AIChE

Multimode Process Monitoring with Bayesian Inference-Based Finite Gaussian Mixture Models

Jie Yu
Dept. of Chemical Engineering, The University of Texas at Austin, Austin, TX 78712

S. Joe Qin
The Mork Family Dept. of Chemical Engineering and Materials Science, Ming Hsieh Dept. of Electrical Engineering, and Daniel J. Epstein Dept. of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089

DOI 10.1002/aic.11515
Published online May 9, 2008 in Wiley InterScience (www.interscience.wiley.com).

For complex industrial processes with multiple operating conditions, the traditional multivariate process monitoring techniques such as principal component analysis (PCA) and partial least squares (PLS) are ill-suited because the fundamental assumption that the operating data follow a unimodal Gaussian distribution usually becomes invalid. In this article, a novel multimode process monitoring approach based on finite Gaussian mixture model (FGMM) and Bayesian inference strategy is proposed. First, the process data are assumed to be from a number of different clusters, each of which corresponds to an operating mode and can be characterized by a Gaussian component. In the absence of a priori process knowledge, the Figueiredo–Jain (F–J) algorithm is then adopted to automatically optimize the number of Gaussian components and estimate their statistical distribution parameters. With the obtained FGMM, a Bayesian inference strategy is further utilized to compute the posterior probabilities of each monitored sample belonging to the multiple components and derive an integrated global probabilistic index for fault detection of multimode processes. The validity and effectiveness of the proposed monitoring approach are illustrated through three examples: (1) a simple multivariate linear system, (2) a simulated continuous stirred tank heater (CSTH) process, and (3) the Tennessee Eastman challenge problem. The comparison of monitoring results demonstrates that the proposed approach is superior to the conventional PCA method and can achieve accurate and early detection of various types of faults in multimode processes.
AIChE J, 54: 1811–1829, 2008
Keywords: multimode process monitoring, fault detection, finite Gaussian mixture model, Bayesian inference, Mahalanobis distance, global probabilistic index, Tennessee Eastman chemical process

Correspondence concerning this article should be addressed to J. Yu at this current address: Shell Global Solutions (US) Inc., Houston, TX 77082; j.yu@shell.com.

Introduction

Over the past two decades, chemical process monitoring has gained tremendous attention in both academia and industry.[1–3] As modern industrial processes become more and

more complex, automatic detection and identification of process and instrument faults are necessary to ensure safe operation, stable product quality, and sustainable profit in chemical plants.[4] Typically, thousands of regulatory loops, sensors, and analyzers are present in a plant and there may exist different kinds of interactions among such a large number of variables.[5,6] Therefore, it is really challenging to monitor the large-scale chemical processes and detect various types of faults accurately. Nowadays, process historian databases are commonly implemented in industrial facilities to provide plenty of operating data for the detection and identification of abnormal conditions.[7,8] Reliable computation methods are thus crucial to accomplish the successful detection of process faults without disturbing the routine operation.

Traditional monitoring methods have been based on mechanistic process models, which are difficult and time-consuming to build.[9–13] As a class of alternatives, multivariate statistical techniques have been widely applied to process monitoring area and shown attractive merits over the model-based methods. This category of approaches, termed as multivariate statistical process monitoring (MSPM), utilize the historical data to extract variable correlations, reduce space dimensionality, and derive normal operating regions. Process faults can then be detected and diagnosed using the induced statistical metrics.[3,14] Two major MSPM methods, principal component analysis (PCA) and partial least squares (PLS), have been intensively investigated and applied to diverse processes.[2,15–19] Both methods project data onto lower-dimensional subspaces and use the Hotelling's T^2 or squared predicted error (SPE) indices to isolate the normal and abnormal conditions.[20,21] Other complementary MSPM approaches, including Fisher discriminant analysis (FDA), canonical variate analysis (CVA), independent components analysis (ICA), have been used to overcome some limitations in PCA/PLS-based monitoring schemes.[22–28] Meanwhile, machine learning techniques, e.g., discriminant analysis (DA), neural network (NN), expert systems, support vector machines (SVM), Bayesian belief network (BBN), and mutual information, have been explored to address the complex process monitoring problems with some success.[29–36]

Despite a rich body of literature in chemical process monitoring, most of the MSPM approaches rely on the assumption that the normal process data come from a single operating region and follow a unimodal distribution. The popular PCA/PLS methods, for example, require the operating data to obey unimodal Gaussian distribution approximately so as to guarantee the valid T^2 and SPE control limits. In industrial processes, however, operating condition shifts are often encountered due to the changes of various factors such as feedstock, product specification, set points, and manufacturing strategy.[37] When a process is running under substantially different operating conditions, mean and covariance changes can be significant in normal operating data. As a result, the multimodality of data distribution would render the conventional MSPM techniques inappropriate.

In literature, some research effort has been reported to approach the multimode process monitoring issue through the modified PCA/PLS methods. Hwang and Han proposed a hierarchical clustering-based super PCA model for real-time monitoring of chemical processes under multiple operating modes.[37] Nevertheless, this solution is applicable only if different operating modes share the common covariance structures. That is, each hyper-ellipsoid corresponding to a mode should possess similar size and orientation. This assumption restricts its application in many cases that experience a wide range of covariance structures. Lane et al. employed a common subspace model to monitor multiproduct semibatch processes.[38] An identical eigenvector subspace is assumed for the covariance of each individual product. Zhao et al. developed multiple PCA/PLS models for multimode process monitoring.[39,40] In the preliminary step, however, a priori process knowledge is required to manually segment the historical operating data into multiple groups that correspond to different operating modes. Moreover, a similarity threshold has to be predefined by user to incorporate the similar data groups. Those conditions are not desirable for automatic process monitoring in industrial practice. Chen and Liu proposed a mixture PCA model to deal with the processes that are distributed in several operating regions.[41] A heuristic smoothing clustering (HSC) technique is adopted to automatically determine the number of clusters and fault detection is then performed on the basis of PCA model in each cluster. The well-known soft independent modeling of class analogy (SIMCA) method similarly identifies global and local PCA models to predict a probable class membership for new observations.[42]

In contrast to PCA/PLS-based approaches, Gaussian mixture model (GMM) has not been explored in multimode process monitoring until very recently. Choi et al. integrated GMM with PCA and DA to detect and isolate faults in processes with nonlinearity, multistates or dynamics.[43] Based on a membership measure, each monitored sample is categorized into a specific Gaussian cluster deterministically and the corresponding local PCA model is then selected for fault detection. Yoo et al. combined a similar strategy with multiway PCA to monitor biological batch processes.[44] However, these methods ignore the possibility that the monitored sample may come from other Gaussian components of lower posterior probabilities, which may lead to biased monitoring results. To overcome this deficiency, Thissen et al. defined a density of mixture modeling (DMM) index for the estimation of overall confidence boundary around multiple clusters.[45] A difficulty of this method, however, lies in its numerical procedure to derive the DMM limit. Monte Carlo simulation-based random sampling method is usually needed to obtain the boundary density value[46] and the intensive computation load is undesirable in industrial implementations. Furthermore, the above mixture models were all estimated via the expectation maximization (EM) algorithm, where a whole set of candidate models have to be built so as to find the optimal number of components.

To address the current issues in multimode process monitoring, a finite Gaussian mixture model (FGMM) and Bayesian inference-based probabilistic approach is developed for fault detection under various modes in this study. The normal process data at each individual operating mode are assumed to follow a multivariate Gaussian distribution and the operation status is treated as a random variable that can be at any of the possible modes with a prior probability. A FGMM is

then constructed to characterize the multiple operating regions, each of which corresponds to a Gaussian component. In this work, the Figueiredo–Jain (F–J) algorithm[47] is adopted to train FGMM and the optimal number of modes is adaptively determined from the data. In addition, the estimated means, covariances, and prior probabilities well describe the multiple operating regions under normal conditions. Subsequently, a Bayesian inference strategy is used to compute the posterior probability of a monitored sample belonging to each Gaussian component. Thus, the Mahalanobis distance metrics for all the Gaussian hyper-ellipsoids can be integrated through the posterior probabilities and a new Bayesian inference-based probability (BIP) index is further defined to detect process faults around different operating regions.

The remainder of the article is organized as follows. Conventional PCA-based process monitoring method is briefly revisited. Then we describe the finite Gaussian mixture model trained by the F–J algorithm to characterize the multimode operating data. The novel process monitoring scheme on the basis of FGMM and Bayesian inference is further developed. The validity and effectiveness of the alternative multimode process monitoring approach are demonstrated through three application examples: an illustrative multivariate system, a simulated continuous stirred tank heater (CSTH) process, and the challenging Tennessee Eastman Process (TEP). Finally, the conclusions and future directions are outlined at the end of the article.

PCA-Based Process Monitoring

PCA serves as the most fundamental MSPM method and has been successfully applied to monitor a wide range of chemical processes. Let $x \in R^m$ represents a m-dimensional sample vector and $X \in R^{n\times m}$ be the normalized data matrix with n samples. The PCA loading matrix P can be obtained by eigenvalue decomposition on the covariance matrix cov(x) as follows:

$$\mathrm{cov}(x) \approx \frac{1}{n-1} X^{\mathrm{T}} X = P \Lambda P^{\mathrm{T}} \tag{1}$$

where $\Lambda = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ denotes the eigenvalue matrix and $P = \left[\hat{P}\tilde{P}\right]$ contains the loading matrices of principal component subspace (PCS) and residual subspace (RS), respectively. Assuming that the first l principal components are retained in the PCS with $\Lambda_l = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_l\}$, two types of statistical indices, SPE and T^2, can be used for process monitoring.[3,48] The SPE index is defined as

$$\mathrm{SPE} = \left\| (I - \hat{P}\hat{P}^{\mathrm{T}}) x \right\|^2 \tag{2}$$

which is capable of detecting the variations in the RS. The SPE control limit under $100(1-\alpha)\%$ confidence level is given by

$$\delta_\alpha^2 = g\chi_{h,\alpha}^2 \tag{3}$$

where $g = \sum_{j=l+1}^{m} \lambda_j^2 / \sum_{j=l+1}^{m} \lambda_j$ and $h = \left(\sum_{j=l+1}^{m} \lambda_j\right)^2 / \sum_{j=l+1}^{m} \lambda_j^2$.

On the other hand, the Hotelling's T^2 index is designed to measure the variability within the PCS as follows

$$T^2 = x^{\mathrm{T}} \hat{P} \Lambda_l^{-1} \hat{P}^{\mathrm{T}} x \tag{4}$$

For large number of samples, the corresponding T^2 index follows an approximate χ^2 distribution with l degrees of freedom. Hence, the $100(1-\alpha)\%$ T^2 control limit is given by

$$T_\alpha^2 = \chi_{l,\alpha}^2 \tag{5}$$

It needs to be stressed that the above results are valid only if the process data follow a multivariate Gaussian distribution approximately.

Finite Gaussian Mixture Model

For the processes running at multiple operating conditions, the assumption of multivariate Gaussian distribution becomes invalid owing to the mean shifts and/or covariance changes. In this situation, however, the local Gaussian distribution is still appropriate to characterize each subset of measurement data from the same operating conditions. Therefore, the finite Gaussian mixture model[49–51] is well suited to representing the data sources driven by different operating modes. Within the mixture model, each Gaussian component corresponds to an individual operating mode and the number of components should be equal to the number of all possible modes through the entire normal operation.

