Recursive Parameter and State Estimation for a Mining Industry Process

Rodrigo A. López*, Juan I. Yuz*, Werner U. Creixell*, and Jaime A. Ramírez†

Abstract— **In this paper we consider the problem of recursive state and parameter estimation of a nonlinear system. We propose an approach where we combine Particle Filter (for state estimation) and Recursive Least Squares (for parameter estimation) The class of nonlinear systems is motivated by a real process of the copper mining industry. The proposed approach is tested with real data.**

I. Introduction

The problem of state and parameter estimation has a long history in Engineering and Statistics [1], [2], [3]. For state estimation the standard tool is the Kalman Filter [1], which is optimal for a linear system with Gaussian noise. For nonlinear systems one can use, for example, the Bayesian estimator for the state vector using input/output measurements of the given system. We consider a state-space model:

$$x_{t+1} = f(x_t, u_t, w_t) \tag{1}$$

$$y_t = g(x_t, v_t) \tag{2}$$

where x_t is the state vector, u_t is the input, y_t is the measured output and w_t, v_t are white noise processes. In particular, we are interested in obtaining an updated estimate of the state vector every time a new measurement is available. This is, if we have an *a-priori* probability density function (pdf) $p(x_t|Y_{t-1})$ of the state vector at time t given the past data, we are interested in $p(x_t|Y_t)$, the *a-posteriori* pdf of the state vector x_t, given the data $Y_t = \{y_i\}_{i=0}^{t}$. Using Bayes's theorem we have that:

$$p(x_t|Y_t) = \frac{p(y_t|x_t)p(x_t|Y_{t-1})}{p(y_t|Y_{t-1})} \tag{3}$$

where the denominator corresponds only to a scaling factor, i.e.,

$$p(y_t|Y_{t-1}) = \int p(y_t|x_t)p(x_t|Y_{t-1})dx_t \tag{4}$$

On the other hand, a new *a-priori* estimate of the state at the next time instant is given by:

$$p(x_{t+1}|Y_t) = \int p(x_{t+1}|x_t)p(x_t|Y_t)dx_t \tag{5}$$

Equations (3)-(5) require an initial distribution $p(x_0|Y_{-1}) = p(x_0)$ for the initial state.

The integrals (4)-(5) are, in general, hard to solve analytically. In the case of linear state-space model with additive Gaussian noise, the solution corresponds to the well known Kalman Filter. In the case of a general nonlinear system, such as (1)-(2) several approximate solutions have been proposed in the literature. For example, the Extended Kalman Filter (EKF) [1], [4] utilizes a linearized model along the estimated state trajectory, and uses the nonlinear model to obtain an *a-priori* state estimate. On the other hand, the Unscented Kalman Filter (UKF), utilizes a nonlinear transformation to approximate the pdf of the state at each time instant by a Gaussian distribution [4]. Notice that both EKF and UKF do not aim at directly solving (4) and (5). In fact, they only try to apply the KF algorithm to the nonlinear case. As a consequence, none of the optimal properties of the KF can be guaranteed. The links between EKF and UKF are discussed, for example, in [5].

A direct approach to the nonlinear state estimation problem is given by the, so called, Particle Filter (PF) [4], [6]. This is a simulation based algorithm where the pdf of the state vector is approximated by a large set of realizations called *particles*. These particles are then used to numerically approximate the integrals (4)-(5). Even though PF is computationally more expensive than, for example, EKF, it usually provides more accurate results (see, for example, page 469 [4]) . The Marginalized Particle Filter (MPF) has been proposed as an alternative to reduce the computational load of the PF [7], [8]. An alternative to compute the conditional probabilities in (4)-(5) is based in vector quantization and the, so called, scenario based approach [9].

For the parameter estimation problem several algorithms have been proposed in System Identification, such as Maximum Likelihood [3] and Prediction Error Methods [10]. However, the most simple algorithm is Least Squares (LS) that can be extended for an online implementation, namely, Recursive Least Squares (RLS) [11]. In this case, the information from new available measurements is used to update the parameter estimate, allowing to model time-varying characteristics of the system of interest.

In this paper we are interested in obtaining a model using data collected from a mining industry process and, using that model, to estimate the state and output distributions. The model structure is motivated by a real processes from the copper mining industry. The main interest is to model the time-varying characteristics of the system, thus, we use RLS for parameter estimation. On the other hand, the available measurements are restricted to the interval $[0, 1]$. As a consequence, the Gaussian assumption underlying in the usual KF may be not appropriate. Thus, we will use PF to estimate the pdf of the state and output signals. To test

*Electronic Engineering Department, Universidad Técnica Federico Santa María (UTFSM), CHILE. Email: rodrigo.lopeza@alumnos.utfsm.cl

†Aplik S.A., CHILE. http://www.aplik.com

Work supported by UTFSM and CONICYT-Chile through grants FONDECYT 1110586 and ACT-053

the proposed approach, real data has been provided by the company Aplik S.A. (`http://www.aplik.com`), which has succesfully developed advanced control applications for the mining industry.

