



On-line forgetting factor adaptation for parameter estimation based diagnosis

Philippe Weber, Sylviane Gentil, Alain Barraud

► To cite this version:

Philippe Weber, Sylviane Gentil, Alain Barraud. On-line forgetting factor adaptation for parameter estimation based diagnosis. Second Conference on Management and Control of Production and Logistics (MCPL 2000), 5-8 Juillet, 2000, Grenoble, France. pp.1-6. hal-00092904

HAL Id: hal-00092904

<https://hal.science/hal-00092904>

Submitted on 12 Sep 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON-LINE FORGETTING FACTOR ADAPTATION FOR PARAMETER ESTIMATION BASED DIAGNOSIS

P. WEBER¹, S. GENTIL, A. BARRAUD

*Laboratoire d'Automatique de Grenoble UMR-CNRS 5528-UJF
BP 46, 38402 Saint Martin d'Hères Cedex - France Phone: (33) 4 76 82 63 85
Philippe.Weber@esstin.uhp-nancy.fr, Sylviane.Gentil@inpg.fr*

Abstract: This paper presents a fault detection method based on a classical transfer function parameter estimation algorithm in the discrete time domain. Non persistently exciting inputs plant an important problem for the convergence of the estimator. Here, the forgetting factor is adapted on-line in order to improve the convergence. Redundant discrete time transfer functions are used to improve the isolation capacity and obtain a signature table. The fault detection and isolation (FDI) is achieved by the exploitation of this table, with a distance computation. *Copyright* © 2000 IFAC

Keywords: Fault detection, fault isolation, diagnosis, parameter estimation, fuzzy logic, decision making.

INTRODUCTION

On-line parameter estimation reflects the process behaviour and therefore allows FDI. Continuous time parameter estimation can be used if the process is not too complex and it is possible to estimate its physical parameters (Isermann 1993). This method brings direct knowledge of the different system elements and simplifies the fault diagnosis task. Nevertheless, it is difficult to obtain the physical model of a complex process. Therefore the aim of this work is to test classical parameter estimation methods as diagnostic tools. These methods are widespread, and can be found in control toolboxes. Unfortunately, using discrete transfer function representation, the parameters cannot be linked to physical elements of the process. Therefore the estimation task must be followed by a classification technique in order to achieve FDI.

This paper is organised as follows: section 1 presents the winding system used as experiment; and section 2 presents the parameter estimation algorithm. Residual generation is described in section 3. Afterwards, section 4 details the technique implemented to adapt the forgetting factor in order to generate residuals robust to poor excitation of the process inputs. Residual fuzzification and aggregation are presented in section 5. Section 6 describes the generation of redundant transfer functions, which allows to obtain a signature table for fault isolation. Finally, results obtained with the winding system simulation are presented in section 7, before concluding remarks.

1. THE WINDING SYSTEM MODEL

The method proposed is applied to a winding machine represented in (Weber 1998). It is composed of three electric motors M_1 , M_2 and M_3 , controlled by the input vector $\mathbf{u}(k) = [u_1(k) \ u_2(k) \ u_3(k)]^T$. The

¹Present address: CRAN-ESSTIN Parc Robert Bentz, 2, rue Jean Lamour 54500 VANDOEUVRE LES NANCY.

$$\hat{A}(q)y(k) = \sum_{i=1}^m \left\{ \hat{B}_i(q) \cdot (q^{-d_i}) \cdot u_i(k) \right\} + \hat{C}(q) \cdot e(k) \quad (1)$$

where $e(k)$ represents the measurement noise, d_i represents the delays, $\hat{A}(q)$, $\hat{B}_i(q)$, and $\hat{C}(q)$ are polynomials in the shift operator q . The model of the winding system can be written as three MISO ARMAX models:

The Givens rotation algorithm enables to transform $\mathbf{W}(k+1)$ to the form defined by eq. (9). $\hat{\boldsymbol{\theta}}(k+1)$ is then calculated using eq. (10).

3. RESIDUAL GENERATION

Estimating parameters on-line with a long time horizon allows to follow their slow variations. These are not considered as faults, but rather due to normal evolution of the process. An on-line ELS algorithm with a forgetting factor equal to 1 computes the long horizon estimates. This algorithm results in a reference parameter estimate vector Es_h^l for each sub-model h of the process ($h=1...3$ for the winding machine experiment, see (2)(3)(4)):

$$Es_h^l = [\Theta_1^h \quad \dots \quad \Theta_{np^h}^h] \quad (14)$$

where np^h is the number of parameters of the model h .

