ADCHEM 2006

International Symposium on Advanced Control of Chemical Processes
Gramado, Brazil – April 2-5, 2006

IFAC

Modified Independent Component Analysis for Multivariate Statistical Process Monitoring

Jong-Min Lee[a], S. Joe Qin[a]* and In-Beum Lee[b]

[a]*Department of Chemical Engineering, The University of Texas at Austin, Austin, TX78712, USA*
[b]*Department of Chemical Engineering, Pohang University of Science and Technology, San 31 Hyoja-Dong, Pohang, 790-784, Korea*

Abstract: In this paper, a modified independent component analysis (ICA) and its application to process monitoring are proposed. The basic idea of this approach is to use the modified ICA to extract some dominant independent components from normal operating process data and to combine them with statistical process monitoring techniques. The proposed monitoring method is applied to fault detection and identification in the Tennessee Eastman process and is compared with the conventional PCA based monitoring method. The monitoring results demonstrate that the proposed method outperforms PCA in terms of the fault detection rate while attaining comparable false alarm rate.

Keywords: Fault Detection; Fault Identification; Statistical Process Control; Independent Component Analysis; Principal Component Analysis

1. INTRODUCTION

In order to extract useful information from a large amount of process data and to detect and diagnose various faults in an abnormal operating situation, a number of multivariate statistical process monitoring (MSPM) approaches based on principal component analysis (PCA) have been developed. PCA is a second-order method, considering only mean and variance of the data. It gives only uncorrelated components, not independent components. PCA performs well in many cases, but gives limited meaningful representations for non-Gaussian data, which can be typical in industrial measurement data (Kermit and Tomic, 2003).
More recently, several MSPM methods based on independent component analysis (ICA) have been proposed (Kano *et al*., 2003, 2004; Lee *et al*., 2003, 2004, Yoo *et al*., 2004; Albazzaz and Wang, 2004). The goal of ICA is to decompose observed data into linear combinations of statistically independent components. In comparison to PCA, ICA involves higher-order statistics, i.e., not only does it decorrelate the data (second order statistics) but also reduces higher order statistical dependencies (Lee, 1998). However, conventional ICA-based monitoring method has some drawbacks for MSPM. First, it is not easy to determine how many independent components (ICs) should be extracted in order to establish a stable ICA model (Kermit and Tomic, 2003). Generally, ICs are extracted up to the dimension of given data, which causes high computational load. Second, the extracted ICs are not ranked in any order as is the case for PCA. In addition, random initialization of the demixing matrix in the whitened space can give different solutions when performing the ICA algorithm.
In this paper, a modified ICA algorithm is proposed to extract dominant ICs from multivariate data. The basic idea is to estimate the variance and the axes of dominant ICs using PCA and then perform ICA to update the dominant ICs while maintaining the variance. This article is organized as follows. The original ICA algorithm is introduced, followed by a modified ICA algorithm and its application to process monitoring. Then, the performance of process monitoring using the modified ICA is illustrated through the Tennessee Eastman process. Finally, a conclusion is given.

* To whom correspondence should be addressed.
E-mail: qin@che.utexas.edu

2. ORIGINAL ICA

The model of ICA is given by

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_m]^T$ is an m-dimensional observation vector, $\mathbf{A} \in R^{m \times p}$ is an unknown mixing matrix and $\mathbf{s} = [s_1, s_2, \ldots, s_p]^T$ is a p-dimensional independent component vector. The objective of ICA is to estimate both $\mathbf{A}$ and $\mathbf{s}$ from only $\mathbf{x}$. This solution is equivalent to finding a demixing matrix $\mathbf{W}$ whose form is such that the elements of the reconstructed vector $\hat{\mathbf{s}}$, given as

$$\hat{\mathbf{s}} = \mathbf{W}\mathbf{x} \tag{2}$$

become as independent of each other as possible.

