CONTROLO
2004

Proceedings

SIXTH PORTUGUESE CONFERENCE ON AUTOMATIC CONTROL

Edited by A.E.RUANO

CONTROLO 2004

UNIVERSITY OF ALGARVE
FARO, PORTUGAL, JUNE 7-9, 2004

Organized by

APCA

Centre for Intelligent Systems

RECURSIVE PARAMETER ESTIMATION ALGORITHMS

J. Boaventura Cunha[a,b], P.B. de Moura Oliveira[a,b], J.P. Coelho[c]

a. Universidade de Trás-os-Montes e Alto Douro, Dep. Engenharias
5001-911 Vila Real, Portugal.
b. Centro de Estudos e Tecnologias do Ambiente e da Vida da UTAD
5001-911 Vila Real, Portugal
c. Instituto Politécnico de Bragança, Escola Superior de Tecnologia e Gestão
5301-854 Bragança, Portugal.

Abstract: Main adaptive control design approaches assume that a suitable dynamic model of the controlled process can be computed. In this way, recursive parameter estimation algorithms play an important role in tracking the time variant parameters of the process dynamic model. This paper describes the major algorithms used to compute the transfer function parameters of time-varying systems. The advantages and limitations of these techniques are illustrated by computing the parameters of a time varying discrete system, with known structure, under the presence of persistent and non-persistent information.

Keywords: Adaptive control, Model based control, Parameter estimation, Recursive Algorithms, Recursive least squares.

1. INTRODUCTION

Model based adaptive controllers require suitable computation of the process dynamic model. Often simple input-output black-box models are used, but statistical relations, rule based models, among others can also be employed. In this way, recursive parameter estimation algorithms play an important role in tracking the time variant parameters of the process dynamic model. Parameter estimation must be seen as one of the key elements to solve a system identification problem, which involves also an experiment design, the selection of a model structure, and the model validation.

Several recursive parameter estimation methods are described in the literature (Ljung, 1987; Söderström and Stoica, 1989; Camacho, E.F. and Bordons, C. 1993). In essence, they can be classified into 3 approaches: direct, indirect and joint input-output. In the first one, the feedback is ignored and the open loop system is identified using the input and output data. In the indirect approach, a closed loop transfer function is firstly identified and, afterwards, the open loop system is determined by using the knowledge of the control law applied. In the last approach, the input and output are viewed jointly as the output from a system driven by the reference signal and noise, and then an algorithm is used to compute the open loop parameters from the estimates of this system.

The recursive parameter estimation algorithms described in this work are based on the data analysis of the input (u) and output (y) signals from the process to be identified. It is considered the case of a system described by the *ARX* model (1) with time varying parameters a, b_1 and b_2 showed in Fig. 1.

$$y(k) = a.y(k-1) + b_1.u(k-1) + b_2.u(k-2) + \xi(k) \quad (1)$$

where $\xi(k)$ is a zero mean white noise with a standard deviation of 0.1.

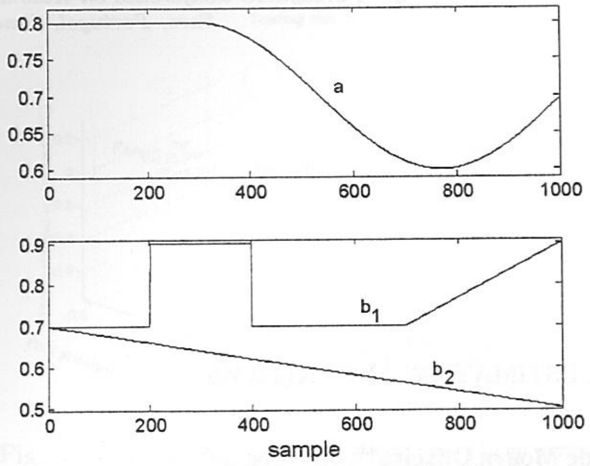

Fig. 1. Parameters of the process described in (1).

