An Improved PCA Fault Detection for the Diagnosis

N. PESSEL[1], J-F. BALMAT[2], F. LAFONT[1] and J. BONNAL
LSIS, UMR CNRS 6168, University of South-Toulon-Var
[1]IUT of Toulon, [2]Faculty of Sciences and Technologies
BP 20 132, 83 957 La Garde Cedex
FRANCE

Abstract: - This paper presents a fault detection and isolation method based on the design of a non linear PCA model and a Fisher Discriminant Analysis (FDA). A new fault detection approach based on the estimation of the prediction error (SPE: Squared Prediction Error) by the non linear PCA model is proposed. This method associates an adaptative thresholding with the study of the dynamic of the SPE. It allows to define several operating regions. The fault isolation is based on the pairwise FDA analysis applied to a class without fault and each class with fault.
In this study, this new diagnosis method is validated in simulation on a quadruple-tank process. Three types of fault are simulated in a sensor: a drift, a bias and a breakdown of sensor.

Keywords: - Fault diagnosis, Neuronal Principal Component Analysis, Fischer Discriminant Analysis, Adaptative thresholding

1 Introduction

Fault detection and isolation (FDI) is a subject of research motivated by the increasing requirements on the design of highly reliable control systems [15].
Different approaches for fault detection using mathematical models have been developed [6] and give valid results. Traditionally, these approaches make use of residual generation and of their evaluation for decision making. Thus, it is possible to detect and localize the default. However, it is often difficult to develop accurate mathematical models that characterize all the physical phenomena occurring in the industrial processes [4].
One alternative is to consider the implicit modeling approaches [10], based on the data-driving techniques such as the linear Principal Component Analysis (PCA) or non linear. These methods are well adapted to emphasize the relationships between the plant variables without the explicit expression of the system model [9].
The proposed approach is based on the design of a non linear PCA model and a Fisher Discriminant Analysis (FDA) [4][12].
This method is composed of three stages:

- in first stage, a data analysis allows to isolate normal and abnormal data clusters (data with and without fault). The detection index SPE (Squared Prediction Error) indicates how much each sample deviates from the PCA model [3],
- secondly, a fault visualization in the principal components 2-D space is effected by performing a global FDA,
- thirdly, a study of the fault directions in pairwise FDA allows to localize the default.

In this paper, the non linear PCA model is designed by a multilayer neural network [8]. We present an improvement of the operating region definition by studying the dynamic of the SPE.
To demonstrate the advantage of the method, we propose a simulation example.
The detection and localization of a drift, bias and breakdown of sensor default are tested. The type of defaults often appears on the sensors in the industrial systems.
This paper is organized as follow: first, we present the principle of non linear PCA modeling based on the neural networks. Then, we summarize the different stages of sensor fault detection and we detail our approach of fault detection. In section 4, we present the application of the method to diagnosis of a quadruple-tank process.

2 Non linear PCA by Neural Networks

We chose to use the approach of Kramer which is based on the implementation of a multilayer neural network [8].
These neural networks are a special class of artificial neural networks which are able to learn the principal components without explicitly solving the eigenvalues and eigenvectors from the sample covariance matrix [9][2]. The extraction of the principal components can be carried out of sequential or parallel way. We have chosen to use this last method. This non linear approach of the PCA boils down to define a neural network with five layers. The apprenticeship is carried out by minimizing the squared error between the inputs and outputs of the network.

The network has feed-forward architecture. It contains a bottleneck layer which allows to represent data in the under-space create by principal components [5][11][14] [13]. The first layer corresponds to m inputs (data matrix) and the fifth layer corresponds to estimated inputs. The data are normalized (centred and reduced). The activation function of the second layer, called encoding layer, and of the fourth layer, called decoding layer, is a sigmoid. The choice of the number of neurons in these layers is bind to the number of the constraints imposed by the set of apprenticeship data (formed of n samples).
The neurons of the bottleneck layer represent the non linear principal components.

