Fault Detection and Monitoring Using Parameters Identification and Principal Component Analysis. Application to Rotary Machines in Skin Pass Process

S. BOUHOUCHE[(1)], M. LAHRECHE[(1)], S. ZIANI[(1)], AND J. BAST[(2)]
[(1)] Iron and Steel Applied Research Unit – URASM
Welding and Control Research Center, CSC
BP 196 Annaba, 23000,
ALGERIA
[(2)] HGUM, Institut für Maschinenbau
TU Bergakademie Freiberg Germany
Cotta Stasse 4, D-9596
GERMANY

Abstract: - A new approach for fault detection and monitoring based on the parameters identification coupled to the Principal Component Analysis (PCA) is proposed in this paper. The proposed Fault Detection and Monitoring consists to apply the PCA method on the dynamic of the identified parameters. Conventional PCA uses the process inputs and outputs as variables which are used in the computing procedure. Using the process parameters behaviour as variables in the PCA computing procedure improve the detect ability by reducing the wrong faults generating by the noise effects. Application on the rotary machines in skin pass machines of cold rolling will be developed in this work.

Keywords:- Identification, Fault Detection and Diagnosis (FDD), parameters estimation, Least Square Recursive Algorithm (RLS), Principal Component Analysis (PCA), Skin Pass Machine, Cold Rolling Process

1 Introduction

Identification techniques are important tools that give an estimation of the process parameters evolutions. According to the importance of the parameters deviations, the monitoring system uses these aspects to adapt the control system, to predict a fault or to shutdown the global functionality if damage will be occur. Many works based on the RLS methods have been developed [1]-[5]. Process identification supposes a persistent and stable input signal to obtain a significant output and optimal convergence of the model parameters [6]-[7].

Fault detection and diagnosis is accomplished by comparing performance determined from measurements such as the model parameters with some expectation of performance. If the deviation exceeds a threshold, then a fault is indicated. Often this process is divided into two steps as depicted in Fig.1. It is considered pre-processing and classification processes. The pre-processor takes measurements from sensors and manipulates them to generate features for classification. This includes parameters identification coupled to PCA. Classifiers then operate on the features to determine faults.
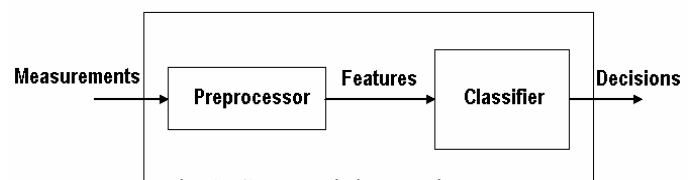

Fig.1: Sequential steps in FDD

The FDD process can be made by modelling; features are obtained from parameters identification using RLS coupled to the PCA. An identification model is used to fault diagnosis and symptoms generation. Fig.2 gives the principle of detection and classification using RLS coupled to PCA.
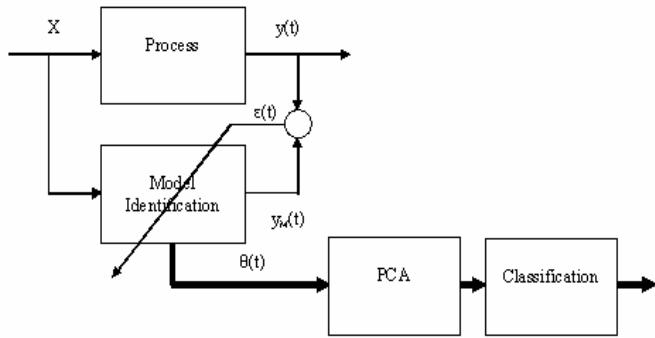
Fig.2: Fault detection and diagnosis based RLS and PCA

2 Modelling and identification using RLS algorithm

The identification technique permits to find the process parameter vector using the minimum least square error between the process output and model output according to the dynamical data. We consider a process with dynamic output and exogenous input; this model is called Autoregressive Moving Average with eXogenous inputs (ARMAX). Each predicted output can be written as:

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})w(t) \tag{1}$$

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2}) + \ldots + a_n q^{-n} \tag{2}$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \ldots + b_m q^{-m} \tag{3}$$

$$C(q^{-1}) = c_0 + c_1 q^{-1} + c_2 q^{-2}) + \ldots + c_p q^{-p} \tag{4}$$

n, m and p is the differentiation order for the output $y(t)$, the control input $u(t)$ and the exogenous input $w(t)$ respectively which are defined according to the process dynamics. The objective is to find optimal values of the process parameters using a least square algorithm. The principle of identification is given in Fig.2.