A second estimator based on a short time horizon allows following fast variations, considered as symptoms of a fault. This estimator results in a tracking model (Basseville 1988). The short horizon estimates are computed by an ELS algorithm similar to (10) with a smaller forgetting factor, and produce the **tracking model parameter estimate** Es_h^s of the model h (eq. (15)). The tracking capability depends on the forgetting factor choice.

$$Es_h^s = [q_1^h \quad \dots \quad q_{p^h}^h] \quad (15)$$

Residuals are computed by the difference between the long horizon estimates and the short horizon ones:

$$r_j^h = \Theta_j^h - q_j^h \quad (16)$$

The residual mean is:

$$r_{rj} = r_{\Theta j} - r_{qj} \quad (17)$$

If no fault occurs r_{rj} should be close to zero, because the ELS algorithm results in unbiased parameter estimate. Residual variance is more complicated to evaluate. In the faulty case, estimates are not correlated, because the models before and after the fault are not related. The residual variance is defined by:

$$s_{rj}^2 = s_{\Theta j}^2 + s_{qj}^2, \quad (18)$$

In the fault free case, the long and short horizon estimates are strongly correlated, and the residual variance is computed by:

$$s_{rj}^2 = s_{\Theta j}^2 + s_{qj}^2 - 2 \cdot E\{\Theta_j^* \cdot q_j^*\} \quad (19)$$

where Θ_j^* and q_j^* represent the centred estimates. Thus (18) and (19) bound the residual variance.

The main problem planted by an implementation on an industrial process is that both estimators may not converge if the inputs are not persistently exciting. A Pseudo Random Binary Signal (PRBS) input is usually used for system identification, but it is evidently impossible to superpose continuously to the inputs a PRBS on-line during the production cycle,

in order to ensure a robust diagnosis. In order to increase the residual robustness, we propose in the following to adapt on-line the forgetting factor to the input excitation. This adaptation is based on the condition number of $\mathbf{X}(k)$ and the estimate variances $\hat{\mathbf{S}}_e^2(k)$ analysis.

4. FORGETTING FACTOR ADAPTATION METHOD

Using the ELS algorithm presented in section 2, the condition number of $\mathbf{X}(k)$, $\mathbf{x}(\mathbf{L}(k)^T \cdot \mathbf{L}(k))$, is given by:

$$\mathbf{x}(\mathbf{L}(k)^T \cdot \mathbf{L}(k)) = \|\mathbf{L}(k)^T \cdot \mathbf{L}(k)\| \cdot \|(\mathbf{L}(k)^T \cdot \mathbf{L}(k))^{-1}\|^2 \quad (20)$$

The matrices $\mathbf{L}(k)^T \cdot \mathbf{L}(k)$ and $(\mathbf{L}(k)^T \cdot \mathbf{L}(k))^{-1}$ are calculated by the ELS algorithm. Nevertheless in order to increase robustness to numerical errors, it is more interesting to compute the condition number of $\mathbf{L}(k)$ which contains the same information. The advantage of this method relies on the fact that $\mathbf{x}(\mathbf{L}(k))$ can be more robustly computed than $\mathbf{x}(\mathbf{L}(k)^T \cdot \mathbf{L}(k))$ which leads to very large numerical values.

$\mathbf{x}(\mathbf{L}(k))$ increases if the excitation of the input decreases. If the excitation is poor, using a forgetting factor leads to a degradation of the estimates due to the multiplication of $\mathbf{L}(k)$ by $\mathbf{I} < 1$. This phenomenon is illustrated in Fig. 1 (a): the condition number increases after 100 sampling periods; the variance of the estimates increases; the estimate diverges.

The solution proposed here is to extend the estimation horizon of the tracking estimator by increasing $\mathbf{I}(k)$ to 1. This can be interpreted as an **elastic estimation horizon**. After 100 sampling periods $\mathbf{x}(\mathbf{L}(k))$ increases, then the adaptation of the estimates are slowed down, the estimate variance increases very slowly and the estimates are not perturbed. These results are presented in Fig. 1 (b). Increasing $\mathbf{I}(k)$ results in a tracking estimator close to the reference estimator, thus residuals are close to zero if the input is not rich enough. If the input excitation increases or if a fault occurs, $\mathbf{I}(k)$ has to be decreased in order to quickly adapt the estimates. Nevertheless the time from which $\mathbf{I}(k)$ must begin to decrease is difficult to determine because if $\mathbf{I}(k)$ is very close to 1, the estimates are less sensitive to the faults. As it can be noted, a fault occurrence leads to a discrepancy between the model and the process outputs, resulting in an increase of $\mathbf{e}(k)$ and $\hat{\mathbf{S}}_e^2(k)$. Fig. 2 allows to compare a fixed forgetting