Let $x \in R^m$ be a m-dimensional sample from a multimode process. Its probability density function can be formulated as FGMM below:

$$p(x|\theta) = \sum_{i=1}^{K} \omega_i g(x|\theta_i) \tag{6}$$

where K is the number of Gaussian components included in FGMM, ω_i denotes the weight of the ith component C_i, $\theta_i = \{\mu_i, \Sigma_i\}$ and $\theta = \{\theta_1, \ldots, \theta_K\} = \{\mu_1, \Sigma_1, \ldots, \mu_K, \Sigma_K\}$ represent the sets of local and global Gaussian model parameters (i.e., mean vector μ_i and covariance matrix Σ_i), respectively. The corresponding multivariate Gaussian density function for the ith component C_i is given by

$$g(x|\theta_i) = \frac{1}{(2\pi)^{m/2}|\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2}(x-\mu_i)^{\mathrm{T}} \Sigma_i^{-1} (x-\mu_i)\right] \tag{7}$$

The mixture density function $p(x|\theta)$ is essentially a weighted sum of multiple local Gaussian components, and the weight ω_i can be interpreted as a prior probability that an arbitrary sample point comes from the ith Gaussian component. For the mixture model, its cumulative density function satisfies

$$\int_{R^m} p(x|\theta)dx = \sum_{i=1}^{K} \omega_i \left(\int_{R^m} g(x|\theta_i)dx \right) = 1 \tag{8}$$

Considering the fact that $\int_{R^m} g(x|\theta_i)dx = 1$ holds for each Gaussian component, we can obtain the following identity

$$\sum_{i=1}^{K} \omega_i = 1 \tag{9}$$

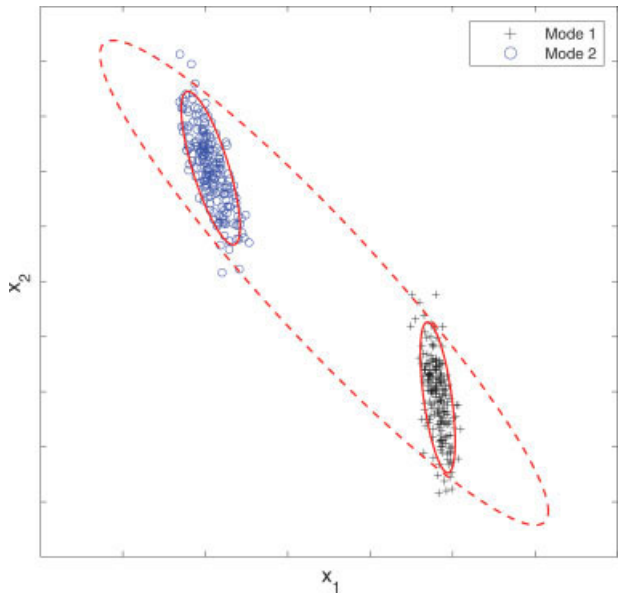

Figure 1. Illustration of multimode operating data along with the estimated local (solid line) and global (dotted line) probability ellipses.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

with $0 \leq \omega_i \leq 1$ for the prior probability. Then the population mean of the operating data from multiple modes is given by

$$\mu_x = \int_{R^m} xp(x|\theta)dx = \sum_{i=1}^{K} \omega_i \left(\int_{R^m} xg(x|\theta_i)dx \right) = \sum_{i=1}^{K} \omega_i \mu_i \tag{10}$$

which indicates that the mixture mean is the convex combination of local Gaussian component means. However, the mixture covariance does not possess such an explicit relationship with the individual component covariances. A set of illustrative data generated in two different operating regions are plotted in Figure 1. The two solid-line ellipses represent the 95% confidence boundaries of the two subsets of data from different modes, whereas the single dashed-line ellipse corresponds to the global confidence boundary of the full set of mixed data. It is obvious that the global probability ellipse encloses the two local ellipses and occupies an even bigger region than the sum of two local ones. The illustration implies that the volume of the global covariance hyper-ellipsoid is inflated and its orientation is rotated relative to the local covariances.

To construct a FGMM, the following list of unknown parameters

$$\Theta = \{\{\omega_1, \mu_1, \Sigma_1\}, \ldots, \{\omega_K, \mu_K, \Sigma_K\}\}$$

need to be estimated. It is noted that Θ involves both the Gaussian model parameters θ and the prior probabilities ω_i ($1 \leq i \leq K$). Because μ_i and Σ_i are $m \times 1$ vector and $m \times m$ matrix, respectively, the total number of scalar parameters to be estimated is $K(\frac{1}{2}m^2 + \frac{3}{2}m + 1) - 1$. A number of learning methods, including maximum likelihood estimation (MLE), EM, and the F–J algorithm, have been proposed for mixture model estimation.[49–51] With a set of training samples $X = \{x_1, x_2, \ldots, x_n\}$, the log-likelihood function can be expressed as

$$\log L(X, \Theta) = \sum_{j=1}^{n} \log \left(\sum_{i=1}^{K} \omega_i g(x_j|\theta_i) \right) \tag{11}$$

and the parameter estimation problem is further formulated as

$$\hat{\Theta} = \arg\max_{\Theta} (\log L(X, \Theta)) \tag{12}$$

The MLE method estimates the parameters Θ analytically by setting the first-order derivatives of the log-likelihood function to zero. A fundamental issue arises that the analytical solution may result in singular or ill-conditioned estimates of Gaussian model parameters in some cases. As a more tractable numerical strategy, the EM algorithm has been extensively used in practice to estimate the maximum likelihood distribution parameters. It is implemented iteratively by repeating the expectation step (E-step) and maximization step (M-step) to compute the posterior probabilities and then the corresponding distribution parameters until a convergence criterion of the log-likelihood function is satisfied. Given the training data X and an initial estimate $\Theta^{(0)} = \left\{ \{\omega_1^{(0)}, \mu_1^{(0)}, \Sigma_1^{(0)}\}, \ldots, \{\omega_K^{(0)}, \mu_K^{(0)}, \Sigma_K^{(0)}\} \right\}$, the iterative E-Step and M-Step are performed as follows:

- E-step:

$$P^{(s)}(C_k|x_j) = \frac{\omega_k^{(s)} g(x_j|\mu_k^{(s)}, \Sigma_k^{(s)})}{\sum_{i=1}^{K} \omega_i^{(s)} g(x_j|\mu_i^{(s)}, \Sigma_i^{(s)})} \tag{13}$$

where $P^{(s)}(C_k|x_j)$ denotes the posterior probability of the jth training sample within the kth Gaussian component at the sth iteration.

- M-step:

$$\mu_k^{(s+1)} = \frac{\sum_{j=1}^{n} P^{(s)}(C_k|x_j) x_j}{\sum_{j=1}^{n} P^{(s)}(C_k|x_j)} \tag{14}$$

$$\Sigma_k^{(s+1)} = \frac{\sum_{j=1}^{n} P^{(s)}(C_k|x_j)(x_j - \mu_k^{(s+1)})(x_j - \mu_k^{(s+1)})^{\mathrm{T}}}{\sum_{j=1}^{n} P^{(s)}(C_k|x_j)} \tag{15}$$

$$\omega_k^{(s+1)} = \frac{\sum_{j=1}^{n} P^{(s)}(C_k|x_j)}{n} \tag{16}$$

where $\mu_k^{(s+1)}$, $\Sigma_k^{(s+1)}$, and $\omega_k^{(s+1)}$ are the mean, covariance, and prior probability of the kth Gaussian component at the $(s+1)$th iteration, respectively.

A major limitation of the basic EM algorithm, however, remains that the number of Gaussian components has to be prespecified by user and cannot be adjusted automatically

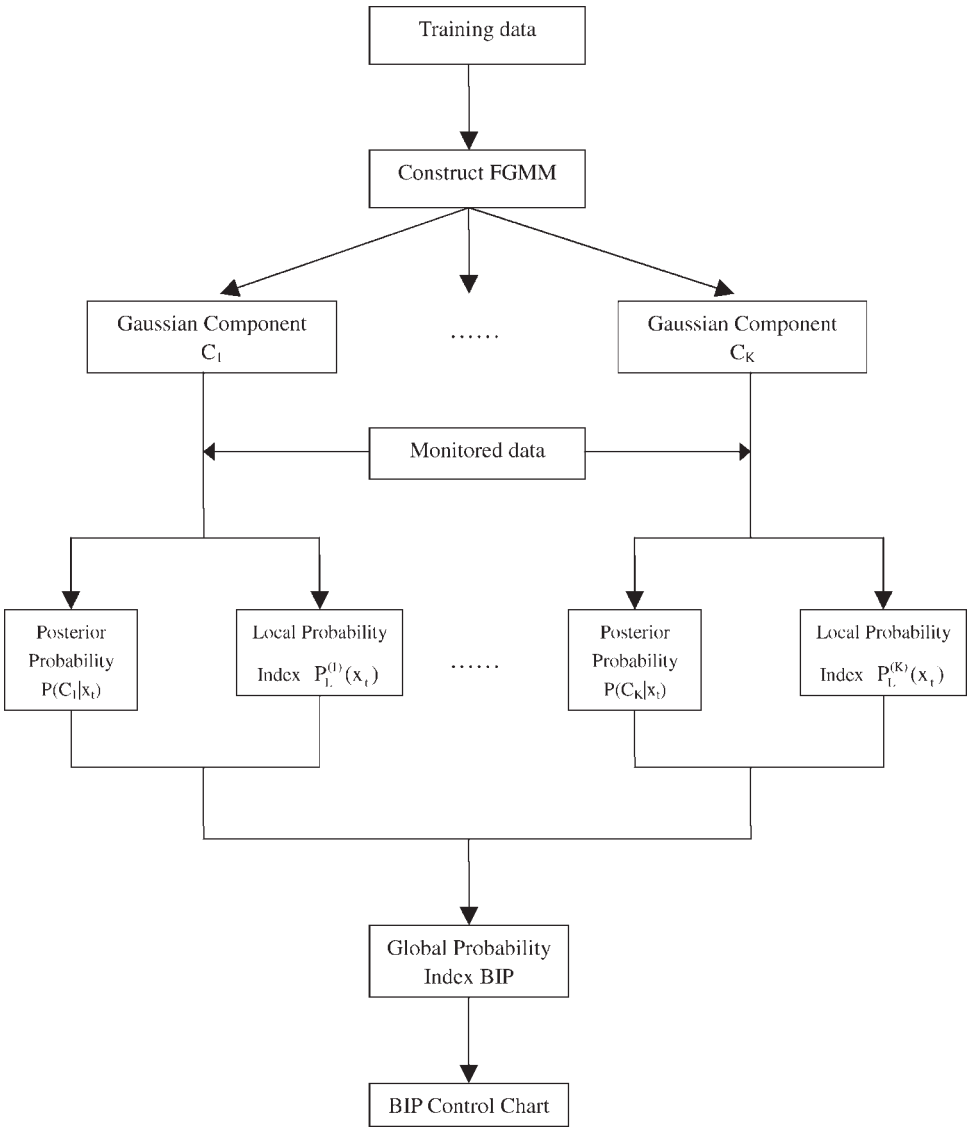

Figure 2. Schematic diagram of the proposed FGMM and Bayesian inference-based process monitoring approach.

during the parameter estimation. To overcome this drawback, the F–J algorithm is employed which starts with an arbitrarily large number of components and adjusts this number adaptively by eliminating the components of insignificant weights.[47,51] The minimum message length (MML) criterion is adopted to lead to the modified objective function as follows

$$M(X,\Theta) = \frac{V}{2}\sum_{\substack{1\le i\le K\\ \omega_i>0}} \log(n\omega_i) + \frac{K_{\mathrm{nz}}}{2}\left(\log\frac{n}{12}+1\right) - \log L(X,\Theta) \qquad (17)$$

where $V = \frac{1}{2}m^2 + \frac{3}{2}m$ denotes the number of scalar parameters specifying each Gaussian component and K_{nz} is the number of effective components with non-zero weights. The EM procedure can be utilized to minimize the objective function in Eq. 17 through the enhanced weighting innovation in the M-step:

$$\omega_k^{(s+1)} = \frac{\max\left\{0, \left(\sum_{j=1}^{n} P^{(s)}(C_k|x_j)\right) - \frac{V}{2}\right\}}{\sum_{i=1}^{K}\max\left\{0, \left(\sum_{j=1}^{n} P^{(s)}(C_i|x_j)\right) - \frac{V}{2}\right\}} \qquad (18)$$

which annihilates the insignificant components by setting the weighting to zero and thus adjust the number of effective components iteratively.

