**Proceedings of the 7th IFAC Symposium on
Fault Detection, Supervision and Safety of Technical Processes
Barcelona, Spain, June 30 - July 3, 2009**

Fault Detection and Diagnosis Based on PCA and a New Contribution Plots ⋆

Baligh Mnassri * **El Mostafa El Adel** * **Bouchra Ananou** * **Mustapha Ouladsine** *

* *Laboratory of Sciences of Information's and Systems (LSIS - UMR CNRS 6168); University Paul Cézanne (Aix-Marseille III) Domaine Universitaire de Saint-Jérôme, Avenue Escadrille Normandie-Niemen; 13397 Marseille Cedex 20, France (Tel: +33(0)491056030; e-mail: {baligh.mnassri, mostafa.eladel, bouchra.ananou, mustapha.ouladsine}@lsis.org).*

Abstract: The fault detection and diagnosis methods based on principal component analysis (PCA) have been developed widely because they need no prior knowledge about the process mechanism model and can detect faults rapidly. The principal components score plot (T2) and squared prediction error (SPE) plot are used to detect the abnormal process operation condition. However to identify the faulty variables, the existing diagnosis algorithms such as contribution plots still have some trouble when they are applied in real industrial processes. In this paper, the idea is grounded on the fact that the classical contribution to SPE, available in the literature, gives erroneous results. The proposed diagnosis method utilizes two new contribution plots to give accurate cause for abnormal process conditions that indicate more precise information about the failure cause. The results show that accurate conclusion could be obtained easily by these methods.

Keywords: Diagnosis; Fault detection; Identification; PCA; SPE.

1. INTRODUCTION

For the successful operation of any process, it is important to detect process upsets, equipment malfunctions, or other special events as early as possible, and then to find and move the factors causing those events. In industrial processes, data-based process monitoring methods, referred to as statistical process control, have been widely used. Conventional Shewhart control charts, cumulative sum (CUSUM) control charts and exponentially weighted moving-average (EWMA) control charts are well established for monitoring univariate processes (de Vargas et al. [2004]). However, multivariate statistical process control (MSPC) has been applied to monitor such multivariable processes. MSPC has been developed as an alternative to theoretical model-based controlling approaches for the analysis of data obtained from diverse categories of process, which are highly noisy, significantly correlated, and redundant. To improve the performance of monitoring, the detection of changes in the process requires a modeling technique that captures the major relations among the process variables. In particular, principal component analysis (PCA) is a powerful MSPC technique which has been widely used to monitor such modern industrial processes. It is a reliable technique for capturing variable correlation.

Historically, the original concept of PCA was introduced by Pearson and developed by Hotelling. The vast majority of applications happened only in the last few decades with the popular use of computers. This technique has wide applications in signal processing, factor analysis, chemical processes data analysis... Several papers have illustrated the use of PCA in process monitoring and statistical process control and fault diagnosis. The applications of this tool vary from batch processes to continuous processes.

The PCA is used to define an orthogonal partition of the measurements space into two orthogonal subspaces: a principal component subspace (PCS) and a residual subspace (RS). PCA-based monitoring can be used to: extract information from regularly obtained data, construct model and residual spaces and determine the control limits of both subspaces. After the construction of the PCA model, test data are projected into the two subspaces. If abnormalities arise in the process operation, PCA monitoring charts as the T2 chart and the squared prediction error (SPE) chart can be used to detect them (Xu and Wang [2006]). After an abnormality is detected in a process operation, a reliable method is needed to identify the variables that cause this abnormality. Thus far, the most popular approach to diagnosis is the contribution plot approach (Kourti [2005], Kourti and MacGregor [1996], Kourti et al. [1995]). Contrary to the reconstructed-based approach (Qin [2003], Lee et al. [2004]), the contribution approach does not require any information about the fault to generate the plots. The contributions are actually the effects of the fault on the observed vector of measurements. In this context, we must clarify that the most common indices used for fault diagnosis with contribution plots are T2 and SPE. This means that each variable participates

⋆ This work was supported in part by the Laboratory of Sciences of Information's and Systems (LSIS - UMR CNRS 6168). http://www.lsis.org

with two contributions that are related respectively to the two detection's distances: T2 and SPE. In earlier work (see Mnassri et al. [2008]), we have proved the limitations of the classical contribution to the SPE available in the literature and we have proposed an improved form for this contribution which is based on the knowledge of the residues. To do this, we can summarize its calculation in the following steps: for each variable, we firstly calculate its contribution to its residue and to all other residues of all other variables because the SPE is the sum of the squares of residues. Secondly, we keep only the non negative values among them.