In the original ICA algorithm, it is assumed that m equals p and all ICs have unit variance for convenience. The initial step in ICA is to remove all the cross-correlation of $\mathbf{x}$, given as

$$\mathbf{z} = \mathbf{\Lambda}^{-1/2}\mathbf{U}^T\mathbf{x} = \mathbf{Q}\mathbf{x} \tag{3}$$

where $E\{\mathbf{z}\mathbf{z}^T\} = \mathbf{I}$, $\mathbf{Q} = \mathbf{\Lambda}^{-1/2}\mathbf{U}^T$, and $\mathbf{U}$ and $\mathbf{\Lambda}$ are eigenvector and eigenvalue matrix, respectively, generated from the eigen-decomposition of $E\{\mathbf{x}\mathbf{x}^T\} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$. Then, Eq. (3) can be expressed as

$$\mathbf{z} = \mathbf{Q}\mathbf{x} = \mathbf{Q}\mathbf{A}\mathbf{s} = \mathbf{B}\mathbf{s} \tag{4}$$

where $\mathbf{B} = \mathbf{Q}\mathbf{A}$ is an orthogonal matrix since

$$\mathbf{I} = E\{\mathbf{z}\mathbf{z}^T\} = \mathbf{B}E\{\mathbf{s}\mathbf{s}^T\}\mathbf{B}^T = \mathbf{B}\mathbf{B}^T. \tag{5}$$

Thus, $\mathbf{s}$ can be estimated from Eq. (4)

$$\hat{\mathbf{s}} = \mathbf{B}^T\mathbf{z} = \mathbf{B}^T\mathbf{Q}\mathbf{x}. \tag{6}$$

From Eq. (2) and Eq. (6),

$$\mathbf{W} = \mathbf{B}^T\mathbf{Q}. \tag{7}$$

To calculate $\mathbf{B}$, each column vector $\mathbf{b}_i$ is randomly initialized and then updated so that the i-th independent component $\hat{s}_i = (\mathbf{b}_i)^T\mathbf{z}$ may have maximized non-Gaussianity. As a measure of non-Gaussianity, negentropy, the difference of the differential entropy between the given data and Gaussian distribution data, has been used. Hyvärinen and Oja (2000) introduced a reliable approximation of negentropy:

$$J(y) \approx [E\{G(y)\} - E\{G(v)\}]^2 \tag{8}$$

where y is assumed to be of zero mean and unit variance, v is a Gaussian variable of zero mean and unit variance, and G is any non-quadratic function. Hyvärinen and Oja (2000) suggested three functions for G:

$$G_1(u) = \frac{1}{a_1}\log\cosh(a_1 u) \tag{9}$$

$$G_2(u) = \exp(-a_2 u^2/2) \tag{10}$$

$$G_3(u) = u^4 \tag{11}$$

where $1 \le a_1 \le 2$ and $a_2 \approx 1$. G_2 and G_3 are more suitable for super-Gaussian and sub-Gaussian components, respectively. G_1 is a good general-purpose contrast function and is therefore selected for use in this paper.

Hyvärinen (1999) introduced a highly efficient fixed-point algorithm for ICA based on the approximate form for the negentropy. The algorithm, called FastICA, calculates each column of the matrix $\mathbf{B}$ one by one and allows the identification of each independent component. More details on the FastICA algorithm are well described in Hyvärinen and Oja (2000), Hyvärinen (1999), Hyvärinen *et al*. (2001).

3. MODIFIED ICA

ICA not only decorrelates the data (second order statistics) but also reduces higher order statistical dependencies; hence it can extract underlying hidden factors efficiently and capture the essential structure of the data. Based on this merit, some researchers have illustrated that applying ICA to process monitoring is useful to detect and identify various faults generated from abnormal situations (Kano *et al*., 2003; Lee *et al*., 2004).

However, the conventional ICA-based monitoring method has some drawbacks. A fundamental assumption behind original ICA is that the number of ICs equals that of variables of given data. In case that the number of measured variables is very large, it has high computational load and may extract additional ICs which are unimportant for detecting faults. Of course, one can reduce data dimension in advance using PCA before performing ICA (Hyvärinen *et al*., 2001). However, much information needed to extract essential ICs is ignored by data reduction with PCA. The second problem in the original ICA algorithm is that ICs are not ordered in the same fashion as with PCA since the variance of extracted ICs is assumed to be all one (Kermit and Tomic, 2003). There is no standard criterion to order ICs. Furthermore, random initialization of demixing matrix $\mathbf{B}$ in the whitened space can lead to different solutions when performing ICA algorithm (Kermit and Tomic, 2003).

In order to solve these problems, a modified ICA algorithm is suggested in this paper. The modified ICA algorithm can extract a few dominant ICs, determine the order of ICs, and give a consistent solution. The basic idea is to first use PCA to estimate initial ICs where the variance of each IC is the same as that of each PC and then to update a few dominant ICs using FastICA algorithm. Here, it is reasonable to expect the space spanned by the major ICs to be essentially similar to the ones associated to the largest principal components (PCs) because ICA can be viewed as a modified PCA (centering and whitening) and an additional iterative process (Kocsor *et al*., 2004).

