On-line Batch Process Monitoring Using Different Unfolding Method and Independent Component Analysis

JONG-MIN LEE[1], CHANGKYOO YOO[2]
AND IN-BEUM LEE[1]
[1]*Department of Chemical Engineering, Pohang University of Science and Technology, Pohang 790-784, Korea*
[2]*BIOMATH, Ghent University, B-9000 Gent, Belgium*

Keywords: Batch Monitoring, Fault Detection, Independent Component Analysis (ICA), Kernel Density Estimation, Principal Component Analysis (PCA), Process Monitoring

In many industries, the effective monitoring and control of batch processes is crucial to the production of high-quality materials. Several techniques using multivariate statistical analysis have been developed for monitoring and fault detection of batch processes. Multiway principal component analysis (MPCA) has shown a powerful monitoring performance in many industrial batch processes. However, it has shortcomings that all batch lengths should be equalized and future values of batches should be estimated for on-line monitoring. In order to overcome these drawbacks and obtain better monitoring performance, we propose a new statistical method for on-line batch process monitoring that uses different unfolding method and independent component analysis (ICA). If the measured data set contains non-Gaussian latent variables, the ICA solution can extract the original source signal to a much greater extent than the PCA solution since ICA involves higher-order statistics and is not based on the assumption that the latent variables follow a multivariate Gaussian distribution. The proposed monitoring method was applied to fault detection and identification in the simulation benchmark of the fed-batch penicillin production, which is characterized by some fault sources with non-Gaussian characteristics. The simulation results clearly show the power and advantages of the proposed method in comparison to MPCA.

Introduction

Batch and semi-batch processes play an important role in most industries in order to produce high-quality materials in the chemical, biological, pharmaceutical, and semi-conductor industries. Batch processes are characterized by pre-described processing of raw materials for a finite duration to convert them to products. The main objective of operating batch processes is to achieve consistent and reproducible qualities with competitive prices. However, some characteristics of batch processes make it difficult to monitor it correctly. Batch processes generally exhibit some batch-to-batch variations in their specified trajectories and small changes in operating conditions during critical periods may impact final product qualities and yields. Furthermore, produced quality variables, the key indicators of the process performance, are often examined off-line in laboratories (Chen and Liu, 2002). Hence, although operators may observe a problem in the quality of the final products of a process, they may not know the source of the problem or when it occurred. Therefore, on-line monitoring and diagnosis of batch processes are crucial to detect faults that can be corrected prior to completion of the batch or can be corrected in subsequent batches. Early detection of problems and correction of deviation by on-line monitoring before the completion of the batch can save the batch and reduce the number of rejected batches.

Several techniques using the multivariate statistical analysis have been developed for monitoring of batch processes. Nomikos and MacGregor (1994, 1995a) have extended the multivariate statistical process control (SPC) methods of Principal Componetn Anaylsis (PCA) to batch processes, where the method is called multiway PCA (MPCA). They also developed the multiway partial least square (MPLS) method for monitoring processes for which both process data and product quality data are available (Nomikos and MacGregor, 1995b). These approaches allow the monitoring of a batch process to be achieved once a model has been developed from nominal or good batch operation. Dong and McAvoy (1996) developed a nonlinear principal component analysis (NLPCA) method based on principal curves and neural networks to monitor batch processes. Rännar *et al.* (1998) suggested an adaptive batch monitoring method that uses

Received on March 28, 2003. Correspondence concerning this article should be addressed to I.-B. Lee (E-mail address: iblee@postech.edu).

 Copyright © 2003 The Society of Chemical Engineers, Japan

hierarchical PCA. This approach has significant benefits in monitoring multi-stage batch processes where the latent variable structure can change at several points during the batch. Chen and Liu (2002) suggested batch dynamic principal component analysis (BDPCA) and batch dynamic partial least square (BDPLS). These methods integrate the time-lagged windows of the process dynamic behavior with PCA and PLS respectively for on-line batch monitoring. Of the methods mentioned above, the application of MPCA to industrial batch monitoring has received the greatest attention (Gallagher and Wise, 1996; Kosanovich *et al.*, 1996; Nomikos, 1996; Gregersen and Jorgensen, 1999; Tates *et al.*, 1999; Westerhuis *et al.*, 1999; Wise *et al.*, 1999; Lennox *et al.*, 2001).

The objective of this article is to present a new method for on-line batch process monitoring using different unfolding method and independent component analysis (ICA). ICA looks for components that are both statistically independent and non-Gaussian from multivariate observed data. Although ICA can be considered as a useful extension of PCA, its objective differs from that of PCA. PCA is a dimensionality reduction technique in terms of capturing the variance of the data, and it is capable of extracting uncorrelated latent variables from highly correlated data. However, its objective is only to decorrelate variables, not to make them independent. PCA can only impose independence up to second order statistics information (mean and variance) while constraining the direction vectors to be orthogonal, whereas ICA has no orthogonality constraint and involves higher-order statistics (Lee, 1998). Conventional statistical process monitoring methods using decorrelation techniques such as PCA are based on the assumption that the measured values of product qualities are normally distributed. However, this assumption is often invalid for measurements from actual chemical processes because of their dynamic and nonlinear nature. As a result, the control limits such as T^2 and squared prediction error (SPE) charts used in PCA monitoring may cause false alarms. In contrast, non-Gaussian variables are usually required for ICA algorithms based on higher order statistics; hence, monitoring based on ICA algorithms will give better results (Kano *et al.*, 2003; Lee *et al.*, 2003a, 2003b).

This paper is organized as follows. Conventional MPCA monitoring and new batch monitoring based on ICA are introduced in the next sections, followed by a brief introduction to the ICA algorithm. The monitoring statistics of the proposed method are then established and an explanation is given for the kernel density estimation used to calculate the confidence limit for non-Gaussian data. The superiority of process monitoring using the proposed method is illustrated by applying it to the simulation benchmark of fed-batch penicillin production. Finally, conclusions are given.

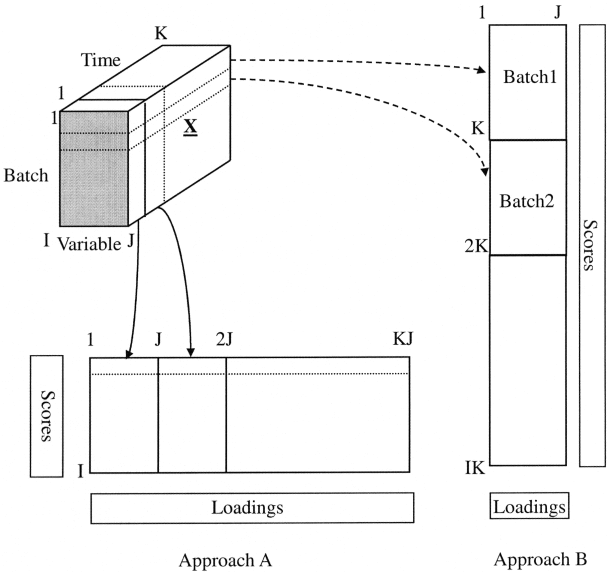


Fig. 1 Two unfolding methods

1. MPCA

Batch process data are naturally represented as a three-way matrix $\mathbf{X}$ ($I \times J \times K$). In a typical batch run, $j = 1, 2, \cdots, J$ variables are measured at $k = 1, 2, \cdots, K$ time intervals throughout the batch. In addition, there exist data from multiple ($i = 1, 2, \cdots, I$) equivalent process batch runs. In MPCA, this three-way matrix must be unfolded to obtain a two-way matrix on which PCA can be performed. There are two main approaches to unfolding the three-way matrix (Dong and McAvoy, 1996; Kosanovich *et al.*, 1996; Albert and Kinley, 2001), which give rise to distinct two-way matrices:

- Approach A: Variables × time for each specific batch
- Approach B: Batches × times for each specific variable

Approach A allows one to analyze the variability among the batches by summarizing the information in the data with respect to both the measured variables and their time variation; in contrast, approach B can be used to obtain information on the variability among the batch variables. Approach B focuses on the local behavior of the process, whilst approach A aims to monitor the final behavior of the batch (Meng *et al.*, 2003). **Figure 1** shows schematic diagrams of the unfolding procedures used in approaches A and B.