In order to achieve our main objective in this paper, we would like first to improve an overview and analysis of recently developed process monitoring methods for fault identification based on contribution plots. Secondly, we would like to propose a other newly form for the contribution to SPE. This new approach is calculated on the principal components (PCs) of the RS. It is advantaged by its simplicity and an important gain in computing time.

The present paper is organized as follows: in section 2, we recall the PCA approach and their two statistics (T2 and SPE) which served to monitoring. After fault detection, the third section involves the step of fault identification based on contribution plots. This is followed by a description of the contributions available in the literature. This section is finished by describing our earlier new approach (Mnassri et al. [2008]) to calculate the contribution to the SPE that is based on residues. As a continuation of this work, we propose another improved approach to calculate the same contribution. To compare their efficiencies with the classical contribution to the SPE, we will present two examples of data for fault diagnosis in section 4. Finally, some concluding remarks are given in section 5.

2. FAULT DETECTION AND DETECTABILITY

Statistical process monitoring relies on the use of normal process data to build process models which include from PCA that is one of the most popular statistical methods for extracting information from measured data. It finds the directions of significant variability in the data by forming linear combinations of variables. In other words, PCA models are predominantly used to extract variable correlation from data.

2.1 Principal component analysis

Let $x \in \Re^m$ denotes a sample vector of m variables. Assuming that there are n samples for each variable, a data matrix $X \in \Re^{n\times m}$ is composed with each row representing a sample. To avoid the scaling problem, we consider that the data are scaled to zero mean and to unit variance. PCA determines an optimal linear transformation of the data X in terms of capturing the variance in the data. Hence, the matrix X can be decomposed into a score matrix T and a loading matrix P as follows:

$$X = TP^T \tag{1}$$

with $T = [t_1\ t_2 \cdots t_m] \in \Re^{n\times m}$, where the vectors t_a are called scores or PCs and the matrix $P = [p_1\ p_2 \cdots p_m] \in \Re^{m\times m}$, where the orthogonal vectors p_a called loading or principal vectors, are the eigenvectors associated to the eigenvalues λ_a of the correlation matrix Σ of X such that:

$$\Sigma = P\Lambda P^T \quad with \quad PP^T = P^T P = I_m \tag{2}$$

$\Lambda = diag\,(\lambda_1 \cdots \lambda_m)$ is a diagonal matrix with elements in the decreasing order.

The partition of eigenvectors and principal components matrices gives respectively:

$$P = [\hat{P}_\ell \mid \tilde{P}_{m-\ell}], \quad T = [\hat{T}_\ell \mid \tilde{T}_{m-\ell}] \tag{3}$$

where ℓ represents the number of the more significant PCs which are sufficient to explain the variability of the process through their data X. This number will be defined bellow.

The first ℓ eigenvectors constitute the representation space or the PCS defined by $S_p = span\{\hat{P}_\ell\}$, whereas the RS is: $S_r = span\{\tilde{P}_{m-\ell}\}$. A sample vector x can be projected on the PCS and RS respectively:

$$\hat{x} = \hat{P}_\ell \hat{P}_\ell^T x = \hat{C}x \in S_p \tag{4}$$

$$\tilde{x} = \tilde{P}_{m-\ell} \tilde{P}_{m-\ell}^T x = \tilde{C}x \in S_r \tag{5}$$

where $\hat{C}$ and $\tilde{C} = (I - \hat{C})$ represent respectively the projection matrices on the PCS and RS with ($\ell < m$). Since S_p and S_r are orthogonal,

$$\hat{x}^T\tilde{x} = \tilde{x}^T\hat{x} = 0 \quad and \quad x = \hat{x} + \tilde{x} \tag{6}$$

The vectors $\hat{x}$ and $\tilde{x}$ represent respectively the modeled and unmodeled variations of x based on ℓ components. The task for determining the number of PCs is to choose ℓ such that $\hat{x}$ contains mostly information and $\tilde{x}$ contains noise.