Figure 2 shows the data used to test the recursive estimation algorithms. The input-output data was obtained using the plant model of eq. (1) under Proportional Integral (PI) control for a reference signal that varies between -2 and 2.

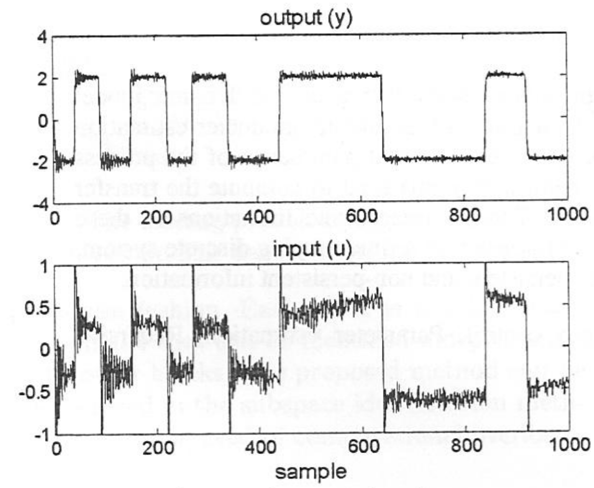

Fig. 2. Process input and output signals.

Several identification algorithms, based on the recursive least squares method, will be presented in the next section in order to compare their performance and point their limitations in solving the problem of parameter estimation in closed loop.

2. RECURSIVE PARAMETER ESTIMATION

In adaptive control, the pairs of input-output data are obtained sequentially on-line. The parameters of the *ARX* model can be computed by performing least squares estimations in a way that the estimates obtained in the previous sample, *k-1*, can be used to derive new ones at the current sample *k*.

To simplify the notation, the model output, $\hat{y}(k)$, of the system described by (1) is expressed as:

$$y(k) = \hat{y}(k) + \varepsilon(k) = \varphi^T(k)\theta + \varepsilon(k)$$

where: $\varphi(k)$ is the data or regression vector and the vector that contains the estimated parameters,

$$\varphi(k) = \left[y(k-1) \quad u(k-1) \quad u(k-2) \right]^T$$

$$\theta = \left[\hat{a} \;\; \hat{b}_1 \;\; \hat{b}_2 \right]^T$$

The recursive computations of the vector parame[ters] can be realized using equations 5 to 7 that is kn[own] as *RLS*-Recursive Least Squares algorithm (Ast[röm] and Wittenmark, 1995),

$$\hat{\theta}(k) = \hat{\theta}(k-1) + K(k)\left[y(k) - \varphi^T(k)\hat{\theta}(k-1) \right]$$

$$K(k) = P(k-1)\varphi(k)\left[I + \varphi^T(k)P(k-1)\varphi(k) \right]^{-1}$$

$$P(k) = \left[I - K(k)\varphi^T(k) \right] P(k-1)$$

in which $K(k)$ denotes a gain matrix and $P(k)$ is covariance matrix of the estimated parameters.

This *RLS* algorithm assumes that the paramete[rs of] the model process are constant. Since the processes have time-varying parameters, the least squares method could lead to inadequate estimat[es]. To cope with time-variant cases some adjust[ment] mechanisms must be introduced in the previous equations. Several implementations have b[een] proposed (Ljung and Söderström, 1983; Ljung and Gunnarsson, 1990; Salgado *et al.*, 1988; Johanss[on] 1992). For instance, in the case where the parameters have abrupt changes, the covari[ance] matrix P can be periodically reset to αI, with α be[ing] a large number. If the real parameters are s[lowly] time-variant, a *RLS* with exponential forgetting [can] be used.

2.1 RLS with exponential forgetting

In the *RLS* with exponential forgetting algorithm [the] least-squares criterion is given by,

$$V(\theta', k) = 0.5\sum_{i=1}^{k} \lambda^{k-i}\left(y(i) - \varphi^T(i)\theta \right)^2$$

where $0<\lambda<1$, is a forgetting factor to account w[ith] time varying weighting of the data.