The objective of the modified ICA can be defined as follows: to find a demixing matrix $\mathbf{W} \in R^{p \times m}$ whose form is such that the elements of the extracted vector $\mathbf{y}$, given as

$$\mathbf{y} = \mathbf{W}\mathbf{x} \tag{12}$$

become as independent of each other as possible and have been ordered by their variances that are the same as the variances of the corresponding PCs.

To solve above problem, first of all, all score components are extracted from PCA

$$\mathbf{t} = \mathbf{U}^T\mathbf{x} \tag{13}$$

where $\mathbf{t}$ is the score vector with $E\{\mathbf{tt}^T\} = \mathbf{\Lambda} = diag\{\lambda_1, \ldots, \lambda_m\} \in R^{m\times m}$ and $\mathbf{U} \in R^{m\times m}$ is the loading matrix obtained from $E\{\mathbf{xx}^T\} = \mathbf{U\Lambda U}^T$, respectively. In some cases, the last a few eigenvalues in $\mathbf{\Lambda}$ are so small that they are close to zero. In that case, the eigenvalues and the corresponding eigenvectors can be excluded. However, it is important to retain as many eigenvalues as possible because the extracted score components give additional information to find essential ICs even though their variances are small.

Eq. (13) can be changed as follows through the whitening transform:

$$\mathbf{z} = \mathbf{Qx} \tag{14}$$

where $\mathbf{z}$ is the normalized score vector, $\mathbf{z} = \mathbf{\Lambda}^{-1/2}\mathbf{t}$, and $\mathbf{Q} = \mathbf{\Lambda}^{-1/2}\mathbf{U}^T$.

From $\mathbf{z} \in R^m$, a few dominant ICs, $\mathbf{y} \in R^p$ satisfying $E\{\mathbf{yy}^T\} = \mathbf{D} = diag\{\lambda_1, \ldots, \lambda_p\}$, should be found such that the elements of $\mathbf{y}$ are as independent of each other as possible, using

$$\mathbf{y} = \mathbf{C}^T\mathbf{z} \tag{15}$$

where $\mathbf{C} \in R^{m\times p}$, $\mathbf{C}^T\mathbf{C} = \mathbf{D}$. $E\{\mathbf{yy}^T\} = \mathbf{D}$ reflects that the variance of each element of $\mathbf{y}$ is the same as that of scores in PCA, hence ICs can be ordered according to their variances.

Eq. (15) can be arranged as a simpler model by multiplying $\mathbf{D}^{-1/2}$ to each side:

$$\mathbf{y}_n = \mathbf{C}_n{}^T\mathbf{z} \tag{16}$$

where $\mathbf{y}_n = \mathbf{D}^{-1/2}\mathbf{y}$, $\mathbf{D}^{-1/2}\mathbf{C}^T = \mathbf{C}_n{}^T$, $\mathbf{C}_n{}^T\mathbf{C}_n = \mathbf{I}$, and $E\{\mathbf{y}_n\mathbf{y}_n{}^T\} = \mathbf{I}$. Consequently, the problem of finding an arbitrary demixing matrix $\mathbf{W}$ is reduced to the simpler problem of finding a matrix $\mathbf{C}_n$ which has fewer parameters to estimate as a result of the orthogonality. Note that $\mathbf{z}$ is the normalized score vector generated from PCA, that is uncorrelated and had been ordered by its original variance. The first p components of $\mathbf{z}$ can be a good initial value of $\mathbf{y}_n$ since statistical dependencies of data have been removed up to the second order (mean and variance) by PCA. To do this, the initial matrix of $\mathbf{C}_n{}^T$ should be set to be

$$\mathbf{C}_n{}^T = \left[\mathbf{I}_p \,\vdots\, \mathbf{0}\right] \tag{17}$$

where $\mathbf{I}_p$ is the p-dimensional identity matrix and $\mathbf{0}$ is p by m-p zero matrix. This initialization is based on the assumption that extracted PCs are good initial estimates of ICs, and thereby can give a consistent solution unlike random initialization.