Multimode Process Monitoring Based on FGMM and Bayesian Inference

With the constructed FGMM, it is necessary to further derive the confidence boundary around the normal operating

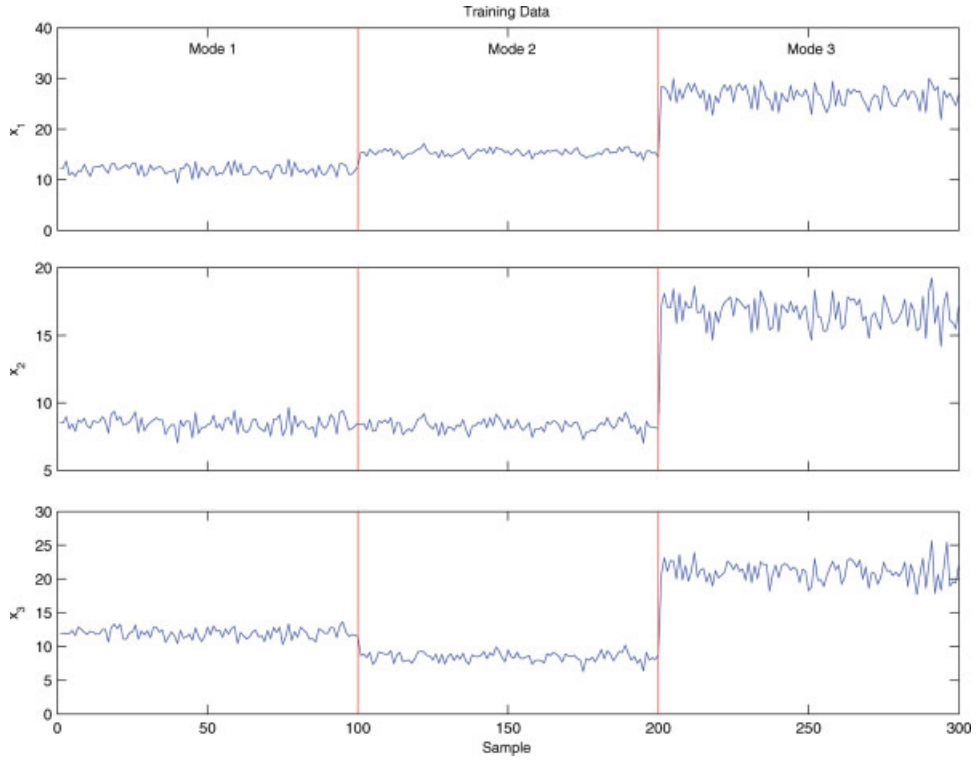

Figure 3. Illustrative example: Training data from three operating modes.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

regions for process monitoring and fault detection. This is similar to the T^2 and SPE confidence limits in PCA-based monitoring scheme. Owing to the multimodality of mixture distribution, it is extremely difficult to acquire the analytical boundary of the density function $p(x|\Theta)$ under a certain confidence level. Monte Carlo simulation has been used to obtain the approximate quantile of probability density, which actually delimits the normal and abnormal data.[46] However, this method is not appealing for large-scale industrial process monitoring because of its expensive computation load.

In the proposed monitoring approach, Bayesian inference strategy is utilized to compute the posterior probability of an arbitrary monitored sample x_t belonging to each Gaussian component as follows:

$$P(x_t \in C_k) = P(C_k|x_t) = \frac{P(C_k) \cdot p(x_t|C_k)}{p(x_t)} = \frac{P(C_k) \cdot p(x_t|C_k)}{\sum_{i=1}^{K} P(C_i) \cdot p(x_t|C_i)} \quad (19)$$

which can be rewritten as

$$P(x_t \in C_k) = \frac{\omega_k g(x_t|\mu_k, \Sigma_k)}{\sum_{i=1}^{K} \omega_i g(x_t|\mu_i, \Sigma_i)} \qquad (k = 1, 2, ..., K) \quad (20)$$

where the denominator $\sum_{i=1}^{K} \omega_i g(x_t|\mu_i, \Sigma_i)$ serves as a scaling factor such that the posterior probability satisfies $\sum_{i=1}^{K} P(x_t \in C_k) = 1$. It is noticed that the computational procedure of posterior probabilities for monitored samples is the same as the expectation step of EM algorithm in Eq. 13.

Considering that each component C_k follows a unimodal Gaussian distribution, the squared Mahalanobis distance of x_t from the center of C_k follows χ^2 distribution provided that x_i belongs to C_k. That is

$$D((x_t, C_k)|x_t \in C_k) = (x_t - \mu_k)^{\mathrm{T}} \Sigma_k^{-1} (x_t - \mu_k) : \chi_m^2 \quad (21)$$

where $D((x_t, C_k)|x_t \in C_k)$ denotes the squared Mahalanobis distance between x_t and the mean center of C_k under the assumption $x_t \in C_k$, and χ_m^2 has m degrees of freedom. When Σ_k is ill-conditioned due to colinearity, the following regularized Mahalanobis distance is utilized instead to avoid too wide confidence regions

$$D_{\mathrm{r}}((x_t, C_k)|x_t \in C_k) = (x_t - \mu_k)^{\mathrm{T}} (\Sigma_k + \varepsilon I)^{-1} (x_t - \mu_k) \quad (22)$$

where the role of ε is to remove the ill condition of covariance matrix Σ_k by adding a positive number to all the diagonal entries.[52] A large ε value tends to convert the Mahalanobis distance to Euclidean distance and the elliptical confidence boundary will become circular. Therefore, a small value of ε is recommended to restrict the bias between the regularized and the actual confidence boundaries. It is noted that a careful selection of ε value is not required as long as it is significantly smaller than the average variance of the data

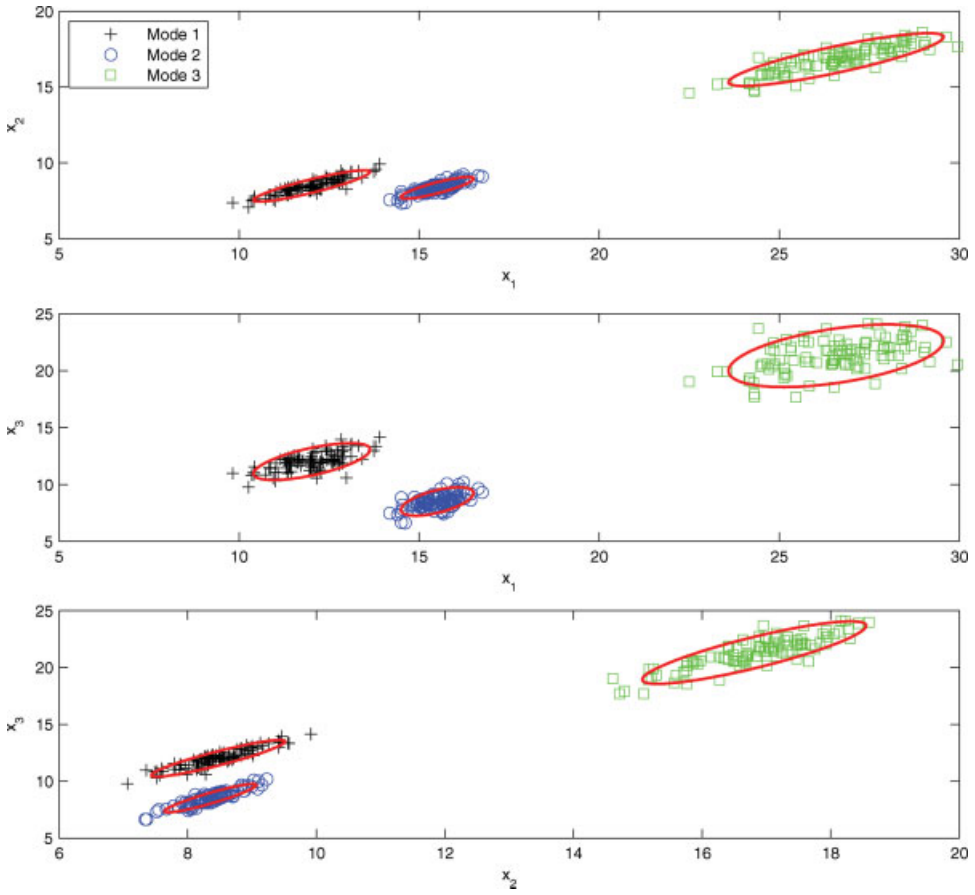

Figure 4. Illustrative example: Scatter plots of training data along with the estimated local probability ellipses.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

samples.[53] Since $(\Sigma_k + \varepsilon I)^{-1}$ is symmetric and positive definite, the regularized distance metric D_r approximately obeys a χ^2 distribution as follows[3,54]

$$D_r : g\chi_h^2 \tag{23}$$

with

$$g = \frac{\mathrm{tr}(\Sigma_k(\Sigma_k+\varepsilon I)^{-1})^2}{\mathrm{tr}(\Sigma_k(\Sigma_k+\varepsilon I)^{-1})} \tag{24}$$

and

$$h = \frac{\left[\mathrm{tr}(\Sigma_k(\Sigma_k+\varepsilon I)^{-1})\right]^2}{\mathrm{tr}(\Sigma_k(\Sigma_k+\varepsilon I)^{-1})^2} \tag{25}$$

For the monitored sample x_t, define a local Mahalanobis distance-based probability index relative to each Gaussian component C_k as

$$P_L^{(k)}(x_t) = \Pr\{D((x,C_k)|x \in C_k) \le D((x_t,C_k)|x_t \in C_k)\} \tag{26}$$

or

$$P_L^{(k)}(x_t) = \Pr\{D_r((x,C_k)|x \in C_k) \le D_r((x_t,C_k)|x_t \in C_k)\} \tag{27}$$

which can be calculated by integrating the χ^2 probability density function with appropriate degree of freedom. Under a given confidence level, this index serves as an indication of whether the monitored sample is normal or faulty provided that it belongs to the corresponding Gaussian component. Considering the random characteristic that each monitored sample may come from multiple Gaussian components with the corresponding posterior probabilities, a global BIP index is further defined to combine the local probability metrics across all the Gaussian clusters. The formulation of BIP index for the monitored sample x_t is given by

$$\mathrm{BIP} = \sum_{k=1}^{K} P(C_k|x_t) P_L^{(k)}(x_t) \tag{28}$$

where the posterior probability $P(C_k|x_t)$ is used to incorporate the contribution of each local Gaussian components to the overall probabilistic index. Without classifying the monitored sample into a single cluster deterministically, such a probabilistic strategy can avoid the potential risk of false detection induced by misclassification. As $0 \le P_L^{(k)}(x_t) \le 1$, we have