The parameters that minimise the loss function of [the] cost function in eq.(8) are obtained recursively fr[om] the parameter estimation law (eqs. 9 to 11).

$$\theta(k) = \theta(k-1) + K(k)(y(k) - \varphi^T(k)\theta(k-1))$$

$$K(k)=P(k)\varphi(k)= \\ =P(k-1)\varphi(k)(\lambda I+\varphi^T(k)P(k-1)\varphi(k))^{-1} \quad (10)$$

$$P(k)=(I-K(k)\varphi^T(k))P(k-1)/\lambda \quad (11)$$

This method has the main disadvantage that when the input is not persistent, and as the old data is discarded in the estimation procedure, the matrix P increases exponentially with rate λ. This is called estimator windup.

To illustrate the effect of poor excitation on the performance of the *RLS* with exponential forgetting algorithm, the parameters of the model defined in eq. (1) were estimated using the input-output signals plotted in Fig 2, where is visible that the reference signal is kept constant for long time periods causing the input to be not persistent. This estimation algorithm was applied using a forgetting factor λ of 0.985, initial parameters values $[a\ b_1\ b_2]$=[0.4 0.4 0.4] and initial covariance matrix P=100.I.

Figures 3 and 4 show the results achieved with the *RLS* with exponential forgetting algorithm.

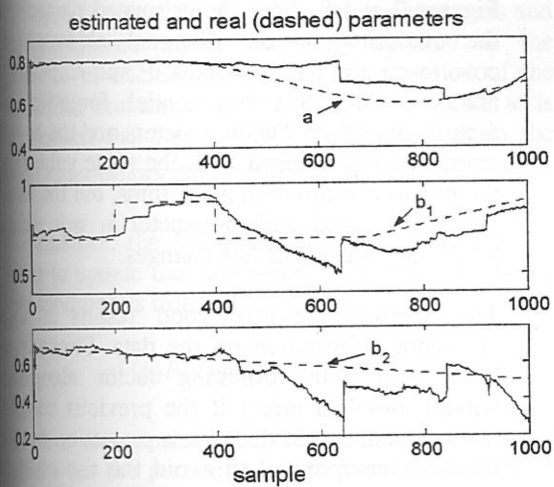

Fig. 3. Real and estimated parameters of the system described in eq. 1 under *PI* control.

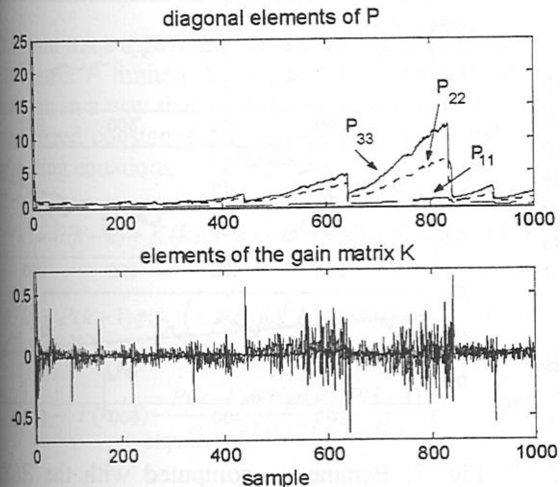

Fig. 4. Diagonal elements of the covariance matrix P and elements of the gain matrix K for the *RLS* exponential forgetting factor algorithm.

These plots point up the problem of estimator windup due to poor excitation. In fact, the elements of the P matrix grow approximately exponentially during the time periods when the control signal is practically constant. This exponential grow of the matrix P and the consequent increase in the gain matrix K have a major impact over the estimated parameters. As it can be observed from Fig. 3, the estimated parameters have large fluctuations and so are inaccurate at the end of the periods where the control signal is approximately constant.