II. The particle filter and RLS estimation

In this section we present the methods for recursive state and parameter estimation used in the paper. Even though the parameters can be included as states to be estimated [12], the PF has, in general, problems when estimating constant parameters [13]. As a consequence, we will split the problem into two parts: When a new measurement is available, given an estimate of the parameters we estimate the states, and then, given the state estimates, we estimate the parameters.

A. Particle filter

The PF is a recursive algorithm that numerically approximates the integrals (4)-(5) of the Bayesian Estimator [4]. It is also referred to as Sequential MonteCarlo methods [14]. The PF has the following main steps:

(PF 1) Initialize the particles for $i = 1, 2, ..., N_p$ generating an initial set of *a-posteriori* state estimates $\{x^{(i)}_{0|0}\}^{N_p}_{i=1}$ from a given initial pdf $p(x_0)$.

(PF 2) For each particle, compute an *a-priori* estimate $x^{(i)}_{t|t-1}$ of the state from the previous *a-posteriori* estimates $x^{(i)}_{t-1|t-1}$ and realizations of the noise process w_t. For this, we use the process equation (1), i.e.

$$x^{(i)}_{t|t-1} = f(x^{(i)}_{t-1|t-1}, u_{t-1}, w^{(i)}_{t-1}) \qquad (6)$$

(PF 3) Compute the likelihoods $q^{(i)}_t$ from the output equation (2) and the available measurement y_t, i.e.,

$$q^{(i)}_t = p(y_t|x^{(i)}_{t|t-1}) \qquad (7)$$

(PF 4) Normalize the weights as:

$$\tilde{q}^{(i)}_t = \frac{q^{(i)}_t}{\sum_{i=1}^{N_p} q^{(i)}_t} \qquad (8)$$

(PF 5) The process of *resampling* allows one to obtain the *a-posteriori* estimate of state $x^{(i)}_{t|t}$. This is given by:

$$p(x^{(i)}_{t|t} = x^{(i)}_{t|t-1}) = \tilde{q}^{(i)}_t \qquad (9)$$

where $x^{(i)}_{t|t-1}$ and $x^{(i)}_{t|t}$ are the *a-priori* and the *a-posteriori* estimates of the states.

(PF 6) To obtain a single state from the set of particles estimate we can, for example, compute the expected value of $E(x_t|y_t)$ as the empirical mean, i.e.,

$$\hat{x}_{t|t} = E(x_t|y_t) \approx \frac{1}{N_p}\sum_{i=1}^{N_p} x^{(i)}_{t|t} \qquad (10)$$

(PF 7) Set $t = t + 1$ and iterate from step 2.

B. Recursive Least Squares

Consider the general linear regression model:

$$y_t = \varphi_t^T \theta \qquad (11)$$

where θ contains the parameters to be estimated and φ_t contains the regressors. We are interested in the Least Squares (LS) estimate

$$\hat{\theta} = \text{Arg}\min_{\theta} \sum_{t=1}^{N} (\varepsilon_t(\theta))^2 = (\Phi^T\Phi)^{-1}\Phi^T Y \qquad (12)$$

where $\varepsilon_t(\theta) = y_t - \varphi_t^T\theta$, and

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \qquad \Phi = \begin{pmatrix} \varphi_1^T \\ \vdots \\ \varphi_N^T \end{pmatrix} \qquad (13)$$

LS provides an off-line method for parameter identification. For on-line identification we can utilize the Recursive Least Squares algorithm (RLS) [11]. This algorithm utilizes the information of each new available measurement to update the parameter estimate. The equations are the following:

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t(y_t - \varphi_t^T\hat{\theta}_{t-1}) \qquad (14)$$

$$K_t = P_t\varphi_t = \frac{P_{t-1}\varphi_t}{1 + \varphi_t^T P_{t-1}\varphi_t} \qquad (15)$$

$$P_t = P_{t-1} - \frac{P_{t-1}\varphi_t\varphi_t^T P_{t-1}}{1 + \varphi_t^T P_{t-1}\varphi_t} \qquad (16)$$

To improve the adaption of the RLS estimate to changes in the parameters, a forgetting factor (FF) λ can be included in the gain and covariance update equation (15)-(16) as follows [15]

$$K_t = \frac{P_{t-1}\varphi_t}{\lambda + \varphi_t^T P_{t-1}\varphi_t} \qquad (17)$$

$$P_t = \frac{1}{\lambda}\left(P_{t-1} - \frac{P_{t-1}\varphi_t\varphi_t^T P_{t-1}}{\lambda + \varphi_t^T P_{t-1}\varphi_t}\right) \qquad (18)$$

This corresponds to an exponential weighting factor that increases the relative weight of recent data compared to past data. The FF is generally chosen as $\lambda \in [0.95, 0.99]$. We will refer to this approach as RLS-FF.

