PCA with Efficient Statistical Testing Method for Process Monitoring

Fangping Mu and Venkat Venkatasubramanian*

Laboratory for Intelligent Process Systems, School of Chemical Engineering
Purdue University, West Lafayette, IN 47907, USA

Abstract

Principal component analysis (PCA) has been used successfully for fault detection and identification in processes with highly correlated variables. The fault detection decision used depends solely on the current sample though the results of previous samples are available and is based on a clear definition of normal operation region, which is difficult to define in reality. In the present work, a novel statistical testing algorithm is integrated with PCA for further improvement of fault detection and identification performance. We use the idea to decompose the scores space and residual space generated by PCA into several subsets so chosen that in each subset the detection problem can be solved with an efficient recursive change detection algorithm based on χ^2-generalized likelihood ratio (GLR) test. *Copyright Ó 2003 IFAC*

Keywords: Principal Component Analysis, Process Monitoring, Sequential Statistical Testing, Contribution Plots

1. INTRODUCTION

Today's chemical processes are becoming heavily instrumented to measure a large number of process variables and data are being recorded more frequently. These process measurements are highly correlated. Identifying and troubleshooting abnormal operating conditions are difficult task with these large amounts of data. The most commonly used technique is principal component analysis (PCA). Process monitoring using PCA is widely based on 'snap shot' Shewhart type control charts, such as T^2- and *SPE*-statistic control charts. The decision depends solely on the current sample though the results of previous sample are available. The implementation of this test is quite simple, but, as one might expect, one pays for this simplicity with rather severe limitations on performance. First, subtle failures are much difficult to detect with this simple scheme. Second, it is difficult to get a tradeoff between false alarm and quick fault detection.

Several extended methods have been proposed for fault detection and identification based on PCA algorithm. Ku et al. (1995) proposed dynamic PCA for process monitoring. Bakshi (1998) combined PCA and wavelet analysis (multiscale PCA) for fault detection. Kano et al. (2001) proposed moving PCA for process monitoring. Kano et al. (2002) described process monitoring based on dissimilarity of process data.

Several problems are not yet solved in these algorithms. First, a clear definition of normal operating condition is needed. This is not the case in reality. In general, there are large gray areas where incipient or small faults occur, while normal process can go to this region by chance. Second, multivariate CUSUM charts, which can detect small changes, are only available for scores. The monitoring of residuals is also very important. Third, though algorithms, which can detect small changes that affect the correlation structure such as Kano's dissimilarity based process monitoring scheme, have been proposed, such schemes cannot detect the variables that are responsible for the fault when the fault occurs.

In this paper, we describe an approach, which integrate PCA with efficient statistical testing algorithm which can solve the problems mentioned above. The outline of the paper is as follows. First, we give a brief introduction of PCA for process monitoring and fault detection. Then, we review

* Corresponding author. Tel: 1-765-494-0734.
E-mail: venkat@ecn.purdue.edu

several statistical testing algorithms. The integration of PCA with an efficient statistical testing algorithm is then presented. Case studies to demonstrate the proposed approach are provided. The paper is concluded with summary.

2. PCA FOR PROCESS MONIOTRING AND FAULT DETECTION

PCA technique is used to develop a model describing the expected variation under normal operating conditions (NOC). An appropriate reference *m*-dimensional data set *X* with *n* samples and *m* variables is chosen which defines the NOC for a particular process. After the data has been properly scaled, PCA can be used to divide the measurement space into two subspaces, one principal component subspace and one residual subspace as,

$$X = TP^T + E = \sum_{i=1}^{A} t_i p_i^T + E$$

where *X* is the normal operating condition data and *E* represents residual error matrix.

