Comparison of Several Filtering Approaches on Water Treatment Processes

António Pedro Aguiar[1], **Oussama Hadj-Abdelkader**[1,2*]

[1]*Research Center for Systems and Technologies (SYSTEC)*
Department of Electrical and Computer Engineering
Faculty of Engineering, University of Porto
4200-465 Porto, Portugal
E-mail: pedro.aguiar@fe.up.pt

[2]*Tlemcen Automatics Laboratory (LAT)*
University of Tlemcen
BP119, 13000 Tlemcen, Algeria
E-mail: hadjabdelkader.oussama@gmail.com

[*]*Corresponding author*

Received: March 27, 2020 **Accepted: November 12, 2020**

Published: September 30, 2021

Abstract: *This paper addresses the state estimation problem of a bioreactor in wastewater treatment processes. The state variables of this process are the concentrations of the organic pollutants and of the bacteria inside the bioreactor. A specific growth rate function is used to describe the variation of the bacteria concentration when the amount of pollutants increases. This rate can also represent the speed of the biological degradation of the pollutants. Most research work in this field uses only deterministic models that do not conveniently account for uncertainties. These models are often obtained using several simplifications during the modeling procedure such as neglecting the measurement noises. In this paper, we consider stochastic models and study the state estimation problem using three approaches: the Extended Kalman filter, the Unscented Kalman filter and the Particle filter. These methods are adapted to the models in study and compared to understand which is the most adequate for this type of processes considering their slow evolution, discrete time measurements and high-intensity noises. Further, we also apply a Multiple Model Adaptive method which adapts the filters to the correct growth rate type. This method is also used to automatically choose the most efficient estimation method for this type of biological processes.*

Keywords: *Filtering, Bioreactor, Particle filter, Extended Kalman filter, Unscented Kalman filter, Continuous-discrete systems, Stochastic differential equations.*

Introduction

Wastewater treatment is currently a major issue, whether it is considered from an environmental or an economic viewpoint. The purpose of such a process is to eventually recover the water for a possible agricultural or urban use or, at least, to give a post-treatment of contaminated water in order to allow its discharge into the natural water reserves. The wastewater treatment process involves two major steps: (i) a mechanical treatment including the filtration, the grit removal, the degreasing and the decantation, and (ii) a biological treatment using mainly the bioreactor and the anaerobic digester [4]. In this work, we focus on the biological reactions taking place inside the bioreactor, that consist in the degradation of the polluting organic substances contained in the water by some bacteria population. This bacteria population evolves very often in an anaerobic (without oxygen) medium. The anaerobic digestion taking place in this case produces energy in the form of biogas together with treated water. The control of a bioreactor involves its stabilization around a stable equilibrium to avoid the extinction of the bacterial population and to provide the best possible water treatment and/or the maximum amount of biogas energy.

An important aspect that must be taken into account in this type of bioreactors is the biomass (bacteria) and the substrate (organic substances) concentrations inside the bioreactor. These

quantities will represent the state variables in the mathematical models. Unfortunately, these concentrations are not directly measurable at the system's output, and therefore it is necessary to use a software sensor to estimate their values. One solution is the use of state observers, which are methods that generally require a relatively accurate knowledge of the system and also require the system to be observable. A detailed study about this last condition is established in [14]. In this field, many of the observer based methods have been applied on deterministic models which may be valid in the well mastered laboratory conditions [9]. However, it is not always the case in reality. In a real water treatment process, the aleatory state variations appear on the system and need to be taken into account [9]. This is also the case in the anaerobic digestion where several types of bacteria intervene in the reactions [1,41]. Hence, it is important that the system at this level of description be modeled by Stochastic Differential Equations (SDEs). This type of models provides a more accurate description of the system by considering the uncertainties in the dynamics. In addition, the measurements are constantly subject to noises and uncertainties, which need to be modeled in the output equation.

We attempt in this work to give the best possible estimates of the system states using the SDE models. For this purpose, we use stochastic (Bayesian) filtering approaches which are better adapted for the considered models. Similar research work appear in the litterature, but are mostly focused on the use of state observers on deterministic bioreactor models. A first state-of-the-art on this topic can be found in [11] and [8]. Other results include the use of invariant observers [12], non-linear Luenberger observers [13], interval observers [17], unknown input observers [34], asymptotic observers [28], and others [16, 32, 37]. However, these methods are exclusively applied for deterministic bioreactor models, which are valid only under some very precise conditions. Kalman filters were also used to address this problem in [8, 22, 25, 39] assuming only additive normal noises on the states and on the output. Unlike the models considered in these last references, the ones used in this paper provide a detailed quantification of the stochasticity in the dynamics and of the output noises. The use of the estimation approaches on this type of models appears in [5, 18, 19], where the Extended Kalman filter (EKF), the Unscented Kalman filter (UKF) and the Particle filter (PF) were applied on the simple second order Chemostat model. We extend these works in this paper by applying these approaches on the more complicated anaerobic digestion models. Other different methods such as Artificial Neural Network can be found in [23]. It is important to stress that all of the works mentioned above, in general have obtained satisfying results, but under relative benign conditions.

In this paper, we study and compare three different but very common filtering approaches for state estimation of nonlinear processes, which are the EKF [2], the UKF [24] and the PF, also known as sequential Monte Carlo method [27]. We apply these methods on the considered stochastic models and compare their performance and robustness against unknown initial conditions, high noise intensities and low frequency observations. We also contrast their performance with respect to a Bayesian Cramér-Rao bound (CRLB) performance indicator.

In addition, we study the use of a Multiple Model Adaptive Estimation (MMAE) framework [20, 21] for two main reasons. One first reason is to see if there is an improvement when the proposed filters are combined (that is, the EKF, UKF, and PF all running in parallel) in a MMAE structure. The MMAE in this case will assign a weight to each filter describing how close are the obtained estimation results from the real state variables. Another purpose is to use the MMAE approach to also estimate the growth rate type of the biomass. Note that the growth rate is by far the most significant parameter within the bioreactor model. It depends on the type of biological reaction taking place in the process and in particular this growth rate function is

related to the type of bacteria used for the treatment. Depending on the applications, there exist in fact many types of functions. In our application, we consider that the growth rate can be either of Monod-type [30] or of Haldane-type function [3], which for water treatment purposes, these growth functions are usually the more appropriate ones.

This paper is organized as follows: Section 2 presents the stochastic bioreactor models used in this work and their implementation using an Euler-Maruyama scheme. The problem statement follows in this same section. In Section 3 we present the application of the classical estimation algorithms EKF, UKF and PF that are mostly used in the stochastic context. We also introduce the MMAE used for the estimation of the bacterial growth rate and for combining the previous estimation methods. The results of all these methods are compared and discussed in Section 4 starting by introducing the Bayesian CRLB used as a performance measure for the filters and comparing the obtained concentrations estimates and the corresponding root mean square errors. The MMAE framework is also analyzed via simulations for the selection of the most likely estimates and growth rates. All these results are commented and discussed along with their presentation. Conclusions and suggestions for future research are summarized in Section 5.

The bioreactor models

We consider three different models for bioreactors of water treatment processes (Fig. 1). The first one, is a simple second order model of the bioreactor called the Chemostat model. This model, that was introduced in [38], captures the reaction of a single type of bacteria that is used to degrade the pollutants. This model is the simplest model that can be used to represent the bioreactor either in aerobic or anaerobic operating conditions and is described as:

$$\begin{aligned} \dot{B}_t &= (\mu(S_t) - D)B_t, \\ \dot{S}_t &= D(s_{in} - S_t) - k_{sc}\mu(S_t)B_t, \end{aligned} \tag{1}$$

where $X_t = (B_t, S_t)^{\mathrm{T}}$ is the state variable with B_t denoting the biomass concentration, and S_t is the substrate concentration at time t. The function $\mu(S_t)$ is a specific growth rate function. The parameter s_{in} denotes substrate concentration in the input, D is the dilution rate and k_{sc} is the yield coefficient.

Next, we consider the fourth order model named AM2 describing the two steps anaerobic digestion process in the bioreactor, see details in [7]. In this model a population of acidogenic bacteria degrade the organic pollutant matters to obtain the fatty volatile Acids. Then, these lasts are degraded by a population of methanogenic bacteria to obtain the biogas. This model is given by:

$$\begin{aligned} \dot{B}_1(t) &= (\mu_1(S_1(t)) - \alpha D) B_1(t), \\ \dot{S}_1(t) &= D(s_{1in} - S_1(t)) - k_1\mu_1(S_1(t)) B_1(t), \\ \dot{B}_2(t) &= (\mu_2(S_2(t)) - \alpha D) B_2(t), \\ \dot{S}_2(t) &= D(s_{2in} - S_2(t)) + k_2\mu_1(S_1(t)) B_1(t) - k_3\mu_2(S_2(t)) B_2(t), \end{aligned} \tag{2}$$

where $B_1(t)$ is the concentration of the acidogenic bacteria, $B_2(t)$ is the concentration of the methanogenic bacteria, $S_1(t)$ is the concentration of the organic pollutants, and $S_2(t)$ is the concentration of the acids. $\mu_1(S_1(t))$ and $\mu_2(S_2(t))$ are the specific growth rate functions; s_{1in} and s_{2in} are the input concentrations of the substrates $S_1(t)$ and $S_2(t)$, respectively; D is the dilution rate; α is the part of $B_1(t)$ and $B_2(t)$ that has left the bioreactor; k_1, k_2 and k_3 are yield coefficients. The third model is a fifth order extension of the AM2 named the AM2b, which