$$0 \le \mathrm{BIP} \le \sum_{k=1}^{K} P(C_k|x_t) = 1 \tag{29}$$

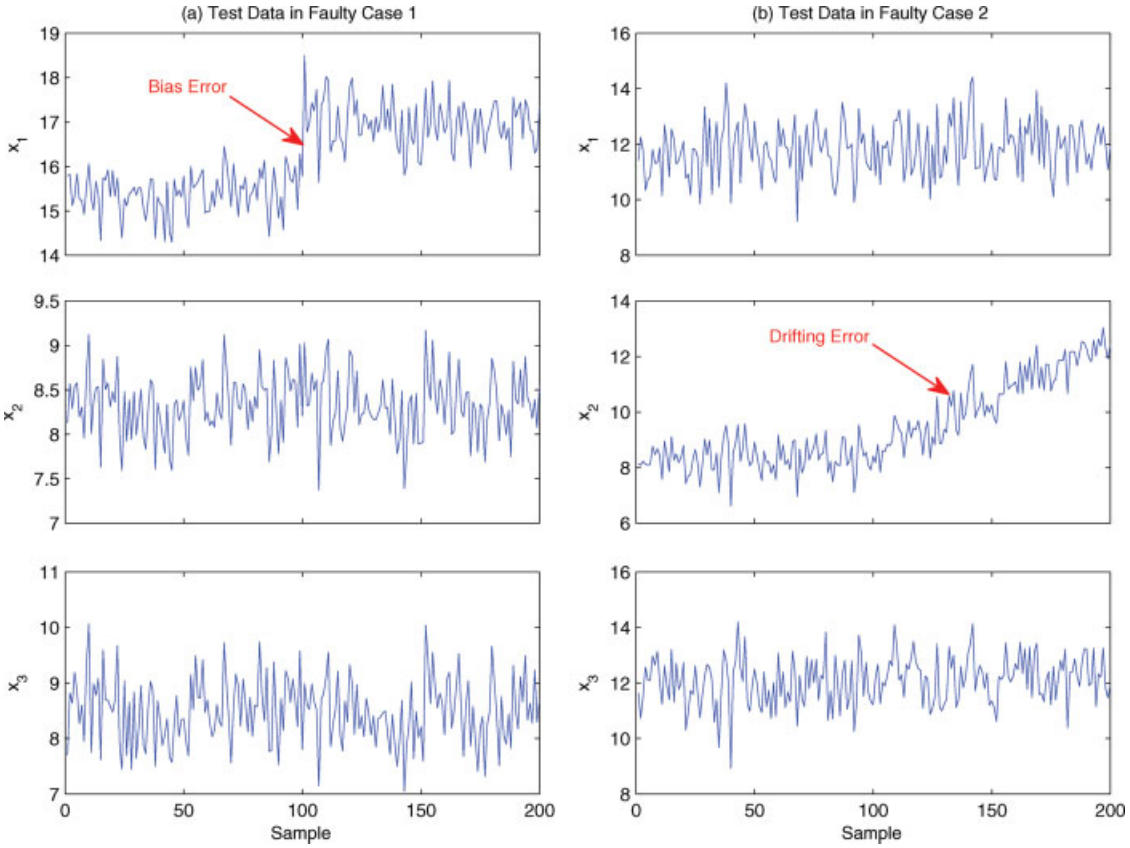

Figure 5. Illustrative example: Test data in faulty Case 1 (a) and 2 (b), respectively.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Under a prespecified confidence level $(1-\alpha)100\%$, the process is determined within normal operation if

$$\text{BIP} \leq 1-\alpha \tag{30}$$

Otherwise, the process operation is treated out of control. The BIP index-based control chart can be generated in a similar way as SPE and T^2 control charts with a constant control limit at the confidence level $1-\alpha$.

The schematic diagram of the proposed approach is shown in Figure 2 and the detailed procedures for multimode process monitoring are summarized below:

1. Collect a set of historical training data under all possible operating conditions.

2. Use the F–J algorithm to learn the Gaussian mixture model and estimate the model parameter set $\Theta = \{\mu_1, \Sigma_1, \omega_1, ..., \mu_K, \Sigma_K, \omega_K\}$ based on the iterative steps in Eqs. 13–15 and 18.

3. For each monitored sample x_t, compute its posterior probabilities belonging to all Gaussian components $P(C_k|x_t)(k = 1, 2, ..., K)$ through Bayesian inference strategy given in Eq. 20.

4. Calculate the local Mahalanobis distance-based probability index $P_{\mathrm{L}}^{(k)}(x_t)$ for the monitored sample x_t within each Gaussian component C_k by Eqs. 26 or 27.

5. Integrate the local probability indices relative to all the Gaussian components into the global BIP index via Eq. 28.

6. Specify a confidence level $(1-\alpha)100\%$, and generate the BIP control chart with the calculated BIP index values for all the monitored samples and the constant control limit $1-\alpha$.

7. Detect the abnormal operating conditions at the monitored samples satisfying BIP$>1-\alpha$.

Application Examples

An illustrative multimode example. A simple illustrative example is used to demonstrate the usage of the proposed multimode monitoring approach. The simulation data are generated from the following multivariate linear system

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3723 & 0.6815 \\ 0.4890 & 0.2954 \\ 0.9842 & 0.1793 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \tag{31}$$

where $[s_1\ s_2]^{\mathrm{T}}$ denote Gaussian distributed data sources and $[e_1\ e_2\ e_3]^{\mathrm{T}}$ are zero-mean white noises with standard deviations of $0.01I$. Three sets of data sources are simulated to represent different operating modes as follows:

$$\text{Mode 1}: s_1 : N(10, 0.8); s_2 : N(12, 1.3)$$

$$\text{Mode 2}: s_1 : N(5, 0.6); s_2 : N(20, 0.7)$$

$$\text{Mode 3}: s_1 : N(16, 1.5); s_2 : N(30, 2.5)$$

where the means and standard deviations of s_1 and s_2 are changed to reflect the shifts of operating regions. The

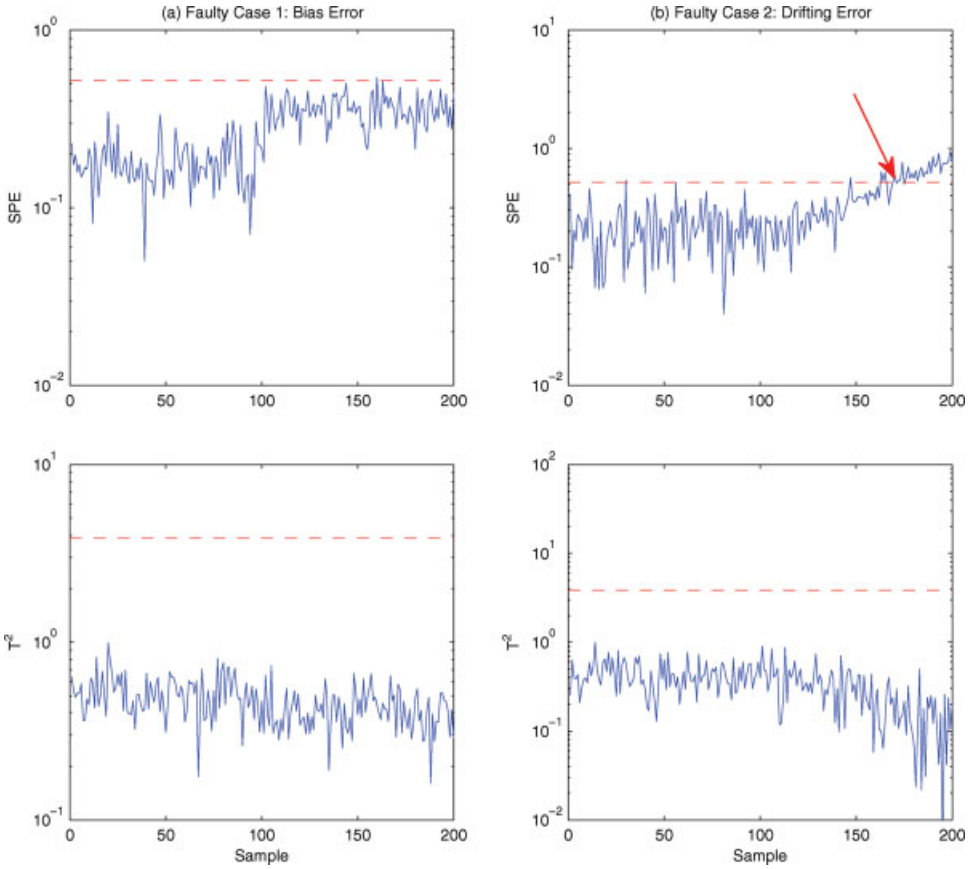

Figure 6. Illustrative example: Conventional PCA-based SPE and T^2 control charts in faulty Case 1 (a) and 2 (b), respectively.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

system is assumed to be running under the above three modes with equal probabilities. In the simulation, we generate 100 samples under each operating mode and use all the 300 samples as training data to construct a FGMM. The time series plots of the training data are displayed in Figure 3. By implementing the F–J algorithm on the training samples, a three-component Gaussian mixture model is obtained and all the statistical parameters (means, covariances, and prior probabilities) are estimated. Figure 4 shows the scatter plots of the multimode training data along with the estimated 95% probability ellipses of the three Gaussian components in all 2D subspaces. The means at different modes are significantly shifted along with the inflated or reduced covariances, which can be inferred from the altered source signals in the simulation design. However, it is noticed that the orientations of the three covariances are essentially the same. This is due to the underlying colinearity of the multivariate system. It can be easily seen from Figure 4 that the established FGMM precisely characterizes the multimodal distribution of the operating data. Furthermore, the estimated prior probabilities of the three Gaussian components are all 0.33, which exactly matches the simulated operation. It is worthwhile to emphasize that a priori knowledge on the number of Gaussian clusters is not required prior to the FGMM estimation.

To illustrate the utility of the proposed monitoring strategy, a period of test data are generated in the following two faulty scenarios, respectively.

- Case 1: The system is initially running at mode 1, and then a bias error of 1.5 is added to x_1 from the 101st through 200th samples.
- Case 2: The system is initially running at mode 2, and a drifting error of $0.04(k - 100)$ with k denoting the serial number of test samples is applied to x_2 between the 101st through 200th samples.

The test data in the above two cases are exhibited in Figures 5a,b where both the bias and the drifting errors are marked by arrows. The conventional PCA-based monitoring procedure is first carried out on the test data, where the PCA model is built using the full training data set from three different modes. The corresponding SPE and T^2 control charts under 95% confidence level in two faulty cases are given in Figures 6a,b, respectively. It is obvious that neither SPE nor T^2 index can detect the bias fault in Case 1 as both of them are below the control limit lines from the 101st through 200th samples. Similarly for Case 2, the T^2 index completely fails, whereas the SPE index does not go beyond the control limit line until the 170th sample. The results indicate that the conventional PCA method is unable to deal with the multimode system because the prerequisite of unimodal Gaussian

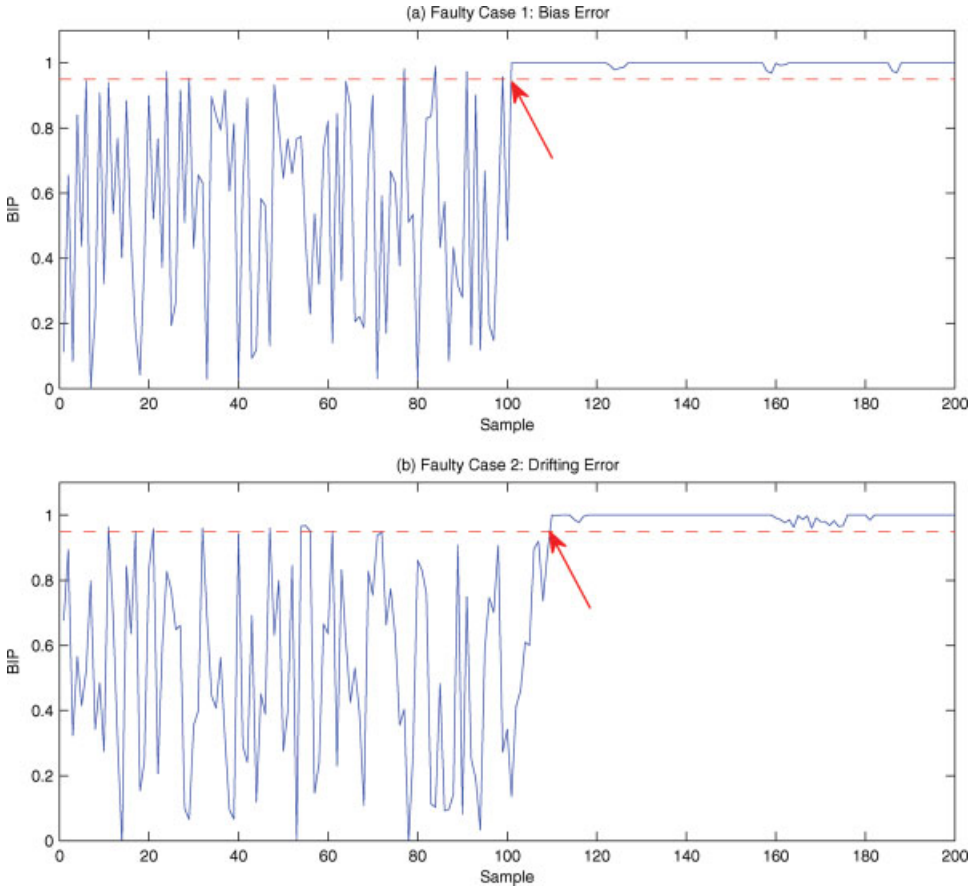

Figure 7. Illustrative example: FGMM-based BIP control charts in faulty Case 1 (a) and 2 (b), respectively.
[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

distribution becomes invalid in this situation. For the data from multiple operating regions, the inflated global covariance makes both SPE and T^2 indices insensitive to different kinds of faults. This fundamental issue cannot be resolved simply by reducing the confidence level because it will cause more normal samples to be misdetermined as faulty ones and increase the type I error accordingly.

