Chinese J. Chem. Eng., **14**(5) 559—568 (2006)

Investigation of Dynamic Multivariate Chemical Process Monitoring*

XIE Lei(谢磊), ZHANG Jianming(张建明) and WANG Shuqing(王树青)**
National Key Laboratory of Industrial Control Technology, Institute of Advanced Process Control, Zhejiang University, Hangzhou 310027, China

Abstract Chemical process variables are always driven by random noise and disturbances. The closed-loop control yields process measurements that are auto and cross correlated. The influence of auto and cross correlations on statistical process control (SPC) is investigated in detail by Monte Carlo experiments. It is revealed that in the sense of average performance, the false alarms rates (FAR) of principal component analysis (PCA), dynamic PCA are not affected by the time-series structures of process variables. Nevertheless, non-independent identical distribution will cause the actual FAR to deviate from its theoretic value apparently and result in unexpected consecutive false alarms for normal operating process. Dynamic PCA and ARMA-PCA are demonstrated to be inefficient to remove the influences of auto and cross correlations. Subspace identification-based PCA (SI-PCA) is proposed to improve the monitoring of dynamic processes. Through state space modeling, SI-PCA can remove the auto and cross correlations efficiently and avoid consecutive false alarms. Synthetic Monte Carlo experiments and the application in Tennessee Eastman challenge process illustrate the advantages of the proposed approach.
Keywords multivariate statistical processes control, subspace identification, false alarms rate, dynamic processes

1 INTRODUCTION

With the advent of improved instrumentation and automation, chemical processes now produce large volumes of information which are highly correlated. Several multivariate statistical methods have been developed to identify the correlations between variables and create a reduced set of variables in the orthogonal axes that capture most of the variability in the collected information. One of most popular MSPC methods is principal component analysis (PCA), which has also been applied to chemical processes[1—6].

However, PCA is based on the assumption that the process variables are independent and identically normally distributed (IID), that is, stationary or uncorrelated in time[5—8]. In practice, this assumption is always violated, as chemical process variables are driven by random noise and disturbances. Due to the feedback control, the impact of disturbances propagates to both the input and output variables. Thus the variables move around the steady state and exhibit some degrees of auto and cross correlation.

In order to monitor the process dynamics, PCA has been extended to include the time-series structures of variables[8—11]. Among these extensions, dynamic PCA (DPCA) by Ku *et al*.[8] is widely adopted and can be treated as a multivariate AR-like time series modeling approach. Although applications in Tennessee Eastman[8] and some batch processes monitoring[12] have demonstrated the efficiencies of dynamic PCA, it is proved recently that dynamic PCA cannot eliminate the auto and cross correlations of variables[13]. If a dynamic PCA is used, the score variables will be auto and cross correlated even when the process variables are neither auto nor cross correlated. In other words, dynamic PCA will always induce the dynamics into the score variables. In order to overcome this question, Kruger *et al*. involved ARMA filters to remove the auto-correlations from the PCA scores. But unfortunately, as demonstrated in this article, the cross-correlations still remain in the filtered score variables and independent assumptions are still not satisfied.

The contributions of this article are as follows. First, the influences of auto and cross correlations of process variables on dynamic PCA are investigated through synthetic Monte Carlo experiments. It is revealed that the presence of auto and cross correlations will cause the false alarms rate (type Ⅰ error rate) to deviate from its theoretic value, but will not always result in higher false alarms rate as mentioned in literatures. More precisely, the average FAR is not affected by the time-series structures. Second, criterions to determine whether variables are auto or cross correlated are introduced. The ARMA filtering approach suggested by Kruger *et al*.[13] is shown to be inefficient to reduce the cross correlations of PCA scores. Third, a subspace identification modeling approach combined with PCA is proposed to remove the entire dynamics from the score variables. A novel information based criteria is presented to determine the order of relative state-space model. Fourth, the effectiveness of proposed approach is demonstrated using the

Received 2005-08-23, accepted 2006-01-13.
* Supported by the National Natural Foundation of China (No.60421002, No.70471052).
** To whom correspondence should be addressed. E-mail: jmzhang@iipc.zju.edu.cn

Tennessee Eastman process. Finally, some remarks and conclusions are presented.

2 PERFORMANCE INVESTIGATION OF DYNAMIC PCA

This section provides a brief overview of dynamic PCA and studies the false alarm rate performance of dynamic PCA.

2.1 Dynamic principal component analysis

Given the measurements, normal operating process data are collected and put in a two-dimensional data matrix $\boldsymbol{X} \in \mathbb{R}^{N\times n}$ with N samples and n variables. PCA decomposes the data matrix $\boldsymbol{X}$ in terms of r linear principal components with $r \leqslant n$:

$$\begin{aligned}\boldsymbol{X} &= \boldsymbol{XPP}^{\mathrm{T}} + \boldsymbol{X}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{P}}^{\mathrm{T}} \\ &= \boldsymbol{TP}^{\mathrm{T}} + \boldsymbol{E}\end{aligned} \tag{1}$$

where $\boldsymbol{P} \in \mathbb{R}^{n\times r}$ and $\boldsymbol{T} = \boldsymbol{XP} \in \mathbb{R}^{N\times r}$ are defined as the principal component loadings and scores, respectively. $\tilde{\boldsymbol{P}}$ contains the retained principal component directions. $\boldsymbol{E}$ is the residual matrix. A more detailed analysis of PCA can be found in Ref.[5].

When PCA is applied to monitor a process, two complementary statistics, that is, Hotelling T^2 and Q are commonly used:

$$\begin{aligned}T_k^2 &= \boldsymbol{x}_k^{\mathrm{T}} \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}_k = \boldsymbol{t}_k^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{t}_k \sim \frac{r(N^2-1)}{N(N-r)}F_{\alpha,r,N-r} \\ Q_k &= \boldsymbol{x}_k^{\mathrm{T}}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{x}_k\end{aligned} \tag{2}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_r]$ and λ_j are the eigenvalues of the correlation matrix of $\boldsymbol{X}$ and the upper α confidence control limit for Q is approximated as

$$CL = \theta_1\left(\frac{c_\alpha\sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2}\right)^{1/h_0} \tag{3}$$

where $\theta_i = \sum_{j=r+1}^{n}\lambda_j^d$ with $i = 1, 2, 3$, $h_0 = 1 - (2\theta_1\theta_3 / 3\theta_2^2)$ and c_α is the standard normal α confidence control limit.