In most articles, the approach A called MacGregor's unfolding is accepted for analyzing batch process data and the PCA analysis on their unfolding method is called simply MPCA. This approach enables simply integrating off-line data records with the unfolded on-line type data. In addition, it is suitable for deriving estimators for final quality measures from process data that are usually not available until the batch terminates. To avoid the problems created by

nonlinearities in the data, the major nonlinear behavior of the process is eliminated from the unfolded matrix by removing the mean trajectory of each variable. This is achieved by subtracting off the mean of each column from the corresponding column in the unfolded matrix (Nomikos and MacGregor, 1994, 1995a). Once the matrix is mean centered and variance scaled, PCA is performed as follows

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} = \sum_{i=1}^{a} \mathbf{t}_i \mathbf{p}_i^T + \mathbf{E} \tag{1}$$

where a is the number of principal components, $\mathbf{T} = [\mathbf{t}_1\ \mathbf{t}_2 \cdots \mathbf{t}_a]$ is a score matrix, and $\mathbf{P} = [\mathbf{p}_1\ \mathbf{p}_2 \cdots \mathbf{p}_a]$ is a loading matrix. The score vector ($\mathbf{t}_i$) contains information about relation of batches, and the loading vectors ($\mathbf{p}_i$) have a weight for each variable at each time; hence they give the history of the process.

For on-line monitoring, two statistics, the T^2 and squared prediction error (*SPE*) of approach A are calculated as follows (Gregersen and Jorgensen, 1999). T^2 can be defined as follows and its confidence limits can be obtained from the following F-distribution:

$$T^2 = \mathbf{t}^T \mathbf{S}^{-1} \mathbf{t} \sim \frac{a\left(I^2 - 1\right)}{I(I-a)} F_{a,I-a} \tag{2}$$

where $\mathbf{t}$ is the vector containing the scores of a given batch from the retained principal components and $\mathbf{S}$ is the covariance matrix of the t-scores calculated during the model development, a is the number of principal components and I is the number of batches used for modeling. The *SPE* is a measure of the lack of fit with the established model and is defined as the sum of the squares of the errors at time interval k

$$SPE_k = \sum_{r=(k-1)J+1}^{Jk} \mathbf{e}_r{}^T \mathbf{e}_r \tag{3}$$

where e_r is the r-th column of the matrix $\mathbf{E}$ in Eq. (1). The distribution of the SPE_k can be approximated by a weighted χ^2 distribution $SPE_k \sim (v_k/2m_k)\,\chi^2_{2m_k^2/v_k}$, where m_k and v_k are the mean and variance of the *SPE* at time instant k obtained for the dataset used for the model development (Nomikos and MacGregor, 1995a).

However, there are two drawbacks for monitoring new batches when approach A is used. When a batch is monitored, we know only the values from the beginning of the batch to the current time, but test data should be completed until the end of the batch for on-line monitoring. Several methods have been proposed for estimating the variable trajectory to the end of the batch. Nomikos and MacGregor (1994, 1995a) suggest three different ways of dealing with this problem.

a) Fill the missing value with zeros.
b) Fill the missing value with the current value
c) Use the ability of PCA that handles missing data

However, anticipating the future observation might cause false alarms because the predicted values without considering the dynamic relationship may distort the data information (Chen and Liu, 2002). Alternative methods such as the recursive hierarchical PCA (HPCA) method have been proposed to eliminate the need for estimating the future portion of the process variable trajectories (Rännar *et al*., 1998). The second drawback of approach A is that all batch lengths are assumed to be of the same length. However, there are many situations in which the total time duration of the batches and the duration of various stages within batches are not the same. One method for synchronizing batches was suggested by Kassidas *et al*. (1998). They used dynamic time warping (DTW) in order to align batches with different duration. Kourti (2003) provides an overview of the various methods that have been proposed for synchronizing batches.

Wold *et al*. (1998) used approach B for modeling and diagnosing batch processes. They introduced the vector $\mathbf{Y}$ to represent the maturity of the batch, and used a PLS model to model the process variables. This approach does not require that all batches be of equal length, nor does it require estimation of future missing values when used for on-line monitoring. However, mean centering of the unfolding matrix generated by this approach cannot remove the batch process trajectory because it is like monitoring the batch around the mean of the variables during all the batches. Furthermore, the resultant loadings will reflect the correlations among, and relative importance of the process variables from an "overall" perspective, without taking into account the time dependency and correlation.

2. A New Batch Monitoring Method

As mentioned in Section 1, the two unfolding methods each have drawbacks as well as advantages. In this paper, we present a new monitoring method that combines the advantages of each approach and that uses ICA instead of PCA. The procedure followed in this new method is presented in **Fig. 2**. The new monitoring method is mainly based on the unfolding approach B. First, the batch data matrix $\mathbf{X}$ ($I \times J \times K$) is unfolded as in Macgregor's approach (approach A). Then, the variables at each time (columns of the unfolded matrix) are mean centered and scaled to remove the batch process trajectory and to attenuate any time correlation. Subtracting the average batch trajectory is an essential step that removes much of the nonlinear dynamics from the batch data. After subtracting off the batch trajectory and scaling the variance at each time, the unfolded matrix ($I \times JK$) is rearranged to the form

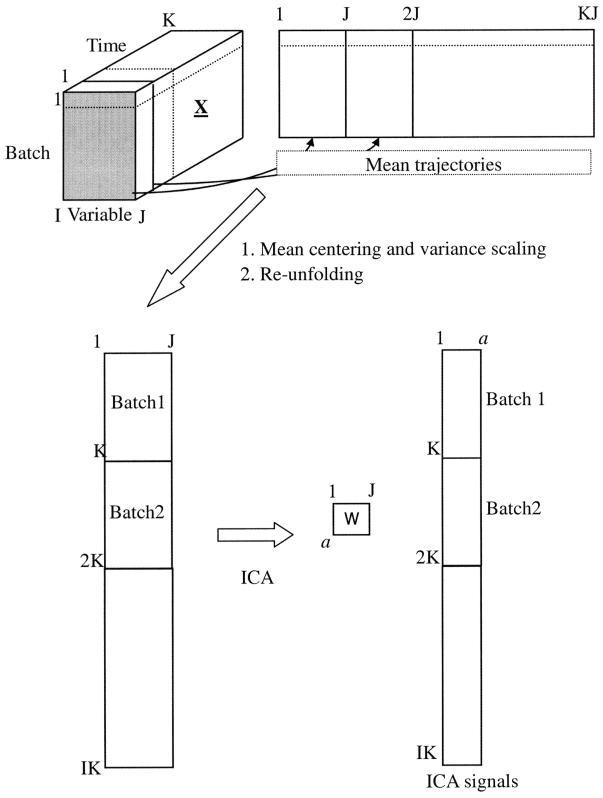


Fig. 2 New monitoring procedure that uses ICA

($J \times IK$) as in Wold's approach (approach B). Then, ICA is used to extract the underlying factors or components from the multivariate statistical data matrix ($J \times IK$) (see Appendix A for details). The proposed procedure has several merits that make it particularly well-suited to on-line monitoring of batch processes: (1) future missing values need not be estimated, (2) the batch lengths of all batches need not be aligned to equal length, and (3) process faults are easily detected because the procedure removes the majority of the nonlinear dynamics and because ICA is better than PCA at identifying the underlying process factors.

3. Independent Component Analysis (ICA)

ICA is a statistical and computational technique for revealing hidden factors in sets of random variables, measurements, or signals. ICA was originally proposed to solve the blind source separation problem, which involves recovering independent source signals (e.g., voice, music, or noise sources) after they have been linearly mixed by an unknown matrix, **A** (Vigário, 1997). Using a simple example, Lee *et al.* (2003a) and Kano *et al.* (2003) showed that if the latent variable follows a non-Gaussian distribution, ICA extracts the original source signal to a much greater extent than PCA. Based on this finding, they proposed an ICA-based method for process monitoring. On the other hand, if the several latent variables are Gaussian, the monitoring performance of ICA may not be superior to that of PCA (Kano *et al.*, 2002). In the case of Gaussian independent components, we can only estimate the ICA model up to an orthogonal transformation (Hyvärinen *et al.*, 2001); that is, the mixing matrix **A** is not identifiable for Gaussian independent components. Kano *et al.* (2002) combined PCA-based SPC and ICA-based SPC to develop a new multivariate statistical process control (MSPC) method. However, the actual distributions of latent variables underlying the measured data are unknown although the several measured variables follow Gaussian distribution. In real processes, the latent variables tend to be non-Gaussian because of their dynamic and nonlinear nature. Therefore, in most cases the basic assumption of ICA, that the independent components are non-Gaussian, is acceptable.