In contrast, the FGMM-based monitoring approach is performed and the obtained BIP control charts for both faulty cases are shown in Figures 7a,b. In the first case of bias error, it can be observed that the BIP index value remains below the control limit line for most of the initial 100 samples. However, it jumps across the control limit right after the 100th sample and then stays above the threshold value until the 200th sample. The BIP trend implies that the fault occurs at the 101th sample and continues through the rest of test period, which coincides with the simulation design of Case 1. Since the confidence level used is 95%, the BIP values are beyond the control limit at ~6% normal samples, which constitute the Type I error of fault detection. A more aggressive setting of confidence level may result in a worse

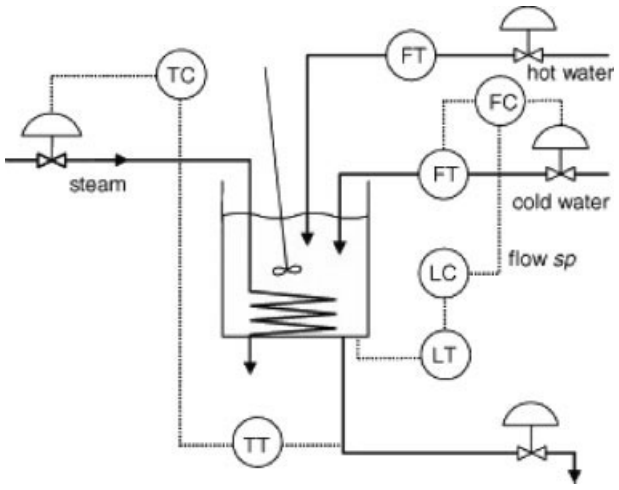

Figure 8. Schematic diagram of the simulated CSTH process from Thornhill et al.[55]

Table 1. CSTH Example: Five Sets of Normal Operating Conditions

Variable	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Level SP	12	16	9	12	12
Temperature SP	10.5	10.5	10.5	13.5	8
HW Valve	5.5	5	4.5	6	4

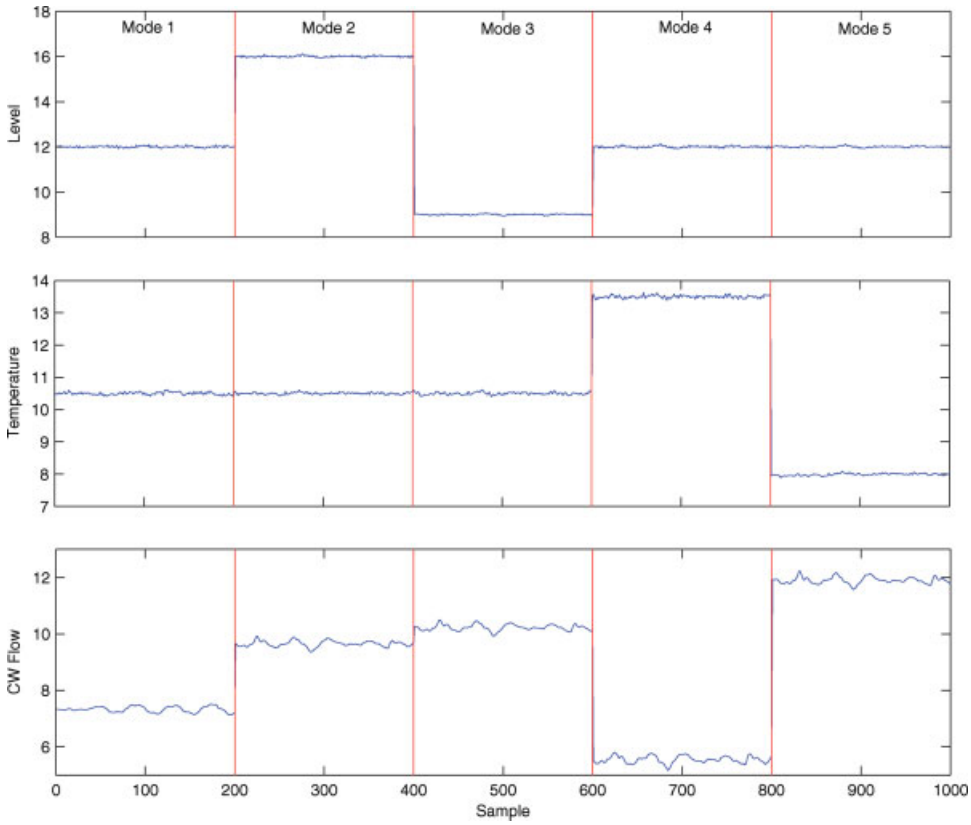

Figure 9. CSTH example: Training data from five operating modes.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Type II error, though the Type I error can be reduced as a consequence. In faulty Case 2, a drifting error of x_2 starts from the 101st sample under operating mode 2. As shown in Figure 7b, the BIP index exceeds the 95% control limit line at the 110th sample and has maintained above the threshold value since then. It indicates the fault occurrence in the second half of test period. The BIP-based fault alarm is triggered with a short delay of 9 samples, which should be attributed to the incipient characteristic of the drifting fault. Since the x_2 measurement is drifting slowly, the first few samples are still located within the statistically normal operating region. As the measurement error is being accumulated, the sample points will move out of the local probability ellipsoid at operating mode 2 and result in the increased BIP values over the control limit. After the initial detection point, all the subsequent faulty samples stand above the control limit line. As opposed to the complete failure of T^2 index and long delay of SPE index in the case of drifting fault, the alternative BIP metric is capable of capturing the slow drifting error in a fairly accurate and fast fashion. The results of two test scenarios in this example demonstrate that the FGMM-based BIP control chart is much more effective than the conventional PCA-based SPE/T^2 control chart in detecting various types of faults under multiple operating modes. Meanwhile, the BIP chart looks similar to the traditional control charts which can be visually inspected by practitioners in a similar manner.

Simulated CSTH example

In this section, the proposed monitoring framework is applied to a simulated continuous stirred tank heater process under multiple operating conditions. The closed-loop simulation of the CSTH pilot plant was originally developed by Thornhill et al.[55] As depicted in Figure 8, the stirred tank is supplied with well mixed hot and cold water (HW and CW) flow and heated by steam. The process inputs are the HW, CW, and steam valve demands, and the outputs include the electronic measurements of the level, temperature, HW, and CW flow that are controlled by multiple PI loops. The normal disturbances that are present in the plant comprise an oscillatory disturbance to the CW flow, a random disturbance to the level, and a measurement noise in the temperature. A full description of the CSTH process can be found in Thornhill et al.[55] and the closed-loop simulation programs are available at their website.[56]

As listed in Table 1, five sets of operating conditions on the level setpoint, temperature setpoint and HW valve are designed to test the performance of the proposed method in monitoring multimode processes. Three controlled variables, i.e., level, temperature, and CW flow, are selected for monitoring in this study. It is noted that all the measurements in this example are in the unit of mA. In the training period, 200 samples are generated under each operating mode with 1-s sampling time and the overall 1000 samples from multi-

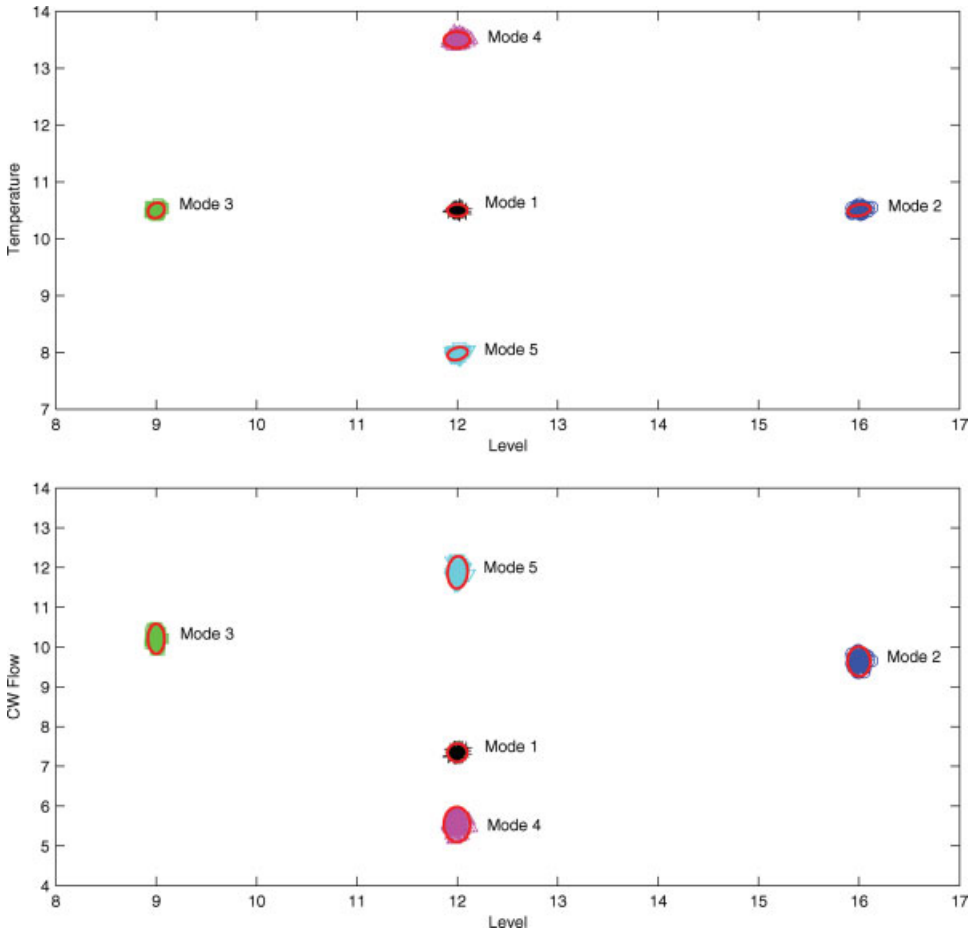

Figure 10. CSTH example: Scatter plots of training data along with the estimated local probability ellipses.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

ple modes are used to construct the nominal FGMM model. The time series plots of the training data is given in Figure 9. By feeding the data into the model learning framework, a five-component GMM is obtained. The estimated prior probability for each component is 0.2, which is in well agreement with the actual process operation. The accuracy of other model parameters (i.e., component means and covariances) is further verified by the projected probability ellipses, as shown in Figure 10. It is obvious that each operating region is well defined by the corresponding Gaussian component. As a critical prerequisite, the precise mixture model enables the effective detection of abnormal operating conditions in the monitoring stage.