Dynamic PCA arranges the process variables to form an autoregressive (AR) structure:

$$\underline{\boldsymbol{X}} = [\boldsymbol{X}_0\ \boldsymbol{X}_{-1} \cdots \boldsymbol{X}_{-d}] \in \mathbb{R}^{(N-d)\times(d+1)n} \tag{4}$$

where $\underline{\boldsymbol{X}}$ is an augmented set of variables, representing an AR model structure of order d and the subscripts 0, 1, d refer to the backshifts applied. PCA is applied to $\underline{\boldsymbol{X}}$ and the corresponding T^2 and Q statistics can be obtained.

It is demonstrated by Kruger *et al*.[13] that the retained scores variables are autocorrelated irrespective of whether the process variables are autocorrelated or not.

2.2 Performance investigation of dynamic PCA

Type Ⅰ error rate or false alarms rate refers to the percentage of statistics violating its confidence bound when monitoring normal operating process. For PCA, the ideal type Ⅰ error rate for T^2 statistic is equal to the significant level. In practice, false alarms rate is one of the most important parameters to determine, too high false alarms rate will result in poor acceptance among shop-floor personnel while too low rate will make the monitoring system insensitive to potential process or sensor faults.

It is argued by Kruger *et al*.[13] that the application of dynamic PCA to process monitoring will result in higher T^2 false alarms due to the auto and cross correlations in process variables. In this section, through the Monte Carlo experiments, it is revealed that, in the sense of average performance, the T^2 type Ⅰ error rate of dynamic PCA coincides with the theoretic value α.

Example 1

For the sake of comparison, the same process described in Ref.[8] is studied:

$$\begin{aligned}\boldsymbol{z}_k &= \boldsymbol{A}\boldsymbol{z}_{k-1} + \boldsymbol{B}\boldsymbol{v}_{k-1} \\ \boldsymbol{y}_k &= \boldsymbol{z}_k + \boldsymbol{f}_k\end{aligned} \tag{5}$$

where $\boldsymbol{A} = \begin{bmatrix} 0.118 & -0.191 \\ 0.847 & 0.264 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$, $\boldsymbol{z}_k$ and $\boldsymbol{f}_k$ are the process outputs/states and measurement noise, respectively. The input variables $\boldsymbol{v}_k$ are defined as follows:

$$\begin{aligned}\boldsymbol{v}_k &= \boldsymbol{C}\boldsymbol{v}_{k-1} + \boldsymbol{D}\boldsymbol{w}_{k-1} \\ \boldsymbol{u}_k &= \boldsymbol{v}_k + \boldsymbol{g}_k\end{aligned} \tag{6}$$

where $\boldsymbol{C} = \begin{bmatrix} 0.811 & -0.226 \\ 0.477 & 0.415 \end{bmatrix}$, $\boldsymbol{D} = \begin{bmatrix} 0.193 & 0.689 \\ -0.320 & -0.749 \end{bmatrix}$ and $\boldsymbol{w}_k$ is a random noise with zero mean and unit/identity variance. The measurement noise $\boldsymbol{g}_k$ is added on the input variables $\boldsymbol{v}_k$ to form measured input $\boldsymbol{u}_k$. The initial states follow a normal distribution, $\boldsymbol{v}_0 \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_v)$, $\boldsymbol{z}_0 \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_z)$ where $\boldsymbol{\Sigma}_v = \begin{bmatrix} 1.7236 & -0.0376 \\ -0.0376 & 1.2572 \end{bmatrix}$ and $\boldsymbol{\Sigma}_z = \begin{bmatrix} 5.0148 & -4.4155 \\ -4.4155 & 38.6601 \end{bmatrix}$ satisfy the following discrete Lyapunov equation[14]:

$$\begin{bmatrix} \boldsymbol{\Sigma}_v & \boldsymbol{\Sigma}_{vz} \\ \boldsymbol{\Sigma}_{vz}^{\mathrm{T}} & \boldsymbol{\Sigma}_z \end{bmatrix} = \begin{bmatrix} \boldsymbol{C} & \boldsymbol{0} \\ \boldsymbol{B} & \boldsymbol{A} \end{bmatrix}\begin{bmatrix} \boldsymbol{\Sigma}_v & \boldsymbol{\Sigma}_{vz} \\ \boldsymbol{\Sigma}_{vz}^{\mathrm{T}} & \boldsymbol{\Sigma}_z \end{bmatrix}\begin{bmatrix} \boldsymbol{C}^{\mathrm{T}} & \boldsymbol{B}^{\mathrm{T}} \\ \boldsymbol{0} & \boldsymbol{A}^{\mathrm{T}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{DD}^{\mathrm{T}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \tag{7}$$

The measurement noises are chosen as normally distributed variables.

The simulations are performed on MATLAB® 6.5 and repeated 400 times. For each simulation, a total number of 2000 samples of $\boldsymbol{y}$ and $\boldsymbol{u}$ are generated. The first 1000 samples $\boldsymbol{Y}=[\boldsymbol{y}_1\ \boldsymbol{y}_2 \cdots \boldsymbol{y}_{1000}]^{\mathrm{T}}$ and $\boldsymbol{U}=[\boldsymbol{u}_1\ \boldsymbol{u}_2 \cdots \boldsymbol{u}_{1000}]^{\mathrm{T}}$ are selected as the reference data and the retained 1000 samples serve as testing data to evaluate the false alarms rate (FAR) of dynamic PCA. The augmented data set $\underline{\boldsymbol{X}}=[\boldsymbol{Y}_0\ \boldsymbol{U}_0\ \boldsymbol{Y}_{-1}\ \boldsymbol{U}_{-1}]\in\mathbb{R}^{999\times 8}$ are constructed to establish dynamic PCA model. Similar to Ku *et al.*[8], five principal components are chosen.

A typical monitoring result of DPCA for the testing normal process data is given in Fig.1. The influence of process dynamics is characterized by the consecutive false alarms between the 200th and 300th sample in T^2 control chart. The simulation FAR results on testing data are provided in Table 1.