Further modifications can be included in the RLS to ensure that the gain K_t does not go to zero and the eigenvalues of the covariance matrix P_t remain bounded. We consider the Exponential Forgetting and Resetting Algorithm (EFRA) [16], where the gain and covariance update equations (17)-(18), are replaced by:

$$K_t = \frac{\alpha P_{t-1}\varphi_t}{1 + \varphi_t^T P_{t-1}\varphi_t} \qquad (19)$$

$$P_t = \frac{1}{\lambda}P_{t-1} - \frac{\alpha P_{t-1}\varphi_t\varphi_t^T P_{t-1}}{1 + \varphi_t^T P_{t-1}\varphi_t} + \beta I - \delta P_{t-1}^2 \qquad (20)$$

where λ is the forgetting factor, generally $\lambda \in [0.9, 0.99]$; α adjusts the gain of the least squares algorithm, generally $\alpha \in [0.1, 0.5]$; β is a small constant related to the minimum eigenvalue of P, generally $\beta \in [0, 0.01]$; δ is a small constant that is inversely related to the maximum eigenvalue of P, generally $\delta \in [0, 0.01]$. We will refer to this approach as RLS-EFRA.

C. The PF combined with Recursive Least Squares

In this subsection we present an approach that combines PF and RLS to estimate the state and parameters of the system. Motivated by the process described later in section IV, we restrict ourselves to the following class of systems:

$$x_{t+1} = A_t x_t + B_t u_t + w_t \tag{21}$$
$$y_t = h(x_t) + v_t \tag{22}$$

where x_t is the state vector, A_t is the matrix of state (of a linear system), B_t is a input matrix (of a linear system), $h(x_t)$ a general known nonlinear function, u_t and y_t, are input/output signals, w_t and v_t are process and measurements noises respectively where $w_t \sim N(0, Q)$ and $v_t \sim N(0, R)$, where Q and R are assumed known.

We combine PF and RLS in an iterative algorithm. This is, given an estimate of the parameter in model (21)-(22) (i.e $\hat{A}_t$ and $\hat{B}_t$), we apply PF to obtain a state estimate. Then, given the state estimate, we apply RLS to equation (21) in the form of a linear regression, in order to obtain an updated estimate of the parameters.

The steps of the proposed algorithm are as follow:

1) Initialization: An initial estimate is given for the parameter vector, i.e, $\hat{\theta}_0 = \begin{pmatrix} \hat{A}_0 & \hat{B}_0 \end{pmatrix}^T$, an initial set of particles is generated (as in step (PF 1)), and initial values for P_0, λ, α, β and δ are set for RLS and its variants.
2) State estimation: Given $\hat{A}_{t-1}$, $\hat{B}_{t-1}$ we perform the steps (PF 2)-(PF 6) to obtain a state estimate $\hat{x}_{t|t}$.
3) Parameter estimation: Given $\hat{x}_{t|t}$ and $\hat{x}_{t-1|t-1}$, we use the linear regression:
$$\hat{x}_{t|t}^T = \begin{pmatrix} \hat{x}_{t-1|t-1}^T & u_{t-1} \end{pmatrix} \begin{pmatrix} \hat{A}_{t-1}^T \\ \hat{B}_{t-1}^T \end{pmatrix} \tag{23}$$
to obtain an updated estimate of
$$\hat{\theta}_{t-1} = \begin{pmatrix} \hat{A}_{t-1}^T \\ \hat{B}_{t-1}^T \end{pmatrix} \tag{24}$$
using either RLS (equations (14)-(16)), RLS-FF (equations (14) and (17)-(18)), or RLS-EFRA (equations (14) and (19)-(20))
4) Set $t = t + 1$ and go back to step 2

III. Example

In this section we consider an example to illustrate the proposed approach for recursive parameter and state estimation. We consider the system given by:

$$x_{t+1} = a_t x_t + b_t u_t + w_t \tag{25}$$
$$y_t = x_t^2 + v_t \tag{26}$$

where w_t y v_t are Gaussian noises with $w_t \sim N(0, 0.4)$ y $v_t \sim N(0, 0.3)$; u_t is the input and y_t is the output of system. In this case $u_t = 2$. The example considers $M = 3000$ samples, and the parameter of the system a_t y b_t are time-varying:

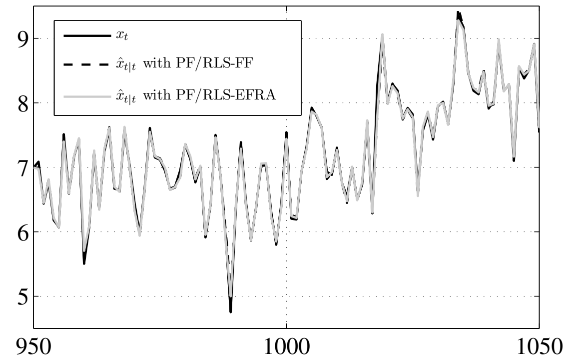

Fig. 1. State Estimate using PF/RLS-FF and PF/RLS-EFRA in the interval $[950, 1050]$

$$a_t = \begin{cases} 0.4 & \text{if } t \leq \frac{M}{3} \\ 0.5 & \text{if } t > \frac{M}{3} \end{cases} \tag{27}$$
$$b_t = \begin{cases} 2 & \text{if } t \leq \frac{2M}{3} \\ 0.5 & \text{if } t > \frac{2M}{3} \end{cases} \tag{28}$$

We consider simultaneous parameter and state estimation of state x_t and the parameter, a_t and b_t using PF/RLS-FF and PF/RLS-EFRA algorithms. Considering the linear regression model as in (11):

$$\hat{x}_{t+1|t+1} = \begin{pmatrix} \hat{x}_{t|t} & u_t \end{pmatrix} \begin{pmatrix} a_t \\ b_t \end{pmatrix} \tag{29}$$

The algorithms use the following initial conditions:

1) PF/RLS-FF: We arbitrarily choose $P_0 = 5I_2$, $\hat{\theta}_0 = \begin{pmatrix} 0.1 & 0.5 \end{pmatrix}^T$ and $\lambda = 0.99$
2) PF/RLS-EFRA: We We arbitrarily choose $P_0 = 5I_2$, $\hat{\theta}_0 = \begin{pmatrix} 0.1 & 0.5 \end{pmatrix}^T$, $\lambda = 0.99$, $\alpha = 0.5$, $\beta = 0.005$ and $\delta = 0.005$.

In both cases, we use $x_{0|0}^{(i)} \sim N(0, 0.4)$ and $N_p = 100$ particles.

Figure 1 shows the state estimation performance using PF/RLS-FF and PF/RLS-EFRA. Similar results are obtaining using both algorithms. The RMS error using PF/RLS-FF is 0.0898 and the RMS error using PF/RLS-EFRA is 0.0910.

Figure 2 shows the parameter estimation performance using RLS-FF and RLS-EFRA algorithm. RLS-FF is more sensitive to parameter's changes. RLS-EFRA is also able to follow the changes in parameters a_t and b_t. For a_t, the RMS error using RLS-FF is 0.1645, and using RLS-EFRA the RMS error is 0.1327. For other hand, the parameter estimation b_t, the RMS error using RLS-FF is 0.4422 and using RLS-EFRA the RMS error is 0.4190.

IV. Application to real process data

In this section we apply the proposed recursive parameter and state estimation approach to real data. The available data correspond to 6 output signals of a process of the mining industry. No manipulated input is available from the

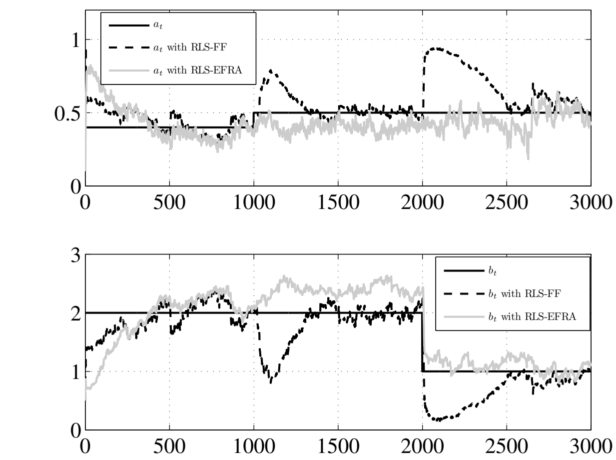

Fig. 2. Parameter estimation a_t and b_t using RLS-FF and RLS-EFRA algorithm in the interval $[0, 3000]$