For the testing purpose, two abnormal scenarios of instrument bias fault and random variation error are considered. In the first case, the initial 100 samples are simulated under normal operating mode 1 and then a sudden step change of -0.5 is introduced into the level measurement from the 101st sample through the 200th sample. For the second case, the initial 100 normal samples are also collected under mode 1, but followed by an inflated random variation in the temperature measurement. The test data in both cases are exhibited in Figures 11a,b. It can be readily seen that the immediate effect of the bias fault in Case 1 is a downward step change in the measured tank level, which leads to a dramatic increase in the CW flow and then the temperature drop. With the control actions to compensate for the level change, the plant is able to move back to the original steady-state operating point at the end of the test period. As shown in Figure 12a, the global PCA-based SPE and T^2 control charts perform poorly in detecting the bias fault in Case 1. The T^2 index is far below its control limit through the whole test period and thus unable to trigger any fault alarm. On the other hand, the SPE index does not exceed the control limit until the 140th sample and returns below the limit line at the 180th sample. It suggests that the fault alarm is delayed by $\sim$40 samples after the first advent of the level step change. The low detection rate and long alarm delay of the global PCA method stem from the inflated model covariance and the reduced fault sensitivity. On the contrary, the proposed BIP control chart in Figure 13a provides a clear and obvious indication of fault occurrence right after the 100th sample. Furthermore, the BIP index value stays above the corresponding control limit until the end of the test period, when the monitored samples are moving back into the original normal operating region. In this case, it is evident that the FGMM-based approach outperforms the global PCA method in terms of fault detection rate and sensitivity for multimode

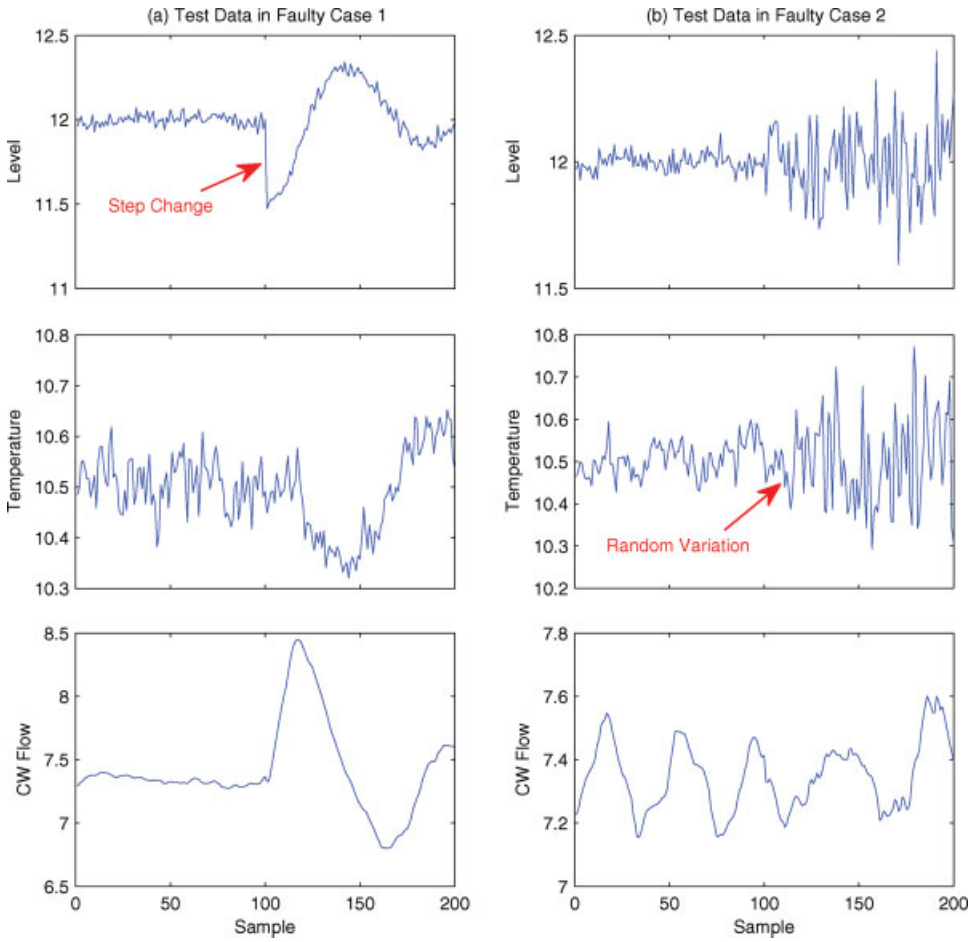

Figure 11. CSTH example: Test data of faulty Case 1 (a) and 2 (b), respectively.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

processes. In the second test scenario, the increased temperature variation happens immediately after the first 100 samples and results in the magnified variation of the level, as illustrated in Figure 11b. The PCA-based SPE and T^2 control charts in Figure 12b completely fail to detect the random variation fault. However, the BIP metric in Figure 13b reports the sensor fault accurately by jumping outside of the confidence limit from the 101st sample without any delay. As the random variation error persists, the BIP index almost remains above the control limit through the end of the test period. The only two exceptional points, i.e., the 114th and 168th samples, constitute a 2% Type II error. In this multimode CSTH example, the FGMM-based monitoring approach is further proven to be highly effective in detecting different types of sensor faults with negligible false alarm rates, whereas the global PCA method appears to be very insensitive to the faults.

Tennessee Eastman challenge problem

As a well-known benchmark process, the Tennessee Eastman process has been widely applied to evaluate and compare the efficiency of process monitoring techniques. The prototype problem is presented by Downs and Vogel[57] and based on an industrial process in Tennessee Eastman Chemical Company. The schematic diagram of the process is illustrated in Figure 14. It consists of five unit operations, i.e., a reactor, a partial condenser, a recycle compressor, a stripper and a vapor/liquid separator. Four gaseous reactants are fed into the reactor to form two products along with a byproduct and an inert, which are separated through the downstream operations. Overall 41 measured output variables and 12 manipulated variables are present in the process. Moreover, there are six modes of process operation, as listed in Table 2. The optimal steady-state conditions can be found from Downs and Vogel.[57] Since the process is essentially open-loop unstable, various control strategies have been developed for this problem.[58–62] In our study, the decentralized control design developed by Ricker[61] is adopted for the closed-loop process operation and the simulation codes can be downloaded online.[63] The sampling time used is 0.05 h and 22 continuous measurements among the 41 output variables are selected as monitored variables. Furthermore, the first three steady-state operating modes are assumed to be present with equal possibilities during the normal process operation.

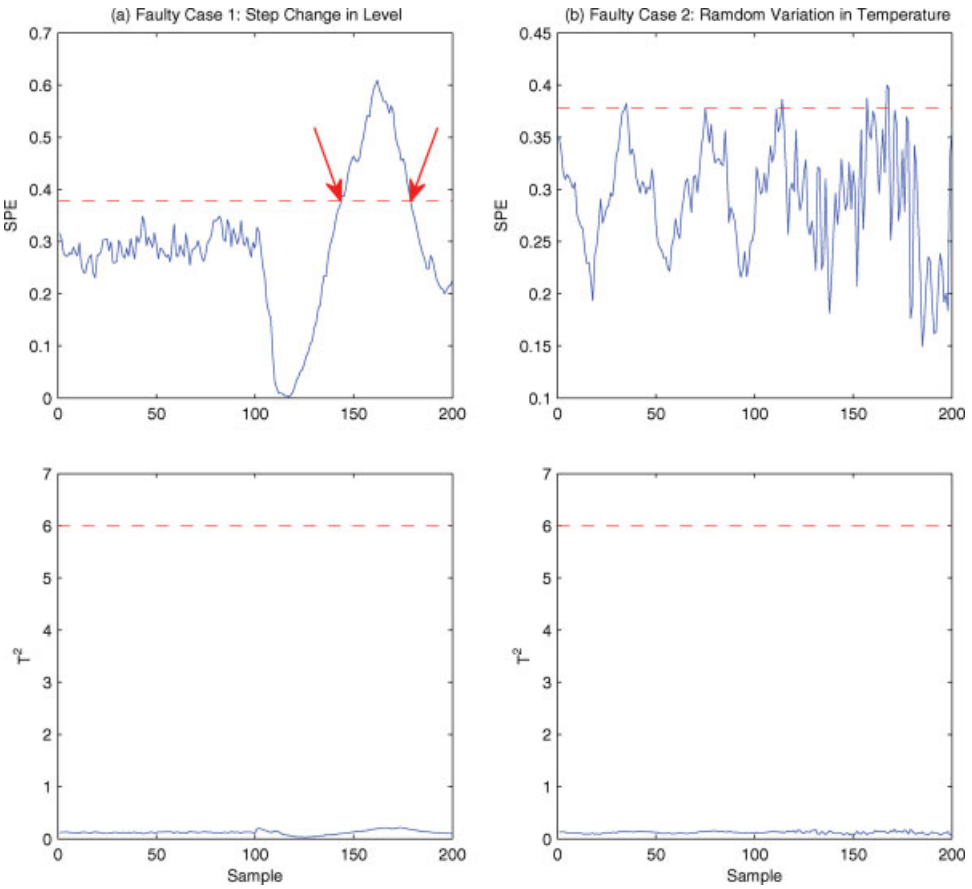

Figure 12. CSTH example: Conventional PCA-based SPE and T^2 control charts in faulty Case 1 (a) and 2 (b), respectively.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

To construct the nominal model for process monitoring, 1000 samples are collected under each of the above three operating modes and total 3000 samples constitute the training data set. The obtained FGMM model consists of three Gaussian components with the prior probability of 0.33 for each component. The projected training data in the 2D PCS along with the 95% probability ellipses corresponding to the three components are exhibited in Figure 15. One can see that the probability density of the operating data under the three modes is well characterized by the estimated Gaussian components. It should be noted that the training data in the original 22D space is directly used for the FGMM estimation and no preliminary step of dimensionality reduction is needed.

To examine the effectiveness of the proposed monitoring approach for complex industrial processes, four test scenarios with different kinds of prespecified faults are designed, as shown in Table 3. The first three test cases are running at a single operating mode with certain introduced fault, while the last one includes a mode shift and then an incoming fault in the operation. The conventional PCA method and the FGMM-based monitoring approach are applied to the four cases and the corresponding control charts are shown in Figures 16 and 17, respectively. In Case 1, it appears that all the three control charts (SPE, T^2, and BIP) are capable of detecting the process fault of feed loss with a short delay of around 24 min. Because of the interactions among the monitored variables, the sudden loss of feed A can cause a series of related variables to move far away from the normal operating regions rapidly. As a result, the accumulated SPE and T^2 indices of faulty data would be beyond the control limits even though the global PCA model inflates the confidence boundary of the multimode data. In this special case, both PCA- and FGMM-based methods capture the dramatic process fault with similar efficiency.

For Case 2, however, the fault alarm from the global PCA method is delayed for more than 150 min in both SPE and T^2 charts. As the global probability ellipse is significantly inflated, the initial period of samples following the step change of condenser cooling water temperature is still within the enlarged normal operating regions, which result in the long delay of fault detection. On the contrary, the proposed probabilistic approach catches the process fault with negligible delay and high detection rate, as shown in Figure 17b. The comparison of detection results in Case 2 reveals that the proposed method has higher sensitivity to the challenging step change fault in multimode processes and can trigger the fault alarm in a timely manner.

In Case 3, the increased random variation of condenser cooling water temperature takes place at mode 3 between the

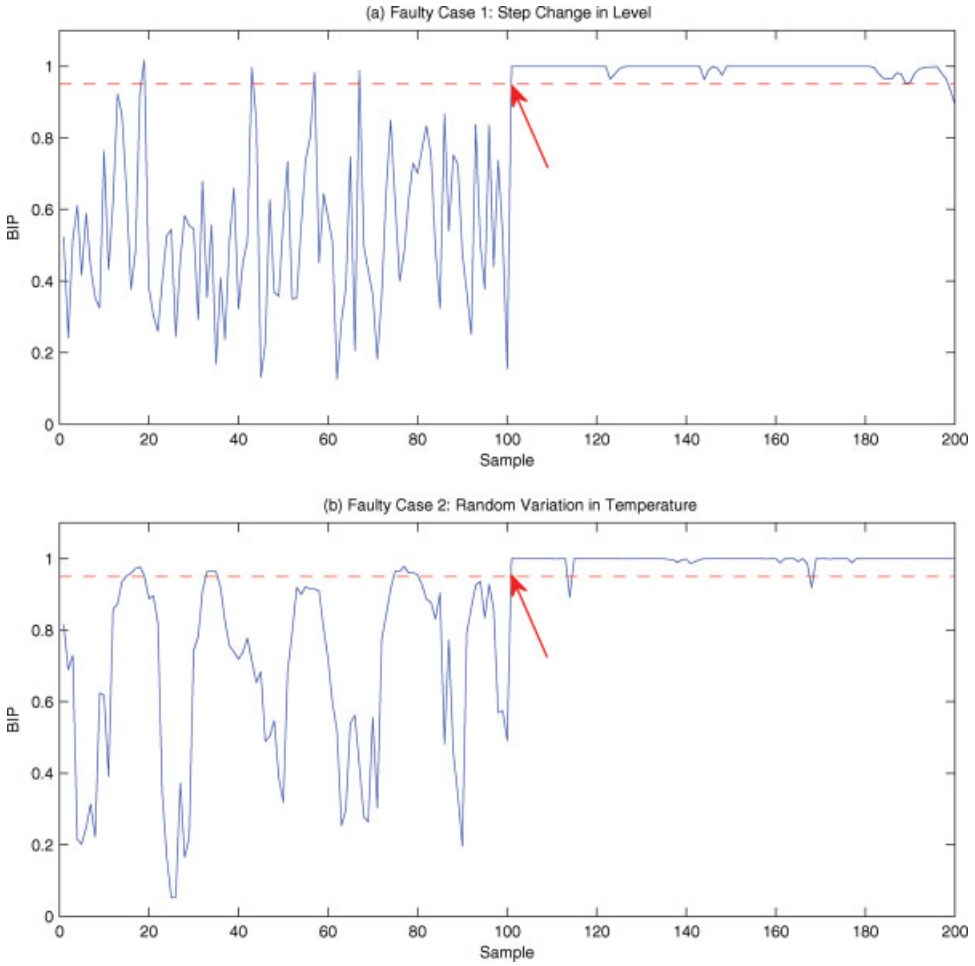

Figure 13. CSTH example: FGMM-based BIP control charts in faulty Case 1 (a) and 2 (b), respectively.
[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