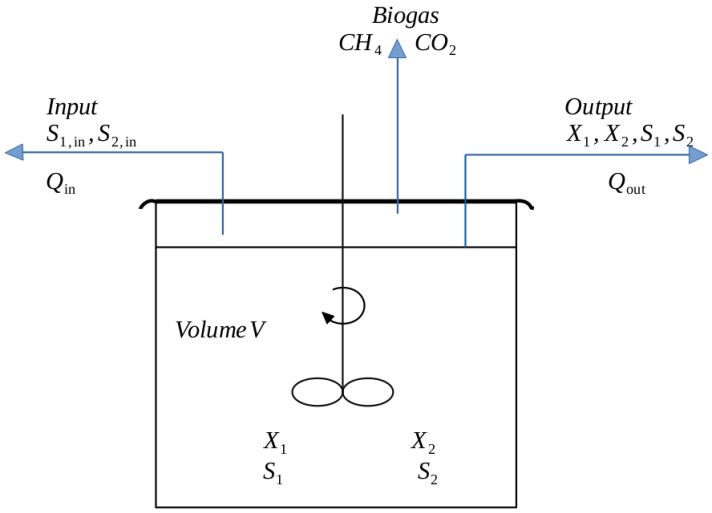

Fig. 1 The anaerobic bioreactor

includes the dynamics of the Soluble Microbian Products (SMP). The dynamics of these SMP provide a description of the membrane fouling phenomenon in the membrane bioreactor, [6]. This model is described as follows:

$$\begin{aligned}
&\dot{B}_1(t) = (\mu_1(S_1(t)) + \mu(S(t)) - D_{wit} - D_{dec}) B_1(t),\\
&\dot{S}_1(t) = D_{in}(s_{1in} - S_1(t)) - k_1\mu_1(S_1(t)) B_1(t),\\
&\dot{B}_2(t) = (\mu_2(S_2(t)) - D_{wit} - D_{dec}) B_2(t),\\
&\dot{S}_2(t) = D_{in}(s_{2in} - S_2(t)) + c_{12}\mu_1(S_1(t)) B_1(t) + c_{02}\mu(S(t)) B_1(t) - k_2\mu_2(S_2(t)) B_2(t),\\
&\dot{S}(t) = (c_{10}\mu_1(S_1(t)) + D_{dec} - k_0\mu(S(t))) B_1(t) + (c_{20}\mu_2(S_2(t)) + D_{dec}) B_2(t) - MS(t).
\end{aligned} \tag{3}$$

The additional variable $S(t)$ represents the SMP concentration. The other parameters are k_0 which represents the degradation rate of $S(t)$; c_{12} and c_{02} are the production rates of the acids and c_{i0} is the production rate of the SMP; D_{in} is the dilution rate; D_{dec} is the biomass decay rate; D_{wit} is the biomass withdraw rate and $D_{out} = D_{in} - D_{wit}$ is the output rate; $M = \beta D_{in} - (1-\beta) D_{wit}$ with β is the amount of SMP which has left the bioreactor.

By considering the stochastic versions of these models proposed by [9], [1] and [41] respectively, all of the three models presented above can be casted as state equations in the form of SDEs and can be generally represented as:

$$dX_t = f(X_t)\, dt + g(X_t)\, dW_t, \tag{4}$$

where X_t is a diffusion process that represents the state vector of the bioreactor, namely, the concentrations of the different bacterial populations and the substrates. It also contains the representation of the SMP in the case of the AM2b. W_t is a vector of brownian motions with the same size as the state vector. The vector field $f(X_t)$ represents the mean of the state dynamics, which is exactly the right hand side of the corresponding deterministic models, and $g(X_t)$ represents the variance of the dynamics and it is the part that was added to the deterministic models to provide a quantification of the stochastic effects on the dynamics. Since the state variables of the process represent concentrations, we suppose that $X_t \geq 0$ in the simulation time $t \in [0, T]$, where $T > 0$ is some final time.

In these models, the growth rate of the bacterial populations is given by the growth rate function $\mu(S_t)$ which depends on the concentration of the corresponding substrate S_t. In wastewater

treatment processes, this growth rate function is generally of a Monod-type or of a Haldane-type. For the former, it takes the form

$$\mu\left(S_t\right) = \mu_{max}\frac{S_t}{K_s + S_t},$$

where μ_{max} is the maximum growth rate and K_s is the half-saturation constant. Whereas for the latter case, it takes the form

$$\mu\left(S_t\right) = \mu_{max}\frac{S_t}{\frac{S_t^2}{K_i} + S_t + K_2},$$

where K_2 and K_i are constants.

For the models output equations, it is supposed that only one of the substrate concentrations noted S_{t_k} is measured at discrete time instants $t_k = k\Delta$, where Δ is the observation time step. This output is subject to a measurement noise v_k of intensity σ. The practical tests had shown that this noise is supposed to be proportional to S_{t_k} [5]. Thus, we obtain the output equation

$$y_k = S_{t_k} + \sigma S_{t_k} v_k, \tag{5}$$

where $v_k \overset{iid}{\sim} \mathcal{N}\left(0,1\right)$ (independent and identically distributed from a Gaussian distribution).

System simulation

To simulate System (4), we use an Euler-Maruyama scheme [26] with time step δ, see Algorithm 1. For a correct simulation, this time step δ is chosen small enough to have a good approximation of the SDE, but not too small to avoid high computational cost. For a simulation time T, and a given number of observations N_{obs} in the time interval $\left[0, T\right]$, we perform N_{sys} iterations of system simulation between every two observations y_k with the simulation step $\delta = T / \left(N_{sys} N_{obs}\right)$, and compute an output value y_k for every step $\Delta = T / N_{obs}$.

The Brownian motion terms are approximated as $dW_{t_n} = \sqrt{\delta} w_{t_n}$, where $t_n = n\delta$, $w_{t_n} \overset{iid}{\sim} \mathcal{N}\left(0,\, I\right)$ and I is the identity matrix.

The Euler-Maruyama approximation of System (4) results in

$$X_{t_n} = X_{t_{n-1}} + f\left(X_{t_{n-1}}\right)\delta + g\left(X_{t_{n-1}}\right)\sqrt{\delta}\, w_{t_n}.$$

The simulation of both the state equation and output equation is given in Algorithm 1. Note that only the non-negative solutions are taken into account because the states represent concentrations, the values are set to 0 whenever they cross the time axis.

Problem formulation

The real concentrations of the biomass and the substrates inside the bioreactor are unknown in practice because presently it is not easy to measure them using an electronic device. One possible way to know their values is by performing a laboratory analysis. This justifies the discrete and low frequency nature of the measurements. Amongst the previously cited works, only [5, 18, 19, 34] considered the discrete nature of the measurements used for the state estimation. This lack of measured information precludes the user from knowing the state of the system. The impact of this issue appears when we would like to use the values of the state variables of the system in the control task. One way to deal with this problem is to estimate these

Algorithm 1 Simulation of the process using Euler-Maruyama scheme

$v_0, \ldots, v_{N_{obs}} \sim \mathcal{N}(0,1)$
$w_0, \ldots, w_{N_{sys}*N_{obs}} \sim \mathcal{N}(0,1)$
initialization
$X_{t_0} \sim \mathcal{N}(\mu_0, Q_0)$
# iterations
For $k = 0, \ldots, N_{obs}$ **do**
  For $n = 1, \ldots, N_{sys}$ **do**
    $\mu\left(S_{t_{n-1}}\right) = \mu_{max}\frac{S_{t_{n-1}}}{K_s + S_{t_{n-1}}}$ or $\mu\left(S_{t_{n-1}}\right) = \mu_{max}\frac{S_{t_{n-1}}}{\frac{S_{t_{n-1}}^2}{K_i} + S_{t_{n-1}} + K_2}$
    $X_{t_n} = \max\left(0, X_{t_{n-1}} + f\left(X_{t_{n-1}}\right)\delta + g\left(X_{t_{n-1}}\right)\sqrt{\delta}\, w_n\right)$
  End For
  $S_k \leftarrow S_{t_n}$
  $y_k = S_k + \sigma S_k v_k$
End For

concentrations online while the process is running. Therefore, in this paper we propose three state estimation methods (EKF, UKF, PF) and show their application to the presented models. The obtained results are compared to find out which method is the most suitable for this type of biological systems given that these systems exhibit nonlinear dynamics and also include other interesting properties such as high noises and infrequent measurements. Such properties represent a challenge for the estimation algorithms.

Another key problem is the estimation of the specific growth rate. This parameter is related to the type of bacteria used in the system. Generally, it can be described either by a Monod function or a Haldane function. The goal is to understand if a MMAE method together with the previous filters can indeed address this issue. To avoid complicated scenarios, we suppose that maximum growth rates, half saturation constants and the other parameters are known in this case. The case of unknown static parameters will be the subject of our future work. Advantages and disadvantages of the considered algorithms are illustrated and discussed through computer simulations to determine the most adequate method.

Nonlinear filtering algorithms for the bioreactor models

The Extended Kalman filter

The Extended Kalman filter is the nonlinear version of the standard Kalman filter (KF) by linearizing the system around the previously obtained state estimation in each iteration. It is considered as the standard algorithm in the theory of non linear state estimation and tracking. Similarly to the usual Kalman filter, the EKF also contains two main parts: the prediction and the update steps. The detailed application of these steps on the system in study is given in the next subsections.

Prediction step

In order to carry out the EKF prediction step for System (4), we need to linearize it around the previous mean estimation $\bar{X}_{t_n}$ and to compute its predicted expectation $\bar{X}_{t_{n+1}} = \mathbb{E}\left[X_{t_{n+1}}\right]$ and its predicted variance $R_{t_{n+1}} = var\left[X_{t_{n+1}}\right]$. After using an Euler scheme with a time step δ, we get

the following expressions of the EKF's prediction equations:

$$\bar{X}_{t_{n+1}} = \bar{X}_{t_n} + f\left(\bar{X}_{t_n}\right)\delta, \tag{6}$$

$$R_{t_{n+1}} = R_{t_n} + R_{t_n}F_{t_n}^{*}\delta + F_{t_n}R_{t_n}\delta + g\left(\bar{X}_{t_n}\right)g\left(\bar{X}_{t_n}\right)^{*}\delta, \tag{7}$$

where $F_{t_n} := \nabla f\left(\bar{X}_{t_n}\right)$ is the Jacobian of $f\left(\bar{X}_{t_n}\right)$.

