ASIA-PACIFIC JOURNAL OF CHEMICAL ENGINEERING
Asia-Pac. J. Chem. Eng. 2014; **9**: 81–92
Published online 18 June 2013 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/apj.1748

Research article

Behavior comparison for biomass observers in batch processes

Adriana Amicarelli, Olga Quintero and Fernando di Sciascio*

Universidad Nacional de San Juan-Instituto de Automática (INAUT), Argentina

Received 28 August 2012; Revised 24 April 2013; Accepted 02 May 2013

ABSTRACT: On-line estimation of biomass concentration in batch biotechnological processes is an active area of research because normally, the biomass is the desired process product output, and also because it is necessary for control purposes to replace the unavailable biomass concentration measurements with reliable and robust on-line estimations. This work presents five different alternatives to face the problem of biomass estimation in a particular batch bioprocess (δ-endotoxins production of *Bacillus thuringiensis*), namely: a phenomenological estimator based on dissolved oxygen balance, an extended Kalman filter estimator, a Gaussian process regression-based estimator, an artificial neural networks-based estimator, and finally, an estimator based on information fusion by a decentralized Kalman filter. Each proposed biomass estimation method has its own advantages and drawbacks according to their ability to take into account the model uncertainties and the measurement errors. First, the design techniques of these five biomass estimators are exposed, and finally, the behavior of each estimation method is compared. The availability of efficient biomass estimators is of great importance for engineers because, on the one hand, it allows developing new control strategies for other bioprocess variables such as for instance: the growth rate of the microorganism, the dissolved oxygen concentration, and so on. On the other hand, it is also important to improve the performance of the bioprocess optimization procedure. This work also aims to show the evolution on biomass estimation techniques from classical to more contemporary approaches, such as the design based on neural networks and Gaussian processes regression. © 2013 Curtin University of Technology and John Wiley & Sons, Ltd.

KEYWORDS: state observers; bioprocess model; biomass estimation; batch processes; Gaussian process

INTRODUCTION

Biomass concentration in a biotechnological process is one of the states that characterize a bioprocess. Moreover, it is generally the main direct or indirectly desired product output. It is well known that the biomass concentration is not normally measured because this measurement is not possible or is economically unprofitable. Therefore, for control purposes, it is necessary to replace the unavailable biomass concentration measurements with reliable and robust on-line estimations. To this aim, several states observers can be found in the literature. A review of commonly used techniques can be found in [1,2] and references therein. Observers can be coarsely divided into two broad classes: first, principles or phenomenological estimators and empirical estimators. The phenomenological estimators can be also subdivided into classical observers and asymptotic observers. Classical observers include extended Kalman filter (EKF), extended Luenberger observer, high-gain observer, nonlinear observers, and full horizon observer. In this class of estimators, a detailed knowledge of the reaction kinetics and associated transport phenomena are required to represent the balance equations. Modeling the biological kinetics reactions is a difficult and time-consuming task, and therefore, the model used by the estimators could differ significantly from reality. This is the main disadvantage of these phenomenological estimators, i.e. their efficiency strongly relies on the model quality. Empirical estimators are based on constructing appropriate nonlinear models of biotechnological processes exclusively from the process input–output data without considering the functional or phenomenological relations between the bioprocess variables.

For the machine learning community, the data-based modeling of the biomass concentration from a finite number of noisy samples (the training dataset) is a supervised learning problem. From this area, in recent years, the artificial neural network (ANN) methodology has become one of the most important techniques applied to biomass estimation, e.g. [3–5] and references therein. Neal's work on Bayesian learning for neural networks [6] shows that many Bayesian regression

**Correspondence to*: Adriana Amicarelli, Universidad Nacional de San Juan- Facultad de Ingeniería - Instituto de Automática. E-mail: amicarelli@inaut.unsj.edu.ar

models based on neural networks converge to a class of probability distributions known as Gaussian processes according as the number of hidden neurons tends to infinity. This fact motivates the idea of replacing parameterized neural networks and work directly with Gaussian process models for the high-dimensional applications to which neural networks are typically applied.[7]

This paper addresses the problem of the biomass estimation in a batch biotechnological process, the *Bacillus thuringiensis* (*Bt*) δ-endotoxins production process, and presents different alternatives that can be successfully used in this sense. The development of the paper includes the design of five biomass concentration estimators. The first two are phenomenological estimators, namely, a phenomenological estimator based on dissolved oxygen (DO) balance, and an EKF estimator. The next two are empirical estimators: a Gaussian process regression-based estimator and an ANN-based estimator. The fifth and last one is an estimator based on the information fusion of two previously available estimators through a decentralized Kalman filter. This paper focuses on the most important aspect of the design of these estimators and their consequences on the basis of the previous work [8] where the tools used were the most important. Finally, the last two sections contain an analysis of results and discussion, and some final conclusions.

BACILLUS THURINGIENSIS δ-ENDOTOXINS PRODUCTION PROCESS

The bioprocess

Bacillus thuringiensis is an aerobic spore-former bacterium, which, during the sporulation, produces insecticidal crystal proteins known as δ-endotoxins. It has two stages on its life span: a first stage characterized by its vegetative growth and a second stage named sporulation phase. When the vegetative growth finalizes, the beginning of the sporulation phase is induced when the mean exhaustion point has been reached. Normally, the sporulation is accompanied by the δ-endotoxin synthesis. After the sporulation, the process is completed with the cellular wall rupture (cellular lysis), and the consequent liberation of spores and crystals to the culture medium.[9–11] The microorganisms used in this work were *Bt* serovar kurstaki strain 172–0451 isolated in Colombia and stored in the culture collection of Biotechnology and Biological Control Unit (CIB).[12] Growth experiments of the fermentation process were performed in a reactor with a nominal volume of 20 L (Fig. 1). The fermentations were developed with an effective volume of 11 L of cultivation medium, and they were inoculated to 10% (v/v) with the microorganism *Bt* culture. The pH medium was adjusted to around 7.0 with potassium hydroxide before its heat sterilization. If the

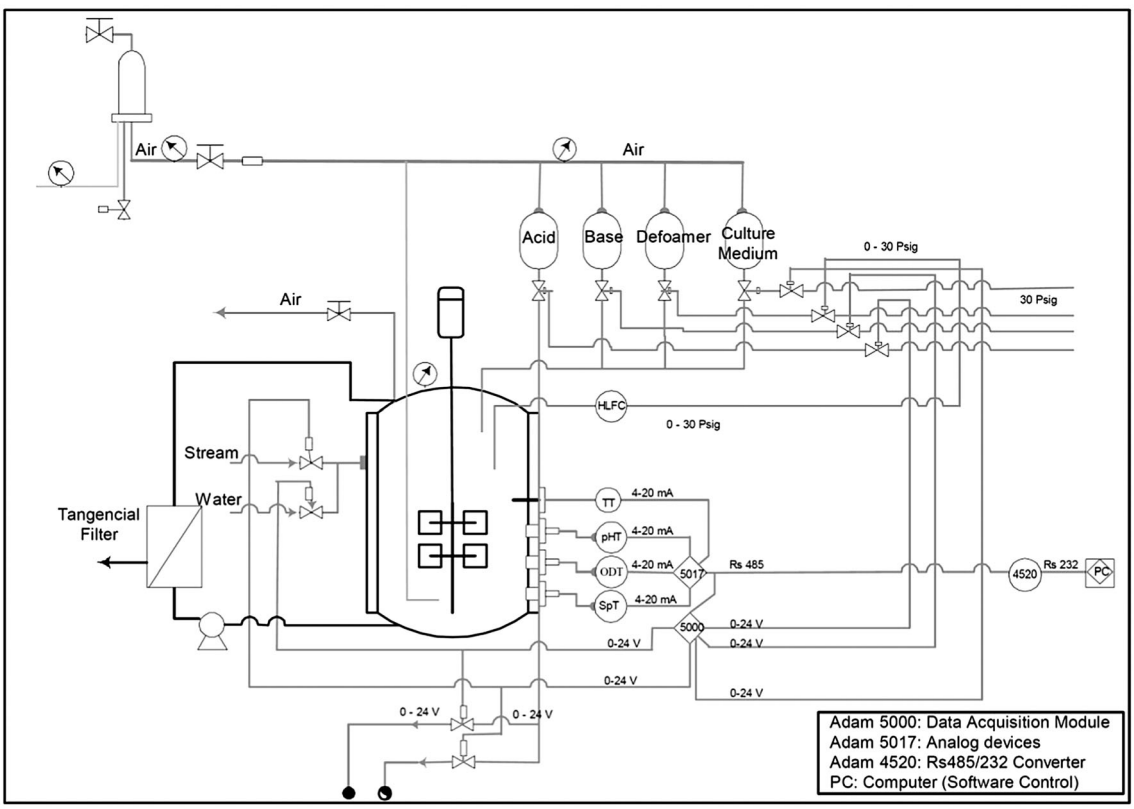

Figure 1. Fermentation pilot plant scheme.