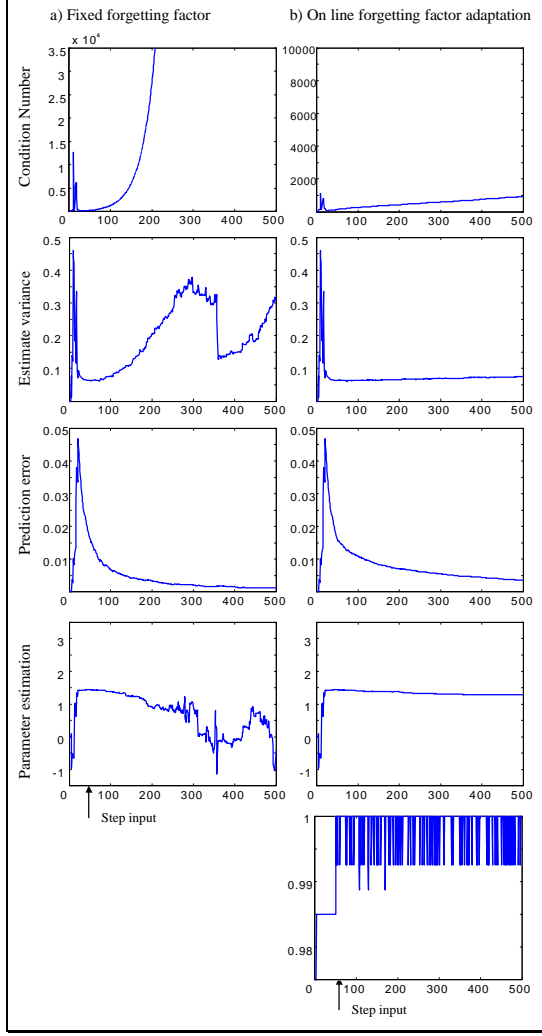


Fig. 1: Comparison between a fixed forgetting factor and an on-line adapted one in a fault free case.

factor and the on-line adaptive solution proposed in this paper. When a fault occurs the variance of the estimates increases then $I(k)$ is decreased.

The forgetting factor adaptation is based on $\hat{\mathbf{S}}_e^2(k)$ and $\mathbf{x}(\mathbf{L}(k))$, both computed with the ELS algorithm. The variation of $\hat{\mathbf{S}}_e^2(k)$ and $\mathbf{x}(\mathbf{L}(k))$ are defined as:

$$\Delta \mathbf{x}(\mathbf{L}(k)) = \mathbf{x}(\mathbf{L}(k)) - \mathbf{x}(\mathbf{L}(k-1)) \quad (21)$$

$$\Delta \hat{\mathbf{S}}_e^2(k) = \hat{\mathbf{S}}_e^2(k) - \hat{\mathbf{S}}_e^2(k-1) \quad (22)$$

$I(k)$ is adapted between I_0 and 1 by the simple following algorithm:

If $\Delta \mathbf{x}(\mathbf{L}(k)) > 0$,

$$I(k+1) = I(k) + (1 - I(k)) / \mathbf{a} \quad (23)$$

If $\Delta \hat{\mathbf{S}}_e^2(k) > 0$,

$$I(k+1) = I(k) - (I(k) - I_0) / \mathbf{b} \quad (24)$$

These formulas are empirical, \mathbf{a} and \mathbf{b} are tuned to obtain the desired sensitivity.