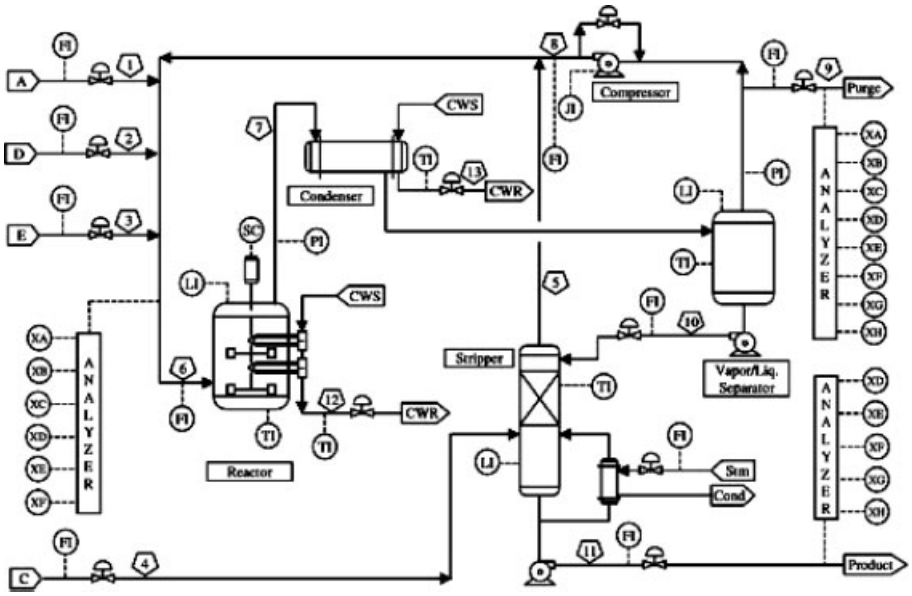

Figure 14. Schematic diagram of the TEP from Downs and Vogel.[57]

Table 2. TEP Example: Six Process Operation Modes from Downs and Vogel[57]

Mode	G/H Mass Ratio	Production Rate (Stream 11)
1	50/50	7038 kg/h G and 7038 kg/h H
2	10/90	1408 kg/h G and 12669 kg/h H
3	90/10	10000 kg/h G and 1111 kg/h H
4	50/50	Maximum
5	10/90	Maximum
6	90/10	Maximum

101st and the 200th samples. The fault can further propagate in the plant and cause the increased variations in other relevant variables to some extent. As observed in Figure 16c, most of the faulty samples fall below the 95% control limits in both SPE and T^2 charts, which thus have extremely low detection rates. The Type II error can be reduced by decreasing the confidence level, whereas the increased Type I error will induce more false alarms. The boundary of normal operating conditions inflates in the faulty directions and therefore makes the SPE and T^2 indices particularly insensitive to this random variation fault. In Figure 17c, nevertheless, the FGMM-based BIP control chart is found to be able to detect the same fault without significant delay or false alarms. The fault detection rate is 98% and the false alarm rate is 4%, both of which demonstrate the high accuracy of the proposed approach in detecting the difficult random variation fault.

For the fourth test case, the process is initially running at mode 3 for 5 h and then switched to mode 1 for additional 5 h. From the 201st sample, a slow drift in reaction kinetics occurs and continues for 5 h. The conventional PCA method still performs poorly in dealing with the slow drifting fault. As shown in Figure 16d, the SPE and T^2 control charts miss most of the faulty samples and thus have high type II errors. In the T^2 chart, for instance, it is obvious that the drifting trend of the statistical metric starts after about 150 min of the first event of drift fault. Even worse, the detection of fault occurrence is further delayed for another 60 min as the T^2 index does not move above the control limit until then. As opposed to the PCA method, the FGMM approach successfully detects the drift fault in an early stage with a delay of 33 min only. Furthermore, the detection rate keeps nearly 99% after the initial effective detection of the fault. The results show that the FGMM-based monitoring approach is far superior to the conventional PCA method in detecting the slow drift fault with a much shorter delay. The different test scenarios in the challenging TEP example demonstrate the outstanding potential of the proposed approach to detect various types of faults in industrial processes of multiple modes.

Conclusions

A novel finite Gaussian mixture model and Bayesian inference-based process monitoring approach is successfully developed in this article. Aimed at the complex industrial processes with multiple operating conditions, the proposed method works when the normal operating data follow a multimodal distribution, which can be characterized by a finite Gaussian mixture model. Based on the multimode operating data, the F–J algorithm is then employed to estimate the optimal number of Gaussian components and their statistical properties, i.e., means, covariances, and prior probabilities. With the constructed mixture model, a Bayesian inference procedure is utilized to compute the posterior probabilities of

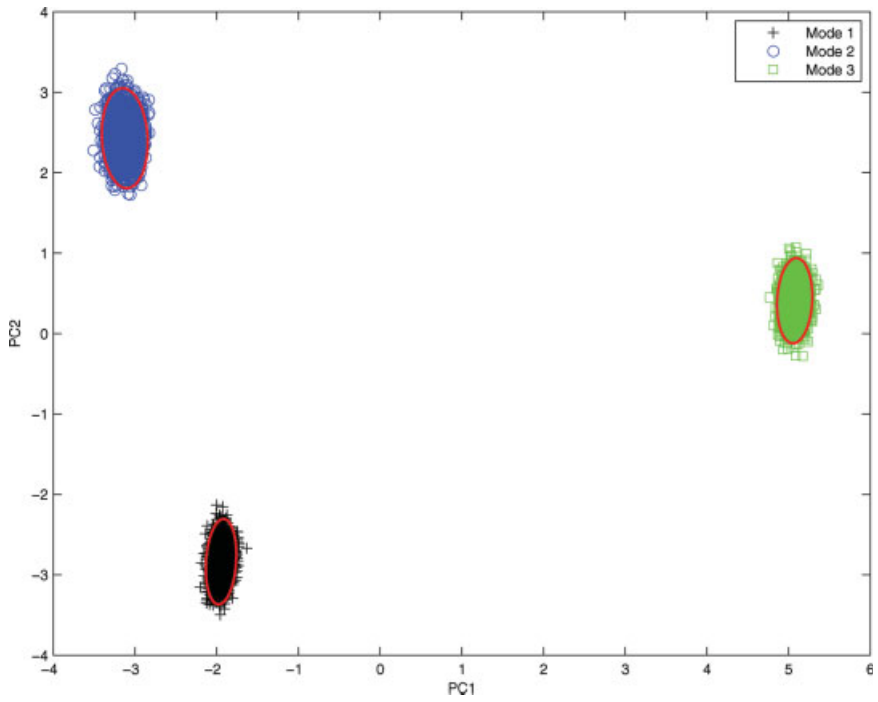

Figure 15. TEP example: Scatter plots of training data along with the estimated local probability ellipses.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 3. TEP Example: Four Test Cases with Different Kinds of Process Faults

Case	Description
1	Normal sample 1–100, Mode 1 Faulty sample 101–200 IDV6: Sudden loss of A feed
2	Normal sample 1–100, Mode 2 Faulty sample 101–200 IDV5: Step change in condenser cooling water temperature
3	Normal sample 1–100, Mode 3 Faulty sample 101–200 IDV12: Random variation in condenser cooling water temperature
4	Normal sample 1–100, Mode 3 Normal sample 101–200, Mode 1 Faulty sample 201–300 IDV13: Slow drift in reaction kinetics

each monitored sample belonging to all the Gaussian components, through which the local Mahalanobis distance metrics are further integrated into a global BIP index. Consequently, the multimode process can be monitored using the BIP control chart.

The presented monitoring approach is applied to the fault detection of three multimode examples, which include an illustrative multivariate system, a simulated CSTH process and the challenging TEP problem. The comparison of monitoring results demonstrates that the presented approach is superior to the conventional PCA method in terms of higher detection rates, lower false alarm rates and shorter delays of fault capture. In addition, it is shown to be consistently effective in detecting different kinds of process faults. Another advantage of the FGMM-based approach lies in its unified monitoring metric (BIP) instead of two separate indices (SPE and T^2) for the PCA-based method. The proposed approach therefore

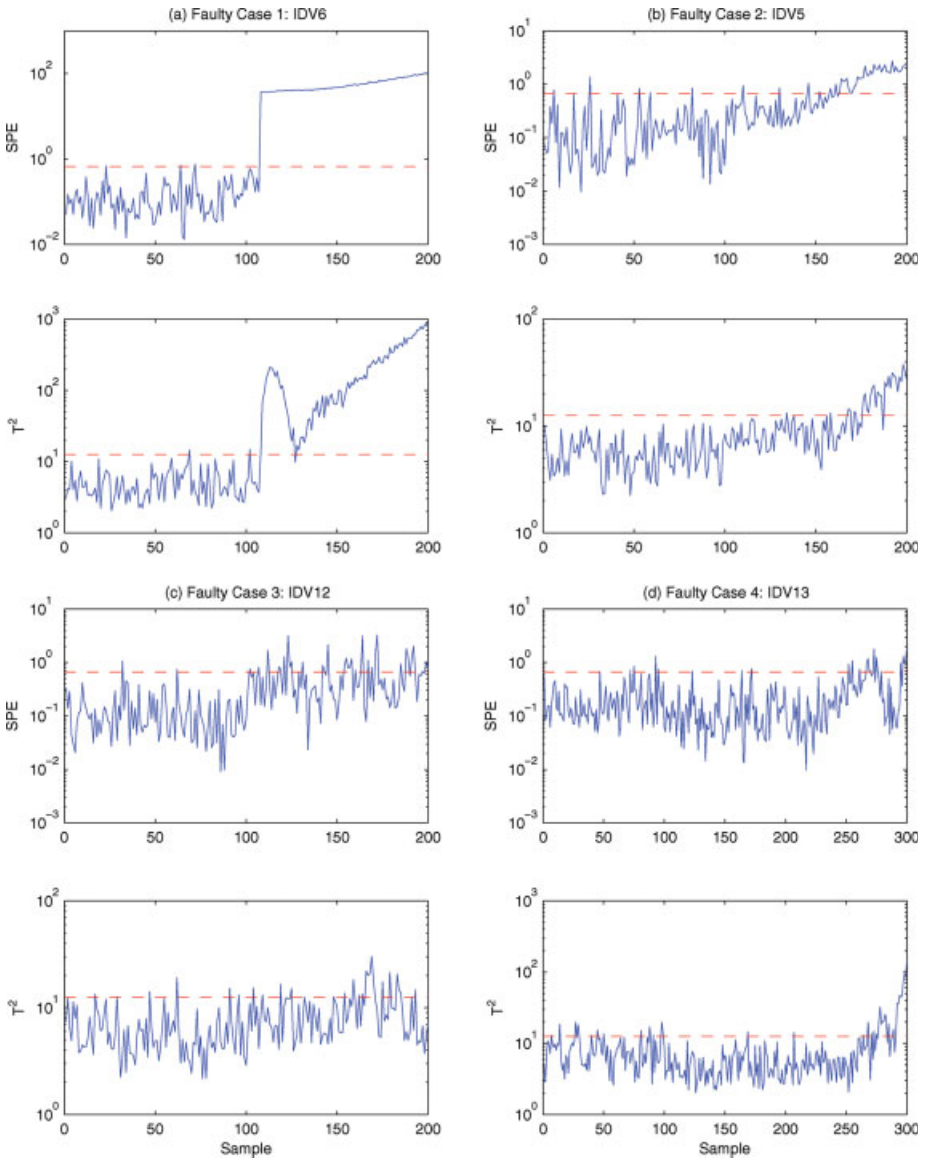

Figure 16. TEP example: Conventional PCA-based SPE and T^2 control charts in faulty Case 1 (a), 2 (b), 3 (c) and 4 (d), respectively.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