The variance of each principal component is determined by the eigenvalues associated with the principal component. T^2 statistic is a Shewhart type chart defined based on principal component subspace as, $T_i^2 = t_i \Lambda^{-1} t_i^T = x_i P \Lambda^{-1} P^T x_i^T$. The matrix Λ is a diagonal matrix containing the eigenvalues associated with the *A* principal components retained in the model. Statistical confidence limits for T^2 can be calculated by means of the F-distribution as follows.

$$T^2_{A,n,a} = \frac{A(n-1)}{n-A} F_a(A, n-A)$$

SPE-statistic is another Shewhart type chart defined on the residual subspace. A general assumption is that the variance is same in all directions, so *SPE*-statistic is defined as,

$$SPE_i = e_i I e_i^T = x_i (I - P_k P_k^T) x_i^T$$

The statistical confidence limits of *SPE*-statistic can be calculated from its approximate distribution. We can also approximate it as,

$$SPE_a = \left(\sum_{i=A+1}^{p} l_i \right) c_a^2 (p - A)$$

When a vector of new data is available, T^2-and *SPE*-statistic can be calculated based on the model generated and are compared with the corresponding confidence limits. If either of the confidence limits is violated, a fault situation is detected.

Contribution plots (Nomikos, 1996) are a PCA approach to fault identification that takes into account the special correlation, thereby improving the univariate statistical techniques. PCA separates the observation space into two subspaces – the reduced space defined by the principal components of the model and the residual subspace. If T^2-statistic or *SPE*-statistic is out of limit, the contribution plots can be used to indicate the variables which are responsible the deviation.

3. STATISTICAL TESTING STRATEGIES

We assume that the measurement of the process follows an independent Gaussian multivariate distribution. For the measurement sequence, $\{X_i\}$, a vector of parameters θ, which is typically the process mean, describes the stochastic behavior of the process. Under the desirable conditions, this vector belongs to the set Θ_0. A control procedure is applied to this process for fault detection and monitoring. If a control procedure triggers a signal under desirable conditions, it is classified as false alarm. At some point in time, the parameters abruptly change to some value that belongs to a rejectable set, Θ_1. The control scheme is then supposed to detect this change as soon as possible.

The criteria of performance of a control scheme are usually related to the behavior of some characteristics of its distribution, most typically the average run length (ARL), which is the average number of observations required for the algorithm to signal that θ has changed. Ideally, the ARL should be large when the process is in control and small when the process is out of control.

An important tool used for fault detection is based on the logarithm of likelihood ratio. It is defined as $s(x) = \ln \frac{p_{q_1}(x)}{p_{q_0}(x)}$. Basseville and Nikiforov (1993) provided a good discussion on log-likelihood ratio strategy for fault detection. Under some general conditions, log-likelihood ratio schemes possess optimality properties in the sense that they provide the best sensitivity for a given rate of false alarms.

3.1 Page's CUSUM algorithm

To improve the sensitivity of the Shewhart charts, Page (1954) modified Wald's theory of sequential hypothesis testing to develop the CUSUM charts that have certain optimality. In this algorithm, post change parameter θ_1 is assumed known and the unknown change point is estimated by maximum likelihood in CUSUM scheme. The CUSUM criterion can be expressed recursively as

$$t = \inf\{n \geq 1 : g_n \geq h\}$$

$$g_n = (g_{n-1} + \log \frac{p_{q_1}(X_n)}{p_{q_0}(X_n)})^+, \; g_0 = 0$$

where $a^+ = a.I\ (a³0)$, $p_q(.)$ is the distribution density function depending on parameter θ.

Moustakides (1986) has shown that Page's CUSUM scheme is optimal in the minmax sense: let *h* be so chosen that $E_0(N) = \gamma$ and let F_γ be the class of all monitoring schemes subject to the constraint $E_0(N) \geq \gamma$, where $E_0(N)$ is the expected ARL when the process is in control. Then the above CUSUM minimize the worst-case expected delay over all rules that belong to F_γ.

3.2 GLR algorithm

The parameter θ_1 after change is generally unknown. An obvious way to modify the CUSUM rule for the case with unknown post change parameter θ is to estimate it by maximum likelihood, leading to the Generalized Likelihood Ratio (GLR) rule.

$$N_G = \inf\{n : \max_{1\le k\le n} \sup_{q\in\Theta} \sum_{i=k}^{n} \log \frac{p_q(X_i)}{p_{q_0}(X_i)} \ge h\}$$

Siegmund and Venkatraman (1995) give asymptotic approximations to the ARL of the GLR under θ_0 and under $\theta \neq \theta_0$, which shows that the GLR rule is asymptotically optimal in the minmax sense. For normal distribution with mean θ and variance 1, they have shown that $\log E_0(N) \sim h$ as $E_0(N) \to \infty$ for the GLR rule. This formula provides an estimation of h given $E_0(N)$.