From equation (1), the model output can be written as:

$$y(t) = X(t)^T \theta(t-1) \tag{5}$$

With:

$$X(t) = [y(t-1), y(t-2),\ldots,y(t-n),u(t-1),\ldots,u(t-m), \tag{6}$$
$$, w(t-1),\ldots, w(t-p)]^T$$

$$\theta(t) = [a_1, a_2,\ldots,a_n, b_1,\ldots,b_m, c_1, c_2,\ldots,c_p]^T \tag{7}$$

The prediction error can be defined as:

$$\varepsilon(t) = y_p(t) - X(t)^T \theta(t-1) \tag{8}$$

The identification objective is to find the process parameters that minimise the sum of errors $\varepsilon(t)$.

$$Min\{J\} = \{\sum_{k=0}^{t} \varepsilon(k)\} \Rightarrow \theta(t)_{optimal} \tag{9}$$

The following form gives the recursive estimation of vector parameters:

$$\theta(t) = \theta(t-1) + P(t)X(t)\varepsilon(t) \tag{10}$$

where:

$$P(t) = \frac{1}{\lambda(t)} \left[P(t-1) - \frac{P(t-1)X(t)X(t)^T P(t-1)}{\lambda(t) + X(t)^T P(t-1)X(t)} \right] \tag{11}$$

The forgetting factor $\lambda(t)$ is usually computed according to the rule,

$$\lambda(t) = \lambda 0 \lambda(t-1) + 1 - \lambda 0 \tag{12}$$

Recursive Least Square (RLS) estimation can be defined as:

Step1: Initialisation

-Define: $\theta \in \Re mx1$, $P \in \Re mxm$, $X \in \Re mx1$, I=Diag(mxm),

-y(0),u(0)…

-P0=I/α, $\alpha \ll 1$

-θ0=[0 0 0…….0]T

-λ0=0.95

Step2: Recursive estimation

Input/output data acquisition

$_{-} X(t) = [y(t-1), y(t-2),\ldots,y(t-n),u(t-1),\ldots,u(t-m),w(t-1),\ldots,w(t-p)]^T$

$_{-} y(t) = X(t)^T \theta(t-1)$

-Compute $\varepsilon(t)$.Equation(8)

-Compute $\theta(t)$ Equation(10)

-Compute $\lambda(t)$ Equation(12)

3 Fault detection using RLS and PCA

By projecting the data into a low-dimensional space that accurately characterises the state of the system, dimensionally reduction techniques can greatly simplify and improve process monitoring procedures. PCA is such a dimensionally reduction technique. It produces a lower-dimensional representation in a way that preserves the correlation structure between the process variables, and is optimal in terms of capturing the data variability. This technique is a linear method of system reduction; it is optimal in term of the capturing of the process variability which is important to detect the process fault.

Given a training set of n observations and m process variables stacked into a matrix X,

$$X_p = \begin{bmatrix} x_{11}, x_{12},\ldots.x_{1m} \\ x_{21}, x_{22},\ldots.x_{2m} \\ . \\ x_{n1}, x_{n2},\ldots.x_{nm} \end{bmatrix} \tag{13}$$

We characterise the measured matrix X_p by:

$$S = \frac{1}{n-1} X^T X = V \Lambda V^T \tag{14}$$

Where Λ is a diagonal matrix:

$$\Lambda = \Sigma^T \Sigma \tag{15}$$

$\Sigma \ \Xi$ Rmxm contains the non-negative real singular values of decreasing magnitude along its main diagonal

($\sigma_1 \geq \sigma_2 \geq,....\sigma_{\min(m,n)} \geq 0$), and zero off diagonal elements. The loading vectors are the orthogonal column vectors in the matrix V, and the variance of the training set projected along the ith column of V is equal to σ_i^2. Solving equation (13) is equivalent to solve an eingenvalue equals the square of the ith singular value (i.e., $\lambda_i = \sigma_i^2$).