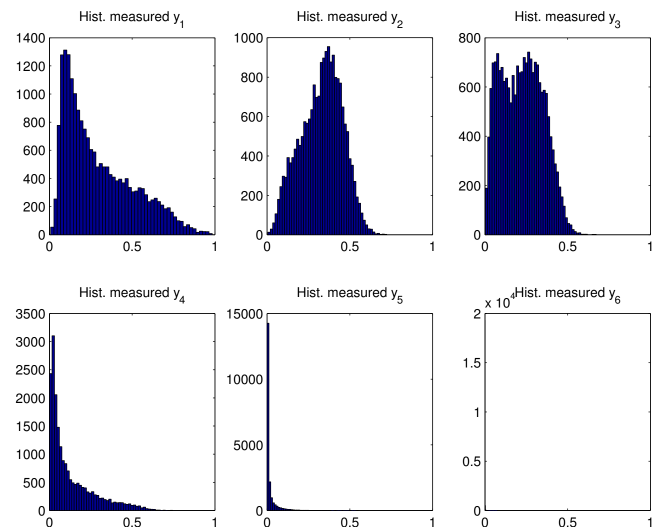

Fig. 3. Histograms of the measurements provided by Aplik S.A. ($M = 20000$ data points).

process. The data was provided by the company Aplik S.A. The process is known to exhibit time-varying characteristics. Thus, for the sake of simplicity we assume a state-space model of order 6 (the same as the number of outputs) with linear dynamics and time-varying matrices, i.e.

$$x_{t+1} = A_t x_t + B_t + w_t \tag{30}$$

In (30) the process noise is assumed Gaussian, zero-mean, and having covariance $0.01 \cdot I_6$.Note that we assume a constant scalar unit input, thus, the matrix B_t can be interpreted as the (time-varying) mean of the process noise. The matrices are estimated using RLS-EFRA, given the estimates of the states, using the (noise-free) linear regression equation,

$$\hat{x}_{t+1|t+1} = \begin{pmatrix} A_t & B_t \end{pmatrix} \begin{pmatrix} \hat{x}_{t|t} \\ 1 \end{pmatrix} \tag{31}$$

where the states are given by the PF.

The main interest of modeling the process is to characterize the pdf of the output variables from the available data. The available measurements correspond to percentages of copper cathodes that are analyzed for quality control purposes. As a consequence, they take values in the interval $[0, 1]$ and add up to 1. Thus, we assume a nonlinear function in the output equation having the form

$$\hat{y}_t = \frac{1}{\sum_{\ell=1}^{6} \hat{x}_{\ell,t}} \hat{x}_t + v_t \tag{32}$$

where $x_{\ell,t}$ represents each individual state. The measurement noise is assumed Gaussian, zero-mean, and having covariance $R = \text{diag}\{0.05, 0.02, 0.02, 0.02, 0.001, 0.001\}$. This covariance matrix was previously estimated as the empirical covariance of a batch of data.

The states (and outputs) are estimated (filtered) using PF, given an estimate of the matrices in the process equation (30). The pdf of the output can then be estimated from the histogram of the output particles.

The available data from the real process contains $M = 20000$ points. The associated histograms are shown in Figure 3 (Note that all the measurements $y_{6,t}$ are close to zero). These histograms are obtained *off-line* using the given data. We obtain an *on-line* estimate of these output distributions using the state estimate given by the PF.

For the state estimation we use PF with $N_p = 1000$ particles. For the parameter estimation, we first use RLS-FF, with $\lambda = 0.995$. In this case, Figure 4 shows the state estimates given by the PF. Notice that, given that the data is a time series collected from a real process, there are no *true* state to compare with. However, the state estimates remain bounded and are smoother than the raw data.

Figure 5 shows the magnitude of the eigenvalues of the estimated matrix A_t. We can see that the combined PF / RLS-FF is able to model time-varying characteristics in the process dynamics.

To estimate *on-line* the output pdf, we use equation (32) (without the noise term) to map the *a-posteriori* state particles given by the PF to the output. Figures 6 and 7 show the associated histograms for $t = 10000$ and $t = 20000$, respectively. We notice that these histograms are similar to the ones in Figure 3 (obtained *off-line*), and they reflect the time-varying nature of the real process. Finally, Figure 8 shows the location of the output particles for different pairs of output signals as a way to illustrate the multivariable output distribution.

We now compare the previous results to the use of RLS-EFRA, with $\lambda = 0.995$, $\alpha = 0.5$, $\beta = 0.01$, and $\delta = 0.001$. Figure 9 shows the state estimates given by the PF. Figures 10 and 11 show the magnitude of the eigenvalues of the estimated matrix A_t (there are complex conjugate pairs) and the components of matrix B_t, respectively.