To introduce the ICA algorithm, let us assume that at sample k the observed d-dimensional data vector, $\mathbf{x}(k) = [x_1(k), \cdots, x_d(k)]^T$, is given by the model (Hyvärinen and Oja, 2000; Li and Wang, 2002):

$$\mathbf{x}(k) = \sum_{j=1}^{m} \mathbf{a}_j s_j(k) = \mathbf{A}\mathbf{s}(k) \tag{4}$$

where $\mathbf{x}(k)$ is linearly mixed with m components of $s_j(k)$. This can be represented in a matrix form as follows:

$$\mathbf{X} = \mathbf{AS} + \mathbf{E} \tag{5}$$

where $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \cdots, \mathbf{x}(n)] \in R^{d \times n}$ is the data matrix, $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_m] \in R^{d \times m}$ is the mixing matrix, $\mathbf{S} = [\mathbf{s}(1), \mathbf{s}(2), \cdots, \mathbf{s}(n)] \in R^{m \times n}$ is the independent component matrix, $\mathbf{E} \in R^{d \times n}$ is the residual matrix, d is the number of variables, n is the number of samples, and m is the number of independent components. Here, we assume $d \geq m$; when d equals m, the residual matrix becomes the zero matrix. In contrast to PCA, ICA employs the transposed data matrix. The basic problem of ICA is to estimate the original components **S** or to estimate **A** from **X** without any knowledge of them. Therefore, the objective of ICA is to calculate a separating matrix **W** so that the components of the reconstructed data matrix $\hat{\mathbf{S}}$, given as

$$\hat{\mathbf{S}} = \mathbf{WX} \tag{6}$$

become as independent of each other as possible. Using the ICA algorithm, we can obtain the rows of $\hat{\mathbf{S}}$ whose norm is 1.

From now on, we assume d equals m unless otherwise specified. The initial step of ICA is whitening, which eliminates all the cross-correlation between

random variables (Hyvärinen and Oja, 2000). The whitening transformation is expressed as

$$\mathbf{z}(k) = \mathbf{Q}\mathbf{x}(k) \tag{7}$$

where $\mathbf{Q}$ can be obtained from the eigen-decomposition of the covariance matrix $\mathbf{R}_x = E(\mathbf{x}(k)\mathbf{x}^T(k))$ (Lee *et al*., 2003a). After the transformation we have

$$\mathbf{z}(k) = \mathbf{Q}\mathbf{x}(k) = \mathbf{Q}\mathbf{A}\mathbf{s}(k) = \mathbf{B}\mathbf{s}(k) \tag{8}$$

where $\mathbf{B} = \mathbf{Q}\mathbf{A}$ becomes an orthogonal matrix. We have therefore reduced the problem of finding an arbitrary full-rank matrix $\mathbf{A}$ to the simpler problem of finding an orthogonal matrix $\mathbf{B}$, which then gives

$$\hat{\mathbf{s}}(k) = \mathbf{B}^T\mathbf{z}(k) = \mathbf{B}^T\mathbf{Q}\mathbf{x}(k) \tag{9}$$

From Eqs. (6) and (9), the relation between $\mathbf{W}$ and $\mathbf{B}$ can be expressed as

$$\mathbf{W} = \mathbf{B}^T\mathbf{Q} \tag{10}$$

To calculate $\mathbf{B}$, it is initialized and then updated so that $\hat{\mathbf{s}}(k) = \mathbf{B}^T\mathbf{z}(k)$ may have great non-Gaussianity. There are two common measures of non-Gaussianity: kurtosis and negentropy (Hyvärinen and Oja, 2000). Hyvärinen (1999a) suggested a fast and robust fixed-point algorithm for ICA that entails maximizing the negentropy under the constraint of $\|\mathbf{b}_i\| = 1$, where $\mathbf{b}_i$ is the i-th column of $\mathbf{B}$. After finding $\mathbf{B}$, the demixing matrix $\mathbf{W}$ can be obtained from Eq. (10).

4. Ordering and Dimension Reduction of ICA

Dimension reduction of ICA is based on the idea that these measured variables are the mixture of some independent variables (Cheung and Xu, 2001; Li and Wang, 2002). An important part of ICA monitoring is the selection of a small number of dominant components from the list of all independent components. In PCA, the order of the score vectors is determined by their variance. Therefore, data dimension can be reduced by selecting dominant score vectors. However, the ordering of components is very difficult in ICA and there is no standard criterion. A number of methods have been suggested to determine the component order (Cardoso and Souloumica, 1993; Back and Weigend, 1997; Cheung and Xu, 1999, 2001; Hyvärinen, 1999b).

In the present study, we used a Euclidean norm (L_2) to sort the rows of the demixing matrix, $\mathbf{W}$, because this method is very simple and gives better results in ICA monitoring than the other methods. Hence, the order of the independent components (ICs) is decided by the L_2 norm of each $\mathbf{w}_i$, the row of $\mathbf{W}$:

$$\arg\underset{i}{Max} \left\|\mathbf{w}_i\right\|_2 \tag{11}$$

That is, ICs are sorted using an L_2 norm in order to show only those ICs that cause dominant changes in the process.

After the ordering of ICs, it is important to select the optimal number of ICs in order to achieve good monitoring and prediction; selecting too many ICs will cause a magnification of noise and poor process monitoring performance. The data dimension can be reduced by selecting a few rows of $\mathbf{W}$ based upon the assumption that the rows with the largest sum of squares coefficient have the greatest effect on the variation of $\hat{\mathbf{S}}$. This approach is based on the idea that the dominant process variation can be monitored by considering the cumulative sums of only the first few dominant ICs (Cheung and Xu, 2001). We used a graphical technique to determine the number of ICs similar to the SCREE test of PCA (Lee *et al*., 2003a, 2003b).

5. Process Monitoring Statistics of ICA

On-line monitoring of measurement variables is carried out with the aim of continuously analyzing and interpreting the measurements in order to detect and isolate disturbances and faults. The implementations of the monitoring statistics of ICA are similar to those for the monitoring statistics of PCA. The ICA model is based on historical data collected during normal operation, i.e., when the product was being manufactured and only common cause variations were present. Future process behavior is then compared against this 'normal' or 'in-control' representation.

In the normal operating condition, designated $\mathbf{X}_{\text{normal}}$ ($J \times IK$), $\mathbf{W}$ as well as $\hat{\mathbf{S}}_{\text{normal}}$ are obtained from the FastICA algorithm ($\hat{\mathbf{S}}_{\text{normal}} = \mathbf{W}\mathbf{X}_{\text{normal}}$) under the assumption that the number of variables is equal to the number of independent components (For the detail monitoring procedure, refer to Appendix). The matrices $\mathbf{B}$, $\mathbf{Q}$ and $\mathbf{A}$ used in Eq. (8) are also obtained by whitening and the FastICA algorithm. As mentioned in the previous section, the data dimension can be reduced by selecting a few rows of $\mathbf{W}$ based upon the assumption that the rows with the largest sum of squares coefficient have the greatest effect on the variation of $\hat{\mathbf{S}}$. The selected a rows of $\mathbf{W}$ constitute a reduced matrix $\mathbf{W}_{\text{d}}$ (deterministic part of $\mathbf{W}$), and the remainder rows of $\mathbf{W}$ constitute a reduced matrix $\mathbf{W}_{\text{e}}$ (excluded part of $\mathbf{W}$). We can construct a reduced matrix $\mathbf{B}_{\text{d}}$ by selecting the columns from $\mathbf{B}$ whose index corresponds to the indices of the selected rows of $\mathbf{W}$. The remainder columns of $\mathbf{B}$ constitute $\mathbf{B}_{\text{e}}$. Then, new independent vectors, $\hat{\mathbf{s}}_{\text{newd}}(k)$ and $\hat{\mathbf{s}}_{\text{newe}}(k)$ can be obtained if new data $\mathbf{x}_{\text{new}}(k)$ is transformed through the

separating matrices $\mathbf{W}_d$ and $\mathbf{W}_e$, i.e., $\hat{\mathbf{s}}_{newd}(k) = \mathbf{W}_d\mathbf{x}_{new}(k)$ and $\hat{\mathbf{s}}_{newe}(k) = \mathbf{W}_e\mathbf{x}_{new}(k)$, respectively.