The detail procedures to find a few dominant ICs are:

1) Determine p, the number of ICs to estimate. Set counter $i \leftarrow 1$.
2) Denote $\mathbf{c}_{n,i}$ as the i-th column vector of $\mathbf{C}_n$ and take the initial vector $\mathbf{c}_{n,i}$ to be i-th column vector of $\mathbf{I}_p$ in Eq. (17)
3) Let $\mathbf{c}_{n,i} \leftarrow E\{\mathbf{z}g(\mathbf{c}_{n,i}{}^T\mathbf{z})\} - E\{g'(\mathbf{c}_{n,i}{}^T\mathbf{z})\}\mathbf{c}_{n,i}$, where g is the first derivative and g' is the second derivative of *G*, where *G* takes the form of Eq. (9), (10) or (11). This step is an approximate Newton iteration procedure for the maximization of the negentropy given in Eq. (8).
4) Do the orthogonalization:
$$\mathbf{c}_{n,i} \leftarrow \mathbf{c}_{n,i} - \sum_{j=1}^{i-1}(\mathbf{c}_{n,i}{}^T\mathbf{c}_{n,j})\mathbf{c}_{n,j}.$$
This step removes the information contained in the solutions already found.
5) Normalize $\mathbf{c}_{n,i} \leftarrow \mathbf{c}_{n,i} / \left\|\mathbf{c}_{n,i}\right\|$
6) If $\mathbf{c}_{n,i}$ has not converged, go back to Step 3).
7) If $\mathbf{c}_{n,i}$ has converged, output the vector $\mathbf{c}_{n,i}$. Then, if $i \le p$ set $i \leftarrow i+1$ and go back to Step 2).

Once $\mathbf{C}_n$ is found, then final demxing matrix $\mathbf{W}$ and mixing matrix $\mathbf{A}$ can be obtained from

$$\mathbf{W} = \mathbf{D}^{1/2}\mathbf{C}_n{}^T\mathbf{Q} \tag{18}$$

$$\mathbf{A} = \mathbf{U\Lambda}^{1/2}\mathbf{C}_n\mathbf{D}^{-1/2} \tag{19}$$

At last, we can obtain some dominant ICs from Eq. (12). The extracted ICs reveal the majority of information and represent a meaningful representation about the observed data $\mathbf{x}$.

4. MODIFIED ICA FOR MONITORING

In the proposed monitoring method, two types of statistics are considered: the *D*-statistic to monitor the systematic part change of the process variation and the *Q*-statistic to monitor the residual part of the process variation. The *D*-statistic, also known as the Hotelling's T^2 statistic, is the Mahalanobis distance defined as follows:

$$T^2 = \mathbf{y}^T\mathbf{D}^{-1}\mathbf{y} \tag{20}$$

where $\mathbf{y}$ is obtained from Eq. (12) and $\mathbf{D}$ is the diagonal matrix of the eigenvalues associated with the retained dominant ICs. In this paper, kernel density estimation is used to define the control limit for T^2 because $\mathbf{y}$ is not Gaussian (Silverman, 1986; Martin and Morris, 1996; Lee *et al.*, 2004).

The *Q*-statistic, also known as the *SPE* statistic is defined as follows:

$$SPE = \mathbf{e}^T\mathbf{e} = (\mathbf{x} - \hat{\mathbf{x}})^T(\mathbf{x} - \hat{\mathbf{x}}) \tag{21}$$

where $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ and $\hat{\mathbf{x}}$ can be calculated as follows:

$$\hat{\mathbf{x}} = \mathbf{Ay} = \mathbf{AWx}. \tag{22}$$

If the number of ICs is chosen such that the majority of non-Gaussianity is included in the ICs, the residual subspace will contain mostly random noise which can be treated as normal distribution. The upper control limit of *SPE* can then be calculated from Jackson and Mudholkar (1979).

The contribution based approach is simple to identify faults and can be generated without prior knowledge (Qin, 2003). In the proposed method, the T^2 statistic can be decomposed as:

$$T^2 = \mathbf{y}^T\mathbf{S}^{-1}\mathbf{y} = \mathbf{y}^T\mathbf{S}^{-1}\mathbf{W}\mathbf{x} = \mathbf{y}^T\mathbf{S}^{-1}\sum_{j=1}^{m}\mathbf{w}_j x_j = \sum_{j=1}^{m}\mathbf{y}^T\mathbf{S}^{-1}\mathbf{w}_j x_j = \sum_{j=1}^{m} c_j \quad (23)$$

Therefore, the contribution to the T^2 statistic for a data $\mathbf{x}$, is given as follows (Westerhuis *et al*., 2000):

$$c_j\left(T^2\right) = \mathbf{y}^T\mathbf{S}^{-1}\mathbf{w}_j x_j \quad (24)$$

where $c_j\left(T^2\right)$ is the contribution of the *j*-th variable to the T^2 statistic, x_j is the *j*-th element of $\mathbf{x}$, and $\mathbf{w}_j$ is the *j*-th row of the demixing matrix $\mathbf{W}$.