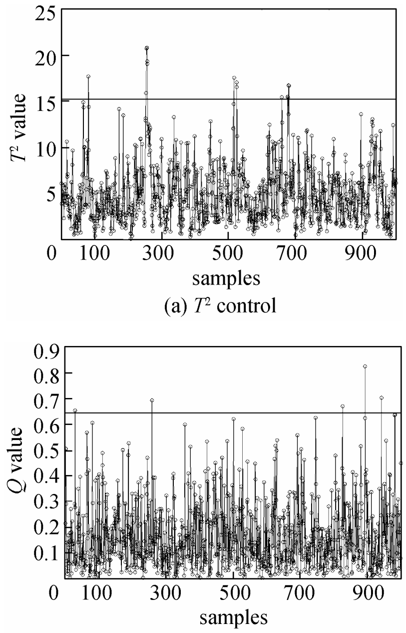

Figure 1 T^2 control and Q control charts for testing data using dynamic PCA (Solid line represents the 99% control limit)

In Table 1, α is the corresponding significance level to determine the control limit. Mean and Std represent the mean and the standard deviation of FAR, respectively. Max and Min denote the maximum and minimum type Ⅰ error rate in the total 400 simulations.

Examining the T^2 FAR values, it can be seen that the mean of FAR agrees with the theoretic value. The presence of auto and cross correlations makes the FAR depart from its mean value apparently. In the worst case, the actual FAR is almost 7.5% less than α (when α=0.15). As α increases, the influence of variables dynamics becomes more and more serious. Comparison given in the following section will revel that dynamic PCA does not outperform conventional PCA method and further improvement of dynamic PCA can be obtained through combining PCA and subspace identification algorithm.

The Q FAR values are also listed in Table 1. The averages of Q FAR are biased because Eq.(3) is an approximation of the real distribution. Nevertheless, compared with T^2 statistic, the Q still performs better considering the value of Std. As defined in Eq.(3), Q is the sum of square of last 8－5=3 components. In the following section, it is shown that the last three components are less correlated than the first five ones. Therefore Q is less affected by the process dynamics.

3 SUBSPACE IDENTIFICATION BASED PCA

This section shows that the use of an ARMA filter for each score variable may not be effective in removing cross-correlations between the score variables. In order to remove the auto and cross correlations between score variables simultaneously, a subspace identification approach is presented. Before that, a discussion on correlation issues is given first in subsection 3.1.

3.1 Auto and cross correlation coefficients and ARMA-based PCA

Consider a set of random time series $\boldsymbol{u}=[u_1, u_2, \cdots, u_r]$, the auto and cross covariance coefficient is defined as

$$\rho_{i,j}(\tau)=\frac{\mathrm{Cov}[u_i(t),u_j(t-\tau)]}{\sqrt{\mathrm{Cov}[u_i(t),u_i(t)]\mathrm{Cov}[u_j(t),u_j(t)]}},\quad 1\leqslant i,j\leqslant r \tag{8}$$

$$\mathrm{Cov}[u_i(t),u_j(t-\tau)]=\begin{cases}\dfrac{1}{N-\tau}\displaystyle\sum_{t=\tau+1}^{N}[u_i(t)-\overline{u}_i][u_j(t-\tau)-\overline{u}_j], & \tau>0\\ \mathrm{Cov}\{u_j(t),u_i[t-(-\tau)]\}, & \tau<0\end{cases}$$

Table 1 FAR of 400 Monte Carlo experiments

α	T^2 FAR				Q FAR			
	Mean	Std	Min	Max	Mean	Std	Min	Max
0.01	0.0103	0.0055	0	0.0340	0.0107	0.0040	0.0020	0.0240
0.02	0.0203	0.0082	0.0020	0.0520	0.0215	0.0060	0.0060	0.0430
0.05	0.0503	0.0142	0.0180	0.1110	0.0539	0.0098	0.0310	0.0890
0.10	0.0998	0.0208	0.0480	0.1660	0.1061	0.0143	0.0690	0.1550
0.15	0.1499	0.0260	0.0750	0.2230	0.1576	0.0173	0.1180	0.2140

$$\overline{u}_i = \frac{1}{N}\sum_{t=1}^{N} u_i(t) \quad (9)$$

where $\mathrm{Cov}[u_i(t), u_j(t-\tau)]$ is the cross covariance between u_i and u_j, N is the number of samples. For multi-normally distributed random variables (independent), $\rho_{i,j}(\tau)=0$ when $\tau \neq 0$.

The auto and cross correlation coefficients can be defined in a similar manner except that the covariance is replaced by correlation. In the following sections, it is assumed that the process is stationary and the variables are zero-centred, so the correlation coefficient equals the covariance coefficient and both of them are denoted as ACFs.

The following two theorems are introduced to determine whether variables are auto or cross correlated.

Theorem 1

Suppose $u_i(t)$ is a Gauss white noise sequence (so independent) and $N \gg M$. Then the auto-correlation coefficient $\rho_{i,i}(\tau)$ is approximately normally distributed, that is, $\sqrt{N}\rho_{i,i}(\tau) \sim N(0,1),\ 1 \leqslant |\tau| \leqslant M$.

Theorem 2

Suppose $u_i(t)$ is a Gauss white noise sequence and independent of $u_j(t)$ and $N \gg M$. Then the cross-correlation coefficient $\rho_{i,j}(\tau)$ is approximately normally distributed, that is, $\sqrt{N}\rho_{i,j}(\tau) \sim \sqrt{\sum_{m=-M}^{M} \rho_{i,i}(m)\cdot\rho_{j,j}(m)} \cdot N(0,1), 1 \leqslant |\tau| \leqslant M$.

The detailed proof of Theorem 1 and 2 refers to Ljung[15]. The condition of Theorem 2 can be relaxed as $u_i(t)$ being the linear combination of time-shifted Gauss white noises. Some other useful criterions to test the auto and cross correlations between random time series are also given in the same reference. Theorem 1 and 2 are adopted because they are easier to implement[16].

In Fig.2, the ACFs of the principal components by dynamic PCA application in Example 1 are given. The shadow area represents 99% confidence intervals which are computed through Theorem 1 and 2. It is evident that the first five principal components (scores, bottom left corner of Fig.2) chosen by dynamic PCA to calculate T^2 are both auto and cross correlated significantly. In contrast, the left three components (top right corner of Fig.2) relating to Q statistics are less correlated as mentioned in section 2.2.