DOI: 10.1002/apj

pH controller has a malfunction, the pH values do not deviate too much from the suitable values for the growth of *Bt* (5.5–8.5 as reported by [13]). An equation for the growth kinetics is used for cases where the pH has a significant effect on cell growth; however, such is not the case for this work.

Bacillus thuringiensis δ-endotoxins production is an aerobic operation, i.e. the cells require oxygen as a substrate to achieve cell growth and product formation.[14]

The temperature was maintained around 30 C° by using an ON/OFF controller, whereas the pH was fixed between 6.5 and 8.5. The airflow was set up at 1320 L/h and the agitation speed at 400 rpm. Manometric pressure in the reactor was set at 41,368 Pa by using a pressure controller. Temperature, pH, DO, and glucose concentration were registered by a data acquisition system using an Advantech® PCL card. DO was measured by a polarographic oxygen sensor InPro 6000 (Mettler Toledo, Switzerland), and glucose concentration with a rapid off-line measurement method through a glucose analyzer (YSI 2700).

Bioprocess model

A simple phenomenological model was proposed by Rivera *et al.*,[15] a modification to the Rivera model was given by Atehortúa *et al.*[16] Afterward, Amicarelli *et al.* [5,17] improved the model process by adding the DO dynamics due to its importance in the biomass estimation problem and the posterior process control. The DO dynamic model is based on unstructured and unsegregated descriptions of the cell population. The microorganism growth is affected by DO concentration (below the 10% critical value). In this work, the DO is considered in excess by an adequate airflow and agitation speed. In a previous work,[17] a DO controller was presented and given an experimentally validated temporal profile of the DO during the fermentation; the controller can maintain the DO level at an optimum level for this process. For the fermentations where the DO concentration is below its critical value, the process with the controller presents an improved behavior.

The following state-space model is the discrete time version of the continuous-time counterpart developed by Amicarelli *et al.*[17]

$$\begin{bmatrix} X_v(k+1) \\ X_s(k+1) \\ S(k+1) \\ DO(k+1) \end{bmatrix} = \begin{bmatrix} ((\mu(k)\text{-}k_s(k)\text{-}k_e(k))\,Ts + 1)\,X_v(k) \\ k_s(k)\,X_v(k)\,Ts + X_s(k) \\ -\left(\dfrac{\mu(k)}{Y_{x/s}} + m_s\right) X_v(k)\,Ts + S(k) \\ (K_1\text{-}K_2Ts)\,X(k)\text{-}K_1\,X(k+1) + DO(k) + K_3\,Q_A\,Ts(DO^*\text{-}DO(k)) \end{bmatrix} \quad (1)$$

where X_v is the vegetative cell concentration, X_S the sporulated cell concentration, $X = X_v + X_S$ is the total cell concentration ($X(k+1) = (\mu(k) - k_e(k))TsX_v(k) + X(k)$), S is the limiting substrate concentration, and DO is the DO concentration.

The following algebraic equations define the specific growth speed μ (model based on Monod equation for each limiting nutrient S and DO), the spore formation rate k_S, and the death cell specific rate k_e.

$$\mu(k) = \mu_{max}\left(\frac{S(k)}{(K_s + S(k))}\frac{DO(k)}{(K_O + DO(k))}\right) \quad (2)$$

$$k_s(k) = k_{smax}\left(\frac{1}{1 + e^{Gs(S(k)\text{-}Ps)}}\right) \text{-}k_{smax}\left(\frac{1}{1 + e^{Gs(S_{initial}\text{-}Ps)}}\right) \quad (3)$$

$$k_e(k) = k_{emax}\left(\frac{1}{1 + e^{Ge(Ts\,k\text{-}Pe)}}\right) \text{-}k_{emax}\left(\frac{1}{1 + e^{Ge(t_{initial}\text{-}Pe)}}\right) \quad (4)$$

The complete notation and model parameter's values are presented in the Nomenclature and in Table 1.

Four batch cultures with different initial glucose concentration (8, 21, 32, and 40 $g.L^{-1}$) and initial vegetative cell concentration (0.424 $g.L^{-1}$ for the first three substrate conditions and 0.605 $g.L^{-1}$ for the last one) were carried out to generate experimental data for model validation and parameters tuning. The initial sporulated cell concentration is 0 $g.L^{-1}$ in all fermentations. In this context, four parameter sets guarantee a representative covering of an intermittent fed-batch culture with total cell retention in the operation space according to the work of Atehortúa *et al.*,[16] see Table 1.

Maximum glucose concentration in the medium (S_{max}) was used as the switching criteria among the estimated batch parameter sets.

BIOMASS ESTIMATORS DESIGN

The duration of the batch fermentation is limited and depends on the initial conditions of the microorganism culture. All the fermentations used in this work were initialized with an inoculum and different substrate concentration conditions.[16] When the medium is inoculated, the biomass concentration increases at the

DOI: 10.1002/apj

Table 1. Model parameters for the intermittent fed-batch culture with total cell retention of *Bacillus thuringiensis serovar. Kurstaki*.

	$S\max < 10\ g.L^{-1}$	$10\ g.L^{-1} < S\max < 20\ g.L^{-1}$	$20\ g.L^{-1} < S\max < 32\ g.L^{-1}$	$S\max > 32\ g.L^{-1}$
$\mu_{max}[h^{-1}]$	0.8	0.7	0.65	0.58
$Yx/s[g.g^{-1}]$	0.7	0.58	0.37	0.5
$K_s[g.L^{-1}]$	0.5	2	3	4
$K_o[g.L^{-1}]$	1×10^{-4}	1×10^{-4}	1×10^{-4}	1×10^{-4}
ms[g/g/h]	5×10^{-3}	5×10^{-3}	5×10^{-3}	5×10^{-3}
$ksmax$	0.5	0.5	0.5	0.5
$Gs[g.L^{-1}]^{-1}$	1	1	1	1
$Ps[g.L^{-1}]$	1	1	1	1
$kemax[h^{-1}]$	0.1	0.1	0.1	0.1
$Ge[h^{-1}]$	5	5	5	5
$Pe[h]$	4	4.7	4.9	6
K_1 dimensionless	9.725×10^{-4}	4.502×10^{-3}	3.795×10^{-3}	1.597×10^{-3}
$K_2[h^{-1}]$	1.589×10^{-4}	0.046×10^{-3}	0.729×10^{-3}	0.561×10^{-3}
$K_3[L^{-1}]$	4.636×10^{-4}	0.337×10^{-3}	2.114×10^{-3}	1.045×10^{-3}
$Ts[h]$	0.1	0.1	0.1	0.1

expense of the nutrients, and the fermentation concludes when the glucose that limits its growth is consumed or when 90% or more of cellular lysis is presented. Note that the latency period is removed (the bioprocess dead time is not considered here) and note also that the duration of each experiment is approximately 16 h.

The collected data from the fermentations are a set of concentration measurements of DO, primary substrate (S), and biomass (X), which have been sampled at different frequencies: 10 samples per hour for the concentrations of DO and glucose and 1 per hour for the biomass concentration that was quantified by cell dry weight method. Practically, DO can be continuously measured, whereas S can be measured up to 20 times per hour. From the bandwidth estimation of system signals by using Fourier frequency analysis, the sampling time $Ts = 1/10$ *hours* has been selected for DO and substrate measurements.[18,19] To design biomass estimators for the *Bt* δ-endotoxins production process, it is proposed a two-stage method.[18] In the first stage, the biomass concentrations dataset is completed to have the same size as the DO concentration and primary substrate (glucose) concentration data sets. For this missing data problem, it was considered a Bayesian Gaussian process regression as an imputation strategy for filling the missing values.[18] In the second stage, different biomass estimators are designed.