Unlike CUSUM rule, the GLR rule doesn't have convenient recursive forms and the memory requirements and number of computations at time n grow to infinity with n.

3.3 c^2-GLR algorithm

If we know the post change parameter magnitude but not the direction, we can design an optimal recursive algorithm for the change detection. Here, the process is an independent Gaussian multivariate (r>1) sequence and its mean vector θ changes at an unknown time n.

$$L(X_t) = \begin{cases} N(q_0, \Sigma), & if \quad t < n \\ N(q_1, \Sigma), & if \quad t > n \end{cases}$$

We know the post change magnitude $(q_1 - q_0)^T \Sigma^{-1} (q_1 - q_0) = b^2$. It has been shown that χ^2-GLR can be calculated in recursive form, which greatly reduce the computational burden. The stopping time of GLR algorithm for this situation can be formulated in recursive form as (Nikiforov 2001),

$$\hat{N} = \inf\{n \ge 1 : S_n \ge h\}$$

$$S_n = -n_n \frac{b^2}{2} + b|c_n|, \quad c_n^2 = V_n^T \Sigma^{-1} V_n$$

$$V_n = \mathbf{1}_{\{\hat{S}_{n-1}>0\}} V_{n-1} + (X_n - q_0), \quad n_n = 1_{\{\hat{S}_{n-1}>0\}} n_{n-1} + 1$$

In this algorithm, the magnitude after change is assumed known, which is not true in practice. To deal with this problem, the GLR algorithm can be used. However, this algorithm is computationally expensive. Nikiforov (2001) proposed a suboptimal scheme to solve the computational burden problem. The idea is to decompose a given parameter space into several subsets so chosen that in each subset the detection problem can be solved with loss of a small part, ε, of optimality by a recursive change detection algorithm.

3.4 e-optimality algorithm

This algorithm is designed for detection of changes over a domain $\Theta_1 = \{q_1 : b_0^2 \le (q_1 - q_0)^T \Sigma^{-1} (q_1 - q_0) \le b_1^2\}$ using a collection of L-parallel recursive tests. Each subset is so chosen that the detection problem can be solved by a recursive χ^2-GLR algorithm.

The ε-optimality algorithm is summarized below.

1) Given the tuning parameters e, b_0, b_1, h, calculate the number of parallel tests L, which is the smallest integer $\ge \log \frac{b_1}{b_0} (\log \frac{1+\sqrt{e}}{1-\sqrt{e}})^{-1}$.
2) For l = 1,…, L compute $a_l = b_0 \frac{(1+\sqrt{e})^l}{(1-\sqrt{e})^{l-1}}$ and initialize the L parallel tests.
3) Take the next observations. For l = 1,.., L, compute $S_n(a_l)$.
4) Check if $\max\{S_n(a_1), \dots, S_n(a_L)\} \ge h$ then declare alarm. Otherwise, go to step 3.

4. PCA WITH EFFICIENT STATISTICAL TESTING ALGORITHM FOR PROCESS MONIOTIRNG

In all the above algorithms, an inverse of covariance matrix Σ is needed for the fault detection procedure. However, when lots of process variables are measured and they are correlated, Σ can be singular or near singular. In such case, PCA can be used to divide the measurement space into two subspaces—a score subspace and a residual subspace.

Based on the PCA model, T^2-statistic is designed to detect abnormality in the scores subspace while *SPE*-statistic is for the residual subspace. In the conventional PCA procedure using T^2 and *SPE* for fault detection, the overall type I error is controlled by the level of α. The type II error will be dependent on the post change parameter. Therefore, it is difficult for the procedure to detect small changes whose T^2 and *SPE* statistics is inside the confidence limits. It is also difficult to get a good tradeoff between false alarm and quick detection based on this procedure. It has been shown that ε-optimality GLR algorithm can be used to detect small faults without increasing the false alarm rate. Here we proposed an algorithm to integrate PCA and ε-optimality GLR statistical testing algorithm for fault detection.