In order to overcome the deficiencies of dynamic PCA, Kruger *et al.*[13] incorporate ARMA filters in the PCA analysis. The ARMA based PCA method applies traditional PCA method to original data matrix $\boldsymbol{X}$ first and r ARMA filters are identified to remove the auto-correlations of each score variable. Although the auto-correlations are efficiently eliminated, the cross-correlations still exist and the filtered scores are not independent yet. In Fig.3, the ACFs of filtered four scores in Example 1 are plotted. The order of ARMA filters are chosen according to Kruger and listed in Table 2. From Fig.3, it can be seen that the filtered Score_1 is still cross-correlated with the other three scores and does not satisfy the independent condition.

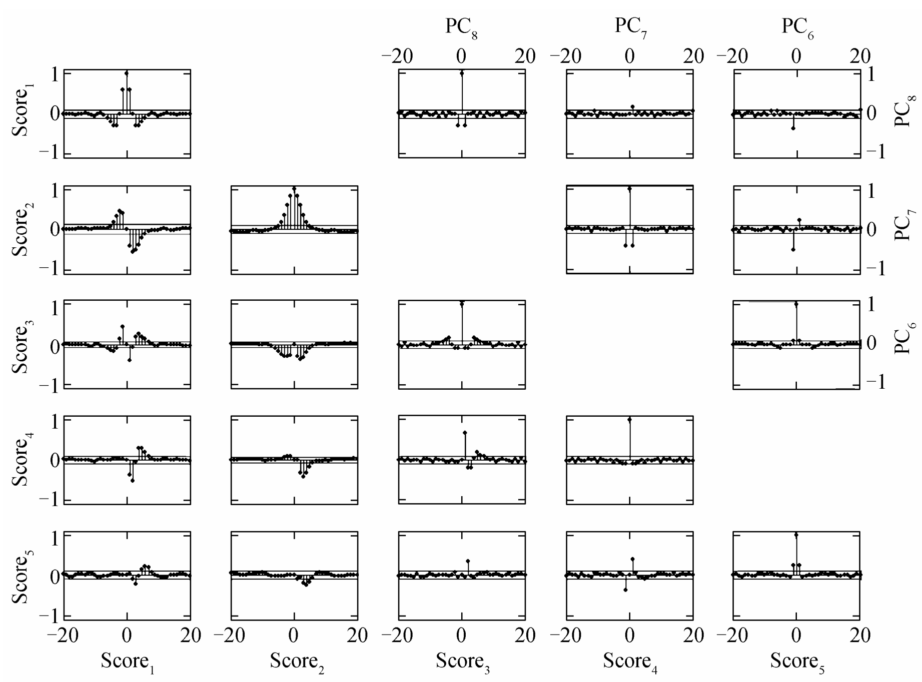

Figure 2 ACFs of scores by dynamic PCA (time lag=1) **in Example 1**

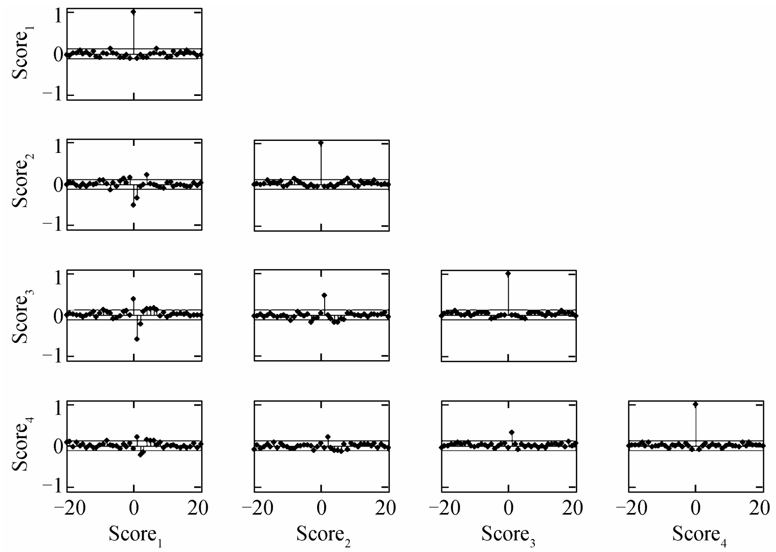

Figure 3 ACFs of filtered scores by ARMA-PCA in Example 1

Table 2 ARMA orders for score variables by ARMA-PCA in Example 1

PC number	AR order	MA order
1	7	0
2	6	1
3	8	3
4	3	1

3.2 Subspace identification-based PCA (SI-PCA)

As mentioned above, Kruger's ARMA-PCA approach ignores the cross-correlations of score variables. This will lead to ARMA models with higher orders than necessary. In Example 1, the underlying model order is 4 while the AR orders in Table 2 are all larger than 4 except for the last score variable.

To describe the auto and cross correlations simultaneously, one should consider that the ARMA approach can be extended to multivariate case and a multivariate time-series model should be established. However, the multivariate ARMA model is much more difficult to analyze because all the coefficients in univariate model will become matrices and the model complexity will grow rapidly as the model order increases[17].

An alternative approach to analyze multivariate time series is state-space modelling. As argued by Ljung[16] that "*Generally speaking, it is preferable to work with state-space models in the multivariate case, since the model structure complexity is easier to deal with*" and any linear model structure (ARX, ARMA *etc*.) can be represented by state-space model.

A state-space model for time series is given by

$$\boldsymbol{z}_{k+1} = \boldsymbol{A}\boldsymbol{z}_k + \boldsymbol{K}\boldsymbol{e}_k$$
$$\boldsymbol{t}_k = \boldsymbol{C}\boldsymbol{z}_k + \boldsymbol{e}_k \qquad (10)$$

where $\boldsymbol{t}_k \in \mathbb{R}^{r\times1}$ is the corresponding time series, $\boldsymbol{z}_k \in \mathbb{R}^{nz\times1}$ is the vector of state variable, $\boldsymbol{A}$, $\boldsymbol{C}$ are the system matrices, $\boldsymbol{K}$ is the Kalman gain. Note that there are no inputs in this model. The residuals of state-space model $\boldsymbol{e}_k \in \mathbb{R}^{r\times1}$, which are assumed IID, are employed to process monitoring instead of correlated scores.

Subspace identification (SI) algorithms have been widely adopted to identify the state-space model from input-output data because it does not need iterative, nonlinear optimization as maximum likelihood method and SI is very easy to implement[18].