First stage design for all estimators – filling the biomass missing data

Suppose that we have a noisy training data set D, which consists of m pairs of n-dimensional input vectors $\{x_i\}$ (regression vector) joined in an $n \times m$ matrix $\boldsymbol{X}$, and m scalar noisy observed outputs $\{y_i\}$ collected in a vector $\boldsymbol{y}$.

$$D = \{(x_i, y_i)|i = 1, L, m\} = \{X, y\} \tag{5}$$

To construct a probabilistic statistical model for D, the following data-generating process is assumed:

$$y_i = f(x_i) + \varepsilon_i \tag{6}$$

where the latent real-valued function f is the deterministic or systematic component of the model, and the additive random term ε is the observation error. The aim of regression is to identify the systematic component f from the empirical observations D.

In this section, the biomass concentration data vector is completed with virtual filtered measurements to have the same size as DO and substrate data vectors. This is a missing data problem, and the Gaussian process regression will be used as imputation method for filling the missing values (note that this task in a deterministic framework, which can be viewed as a curve-fitting or interpolation problem).

For all experimental fermentations, the data-generating model for biomass concentration is:

$$X(tk) = \hat{X}(tk) + \varepsilon(tk) \tag{7}$$

The training data set D consists of 18 pairs of time inputs $t = \{tk\} = \{1, \ldots, 18\}$ (in hours), and noisy biomass measurements outputs $X = \{X_k\} = \{X(t_1), \ldots, X(t_{18})\}$. The latent functions $\hat{X} = \left\{\hat{X}_k\right\} = \left\{\hat{X}(t_1), \ldots, \hat{X}(t_{18})\right\}$ are the estimated biomass concentrations.

The expression 'Gaussian process regression model' refers to the use of a Gaussian process as a prior on **f**.

DOI: 10.1002/apj

This means that every finite-dimensional marginal joint distributions of function values **f** associated to any input subset of **X** is multivariate Gaussian.

$$p(f|X,\theta_P) = N(m(X), K(X,\theta_P)) \tag{8}$$

A Gaussian process is fully specified by a mean function $m(X) = [m(x_1), L, m(x_m)]^T$ and a positive definite covariance matrix $K(X, \theta_P)$, and it can be viewed as a generalization of the multivariate Gaussian distribution to infinite dimensional objects. Choosing a particular form of covariance function, the hyperparameters θ_P may be introduced to the Gaussian process prior. Depending on the actual form of the covariance function $K(X, \theta_P)$, the hyperparameters θ_P can control various aspects of the Gaussian process.

In this work, the elements of the parameterized covariance matrix, $C(X, \theta_P, \sigma^2)$, are denoted $C_{ij} = C(x_i, x_j)$, and they are functions of the training input data *X*, because these data determine the correlation between the training data outputs *y*. A suitable parametric form of the covariance function is:[18]

$$Cij = \theta_0 + \theta_1 \exp\left[-\frac{1}{2}\sum_{l=1}^{n}\frac{\left(x_i^{(l)}-x_j^{(l)}\right)^2}{r_l^2}\right] + \theta_2\,\delta(i,j) + \sum_{l=1}^{n}\alpha_l\, x_i^{(l)}\, x_j^{(l)} \tag{9}$$

where $x_i^{(l)}$ is the l^{th} dimension of the input vector, x_i.

From the training data *D*, and by means of a conjugate gradient routine $\#\theta = 5$ hyperparameters, and the matrix *C* are determined recursively through:

$$\log\theta = [\log\theta_0, \log\theta_1, \log r_1, .., \log r_n, \log\theta_2, \log\alpha_1, .., \log\alpha_n]^T \tag{10}$$

and

$$\begin{cases} L = -\frac{1}{2}\log|C| - \frac{1}{2}y^T C^{-1} y - \frac{m}{2}\log 2\pi + \log p(\theta) + c \\ \frac{\partial L}{\partial\theta_i} = -\frac{1}{2}\,\mathrm{trace}\left(C^{-1}\frac{\partial C}{\partial\theta_i}\right) + \frac{1}{2}y^T C^{-1}\frac{\partial C}{\partial\theta_i}C^{-1}y + \frac{\partial \log p(\theta)}{\partial\theta_i} \end{cases} \tag{11}$$

Afterward, at different times, $t_* = 0.1, 0.2, \ldots, 17.9, 18$, by (12)

$$\begin{aligned} \hat{f}_* &= E(f_*|D, x_*, K, \sigma^2) = k_*{}^T C^{-1} y \\ \sigma^2_{\hat{f}_*} &= k_{**} - k_*{}^T C^{-1} k_* \end{aligned} \tag{12}$$

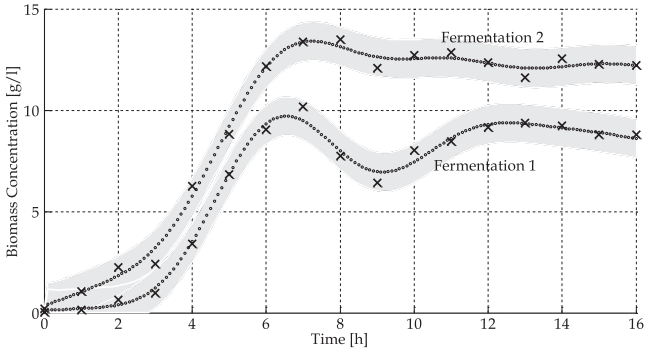

Figure 2. Example of completion of biomass missing data for Fermentations 1 and 2. The crosses being the biomass concentration measurements (training data), the small circles represent the biomass estimated (virtual filtered biomass measurements), and the gray region depicts the 95% confidence interval for the estimations (±2 standard deviations) (from [17]).

the latent functions $\hat{X}_* = \{\hat{X}_*\} = \{\hat{X}(t_*)\}$ and the variance $\sigma^2_{\hat{X}_*}$ are estimated. The expression 'virtual filtered measurements' refers to the latent functions $\hat{X}_*$, because the additive normal noise ε has been removed (filtered) from the 'virtual measurement' X_* in the data-generating model (7). Figure 2 gives an example of completion of biomass missing data for two fermentations (Fermentation 1 and Fermentation 2).

Second stage design – implementation of estimators

In this second stage of biomass estimators design, five different estimators are implemented. The first two are phenomenological estimators, namely, a phenomenological estimator based on DO balance and an EKF standard estimator. The next two are empirical estimators: a Gaussian process regression-based estimator and an ANN-based estimator. The fifth and last one is an estimator based on the information fusion of two (or more) previously available estimators through a decentralized Kalman filter.

Phenomenological biomass estimator based on dissolved oxygen balance

To design a phenomenological biomass estimator, consider the discrete time state-space model described previously in Section Bioprocess model. The fourth equation of the bioprocess model (1) explains the dynamic of the DO balance, which is repeated here for convenience.

$$DO(k+1) = (K_1 - K_2 Ts)\,X(k) - K_1\,X(k+1) + DO(k) + K_3\,Q_A\,Ts(DO^* - DO(k)) \tag{13}$$

DOI: 10.1002/apj

Rewriting this difference equation, and replacing the unknown biomass X by their estimates $\hat{X}$, we obtain our proposed phenomenological biomass estimator:

$$\hat{X}(k) = \frac{1}{K_1}(K_1\text{-}K_2 Ts)\,\hat{X}(k\text{-}1) - \frac{1}{K_1}DO(k) + \frac{1}{K_1}[1\text{-}K_3 Q_A Ts]\,DO(k\text{-}1) + \frac{K_3 Q_A Ts}{K_1}DO^* \quad (14)$$

Note that (14) is a linearly parameterized estimator. If θ denotes the known parameter vector and $\varphi(k)$ describes the regressor vector of measured (or known) signals, then the phenomenological biomass estimator (14) can be written compactly as:

$$\hat{X}(k) = \theta^T \varphi(k)$$

where $\theta = \left[\frac{K_1\text{-}K_2 Ts}{K_1} \quad -\frac{1}{K_1} \quad \frac{1\text{-}K_3 Q_A Ts}{K_1} \quad \frac{K_3 Q_A Ts\, DO^*}{K_1}\right]^T$

and $\varphi(k) = \left[\hat{X}(k\text{-}1) \quad DO(k) \quad DO(k\text{-}1) \quad 1\right]^T$.