First, capability to detect changes of extremely high magnitude can frequently be improved by introducing an additional signal criterion, which calls for a signal at the moment k if testing statistic of a single observation x_k exceeds c, which is a predefined value. Here we choose 99.99% confidence limit for T^2 and *SPE*-statistics as the c value for T^2 and *SPE* statistics, respectively. We define the area between 68% and 99.99% confidence of T^2- and *SPE*-statistic as gray area in the scores and residuals subspace, respectively. Several parallel recursive tests based on ε-optimality algorithm can be designed for the gray area. The following is a summary of the proposed algorithm.

Offline stage

1) Collect normal operating condition (NOC) data X and build PCA model based on NOC data $X = \sum_{i=1}^{A} t_i p_i + E$, where A is the number of principal components used in the model.
2) Based on the PCA model, calculate the 68% and 99.99% confidence limits for T^2-statistic as T_{68} and $T_{99.99}$, and for SPE-statistic as SPE_{68} and $SPE_{99.99}$.
3) Given ε, calculate the number of parallel test for scores as $L_T = ceil(\log\sqrt{\frac{T_{99.99}}{T_{68}}}(\log\frac{1+\sqrt{e}}{1-\sqrt{e}})^{-1})$ and the number of parallel test for residuals as $L_{SPE} = ceil(\log\sqrt{\frac{SPE_{99.99}}{SPE_{68}}}(\log\frac{1+\sqrt{e}}{1-\sqrt{e}})^{-1})$, where *ceil(x)* rounds the elements of X to the nearest integers towards infinity.
4) Calculate the L optimal subdivisions for the test of scores and residuals. For $l = 1,\ldots, L$, compute optimal subdivisions for T^2-statistic as $T_l = \sqrt{T_{68}}\frac{(1+\sqrt{e})^l}{(1-\sqrt{e})^{l-1}}$ and for *SPE*-statistic as $SPE_l = \sqrt{SPE_{68}}\frac{(1+\sqrt{e})^l}{(1-\sqrt{e})^{l-1}}$.
5) Given $E_0(N)$, which is the expected ARL when the process is in control, calculate the threshold for the parallel tests. For parallel tests of scores, $h_T = A\{\log(E_0(N))\}$. For parallel test of residuals, $h_{SPE} = \{\log(E_0(N))\}(\sum_{j=A+1}^{m} l_j)$, where l_j is the jth eigenvalue of covariance or correlation matrix of X.

Online stage

1) When new measurements x_i are available, calculate scores t_i and residuals e_i as $t_i = Px_i$, $e_i = x_i - t_i P$.
2) Calculate the T^2 and *SPE*-statistic for the new data based on scores and residuals as $T_i^2 = t_i^T \Lambda^{-1} t_i$ and $SPE_i = e_i^T e_i$. If $T_i^2 > T_{99.99}$ and/or $SPE_i > SPE_{99.99}$, an alarm is triggered.
3) Otherwise, calculate the testing statistic for each parallel test for scores and residuals.

   For each $l = 1,\ldots, L$, compute $S_i(T_l)$ as

   $i := i+1,\; i_i = \mathbf{1}_{\{S_{i-1}(T_l)>0\}} i_{i-1} + 1$

   $V_{T^2,i} = 1_{\{S_{i-1}(T_l)>0\}} V_{T^2,i-1} + t_i, \quad c_i^2 = V_{T^2,i}^T \Lambda^{-1} V_{T^2,i}$

   $S_i(T_l) = -i_i \frac{T_l^2}{2} + T_l |c_i|$

   For each $l = 1,\ldots, L$, compute $S_i(SPE_l)$ similarly as,

   $i := i+1, \quad i_i = \mathbf{1}_{\{S_{i-1}(SPE_l)>0\}} i_{i-1} + 1$

   $V_{SPE,i} = 1_{\{S_{i-1}(SPE_l)>0\}} V_{SPE,i-1} + e_i, \quad c_i^2 = V_{SPE,i}^T V_{SPE,i}$

   $S_i(SPE_l) = -i_i \frac{SPE_l^2}{2} + SPE_l |c_i|$

   If $\max\{S_i(T_1),\ldots,S_i(T_L)\} \geq h_T$ and/or $\max\{S_i(SPE_1),\ldots,S_i(SPE_L)\} \geq h_{SPE}$, then an alarm is triggered.