T^2 and *SPE* charts are useful monitoring methods when PCA is used. Two types of statistics are calculated from the process model in normal operation: the *D*-statistic for the systematic part of the process variation and the *Q*-statistic for the residual part of the process variation. The *D*- and *Q*-statistics calculated for the new data object should be within the confidence limits for the data to be in control. Similarly, these methods can be applied to ICA monitoring.

The *D*-statistic for sample *k*, also known as the I^2 statistic, is the sum of the squared independent scores and is defined as follows:

$$I^2(k) = \hat{\mathbf{s}}_{newd}(k)^T \hat{\mathbf{s}}_{newd}(k) \tag{12}$$

The *Q*-statistic for the nonsystematic part of the common cause variation of new data, also known as the *SPE* statistic, can be visualized in a chart with confidence limits. The *SPE* statistic at sample *k* is defined as follows:

$$SPE(k) = \sum_{j=1}^{d} \left(x_{new,j}(k) - \hat{x}_{new,j}(k)\right)^2 \tag{13}$$

where *d* is the number of variables, $x_{new,j}(k)$ is the *j*-th element of $\mathbf{x}_{new}(k)$, and $\hat{x}_{new,j}(k)$ is the *j*-th element of $\hat{\mathbf{x}}_{new}(k)$. $\hat{\mathbf{x}}_{new}(k)$ can be predicted as follows.

$$\hat{\mathbf{x}}_{new}(k) = \mathbf{Q}^{-1}\mathbf{B}_d\hat{\mathbf{s}}_{newd}(k) = \mathbf{Q}^{-1}\mathbf{B}_d\mathbf{W}_d\mathbf{x}_{new}(k) \tag{14}$$

Lee *et al*. (2003a) proposed another statistics, that is, the I^2 metrics of the excluded independent components. Monitoring the non-systematic part of the measurements provides an additional fault detection tool, which can detect special events entering the system. I_e^2 statistics has the further advantage that it can compensate for the error that results when an incorrect number of ICs is selected for the deterministic part. The use of I^2 and I_e^2 statistics allows the whole space spanned by the original measured variables to be monitored through a new basis of dependent components. I_e^2 statistic can be defined as follows:

$$I_e^2(k) = \hat{\mathbf{s}}_{newe}(k)^T \hat{\mathbf{s}}_{newe}(k) \tag{15}$$

where $\hat{\mathbf{s}}_{new}(k) = \mathbf{W}_e\mathbf{x}_{new}(k)$.

In PCA monitoring, the confidence limit is approximated by a specified distribution based upon the assumption that the latent variables follow a Gaussian distribution. However, the monitoring charts indicate many false alarms since this assumption is not valid in some cases (Martin *et al*., 1996). On the other hand, the independent components over some period do not conform to a multivariate Gaussian distribution; hence, the confidence limits of the I^2, I_e^2 and *SPE* statistics cannot be determined directly from a particular approximate distribution. Consequently, we need to find an alternative method. The confidence limits of the three statistics, I^2, I_e^2 and *SPE*, can be obtained by kernel density estimation, which will be explained in the next section.

6. Confidence Bounds

A univariate kernel estimator with kernel *K* is defined by:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left\{\frac{x - x_i}{h}\right\} \tag{16}$$

where *x* is the data point under consideration, x_i is an observation value from the data set, *h* is the window width (also known as the smoothing parameter), *n* is the number of observations, and *K* is the kernel function. The kernel estimator is therefore a sum of 'bumps' located at the observations. The kernel function *K* determines the shape of the bumps and satisfies the condition

$$\int_{-\infty}^{\infty} K(x)dx = 1 \tag{17}$$

There are a number of possible kernel functions. In practice, the form of the kernel function is not important, and the Gaussian kernel function is the most commonly used (Silverman, 1986). The Gaussian kernel is also employed in the present study.

Many measures have been proposed for the estimation of *h*, the window width or smoothing parameter. The problem of choosing how this smoothing parameter is of crucial importance in density estimation. In this paper, in order to reduce the complexity, we use a simple method to select *h*, $h = 1.06sn^{-1/5}$, where *s* is the standard deviation of samples and *n* is the number of samples. For more details regarding kernel density estimation, refer to the books of Silverman (1986) and Wand and Jones (1995).

One major advantage of the confidence region obtained by kernel density estimation is that it follows the data more closely and is less likely to incorporate regions of unknown operation, compared with the confidence regions obtained on the basis of Hotelling's T^2 statistics (Chen *et al*., 2000).

7. Contribution Plots

In the previous section it was stated that process faults are detected by computing three multivariate control charts. However, the monitoring charts give no information on what is wrong with the process, or which process variables caused the process to be out of control. Once a fault is detected by the statistical monitoring method, the key approach to fault identification using the ICA model is the use of contribution plots. By interrogating the underlying process model at the point where an event has been detected, contribution plots may reveal the group of process variables that most influence the model or the residual (MacGregor *et al*., 1994; Rosen, 1998; Teppola, 1999; Westerhuis *et al*., 2000).

Lee *et al*. (2003a) suggest an equation to calculate the variable contribution for $I_{\text{newd}}{}^2(k)$ and $I_{\text{newe}}{}^2(k)$, respectively. These are as follows.

• Variable contribution to $I^2(k)$

$$\mathbf{x}_{\text{cd}}(k) = \frac{\mathbf{Q}^{-1}\mathbf{B}_{\text{d}}\hat{\mathbf{s}}_{\text{newd}}(k)}{\left\|\mathbf{Q}^{-1}\mathbf{B}_{\text{d}}\hat{\mathbf{s}}_{\text{newd}}(k)\right\|}\left\|\hat{\mathbf{s}}_{\text{newd}}(k)\right\| \tag{18}$$

• Variable contribution to $I_{\text{e}}{}^2(k)$

$$\mathbf{x}_{\text{ce}}(k) = \frac{\mathbf{Q}^{-1}\mathbf{B}_{\text{e}}\hat{\mathbf{s}}_{\text{newe}}(k)}{\left\|\mathbf{Q}^{-1}\mathbf{B}_{\text{e}}\hat{\mathbf{s}}_{\text{newe}}(k)\right\|}\left\|\hat{\mathbf{s}}_{\text{newe}}(k)\right\| \tag{19}$$

where $\|\ \|$ is the L_2 norm of a vector.

Similarly, variable contributions can also be computed for the *SPE* statistic, i.e., the variable contribution of the residuals. Generally, the aberrant variables will have the largest residuals. The residual at sample k, $SPE(k)$, is defined as the sum of the squares of $\mathbf{e}(k)$. Thus, the vector $\mathbf{e}(k)$ contains information on the individual prediction errors of each process variable for sample k. By plotting $\mathbf{e}(k)$ as a bar graph, the contributions to $SPE(k)$ can be viewed. The relative size of the bars indicates the contribution from each variable to the prediction error, or the lack of fit of a sample to the model. In some situations, especially those with varying *SPE* statistic, it is a good idea to use a mean of the contribution to *SPE* over some period. This is done by replacing the vector $\mathbf{e}(k)$ by a vector expressing the mean error over a period of length l.

8. Case Study (Comparison of the Proposed Method with MPCA)

PCA uses only the information contained in the covariance matrix of the data vector $\mathbf{x}$, whereas ICA uses information on the distribution of $\mathbf{x}$ that is not contained in the covariance matrix. Hence, in comparison to PCA, ICA may give more sophisticated monitoring results because it uses the independent components rather than the principle components. In this paper, the FastICA algorithm developed by Hyvärinen and Oja (2000) was applied to the detection and diagnosis of faults during monitoring. The proposed monitoring method was applied to fault detection and diagnosis in fed-batch penicillin fermentation, a process that is known to have a variety of fault sources. The simulation results demonstrate the power and advantages of the proposed monitoring technique in comparison to the MPCA-based method.

8.1 Fed-batch penicillin fermentation

The production of secondary metabolites has been examined in numerous studies on account of its importance in academic and industrial contexts. Filamentous microorganisms are used in the commercial production of secondary metabolites such as antibiotics. It has been shown that the formation of the target product (i.e., the antibiotic) is usually not associated with cell growth. For this reason, it is common practice to first grow the microorganisms in a batch culture, and then to carry out a fed-batch operation to promote the synthesis of the antibiotic (Birol *et al*., 2002a, 2002b). In typical penicillin fermentation, most of the necessary cell mass is obtained during the first 40–45 hours. The penicillin starts to be produced at the exponential growth phase and continues to be produced until the stationary phase. Growth must continue at a certain minimum rate to maintain high levels of penicillin production. For this reason, glucose is fed continuously during fermentation instead of being added only at the beginning.