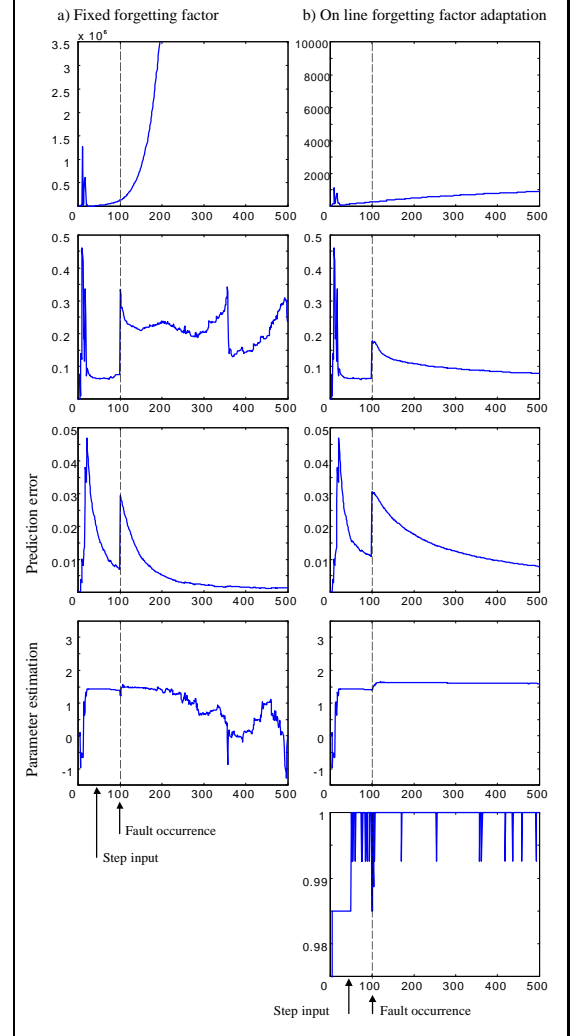


Fig. 2: Comparison between a fixed forgetting factor and an on-line adapted one under fault occurrence.

5. SYMPTOM GENERATION

The symptoms are generated in two steps: the residual fuzzification, and the aggregation of the residual fuzzy descriptions. More details can be found in (Weber and Gentil 1998).

5.1 Residual fuzzification

In order to bypass the unknown residual probability distribution, fuzzy set theory is used. Two fuzzy sets are defined in \mathcal{R}_+ , the universe of discourse of the residual absolute values: Z for Zero; and NZ for Not Zero. The fuzzy set Z is the complement of NZ . The membership functions are dynamically adapted thanks to the residual variance which is chosen by linear interpolation between the equations (18) and (19), depending on the NZ membership degree computed at the previous step. Thus the membership functions to the two sets are defined by:

$$\mathbf{m}_{NZ}(|r_j(k)|) = \min \left(1, \max \left(0, \frac{|r_j(k)| - a \cdot \hat{\mathbf{S}}_{r_j}(k)}{b \cdot \hat{\mathbf{S}}_{r_j}(k) - a \cdot \hat{\mathbf{S}}_{r_j}(k)} \right) \right) \quad (25)$$

$$\mathbf{m}_z(|r_j(k)|) = 1 - \mathbf{m}_{NZ}(|r_j(k)|) \quad (26)$$

where a and b are chosen to tune the sensitivity of the residual fuzzy description to the residual variations (in the following example $a=2$ and $b=3$).

5.2 Aggregation

For fault detection and isolation, the relevant information is a *global perturbation* of all parameter estimates Es_h^s . Note that estimate vectors Es_h^s have not necessarily the same dimension. Thus a global perturbation degree is computed for each vector Es_h^s , using residual membership degrees to the fuzzy sets. The symptom vector s_h representing the state of the model h , is defined on the universe $\mathfrak{R}_+^{np^h}$ which is the cartesian product of the residual universes of discourse, where np^h is the dimension of Es_h^s .

$$s_h = \left[\begin{array}{c|c|c} |r_1^h| & \dots & |r_{np^h}^h| \end{array} \right] \quad \text{where } s_h \in \mathfrak{R}_+^{np^h} \quad (27)$$

Two fuzzy sets are defined on the $\mathfrak{R}_+^{np^h}$ universe: Globally Perturbed (*GP*); and Not Perturbed (*NP*). The membership degrees of s_h to the sets *GP* and *NP* are computed by aggregation of the residual fuzzy descriptions. A mean aggregation operator is chosen as a compromise between a T-norm and a T-conorm. Membership functions to *GP* and *NP* sets are thus defined by:

$$\mathbf{m}_{GP}(s_h) = \frac{1}{np^h} \sum_{j=1}^{np^h} \mathbf{m}_{NZ}(|r_j^h|) \quad (28)$$

$$\mathbf{m}_{NP}(s_h) = 1 - \mathbf{m}_{GP}(s_h) \quad (29)$$

6. FAULT ISOLATION

6.1 Signature table generation

When a fault occurs, all the residuals related to the estimates of the models linked to the faulty measurement are sensitive to the fault. Thus, only the estimation of the transfer functions uncoupled to the faulty measurements remains unchanged.