2.2 Fault detection indices

Fault detection is usually the first step in MSPC. A typical statistics for detecting abnormal conditions are the SPE (or Q-statistic) and Hotelling's T2 (or D-statistic) which represent the variability in the RS and PCS respectively (Westerhuis et al. [2000]).
Hotelling's T2 statistic measures variations in the PCS and is expressed by using the estimated value $\hat{\Sigma}$ of the correlation matrix as follows:

$$T2 = x^T \hat{\Sigma}^{-1} x = \hat{t}^T \hat{\Lambda}^{-1} \hat{t} = \sum_{a=1}^{\ell} \frac{t_a^2}{\lambda_a} \tag{7}$$

Under the condition that the process is normal and the data follow a multivariate normal distribution, the T2 can be well approximated by a Chi2 distribution with ℓ degrees of freedom and a significance level or quantile $(1 - \alpha)$, see Hubert et al. [2005], Verboven and Hubert [2005]. The process operation is considered normal if:

$$T2 \leq \chi^2_{\ell,1-\alpha} \tag{8}$$

A change in variable correlation indicates an unusual situation because the variables do not conserve their normal relations. Under this situation, the sample x increases its projection to the RS. As a result, the magnitude of $\tilde{x}$ reaches unusual values compared to those obtained during normal conditions (Xu and Wang [2006]). SPE is the magnitude of $\tilde{x}$. Hence its expression is given by:

$$SPE = ||\tilde{x}||^2 = \sum_{j=1}^{m} \tilde{x}_j^2 = ||\tilde{C}x||^2 = \tilde{t}^T\tilde{t} = \sum_{a=\ell+1}^{m} t_a^2 \quad (9)$$

The process is considered normal if the SPE statistic is under its control limit which is expressed as follows, see Engelen et al. [2005], Hubert et al. [2005]:

$$SPE \leq (\hat{\mu} + \hat{\sigma} z_{1-\alpha})^3 \quad (10)$$

with $z_{1-\alpha}$ as the $(1-\alpha)\%$ quantile of Gaussian distribution Φ. $\hat{\mu}$ and $\hat{\sigma}$ are respectively the estimated mean and standard deviation of $SPE^{2/3}$.

3. FAULT IDENTIFICATION

Contribution plots are well known diagnostic tools for fault identification. Contribution plots on PCA scores and SPE indicate the significance of the effect of each variable on the T2 and SPE indices respectively. The variables with the largest contribution are considered major contributors to the fault. In this section, we focus on the contribution-based approaches to T2 and SPE.

3.1 Contribution to T2

For an observation x and according to (1), each PCs (score t_a) is expressed as follows:

$$t_a = p_a^T x = \sum_{j=1}^{m} p_{aj} x_j \quad (11)$$

where p_{aj} is the jth component of the eigenvector p_a corresponding to the eigenvalue λ_a. This implies that the contribution of the variable x_j to the calculation of the normalized scores $(t_a/s_a)^2$, where s_a is the standard deviation of the component t_a, is:

$$c_j^{(t_a/s_a)^2} = \frac{t_a}{\lambda_a} p_{aj} x_j \quad with \quad s_a^2 = \lambda_a \quad (12)$$

Hence the total contribution of the variable x_j in the calculation of all normalized PCs is equal to its contribution to the $T2$ statistic, and it is defined by:

$$C_j^{T2} = \sum_{a=1}^{\ell} c_j^{(t_a/s_a)^2} \quad (13)$$

However, as in (Kourti [2005], Kourti and MacGregor [1996]), an improved idea has been proposed to select the more important normalized scores. The total contribution of the variable x_j is calculated on the q (among ℓ) biggest elevated normalized PCs. It is expressed as follows:

$$C_j^{T2} = \sum_{a=1}^{q \in \{1 \cdots \ell\}} c_j^{(t_a/s_a)^2} \quad (14)$$

According to the same references above (Kourti [2005], Kourti and MacGregor [1996]), the contribution to the normalized score (12) is set to zero if its sign is negative (its sign is opposed to the sign of the normalized score).