On the other hand, the online histograms obtained using RLS-EFRA (obtained *off-line*) and the ones in Figures 6 and 7 (obtained using RLS-FF), and they are able to adequately

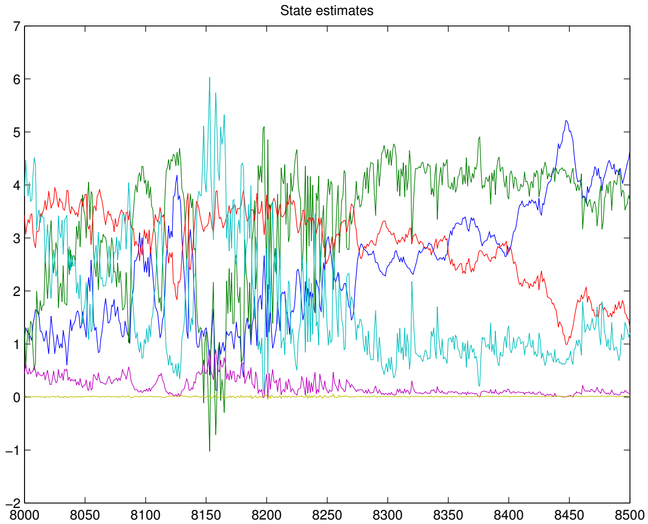

Fig. 4. State estimates given by the PF combined with RLS-FF in the interval $[8000, 8500]$.

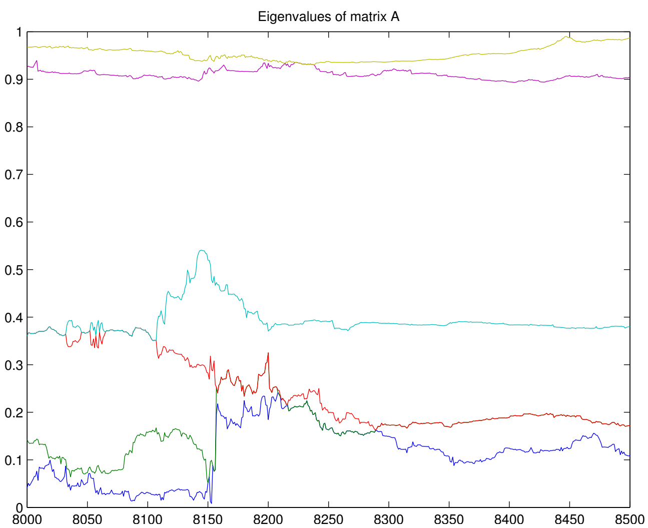

Fig. 5. Magnitude of the eigenvalues of the estimated matrix A_t when using RLS-FF in the interval $[8000, 8500]$

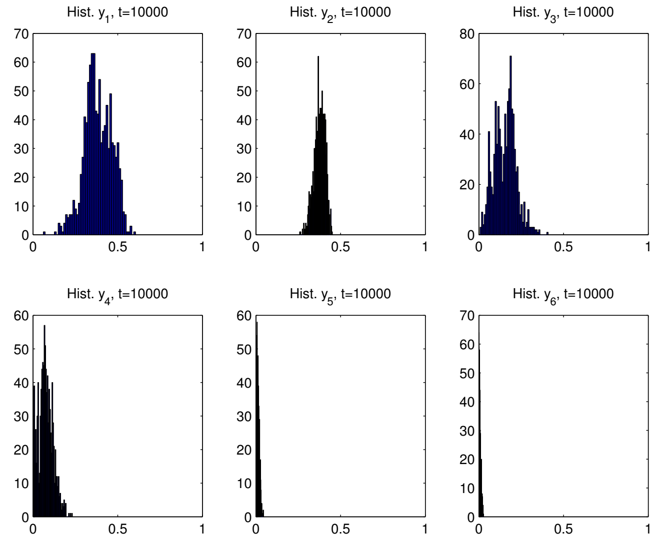

Fig. 6. Histogram of the output particles at time $t = 10000$ for the PF/RLS-FF algorithm.