The EKF prediction step of System (4) is given in Algorithm 2, where $\hat{X}_{t_n}^{-}$ and $R_{t_n}^{-}$ denote, respectively, the predicted mean value of X_{t_n} and its covariance matrix at time t_n, whereas $\hat{X}_{t_n}$ and R_{t_n} denote the estimated mean value of X_{t_n} and its covariance at time t_n.

Algorithm 2 Prediction step of the EKF for the bioreactor

initialization
$\delta = T/(N_{sys} * N_{obs})$
$X_0 \sim \mathcal{N}(\mu_0, Q_0)$
$\hat{X}_{t_0} = \mu_0$
$R_{t_0} = Q_0$
# iterations
For $k = 0, \ldots, N_{obs}$ **do**
# prediction step
  For $n = 1, \ldots, N_{sys}$ **do**
    $\hat{X}_{t_n}^{-} \leftarrow max\left(0, \hat{X}_{t_{n-1}} + f\left(\hat{X}_{t_{n-1}}\right)\delta\right)$
    $R_{t_n}^{-} \leftarrow R_{t_{n-1}} + \left(R_{t_{n-1}}F_{t_{n-1}}^{*} + F_{t_{n-1}}R_{t_{n-1}} + g\left(X_{t_{n-1}}\right)g\left(X_{t_{n-1}}\right)^{*}\right)\delta$
    $\hat{X}_{t_n} \leftarrow \hat{X}_{t_n}^{-}$
    $R_{t_n} \leftarrow R_{t_n}^{-}$
  End For
End For

Update step

In this step, the same method of the standard KF is used to calculate the updated values of the mean $\hat{X}_{t_n}$ and covariance matrix R_{t_n}. To this end, we first have to apply a convenient change of coordinates given the specific nonlinearity between the state variable S_{t_k} and the noise variable v_k in the output Eq. (5), which does not enter only in an additive form. More precisely, from Eq. (5) we obtain

$$y_k = S_{t_k}\left(1 + \sigma v_k\right)$$

by setting

$$\tilde{y}_k = \log\left(y_k\right),$$

$$\tilde{y}_k = \log\left(S_{t_k}\right) + \log\left(1 + \sigma v_k\right)$$

and using a first order Taylor series development of $\log\left(1 + \sigma v_k\right)$ at point 1, we obtain a linear output equation

$$\tilde{y}_k = \log\left(S_{t_k}\right) + \log\left(1\right) + \sigma v_k \log^{'}\left(1\right),$$

$$\tilde{y}_k = \log\left(S_{t_k}\right) + \sigma v_k. \tag{8}$$

The following notation is used in Algorithm 3:

$h(X_{t_k}) = \log(S_{t_k})$,

$H_{t_k} = \nabla h(X_{t_k})$.

Notice that replacing y_k by $\tilde{y}_k = \log(y_k)$ does not affect the estimation algorithm or its quality since the same information is acquired in both cases. This substitution is carried-out only in the update step: the system simulation still gives an output value y_k. However, this also implies considering $\log(y_k)$ instead of y_k in the EKF when updating the estimation of $\hat{X}_{t_k}$. The obtained update step of the EKF is given in Algorithm 3.

Algorithm 3 Update step of the EKF for the bioreactor

For $k = 0, \ldots, N_{obs}$ **do**
# update step
$\hat{X}_{t_k}^- \leftarrow \hat{X}_{t_n}^-$
$R_{t_k}^- \leftarrow R_{t_n}^-$
$K_k = R_{t_k}^- H_{t_k}^* \left(H_{t_k} R_{t_k}^- H_{t_k}^* + \sigma^2\right)^{-1}$ # Kalman gain
$\hat{X}_{t_k} \leftarrow \hat{X}_{t_k}^- + K_k \left(\log(y_k) - h\left(\hat{X}_{t_k}^-\right)\right)$
$R_{t_k} \leftarrow (I - K_k H_{t_k}) R_{t_k}^-$
End For

The Unscented Kalman filter

Because of the non-linearity between the noise and the state variables in the System (4) and the output Eq. (5), the UKF algorithm with a non-additive noise, presented in [35, 36, 43], appears to be a good choice. The application of this algorithm to the stochastic models of the bioreactor is given in two steps as detailed in the next subsections.

Prediction step

In the prediction step, since the model is a non-linear function of both the state and the Brownian motion terms, following the approach in [35], we add the noise variables in the state vector to create an augmented state variable with mean $\tilde{X}_t = \begin{bmatrix} \hat{X}_t & 0 \end{bmatrix}^{\mathrm{T}}$ and covariance matrix

$$\tilde{R}_t = \begin{bmatrix} R_t & 0 \\ 0 & I \end{bmatrix}.$$

The UKF uses the unscented transform [24] to approximate the target distribution. The idea of the unscented transform is to create a fixed number of sigma points $\tilde{\mathcal{X}}_{t,i}$ from the previous estimation of the mean $\tilde{X}_t$ and covariance matrix $\tilde{R}_t$. The parameters α, β and κ are tuned to create these points and to determine their spread around the mean. These parameters are also used to compute the weights W of the sigma points which are computed according to

$$W_0^{m,j} = \frac{\lambda_j}{n_j + \lambda_j},$$

$$W_0^{c,j} = \frac{\lambda_j}{n_j + \lambda_j} + (1 - \alpha^2 + \beta),$$

$$W_i^{m,j} = \frac{\lambda_j}{2(n_j+\lambda_j)}, \qquad i = 1, \ldots, 2n_j,$$

$$W_i^{c,j} = \frac{\lambda_j}{2(n_j+\lambda_j)}, \qquad i = 1, \ldots, 2n_j,$$

where $\lambda_j = \alpha^2(n_j+\kappa) - n_j$, $j = 1$ in the prediction step and $j = 2$ in the update step.

Then, these sigma points are propagated through the non-linearity in the state equation to obtain new sigma points $\tilde{\mathcal{X}}_{t^-,i}$, and the predicted mean and covariance of the current state are estimated from these sigma points using the weights. The prediction step of the UKF is summarized in Algorithm 4.

Algorithm 4 Prediction step of the UKF for the bioreactor

initialization
$X_0 \sim \mathcal{N}(\mu_0, Q_0)$
$\hat{X}_0 = \mu_0$
$R_0 = Q_0$
# iterations
For $k = 0, \ldots, N_{obs}$ **do**
  For $t = \delta, 2\delta, \ldots, N_{sys}\delta$ **do**
  # prediction step
  #creating sigma points
    $\tilde{\mathcal{X}}_{t-\delta,0} = \tilde{X}_{t-\delta}$
    $\tilde{\mathcal{X}}_{t-\delta,i} = \tilde{X}_{t-\delta} + \sqrt{n_1+\lambda_1}\left[\sqrt{\tilde{R}_{t-\delta}}\right]_i, \quad i = 1, \ldots, n_1$
    $\tilde{\mathcal{X}}_{t-\delta,i+n_1} = \tilde{X}_{t-\delta} - \sqrt{n_1+\lambda_1}\left[\sqrt{\tilde{R}_{t-\delta}}\right]_i, \quad i = 1, \ldots, n_1$
  # propagating the sigma points in the state equation
    $\tilde{\mathcal{X}}_{t,i} = \tilde{\mathcal{X}}^x_{t-\delta,i} + f(\tilde{\mathcal{X}}^x_{t-\delta,i})\delta + g(\tilde{\mathcal{X}}^x_{t-\delta,i})\sqrt{\delta}\,\tilde{\mathcal{X}}^w_{t-\delta,i}, \quad i = 0, \ldots, 2n_1$
where $\tilde{\mathcal{X}}^x_{t-\delta,i}$ denotes the n_x first components of $\tilde{\mathcal{X}}_{t-\delta,i}$ and $\tilde{\mathcal{X}}^w_{t-\delta,i}$ denotes its n_w last components
  # prediction of the state's mean and covariance matrix
    $\hat{X}_{t^-} = \sum_{i=0}^{2n_1} W_i^{m,1}\tilde{\mathcal{X}}_{t,i}$
    $R_{t^-} = \sum_{i=0}^{2n_1} W_i^{c,1}(\tilde{\mathcal{X}}_{t,i} - \hat{X}_{t^-})(\tilde{\mathcal{X}}_{t,i} - \hat{X}_{t^-})^{\mathrm{T}}$
    $R_{t-\delta} \leftarrow R_{t^-}$
    $\hat{X}_{t-\delta} \leftarrow \hat{X}_{t^-}$
  End For
End For

Update step

In the update step, we create an augmented state variable with mean $\bar{X}_{k^-} = \begin{bmatrix} \hat{X}_{k^-} & 0 \end{bmatrix}^{\mathrm{T}}$ and covariance matrix $\bar{R}_{k^-} = \begin{bmatrix} R_{k^-} & 0 \\ 0 & 1 \end{bmatrix}$.

The same method described in the previous subsection is used in this step, that is, we propagate another set of sigma points $\tilde{\mathcal{X}}_{k^-,i}$, which are created from the predicted mean $\bar{X}_{k^-}$ and covariance matrix $\bar{R}_{k^-}$ through the non-linearity in the output equation. The resulting sigma points $\tilde{\mathcal{Y}}_{k,i}$

together with the weights W are used to compute the estimated output mean μ, its covariance matrix S and its cross covariance with the state C. These two lasts are used to compute the Kalman gain K and finally the estimation of the current state's mean $\hat{X}_k$ and covariance matrix R_k.

The UKF update step is given in Algorithm 5, where the following notation is used in this step

$$h(\hat{X}_k, v_k) = \hat{S}_k + \sigma \hat{S}_k v_k$$

and $[.]_i$ denotes de i^{th} column of the matrix in between brackets.