2.2 RLS with directional exponential forgetting

To solve the problem of estimator windup, an estimator that forgets the information only in the directions in which new information is gathered, such as the described by Parkum (1992), Eqs (12) to (19), assures the convergence of the estimations and avoids large changes in the parameters.

$$\alpha(k)=P(k)\varphi(k+1) \quad (12)$$

$$\beta(k)=\varphi^T(k+1)\alpha(k) \quad (13)$$

$$\begin{cases}\xi(k)=\dfrac{\beta(k)}{\lambda\beta(k)-(1-\lambda)} & ,\beta(k)>0\\ \xi(k)=1 & ,\beta(k)\le 0\end{cases} \quad (14)$$

$$\delta(k)=\frac{\alpha(k)}{\xi(k)+\beta(k)} \quad (15)$$

$$K(k)=\delta(k)\varphi^T(k+1) \quad (16)$$

$$P(k+1)=[I-K(k)]P(k)[I-K^T(k)]+\delta(k)\delta^T(k) \quad (17)$$

$$\varepsilon(k+1)=y(k+1)-\varphi^T(k+1)\cdot\hat{\theta}(k) \quad (18)$$

$$\hat{\theta}(k+1)=\hat{\theta}(k)+P(k+1)\varphi(k+1)\cdot\varepsilon(k+1) \quad (19)$$

Figures 5 and 6 show the estimated parameters and the evolution of the diagonal elements of the covariance matrix P, as well as the elements of the gain matrix K, respectively, achieved with this directional forgetting factor algorithm. In this recursive algorithm the parameters, the covariance matrix and the forgetting factor were initialised with the same values used in the previous estimation algorithm.

From the plot of the evolution of the estimated parameters, it can be concluded that they are not significantly affected when the signal input is not persistent. However, there is a significant parameter tracking error, namely for the estimates of parameters b_1 *and* b_2.

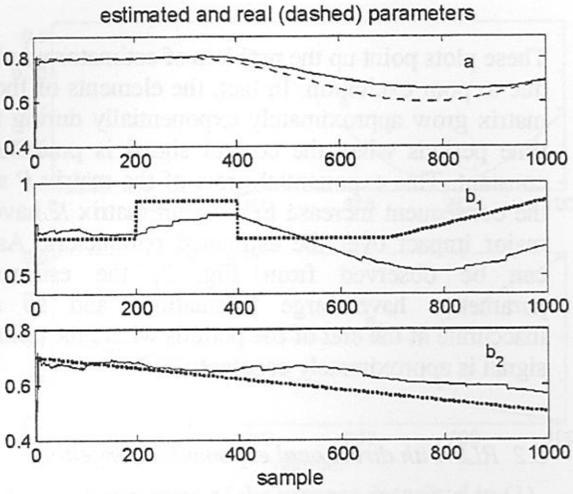

Fig. 5. Parameters computed with the *RLS* with directional exponential forgetting.

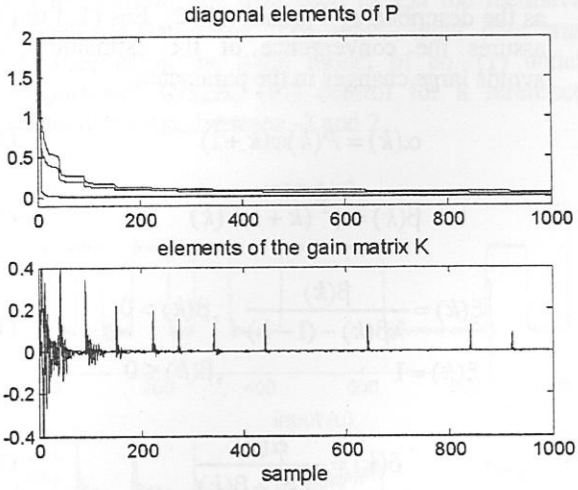

Fig. 6. Elements of P and K matrixes for the *RLS* with directional exponential forgetting.

The reason for this behaviour is due to the fact of the covariance and gain matrixes being only updated in the presence of proper excitation signals, which for this simulation example only occurs when the reference signal changes between the two levels -2 and 2.