For subspace algorithms, it is crucial to estimate the state $\boldsymbol{z}_k$ which is defined as a linear combination of past outputs

$$\boldsymbol{p}_k = [\boldsymbol{t}_{k-1}^{\mathrm{T}}\ \boldsymbol{t}_{k-2}^{\mathrm{T}} \cdots \boldsymbol{t}_{k-d}^{\mathrm{T}}]^{\mathrm{T}}$$
$$\boldsymbol{z}_k = \boldsymbol{J}\boldsymbol{p}_k \qquad (11)$$

where d is the number of lags as mentioned in dynamic PCA. Once $\boldsymbol{J}$ is determined, the z_k can be calculated by Eq.(11) and the state-space matrices can be estimated by linear squares regression. Different approach to calculate $\boldsymbol{J}$ distinguishes various derivations of subspace algorithms including CVA, N4SID and PLS *etc*. For example, CVA calculates $\boldsymbol{J}$ from the canonical loading between the conditional future and the past outputs. More details about subspace identification algorithms refer to Ref.[14].

To determine the order of system in Eq.(10), a number of approaches have been proposed. For instance, N4SID determines the system order by checking the singular values. Akaike information criterion (AIC) is also employed to determine the model order automatically. In this context, however, the purpose of modeling is to remove the auto and cross correlations from the score variables as much as possible, that is, reduce the ACFs $\rho_{i,j}(\tau)$ of e_k as close to zero as possible when $\tau \neq 0$. To this end, the following Akaike-like information criterion is suggested

$$AICx = \lg[V(nz)] + \frac{2p}{N} \tag{12}$$

$$V(nz) = \frac{1}{2Mr^2} \sum_{i=1}^{r} \sum_{j=1}^{r} \sum_{\substack{\tau=-M \\ \tau \neq 0}}^{M} \rho_{i,j}^2(\tau) \tag{13}$$

where p is the number of estimated parameters in Eq.(10), N is the length of modeling data and M=3×nz is the maximum time lag. $\frac{2p}{N}$ is included to avoid over-fitting. The system order nz is then selected to minimize the $AICx$ objective function.

Once the state-space model is identified and the residual e_k is generated, we have the following T^2 statistic

$$T_k^2 = \boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{S}_e^{-1} \boldsymbol{e}_k \sim \frac{r(N^2-1)}{N(N-r)} F_{\alpha,r,N-r} \tag{14}$$

where $\boldsymbol{S}_e = \frac{1}{N-1}\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}}$ is the correlation matrix of $\boldsymbol{e}_k$.

Note that not all scores are necessarily included in the state-model. If a score is independent on itself and other scores, it should be excluded to reduce the complexity of state-space model. In this situation, the T^2 statistic will become

$$T_k^2 = \begin{bmatrix} \hat{\boldsymbol{e}}_k \\ \tilde{\boldsymbol{t}}_k \end{bmatrix}^{\mathrm{T}} \boldsymbol{S}_{et}^{-1} \begin{bmatrix} \hat{\boldsymbol{e}}_k \\ \tilde{\boldsymbol{t}}_k \end{bmatrix} \sim \frac{r(N^2-1)}{N(N-r)} F_{\alpha,r,N-r} \tag{15}$$

where $\boldsymbol{S}_{et} = \frac{1}{N-1}\sum_{i=1}^{N} \begin{bmatrix} \hat{\boldsymbol{e}}_i \\ \tilde{\boldsymbol{t}}_i \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{e}}_i \\ \tilde{\boldsymbol{t}}_i \end{bmatrix}^{\mathrm{T}}$ is the correlation matrix of $\hat{\boldsymbol{e}}_k$ and $\tilde{\boldsymbol{t}}_k$. $\hat{\boldsymbol{e}}_k$ is the residual of auto and cross correlated scores $\hat{\boldsymbol{t}}_k$ and $\tilde{\boldsymbol{t}}_k$ is the independent part.

3.3 Dynamic process monitoring framework based on SI-PCA

The procedures of offline and online monitoring using SI-PCA are as follows.

Offline: develop the normal operating condition model (NOC)

(1) Collect the operating data set during normal operation in $\boldsymbol{X}$.

(2) Apply PCA to $\boldsymbol{X}$ and obtain the score variables $\boldsymbol{t}_k = \begin{bmatrix} \hat{\boldsymbol{t}}_k \\ \tilde{\boldsymbol{t}}_k \end{bmatrix}$. The number of components can be determined by cross validation or other criteria. The independence of excluded principal components relative to Q statistic should also be checked. If dynamics exist, state-space model can also be employed.

(3) Subspace identification method in section 3.2 is employed to remove the dynamics of $\hat{\boldsymbol{t}}_k$. The confidence interval of T^2 statistic is determined based on the residual $\hat{\boldsymbol{e}}_k$, $\tilde{\boldsymbol{t}}_k$ from Eq.(15) and Q statistics from Eq.(2).

Online monitoring

(1) For new observation, obtain the score values *via* $\boldsymbol{t}_k = \boldsymbol{P}^{\mathrm{T}} \boldsymbol{x}_k$.

(2) Apply the identified subspace model to calculate the residual $\hat{\boldsymbol{e}}_k$.

(3) Determine the T^2 and Q statistics and compare with the confidence intervals.

4 APPLICATION

4.1 Application in Example 1

To investigate the performance of SI-PCA, the new approach is also applied to monitoring the process described in Example 1.

Since all the principal components are auto and cross correlated as depicted in Fig.4, all four scores are chosen and there is no Q statistic for Example 1. Half of the training data (500 samples) are used to estimate the values of parameters and the other half serve as validating data to determine the model order and the residual covariance matrix by applying Eq.(12) and Eq.(15).