Similarly, the contribution of process variable *j* at given time to the *SPE* statistic is defined as follows:

$$c_j\left(SPE\right) = e_j^{\ 2} \quad (25)$$

where e_j is the *j*-th variable of $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$.

In this paper, the upper control limits for $c_j\left(T^2\right)$ are calculated as the mean of the contributions plus three standard deviations of the contributions for each process variable (Westerhuis *et al*., 2000). Control limits for $c_j\left(SPE\right)$ are calculated the same way as the Q-statistic control limit (Westerhuis *et al*., 2000).

5. CASE STUDY

In this section, the proposed method is applied to the Tennessee Eastman process simulation data and is compared with PCA monitoring results. The details on the process description are well explained in Chiang *et al*. (2001). A total of 33 variables listed in Table 1 are used for monitoring in this study. A sampling interval of 3 minutes was used to collect the simulated data. Both the training and testing data sets for each fault are composed of 960 observations. A set of programmed faults (Fault 1-21) is listed in Table 2. All faults in the test data set were introduced from sample 160. The data can be downloaded from http://brahms.scs.uiuc.edu (Chiang *et al*., 2001).

All the data were auto-scaled prior to the application of PCA and the modified ICA. In the modified ICA, 30 whitened vectors are extracted from Eq. (13) to update and find ICs. 9 PCs are selected for the PCA by cross-validation and the same number of ICs is chosen for fair comparison.

The false alarm rates and the fault detection rates of the two multivariate methods, PCA and modified ICA, for all 21 fault data were computed and tabulated in Table 3. For the data obtained after the fault occurrence, the percentage of the samples outside the 99% control limits was calculated in each simulation and termed as detection rate. Maximum detection rate achieved for each fault is marked with a bold number. With 9 PCs and 9 ICs, false alarm rates of PCA and modified ICA are comparable though they are different for each fault data. As shown in Table 3, the modified ICA can detect most faults more effectively than PCA except Fault 4 and 11. For Faults 10 and 16, the detection rate of the proposed method is more than twice as high as that of PCA, which shows that the modified ICA with acceptable false alarm rate can detect small events that are difficult to detect by PCA. One thing that needs to be noted is T^2 ability of the proposed method for detecting faults. For all cases, the detectability of T^2 is considerably enhanced by the proposed method. It means the proposed method can extract essential features in a process much more sensitively than PCA. This result demonstrates that the proposed method is expected to be more effective than PCA to diagnose fault patterns in the feature space.

The monitoring charts of PCA and modified ICA in the case of Fault 10 are shown in Fig. 1. PCA can detect the fault from about sample 200, however, there are lots of samples below the 99% control limit despite the presence of the fault. On the other hand, the modified ICA detects the fault earlier than PCA by 11 samples and gives a consistent fault alarm up to the end of the processing time. Also, the random pattern changes caused by the fault are reflected well in the proposed method. The results of this example indicate the proposed method has a superior capability in detecting faults that are difficult to detect by the conventional method. Fig. 2 shows contribution plots to T^2 and *SPE* at sample 195, respectively, in the case of Fault 10. From this figure, variables 16 (Stripper pressure) and 18 (Stripper temperature) make the largest contribution to the T^2 statistic while variables 19 (Stripper steam flow) and 31 (Stripper steam valve) give dominant effects on *SPE* statistic. This contribution plot correctly indicates the major variable groups affected by the fault. Thus, the fault detection and identification ability of the proposed method is much worthy of consideration.