2.3 *RLS with exponential forgetting matrix*

The previous techniques are not suited to cope with the cases where the parameters have distinct rates of change in time, since the same value of the forgetting factor is applied in all directions of the parameters space. Here, is described a recursive estimation algorithm with exponential forgetting matrix factors in order to provide distinct information discounts for each parameter.

The *RLS* with exponential forgetting matrix is governed by the following equations, (Ljung, 1987):

$$\Lambda(k) = \Omega \cdot P(k) \cdot \Omega^T \qquad (20)$$

$$K(k) = \frac{\Lambda(k)\cdot\varphi^T(k+1)}{I+\varphi^T(k+1)\cdot\Lambda(k)\cdot\varphi(k+1)}$$

$$\varepsilon(k+1) = y(k+1) - \varphi^T(k+1)\cdot\hat{\theta}(k)$$

$$\hat{\theta}(k+1) = \hat{\theta}(k) + K(k)\cdot\varepsilon(k+1)$$

$$P(k+1) = \Lambda(k)\cdot\left[I - \frac{\varphi(k+1)\cdot\varphi^T(k+1)\cdot\Lambda(k)}{I+\varphi^T(k+1)\cdot\Lambda(k)\cdot\varphi(k+1)}\right]$$

with,

$$\Omega = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sqrt{\lambda_n}} \end{bmatrix}$$

representing a matrix with diagonal elements equal square roots of the forgetting factors associated each column of the regression vector φ.

Figures 7 and 8 show the estimated parameters the evolution of the diagonal elements of covariance and gain matrixes, P and K, respective achieved with this exponential forgetting ma factors algorithm. The parameters and the covaria matrix were initialised with the same values used the previous estimation algorithms, but the forgett factor associated with parameter b_1 was reduced 0.97, since it exhibits fast changes.

This method provides good results if there persistent information on the data. Similarly to basic *RLS* with forgetting factor algorithm windup problem arises if the previous conditio not satisfied, and in this case a protection mechan must be incorporated to avoid the rise of P and matrixes whenever the data is stationary.

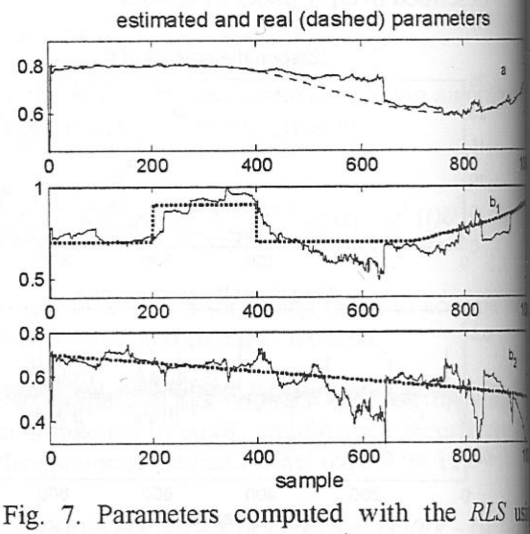

Fig. 7. Parameters computed with the *RLS* us exponential forgetting matrix.

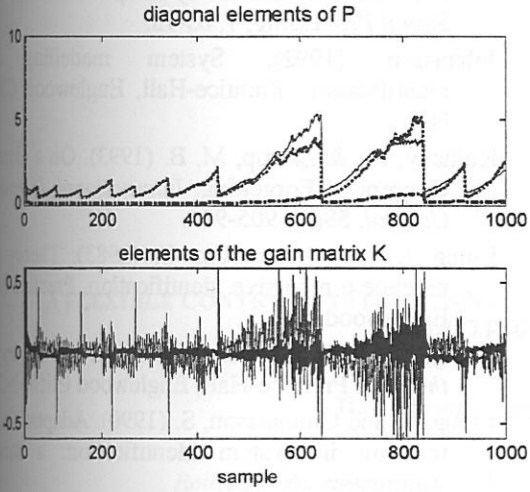

Fig. 8. Elements of P and K matrixes for the *RLS* with exponential forgetting matrix.