From (14), it can be inferred that the total biomass concentration can be estimated on-line with experimental data of DO concentration $(DO(k), k = 0, 1, \cdots)$ and with the biomass initial values $X(0)$ (obtained by dry weight method). Figure 3 shows the model structure for the phenomenological biomass estimator.

Figure 4 shows the phenomenological estimation results. This observer can approximate the biomass concentration better than the first model proposed by Atehortúa *et al.*[16] This is because, this estimator includes the DO consumption for growth and maintenance of the microorganism on its structure and through the experimental data of DO available on-line (Fig. 3). Figure 5 and Fig. 6 show the time evolution of DO percentages and S concentrations for both fermentations respectively. Moreover, Figure 4 shows satisfactory results and a correct behavior of the phenomenological estimator for two different fermentations. It is important to remark that the estimator involves in its structure the original model of vegetative and sporulated cells, whereas the consideration of the DO influence on the microorganism concentration improves the biomass estimation performance. When the DO influence is not significant, the biomass estimation achieved with the model without the DO dynamics and the phenomenological estimators are comparable (see Fermentation 1 in Fig. 4.). However, for those cases in which the DO approaches critical values (see Fermentation 2 in Fig. 4), the phenomenological observer gives better estimations (Fermentation 2).

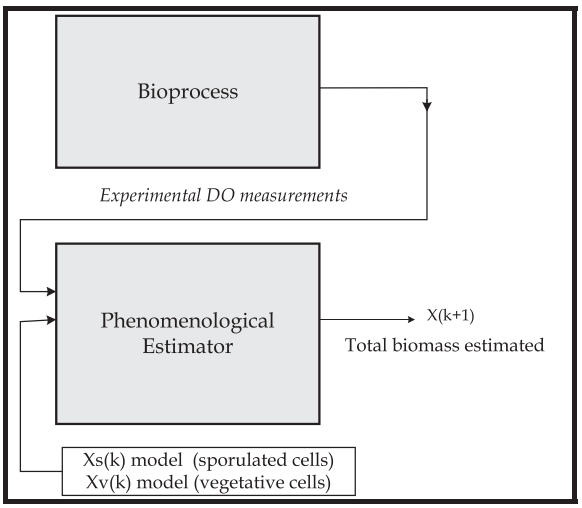

Figure 3. Simulated output model structure for the phenomenological biomass estimator.

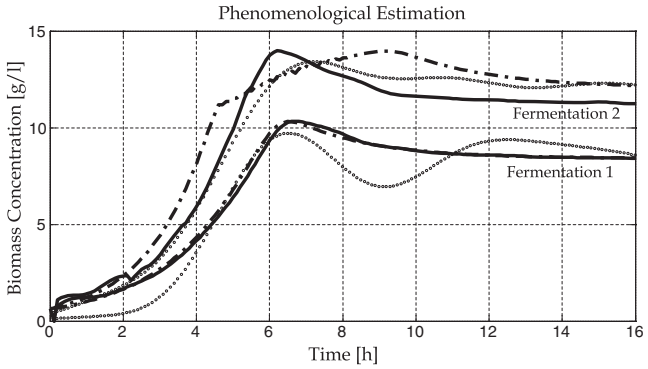

Figure 4. Biomass estimator performance. The dash-dot line describes the behavior of biomass when considering the model (1), the solid line depicts the phenomenological estimator behavior based on DO dynamics, and the real biomass measurements are represented by small circles.

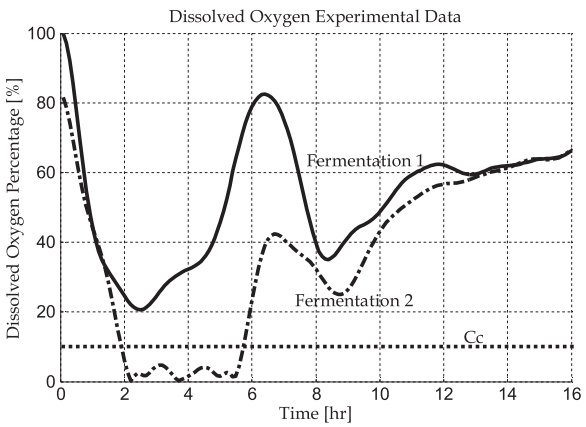

Figure 5. Dissolved oxygen experimental data. The solid line describes the dissolved oxygen behavior for Fermentation 1, the dash-dot line depicts the dissolved oxygen behavior for Fermentation 2, and the dotted line corresponds to the percentage of dissolved oxygen for the critical dissolved oxygen concentration for this process.

DOI: 10.1002/apj

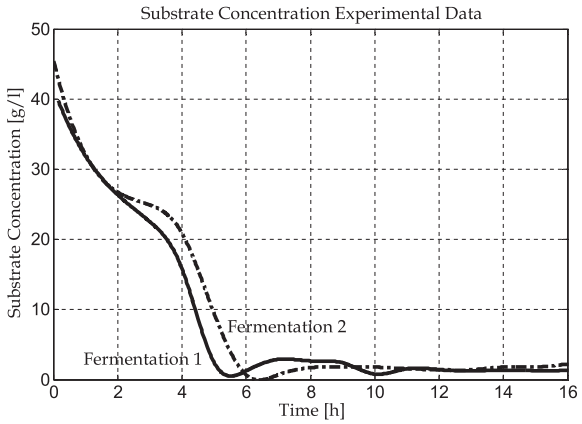

Figure 6. Substrate concentration experimental data. The solid line describes the substrate concentration for Fermentation 1 and the dash-dot line depicts the substrate concentration for Fermentation 2.

Extended Kalman filter standard estimator

The basic or standard EKF is the most commonly used state estimator for nonlinear systems. The quality of the estimation achieved by an EKF depends on the process model accuracy as well as on the available noisy measurements. There are numerous improvements to the standard EKF scheme, for example, by using different coordinate systems for the design, or by using different factorizations of the covariance matrix, or by using second or higher order Taylor series corrections to the state vector prediction, and so on.[20] The underlying theory of the EKF is largely known in the literature devoted to filtering, estimation, and control; see, for example, the classic books by [21]; [22], or most recently, the book by Simon.[23] Therefore, in this work, only brief explanations of the specific EKF implementation for the *Bt* fermentation process are given. In the EKF framework, the state transition and observation models are nonlinear differentiable states functions.

State transition model:

$$\mathrm{x(k+1)} = \mathrm{f(x(k), u(k), k)} + \mathrm{w(k)} \qquad (15)$$

Measurements model:

$$\mathrm{y(k)} = \mathrm{h(x(k), k)} + \mathrm{v(k)} \qquad (16)$$

where $\mathrm{f}(\times,\times)$ is the state transition function, $\mathrm{h}(\times,\times)$ is the measurement function, x(k) is the system state vector with initial condition $\mathrm{x(0)} \sim \mathrm{N}(\mathrm{x_0}, \mathrm{Q_0})$ (as is usual in statistical literature the symbol (~) means 'distributed according to'), u(k) is the input or control vector, y(k) is the observation vector, w(k) is a discrete time normal white noise process (process noise) with null mean and covariance matrix Q, i.e. $\mathrm{w(k)} \sim N(0,Q)$, and v(k) is a discrete time normal white noise process (measurements noise) with null mean and covariance matrix R, i.e. $\mathrm{v(k)} \sim N(0,R)$. The initial condition x(0) and the sequences w(k) and $v(k)$ are uncorrelated for all time shifts.