The choice of the neurons number of the bottleneck layer is based on the estimation of an index ε which depends of the estimation error:

$$\varepsilon = \sqrt{\frac{\left\|\hat{X} - X\right\|^2}{\left\|X - \overline{X}\right\|^2}} \qquad (1)$$

where X represents the input data matrix, $\hat{X}$ the outputs of the network and $\overline{X}$ is a matrix of which the rows are composed of the X mean vector.

3 Fault detection

In this part, we present the fault detection method inspired of the approach proposed by He, *et al*. [4]. It is based on a linear PCA model. Our first contribution consists to extend this method to the non linear systems. Thus, knowing that the linear PCA isn't adapted, we propose to use a Neuronal Non-Linear PCA (NNLPCA). Our second contribution is located in the used method to determine the presence of a fault and its time of beginning and of end. For this, we propose a new error study method between the input data and the outputs of the NNLPCA.

3.1 Data analysis

In this part, we develop a data analysis method which allows to determine the number of the classes with and without default.
Sensor fault detection is based on a similarity test between the measured and the estimated data with the NNLPCA model.
Thus, to detect the errors, the typical statistic used is the Squared Prediction Error (SPE):

$$SPE = \sum_{i=1}^{m} (X_i - \hat{X}_i)^2 \qquad (2)$$

where m is the number of the variables.

The detection of an anomaly is carried out by comparing the SPE with a threshold δ defined starting from the average error of the model estimation. From this SPE, with regard to δ, the existence of an abnormal behavior of sensor is detected when :

$$SPE > \delta \qquad (3)$$

He, *et al.* [4] use the δ^2 expression developed by Jackson and Mudholkar. In this paper, we propose a new fault detection method based on the SPE. We implement, in parallel, two algorithms.
The first algorithm is based on an adaptive threshold (a localized average of the SPE) which is compared with the envelope of the SPE (minimum and maximum localized of the SPE). This approach allows to detect a fault when the localized average of the SPE is outside the envelope (see Fig. 1). However, when the signal variation is weak, the envelope is near of the SPE signal and the false alarm rate is important.

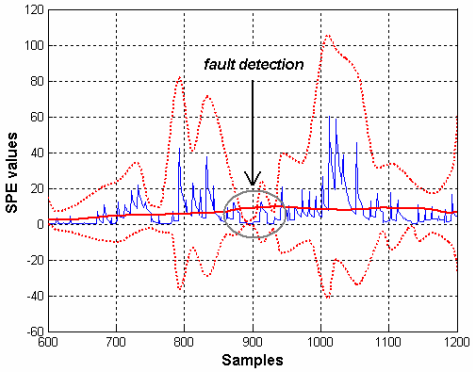

Fig. 1. Fault detection illustration

The second algorithm consists to apply a classical gradient convolution mask to the SPE signal. This method allows to detect the strong variations of the SPE. The final result is the intersection of the two methods. Thus the false alarm rate is reduced. The Fig. 2 presents the principle of the fault detection method.

This first analysis allows to define working classes by keeping the data chronology. The k classes are viewable on the sample projection in the non linear principal component plane.
The proximity of the classes varies in function of the anomaly importance. He, *et al*. [4], propose to use the k-means algorithm to separate a set of data into mutually exclusive clusters and thereby, to visualize the classes dispersion [1].
The samples of the transition area between many classes will be removed of the data set. So, we obtain k distinct classes of samples: C_1, C_2, … , C_k.

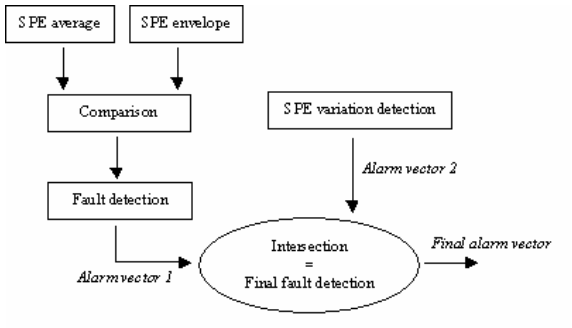

Fig. 2 Principle of the fault detection method

3.2 Fault visualization

In this section, we develop the criterions of the discrimination retained for the isolation of the class(es) presenting a sensor fault.