2.4 Other RLS estimation methods

Numerous methods have been proposed to avoid the estimator windup problem), such as the conditional or dead zone algorithm (Aström and Wittenmark, 1995; Wellstead and Zarrop, 1991; Kulhavý, R. and Zarrop, M., 1993). In this method the parameters and the covariance matrix are only updated in the presence of excitation in data. The detection of the excitation condition can be based on simple tests concerning the magnitudes of the variations in the plant input-output data or the signals ξ and $\varphi^T P\varphi$.

For instance, the *RLS* estimator can be modified in order to update the parameters only if the following test condition is fulfilled,

$$\varphi^T(k)P(k)\varphi(k)^{-1} > c.(1-\lambda) \tag{26}$$

where c is a positive number that must be proper chosen for the particular application. If its value is very low the covariance windup will appear and if it is too high, the parameters will be updated infrequently leading to poor estimations.

Constant trace algorithms could also be used to keep the matrix P limited, by scaling the matrix at each iteration in a way that the trace of P is constant. The regularized constant-trace algorithm is given by the following equations:

$$\theta(k)=\theta(k-1)+K(k)(y(k)-\varphi^T(k)\theta(k-1)) \tag{27}$$

$$K(k)=P(k-1)\varphi(k)\left(\lambda+\varphi(k)^T P(k-1)\varphi(k)\right)^{-1} \tag{28}$$

$$\bar{P}(k)=\frac{1}{\lambda}\left(P(k-1)-\frac{P(k-1)\varphi(k)\varphi(k)^T P(k-1)}{(1+\varphi(k)^T P(k-1)\varphi(k)}\right) \tag{29}$$

$$P(k)=c_1\frac{\bar{P}(k)}{tr(\bar{P}(k))}+c_2 I \tag{30}$$

in which c_1 and c_2 have positive values given by,

$$\frac{c_1}{c_2}\simeq 10000,\quad \varphi^T\varphi c_1 >> 1 \tag{31}$$

The parameters computed with this algorithm are plotted in Fig. 9.

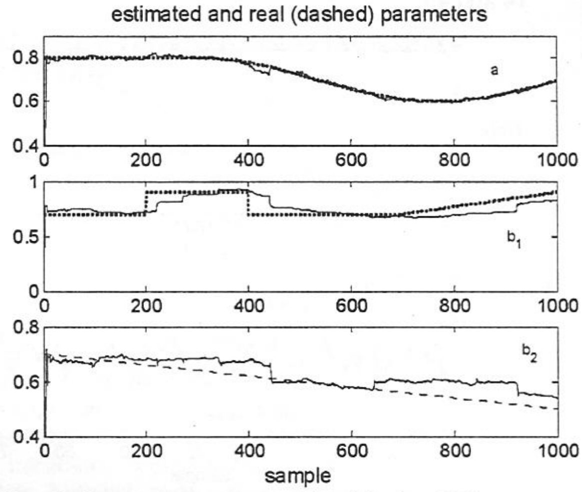

Fig. 9. Parameters computed with the *RLS* constant trace algorithm.

When compared to the previous results, and from the point of view of the quality of the estimated parameters, this method presents a better agreement between the real and computed parameters. Also, the mean squared errors, *MSE*, computed using the one step ahead model predicted output and the real output signals are better in this last case. Table 1 shows the *MSE* values obtained with the four algorithms tested for the time-variant model of eq. (1).

Table 1 Mean squared errors computed using the one step ahead prediction model outputs with the parameters of the four *RLS* algorithms

RLS algorithm	*MSE*
Forgetting factor	0,086
directional exponential forgetting	0,085
exponential forgetting matrix	0,081
constant-trace	0,039

In practical applications, the performance of the *RLS* algorithms could be improved if the prediction errors used in the estimator law, $\xi(k)=y(k)-\varphi^T(k)\theta(k-1)$, are replaced by a function, $f\left(\xi(k)\right)$, that reduces the effects of infrequent large noise signals over the computed parameters, such as in the case of intermittent sensors failures. In this way the parameter adaptation law can be replaced by,

$$\theta(k)=\theta(k-1)+K(k)f(\xi(k)) \tag{29}$$

where $f\left(\xi(k)\right)$ is a linear function for small prediction errors but increases slowly for large ones,

$$f(\varepsilon(k)) = \frac{\xi(k)}{1 + d\left|\xi(k)\right|} \quad (30)$$

in which d is a positive constant.