In our case, the nominal state transition model (without the process noise $w(k)$) is obtained by introducing (2), (3), and (4) in (1).

$$x(k+1) = f(x(k), k) \qquad (17)$$

The system state vector is $\mathrm{x(k)} = [\mathrm{X_V(k)} \quad \mathrm{X_S(k)} \quad \mathrm{S(k)} \quad \mathrm{DO(k)}]^{\mathrm{T}}$, the input vector is $u(k) = 0$ (the bioprocess has no external input), and the bioprocess outputs (observation vector) is $\mathrm{y(k)} = [\mathrm{S(k)} \quad \mathrm{DO(k)}]^{\mathrm{T}}$ (Fig. 7). The measurement model is linear in the states:

$$y(k) = Hx(k) \qquad (18)$$

where

$$H = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Taking into account the scales of the outputs, a balanced linear combination of S(k) and DO(k) can be considered as an alternative measurement model.

$$\mathrm{y'(k)} = \mathrm{H'x(k)} = \alpha\,\mathrm{S(k)} + \beta\,\mathrm{DO(k)} \qquad (19)$$

In this measurement model $H' = [0 \quad 0 \quad \alpha \quad \beta]$ where:

$\alpha = \mathrm{DOmax}/(\mathrm{Smax} + \mathrm{DOmax})$ $\beta = \mathrm{Smax}/(\mathrm{Smax} + \mathrm{DOmax})$

The next step is to obtain the Jacobian matrices $\frac{\partial \mathrm{f(x(k),k)}}{\partial \mathrm{x}}$ and $\frac{\partial \mathrm{h(x(k),k)}}{\partial \mathrm{x}}$ evaluated at $\hat{\mathrm{x}}(\mathrm{k\text{-}1}|\mathrm{k\text{-}1})$.

$$\mathrm{A(k)} = \left.\frac{\partial \mathrm{f(x(k),k)}}{\partial \mathrm{x}}\right|_{\hat{\mathrm{x}}(\mathrm{k\text{-}1}|\mathrm{k\text{-}1})}$$

$$\mathrm{A(k)} = \begin{bmatrix} \left.\frac{\partial \mathrm{f}_1\mathrm{(x(k),k)}}{\partial \mathrm{x}_1}\right|_{\hat{\mathrm{x}}(\mathrm{k\text{-}1}|\mathrm{k\text{-}1})} & \cdots & \left.\frac{\partial \mathrm{f}_1\mathrm{(x(k),k)}}{\partial \mathrm{x}_4}\right|_{\hat{\mathrm{x}}(\mathrm{k\text{-}1}|\mathrm{k\text{-}1})} \\ \vdots & \vdots & \vdots \\ \left.\frac{\partial \mathrm{f}_4\mathrm{(x(k),k)}}{\partial \mathrm{x}_1}\right|_{\hat{\mathrm{x}}(\mathrm{k\text{-}1}|\mathrm{k\text{-}1})} & \cdots & \left.\frac{\partial \mathrm{f}_4\mathrm{(x(k),k)}}{\partial \mathrm{x}_4}\right|_{\hat{\mathrm{x}}(\mathrm{k\text{-}1}|\mathrm{k\text{-}1})} \end{bmatrix} \qquad (20)$$

$$\mathrm{H(k)} = \left.\frac{\partial \mathrm{h(x(k),k)}}{\partial \mathrm{x}}\right|_{\hat{\mathrm{x}}(\mathrm{k}|\mathrm{k\text{-}1})} = \left.\frac{\partial \mathrm{Hx(k))}}{\partial \mathrm{x}}\right|_{\hat{\mathrm{x}}(\mathrm{k}|\mathrm{k\text{-}1})} = \mathrm{H} \qquad (21)$$

Finally, initializing the elements of the matrices $\boldsymbol{P}$, Q, and $\boldsymbol{R}$, we have all the components of the EKF algorithm. To obtain the best possible fit of the EKF

DOI: 10.1002/apj

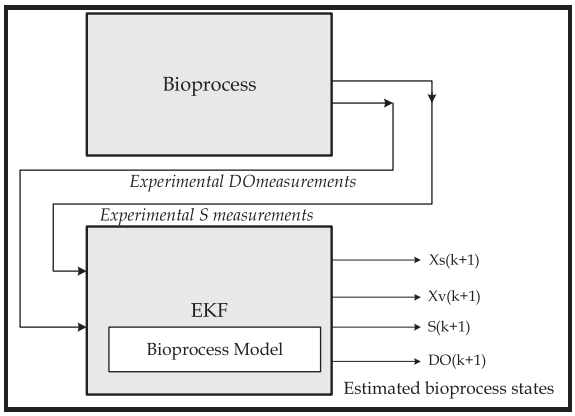

Figure 7. Simulated output model structure for EKF biomass estimator.

to the experimental data, the elements of the matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ are empirically adjusted by simulations.

The experimental DO percentages and substrate concentration data employed are shown in Figs. 5 and 6. Figure 8 shows results for two different fermentations. It is performed a comparison between this estimator and the phenomenological observer based on DO dynamics previously presented. It can be concluded that the performance of the standard EKF estimator is adequate. This of course does not mean that the performance of the EKF cannot be meaningfully enhanced by using a better model of the bioprocess or by some of the improvements pointed out at the beginning of this section.

Estimator based on Bayesian Gaussian process regression

The first step in the design is to select the regressor variables, i.e. the components of the input vector x. This is a laborious task, and has been performed heuristically, chosen from numerous alternatives. The best empirical results have been achieved with:

$$\mathrm{x(kTS)} = \left[\mathrm{DO(kTs)}, \mathrm{S(kTs)}, \hat{\mathrm{X}}\mathrm{((k\text{-}1)Ts)},\ \hat{\mathrm{X}}\mathrm{((k\text{-}2)Ts)}\right]^{\mathrm{T}} \quad (22)$$

where $\mathrm{k} = \{1,L,180\}$ is the time index, $\mathrm{Ts} = 1/10$ *hours* is the sampling time, $DO(.)$ is the DO concentration, $S(.)$ is the substrate concentration, and $\hat{\mathrm{X}}(.)$ is the virtual filtered biomass measurement. In this case, the training data set D consists of 180 pairs of input vectors $\{x(kTs)\} = \{x_k\} \in \mathrm{R}^4$ collected in a matrix $\mathrm{X} \in \mathrm{R}^{4\times 180}$, and scalars outputs $\left\{\hat{\mathrm{X}}(\mathrm{kTs})\right\} = \left\{\hat{\mathrm{X}}_\mathrm{k}\right\}$ collected in a vector $\hat{\mathrm{X}} \in \mathrm{R}^{180}$ (note that in this section, the virtual filtered biomass measurements $\left\{\hat{\mathrm{X}}_\mathrm{k}\right\}$ are considered as true observed measurements).

The data-generating process is $\hat{\mathrm{X}}_\mathrm{k} = \hat{\hat{\mathrm{X}}}_\mathrm{k} + \varepsilon_\mathrm{k}$, being the latent function $\hat{\hat{\mathrm{X}}}_\mathrm{k}(.)$, and the additive normal noise ε. Once again, the $\#\theta = 11$ hyperparameters and the new covariance matrix C Eqn. (9) are determined via a conjugate gradient routine from (11) and:

$$Cij = \theta_0 + \theta_1 \exp\left[-\frac{1}{2}\sum_{\mathrm{l}=1}^{\mathrm{n}} \frac{\left(\mathrm{x}_\mathrm{i}^{(\mathrm{l})} - \mathrm{x}_\mathrm{j}^{(\mathrm{l})}\right)^2}{\mathrm{r}_\mathrm{l}^2}\right] + \theta_2\,\delta(\mathrm{i},\mathrm{j}) + \sum_{l=1}^{\mathrm{n}} \alpha_l\, \mathrm{x}_\mathrm{i}^{(\mathrm{l})}\, \mathrm{x}_\mathrm{j}^{(\mathrm{l})} \quad (23)$$

Furthermore, by (12), the biomass concentration $\hat{\hat{\mathrm{X}}}_* = \hat{\hat{\mathrm{X}}}(t_*)$ and the variance $\sigma_{\hat{\hat{X}}*}^{\ 2}$ are estimated for a set of different times $\{t^*\}, < t^* < 16$ hours.