The fault visualization exposed by He, *et al* [4], allows to project the classes in a low-dimensional space by using the Fisher Discriminant Analysis (FDA). The discrimination nature of FDA allows to emphasize the separation between the classes with and without fault. This separation is carried out in determining the Fisher optimal discriminant vector such that the Fisher criterion function is maximized. This estimation is based on the between-class and without-class scatter matrices. Thus, it is possible to isolate the classes with fault.

These two first steps of the diagnosis method allow to detect the presence of an anomaly and to locate it in the time. The last step consists in locating the variable which presents a fault.

3.3 Fault localization

The fault localization is based on the pairwise FDA analysis applied to a class without fault and the considered class with fault.

For each association of two classes, we search the Fisher direction which separates the data with fault, denoted C_i ($i = 1 \ldots k-1$) from data without fault denoted C_0 called normal data.

The characteristic Fisher direction for the class i defines the direction of the fault for the class C_i. The j^{th} element ν_j of the vector $\omega_i = [\nu_1, \cdots, \nu_j, \cdots, \nu_q]^T$ is the contribution of the j^{th} variable in the class presenting a fault.

Through Fisher direction vectors ω estimated for each class with fault, we can determine the incriminated variable.

4 Experiments

The validation of the diagnosis method has been carried out in simulation on a benchmark: a hydraulic process.

We describe in a first section, the benchmark used. In the second section are exposed many results obtained using the diagnosis method.

4.1 Process description

The approach proposed previously has been validated on the quadruple-tank process. The quadruple-tank process was developed by Johansson [7] as a novel multivariate laboratory process. This process consists of four interconnected water tanks, two pumps and two associated valves (Fig 3).

The inputs are the voltages supplied to the pumps, υ_1 and υ_2, and the outputs are the water levels h_1–h_4. The flow to each tank is adjusted using the associated valves γ_1 and γ_2.

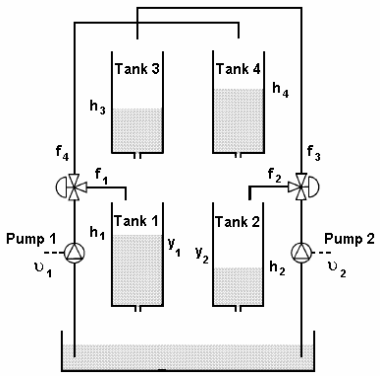

Fig 3. An hydraulic process

A non linear model is derived based on mass balances and Bernoulli's law:

$$\frac{dh_1}{dt} = -\frac{a_1}{A_1}\sqrt{2gh_1} + \frac{a_3}{A_1}\sqrt{2gh_3} + \frac{\gamma_1 k_1}{A_1}\upsilon_1 \qquad (4)$$

$$\frac{dh_2}{dt} = -\frac{a_2}{A_2}\sqrt{2gh_2} + \frac{a_4}{A_2}\sqrt{2gh_4} + \frac{\gamma_2 k_2}{A_2}\upsilon_2 \qquad (5)$$

$$\frac{dh_3}{dt} = -\frac{a_3}{A_3}\sqrt{2gh_3} + \frac{(1-\gamma_2)k_2}{A_3}\upsilon_2 \qquad (6)$$

$$\frac{dh_4}{dt} = -\frac{a_4}{A_4}\sqrt{2gh_4} + \frac{(1-\gamma_1)k_1}{A_4}\upsilon_1 \qquad (7)$$

where A_i is the cross-section of tank i, a_i is the cross-section of the outlet hole, and h_i is the water level.

Parameter	Unit	Value
A_1, A_3	cm^2	28
A_2, A_4	cm^2	32
a_1, a_3	cm^2	0.071
a_2, a_4	cm^2	0.057
k_c	$v.cm^{-1}$	0.50
g	$cm.s^{-2}$	981

Table 1 Simulation parameters

The voltage applied to pump i is υ_i and the corresponding flow is $k_i\upsilon_i$. The parameters γ_1, $\gamma_2 \in (0, 1)$ are determined from how the valves are set prior to an experiment. The water flow to tank 1, f_1, is $\gamma_1 k_1 \upsilon_1$ and the flow to tank 4, f_4, is $(1-\gamma_1)k_1\upsilon_1$ and similarly for tanks 2 and 3. The acceleration of gravity is denoted g. The measured level signals are $k_c h_1$ and $k_c h_2$. The parameter values of the laboratory process are given in Table 1. The model and control of the quadruple-tank process are studied at an operating point. The parameters values are specified in Table 2.