The identified state-space matrices are

$$\boldsymbol{A} = \begin{bmatrix} 0.8712 & 0.3598 & 0.0145 & -0.1322 \\ -0.4414 & 0.5138 & -0.1168 & -0.3705 \\ -0.0387 & 0.4187 & 0.3546 & 0.2463 \\ 0.0117 & -0.0158 & -0.2218 & 0.3195 \end{bmatrix}$$

$$\boldsymbol{C} = \begin{bmatrix} -6.6158 & 21.6087 & -3.6322 & -9.4871 \\ -28.4087 & -0.8583 & 2.1548 & 1.8765 \\ -0.7248 & -5.0136 & 0.4425 & -3.2668 \\ -0.1470 & 0.0674 & -4.0356 & -1.2017 \end{bmatrix}$$

$$\boldsymbol{K} = \begin{bmatrix} 0.0100 & -0.0134 & 0.0006 & -0.0051 \\ 0.0221 & -0.0021 & 0.0035 & -0.0024 \\ -0.0076 & 0.0067 & -0.0071 & 0.0006 \\ 0.0029 & -0.0096 & -0.0331 & 0.0353 \end{bmatrix}$$

In contrast to ARMA-PCA, a four-order state-space model is enough to remove the auto and cross correlations. The ACFs of residuals are plotted in Fig.5. Fig.6 shows a typical monitoring result of SI-PCA of normal operating process. There are no longer consecutive false alarms in T^2 control chart as dynamic PCA does.

Four hundred Monte Carlo experiments are also performed and the T^2 FAR results for normal operating data of PCA, dynamic PCA, ARMA-PCA and

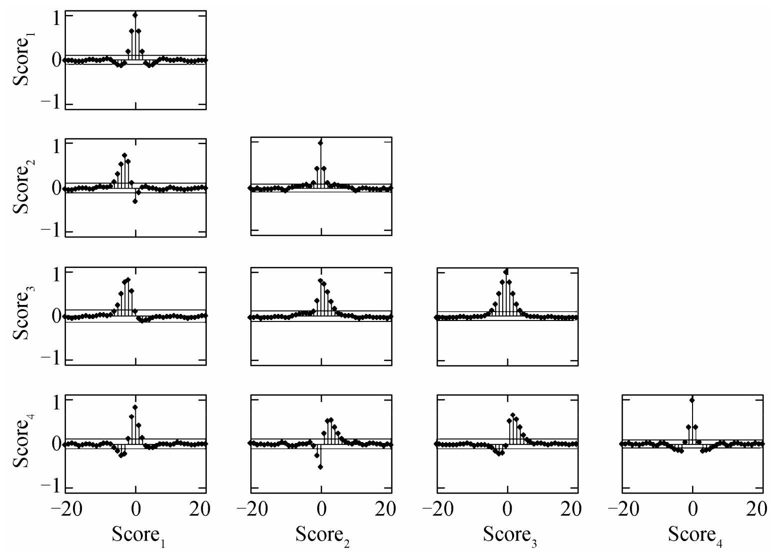

Figure 4 ACFs of PCA score variables in Example 1

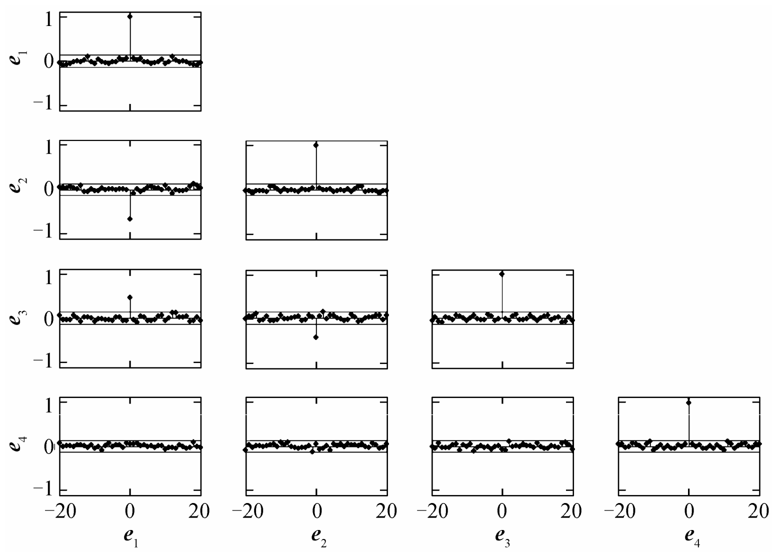

Figure 5 ACFs of residuals by SI-PCA in Example 1

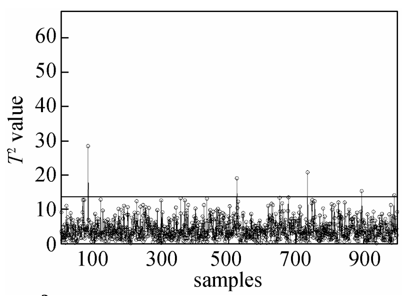

Figure 6 T^2 **control chart for testing data in Example 1 using SI-PCA** (Solid line is the 99% control limit)

SI-PCA are listed in Table 3. The numbers of principal components for each method are 4, 5, 4, and 4, respectively. Hence there are no Q statistics for PCA, ARMA-PCA and SI-PCA approaches.

It is interesting to note that all methods' average FAR coincide with the theoretic value and the conclusion might be misleading if we compare the performance of two approaches only by a single experiment. With respect to the standard deviation of FAR, dynamic PCA does not outperform PCA because the scores of DPCA are still significantly auto and cross correlated as depicted in Fig.2. The ARMA-PCA has a better performance compared with the PCA and the dynamic PCA due to the fact that auto correlations are eliminated in the filtered scores. The proposed SI-PCA approach outperforms all the other methods at any significance level α. In general, the actual FAR rate of SI-PCA is closer to theoretic value.

In the following section, a realistic application of SI-PCA to detect the process abnormal behavior is presented.

4.2 Tennessee Eastman process

Since the introduction in 1993 by Downs and Vogel[19], Tennessee Eastman (TE) process has been widely studied in the literatures[8,20—23]. The TE model

Table 3 T^2 FAR of 400 Monte Carlo experiments

α	Mean (FAR)				Std (FAR)			
	PCA	Dynamic PCA	ARMA-PCA	SI-PCA	PCA	Dynamic PCA	ARMA-PCA	SI-PCA
0.01	0.0098	0.0103	0.0105	0.0100	0.0050	0.0055	0.0042	0.0037
0.02	0.0196	0.0203	0.0205	0.0199	0.0074	0.0082	0.0063	0.0054
0.05	0.0485	0.0503	0.0511	0.0499	0.0124	0.0142	0.0116	0.0096
0.10	0.0973	0.0998	0.1015	0.0997	0.0186	0.0208	0.0172	0.0135
0.15	0.1470	0.1499	0.1522	0.1498	0.0235	0.0260	0.0212	0.0165