For the training stage, a one-step ahead predicted output schema is performed, i.e. the input measurements, DO(k)S(k), and the previous output measurements $\hat{\mathrm{X}}(\mathrm{k\text{-}1}), \hat{\mathrm{X}}(\mathrm{k\text{-}2})$ are used as repressors in:

$$\hat{\hat{\mathrm{X}}}(\mathrm{k}) = \hat{\hat{\mathrm{X}}}\left(\mathrm{DO(k)}, \mathrm{S(k)}, \hat{\mathrm{X}}(\mathrm{k\text{-}1}), \hat{\mathrm{X}}(\mathrm{k\text{-}2})\right) \quad (24)$$

For on-line estimation, the implemented estimator is the simulated output schema, i.e. only input measurements DO(k), S(k) are used. The simulated output is obtained as previously, by replacing the measured outputs by the simulated output from the previous steps, i.e. previous outputs from the model have to be fed back into the model computations on-line (Fig. 9).

$$\hat{\hat{\mathrm{X}}}(\mathrm{k}) = \hat{\hat{\mathrm{X}}}\left(\mathrm{DO(k)}, \mathrm{S(k)}, \hat{\hat{\mathrm{X}}}(\mathrm{k\text{-}1}), \hat{\hat{\mathrm{X}}}(\mathrm{k\text{-}2})\right) \quad (25)$$

This one-step ahead predicted output scheme is known as nonlinear auto-regressive with exogenous input model, or as series–parallel model.[24,25] Furthermore,

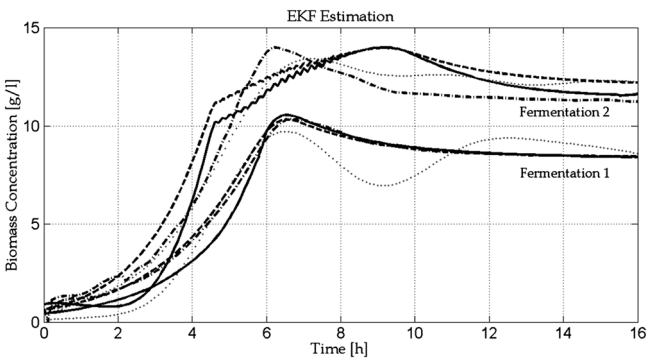

Figure 8. Biomass estimator performance. The dashed line describes the biomass evolution obtained from the original model (1), the solid line depicts the EKF behavior, the dashed-dotted line depicts the phenomenological estimator behavior based on DO dynamics, and the real biomass measurements are represented by small circles.

DOI: 10.1002/apj

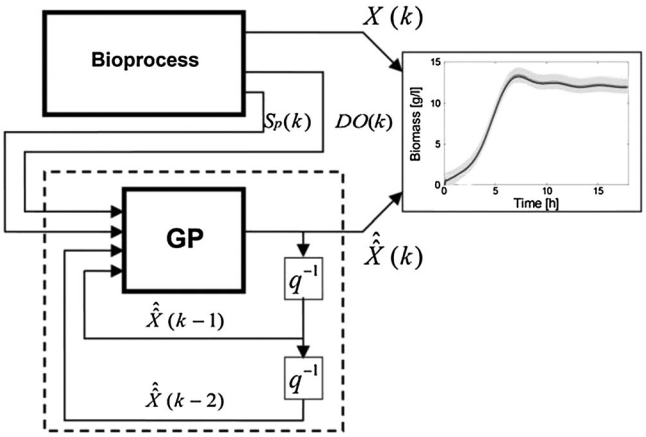

Figure 9. Simulated output model structure of proposed biomass estimator (from [18]).

the simulated output schema is known as nonlinear output error model, or as parallel model.[24,25]

The biomass concentration of Fermentations 1 and 2 from the preceding section (see Fig. 2) has been adopted as training and validation data, respectively. Figures 5 and 6 show the measurements of DO percentages (DO) and glucose concentration (S), respectively. Both signals have been filtered with a low-pass filter with a 1/36 Hz corner frequency.

Figure 10 shows the evolutions of the true biomass measurements, the virtual filtered biomass measurements, and our proposed biomass estimator within its 95% confidence intervals. This figure clearly shows the correct behavior of the biomass estimator based on Gaussian process regression. The estimated biomass follows closely the true and the virtual filtered biomass measurements. This performance is achieved setting the 11 hyperparameters of the covariance function.

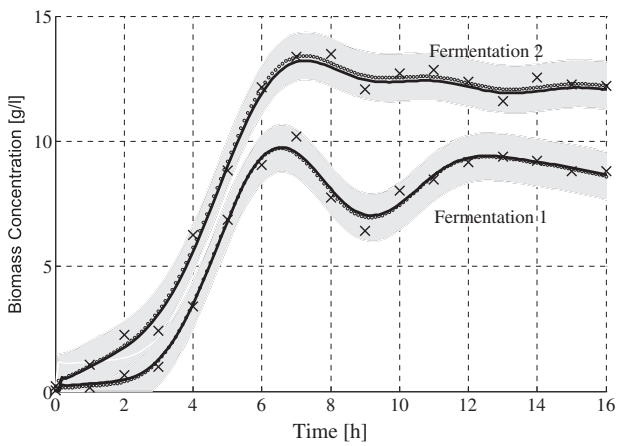

Figure 10. Biomass estimator performance. The bold solid line describes the behavior of the proposed biomass estimator ($\hat{X}(kTs)$), the crosses are the true biomass measurements, the virtual filtered biomass measurements ($\hat{X}(kTs)$) are represented by small circles, and the gray region depicts the 95% confidence interval (from [18]).

Artificial neural networks-based estimator

Through ANN, the empirical knowledge (set of measurements) that characterizes a phenomenon of interest can be adequately codified. Because of the high degree of parallelism, the high generalization capability, and the possibility to use architecture of multiple inputs and outputs, the ANNs can provide a satisfactory solution to the problems of models identification, variables estimation, pattern recognition, and functions approximation among others. ANNs have the ability to abstract automatically essential characteristics of the experimental data, and to generalize from the previous experience; this allows the identification of the model process at lower cost.

Supervision and control techniques require optimizing the fermenter operation and the monitoring of all variables on-line is the best solution, because the methods offline delay the possibility of getting results and generally require more effort.

The ANN employed in this work is a recurrent multilayer perceptron with one hidden layer of 30 neurons and one output in the output layer. The activation functions of the hidden layer were hyperbolic tangent and a linear function for the output layer. The input vector to the ANN is the same as the previous section, i.e. $\left[\mathrm{DO(kTs)}, \mathrm{S(kTs)}, \hat{\mathrm{X}}\mathrm{((k\text{-}1)Ts)}, \hat{\mathrm{X}}\mathrm{((k\text{-}2)Ts)}\right]^T$, and the scalar output is the biomass estimate $\hat{\mathrm{X}}\mathrm{(kTs)}$. This one-step ahead predicted output scheme $\hat{\mathrm{X}}\mathrm{(kTs)} = \mathrm{f}\left(\mathrm{DO(kTs)}, \mathrm{S(kTs)}, \hat{\mathrm{X}}\mathrm{((k\text{-}1)Ts)}, \hat{\mathrm{X}}\mathrm{((k\text{-}2)Ts)}\right)$ is the same nonlinear auto-regressive with exogenous input model or series–parallel model presented in previous section.

For the training stage, the back-propagation algorithm [25] was employed. The network was trained with data from a fermentation identified as 'Fermentation 1' (see Fig. 11) and was generalized with other set of experimental data 'Fermentation 2' (see Fig. 12).

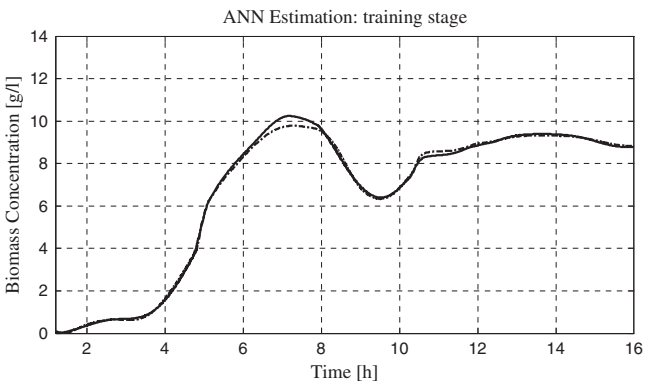

Figure 11. Biomass estimator performance. The dashed line describes the biomass evolution obtained by the ANN in the training stage and the real biomass measurements are represented by the solid line. The perceptual training error e = 0.16%.