Algorithm 5 Update step of the UKF for the bioreactor

For $k = 0, \ldots, N_{obs}$ **do**

update step

creating sigma points

$\bar{\mathscr{X}}_{k^-,0} = \bar{X}_{k^-}$

$\bar{\mathscr{X}}_{k^-,i} = \bar{X}_{k^-} + \sqrt{n_2 + \lambda_2}\left[\sqrt{\bar{R}_{k^-}}\right]_i, \quad i = 1, \ldots, n_2$

$\bar{\mathscr{X}}_{k^-,i+n_2} = \bar{X}_{k^-} - \sqrt{n_2 + \lambda_2}\left[\sqrt{\bar{R}_{k^-}}\right]_i, \quad i = 1, \ldots, n_2$

propagating the sigma points in the output equation

$\tilde{\mathscr{Y}}_{k,i} = h(\bar{\mathscr{X}}^x_{k^-,i}, \bar{\mathscr{X}}^v_{k^-,i}), \quad i = 0, \ldots, 2n_2$

where $\bar{\mathscr{X}}^x_{k^-,i}$ denotes the n_x first components of $\bar{\mathscr{X}}_{k^-,i}$ and $\bar{\mathscr{X}}^v_{k^-,i}$ denotes its n_v last components

estimation of the output's mean and covariance matrix

$\mu = \sum_{i=0}^{2n_2} W_i^{m,2} \tilde{\mathscr{Y}}_{k,i}$

$S = \sum_{i=0}^{2n_2} W_i^{c,2} (\tilde{\mathscr{Y}}_{k,i} - \mu)(\tilde{\mathscr{Y}}_{k,i} - \mu)^{\mathrm{T}}$

$C = \sum_{i=0}^{2n_2} W_i^{c,2} (\bar{\mathscr{X}}_{k^-,i} - \hat{X}_{k^-})(\tilde{\mathscr{Y}}_{k,i} - \mu)^{\mathrm{T}}$

computing the Kalman gain

$K = CS^{-1}$ % Kalman gain

estimation of the state's mean and covariance matrix

$\hat{X}_k \leftarrow \hat{X}_{k^-} + K(y_k - \mu)$

$R_k \leftarrow R_{k^-} + KSK^{\mathrm{T}}$

End For

In the case of dimension higher than 3, the unscented transform fails to construct the set of sigma points to approximate the state variable. Many extensions had been proposed to solve the unscented Kalman filtering problem for high dimensional systems, e.g., [10, 29, 31, 33]. To estimate the state variables for the AM2 and AM2b stochastic models, we can apply the approach proposed by [31] which consists in reducing the rank of the covariance matrix R using singular values decomposition (SVD) $R = U\Sigma V^{\mathrm{T}}$. This matrix reduction allows to decrease the number of sigma points used for the approximation. A reduced approximation of R is obtained by retaining the first p singular values σ_j. We truncate the dimensions of U, Σ and V to obtain the reduced matrices U_r, Σ_r and V_r, then we can form the reduced covariance matrix given by $R = U_r \Sigma_r V_r^{\mathrm{T}}$. This last matrix is used in the UKF to create the sigma points and one can proceed to the remaining UKF steps.

The Particle filter

The application of the PF to the stochastic models of the bioreactor is given in Algorithm 6. The particle filter gives a Monte Carlo approximation to the conditional probability density function $p_{X_{t_k}|Y_{0:k}}(x_{t_k}|y_{0:k})$ of the form

$$p_{X_{t_k}|Y_{0:k}}(x_{t_k}|y_{0:k}) \approx \frac{1}{N}\sum_{i=1}^{N}\delta_{\xi^i}(x_{t_k}),$$

where $y_{0:k}$ denotes the system's output; ξ^i are the particles, which are distributed according to $p_{X_{t_k}|Y_{0:k}}(x_{t_k}|y_{0:k})$; and δ_{ξ^i} is the Dirac mesure at point ξ^i. Notice, however, that since the density $p_{X_{t_k}|Y_{0:k}}(x_{t_k}|y_{0:k})$ is not available, we compute an approximation of it by drawing several particles from a proposal density $q_{X_{t_k}|Y_{0:k}}(x_{t_k}|y_{0:k})$ and correct them using weights ω^i. The simplest method to do this is to consider the transition density $p_{X_t|X_{t-\delta}}(x_t|x_{t-\delta})$ as a proposal density and compute the weights from the likelihood function $p_{Y_k|X_{t_k}}(y_k|\tilde{\xi}^i)$. From this approximation, we get a two-steps algorithm.

Prediction step

In the prediction step, we use the state Eq. (4) as an approximation of the transition density $p_{X_t|X_{t-\delta}}(x_t|x_{t-\delta})$ and draw the particles ξ^i. This PF prediction step is given in Algorithm 6.

Algorithm 6 Prediction step of the Bootstrap PF for the bioreactor

Initialization
$\delta = T/(N_{sys} * N_{obs})$
$N =$number of particles
$\xi^{1:N} \overset{iid}{\sim} \mathbb{P}(X_0)$
$\omega^i = \frac{1}{N}, \quad i = 1, \ldots, N$ # importance weights
$\hat{X}_0 = \frac{1}{N}\sum_{i=1}^{N}\xi^i$
# iterations
For $k = 0, \ldots, N_{obs}$ **Do**
  For $n = 1, \ldots, N_{sys}$ **Do**
# prediction step
    $\tilde{\xi}^i = max\left(0, \xi^i + f(\xi^i)\,\delta + g(\xi^i)\sqrt{\delta}\,w_n^1\right), \; i = 1, \ldots, N$
    $\xi^i \leftarrow \tilde{\xi}^i$
  End For
End For

Update step

In the update step, we compute the weights from the likelihood function $p_{Y_k|X_{t_k}}(y_k|\tilde{\xi}^i)$ given by

$$p_{Y_k|X_{t_k}}(y_k|\tilde{\xi}^i) = \frac{1}{\sigma S_{t_k}\sqrt{2\pi}}\exp\left(-\frac{(y_k - S_{t_k})^2}{2(\sigma S_{t_k})^2}\right).$$

A resampling procedure is performed within this step to replace the unsignificant low-weight particles by the high-weight ones. The reason for the resampling is to have a large number of efficient particles, that is, to assure that most of the particles converge to the real state instead of only one particle. There exist several resampling methods, see [15] for a comparison of a set of

such methods. In our algorithm, we used the residual resampling method. The PF update step is given in Algorithm 7. The obtained algorithm composed by Algorithm 6 and 7, is called the Bootstrap PF.

Algorithm 7 Update step of the Bootstrap PF for the bioreactor

For $k = 0, \ldots, N_{obs}$ **Do**

update step

$\omega^i \leftarrow p_{Y_k|X_k}(y_k|\tilde{\xi}^i) \quad i = 1, \ldots, N$ # weights

$\omega^i \leftarrow \omega^i / \sum_{j=1}^{N} \omega^j \quad i = 1, \ldots, N$ # normalization

$\xi^{1:N} \leftarrow resample\left(\tilde{\xi}^{1:N}, \omega^{1:N}\right)$ # resampling

$\hat{X}_k = \frac{1}{N}\sum_{i=1}^{N} \xi^i$

End For

Multiple model adaptive estimation

We propose a Multiple Model Adaptive method to estimate the type of the specific growth rate function $\mu(S_t)$. This growth rate could be either a Monod-type function or a Haldane-type function. The idea of the MMAE approach is to run in parallel a bank of n filters. In this case, a simple approach is to set $n = 2$, that is, one tuned for the Monod-type growth function and another for the Haldane-type growth function. Then, a weighting algorithm is used to assign a weight $p_{i,k}$ to each state estimate by comparing the residuals, that is, the estimated outputs $\hat{y}_{i,k}$ of each filter i with the system's output y_k at each discrete time t_k. These weights are usually computed according to

$$p_{i,k} = f(\hat{y}_{i,k}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\hat{y}_{i,k} - y_k)^2}{2\sigma^2}\right). \tag{9}$$

Eq. (9) represents a normal distribution with mean y_k and variance $\sigma^2 = var(\hat{y}_{i,k})$. Afterwards, these weights are normalized using $p_{i,k} = \frac{p_{i,k}}{\sum_{i=1}^{n} p_{i,k}}$, where n is the number of filters.

The final state estimate is then given by the sum of all the weighted estimates, that is $\hat{x}_t = \sum_{i=1}^{n} p_{i,k}\hat{x}_{i,t}$ where $\hat{x}_{i,t}$ is the estimate given by each filter "i" at time t and $p_{i,k}$ is the corresponding weight. Given the discrete nature of the measurements y_k, the weights are computed only when an output value becomes available. Otherwise, they are kept constant and equal to $p_{i,k}$ along the interval $[t_k \; t_{k+1}]$. The obtained results are given in Subsection *Growth rate estimation using a multiple model adaptive method*. We can use the EKF in a first simulation the UKF in the second and the PF in the third or we can put all three filters together in the same MMAE algorithm, and use the multiple model method to automatically choose which one of these algorithms gives the best estimation of the state and also readapts itself better to the change of the growth function. This was indeed done, and the obtained results are presented in the next section.

Comparison of the algorithms

This section illustrates the performance of each estimation algorithm described in the previous sections through simulations. We implemented the Algorithms 1-7 in MATLAB 2017a within a MMAE method to simulate the obtained estimation architecture. We consider a simulation time interval for the Chemostat model going from zero to $T = 1000$ hours which contains $N_{obs} = 1000$ measurements, that is, N_{obs} iterations of the update step in the filters, whereas for the AM2

and AM2b model, $T = 100$ hours and we have $N_{obs} = 100$ measurements. The simulation time is discretized with a time step of $\delta = 0.1$ [ht]. Thus, between every two successive update steps, there is $N_{sys} = 10$ iterations of the prediction step. The initial state distributions of the models are Gaussian. The obtained results are given in Fig. 2 for the Chemostat model, in Fig. 3 for the AM2 model and in Fig. 4 and Fig. 5 for the AM2b model.