If an alarm is triggered, variable contribution can be used to determine the process variable(s) that are responsible for the alarm. PCA divides the variable space into the score subspace and the residual subspace. Therefore, the variable contribution to T^2 should just use the information in the subspace captured by PCs. According to our knowledge, all of the definition of variable contribution to T^2 uses the information in the whole variable space. Here we provide a new definition of variable contribution to T^2 which using only the information in the subspace spanned by PCs. Given that $t = xP$, $\hat{x} = tP^T$ where $\hat{x}$ is the prediction based on PCA model,

$$T^2 = t^T\Lambda^{-1}t = \hat{x}^T P\Lambda^{-1}P^T\hat{x} = \left\|\Lambda^{-1/2}P^T\hat{x}\right\|^2 = \left\|\sum_{k=1}^{m}\Lambda^{-1/2}P_k\hat{x}_k\right\|^2,$$

so we can define the variable contribution to T^2 as

$$T_k^2 = \left\|\Lambda^{-1/2}P_k\hat{x}_k\right\| = \sum_{i=1}^{a} p_{k,i}^2\hat{x}_k \,/\, l_i \,.$$

If the alarm is triggered by T^2-statistic out of 99.99% confidence limits, the new definition of variable contribution to T^2 can be used to determine the variables that are most affected by the fault. If the alarm is triggered by one of the parallel tests in scores space. The following variable contribution definition to cumulative scores can be used.

$$V_{T^2,k}^2 = \sum_{j=1}^{a} p_{k,j}^2 (V_{T^2}P)_k^2 \,/\, l_i$$

If the alarm is triggered by *SPE*-statistic out of 99.99% confidence limits, variable contribution to SPE statistic can be used for fault identification. If the alarm is triggered by one of the parallel tests for residuals, we can define variable contributions based on the cumulative residuals as follows and use them for fault identification.

$$V_{SPE,k}^2 = (V_{SPE})_k^2$$

If $\max\{S_i(T_1),\ldots,S_i(T_L)\} \geq h_T$, and $V_{T^2,k}^2$ is large compare to others, then the k^{th} variable is heavily affected by the fault. Similarly, if $\max\{S_i(SPE_1),\ldots,S_i(SPE_L)\} \geq h_{SPE}$ and $V_{SPE,k}^2$ is large than the others, then the k^{th} variable is heavily affected by the fault.

When the proposed scheme detects a fault, it also provides a rough estimation of the fault magnitude based on the information which test is above the confidence limit.

Note that the proposed scheme is different from *L*-parallel tests of T^2- and *SPE*-statistic. In this scheme the multivariate nature of the process is considered during the design of the algorithm.

5. CASE STUDIES

5.1 AR process

In this section, we will demonstrate the use of the proposed algorithm for process monitoring of a simple multivariate process. The simple process is used to obtain statistically meaningful results. The data for this example are generated from a model suggested by Ku et al. (1995).

$$x(k) = \begin{bmatrix} 0.118 & -0.191 \\ 0.847 & 0.264 \end{bmatrix} x(k-1) + \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} u(k-1),$$

$$y(k) = x(k) + v(k)$$

where u is the correlated input:

$$u(k) = \begin{bmatrix} 0.811 & -0.226 \\ 0.477 & 0.415 \end{bmatrix} u(k-1) +$$

$$\begin{bmatrix} 0.193 & 0.689 \\ -0.320 & -0.749 \end{bmatrix} w(k-1)$$

The input w is a random noise with zero mean and variance 1. The output y is equal to x plus the random noise v(k), with zero mean and variance 0.1. Both input u and output y are measured but v and w are not. Normal operating condition data consists of 200 measurements. 2 principal components are used to build the monitoring model. For the proposed algorithm, ε = 0.05, $E_0(N)$ = 10,000. 4 parallel tests are used for scores space and residuals space, respectively.

Case 1: This case is to monitor the normal process. 1000 normal operating condition data are simulated and used for monitoring based on the conventional PCA model and the proposed algorithm. T^2- and *SPE*-statistic for the conventional PCA model is shown in Figure 1. Though the process is normal, 36 samples are above the warning limit (95%) of T^2-statistic and 4 are above action limit (99%). For the SPE-statistic, 43 samples are above warning limit and 6 are above action limit. The proposed algorithm is used for the normal data. The results are shown in Figure 2. No alarms are generated for those samples.