3.2 Contribution to SPE based on residues

Firstly, let us describe the classical formula of the contribution to SPE available in the literature. It is simply breaking down the summation of SPE into each element (Qin [2003], Kourti [2005], Kourti and MacGregor [1996], Kourti et al. [1995]). This will enable us to define the classical contribution of each variable to the SPE as follows:

$$CC_j^{SPE} = \tilde{x}_j^2 \quad (15)$$

However, we have proved (see Mnassri et al. [2008]) that this expression can not identify precisely the faulty variables. It indicates if each variable is well estimated or not through PCA model. Consequently, we can propose two new ways to define two new formulas of contributions to the SPE that are different to the classical method. In the same reference above, we have proposed a first improved form to this contribution. Because the SPE is the sum of m residues of all variables, we have shown that the contribution of the jth variable is the sum of all theirs contributions to all residues. That's why, it is called: contribution to SPE based on residues. Let us define a common term called "relative contribution" given by:

$$R_{uv} = -x_v \tilde{x}_u c_{uv} \quad (16)$$

where c_{uv} is the uth component of the vth column of the matrix $\hat{C}$.

The contribution of the variable x_j to the calculation of its residue $\tilde{x}_j^2$ is:

$$C_j^{\tilde{x}_j^2} = x_j\tilde{x}_j - x_j\tilde{x}_j c_{jj} = x_j\tilde{x}_j + R_{jj} \quad (17)$$

The contribution of the variable x_j to the calculation of each residue $\tilde{x}_r^2$, $(r = 1, 2, \ldots m \quad and \quad r \neq j)$, is:

$$c_j^{\tilde{x}_r^2} = -x_j \tilde{x}_r c_{rj} = R_{rj} \quad (18)$$

The total contribution of the variable x_j to SPE is the sum of m terms: the contribution to its residue $\tilde{x}_j^2$ given by (17), and all other $(m-1)$ residues $\tilde{x}_r^2$ deduced from (18). This total contribution is given as follows:

$$CI_j^{SPE} = x_j\tilde{x}_j + \sum_{r=1}^{m} R_{rj} \quad (19)$$

The negative contribution hasn't a physical sense. That's why, we make the same assumptions as in (Kourti [2005], Kourti and MacGregor [1996]) to their contribution approach to T2. So, we consider that the contributions given by (17) and (18) are respectively set to zero if their signs are negatives. To achieve these conditions, we assume that:

$$\begin{cases} if \quad r = j \;\; \& \;\; R_{jj} < -x_j\tilde{x}_j & \Rightarrow R_{jj} = -x_j\tilde{x}_j \\ if \quad r \neq j \;\; \& \;\; R_{rj} < 0 & \Rightarrow R_{rj} = 0 \end{cases} \quad (20)$$

Based on this method, variables having high contribution must be investigated.

3.3 Contribution to SPE based on PCs of the RS

In this subsection, we propose a second new approach to calculate the contribution to SPE which is based on all PCs of the RS. According to equation (9), the SPE index is expressed with the PCs of the residual subspace. Similarly to the contribution to T2, we can define a contribution to SPE with the following tasks: from (11), we deduce that the variable x_j contributes to the calculation of the square score t_a^2 by:

$$c_j^{t_a^2} = t_a p_{aj} x_j \tag{21}$$

From (9), SPE is the sum of the squares PCs of the RS. This implies that the total contribution for a variable x_j to SPE is the sum of its contributions to all $(m - \ell)$ last squares PCs as follows:

$$CII_j^{SPE} = \sum_{a=\ell+1}^{m} c_j^{t_a^2} = \sum_{a=\ell+1}^{m} t_a p_{aj} x_j \tag{22}$$

To not count the negative contribution's values, we set to zero the contribution given by (21) if its sign is negative because it is opposed to the sign of the score t_a^2. We conclude that the variables with the largest contribution are considered major contributors to the fault.