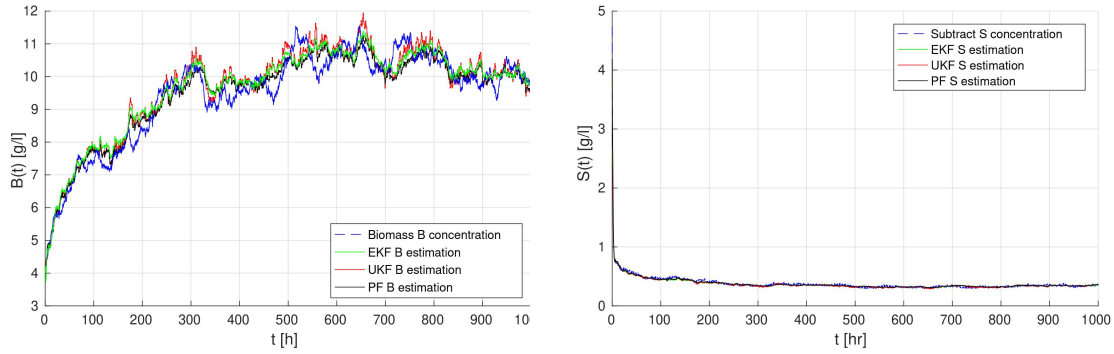

Fig. 2 Time evolution of the estimation of biomass $B(t)$ and substrate $S(t)$ concentrations of the Chemostat model

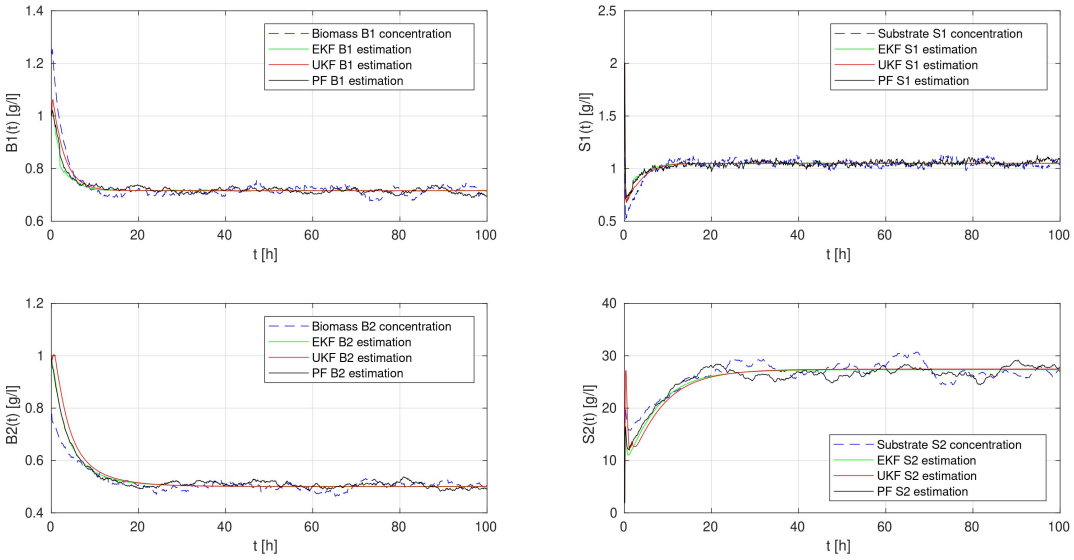

Fig. 3 Time evolution of the estimation of $B_1(t)$, $S_1(t)$, $B_2(t)$ and $S_2(t)$ of the AM2 model representing respectively, the acidogenic bacteria, the organic pollutants, the methanogenic bacteria and the acids

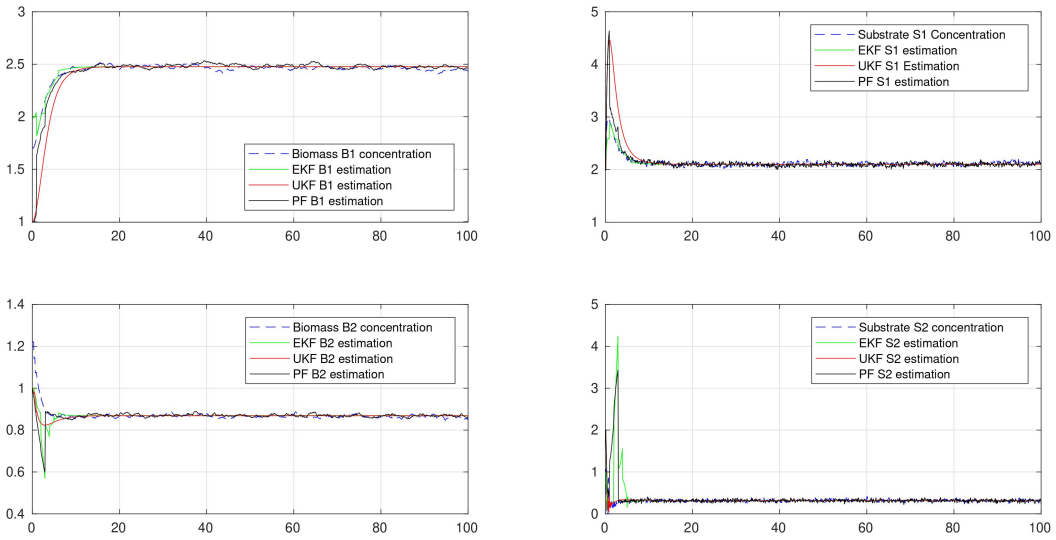

Fig. 4 Time evolution of the estimation of $B_1(t)$, $S_1(t)$, $B_2(t)$ and $S_2(t)$ of the AM2b model representing respectively, the acidogenic bacteria, the organic pollutants, the methanogenic bacteria and the acids

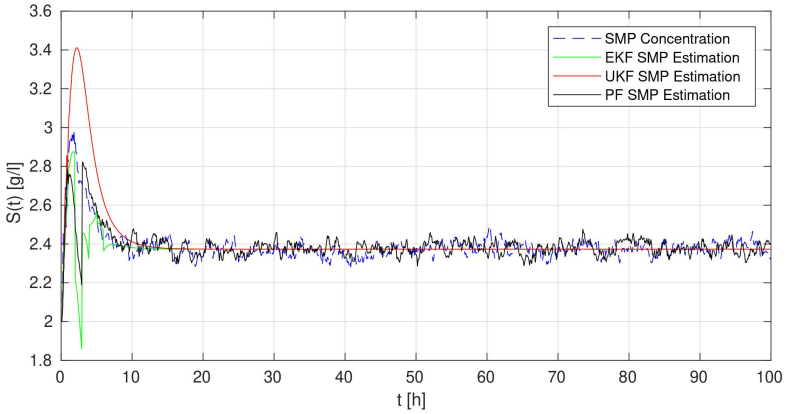

Fig. 5 Time evolution of the estimation of the SMP $S(t)$ in the AM2b model

Performance criteria

To compare the performance of the proposed filters we used the following indicators that are computed using the average of 100 runs:

- The evolution along time of the Root Mean Square (RMS) estimation errors, that is, $RMS(t) = \sqrt{\mathbb{E}\left(\left(\hat{X}_t - X_t\right)^2\right)}$, where $\hat{X}_t$ is the estimation produced by the filter and X_t is the state of the system.
- The average with respect to time of these RMS estimation errors.
- The comparison of the RMS estimation errors with the CRLB. The CRLB corresponds to the best possible filter accuracy that can be achieved for unbiased estimators [40]. We

followed the approach proposed in [42] to compute it.

- The average computation time required by each algorithm in one iteration of simulation.

Comparison

Fig. 2 shows the time evolution of the biomass $B(t)$ and substrate $S(t)$ concentrations and the respective estimated ones produced by three algorithms (EKF, UKF, PF) for the Chemostat model, where it can be seen that all the methods perform well and provide a good estimation. For a better comparison of the performance of the filters, Fig. 6 displays the time evolution of the RMS state estimation errors and the CRLB. The RMS errors and the CRLB in the figure are the results of using 100 Monte Carlo simulations. It can be seen that the PF performs slightly better than the EKF and both of them much better than the UKF.

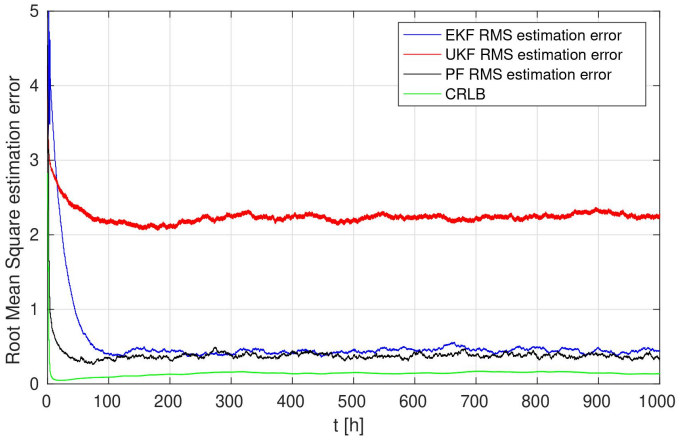

Fig. 6 Time evolution of the RMS state estimation errors of the EKF, the UKF and the PF for the Chemostat model

On the other hand, for the AM2 and AM2b models, it can be seen in Figs. 3, 4, and 5 that the estimations provided by Kalman-based filters stay arround the mean values of the variables while the particle filter produces good estimation results. This can mean that the EKF and the UKF failed to estimate the covariance of the state variables. Fig. 7 and Fig. 8 show the corresponding RMS errors. For the AM2 model, we can see that the estimation errors are almost the same, whilst for the AM2b case, it can be seen that the EKF exhibits an oscillation between the desired equilibrium and another equilibrium of the system. The other two estimation methods produced a higher error than in the AM2 case because of the relatively high variance of the AM2b state variables.