In order to optimally capture the variations of the data while minimizing the effect of random noise corrupting the PCA representation, the loading vectors corresponding to the a largest singular values are typically retained.

Selecting the columns of the loading matrix $P \in R^{mxa}$ to correspond to the loading vectors associated with the first a singular values, the projection of the observations in X into lower-dimensional space are contained in the score matrix,

$$T = XP \tag{16}$$

and the projection of T back into m-dimensional observation space,

$$\hat{X} = TP^T \tag{17}$$

The difference between X and $\hat{X}$ is the residual matrix E

$$E = X - \hat{X} \tag{18}$$

The residual matrix captures the variations in the observation space spanned by the loading vectors associated with the m-smallest singular values. The subspaces spanned by $\hat{X}$ and X are called the score space and residual space respectively.

Process variability is characterised by the dynamic of the process parameters, PCA can be extended to analysis of the model parameters estimated by the RLS algorithm. In this way, the process parameters variability can be defined by the analysis of the following observations matrix X_p

$$X_p = \begin{bmatrix} a_1(t), a_1(t-1),........................a_1(t-n) \\ a_2(t-1), a_2(t-2),.................a_2(t-n) \\ . \\ c_p(t-1), c_p(t-1),.................c_p(t-n) \end{bmatrix} \tag{19}$$

T2 statistic can be used to detect faults for multivariate process parameters data. Given an observation vector x and assuming that $\Lambda = \Sigma^T \Sigma$ is invertible, the T^2 statistic can be calculated directly from the PCA representation as:

$$T^2 = x_p^T V(\Sigma^T \Sigma)^{-1} V^T x_p \tag{20}$$

x is defined as:

$$x_p = [x_{i1}, x_{i2},....x_{im}], \text{ i=1 to n} \tag{21}$$

By including in the matrix P the loading vectors associated only with the a largest singular values, the T2 statistic for the lower-dimensional space can be computed:

$$T^2 = x_p^T P \Sigma P^T x_p \tag{22}$$

An optimal T2 statistic can be computed using the normal process parameters, fault can be detected when the T2 statistic related to actual operating conditions exceeds a fixed threshold.

Another indicator such Q statistic can be used. It is defined as:

$$Q = r^T r \tag{23}$$

Where r is the residual vector, a projection of the observation x into the residual space. Since Q statistic does not suffer from an over sensitivity to inaccuracies in the smaller singular values. The Q statistic, also known as the squared prediction error. Computational method is summarised in the following algorithm:

Step1: Estimation of the model parameters
Use the RLS algorithm given in section (2) to estimate the process parameters $\theta(t)$

Step2: Fault detection using PCA

- Define T_{sh}^2 and Q_{sh} threshold

- Define the observation matrix

$$X_p = \begin{bmatrix} a_1(t), a_1(t-1),........................a_1(t-n) \\ a_2(t-1), a_2(t-2),.................a_2(t-n) \\ . \\ c_p(t-1), c_p(t-1),.................c_p(t-n) \end{bmatrix} \tag{24}$$

- Compute: $S = \dfrac{1}{n-1} X_p^T X_p$

- Compute V, the eigenvalues of S
- Compute P as reduced eigenvectors,
- P=V(:,id-s); where id is the maximum dimension of V and s is number of insignificant eigenvalues that constitute appreciatively 3% of the maximum eigenvalue.
- T=Xp.P;
- Compute the estimation of Xp ,
- Xes=T*P';
- Compute de residual,
- E=Xp-Xes
- Compute T2 statistic
- T2 =Xes V Inv(D)*VT*XesT;
- Compute Q Statistic
- Q=ET*E, $Q \leq Q_{sh}$
- If $T^2 \leq T^2_{sh}$, there are no fault go to step 1 for new identification process, else fault occur
- If $Q \leq Q_{sh}$, there are no faults go to step 1 for new identification process, else fault occur
- Continue