The Monitoring and Control Group of the Illinois Institute of Technology (Ündey, C., Birol, G., and Cinar, A.) has developed a simulator (PenSim v2.0) that is capable of simulating concentrations of biomass, CO_2, hydrogen ion, penicillin, carbon source, and oxygen, as well as heat generation, under various operating conditions. In their model, batch growth of cells followed by a fed-batch operation was considered to be consistent with the data available in literature (Ündey *et al*., 2000). The flow diagram of the penicillin fermentation process is illustrated in **Fig. 3**. For more detailed information about this simulator, please refer to the paper of Birol *et al*. (2002b) and the website of the Çinar group (http://www.chee.iit.edu/~cinar). The variables generated by the penicillin simulator can be separated into the following categories:

- Load variables: aeration rate, agitator power, substrate feed rate, and substrate feed temperature
- Manipulated variables: acid/base and heating/cooling water flow rates
- Internal state variables: culture volume, generated heat, carbon dioxide, dissolved oxygen, biomass, penicillin and substrate feed concentrations

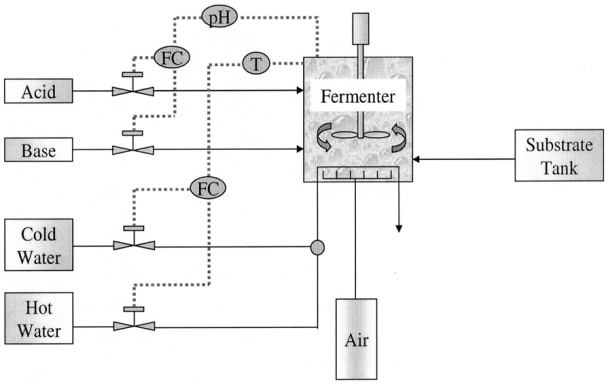


Fig. 3 Flow chart of the penicillin fermentation process

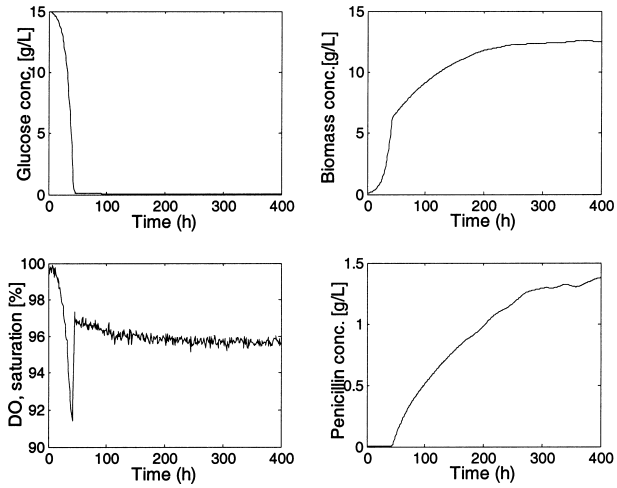


Fig. 4 Typical time profiles of glucose, penicillin, biomass and oxygen for a typical batch run

- Controlled variables: pH and bioreactor temperature

The time profiles of glucose, penicillin, biomass and oxygen during a typical batch run are shown in **Fig. 4**.

8.2 Equal batch length

In our study, a total of 50 batches that showed normal operation were used to develop the MPCA and the proposed model, providing the necessary information for on-line monitoring. The 11 variables considered for the monitoring are shown in **Table 1**. The duration of each batch was 400 h, comprising a preculture stage of about 45 h and a fed-batch stage of about 355 h. The sampling interval was 1 h. Small variations were added to the simulation input data to mimic process variations under normal operating conditions. Measurement noise was also added to each of the 11 monitored variables. Four principal components were selected by cross validation for MPCA modeling, while three independent components were chosen by the SCREE test for modeling the deterministic part of the proposed method. Using four principal components,

Table 1 Variables used in the monitoring of the benchmark model

No.	Variables
1	Aeration rate [L/h]
2	Agitator power [W]
3	Substrate feed rate [L/h]
4	Substrate feed temperature [K]
5	Dissolved oxygen concentration [g/L]
6	Culture volume [L]
7	Carbon dioxide concentraion [g/L]
8	pH
9	Fermentor temperature [K]
10	Generated heat [kcal]
11	Cooling water flow rate [L/h]

MPCA explained 63.6% of the variation. To fill-in the future values in $\mathbf{X}_{new}$, we used the ability of PCA to handle missing data (i.e., the third approach suggested by Nomikos and Macgregor (1995a)). Then, the models constructed using MPCA and the proposed method were tested against a new batch with a 99% control limit. The abnormal batch was monitored at every time point k with the monitoring charts. The results obtained using MPCA and the proposed method are shown in **Figs. 5** and **6**, respectively. From the beginning of a batch up to about 45 h, the control limits of each monitoring chart (except the T^2 chart) fluctuate due to batch-to-batch variations in the startup operating conditions. Furthermore, the control limits fluctuate to the greatest extent when the preculture stage ends and the fed-batch stage begins. The fluctuating control limits of the ICA monitoring charts were obtained by kernel density estimation based on the extracted independent components or residuals generated from normal batches at each time. Hence, the fluctuations of the control limits reflect the dynamic characteristics of normal operating batches and compensate for any disadvantage arising from the assumption that the correlations among the variables remain constant throughout the entire batch.

To create an abnormal batch, a fault was imposed on the batch; specifically, the substrate feed rate was decreased by 10% at time 50 h and then maintained at that lower value until the end of the batch. As shown in Fig. 5, the MPCA monitoring charts do not detect this fault. (This chart does briefly exceed the control limit at about time 195 h, but this does not represent the unequivocal detection of a fault given the meaning of 99% probability.) We also tested the monitoring performance of MPCA using the filling method in which the missing values are made equal to the current value; however, MPCA still did not detect the fault (figure not shown). However, in comparison to MPCA, the proposed method does detect the abnormal situation although the dynamicity of the process is strong.

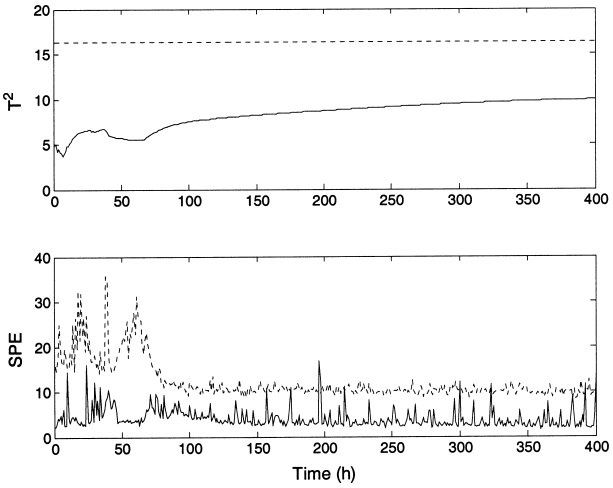


Fig. 5 T^2 and *SPE* charts for on-line monitoring using MPCA

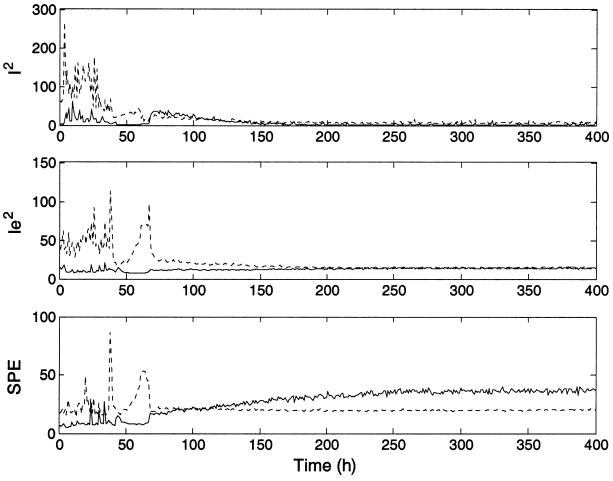


Fig. 6 I^2, I_e^2 and SPE charts for on-line monitoring using the proposed method