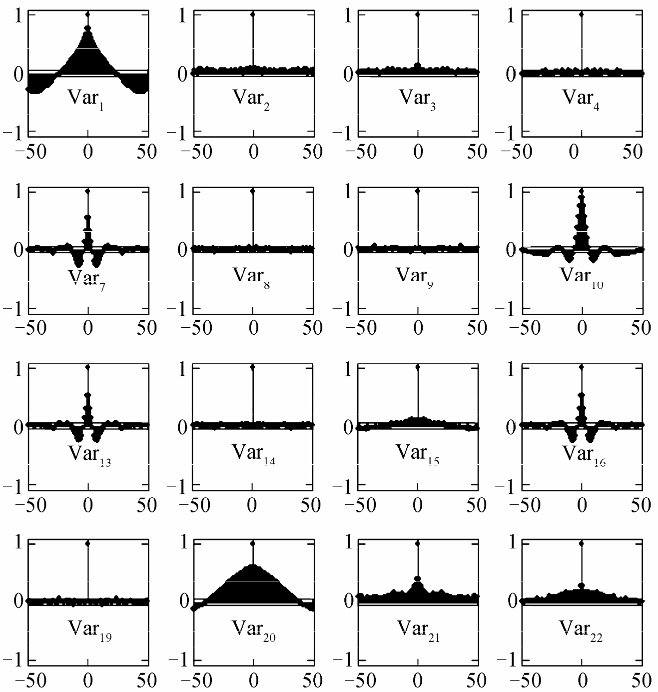

Figure 7 Auto-correlation plots of original variables for TE process

includes five process units: a reactor, a condenser, a vapour-liquid separator, a recycle compressor and a product stripper. There are 41 measurements and 12 manipulated variables. The process is open-loop unstable and requires regulatory controllers. TE process includes 20 programmed disturbances including composition step change in reactants, random reactor cooling water inlet temperature random disturbance and valve sticking *etc*. Detailed description about the operation of the TE process can be found in Ref.[19].

In this article, Ricker's TE simulator based on MATLAB® 6.5 was used to generate the dataset. The closed-loop control strategy of Ricker[22] was also adopted. The reference dataset to construct NOC model includes 2000 samples of 22 continuously measured variables which were recorded at 0.1h interval.

It is somewhat cumbersome to plot the ACFs of all the 22 variables. In Fig.7, only the auto-correlation coefficients of variables are illustrated. It can be seen that ten variables are strongly auto-correlated including feed A (Var_1), reactor pressure (Var_7), purge rate (Var_{10}), separator temperature (Var_{11}), separator pressure (Var_{13}), stripper pressure (Var_{16}), stripper temperature (Var_{18}), compressor power (Var_{20}), reactor outlet coolant temperature (Var_{21}), and separator outlet coolant temperature (Var_{22}).

PCA was first applied to the NOC data and 15 scores which explain that 87.6% of total variation were retained. The auto-correlation coefficients of scores are illustrated in Fig.8, not only the variation information but also the dynamic information are concentrated in the first five scores.

Based on the criterion presented in section 3.2, a five-order state-space model was identified to remove the dynamics efficiently. Fig.9 confirms that residuals are no longer auto or cross correlated. In contrast, the ARMA-PCA had to use 20-order filters to remove the auto-correlations of the first two scores.

To demonstrate the capability of SI-PCA to detect abnormal behavior, excessive variation of the reactor cooling water temperature was simulated (disturbance type 11) to generate a fault dataset. The fault dataset also contains 2000 samples and the fault was

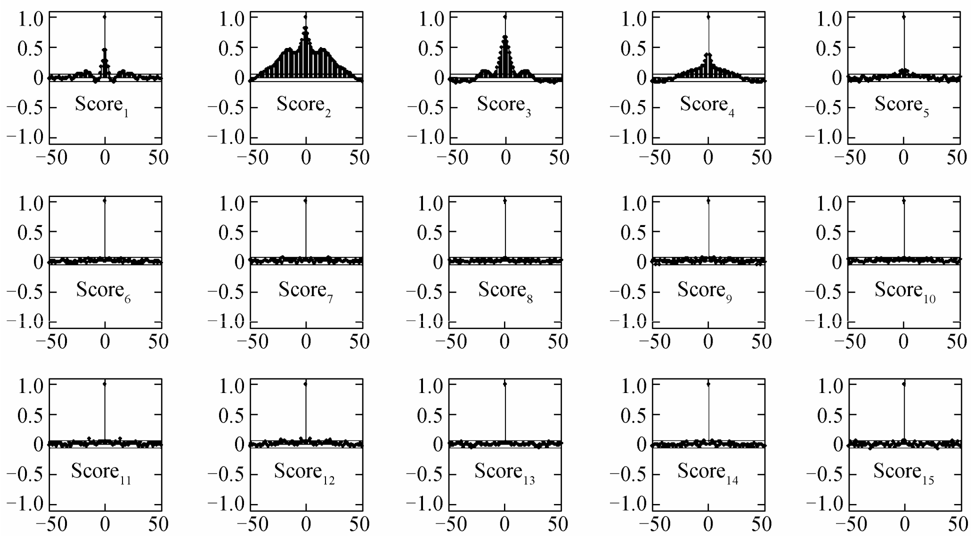

Figure 8 Auto-correlation plots of PCA scores for TE process

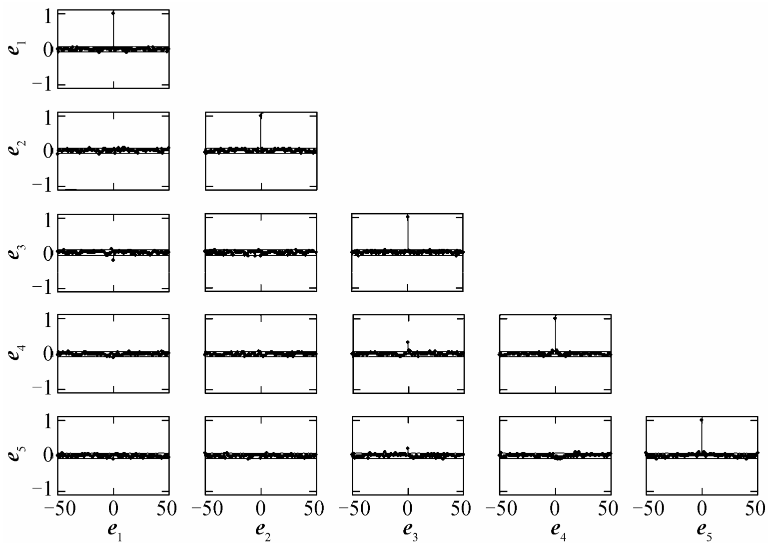

Figure 9 ACFs of residuals by SI-PCA for process

injected after the 1000 sample.