Table 1 Variables in the Tennessee Eastman process

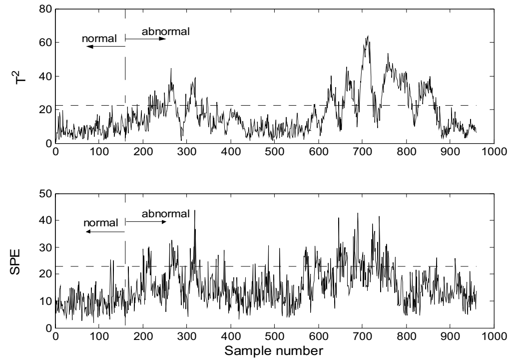
1	*A* feed	18	stripper temperature
2	*D* feed	19	stripper steam Flow
3	*E* feed	20	compressor work
4	total feed	21	reactor cooling water outlet temperature
5	recycle flow	22	separator cooling water outlet temperature
6	reactor feed rate	23	*D* feed flow valve
7	reactor pressure	24	*E* feed flow valve
8	reactor level	25	*A* feed flow valve
9	reactor temperature	26	total feed flow valve
10	purge rate	27	compressor recycle valve
11	product separator temperature	28	purge valve
12	product separator level	29	separator pot liquid flow valve
13	product separator pressure	30	stripper liquid product flow valve
14	product separator underflow	31	stripper steam valve
15	stripper level	32	reactor cooling water flow
16	stripper pressure	33	condenser cooling water flow
17	stripper underflow		

Table 2 List of Process faults for the Tennessee Eastman process

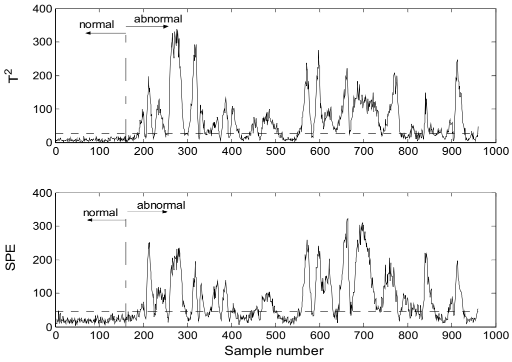
No.	Description	Type
1	*A*/*C* feed ratio, *B* composition constant	Step
2	*B* composition, *A*/*C* ratio constant)	Step
3	*D* feed temperature	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	*A* feed loss	Step
7	C header pressure loss - reduced availability	Step
8	*A*, *B*, *C* feed composition	Random
9	*D* feed temperature	Random
10	*C* feed temperature	Random
11	Reactor cooling water inlet temperature	Random
12	Condenser cooling water inlet temperature	Random
13	Reaction kinetics	Slow drift
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16 ~ 20	Unknown	
21	The valve for Stream 4 was fixed at the steady state position	Constant Position

Table 3 Representative detection rates of PCA and modified ICA

Faults	False alarm rate				Detection rate			
	PCA		Modified ICA		PCA		Modified ICA	
	T^2	*SPE*	T^2	*SPE*	T^2	*SPE*	T^2	*SPE*
1	0	1.25	0	1.88	99	**100**	**100**	**100**
2	0.63	0	0	0.63	**98**	96	**98**	**98**
3	0.63	0.63	0	0	2	1	1	1
4	0.63	0	0	1.88	6	**100**	65	96
5	0.63	0	0	1.88	**24**	18	**24**	**24**
6	0.63	0	0	0	99	**100**	**100**	**100**
7	0	0.63	0	0.63	42	**100**	**100**	**100**
8	0	1.88	0	0	97	89	97	**98**
9	1.25	0.63	0.63	3.75	1	1	1	2
10	0.63	1.25	0	0.63	31	17	**70**	64
11	0.63	0	0	0	21	**72**	43	66
12	0.63	1.25	0	0	97	90	**98**	97
13	0	0.63	0	0	93	**95**	**95**	94
14	0	1.88	0	0.63	81	**100**	**100**	**100**
15	0.63	1.25	0.63	0	1	2	1	2
16	3.13	0.63	1.25	1.25	14	16	**76**	73
17	0	1.88	0	1.25	74	93	87	**94**
18	0	1.88	0.63	1.88	89	**90**	**90**	**90**
19	0	0	0	0	0	**29**	25	**29**
20	0	1.88	0	0	32	45	**70**	66
21	0	0	1.25	0.63	33	46	**54**	19



(a) PCA



(b) Modified ICA

Fig. 1. Monitoring charts of a) PCA and b) Modified ICA for Fault 10.