As shown in Fig. 6, the I^2 chart displays a significant deviation from time 68 h onwards, which represents a delay of about 18 h from the onset of the fault. This delay reflects the fact that the effect of the glucose substrate feed rate is only slowly propagated to other variables. If the dynamic characteristics of batch data are removed by other processing in advance, the lag between event and detection will be shortened. Another interesting feature of Fig. 6 is that the *SPE* values exceed the 99% control limit from about 110 h until the end of the abnormal batch whereas the I^2 chart does not indicate any deviation after time 130 h. During the period shortly after the onset of the fault, the disturbance changes a systematic part of the process variation in the ICA domain; however, its effect is propagated and causes a change in the residual part of the process variation from 110 h onwards. Consequently,

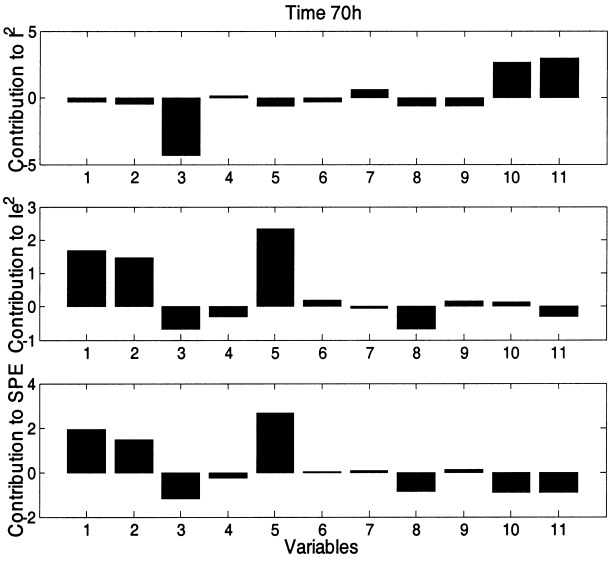


Fig. 7 Contribution plots to I^2, I_e^2 and *SPE* charts at time 70 h

compared with the on-line prediction of the MPCA model, the proposed method without estimating the future measurements yields better results. Although the unfolding method used in the proposed model differs from that used in MPCA, the monitoring results along time are not distorted because the time correlations of variables are attenuated by the scaling of variables at each time in advance. If the time correlations of the variables are very strong, it may be appropriate to use a separate model for each stage of a batch (Dong and McAvoy, 1996).

Once a fault or special event has been detected, it is important to diagnose the event to find an assignable cause. One aid to assigning the cause of a disturbance is to display the contribution of each measurement variable to the deviations observed in the monitoring metric. Such contribution (or diagnostic) charts can be displayed on-line by the operator immediately after a special event is detected. Although these charts may not provide an unequivocal diagnosis, they will at the very least clearly show the group of variables that are primarily responsible for the detected deviations. If, however, there is a lag between an event and its detection, diagnosis becomes more complicated because other variables correlated with the fault may deviate significantly from their expected values (Birol *et al.*, 2002b). Hence, the contribution plots should be analyzed as soon as the fault is detected. **Figure 7** shows the contribution plots for I^2, I_e^2, and *SPE* at 70 h. At 70 h, only the I^2 chart has successfully detected the out-of-control situation; hence, we can conclude that variables 10 (generated heat) and 11 (cooling water flow rate) as well as variable 3 (substrate feed rate) make the greatest contributions to the I^2 statistics. It can also be inferred that the deviation

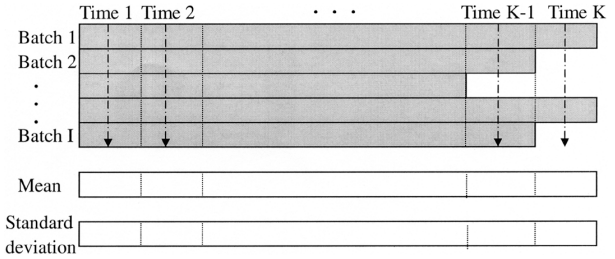


Fig. 8 Simple procedure for on-line monitoring before using the proposed method when batch lengths are different

of variable 3 affects other variables (variables 10 and 11).

From these contribution plots, an 'out of control' situation is identified when the contributions of some variables are larger than anticipated. Identification of the variables that have experienced the greatest change, in conjunction with the expert knowledge of the production engineer and operator, makes it possible to relate a particular sequence of changes to a particular process malfunction. This information provides sufficient information to allow operational personnel to narrow down the potential causes of the process problem.

8.3 Unequal batch lengths

MPCA-based methods require that the length of the modeling batches be the same; however, in many situations the total batch duration, or the durations of particular stages, may vary within a series of batches. To solve this problem, the batches must be aligned. In situations where batch durations vary, the most widely accepted alignment method is to mark the beginning and end of each batch by referring to a variable that grows monotonically during the batch and that has the same value at the end of all batches. A good example of such a variable is conversion (Neogi and Schlags, 1998). Another method for synchronizing batches is to use the dynamic time warping (DTW) approach proposed by Kassidas *et al*. (1998). However, DTW is cumbersome for modeling control charts and on-line monitoring. The approach proposed in the present work provides a simple method for handling batches with different durations; it avoids the need to equalize the batch length by using another unfolding method.

For illustration purposes, we use penicillin data with different batch lengths. The simple procedure is shown in **Fig. 8**. First, the data of each batch are arranged into a vector $(1 \times JK(i))$, where $K(i)$ is the time index of the i-th batch. Any of the longest batches is posed at the first row. Then, at each time the variables are mean centered and variance scaled, as mentioned earlier. After mean centering and variance scaling, each vector $(1 \times JK(i))$ is rearranged to a matrix of form $(J \times \sum_{i=1}^{I} K(i))$. The remainder of the procedure is similar to that described in the Appendix. In our example, a total of 67 batches were generated: 20 batches (time 400), 10 batches (time 399), 10 batches (time 398), 5 batches (time 397), 5 batches (time 396), 5 batches (time 395), 5 batches (time 394), 5 batches (time 393), and 2 batches (time 392). A test batch (time 395) containing a disturbance was also generated. Three independent components were chosen by the SCREE test for modeling the deterministic part of the proposed method. The disturbance was imposed by reducing the agitation power by 10% at 300 h and then leaving it at that lower power until the end of the batch. **Figures 9** and **10** show the monitoring results and contribution plots, respectively. The disturbance is detected very well by the I_e^2 and *SPE* charts from 300 h onwards.

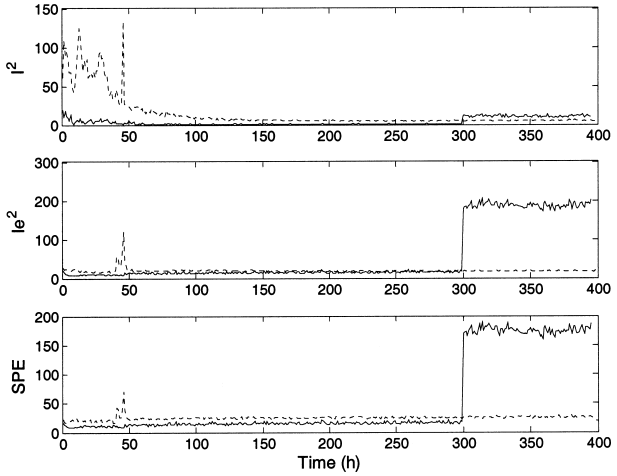


Fig. 9 I^2, I_e^2 and SPE charts for on-line monitoring using the proposed method when batch lengths are different

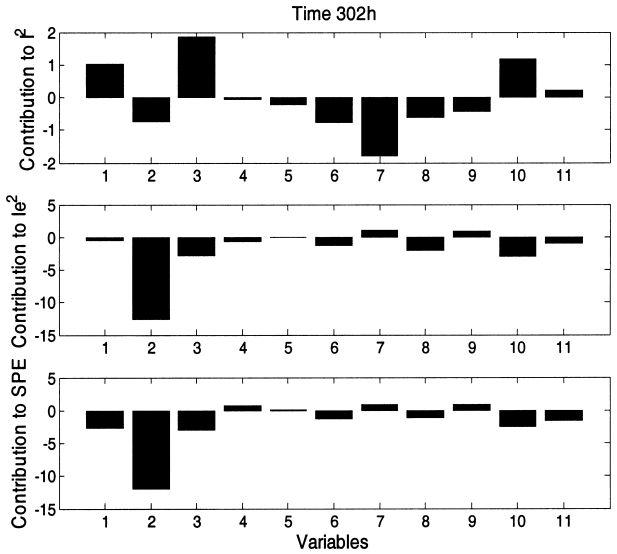


Fig. 10 Contribution plots to I^2, I_e^2 and *SPE* charts at time 302 h

The I_e^2 and *SPE* contribution plots at 302 h indicate that the second variable (agitation power) is the principal source of the large deviation. If the batch lengths are not made equal, MPCA monitoring is impossible. However, as shown in Fig. 9, the proposed method can be used to monitor batch processes in which the batch length varies. The monitoring performance achieved using the proposed method will be best if the batch lengths are similar and the time correlations are weak.