4 Applications

4.1 Process description

Fig. 3 shows a part of skin pass process; the main objective is to obtain a uniform pression on the metal sheet surface. The metal sheet characterised by its thickness is maintained between the upper and lower rolls. Pression on the surface is obtained by a mechanism which is controlled by forces applied by two motors. Motors M1 and M2 must have a same rotary speeds in order to generate an equal pressions in left and right sides.
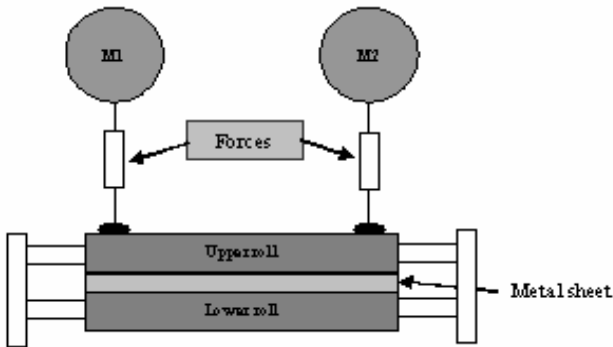
Fig. 3: Principle of skin passes process

When a fault appears on the equipment of the rotary control system of motor M1 or motor M2, the rotary synchronism is then disturbed generating an important defect on the left or on the right side of the metal sheet. This cause also a shutdown of the process.

4.2 Speed control

Fig. 4 defines the principle of speed control of motors M1 and M2. Only motor M1 is controlled in closed loop using the feedback loop, motor M2 has the same characteristics of motor M1, its control is obtained by a simple connexion of the control law given by the feed back.
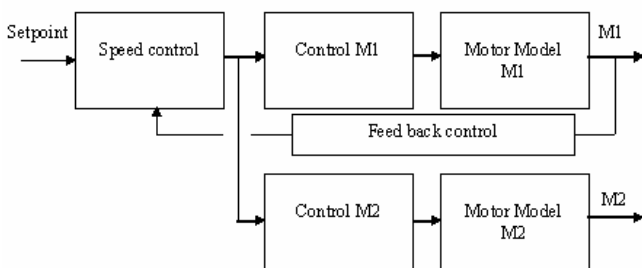
Fig. 4: Principle of length loop control

4.3 System Modelling

Fig. 5 defines the principle of data acquisition by interfacing the analog control system and the process computer by means of data acquisition package. Data has been stored on the hard disk of process computer in real time.
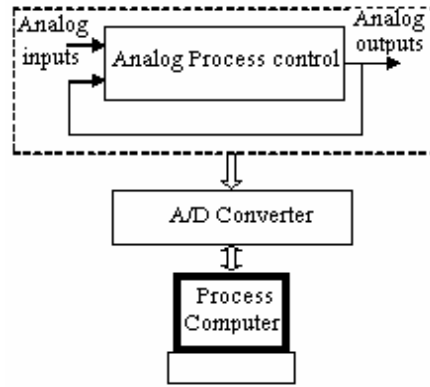
Fig. 5: Principle of data acquisition

We consider in this part the identification of the model defined by:
The input is the control signal of control card M1 named "Control M1", see Fig.4.
The output is the output signal of "Control M1"
Modelling process uses the stored input – output data from normal operating conditions to find the dynamic relationship between the input and output variables. RLS identification algorithm is then used. We define the dynamic interactions between the input and output by the following equation.

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + b_1 u(t-1) + b_2 u(t-2) \qquad (25)$$

$u(t)$ and $y(t)$ are the model input and output given in Fig.6. According to the dynamic data of $u(t)$ and $y(t)$, it appears that we can consider a deterministic model without random noise.

Using the RLS algorithm developed in section (II), the results given in Fig.6c and Fig.6d have been obtained. It has been used the final value of $\Theta(t)$, with $\langle \lim \theta(t) \to \theta_0, if, t \to \infty \rangle$ to calculate the model output. $\Theta_0 = \Theta(t=610)$, is then used to calculate the computed output (Fig.6a).

In normal situation $\theta(t) \in D_N$, θ is bounded in an admissible domain.
Faulty situation is defined by the variation of $\Theta(t)$ in a faulty domain D_F, This is caused by the model changes including the structural and parameters variations $\Delta\theta(t) = \theta(t) - \theta_0$.