DOI: 10.1002/apj

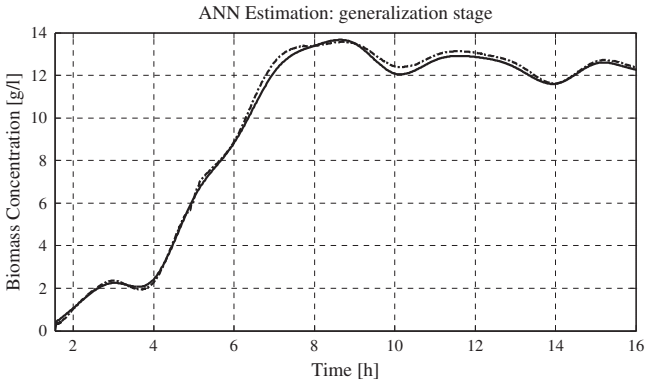

Figure 12. Biomass estimator performance. The dashed line describes the biomass evolution obtained from the ANN in the generalization stage and the real biomass measurements are represented by the solid line. The perceptual generalization error e = 0.25%.

Fusion through decentralized Kalman filter

The aim of this section is to obtain an improved estimate of the biomass value for the process of *Bt* by using classical data fusion techniques. Recall that the data fusion approach to estimation problem combines independent data from multiple estimators to produce an improved estimate that could not be achieved by the use of only a single estimator. For simplicity, in what follows, the data fusion technique is exemplified by considering only two independent sequences of biomass estimates. The first one is provided by the phenomenological estimator based on DO balance designed in Section Phenomenological biomass estimator based on dissolved oxygen balance, and the second one is given by the ANN estimator designed in the preceding section. To refine the accuracy of the biomass estimation, these two sequences of estimates are fused through a decentralized Kalman filter.[26]

Assuming that the estimations are the optimum value for each sequence in time and the relationship between these values is given by:

$$X^i = X^{iOPT} + v^i \tag{26}$$

where v^i is a random variable with zero mean and covariance R^i. In a basic approach of the decentralized Kalman filter, each local filter operates autonomously. Each local filter has its own set of measurements, and there is no sharing of measurements. Note that this is inherently a cascaded operation mode, because the outputs of one or more of the local filters are acting as inputs to the master filter. The local filters (one for each sequence of measurements), the master filter, and the different variables involved can be appreciated in Fig. 13.

The mean and covariance for each sequence of measurements are calculated recursively according to:

$$\hat{X}^i(k+1) = \hat{X}^i(k) + \mu\left(X^i(k) - \hat{X}^i(k)\right) \tag{27}$$

$$R^i = R^i + \mu\left(\left(X^i - \hat{X}^i\right)^2 - R^i\right) \tag{28}$$

where $\hat{X}^i$ is the average sequence value of X^i and $0 < \mu < 1$ is a design constant. Then, each sequence is individually filtered:

$$\left(P^i\right)^{-1} = \left(M^i\right)^{-1} + \left(R^i\right)^{-1} \tag{29}$$

$$X_v^{i.OPT} = P^i\left[m^i\left(M^i\right)^{-1} + \left(R^i\right)^{-1}X_v^i\right] \tag{30}$$

Equation (29) provides the updated information matrix, whereas (30) are the states estimated updates, M^i and m^i are the covariance error and the previous estimation values for the measurements sequences X^i, respectively. All values are merged to obtain the optimum value of the estimated biomass. In Fig. 14, the results achieved with this approach can be observed.

$$X = P\left[\frac{m}{M} + \sum_i\left(\frac{X_l^{i.OPT}}{P^i} - \frac{m^i}{M^i}\right)\right] \tag{31}$$

$$P^{-1} = M^{-1} + \sum_i\left[\left(P^i\right)^{-1} - \left(M^i\right)^{-1}\right] \tag{32}$$

$$\begin{aligned} M^i &= P^i \\ m^i &= X^{i.OPT} \\ m &= X \\ M &= P \end{aligned} \tag{33}$$

This architecture allows the complete autonomy of the local filters. The system achieves optimality in each individual local filter and global optimality in the primary filter.

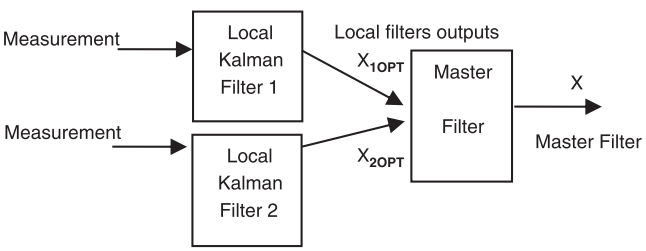

Figure 13. Fusion scheme through a decentralized Kalman filter.

DOI: 10.1002/apj

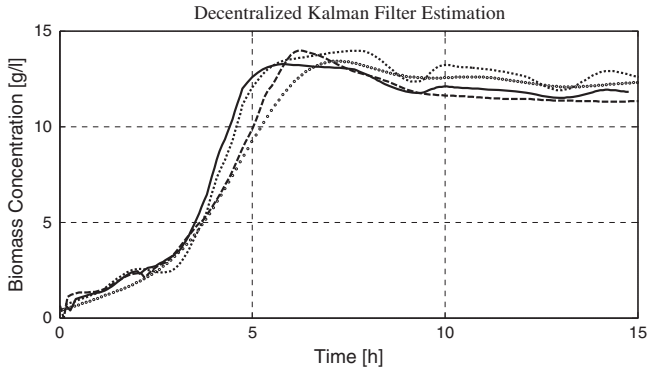

Figure 14. Biomass estimator performance. The dashed line describes the biomass evolution obtained from the phenomenological estimator and the dot line describes a biomass estimation obtained from the ANN. The solid line describes the biomass evolution obtained through the decentralized Kalman filter. The real biomass measurements are represented by small circles.

RESULTS, ANALYSIS, AND DISCUSSION

From Figs. 4, 8, 10, 12, and 13, it can be seen that the five proposed estimators follow more or less closely the true and the virtual filtered biomass measurements. Similar results can be obtained for almost all fermentations except for some atypical fermentations in which the DO concentration decreases markedly due to the excessive aggregation of antifoam during the batch evolution. This undesirable behavior can be avoided or at least minimized by an adequate control of the antifoam dosage.

As was pointed out in the introduction, the performance of phenomenological estimators depends strongly on the accuracy of the model used to describe the bioprocess. In the case of the phenomenological biomass estimator based on DO balance, it is important only that the model describe accurately the dynamics of DO concentration (see (13). However, in the case of EKF, it is necessary the accuracy of the full model described in Section Bioprocess model (see (15 to 21). This fact and the simplicity of the estimator based on DO balance are the main advantages over the standard EKF (other drawbacks of the EKF are that is difficult to tune and is reliable only for systems that are almost linear).

Another interesting point related to the previous paragraph is as follows: to ensure an efficient production of δ-endotoxin of *Bt*, it is important to keep throughout the fermentation an optimal profile of DO concentration. In other words, the DO concentration should be controlled in such a way that it never falls below the minimal critical value for the microorganism growth (for most of the microorganisms the effect of oxygen limitation ranges between 0.1 and 1 mg/L), or exceeds certain threshold level of DO concentration (high levels of DO concentrations promote the formation of toxic O2 compounds for *Bt*). This means that if the DO concentration is adequately controlled, then the linear model of the biomass estimator based on the DO balance (see (13) is approximately valid throughout the fermentation.

Regarding biomass empirical estimators, recall that they are constructed exclusively from the process input–output noisy data without considering the functional relations between the bioprocess variables. The time evolution of biomass is conceived as a dynamic system perturbed by a certain process noise. The estimation of the biomass concentration is obtained indirectly through the observation of noisy measurements. This means that the biomass empirical estimation problem can be formulated as a filtering problem.

For the two proposed empirical biomass estimators, i.e. an estimator based on Bayesian regression with Gaussian process and an ANN biomass estimator of the same set of regressors have been selected. This selection is a problem related to the choice of states in a state-space representation of the system, and they are chosen as finite-dimensional projections of past data. Finding a set of 'good' regressors for biomass estimation is a nontrivial task. In this work, this selection has been performed heuristically by trial and error between numerous alternatives.