| Table 1 | | | |
|--------------------------|----------|----------|----------|
| \mathbf{D}_1 | Es_1^s | Es_2^s | Es_3^s |
| \mathbf{Sg}_{T1} | 1 | 0 | 0 |
| $\mathbf{Sg}_{\Omega 2}$ | 1 | 1 | 1 |
| \mathbf{Sg}_{T3} | 0 | 0 | 1 |

In the following, an incidence matrix $\mathbf{D}_1(n, h)$ (Gertler 1998) is proposed: ‘1’ represents the Es_h^s vectors affected by a fault, and ‘0’ represents the

ones not affected. The columns are related to the models. The rows represent the fault signatures \mathbf{Sg}_n for a fault in T_1 , Ω_2 and T_3 (Table 1).

Additional transfer function estimation allows extending the incidence matrix (Table 1), by adding new symptoms. They can be obtained using eq. (2), (3) and (4) as parity relations (Gertler 1998). Substituting eq. (3) into (2) leads to a model that links T_1 to u_1 and u_2 ; with this model, T_1 and Ω_2 are no longer coupled. The parameters estimated in this way will not be sensitive to the faults on Ω_2 . The same procedure can be applied for (4) and (3) and so, the vector of estimates Es_4^s and Es_5^s can be obtained from the new models:

$$\hat{A}_{11}(q) \cdot T_1(k) = \hat{B}_{11u2}(q) \cdot u_2(k - d_{11u2}) + \hat{B}_{11u1}(q) \cdot u_1(k - d_{11u1}) + \hat{C}_{11}(q) \cdot e_{11}(k) \quad (30)$$

$$\hat{A}_{33}(q) \cdot T_3(k) = \hat{B}_{33u2}(q) \cdot u_2(k - d_{33u2}) + \hat{B}_{33u3}(q) \cdot u_3(k - d_{33u3}) + \hat{C}_{33}(q) \cdot e_{33}(k) \quad (31)$$

The incidence matrix $\mathbf{D}_2(n, h)$, is represented in Table 2 using the vector of estimates Es_4^s , Es_2^s and Es_5^s .

| Table 2 | | | |
|--------------------------|----------|----------|----------|
| \mathbf{D}_2 | Es_4^s | Es_2^s | Es_5^s |
| \mathbf{Sg}_{T1} | 1 | 0 | 0 |
| $\mathbf{Sg}_{\Omega 2}$ | 0 | 1 | 0 |
| \mathbf{Sg}_{T3} | 0 | 0 | 1 |

6.2 Isolation function

The vector \mathbf{S} of *GP* membership degrees is defined as:

$$\mathbf{S} = [\mu_{GP}(s_1), \dots, \mu_{GP}(s_H)] \quad (32)$$

If this vector is close to zero, no fault is detected. Otherwise a decision procedure isolates the fault.

The isolation function noted $F_I(\mathbf{S}, \mathbf{Sg}_n)$ is achieved comparing \mathbf{S} to the signatures \mathbf{Sg}_n . This function measures the similarity between \mathbf{Sg}_n and \mathbf{S} . This similarity can be determined using the Hamming distance calculated in a sup-space sensitive to the fault n (Weber and Gentil 1998):

$$F_I(\mathbf{S}, \mathbf{Sg}_n) = 1 - \frac{1}{W_n} \sum_{h=1}^H \{ |\mathbf{m}_{GP}(s_h) - \mathbf{D}(n, h)| \cdot \mathbf{D}(n, h) \} \quad (33)$$

such that $F_I(\mathbf{S}, \mathbf{Sg}_n) \in [0, 1]$, where W_n is the number of elements $\mathbf{D}(n, h) \neq 0$.

7. APPLICATION

The above-described method has been applied to the winding machine simulator using the redundant signature table (Table 2).