We note that this last new formula of contribution (22) gives exactly the same results as the contribution (19). However, the difference will be appeared after application of the assumptions for avoiding the negative values. In addition, it is clear that the algorithm of the second proposed formula shows a gain in computation time.

4. FAULT DIAGNOSIS

In this section, the proposed methods are applied to two models with several variable faults: a simple simulated model and a real case which represents data measures of wafers of the STMicroelectronics-Rousset society.

Before starting the two examples, let us discuss how to select the number of PCs (ℓ) which will be retained to build an efficient PCA model. A key issue in developing a PCA model is to choose the adequate number of PCs to represent the system in an optimal way. If fewer PCs are selected than required, a poor model will be obtained and an incomplete representation of the process results. On the contrary, if more PCs than necessary are selected, the model will be over parameterized and will be include noise. For this, we use and compare two approaches available in the literature.

The first criterion is the cumulative percent variance (CPV), see Valle et al. [1999]. It is a measure of the percent variance captured by the ℓ PCs:

$$CPV(\ell) = 100 \left(\frac{\sum_{j=1}^{\ell} \lambda_j}{\sum_{j=1}^{m} \lambda_j} \right) \% \tag{23}$$

with this criterion one selects a desired CPV, e.g., 90%, 95%, or 99% which is very subjective. While we would like to account for as much of the variance as possible, we want to retain as few PCs as possible. The second criterion is the variance of the reconstruction error (VRE) (Qin [2003], Qin and Dunia [2000], Valle et al. [1999]). When the PCA model is used to reconstruct missing values or faulty variables, the reconstruction error is a function of the number of PCs. The minimum found in the VRE calculation directly determines the number of PCs. The VRE of the jth variable is a function of ℓ and is defined as follows:

$$\sigma_j(\ell) = var\{\xi_j^T (x - x^j)\} = \frac{\tilde{\xi}_j^T \Sigma \tilde{\xi}_j}{\left(\tilde{\xi}_j^T \tilde{\xi}_j\right)^2} \tag{24}$$

where $\tilde{\xi}_j = \tilde{C} \xi_j$ and ξ_j corresponds to the jth column of an identity matrix. x^j is the sample vector whose the jth variable is reconstructed. The problem to find the number of PCs is to minimize the $\sigma_j(\ell)$ with respect to the number of PCs. Considering all possible faults, the VRE to be minimized is defined as:

$$VRE(\ell) = \underbrace{min}_{\ell} \sum_{j=1}^{m} \frac{\sigma_j(\ell)}{var\{\xi_j^T x\}} = \underbrace{min}_{\ell} \sum_{j=1}^{m} \frac{\sigma_j(\ell)}{\xi_j^T \Sigma \xi_j} \tag{25}$$

4.1 Example 1: simulated data

In this example, we have simulated a simple five-variable model (Lee et al. [2004]). These variables represent five sensors which are structured as follows:

$$X = As \tag{26}$$

where

$$A = \begin{bmatrix} -0.07 & 0.63 & -0.32 & -0.72 & -0.60 \\ -0.45 & -1.05 & -1.08 & -0.72 & 0.01 \\ 0.97 & 0.11 & -0.47 & -0.62 & 0.51 \\ 0.65 & 0.26 & -0.24 & 0.96 & -0.01 \\ -1.02 & 0.19 & -0.17 & 0.13 & -0.03 \end{bmatrix}$$