If $\lim \Delta\theta(t) \to 0$ when $t \to 0$: The fault is steady state
If $\lim \Delta\theta(t) \neq 0$ when $t \to 0$: The fault is static
Static fault generates by the control system of the rotary machine M1 or M2 induces a synchronism error. When motor M1 and M2 have not a same rotary speed, the pression applied on the right and left sides are not equal , this generates an important defect that induce a shutdown of the process.

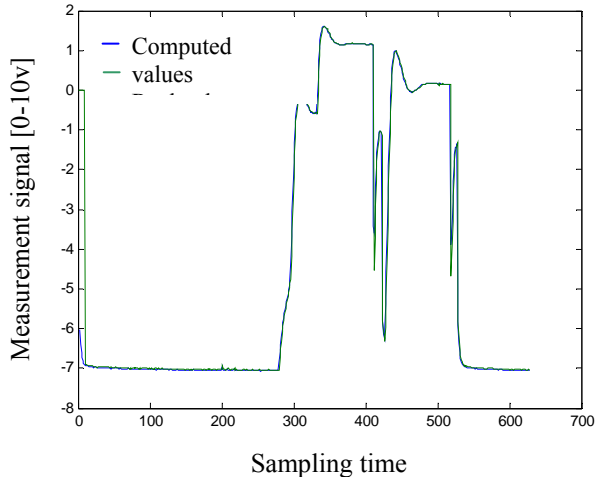Fig. 6a: Computed and real outputs
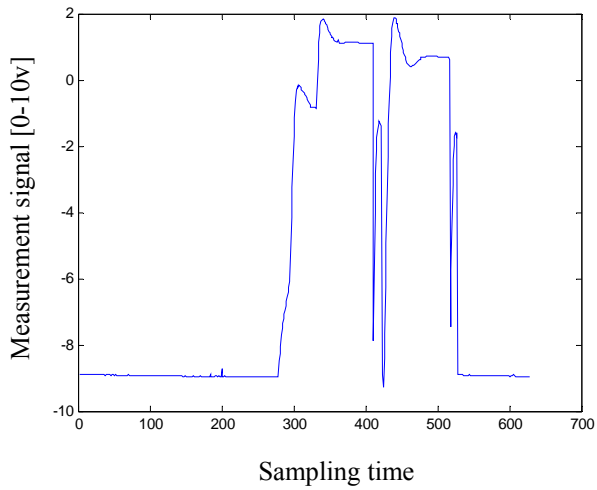
Fig.6b: Input signal

Fig.6c: Estimated parameters

Qualitative feature extraction is the abstraction of trend information. Trend analysis and prediction are important components of process monitoring and supervisory control. Trend modelling can be used to explain the various important events happening in the process, do malfunction diagnosis and predict futures states. We consider in this section a comparative study of fault detection using normal PCA and PCA coupled to RLS algorithm. Application has been made on the analysis of faulty situation of the control system of the motor M2 in Fig.4. The fault is characterised by a deviation of the rotary speed dynamic between motor M1 and M2. Motors M1 and M2 have the same mechanical and electrical characteristics. Motor M1 is controlled in closed loop, motor M2 is controlled in open loop by the control card named "Control M2" in Fig.4. This electronic card has a same input signal than that applied to the speed control system of motor M1.

Although the motors M1 and M2 have a same input signal and characteristics a difference in outputs appears see Fig.7. This fault is generating by a defect on the control module of the motor M2. Fault detection is carried out by two approaches, one based on the conventional PCA and the other is based on the PCA coupled to RLS.

Application of the algorithm developed in section (II) to the conventional PCA and PCA coupled to RLS respectively, gives the results shown in Fig.8.

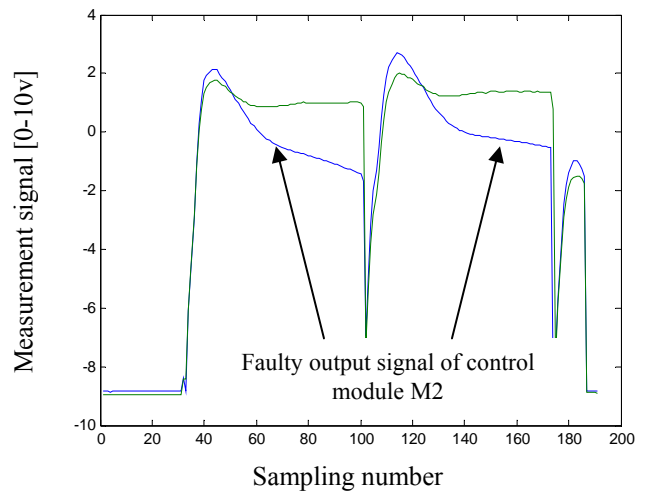Fig. 8a: Outputs of control modules M1 and M2