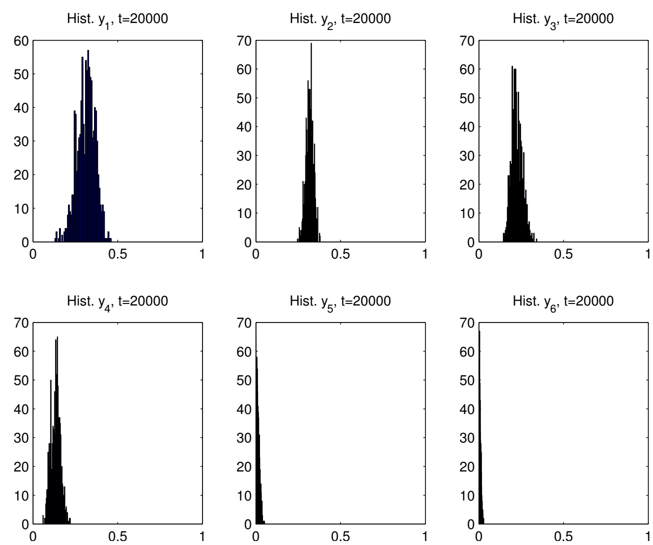

Fig. 7. Histogram of the output particles at time $t = 20000$ for the PF/RLS-FF algorithm.

reflect the time-varying characteristics of the real process.

V. CONCLUSIONS

In this paper we are considered the problem of recursive state and parameter estimation of an industrial process of the mining industry. From the information available from the process (in particular, by the nature of the real output-data provided by Aplik S.A.) a simple class of state-space models was considered, namely, a linear dynamics process equation followed by a nonlinear output equation.

The proposed approach combines a simulation-based technique for the state estimation problem (i.e. the particle filter) with a recursive least squares algorithm (with exponential resetting/forgetting) to estimate the system matrices.

The results of the initial work presented here are encouraging. They show that the proposed approach can deal with the nonlinear and non-Gaussian nature of the available measurements, and it makes possible to model the time-varying characteristics of the system. In particular, from the *ouput particles* it is possible to obtain an *on-line* estimate of the output probability density functions.

In the proposed approach there and still, several open issues that are subject of current and future research. Firstly, the proposed approach requires the tuning of several parameters (in RLS and PF) and initial conditions for states and matrices estimates. Secondly, the results presented here are expected to be contrasted with process data in order to validate states and parameters estimates.

REFERENCES

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, T. Kailath, Ed. Prentice- Hall, 1979.

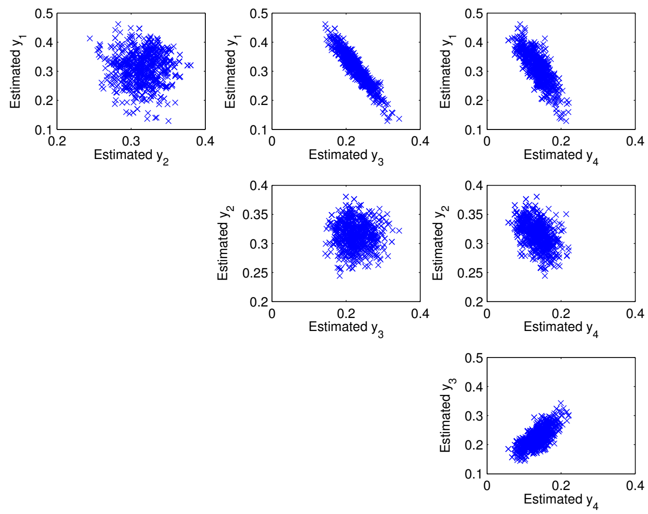

Fig. 8. Particles for different pairs of output signals, for $t = 20000$ when using the PF / RLS-FF algorithm.

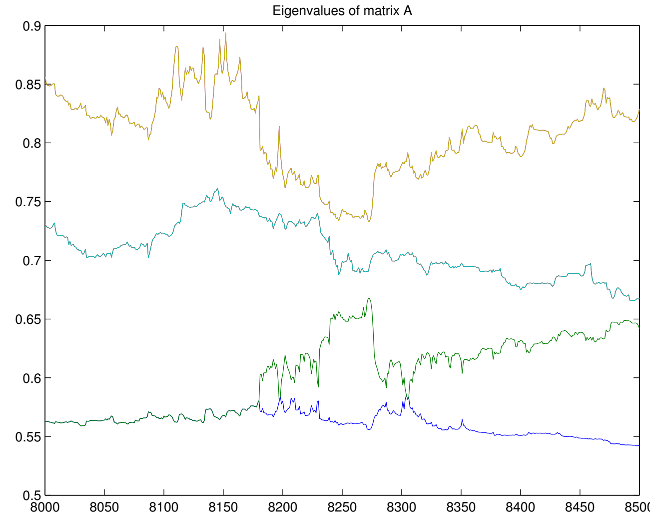

Fig. 10. Magnitude of the eigenvalues of the estimated matrix A_t when using RLS-EFRA in the interval $[8000, 8500]$