and the elements of s are uncorrelated random signals with zero means and unity variances. These data are constituted of 200 samples and consist of normal and abnormal parts. The normal part consists of 100 sample vectors. The second half of data represents the abnormal part which is generated by adding a random vector. This last was normally distributed with a zero mean and a standard deviation of 5. The data X was constructed to simulate a precision degradation fault in the fourth variable (Fig. 1). To build a PCA model, the two criterions (CPV and VRE) indicate that almost three PCs can be sufficient and capture about 90% of the variability (Fig. 2). The PCA-based monitoring results for these data are shown in figure 3. Both monitoring indices, T2 and SPE, catch the abnormalities of the data in the PCS and RS respectively. To identify the faulty variable, we can choose freely two outliers. Without loss of generality, we consider the $140th$ and $180th$ outliers that are near the control limits. In figure 4, each variable has a group of four bars (A, B, C, and D) which represent respectively the following contributions: contribution to T2, classical contribution to SPE, new contribution (form I) to SPE, and new contribution (form II) to SPE. We know that the only faulty variable is x_4. Hence, we can notice easily, from figure 4, that our two new proposed approaches are better than the classical contribution to SPE and even the contribution to T2. Our two new methods suspected only the faulty variable: x_4, while the other classical approaches (C_j^{T2} and CC_j^{SPE}) consider that x_1 and x_2 (Fig 4.a) and x_2 (Fig 4.b) are also contaminated.

To calculate the contribution to the SPE, we realized a gain in computing time equal to eight. We found that the algorithm of the second new approach (contribution to SPE based on PCs of the RS) is eight times faster than the algorithm of the first new approach (contribution to SPE based on residues).

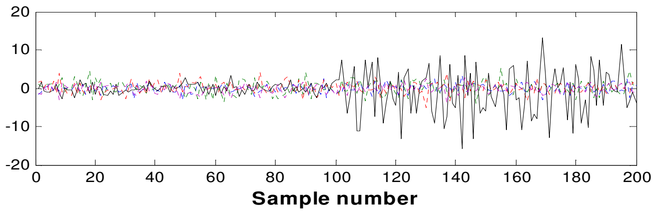

Fig. 1. Test data (solide line, $4th$ variable; dotted lines, the other variables).

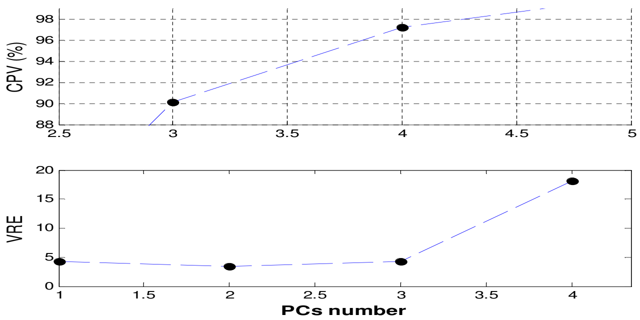

Fig. 2. CPV and VRE indices.

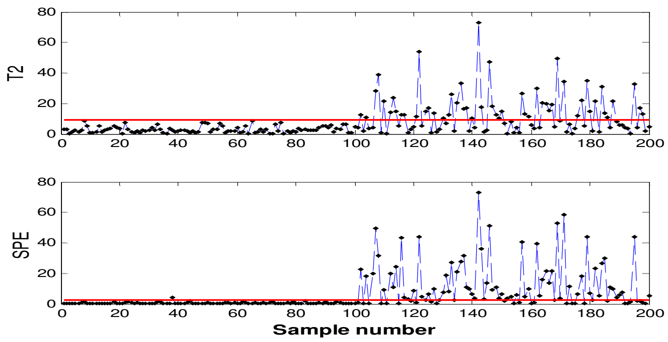

Fig. 3. T2 and SPE indices.

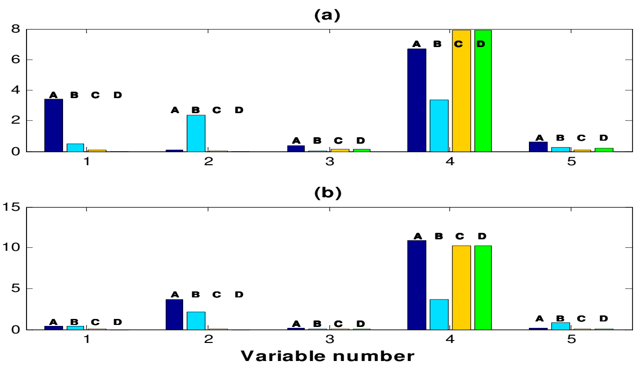

Fig. 4. Contribution plots for outliers: (a) $140th$, (b) $180th$; each variable has a group of four bars representing: (A) contribution to T2, (B) classical contribution to SPE, (C) new contribution to SPE (I), and (D) new contribution to SPE (II) respectively.