Conclusions

A new approach to monitoring the progress of batch processes has been suggested. This approach utilizes only the information contained in the historical database of past batches. The proposed method uses different unfolding method and ICA. ICA provides better monitoring results than PCA because it is based on the assumption that the latent variables are not normally distributed and because it can extract underlying independent factors from correlated data. In on-line monitoring of an evolving batch, the values of the variables from the current time until the end of the batch are unknown. When MPCA is used for on-line monitoring, it is necessary to complete the data sets with predictions of these future observations. However, anticipating the future observations might cause false alarms because the inclusion of values predicted without considering dynamic relationships may distort the data information. To handle this problem, we proposed the use of a different unfolding method. Under the proposed approach, Macgregor's unfolding method is first used to eliminate the batch process trajectory and to attenuate the time correlation and then, after scaling, the data matrix is rearranged to another form. The proposed monitoring method was applied to fault detection and identification in fed-batch penicillin fermentation. The control limits of the monitoring charts were obtained from kernel density estimation for normal batches at each time; hence they themselves involve process dynamicity, which compensates for any disadvantage that may arise from the assumption that the correlations among the variables remain constant throughout the entire batch. The simulation results clearly demonstrated the power and advantages of the proposed monitoring performance in comparison to MPCA though the dynamicity of the process is strong. In particular, the proposed method can be used to monitor batch processes in which the batch length varies. Although the unfolding method used in the proposed model differs from that used in MPCA, the monitoring results along time are not distorted because the time correlations of the variables are attenuated by subtracting the mean trajectory and scaling the variables at each time in advance, and because the control limits are established at each time. In processes whose variables show strong time correlations, more than one model needs to be derived to represent different phases of the batch.

Acknowledgment

This work was supported by a grant (No. R01-2002-000-00007-0) from Korea Science & Engineering Foundation. This work was also supported by the Brain Korea 21 project.

Appendix: The proposed on-line monitoring procedures

A. Develop normal operating condition (NOC) model

1. Unfold $\mathbf{X}$ ($I \times J \times K$) to $\mathbf{X}$ ($I \times JK$)
2. The data $\mathbf{X}$ ($I \times JK$) are normalized using the mean and standard deviation of each variable at each time in the batch cycle over all batches.
3. Rearrange the scaled $\mathbf{X}$ ($I \times JK$) to $\mathbf{X}_{normal}$ ($J \times IK$)
4. Carry out the whitening procedure

$$\mathbf{Z}_{normal} = \mathbf{Q}\mathbf{X}_{normal}$$

5. Carry out the ICA procedure
Obtain $\mathbf{W}$, $\mathbf{B}$, and $\hat{\mathbf{S}}_{normal}$ from $\hat{\mathbf{S}}_{normal} = \mathbf{W}\mathbf{X}_{normal} = \mathbf{B}^T\mathbf{Z}_{normal}$.
6. Calculate the norm of the row vectors of $\mathbf{W}$ and separate $\mathbf{W}$ into the deterministic part and the excluded part based on the magnitude of norms. $\mathbf{B}$ and $\hat{\mathbf{S}}_{normal}$ can be separated with the same criterion.

$$\mathbf{W} \rightarrow \mathbf{W}_d, \mathbf{W}_e$$

$$\mathbf{B} \rightarrow \mathbf{B}_d, \mathbf{B}_e$$

$$\hat{\mathbf{S}}_{normal} \rightarrow \hat{\mathbf{S}}_d, \hat{\mathbf{S}}_e$$

7. Calculate I^2, I_e^2, and *SPE* metrics.

$$I^2(n) = \hat{\mathbf{s}}_d(n)^T \hat{\mathbf{s}}_d(n)$$

$$I_e^2(n) = \hat{\mathbf{s}}_e(n)^T \hat{\mathbf{s}}_e(n)$$

$$SPE(n) = \sum_{j=1}^{d} \left(x_j(n) - \hat{x}_j(n) \right)^2$$

where n is a value from 1 to IK and $\hat{\mathbf{x}} = \mathbf{Q}^{-1}\mathbf{B}_d\hat{\mathbf{s}}_d = \mathbf{Q}^{-1}\mathbf{B}_d\mathbf{W}_d\mathbf{x}_{normal}$.
8. Rearrange I^2 ($1 \times JK$), I_e^2 ($1 \times JK$) and *SPE* ($1 \times JK$) to I^2 ($I \times K$), I_e^2 ($I \times K$) and *SPE* ($I \times K$), respectively.
9. Obtain control limits of I^2, I_e^2 and *SPE* metrics at each time using kernel density estimation.

B. On-line monitoring

1. For new batch data at time k, $\mathbf{x}_{new}(k)$, apply the same scaling used in modeling.
2. Calculate $\hat{\mathbf{s}}_{newd}(k)$ and $\hat{\mathbf{s}}_{newe}(k)$ from $\hat{\mathbf{s}}_{newd}(k) = \mathbf{W}_d\mathbf{x}_{new}(k)$, $\hat{\mathbf{s}}_{newe}(k) = \mathbf{W}_e\mathbf{x}_{new}(k)$.
3. Calculate $I^2(k)$, $I_e^2(k)$, and $SPE(k)$ for the new data

$$I^2(k) = \hat{\mathbf{s}}_{newd}(k)^T \hat{\mathbf{s}}_{newd}(k),\ I_e^2(k) = \hat{\mathbf{s}}_{newe}(k)^T \hat{\mathbf{s}}_{newe}(k)$$

$$SPE(k) = \sum_{j=1}^{d} \left(x_{new,j}(k) - \hat{x}_{new,j}(k) \right)^2$$

where $\hat{\mathbf{x}}_{new}(k) = \mathbf{Q}^{-1}\mathbf{B}_d\hat{\mathbf{s}}_{newd}(k) = \mathbf{Q}^{-1}\mathbf{B}_d\mathbf{W}_d\mathbf{x}_{new}(k)$.
4. Monitor whether $I^2(k)$, $I_e^2(k)$, or $SPE(k)$ exceeds its control limit calculated in the modeling procedure.

C. Contribution plot

1. Variable contribution to $I^2(k)$

$$\mathbf{x}_{cd}(k) = \frac{\mathbf{Q}^{-1}\mathbf{B}_d\hat{\mathbf{s}}_{newd}(k)}{\left\|\mathbf{Q}^{-1}\mathbf{B}_d\hat{\mathbf{s}}_{newd}(k)\right\|}\left\|\hat{\mathbf{s}}_{newd}(k)\right\|$$

2. Variable contribution for $I_e^2(k)$

$$\mathbf{x}_{ce}(k) = \frac{\mathbf{Q}^{-1}\mathbf{B}_e\hat{\mathbf{s}}_{newe}(k)}{\left\|\mathbf{Q}^{-1}\mathbf{B}_e\hat{\mathbf{s}}_{newe}(k)\right\|}\left\|\hat{\mathbf{s}}_{newe}(k)\right\|$$

3. Variable contribution for $SPE(k)$

$$\mathbf{x}_{cspe}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k)$$

Literature Cited

Albert, S. and R. D. Kinley; "Multivariate Statistical Monitoring of Batch Processes: An Industrial Case Study of Fermentation Supervision," *TRENDS in Biotechnolgy*, **19**, 53–62 (2001)

Back, A. D. and A. S. Weigend; "A First Application of Independent Component Analysis to Extracting Structure from Stock Returns," *International Journal of Neural Systems*, **8**, 473–484 (1997)