The T^2 control chart is presented in Fig.10 and the data around the control limit are zoomed in Fig.11 for clear illustration. The last 1000 values of T^2 statistic produce an excessive number of violations. Therefore, SI-PCA detects the out-of-control situation correctly.

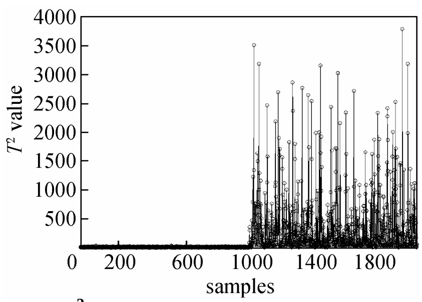

Figure 10 T^2 control chart for reactor cooling water excessive variation by SI-PCA

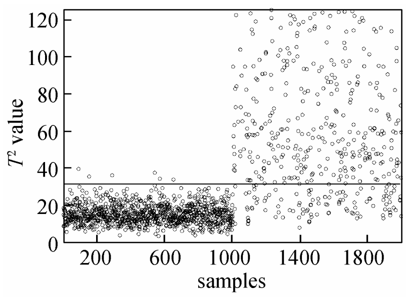

Figure 11 Zooming in the data of Fig.10 around the 99% control limit (solid line)

5 CONCLUSIONS

This article studies the influences of process dynamics on FAR of statistical process control. It is shown that the presence of auto and cross correlation will cause the actual FAR to deviate from its theoretic value. Applications illustrate that two improved PCA methods, dynamic PCA and ARMA-PCA, do not remove the effects of time-series structures of variables.

Motivated by the above facts, a subspace identification approach-based PCA (SI-PCA) is proposed to remove the auto and cross correlations of score variables simultaneously. Akaike information-like criterion is introduced to determine the state-space model. Numerical examples demonstrate that SI-PCA can remove the auto and cross correlations efficiently and the FAR of SI-PCA coincides more with the theoretic value respecting to the standard deviation of FAR.

Application in TE process fault detection reveals that the SI-PCA is also efficient in detecting the process abnormal behavior, that is, SI-PCA does not compromise the type Ⅱ error or missing fault detection performance.

REFERENCES

1 Nomikos, P., MacGregor, J.F., "Multivariate SPC charts for monitoring batch processes", *Technometrics*, **37**, 41—59(1995).
2 Wise, B., Gallagher, N.B., "The process chemometrics approach to process monitoring and fault detection", *J. Process Contr*., **6**, 329—348(1996).
3 Cinar, A., Ündey, C., "Statistical process and controller performance monitoring: A tutorial on current methods and future direction", In: Proceedings of American Control Conference, Ⅳ, 2625—2639(1999).
4 Xie, L., Zhang, J., Wang, S., "A robust statistical batch process monitoring framework and its application", *Chin. J. Chem. Eng*., **12**, 682—687(2004).
5 Zhao, L.J., Cai, T.Y., Wang, G., "Double moving window MPCA for online adaptive batch monitoring", *Chin. J. Chem. Eng*., **13**, 649—655(2005).
6 Zhou, S., Xie, L., Wang, S., "On-line fault diagnosis in industrial processes using variable moving window and Hidden Markov Model", *Chin. J. Chem. Eng*., **13**, 388—395(2005)
7 Jackson, J.E., A User's Guide to Principal Components, Wiley Inter-science, New York (1991).
8 Ku, W., Storer, R.H., Georgakis, C., "Disturbance rejection and isolation by dynamic principal component analysis", *Chemometr. Intell. Lab*., **30**, 179—196(1995).
9 Negiz, A., Cinar, A., "Statistical monitoring of multivariate dynamic processes with state space models", *AIChE J*., **43**, 2002—2020(1997).
10 Callao, M.P., Rius, A., "Time series: a complementary technique to control charts for monitoring analytical systems", *Chemometr. Intell. Lab*., **66**, 79—87(2003).
11 Simoglou, A., Martin, E.B., Morris, A.J., "Statistical performance monitoring of dynamic multivariate processes using state space modeling", *Comput. Chem. Eng.*, **26**, 909—920(2002).
12 Chen, J., Liu, K., "On-line batch process monitoring using dynamic PCA and dynamic PLS models", *Chem. Eng. Sci*., **57**, 63—75(2002).
13 Kruger, U., Zhou, Y., Irwin G.W., "Improved principal component monitoring of large scale processes", *J. Process Contr.*, **14**, 879—888(2004).
14 Åström, K.J., Introduction to Stochastic Control Theory, Academic Press, New York (1970).
15 Ljung, L., System Identification-Theory for the User, Prentice Hall, Englewood Clifs (1999).
16 Ljung, L., System Identification Toolbox-for Use with MATLAB, Mathwork, Natick (2002).
17 George, B., Gwilym, M.J., Gregory, R., Time series analysis: forecasting & control, 3rd edition, Prentice Hall, Englewood Clifs (1994).
18 Van, Overschee, P., De, Moor, B., Subspace Identification for Linear Systems: Theory Implementation Applications, Kluwer Academic Publishers, Boston-London-Dordrecht (1996).
19 Downs, J.J., Vogel, E.F., "A plant-wide industrial process control problem", *Comput. Chem. Eng*., **17**, 245—255(1993).
20 Chiang, L.H., Russell, E.L., Braatz, R.D., Fault Detection and Diagnosis in Industrial Systems, Springer-Verlag, Heidelberg (2000).
21 Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., "Comparison of statistical process monitoring method: application to the Eastman challenge problem", *Comput. Chem. Eng*., **24**, 175—181(2000).
22 Ricker, N.L., "Decentralized control of the Tennessee Eastman challenge process", *J. Process Contr*., **6**, 205—221(1996).
23 Russell, E.L., Chiang, L.H., Braatz, R.D., "Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis", *Chemometr. Intell. Lab*., **51**, 81—93(2000).