4.2 Example 2: real data

This example consists to diagnosing a high dimension data which represent real measurements resulting from a production's chain of wafers of the company STMicroelectronics Rousset. These data are constituted of 70 variables and more than 2700 samples. Using the two selections criterions, we have a PCA model represented by 18 PCs. According to figure 5, we note that the majority of faults are apparently projected into the RS. Because of the high dimension of the data, we cannot treat all outliers in this paper. For this reason, we chose to study an observation that has a large T2 and SPE, as the $2717th$ sample. Our objective is to compare the two new approaches (figures: 8.a and 8.b) to the classical contribution to SPE (Fig. 7). We have verified the variables that are suspected by all approaches. We proved that the classical contribution to SPE suspects many more variables that are not really faulty variables, as x_{20} and x_{59}, compared to those obtained by the contribution to T2 and the two new contributions to SPE. We can note in addition that our two approaches give similar results which are mostly the same as those given by the contribution to T2 (Fig. 6).

We saw that the calculation of the contribution to T2 is based on the more important normalized PCs of the PCS. In our case, we must select some components among 18 PCs which spanned the PCS. We note that when the dimension of data is very large, this choice cannot be always optimal. Consequently, we cannot entrust the fault identification to the calculation of the contribution to T2.

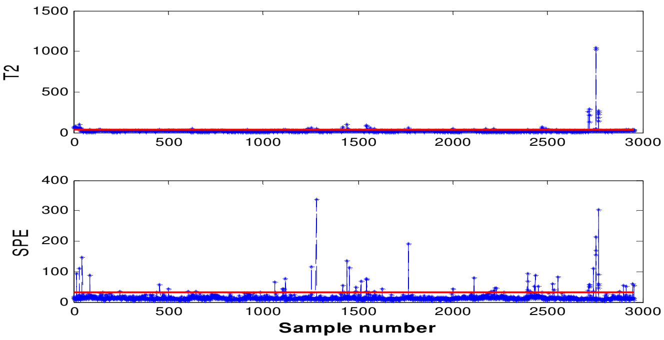

Fig. 5. T2 and SPE indices.

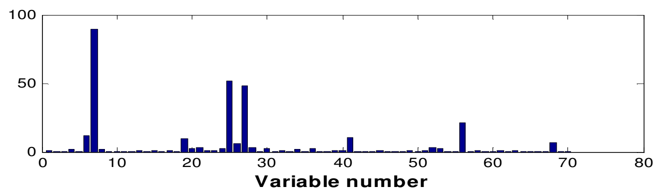

Fig. 6. Contribution to T2.

5. CONCLUSION

In this paper we have discussed basic and advanced issues in statistical process monitoring based on PCA including fault detection and identification. Two fault detection indices (T2 and SPE) are presented and fault identification

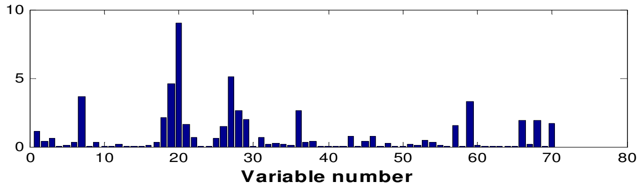

Fig. 7. Classical contribution to SPE.