The data are generated by equations 4 to 7. The parameters υ_i are independently corrupted by Gaussian white noise with zero mean and standard deviation of 0.1. The measured water levels h_i are corrupted by Gaussian distributed white noise with zero mean and standard deviation of 0.2.

Parameter	Unit	Value
(h_1^0, h_2^0)	cm	(12.4, 12.7)
(h_3^0, h_4^0)	cm	(1.8, 1.4)
$(\upsilon_1^0, \upsilon_2^0)$	V	(3.00, 3.00)
(k_1, k_2)	$cm^3.V^{-1}.s^{-1}$	(3.33, 3.35)
(γ_1, γ_2)		(0.70, 0.60)

Table 2 Operating point parameters

4.2 Results

We strictly concentrated on the diagnosis of sensor defaults. The considered sensors are the four heights h_1-h_4 and the four water flows f_1–f_4.

The proposed approach is based on a NNLPCA model with five layers with five neurons for the encoding layer, four for the bottleneck layer and five neurons for the decoding layer.

This study presents results carried out from simulated data files modified with simulated defaults and noise. We generated 2000 normal data samples. We were interested in sensor default of drift, bias and breakdown types. Many tests were carried out; each type of default is applied on the sensor h_4.

4.2.1 Sensor default of drift type

For this first experiment, we simulated a sensor fault of drift type in the fourth height h_4. The slope of the drift is equal to -0,3 % from the sample 1000 to the end of the simulation.

The Fig. 4 shows the time series data of the process variables: water levels h_1–h_4 and water flows f_1–f_4 with a drift in h_4 from the sample 1000. We can see the noise introduced for each data.

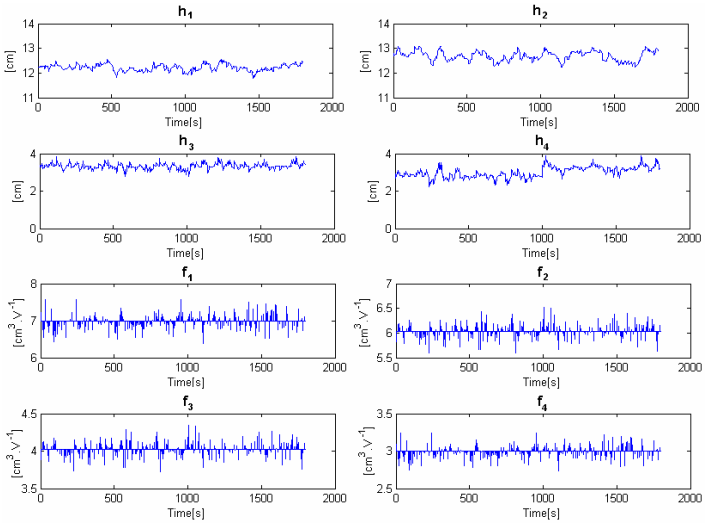

Fig. 4. Process time series data with sensor fault of type drift in h_4

The detection of the fault sensor is carried out using the Squared Prediction Error associated to the "envelope method".

The "envelope method" illustrated by the Fig. 5 allows to detect the presence of two strong variations of the SPE, at the sample 1104 and close to the end of the simulation. So, we note the existence of 3 operating regions. The difference of the detection time is inversely proportional at the percent of the slope. The beginning and end temporal information of each operating region allow to project independently the samples associated to each class in the non linear PCA score space.