4.4 Application to Fault Detection and Monitoring

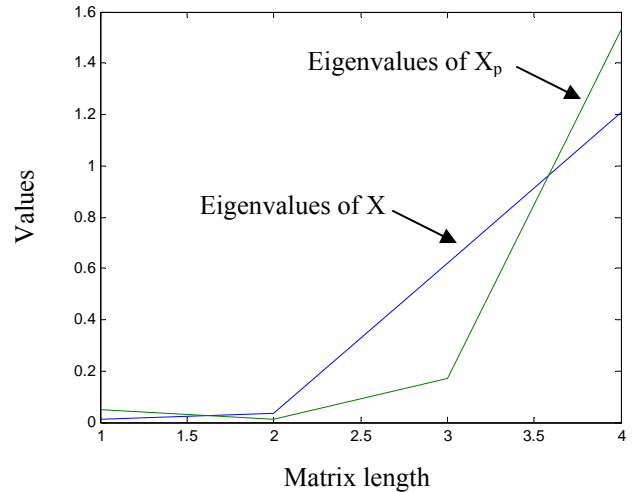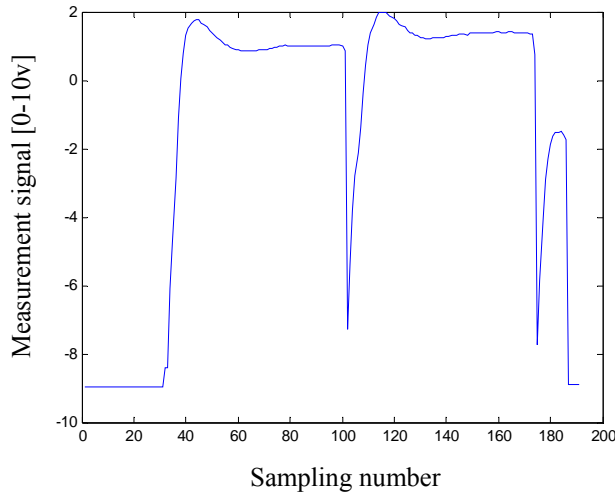Fig.8b: Input signal of control modules M1 and M2

Sampling number

Matrix length

Fig.8a shows the faulty status defined y a divergence of the signal output of the control module M2. Fault analysis is obtained by two approaches as defined in Fig.9. The first approach computes the Q statistic (Q_x) using the variability matrix X with input - output data, the second is similar, it computes the Q statistic (Q_p) using the matrix parameters X_p defined by equation (24). Performances are given in Fig.9 and Fig.10. Q statistic of Q_p characterising the defect is more stable out of the acceptable range. This is confirmed also confirmed by the analysis of the eingenvalues of the matrix X_p where more of 80% of the variations are expressed by the two last values against the three last using the matrix X (Fig.10).
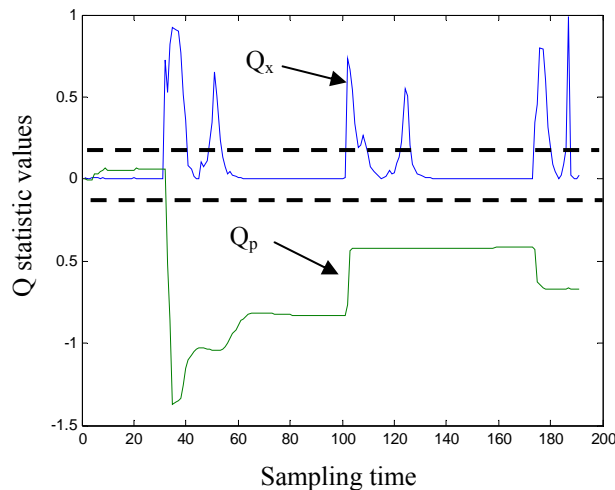
Fig. 9: Q statistic

Sampling time

Fig. 10: Eigenvalues of X and X_p

5 Conclusion

We developed in this work an extension of conventional PCA approach operating on the input output process data to the PCA coupled to RLS. The obtained results show that the extension of such approach improves considerably the fault detection stability compared to the conventional PCA method. Results have been confirmed by the application on a real fault of rotary machine control system in skin pass process.

