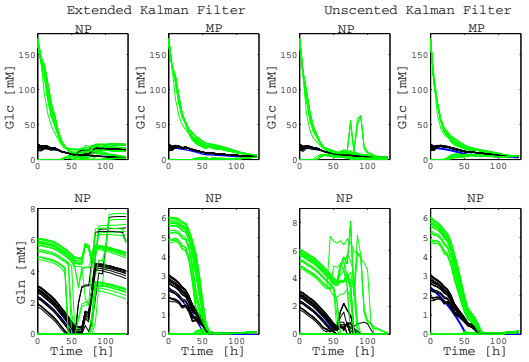

Fig. 4. Glucose and glutamine estimation with EKF and UKF using nominal and modified parameters values (20 runs by varying initial conditions - black curves) for a fed-batch culture with exponential feeding. In blue: model evolution. In green: Confidence intervals at 95%.

Table 2. RMSE of glucose and glutamine obtained for UKF and EKF using nominal and modified parameters.

	RMSE Glucose		RMSE Glutamine	
Continuous culture				
	Nominal values	Modified values	Nominal values	Modified values
EKF	0.9058	1.3631	0.2204	0.2180
UKF	0.6695	1.3816	0.2012	0.1990
Fed-Batch culture with constant feeding				
EKF	0.8493	1.5107	0.2072	0.1930
UKF	0.8389	2.0763	0.2070	0.2072
Fed-Batch culture with exponential feeding				
EKF	5.3263	1.8843	2.0069	0.2384
UKF	1.3832	2.4647	0.3701	0.2873

In figure 1, biomass, lactate and ammonia measurements simulated with nominal parameter values, for the three tests are shown. Glucose and glutamine estimations for the continuous, fed-batch with constant and exponential feedings cultures are shown in figures 2, 3 and 4, respectively. Whatever the type of feeding using the set of nominal parameters (NP), the UKF exhibits better results in terms of convergence of the unmeasured state estimates towards the true values as can be seen from the RMSE values which are (in the average) about 32 % smaller in the case of an UKF in comparison of the use of the EKF in the same conditions. This can of course be explained on the basis

that the UKF corresponds to a true nonlinear solution on the contrary to the EKF which is based on a model linearization.

From figure 4 it can be seen that the use of the NP can lead to divergent state estimates in the case of an exponential feeding. This phenomenon is observed when the sensitivity of the unmeasured states with respect to the measured ones becomes very low (which happens here when the glutamine is almost depleted) and, simultaneously, the corresponding state estimates are still far from the true values. This only happens with the exponential feeding profile because it is applied much later (after 52h) than the constant feeding (35h in both continuous and fed-batch cases) and with the estimates which correspond to the highest overestimated initial values of the glutamine concentration. As a consequence, the state estimates enter the region of low sensitivity much faster while they are still far from converging to the true values given the high initialization error. Even though, from table 2 it can be seen that the RMSE values are (in the average) 78 % smaller when using the UKF (NP) instead of the EKF (NP). This phenomenon is not observed with the modified parameters as they lead to a significantly higher state estimation sensitivity.

7. CONCLUSION

In this study, EKF anf UKF algorithms are applied in order to estimate glucose and glutamine from biomass, lactate and ammonia measurements. Due to the poor glutamine estimation, a set of modified parameters (the ones obtained by a cost function combining the usual least-squares criterion with a state estimation sensitivity criterion) is used and compared with a set of nominal parameters identified by a classical least-squares method.

From the presented results two main conclusions can be deduced. The first one is that the UKF achieves a better level of accuracy than the EKF when nominal parameters are used. The second one is that the state estimations based on the set of modified parameters is notably useful when the sensitivity of the unmeasured states with respect to the measured ones becomes very low, which is the case of the fed-batch culture with exponential feeding.

ACKNOWLEDGEMENTS

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 Science Policy Office. This work has been supported by FEDER (European Union and Wallonia Region in the Hainaut-Biomed program (OCPAM project).

REFERENCES

Amribt, Z., Dewasme, L., Wouwer, A.V., and Bogaerts, P. (2014). Parameter Identification for State Estimation: Design of an Extended Kalman Filter for Hybridoma Cell Fed-Batch Cultures. In *Proceedings of the 19th IFAC World congress*.

Amribt, Z., Niu, H., and Bogaerts, P. (2013). Macroscopic modelling of overflow metabolism and model based optimization of hybridoma cell fed-batch cultures. *Biochemical engineering journal*, 70, 196–209.

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

Dewasme, L., Goffaux, G., Hantson, A.L., and Wouwer, A.V. (2013). Experimental validation of an Extended Kalman Filter estimating acetate concentration in E. coli cultures. *Journal of Process Control*, 23(2), 148–157.

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

Hitzmann, B., Broxtermann, O., Cha, Y.L., Sobieh, O., Stärk, E., and Scheper, T. (2000). The control of glucose concentration during yeast fed-batch cultivation using a fast measurement complemented by an extended Kalman filter. *Bioprocess Engineering*, 23(4), 337–341.

Julier, S.J. and Uhlmann, J.K. (1997). New extension of the Kalman filter to nonlinear systems. In *AeroSense'97*, 182–193. International Society for Optics and Photonics.

Julier, S.J. and Uhlmann, J.K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3), 401–422.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1), 35–45.

Kandepu, R., Foss, B., and Imsland, L. (2008). Applying the unscented Kalman filter for nonlinear state estimation. *Journal of Process Control*, 18(7), 753–768.

Rodrigues, M.E., Costa, A.R., Henriques, M., Azeredo, J., and Oliveira, R. (2010). Technological progresses in monoclonal antibody production system. *Biotechnology Progess*, 26, 332–351.

Simon, D. (2006). *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons.

Veloso, A.C., Rocha, I., and Ferreira, E. (2009). Monitoring of fed-batch E. coli fermentations with software sensors. *Bioprocess and biosystems engineering*, 32(3), 381–388.

Wan, E.A. and Van Der Merwe, R. (2000). The unscented Kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, 153–158. IEEE.

Wurm, F.M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology*, 22, 1393–1397.

Zeitz, M. (1984). Observability canonical (phase-variable) form for non-linear time-variable systems. *International journal of systems science*, 15(9), 949–958.

Zhu, X. and Feng, E. (2012). Joint estimation in batch culture by using unscented kalman filter. *Biotechnology and Bioprocess Engineering*, 17(6), 1238–1243.