Case 2: This case is to simulate the mean of w_1 shift from 0.0 to 0.5 introduced at sample 100. T^2- and *SPE*-statistic for conventional PCA model are shown in Figure 3. The conventional PCA cannot detect the fault effectively. The results of the 4 parallel tests for the scores and residuals subspace are shown in Figure 4. The fault is detected at sample 135 by tests in

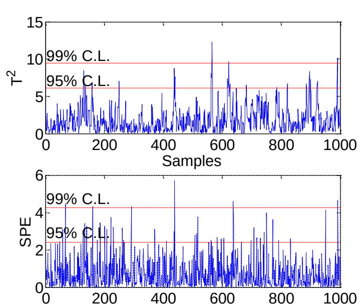

Figure 1. T^2 and SPE-statistic for conventional PCA.

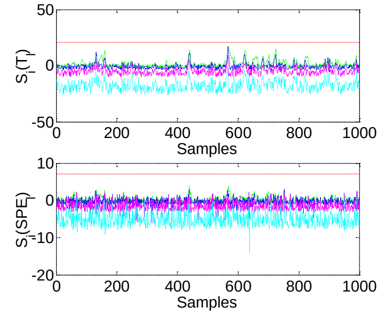

Figure 2. 4 parallel tests for scores and residuals subspace.

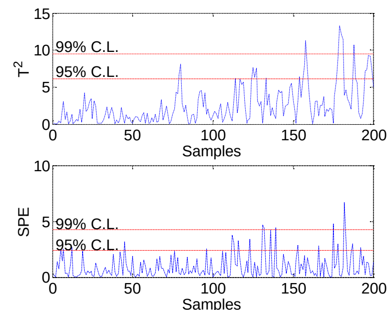

Figure 3. T^2 and SPE-statistic for conventional PCA.

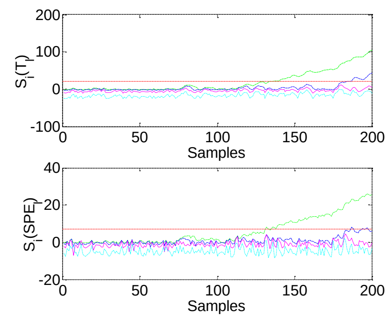

Figure 4. 4 parallel tests for scores and residuals subspace.

scores subspace and at sample 132 by tests in residuals subspace. This scheme can also provide an estimation of the magnitude of the fault.

5.2 Tennessee Eastman Process

The Tennessee Eastman challenge problem is a simulation of a real chemical plant provided by the Eastman Company (Downs and Vogel, 1993). The process has five major units: the reactor, the product condenser, a vapor-liquid separator, a recycle compressor and a product stripper. The control system used for dynamic simulations is the decentralized PID control system designed by McAvoy and Ye (1994). A total of 16 variables, selected by Chen and McAvoy (1998) for monitoring purposes, are used for monitoring in this study. PCA model is built based on 48 hours of steady state simulation data. The sampling interval of the process variable is 3 min. 11 principal components are used to build the model. For the proposed algorithm, ε = 0.05, $E_0(N)$ = 10,000. 2 and 3 parallel tests are used for scores subspace and residuals subspace, respectively.

Case 1: This is the 3rd process disturbance designed in the original paper. It is to simulate a step change in the D feed temperature. The total simulation time is 48 hours and the disturbance is introduced into the system after 36 hours of steady state simulation.

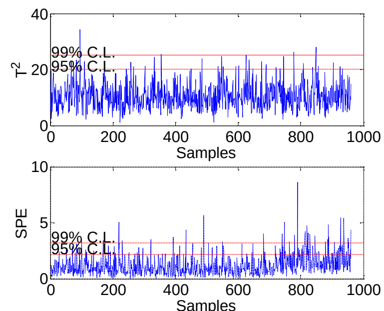

Figure 5. T^2 and SPE-statistic for conventional PCA.