Birol, G., C. Ündey, S. J. Parulekar and A. Çinar; "A Morphologically Structured Model for Penicillin Production," *Biotechnology & Bioengineering*, **77**, 538–552 (2002a)

Birol, G., C. Ündey and A. Çinar; "A Modular Simulation Package for Fed-batch Fermentation: Penicillin Production," *Computers and Chemical Engineering*, **26**, 1553–1565 (2002b)

Cardoso, J.-F. and A. Soulomica; "Blind Beamforming for Non-Gaussian Signals," *IEE Proc. F.*, **140**, 362–370 (1993)

Chen, J. and K.-C. Liu; "On-line Batch Process Monitoring Using Dynamic PCA and Dynamic PLS Models," *Chemical Engineering Science*, **57**, 63–75 (2002)

Chen, Q., R. J. Wynne, P. Goulding and D. Sandoz; "The Application of Principal Component Analysis and Kernel Density Estimation to Enhance Process Monitoring," *Control Engineering Practice*, **8**, 531–543 (2000)

Cheung, Y. M. and L. Xu; "An Empirical Method to Select Dominant Independent Components in ICA Time Series Analysis," Proceedings of International Joint Conference on Neural Networks (IJCNN'99), pp. 3883–3887, Washington, D.C., USA (1999)

Cheung, Y. M. and L. Xu; "Independent Component Ordering in ICA Time Series Analysis," *Neurocomputing*, **41**, 145–152 (2001)

Dong, D. and T. J. McAvoy; "Batch Tracking via Nonlinear Principal Component Analysis," *AIChE J.*, **42**, 2199–2208 (1996)

Gallagher, N. B. and B. M. Wise; "Application of Multi-way Principal Components Analysis to Nuclear Waste Storage Tank Monitoring," *Computers and Chemical Engineering*, **20**, S739–S744 (1996)

Gregersen, L. and S. B. Jorgensen; "Supervision of Fed-Batch Fermentations," *Chem. Eng. J.*, **75**, 69–76 (1999)

Hyvärinen, A.; "Fast and Robust Fixed-point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks*, **10**, 626–634 (1999a)

Hyvärinen, A.; "Survey on Independent Component Analysis," *Neural Computing Surveys*, **2**, 94–128 (1999b)

Hyvärinen, A. and E. Oja; "Independent Component Analysis: Algorithms and Applications," *Neural Networks*, **13**, 411–430 (2000)

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

Kano, M., S. Tanaka, S. Hasebe, I. Hashimoto and H. Ohno; "Combination of Independent Component Analysis and Principal Component Analysis for Multivariate Statistical Process Control," Proceedings of International Symposium on Design, Operation and Control of Chemical Plants (PSE Asia 2002), pp. 319–324, Taipei, Taiwan (2002)

Kano, M., S. Tanaka, S. Hasebe, I. Hashimoto and H. Ohno; "Monitoring Independent Components for Fault Detection," *AIChE J.*, **49**, 969–976 (2003)

Kassidas, A., J. F. MacGregor and P. A. Htaylor; "Synchronization of Batch Trajectories Using Dynamic Time Warping," *AIChE J.*, **44**, 864–875 (1998)

Kosanovich, K. A., K. S. Dahl and M. J. Piovoso; "Improved Process Understanding Using Multiway Principal Component Analysis," *Ind. Eng. Chem. Res.*, **35**, 138–146 (1996)

Kourti, T.; "Multivariate Dynamic Data Modeling for Analysis and Statistical Process Control of Batch Processes, Start-ups and Grade Transitions," *J. Chemometrics*, **17**, 93–109 (2003)

Lee, J.-M., C. K. Yoo and I.-B. Lee; "Statistical Process Monitoring with Independent Component Analysis," *J. Process Control* (2003a) (in press)

Lee, J.-M., C. K. Yoo and I.-B. Lee; "Statistical Process Monitoring with Multivariate Exponentially Weighted Moving Average and Independent Component Analysis," *J. Chem. Eng. Japan*, 36, 563–577 (2003b)

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

Lennox, B., G. A. Montague, H. G. Hiden, G. Kornfeld and P. R. Goulding; "Process Monitoring of an Industrial Fed-Batch Fermentation," *Biotechnology & Bioengineering*, **74**, 125–135 (2001)

Li, R. F. and X. Z. Wang; "Dimension Reduction of Process Dynamic Trends Using Independent Component Analysis," *Computers and Chemical Engineering*, **26**, 467–473 (2002)

MacGregor, J. F., C. Jaeckle, C. Kiparissides and M. Koutoudi; "Process Monitoring and Diagnosis by Multiblock PLS Methods," *AIChE J.*, **40**, 826–838 (1994)

Martin, E. B., A. J. Morris, M. C. Papazoglou and C. Kiparissides; "Batch Process Monitoring for Consistent Production," *Computers and Chemical Engineering*, **20**, S599–S604 (1996)

Meng, X., A. J. Morris and E. B. Martin; "On-line Monitoring of Batch Processes Using a PARAFAC Representation," *J. Chemometrics*, **17**, 65–81 (2003)

Neogi, D. and C. E. Schlags; "Multivariate Statistical Analysis of an Emulsion Batch Process," *Ind. Eng. Chem. Res.*, **37**, 3971–3979 (1998)

Nomikos, P.; "Detection and Diagnosis of Abnormal Batch Operations Based on Multi-way Principal Component Analysis," *ISA Transactions*, **35**, 259–266 (1996)

Nomikos, P. and J. F. MacGregor; "Monitoring Batch Processes Using Multiway Principal Component Analysis," *AIChE J.*, **40**, 1361–1375 (1994)

Nomikos, P. and J. F. MacGregor; "Multivariate SPC Charts for Monitoring Batch Processes," *Technometrics*, **37**, 41–59 (1995a)

Nomikos, P. and J. F. MacGregor; "Multi-way Partial Least Square in Monitoring Batch Processes," *Chemo. Intell. Lab. Sys.*, **30**, 97–108 (1995b)

Rännar, S., J. F. MacGregor and S. Wold; "Adaptive Batch Monitoring Using Hierarchical PCA," *Chemo. Intell. Lab. Sys.*, **41**, 73–81 (1998)

Rosen, C.; Monitoring Wastewater Treatment Systems, M.S. Thesis, Lund Univ., Sweden (1998)

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

Tates, A. A., D. J. Louwerse, A. K. Smilde, G. L. M. Koot and H. Berndt; "Monitoring a PVC Batch Process with Multivariate Statistical Process Control Charts," *Ind. Eng. Chem. Res.*, **38**, 4769–4776 (1999)

Teppola, P.; Multivariate Process Monitoring of Sequential Process Data—A Chemometric Approach, Ph.D. Thesis, Lappeenranta Univ. of Tech., Finland (1999)

Ündey, C., E. Tatara, B. A. Williams, G. Birol and A. Çinar; "Online Real Time Monitoring of Penicillin Fermentation," International Symposium on advanced Control of Chemical Processes, ADCHEM2000, pp. 243–247, 14–16, Pisa, Italy (2000)

Vigário, R. N.; "Extraction of Ocular Artifacts from EEG using Independent Component Analysis," *Electroencephalography and Clinical Neurophysiology*, **103**, 395–404 (1997)

Wand, M. P. and M. C. Jones; Kernel Smoothing, Chapman&Hall, London, UK (1995)

Westerhuis, J. A., T. Kourti and J. F. Macgregor; "Comparing Alternative Approaches for Multivariate Statistical Analysis of Batch Process Data," *J. Chemometrics*, **13**, 397–413 (1999)

Westerhuis, J. A., S. P. Gurden and A. K. Smilde; "Generalized Contribution Plots in Multivariate Statistical Process Monitoring," *Chemo. Intell. Lab. Sys.*, **51**, 95–114 (2000)

Wise, B. M., N. B. Gallagher, S. W. Butler, D. D. White, Jr. and G. G. Barna; "A Comparison of Principal Component Analysis, Multiway Principal Component Analysis, Trilinear Decomposition and Parallel Factor analysis for Fault Detection in a Semiconductor Etch Process," *J. Chemometrics*, **13**, 379–396 (1999)

Wold, S., N. Kettanhe, H. Friden and A. Holmberg; "Modelling and Diagnostics of Batch Processes and Analogous Kinetic Experiments," *Chemo. Intell. Lab. Sys.*, **44**, 331–340 (1998)