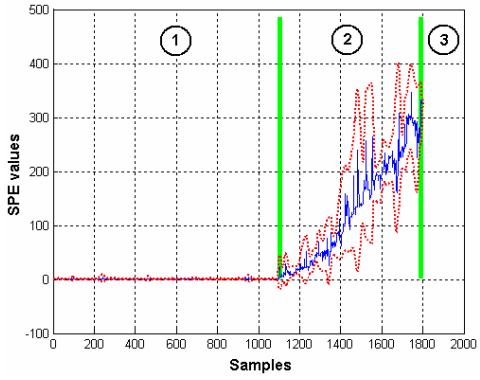

Fig. 5. SPE

The Fig. 6 presents the projections of the remaining samples in the non linear PCA score space. The samples situated in the transition areas are removed. The detachment of the second and the third operating regions from the first, translates the presence of a fault on these samples.

To determine the root cause of the fault, we examine the contribution of each variable in the second and third operation regions (Fig. 7).

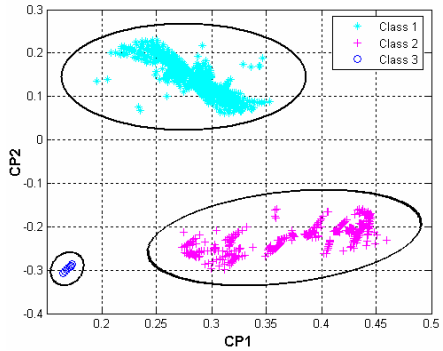

Fig. 6. Clusters in the non linear PCA score space after deleting transitional samples.

This figure clearly shows that the variable 4, associated to the sensor h_4, is the only root cause of the fault in the two operating regions 2 and 3. In conclusion of this first experiment, we can say the sensor h_4 presents a default from the sample 1107 to the end of the simulation.

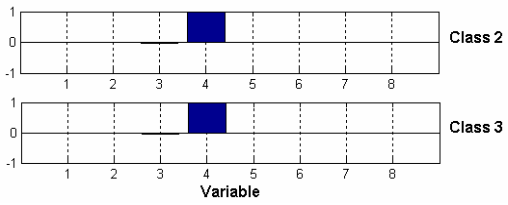

Fig. 7. Contributions of each variable in the classes 2 and 3

4.2.2 Sensor default of bias type

In this case, a bias in the sensor h_4 bas been simulated from the sample 1000. The amplitude of this bias is equal of $\Delta h_4 = 0.4$.

The Fig. 8 shows the time series data of the process variables: water levels h_1–h_4 and water flows f_1–f_4 when a bias of 0.4 is applied in h_4. The presence of this default modifies very lightly the h_4's plot.

The "envelope method" coupled to the SPE allows to detect three operating regions Fig. 9. The second begins to the sample 1013 and the third to the sample 1310.

After this first step, we project the samples taking into account the temporal information of each operating regions, in the non linear PCA score space is presented Fig. 10.

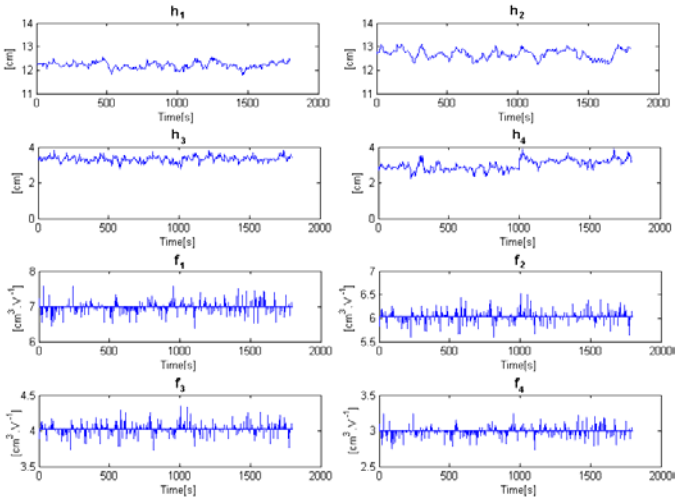

Fig. 8. Process time series data with sensor fault of type bias in h_4

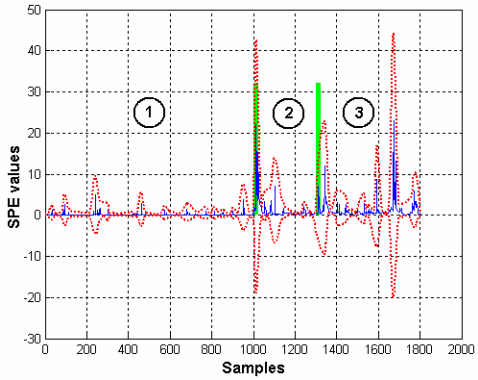

Fig. 9. SPE

The Fig. 10 allows to note that the operating regions 2 and 3 present the same fault, because they are confused and separate from the operating region composed of "normal data". The formation of two operating regions from the sample 1013, defined by the "envelope method" applied to the SPE, is a mistake may be due to a modeling error.