The average of these RMS errors are given in Table 1. Note that the initial condition values and the peak values resulting from the abrupt change of growth rate (presented in the next subsection) are taken into account while computing the average RMS errors which explains the high values shown in the table.

About the computation times, we can see in Table 1 that the EKF algorithm has the lowest computational cost compared to the others. However, all these three algorithms are convenient for a real time implementation using a mid-range computer since our system has a very slow evolution.

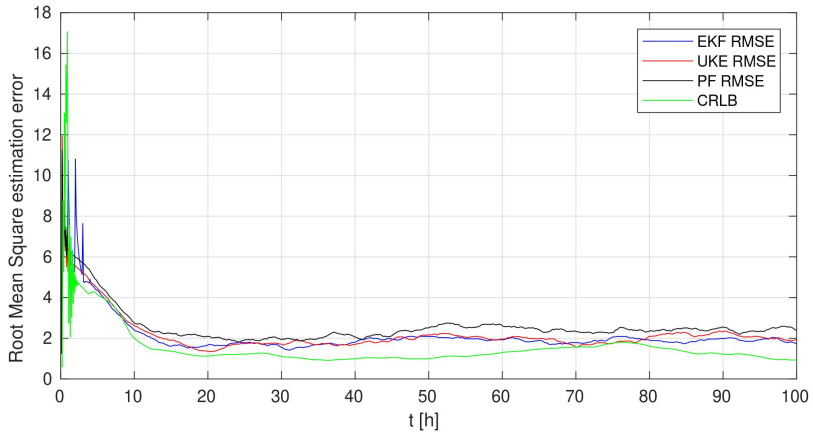

Fig. 7 Time evolution of the RMS state estimation errors of the EKF, the UKF and the PF for the AM2 model

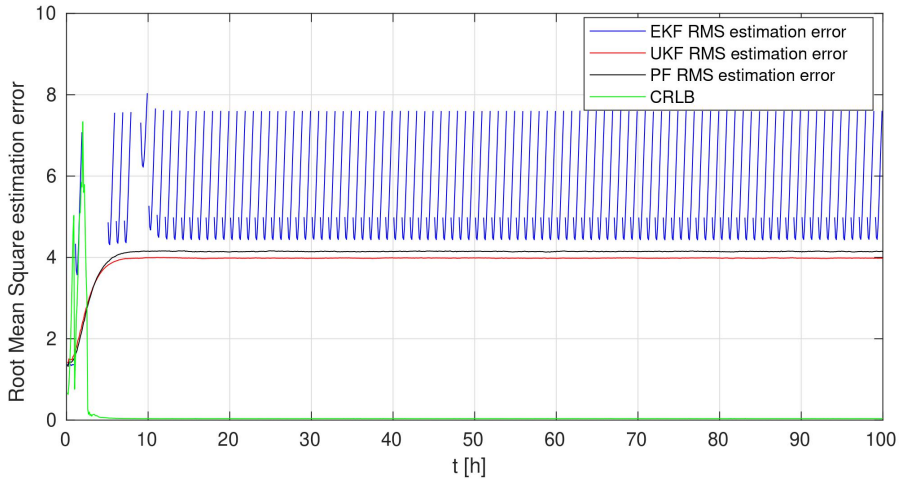

Fig. 8 Time evolution of the RMS state estimation errors of the EKF, the UKF and the PF for the AM2b model

Growth rate estimation using a multiple model adaptive method

The results when the growth rate functions change are shown in Figs. 9-14. In this set of simulations, the type of the growth rate functions is changed at $t = 500$ h for the Chemostat model and at $t = 50$ h for the other models. In the Chemostat model, $\mu(S_t)$ is of Monod-type during the time interval $[0, 500]$ and of Haldane-type during the interval $[500, 1000]$. Whereas in the AM2 model, $\mu_1(S_t)$ is of Monod-type and $\mu_2(S_t)$ is of Haldane-type during the interval $[0, 50]$ then they are interchanged during the interval $[50, 100]$. Moreover, in the AM2b model, the growth rate function in the SMP equation $\mu(S_t)$ is also switched from a Monod-type in $[0, 50]$ to a Haldane-type in $[50, 100]$. Looking at the figures, we can see in the weights plots that the MMAE method always give all the importance to the PF estimation especially in the case of high dimensions.

Table 1. Comparison of the average of the state RMS estimation errors and of the computation time

Model	Criteria	EKF	UKF	PF	MMAE EKF	MMAE UKF	MMAE PF	MMAE EKF & UKF & PF
Chemostat	Average RMS estimation errors	0.5324	2.2460	0.3877	0.9824	14.1705	0.5425	0.5926
AM2		2.142	1.972	2.197	2.432	10.583	2.356	2.384
AM2b		6.072	3.985	4.241	6.642	12.341	4.432	4.451
Chemostat	Average computation time, s	0.020	0.38	0.040	0.023	0.049	0.056	0.189
AM2		0.032	0.045	0.048	0.039	0.061	0.067	0.195
AM2b		0.037	0.053	0.055	0.041	0.064	0.069	0.207

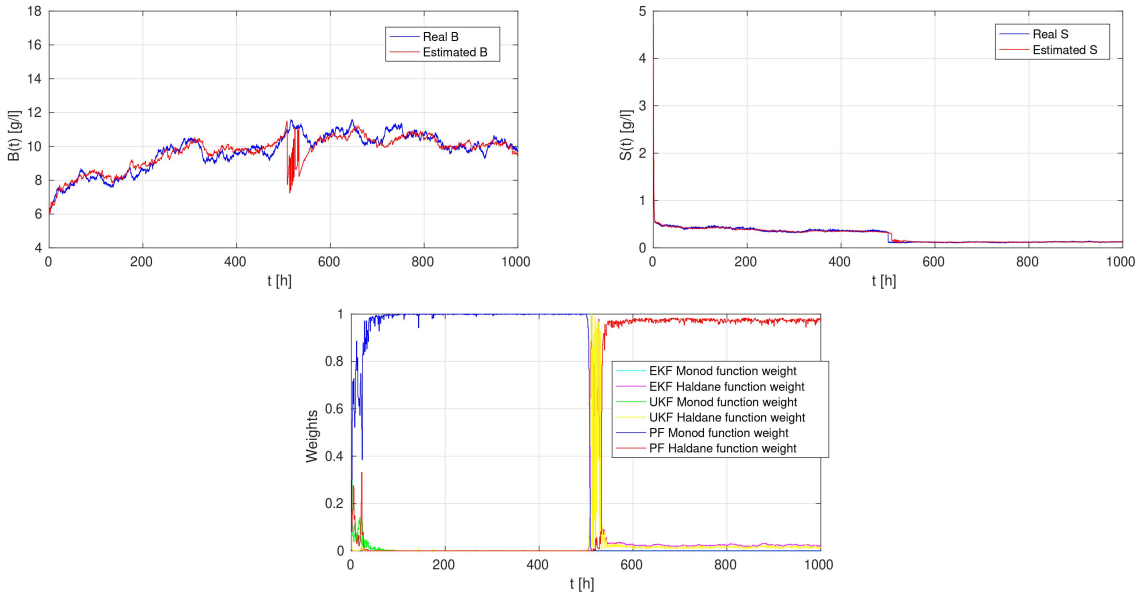

Fig. 9 Estimation of biomass $B(t)$ and substrate $S(t)$ concentrations of the Chemostat model using EKF, UKF and PF in the MMAE method.
Left: Estimation of the biomass $B(t)$, Right: Estimation of the substrate $S(t)$,
Bottom: Weights of the Monod and Haldane growth functions.

To compare the previously used MMAE algorithms, we can compute the average of the RMS estimation errors. These RMS estimation errors are given in Table 1, the smallest RMS estimation error is the one of the MMAE with PF only. In this context, and considering the results of the Monte Carlo simulations, we may conclude that the MMAE-PF algorithm allows a significant reduction in the state estimation errors and gives the best state estimation and the best adaptation to the change of the growth function.

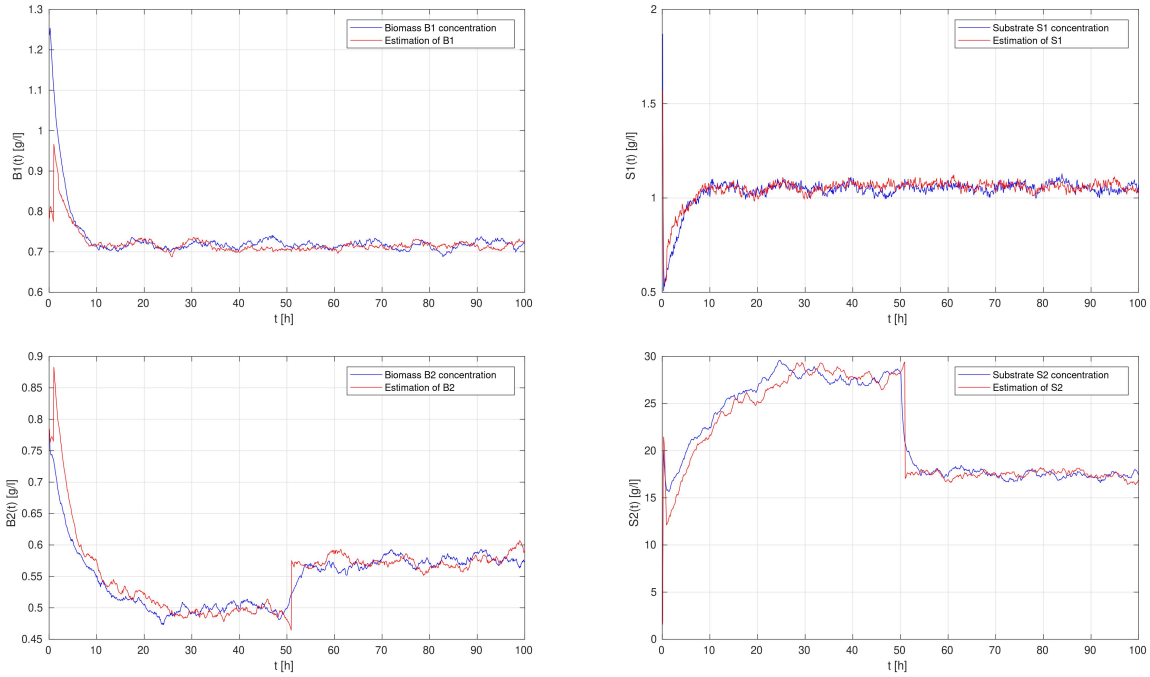

Fig. 10 Estimation of $B_1(t)$, $S_1(t)$, $B_2(t)$ and $S_2(t)$
of the AM2 model using EKF, UKF and PF in the MMAE method