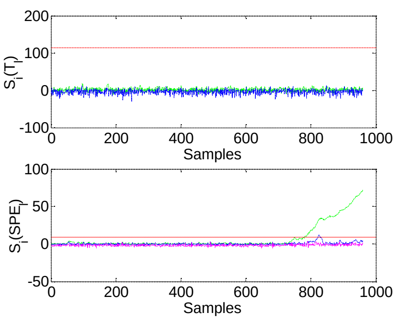

Figure 6. 2 parallel tests for scores subspace and 3 for residuals subspace.

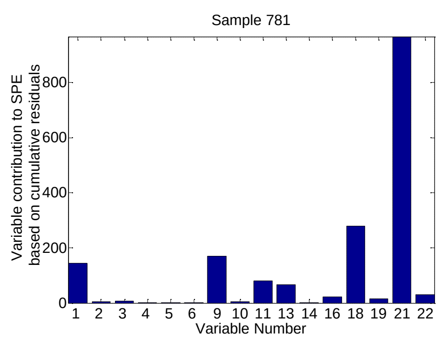

Figure 7. Variable contributions to the cumulative residuals.

T^2 and *SPE*-statistic for conventional PCA are shown in Figure 7. For the 36 hours of steady state simulation, 20 samples are above the warning limit (95%) of T^2-statistic and 2 are above action limit (99%). For the *SPE*-statistic, 54 samples are above warning limit and 10 are above action limit. Conventional PCA cannot detect the fault effectively. Results using the proposed algorithm are shown in Figure 8. There is no false alarm during the 36 hours of steady state simulation. The fault is identified at sample 781 by parallel tests of residuals subspace. Variable contribution to *SPE* based on cumulative residuals is shown in Figure 9. Based on this plots, we can find that variables 21 (Reactor cooling water outlet temperature), 18 (Stripper temperature) and 9 (Reactor temperature) contribute most to the out of control situation.

6. SUMMARY

An approach to integrate PCA with efficient statistical testing algorithm for process monitoring and fault detection has been presented. The fault detection decision depends not only on the current sample but the results of previous sample. A clear definition of normal operating condition is not needed. PCA can separate the observation space into a score subspace and a residual subspace. The two subspaces are divided into several subsets so chosen that in each subset the detection problem can be solved with an efficient recursive change detection algorithm based on χ^2-GLR test. Simulations show that the proposed algorithm can effectively suppress the false alarm and detect small changes in the process.

REFERENCE

Bakshi, B.R. (1998). Multiscale PCA with Application to Multivariate Statistical Process Monitoring. *AIChE J.*, **44** 1596-1610

Basseville, M. and I. Nikiforov (1993). *Detection of Abrupt Changes*, Prentice-Hall, Englewood Cliffs, NJ

Chen, G. and T.J. McAvoy (1998). Predictive on-line monitoring of continuous processes. *J. of Process Control*, **8**, 409-420.

Downs, J.J. and E.F., Vogel (1993). A plant-wide Industrial Process Control Problem. *Comput. Chem. Engng*, **17**, 245-255

Kano, M., S. Hasebe, I. Hashimoto, and H. Ohno (2002). Statistical Process Monitoring Based on Dissimilarity of Process Data. *AIChE J.*, Vol.48, pp.1231-1240.

Kano. M., S. Hasebe, I. Hashimoto and H. Ohno (2001). A new multivariate statistical process monitoring method using principal component analysis. *Comput. Chem. Engng.*, **25** 1103-1113

Ku, W., R.H. Storer and C. Georgakis (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemom. Intell. Lab. Syst.*, **30**, 179-196

McAvoy, T.J. and N. Ye (1994). Base Control for the Tennessee Eastman Problem. *Comput. Chem. Engng*, **18** 383-413

Moustakides, G.V. (1986). Optimal stopping for detecting change in distribution. *Ann. Statist.*, **14**, 1379-1387

Nikiforov, I. (2001). A simple change detection scheme. *Signal Processing*, **81** 149-172

Nomikos, P. (1996). Detection and diagnosis of abnormal batch operations based on multi-way principal component analysis. *ISA Trans.*, Vol.35, pp.259-266.

Page, E.S. (1954). Continuous inspection schemes. *Biometrika*, **41**,100-114

Siegmund, D. and E.S. Venkatraman (1995). Using the generalized likelihood ratio statistics for sequential detection of a change-point. *Ann. Statist.*, **23** 255-271