References:
[1] S.Bouhouche, Contribution to Quality and Process Optimisation using Mathematical Modelling, *Ph.D thesis, Institut fûr Maschinenbau,* ISSN :1617-3309, ID-Nr :104 TU Bergakademie Freiberg Germany, 2002
[2] M.Norgaard, O. Ravn, N.K. Poulsen and L.K. Hansen,Neural Nerworks for Modelling and Control of Dynamic Systems, *Springer-Verlag* London Berlin *Heidelberg Second Edition*, ISBN 1-85233-227-1, 2001,
[3] L.H.Chiang, E.L. Russel land R.D. Braatz, Fault Detection and Diagnosis in Industrial Systems, *Springer-Verlag* London Berlin Heidelberg Second Edition, ISBN 1-85233-327-8, 2001
[4] S.Bouhouche, M.S.Boucherit and M.Lahreche, Improvement of Breakout Detection System in Continuous Casting Process using Neural Networks, *IEEE Proceedings on Advanced Process Control - Applications for Industry Workshop*, Vancouver, Canada, 63, 2001, pp 53-
[5] G.Sorgrl, T.Poppe and M.schlang, Real-time Control with Neural Networks in Steel Processing, *European Commission for Technical Steel Research, ECSC Workshop, Proceedings of Application of Artificial Neural Network Systems in the Steel Industry*, Brussels 22-23 January, 1998, pp 209-225
[6] D.Pham and X.Liu, State Space Identification of

Dynamic Systems using Neural Networks, *Engineering Application in Artificial Intelligence*, (3), , 1990, pp 198-203

[7] D.Lee, J.S.Lee and T.Kang, Adaptive Fuzzy Control of the Molten Steel Level in a Strip Casting Process, *Control Engineering Practices*, 4, (11), 1996, pp 1511-1520

[8] C.Harris, M.Brown, K.M.Bossley, D.J.Mills and F.Ming, Advances in Neuro-Fuzzy Algorithms for Real-time Modelling and Control, *Engineering Application of Artificial Intelligence*, 9, (1), 1996, pp 1-16

[9] T.Kim and S.R.T.Kumara, Boundary Defect Recognition using Neural Networks, *International Journal of Production Research*, 35, (9), 1997, pp 2397-2412

[10] A.P.Loh, K.O.Looi and K.F.Fong, Neural Network Modelling and Control Strategies for a pH Process, *Journal of Process Control*, 15 (6), , 1995, pp 355-362

[11] W.Zhenni, D.Christine, T.Ming and J.A.Morris, A Procedure for Determining the Topology of Multilayer Feed-Forward Neural Networks, *Neural Networks,* 7, (2), 1994, pp 291-300

Dynamic Systems using Neural Networks, Engineering Application in Artificial Intelligence, (3), 1990, pp 198-203.

[13] S Bouhouche, MS Boucherit, M Lahreche and J Bast, Controlled Solidification in Continuous Casting using Neural Networks, *Book EPMESC IX Conference*, Hong Kong November 2003

[14] R. Isermann and P. Ballé, Trends in the Application of Model- Based Fault Detection and Diagnosis of Technical Processes, *Control Engineering Practice*, Vol5, N°5, , 1997, pp 709-719

[15] C. Comstock Mathew and E. B James, Literature Review for Application of Fault Detection and Diagnosis Methods to Vapour Compression Cooling Equipment, *HL 99-19, Report 4036-2*, 1999.

[16] Bouhouche.S, Lahreche M, Ziani S and Bast J, Fault Detection and Monitoring of Length Loop Control System in Pickling Process, *CIMCA2005, IEEE Conference on computational intelligence in Modelling, Control and Automation*, VIENNA 28-30 NOV. 2005.

[12] D.Pham and X.Liu, State Space Identification of