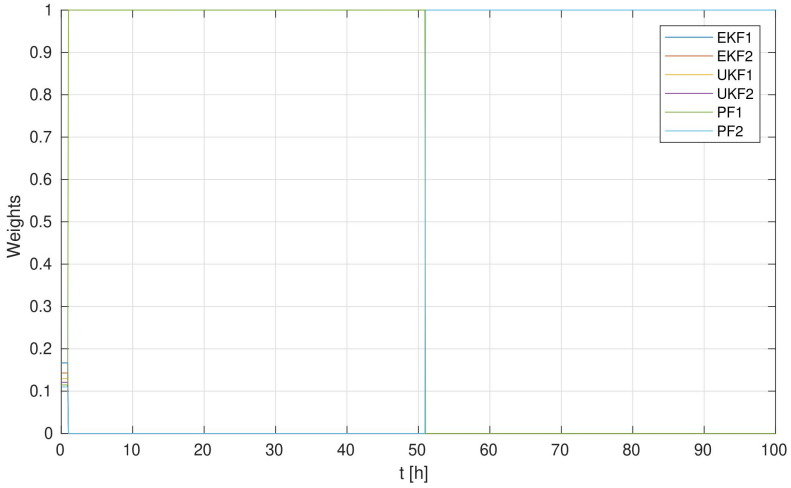

Fig. 11 Weights of the different growth functions of the AM2 model obtained
by using EKF, UKF and PF in the MMAE method

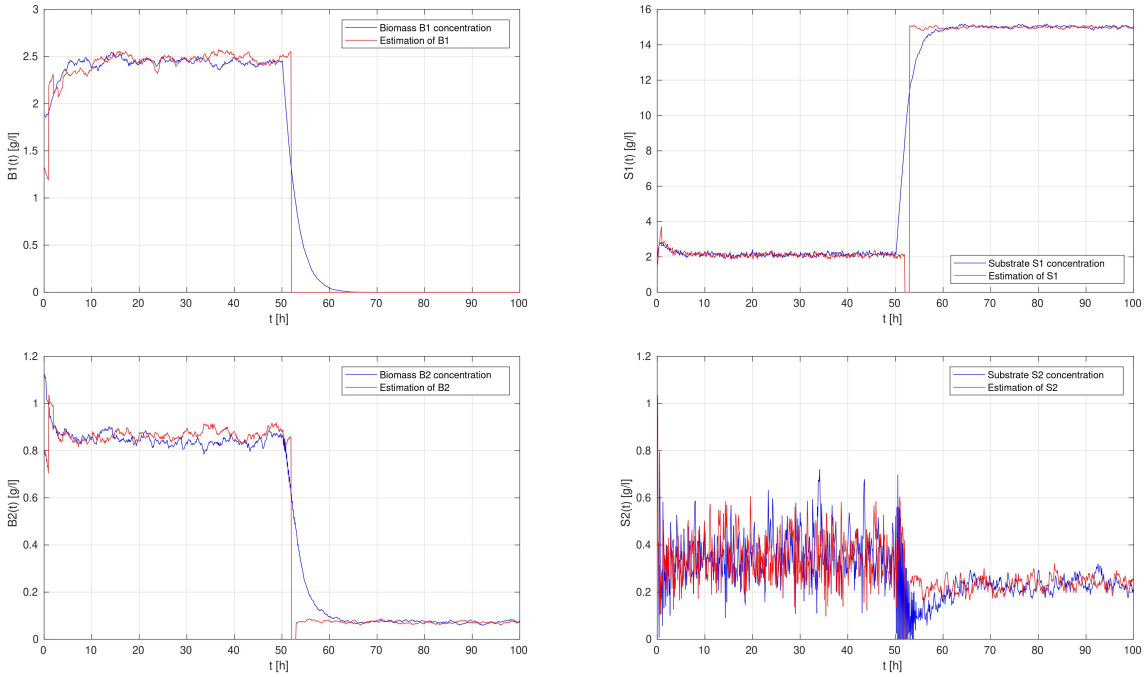

Fig. 12 Estimation of $B_1(t)$, $S_1(t)$, $B_2(t)$ and $S_2(t)$ of the AM2b model using EKF, UKF and PF in the MMAE method

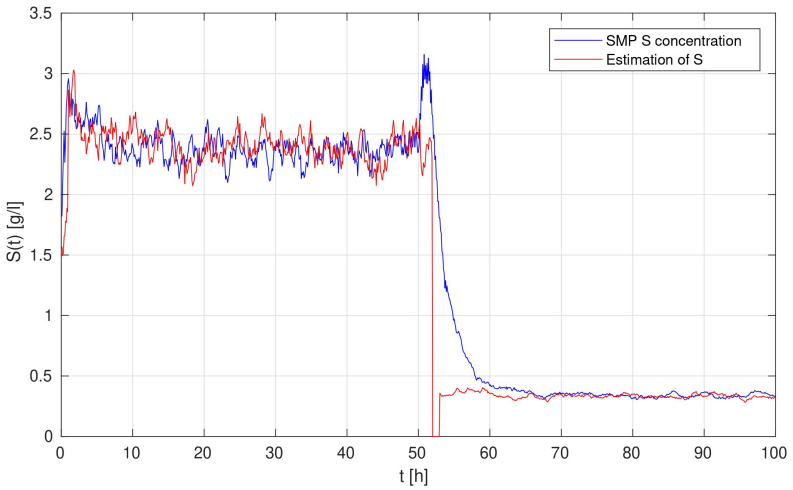

Fig. 13 Estimation of the SMP concentration $S(t)$ of the AM2b model using EKF, UKF and PF in the MMAE method

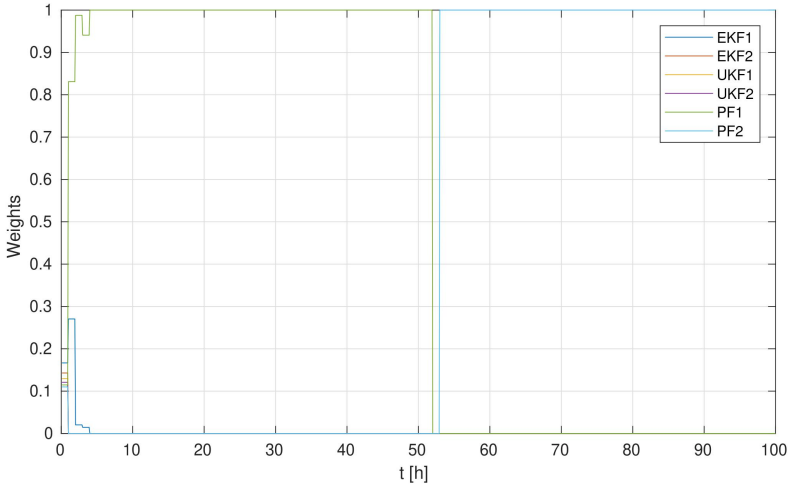

Fig. 14 Weights of the different growth functions of the AM2b model obtained by using EKF, UKF and PF in the MMAE method

Conclusion

In this paper, we presented three methods for the state estimation of stochastic bioreactor models: the Extended Kalman filter, the Unscented Kalman filter and the Particle filter. Some particular adaptations were required to correctly implement these methods. Those included for the EKF the linearization of the models and the substitution of the output equation by an equivalent one to address the problem that the noise does not enter only in an additive form. Also, the addition of the noise variables in the state vector to apply the UKF algorithm, and reducing the dimension of the covariance matrix to obtain a reduced number of sigma points in the case of high dimensional models, and finally the replacement of the proposal density by the Euler approximation of the state equation to simplify the application of the PF. The obtained results were acceptable even in the presence of significant noises and uncertainties acting on the system, and when compared to similar work in the literature. Furthermore, we also compared the filters RMS estimation errors to understand how close or far away they are from the theoretical possible bound. Regarding the problem of growth rate type of the bacteria, we proposed an adaptive multiple-model approach together with the EKF, UKF, and PF to be able to estimate the type of the growth rate functions whether they were a Monod or a Haldane one. Moreover, we used the three filters together in the same MMAE algorithm to choose which one of them gives the best estimation. By comparing the results, we could conclude that the PF provides the best estimation for the system's state and also the smallest error in the growth rate estimation. On another side, the EKF and UKF can have a fast convergence comparing to the PF, which can be seen from the weights plots in the MMAE algorithm when it was applied to the Chemostat model. In spite of this, the PF has proven its superiority in the cases of high dimensional models and high intensity noises. A future study would be the online estimation of the system's parameters together with the state variables using different approaches.

References

1. Abdelkader O. H., A. H. Abdelkader (2019). Modeling Anaerobic Digestion Using Stochastic Approaches, Trends in Biomathematics: Mathematical Modeling for Health, Harvesting, and Population Dynamics, Springer, 373-396.
2. Anderson B. D., J. B. Moore (2012). Optimal Filtering, Prentice-Hall Electrical Engineering Series, Courier Corporation.