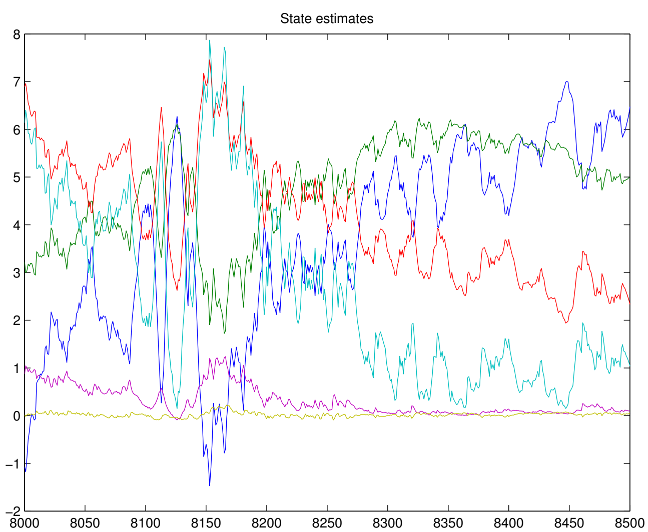

Fig. 9. State estimates given by the PF combined with RLS-EFRA in the interval $[8000, 8500]$.

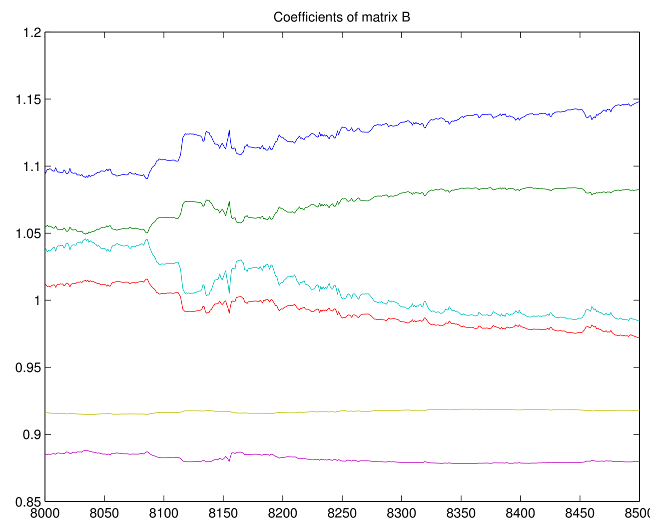

Fig. 11. Components of the estimated matrix B_t when using RLS-EFRA in the interval $[8000, 8500]$

[2] M. Kendall, A. Stuart, and K. J. Ord, *Advanced Theory of Statistics: Classical Inference and Relationship*, 6th ed. Oxford, UK: Oxford University Press, 1991, vol. 2.

[3] G. C. Goodwin and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press, Inc, 1977.

[4] D. Simon, *Optimal State Estimation*. John Wiley & Sons, Inc., 2006.

[5] F. Gustafsson and G. Hendeby, "Some Relations Between Extended and Unscented Kalman Filters," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 545–555, 2012.

[6] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *Radar and Signal Processing, IEE Proceedings F*, vol. 140, no. 2, pp. 107 –113, apr 1993.

[7] T. Schon, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *Signal Processing, IEEE Transactions on*, vol. 53, no. 7, pp. 2279 – 2289, july 2005.

[8] R. Karlsson, T. Schon, and F. Gustafsson, "Complexity analysis of the marginalized particle filter," *Signal Processing, IEEE Transactions on*, vol. 53, no. 11, pp. 4408 – 4411, nov. 2005.

[9] D. Marelli, B. Godoy, and G. Goodwin, "A scenario-based approach to parameter estimation in state-space models having quantized output data," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, 2010, pp. 2011–2016.

[10] L. Ljung, *System Identification*. John Wiley & Sons, Inc., 2001.

[11] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Massachusetts and London: The MIT Press, 1983.

[12] T. Schon and F. Gustafsson, "Particle filters for system identification of state-space models linear in either parameters or states," in *13th IFAC Symposium on System Identification (SYSID)*, Rotterdam, The Netherlands, sep 2003.

[13] C. Andrieu, A. Doucet, S. Singh, and V. Tadic, "Particle methods for change detection, system identification, and control," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 423–438, 2004.

[14] A. Doucet, N. Freitas, N. Gordon, and A. Smith, *Sequential Monte Carlo Methods in Practice*, ser. Information Science and Statistics Series. Springer, 2010.

[15] T. Söderström and P. Stoica, *System Identification*. Prentice Hall, 1989.

[16] M. E. Salgado, G. C. Goodwin, and R. H. Middleton, "Modified least squares algorithm incorporating exponential resetting and forgetting," *International Journal of Control*, vol. 47, no. 2, pp. 477–491, 1988.