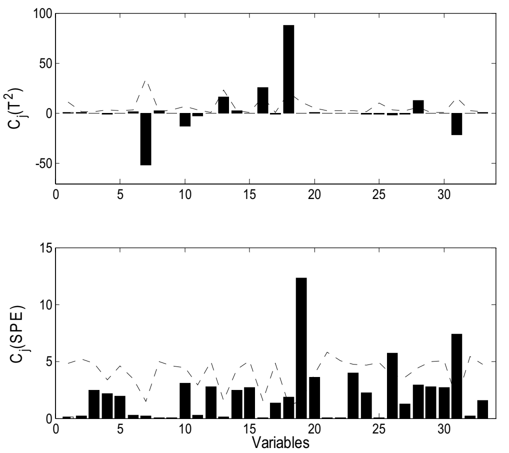


Fig. 2. Variables contribution plots to T^2 and *SPE* at sample 195 for Fault 10

6. CONCLUSION

This paper proposes a novel approach to process monitoring that uses modified ICA. Some problems of original ICA are analyzed and a modified ICA algorithm is developed and applied to MSPM. Compared to original ICA, the proposed algorithm has the following advantages: (1) It extracts a few dominant factors needed for process monitoring; (2) High computational load is attenuated by extracting a few dominant ICs, not all ICs; (3) The ordering of ICs is considered; (4) It gives a consistent solution.

The proposed method was applied to the fault detection and identification of Tennessee Eastman process. The fault detection performance was evaluated and compared with that of conventional PCA-based monitoring. This example demonstrates that the proposed method can detect various faults more efficiently than PCA. In particular, the extracted dominant ICs are expected to be more useful to diagnose fault patterns in the feature space. In addition, contribution plots of the proposed method can reveal the group of process variables responsible for the process to go out of control.

ACKNOWLEDGEMENT

This work was supported by the Korea Research Foundation Grant funded by Korea Government (MOEHRD, Basic Research Promotion Fund) (KRF-2005-214-D00027) and the Texas-Wisconsin Modeling and Control Consortium.

REFERENCES

Albazzaz, H., & Wang, X. Z. (2004). Statistical Process Control Charts for Batch Operations Based on Independent Component Analysis", *Industrial and Engineering Chemistry Research*, **43**(21), 6731-6741.

Chiang, L. H., Russell, E. L., & Braatz, R. D. (2001). *Fault detection and diagnosis in industrial systems*, Springer, London

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, **10**, 626 -634.

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, **13**(4-5), 411-430.

Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent Component Analysis*. John Wiley & Sons, Inc., New York.

Jackson, J. E., & Mudholkar, G. S. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, **21**(3), 341-349.

Kocsor, A., & Tóth, L. (2004) Kernel-Based Feature Extraction with a Speech Technology Application. *IEEE Transactions on Signal Processing*, **52**(8), 2250-2263

Kano, M., Tanaka, S., Hasebe, S., Hashimoto, I., & Ohno, H. (2003). Monitoring independent components for fault detection. *A.I.Ch.E. Journal*, **49**(4), 969-976.

Kano, M., Hasebe, S., Hashimoto, I., & Ohno, H. (2004). Evolution of multivariate statistical process control: independent component analysis and external analysis. *Computers & Chemical Engineering*, **28**(6-7), 1157-1166.

Kermit, M., & Tomic, O. (2003). Independent Component Analysis Applied on Gas Sensor Array Measurement Data. *IEEE Sensors Journal*, **3**(2), 218-228.

Lee, T. (1998). *Independent Component Analysis: Theory and Applications*. Kluwer Academic Publishers, Boston.

Lee, J.-M., Yoo, C. K., & Lee, I.-B. (2003). New monitoring technique with ICA algorithm in wastewater treatment process. *Water Science and Technology*, **47**(12), 49-56.

Lee, J.-M., Yoo, C. K., & Lee, I.-B. (2004). Statistical process monitoring with independent component analysis. *Journal of Process Control*, **14**, 467-485.

Martin, E. B., & Morris, A. J. (1996). Non-parametric confidence bounds for process performance monitoring charts. *Journal of Process Control*, **6**(6), 349-358.

Qin, S. J. (2003). Statistical Process Monitoring: Basics and Beyond. *Journal of Chemometrics*, **17**, 480-502 .

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, UK.

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

Yoo, C.K., Lee, J.-M., Vanrolleghem, P. A., & Lee, I.-B. (2004). On-line Monitoring of Batch Processes Using Multiway Independent Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, **71**, 151-163.