Figure 10 shows the effect of this transformation, with d=0.4, over a pseudo random signal, where a large amplitude changing was introduced in samples 14 and 49.

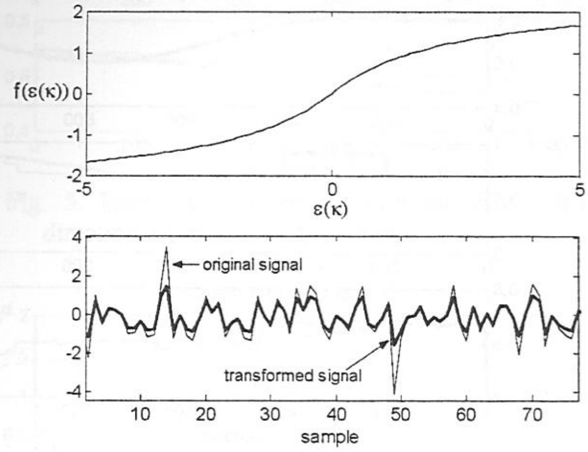

Fig. 10. Graph of the non-linear transformation function $f(\xi(k))$ and its effect over a signal.

3. CONCLUSIONS

Recursive least squares estimation algorithms were reviewed and applied to estimate, recursively the parameters of a first-order *ARX* dynamic model under closed loop control. It has been illustrated that these techniques could lead to robust estimates in the cases that the system to be identified is low or fast time-variant. In the presence of data sets with poor excitation the directional forgetting and constant-trace algorithms provide more robust results, since the diagonal elements of the covariance matrix, P, do not present an exponential rise when there is no persistent excitation.

The different approaches of the *RLS* algorithms presented here, as well others, that were not described due to lack of space, will be implemented to run in parallel. Afterwards, it will be implemented and tested a decision making process, based on the statistical analysis of the recent past prediction errors computed with each algorithm, to choose the better models to be used at each sampling time.

REFERENCES

Aström, K. J. and B. Wittenmark, B. (1995). Adaptive control, second edition. Addison-Wesley, Massachusetts.

Camacho, E.F. and Bordons, C. (1993). Implementation of self-tuning generalised predictive controllers for the process industry. *International Journal of Adaptive Control* [...] *Signal Processing*, 7, 63-73.

Johansson (1992). System modelling [...] identification. Prentice-Hall, Englewood C[...] N.J.

Kulhavý, R., & Zarrop, M. B. (1993). On a Gen[...] Concept of Forgetting. *International Journa[...] Control*, **58(4)**, 905-924.

Ljung, L. and Söderström, T. (1983). Theory [...] practice o recursive identification. Prentice-H[...] Englewood clifs.

Ljung, L. (1987). *System identification - theory [...] the user.* Prentice-Hall, Englewood Cliffs, N.J.

Ljung, L. and Gunnarsson, S. (1990). Adaptation [...] tracking in system identification: a surv[...] Automatica 26, 7-21.

Parkum, J.E., (1992). Recursive Identification [...] Time-Varying Systems, PhD Thesis, Lyngby.

Salgado, M., Mario, E., Graham C., Middleton, [...] (1988). Modified least squares algori[...] incorporating exponential resetting [...] forgetting, Int. Journal of Control, 21, 477-49[...]

Söderström T., Stoica, P. (1989). Sys[...] identification. Prentice-Hall New Jersey.

Wellstead, P. E. and Zarrop, M.B. (1991). *Self-tun[...] Systems Control and Signal Processing*, J[...] Wiley & Sons. Chichester, UK.