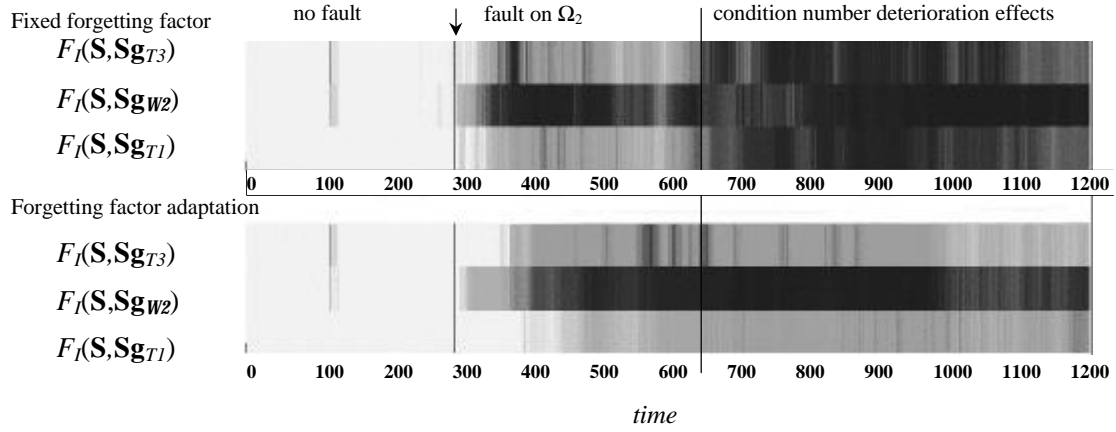


Fig. 3: Fault isolation results

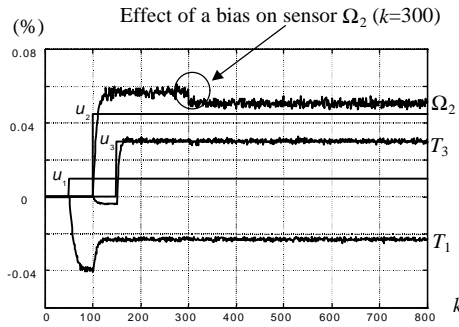


Fig. 4: Inputs and outputs around the operating point.

The inputs u_1 , u_2 and u_3 are step inputs at sample 50, 100 and 150 and the sampling period is equal to 0.1s. The signal to noise ratio was fixed to 31 dB. The fault was simulated as a 10 % bias on the sensor Ω_2 , at time 300 (Fig. 4).

Fig. 3 presents the isolation functions related to Table 2. $F_I(\mathbf{S}, \mathbf{Sg}_{\Omega_2})$ (33) is greater than 0.5, 50 sampling periods after the fault occurrence, thus the fault is isolated. With a fixed forgetting factor method, a lot of false alarms are generated, explained by the degradation of the estimates due to the problem of poor input excitation. The proposed forgetting factor adaptation method results in good isolation of the fault in spite of the poor input excitation.

CONCLUSION

This paper proposes a method for the fault detection and isolation based on parameter estimation.

A classical identification method is used to estimate discrete time transfer function parameters. The on-line implementation of the ELS algorithm, is achieved by an orthogonal transformation in order to guarantee good numerical properties, under this conditions the estimations used to compute the symptoms lead to a decision robust to the numerical errors.

Non persistently exciting inputs plant an important problem for the convergence of the estimator. The

on-line adaptation of the forgetting factor related to the condition number of the observation matrix and the estimate variances improve the convergence of the estimates. This method leads to a decision robust to the poor input excitation.

Several sub-models of the system are used in order to generate structured residuals. The redundancy allows the generation of a procedure for fault isolation. This decision is based on fuzzy sets to support the aggregation of symptoms and a distance to classify the symptoms in the signature space.

The application on the winding process simulation proved that this method decreases the false alarms.

REFERENCES

- Basseville, M. (1988). Detecting changes in signals and systems - A survey. *Automatica*, Vol. **24**, No. **3**, pp. 309-326.
- Gertler, J. (1998). *Fault detection and diagnosis in Engineering systems*. Marcel Dekker, Inc., New York, USA.
- Golub, G.H., Van Loan C.F. (1983). *Matrix computations*. Third Edition, The Johns Hopkins University Press, London.
- Isermann, R. (1993). Fault diagnosis of machines via parameter estimation and knowledge processing - Tutorial paper. *Automatica*, Vol. **29**, No. **4**, pp. 815-835.
- Ljung, L. (1987). *System identification: Theory for the user*. Prentice-Hall Englewood Cliffs Information and system science series, New Jersey.
- Weber, P., Gentil S. (1998). Fault detection and isolation using Multiple model parameter estimation. *CESA'98 IMACS, Multiconference, Computational Engineering in systems applications*, Nabeul-Hammamet, Tunisia.
- Weber, P. (1999). *Diagnostic de procédé par l'analyse des estimations paramétriques de modèles de représentation à temps discret*. Phd These, Institut National Polytechnique de Grenoble, France.