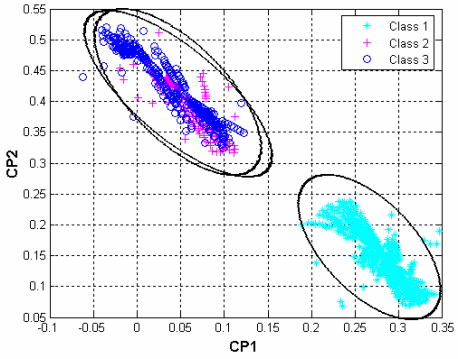

Fig. 10. Clusters in the non linear PCA score space after deleting transitional samples

The contributions of each variable in the classes 2 and 3 (Fig. 11), allow to underline the fourth variable.

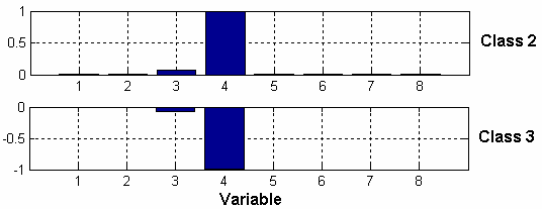

Fig. 11. Contributions of each variable in the classes 2 and 3

In conclusion, this second study shows the presence of a default in the sensor h_4 from the sample 1013 to the end of the simulation.

4.2.3. Breakdown sensor

In this last simulation, we study the breakdown of the sensor h_4 from the sample 1000.

The Fig. 12 presents the detection of a default via the SPE associated to the "envelope method". We clearly see two operating regions. By comparing the Fig. 9 to this of the Fig. 12, we remark that the SPE amplitude has increased of 40 times. The increase is due to the important fall of the h_4 values from the sample 1007.

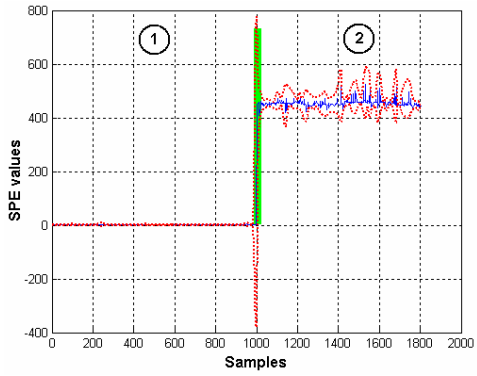

Fig. 12. SPE

The Fig. 13 presents the projections of the remaining samples in the non linear PCA score space. The SPE amplitude jump corresponds to the important separation between the two classes (Fig. 13). This remark confirms that the second operating region present a default.

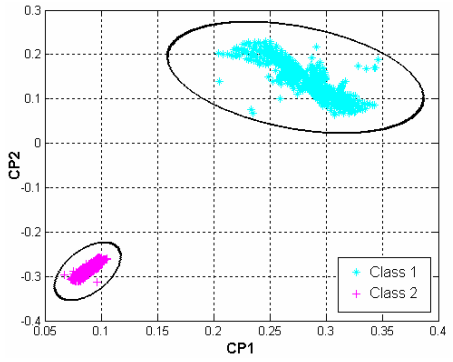

Fig. 13. Clusters in the non linear PCA score space after deleting transitional samples

From the contributions of each variable in the class 2 (Fig. 14), the variable 4 is the root cause of the fault.