For the same training data set D, in the case of the estimator based on Bayesian Gaussian process regression, an acceptable performance is achieved setting the 11 hyperparameters of the covariance function (see (23) in Section Estimator based on Bayesian Gaussian process regression). To obtain a similar performance with a multilayer feedforward neural network-based estimator with four inputs to the input layer (the dimension of the regressor vector), one hidden layer of 30 neurons, and one neuron in the output layer (see Section Artificial neural networks-based estimator) hundreds of parameters can be calculated during the training phase. Due to the higher number of ANN parameters, it is very probable that the variance of the ANN biomass estimator over different experiments (fermentations) is higher than the Gaussian process estimator.

Finally, with regard to the estimator based on information fusion by a decentralized Kalman filter. This is an example to show how to refine the accuracy of the biomass estimation by using two sequences of independent estimates. The first one is provided by the phenomenological estimator based on DO balance and the second one by the ANN-based estimator. These two sequences are 'fused' by a decentralized Kalman filter, and as a result, we obtain a slightly improved estimation performance (see Fig. 14.). This result holds because when two or more estimators of the same quality are fused, the resulting estimator shows only a marginal improvement. In this situation, the advantages of this estimator are not apparent. The major advantage of this scheme is the relative insensitivity to errors in the fused

DOI: 10.1002/apj

estimators, for example, in case of one of the fused estimators in the previous situation fail suddenly; in this new situation, the resulting estimator shows only a little degradation of performance. This property is known as 'fault-tolerance' or 'graceful degradation' property.

CONCLUSIONS

In this paper, it has been addressed the problem of on-line estimation of the biomass concentration in a particular batch bioprocess (δ-endotoxins production of *Bt*). This research topic has been primarily motivated by the need to replace in the process control system the actual (unavailable) biomass concentration measurements by reliable and robust on-line estimations. Five different alternatives to face the problem of biomass concentration estimation have been developed and discussed. If a reliable and accurate model of the bioprocess is available (as in our case study), then simple phenomenological estimators (as the biomass estimator based on DO balance) are the first option.

From the process control point of view, the graceful degradation property of the estimator based on information fusion is of great importance. This is so because the insensitivity of the control system relative to estimator errors is increased, thereby increasing the overall system robustness. This fact motivates the inclusion of robust biomass estimators in the control loop. This task will be addressed in future works.

NOMENCLATURE

Symbol	Description
S	Limiting substrate concentration [g. L^{-1}]
Ts	Sampling time [h]
X_s	Sporulated cells concentration [$g. L^{-1}$]
X_v	Vegetative cells concentration [g. L^{-1}]
μ	Specific growth rate [h^{-1}]
μ_{max}	Maximum specific growth rate [h^{-1}]
m_s	Maintenance constant [*g substrate*. [g cells. h^{-1}]$^{-1}$]
k_s	Kinetic constant representing the spore formation [h^{-1}]
k_e	Death cell specific rate [h^{-1}]
$Y_{X/S}$	Growth yield [g cells. g substrate^{-1}]
K_s	Substrate saturation constant [g. L^{-1}]
K_O	Oxygen saturation constant [g. L^{-1}]
K_1	Oxygen consumption constant by growth (dimensionless)
K_2	Oxygen consumption constant for maintenance [h^{-1}]
K_3	Ventilation constant [L^{-1}]
DO*	O2 saturation concentration (DO concentration in equilibrium with the oxygen partial pressure of the gaseous phase) [g. L^{-1}]
Q_A	Air flow that enters the bioreactor [L. h^{-1}]

REFERENCES

[1] G. Bastin, Dochain, D. *On-Line Estimation and Adaptive Control of Bioreactors*, Amsterdam, Elsevier, ISBN-13: **1990**; pp. 978–0444884305.
[2] D. Dochain. *J. Process Control*, **2003**; *13*, 801–818.
[3] R. Leal Ascencio. Artificial neural networks as a biomass virtual sensor for a batch process. *In Proceedings of the 2001 IEEE international symposium on intelligent control*, **2001**.
[4] B. Li. Artificial neural network based software sensor for biomass during microorganism cultivation. *PhD thesis*. South China University of Technology, **2003**.
[5] A. Amicarelli, F. di Sciascio, H. Álvarez, O. Ortiz. (In Spanish) estimación de biomasa en un proceso batch: aplicación a la producción de δ-endotoxinas de Bt. In: *XXII Interamerican congress of chemical engineering*, **2006**.
[6] R. M. Neal. *Bayesian Learning for Neural Networks*, Springer-Verlag, New York, ISBN 0-387-94724-8, **1996**.
[7] R. M. Neal. Monte Carlo implementation of GP models for Bayesian regression and classification. *Tech. Rep*. No. 9702. Toronto: Department of Statistics, University of Toronto, **1997**.
[8] A. Amicarelli, F. di Sciascio, O. Quintero, O. Ortiz "On-line biomass estimation in a batch biotechnological process: bacillus thuringiensis δ -endotoxins production." Biomass, Book edited by: Maggie Momba and Faizal Bux, ISBN 978-953-307-113-8, pp. 202, September **2010**, Sciyo.
[9] M. Starzak, R. Bajpai. A structured model for vegetative growth and sporulation in Bacillus thuringiensis. *Appl. Biochem. Biotechnol.*, **1991**; *28/29*, 699–718.
[10] A.I. Aronson. *Mol. Microbiol.*, **1993**; *7*, 489–496.
[11] B.L. Liu, Y.M. Tzeng. *Biotechnol. Bioeng.*, **2000**; *68*, 11–17.
[12] F. Vallejo, A. González, A. Posada, A. Restrepo, S. Orduz. *Biotechnol. Tech.*, **1999**; *13*, 279–281.
[13] P. Caballero, J. Iriarte. Biología y ecología de Bacillus thuringiensis. Bioinsecticidas: fundamentos y aplicaciones de Bacillus thuringiensis en el control integrado de plagas. *Editorial Phytoma-España*, **2001**; *1*, 15–44 318 p.
[14] D. Ghribi, N. Zouari, H. Trabelsi, S. Jaoua. *Enzyme Microb. Technol.*, **2007**; *40*(4), 614–622.
[15] D. Rivera, A. Margaritis, H. De Lasa. *Can. J. Chem. Eng.*, **1999**; *77*, 903–910.
[16] P. Atehortúa, H. Alvarez, S. Orduz. *Bioprocess Biosyst. Eng.*, **2007**; *30*, 447–456.
[17] A. Amicarelli, F. di Sciascio, J.M. Toibero, H. Álvarez. *Braz. J. Chem. Eng.*, **2010**; *27*(01), 41–62.
[18] F. di Sciascio, A. Amicarelli. *Comput. Chem. Eng.*, **2008**; *32*, 3264–3273.
[19] A. Amicarelli. Modelado, identificación y control de bioprocesos. (In Spanish) *PhD thesis*. Universidad Nacional de San Juan Argentina – Instituto de Automática. ISBN978-987-05-7087-5, **2009**:1–191.
[20] F. Daum. Nonlinear filters: beyond the Kalman filter. *IEEE Aerosp. Electron. Syst. Mag.*, **2005**; 20:8:Part 2, 57–69.
[21] A.H. Jazwinski. *Stochastic Processes and Filtering Theory*, Academic Press, Inc. New York, **1970**.
[22] B.D.O. Anderson, J.B. Moore. Optimal filtering. Prentice-Hall, Englewood Cliffs, New Jersey. ISBN 0136381227 : 0136381227, **1979**.
[23] D. Simon. *Optimal State Estimation: Kalman, H∞, and Nonlinear Approaches*. Hoboken, New Jersey, John Wiley & Sons, Inc., **2006**.
[24] K. Narendra, K. Parthasarathy. Identification and control of dynamical systems using neural networks. *IEEE Trans. Neural Netw.*, **1990**; *1*, 4–27.
[25] S. Haykin. *Neural Networks: A Comprehensive Foundation*. 2nd Ed. Prentice-Hall, Inc., New Jersey, **1999**.
[26] R. Brown, P. Hwang. *Introduction to Random Signals and Applied Kalman Filtering*, 3rd edn, New York, John Wiley & Sons, **1997**; pp. 371–375.