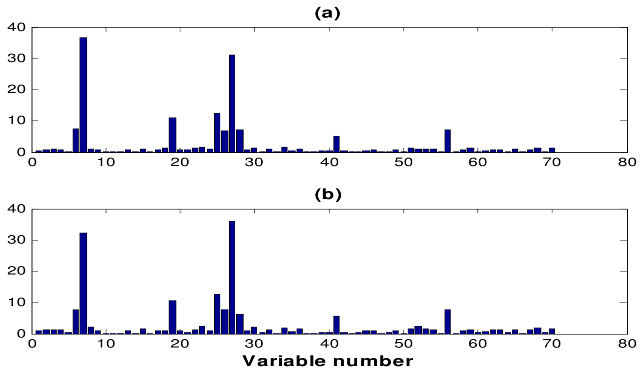

Fig. 8. New contributions to SPE: (a) form I, (b) form II.

via contribution plots is discussed by analysis and examples. We have provided an overview and analysis of the contribution to T2 and the classical contribution to SPE. We have proved that this last is not enabling to identify precisely the faulty variables. For this reason and relying on the fact that the failure is projected into the RS, we have proposed two improved methods to the calculation of the contribution to SPE and we have illustrated their effectiveness from two different examples of process.

The calculation of the contribution to T2 is based on the selection of the more important normalized scores of the PCS. However, when we deal with a high dimension data, the size of the PCS is also relatively large. This makes the selection uninsured which can influence the results of fault identification based on T2. As against, the calculation of the contribution to SPE is more assured because it is calculated on all PCs of the RS. In addition, the process monitoring and fault identification must be based on SPE better than T2 since the failure is projected into the RS. Taking account into these arguments, our two new proposed approaches of fault identification based on the contribution to SPE present the following advantages:

- A higher confidence to fault diagnosis based on the SPE than T2
- A powerful efficiency and accuracy of fault identification compared to those obtained by the classical contribution to SPE
- The second new proposed approach (contribution to SPE based on PCs of the RS) is more economical in computing time.

REFERENCES

V.C.C. de Vargas, L.F.D. Lopes, and A.M. Souza. Comparative study of the performance of the cusum and ewma control charts. *Computers & Industrial Engineering*, 46(4):707–724, 2004.

S. Engelen, M. Hubert, and B.K. Vanden. A comparison of three procedures for robust pca in high dimensions. *Austrian Journal of Statistics*, 34(2):117–126, 2005.

M. Hubert, P.J. Rousseeuw, and K.V. Branden. Robpca: A new approach to robust principal component analysis. *Technometrics*, 47(1):64–79, 2005.

T. Kourti. Application of latent variable methods to process control and multivariate statistical process control in industry. *International Journal of Adaptive Control and Signal Processing*, 19(4):213–246, 2005.

T. Kourti and J.F. MacGregor. Multivariate spc methods for processes and product monitoring. *Journal of Quality Technology*, 28(4):409–428, 1996.

T. Kourti, P. Nomikos, and J.F. MacGregor. Analysis, monitoring and fault diagnosis of batch process using multiblock and multiway pls. *Journal of Process Control*, 5(4):277–284, 1995.

C. Lee, S.W. Choi, J-M. Lee, and I-B. Lee. Sensor fault identification in mspm using reconstructed monitoring statistics. *Ind. Eng. Chem. Res.*, 43(15):4293–4304, 2004.

B. Mnassri, E.M. El Adel, and M. Ouladsine. Fault localization using principal component analysis based on a new contribution to the squared prediction error. In *16th Mediterranean Conference on Control and Automation*, pages 65–70, Ajaccio, France, 2008.

S.J. Qin. Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 17(8–9):480–502, 2003.

S.J. Qin and R. Dunia. Determining the number of principal components for best reconstruction. *Journal of Process Control*, 10(2):245–250, 2000.

S. Valle, W. Li, and S.J. Qin. Selection of the number of principal components: The variance of the reconstruction error criterion with a comparison to other methods. *Ind. Eng. Chem. Res.*, 38(11):4389–4401, 1999.

S. Verboven and M. Hubert. Libra: a matlab library for robust analysis. *Chemometrics and Intelligent Laboratory Systems*, 75(2):127–136, 2005.

J. A. Westerhuis, S.P. Gurden, and A.K. Smilde. Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 51(1):95–114, 2000.

T. Xu and Q. Wang. Application of mspca to sensor fault diagnosis. *ACTA AUTOMATICA SINICA*, 32(3):417–421, 2006.