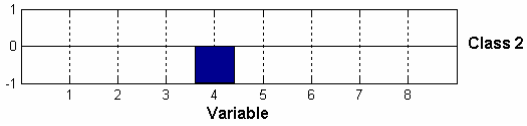

Fig. 14. Contributions of each variable in the class 2

In conclusion of this third experiment, we can say that the sensor h_4 presents a default from the sample 1007 to the end of the simulation.

The results presented in this part allow to validate the method of detection in the simulation case of a sensor default of drift, bias and breakdown types.

5 Conclusions

This paper presents a detection and isolation method for the systems with a non linear behavior. The proposed approach is based on the diagnosis method developed by He, *et al.* [4]. We propose in this study two evolutions. Our first contribution is to replace the linear PCA model with a Neuronal Non Linear PCA model in order to adapt this diagnosis method to the Non Linear systems [2]. The second contribution is to modify the method used to detect the fault presence. We replace the comparison of the SPE with a fix threshold and we define an adaptative thresholding associates to the study of variations of the SPE.

The results obtained by this method are very satisfactory. To through the application on a quadruple-tank process, we come to show that it is possible to detect and locate a default of sensor of the bias type, drift or even total breakdown.

References:

[1] P. N. Belhumeur, J. P. Hespanka and D. J. Kriegman, Eigenvalues vs. Fisherfaces: Recognition Using Class Specific Linear projection, *IEEE Transactions on Pattern analysis and machine intelligence*, Vol. 19, n°7, pp. 711-720, 1997.

[2] M. F. Harkat, G. Mourot and J. Ragot, An improved PCA scheme for sensor FDI: Application to an air quality monitoring network, *Journal of Process Control*, Vol. 16, pp 625-634, 2006.

[3] T. Hastie and W. Stuetzle, Principal Curves, Journal of the American Statistical Society, Vol. 84, n° 406, pp. 502-516, 1989.

[4] Q. P. He, S. J. Quin and J. Wang, A new fault diagnosis method using fault directions in Fisher discriminant analysis, *In AIChE Journal*, Vol. 51, Issue 2, pp. 555-571, 2004.

[5] W.W. Hsieh, Nonlinear principal component analysis and extensions, Dept. of Earth & Ocean Sciences, University of British Columbia, 2004.
[6] R. Isermann, *Model-based fault-detection and diagnosis – status and applications*, Annual Reviews in Control, 29, pp. 71-85, 2004.
[7] K. H. Johansson, The Quadruple-Tank Process: A Multivariable Laboratory Process with an Adjustable Zero, IEEE Transactions on control Systems Technology, Vol. 8, n° 3, May 2000.
[8] M. A. Kramer, Nonlinear principal component analysis using autoassociative neural networks, *AIChe Journal*, Vol. 37, n° 2, pp. 233-243, 1991.
[9] J. V. Kresta, J. F. MacGregor and T. E. Marlin, Multivariate statistical monitoring of process operating performance, *The Canadian Journal of Chemical Engineering*, Vol. 69, n°1, pp. 35-47, 1991.
[10] R. Linker, P. O. Gutman and I. Seginer, Robust model-based failure detection and identification in greenhouses, *Computers and Electronics in Agriculture*, Vol. 26, pp 255-270, 2000.
[11] E. C. Malthouse, Limitations of Nonlinear PCA as Performed with Generic Neural Networks, IEEE Transactions on Neural Networks, Vol. 9, n° 1, pp. 165-173, 1998.
[12] N. Pessel, J-F. Balmat and N. K. M'Sirdi, Analyse discriminante pour le diagnostic. Application à une serre agricole expérimentale, In Proceedings of the IEEE Conférence Internationale Francophone d'Automatique, 2006.
[13] M. Scholz and R. Vigario, Nonlinear PCA: a new hierarchical approach, In Proc. European Symposium on Artificial Neural Networks, pp. 439-444, 2002.
[14] S. K. Sengupta and J. S. Boyle, Nonlinear Principal Component Analysis of Climate Data, PCMDI Program for Climate Model Diagnosis and Intercomparison, Report n°29, 1995.
[15] A. Xu and Q. Zhang, *Nonlinear system fault diagnosis based on adaptive estimation,* Automatica, Volume 40, Issue 7 , July 2004, Pages 1181-1193.