3. Andrews J. F. (1968). A Mathematical Model for the Continuous Culture of Microorganisms Utilizing Inhibitory Substrates, Biotechnology and Bioengineering, 10(6), 707-723.
4. Benyahia B. (2012). Modélisation et observation des bioprocédés á membranes: Application á la digestion anaérobie (Unpublished doctoral dissertation).
5. Benyahia B., F. Campillo, B. Cherki, J. Harmand (2012). Particle Filtering for the Chemostat, Proceedings of the 20th Mediterranean Conference on Control and Automation – MED 2012, Barcelona, Spain, July 3-6, 364-371.
6. Benyahia B., T. Sari, B. Cherki, J. Harmand (2013). Anaerobic Membrane Bioreactor Modeling in the Presence of Soluble Microbial Products (SMP) – The Anaerobic Model AM2b, Chemical Engineering Journal, 228, 1011-1022.
7. Bernard O., Z. Hadj-Sadok, D. Dochain, A. Genovesi, J.-P. Steyer (2001). Dynamical Model Development and Parameter Identification for an Anaerobic Wastewater Treatment Process, Biotechnology and Bioengineering, 75(4), 424-438.
8. Bogaerts P., A. V. Wouwer (2003). Software Sensors for Bioprocesses, ISA Transactions, 42(4), 547-558.
9. Campillo F., M. Joannides, I. Larramendy-Valverde (2011). Stochastic Modeling of the Chemostat, Ecological Modelling, 222(15), 2676-2689.
10. Chang L., B. Hu, A. Li, F. Qin (2013). Transformed Unscented Kalman Filter, IEEE Transactions on Automatic Control, 58(1), 252-257.
11. de Assis A. J., R. Maciel Filho (2000). Soft Sensors Development for On-line Bioreactor State Estimation, Computers & Chemical Engineering, 24(2), 1099-1103.
12. Didi I., H. Dib, B. Cherki (2014). An Invariant Observer for a Chemostat Model, Automatica, 50(9), 2321-2326.
13. Didi I., H. Dib, B. Cherki (2015). A Luenberger-type Observer for the AM2 Model, Journal of Process Control, 32, 117-126.
14. Dochain D., L. Chen (1992). Local Observability and Controllability of Stirred Tank Reactors, Journal of Process Control, 2(3), 139-144.
15. Douc R., O. Cappé (2005). Comparison of Resampling Schemes for Particle Filtering, Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis (ISPA 2005), 64-69.
16. Gauthier J., H. Hammouri, S. Othman (1992). A Simple Observer for Nonlinear Systems Applications to Bioreactors, IEEE Transactions on Automatic Control, 37(6), 875-880.
17. Gouzé J.-L., A. Rapaport, M. Z. Hadj-Sadok (2000). Interval Observers for Uncertain Biological Systems, Ecological Modelling, 133(1), 45-56.
18. Hadj-Abdelkader O., A. Hadj-Abdelkader (2017). State Estimation for a Chemostat Model by the Unscented Kalman Filtering Approach, Electrotehnica, Electronica, Automatica, 65(2), 104-108.
19. Hadj-Abdelkader O., A. Hadj-Abdelkader (2019). Estimation of Substrate and Biomass Concentrations in a Chemostat Using an Extended Kalman Filter, Int J Bioautomation, 23(2), 215-232.
20. Hassani V., A. P. Aguiar, M. Athans, A. M. Pascoal (2009). Multiple Model Adaptive Estimation and Model Identification Using a Minimum Energy Criterion, Proceedings of the American Control Conference (ACC'09), 518-523.
21. Hassani V., A. P. Aguiar, A. M. Pascoal, M. Athans (2009). A Performance Based Model-set Design Strategy for Multiple Model Adaptive Estimation, Proceedings of the European Control Conference (ECC 2009), 4516-4521.
22. Haugen F., R. Bakke, B. Lie (2014). State Estimation and Model-based Control of a Pilot Anaerobic Digestion Reactor, Journal of Control Science and Engineering, 2014, Article ID 572621.

23. Hocalar A., M. Türker, C. Karakuzu, U. Yüzgeç (2011). Comparison of Different Estimation Techniques for Biomass Concentration in Large Scale Yeast Fermentation, ISA Transactions, 50(2), 303-314.
24. Julier S. J., J. K. Uhlmann (1997). New Extension of the Kalman Filter to Nonlinear Systems, Proceedings of the Signal Processing, Sensor Fusion, and Target Recognition VI, Aerosense'97, Orlando, FL, United States, Vol. 3068, 182-193.
25. Kalchev B., I. Simeonov, N. Christov (2011). Kalman Filter Design for a Second-order Model of Anaerobic Digestion, Int J Bioautomation, 15(2), 85-100.
26. Kloeden P. E., E. Platen (1992). Numerical Solution of Stochastic Differential Equations, Springer-Verlag.
27. Liu J. S., R. Chen (1998). Sequential Monte Carlo Methods for Dynamic Systems, Journal of the American Statistical Association, 93(443), 1032-1044.
28. Lopez I., L. Borzacconi (2009). Modelling a Full Scale UASB Reactor Using a COD Global Balance Approach and State Observers, Chemical Engineering Journal, 146(1), 1-5.
29. Moireau P., D. Chapelle (2011). Reduced-order Unscented Kalman Filtering with Application to Parameter Identification in Large-dimensional Systems, ESAIM: Control, Optimisation and Calculus of Variations, 17(2), 380-405.
30. Monod J. (1949). The Growth of Bacterial Cultures, Annual Reviews in Microbiology, 3(1), 371-394.
31. Padilla L. E., C. W. Rowley (2010). An Adaptive-covariance-rank Algorithm for the Unscented Kalman Filter, Proceedings of the 49th IEEE Conference on Decision and Control, 1324-1329.
32. Patarinska T., V. Trenev, S. Popova (2010). Software Sensors Design for a Class of Aerobic Fermentation Processes, Int J Bioautomation, 14(2), 99-118.
33. Ponomareva K., P. Date, Z. Wang (2010). A New Unscented Kalman Filter with Higher Order Moment-matching, Proceedings of Mathematical Theory of Networks and Systems (MTNS 2010), Budapest, 1-5.
34. Rocha-Cózatl E., M. Sbarciog, L. Dewasme, J. Moreno, A. V. Wouwer (2015). State and Input Estimation of an Anaerobic Digestion Reactor Using a Continuous-discrete Unknown Input Observer, IFAC-PapersOnLine, 48(8), 129-134.
35. Särkkä S. (2007). On Unscented Kalman Filtering for State Estimation of Continuous-time Nonlinear Systems, IEEE Transactions on Automatic Control, 52(9), 1631-1641.
36. Särkkä S. (2013). Bayesian Filtering and Smoothing, Vol. 3, Cambridge University Press.
37. Sbarciog M., J. Moreno, A. V. Wouwer (2014). Application of Super-twisting Observers to the Estimation of State and Unknown Inputs in an Anaerobic Digestion System, Water Science and Technology, 69(2), 414-421.
38. Smith H. L., P. E. Waltman (1995). The Theory of the Chemostat: Dynamics of Microbial Competition, Cambridge University Press.
39. Sotomayor O. A., S. W. Park, C. Garcia (2002). Software Sensor for On-line Sstimation of the Microbial Activity in Activated Sludge Systems, ISA Transactions, 41(2), 127-143.
40. Tichavsky P., C. H. Muravchik, A. Nehorai (1998). Posterior Cramér-Rao Bounds for Discrete-time Nonlinear Filtering, IEEE Transactions on Signal Processing, 46(5), 1386-1396.
41. Toumi S., M. Chebbi, F. Campillo (2019). Stochastic Modeling for Biotechnologies Anaerobic Model AM2b, Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées, INRIA, Vol. 28 - 2017 - Mathematics for Biology and the Environment, 13-23.
42. Van Trees H. L., K. L. Bell (2007). Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking, Wiley-IEEE Press.

43. Wan E. A., R. Van Der Merwe (2000). The Unscented Kalman Filter for Nonlinear Estimation, Proceedings of the IEEE Symposium of the Adaptive Systems for Signal Processing, Communications, and Control, 153-158.

Assoc. Prof. António Pedro Aguiar, Ph.D.
E-mail: pedro.aguiar@fe.up.pt

António Pedro Aguiar received the Licenciatura, M.Sc. and Ph.D. in Electrical and Computer Engineering from the Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal in 1994, 1998 and 2002, respectively. Currently, Dr. Aguiar holds an Associate Professor position with the Department of Electrical and Computer Engineering, Faculty of Engineering, University of Porto. From 2002 to 2005, he was a Post-doctoral Researcher at the Center for Control, Dynamical-systems, and Computation at the University of California, Santa Barbara. From 2005 to 2012, he was a Senior Researcher with the Institute for Systems and Robotics at IST (ISR/IST), and an invited Assistant Professor with the Department of Electrical and Computer Engineering, IST. His research interests include modeling, control, navigation, and guidance of autonomous robotic vehicles, nonlinear control, switched and hybrid systems, tracking, path-following, performance limitations, nonlinear observers, the integration of machine vision with feedback control, networked control, and coordinated/cooperative control of multiple autonomous robotic vehicles. Further information related to Dr. Aguiar's research can be found at the Cyber-physical Control Systems and Robotics Lab and http://paginas.fe.up.pt/apra.

Res. Oussama Hadj-Abdelkader, Ph.D.
E-mail: hadjabdelkader.oussama@gmail.com

Oussama Hadj-Abdelkader was born in Tlemcen, Algeria, in 1991. He graduated from the Electrical Engineering Department at the University of Tlemcen, Algeria, in 2013. Hadj-Abdelkader obtained his Ph.D. degree from the same university. He is currently working as a Researcher in SYSTEC Laboratory at the University of Porto. His main research interests are Monte Carlo methods, stochastic systems identification, control and modeling.