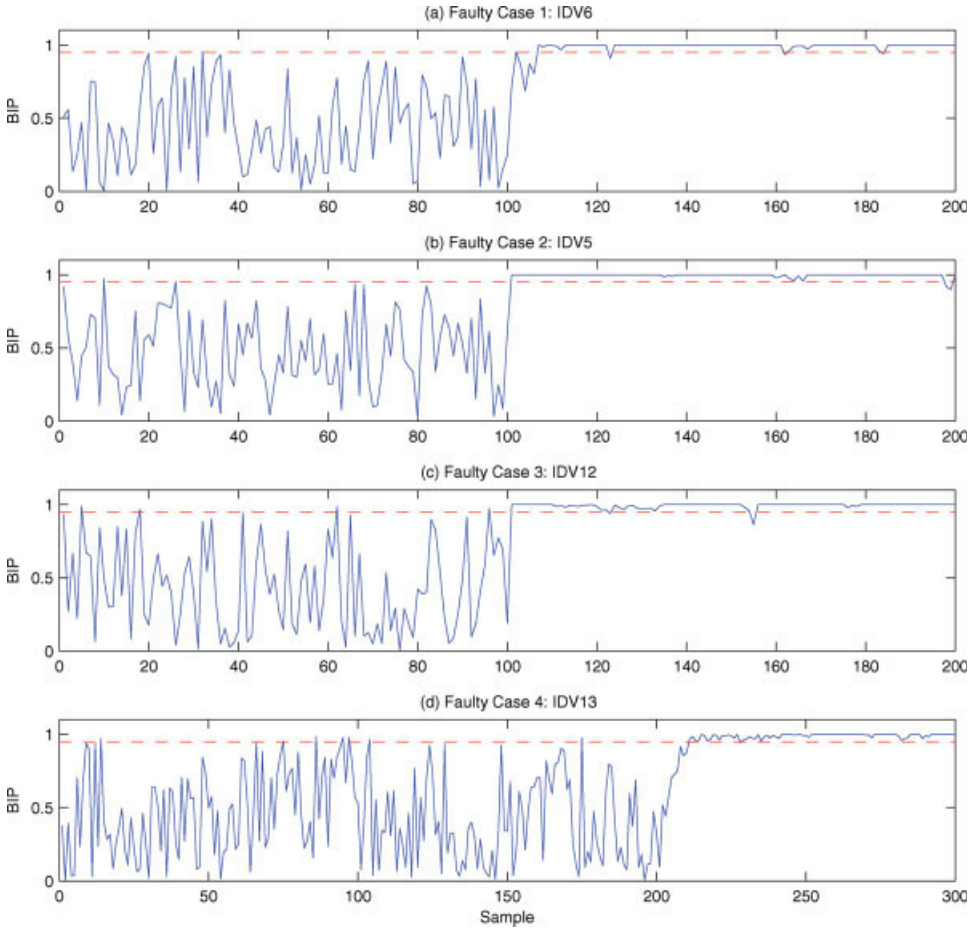

Figure 17. TEP example: FGMM-based BIP control charts in faulty Case 1 (a), 2 (b), 3 (c) and 4 (d), respectively.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

provides a promising tool to monitor the complex industrial processes under multiple operating conditions without needing a priori process knowledge. However, it should be pointed out that the proposed BIP index, like the SPE and T^2 indices, is sensitive to data autocorrelation due to system dynamics and thus requires each normal operating mode to be at steady state approximately. Our current research is focused on fault detection of continuous processes in chemical industry. Future work may extend the proposed methodology to the batch or semibatch processes. Furthermore, the fault diagnosis aspect of multimode processes also deserves investigation to identify the root cause variables.

Acknowledgments

The authors appreciate the valuable comments and suggestions of the anonymous reviewers.

Literature Cited

1. Kourti T, MacGregor JF. Multivariate SPC methods for process and product monitoring. *J Qual Technol*. 1996;28:409–428.
2. Chiang LH, Russell EL, Braatz RD. *Fault Detection and Diagnosis in Industrial Systems. Advanced Textbooks in Control and Signal Processing*. London, Great Britain: Springer-Verlag, 2001.
3. Qin SJ. Statistical process monitoring: basics and beyond. *J Chemom*. 2003;17:480–502.
4. Yélamos I, Graells M, Puigjaner L, Escudero G. Simultaneous fault diagnosis in chemical plants using a multilabel approach. *AIChE J*. 2007;53:2871–2884.
5. Paulonis MA, Cox JW. A practical approach for large-scale controller performance assessment, diagnosis, and improvement. *J Process Control*. 2003;13:155–168.
6. Qin SJ, Yu J. Recent developments in multivariable controller performance monitoring. *J Process Control*. 2007;17:221–227.
7. Singhal A, Seborg DE. Evaluation of a pattern matching method for the Tennessee Eastman challenge process. *J Process Control*. 2006;16:601–613.
8. Yu J, Qin SJ. Statistical MIMO controller performance monitoring. II. Performance diagnosis. *J Process Control*. 2008;18:297–319.
9. Isermann R. Process fault detection based on modeling and estimation methods—a survey. *Automatica*. 1984;20:387–404.
10. Basseville M. Detecting changes in signals and systems—a survey. *Automatica*. 1988;24:309–326.
11. Gertler JJ. Survey of model-based failure detection and isolation in complex plants. *IEEE Control Syst Mag*. 1988;8:3–11.
12. Frank PM. Fault diagnosis in dynamic systems using analytical and knowledge based redundancy—a survey and some new results. *Automatica*. 1990;26:459–474.
13. Yoon S, MacGregor JF. Fault diagnosis with multivariate statistical models. I. Using steady state fault signatures. *J Process Control*. 2001;11:387–400.
14. MacGregor JF, Kourti T. Statistical process control of multivariate processes. *Control Eng Pract*. 1995;3:403–414.

15. MacGregor JF, Jaeckle C, Kiparissides C, Koutoudi M. Process monitoring and diagnosis by multiblock PLS methods. *AIChE J*. 1994;40:826–838.
16. Dunia R, Qin SJ, Edgar TF, McAvoy TJ. Identification of faulty sensors using principal component analysis. *AIChE J*. 1996;42:2797–2812.
17. Raich A, Cinar A. Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. *AIChE J*. 1996; 42:995–1009.
18. Bakshi BR. Multiscale PCA with application to multivariate statistical process monitoring. *AIChE J*. 1998;44:1596–1610.
19. Qin SJ. Recursive PLS algorithms for adaptive data modeling. *Comput Chem Eng*. 1998;22:503–514.
20. Wold S, Esbensen K, Geladi P. Principal component analysis. *Chemom Intell Lab Syst*. 1987;2:37–52.
21. Tong H, Crowe CM. Detection of gross errors in data reconciliation by principal component analysis. *AIChE J*. 1995;41:1712–1722.
22. Negiz A, Cinar A. Statistical monitoring of multivariable dynamic processes with state space models. *AIChE J*. 1997;43:2002–2020.
23. Chiang LH, Russell EL, Braatz RD. Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis. *Chemom Intell Lab Syst*. 2000;50:243–252.
24. Simoglou A, Martin EB, Morris AJ. Statistical performance monitoring of dynamic multivariate processes using state space modeling. *Comput Chem Eng*. 2002;26:909–920.
25. Lee JM, Yoo CK, Lee IB. Statistical process monitoring with independent component analysis. *J Process Control*. 2004;14:467–485.
26. He QP, Qin SJ, Wang J. A new fault diagnosis method using fault directions in Fisher discriminant analysis. *AIChE J*. 2005;51:555–571.
27. Lee C, Choi SW, Lee IB. Variable reconstruction and sensor fault identification using canonical variate analysis. *J Process Control*. 2006;16:747–761.
28. Lee JM, Qin SJ, Lee IB. Fault detection and diagnosis based on modified independent component analysis. *AIChE J*. 2006;52:3501–3514.
29. Kramer MA. Autoassociative neural networks. *Comput Chem Eng*. 1992;16:313–328.
30. Vinson JM, Ungar LH. Dynamic process monitoring and fault diagnosis with qualitative models. *IEEE Trans Syst Man Cybern*. 1995;25:181–189.
31. Rengaswamy R, Venkatasurbramanian V. A fast training neural network and its updation for incipient fault detection and diagnosis. *Comput Chem Eng*. 2000;24:431–437.
32. Mehranbod N, Soroush M, Piovoso M, Ogunnaike BA. Probabilistic model for sensor fault detection and identification. *AIChE J*. 2003;49:1787–1802.
33. Chiang L, Kotanchek M, Kordon A. Fault diagnosis based on Fisher discriminant analysis and support vector machines. *Comput Chem Eng*. 2004;28:1389–1401.
34. Chu Y, Qin SJ, Han C. Fault detection and operation mode identification based on pattern classification with variable selection. *Ind Eng Chem Res*. 2004;43:1701–1710.
35. Mehranbod N, Soroush M, Panjapornpon C. A method of sensor fault detection and identification. *J Process Control*. 2005;15:321–339.
36. Verron S, Tiplica T, Kobi A. Fault detection and identification with a new feature selection based on mutual information. *J Process Control*. 2008;18:479–490.
37. Hwang DH, Han C. Real-time monitoring for a process with multiple operating modes. *Control Eng Pract*. 1999;7:891–902.
38. Lane S, Martin EB, Kooijmans R, Morris AJ. Performance monitoring of a multiproduct semi-batch process. *J Process Control*. 2001;11:1–11.
39. Zhao SJ, Zhang J, Xu YM. Monitoring of processes with multiple operating modes through multiple principal component analysis models. *Ind Eng Chem Res*. 2004;43:7025–7035.
40. Zhao SJ, Zhang J, Xu YM. Performance monitoring of processes with multiple operating modes through multiple PLS models. *J Process Control*. 2006;16:763–772.
41. Chen J, Liu J. Mixture principal component analysis models for process monitoring. *Ind Eng Chem Res*. 1999;38:1478–1488.
42. Beebe KR, Pell RJ, Seasholtz MB. *Chemometrics: A Practical Guide*. New York, NY: Wiley, 1998.
43. Choi SW, Park JH, Lee IB. Process monitoring using a Gaussian mixture model via principal component analysis and discriminant analysis. *Comput Chem Eng*. 2004;28:1377–1387.
44. Yoo CK, Villez K, Lee IB, Rosén C, Vanrolleghem PA. Multi-model statistical process monitoring and diagnosis of a sequencing batch reactor. *Biotechnol Bioeng*. 2007;96:687–701.
45. Thissen U, Swierenga H, deWeijer AP, Wehrens R, Melssen WJ, Buydens LMC. Multivariate statistical process control using mixture modeling. *J Chemom*. 2005;19:23–31.
46. Hyndman RJ. Computing and graphing highest density regions. *Am Stat*. 1996;50:120–126.
47. Figueiredo MAF, Jain AK. Unsupervised learning of finite mixture models. *IEEE Trans Pattern Anal Mach Intell*. 2002;24:381–396.
48. Nomikos P, MacGregor JF. Multivariate SPC charts for monitoring batch processes. *Technometrics*. 1995;37:41–59.
49. Bishop CM. *Neural Networks for Pattern Recognition*. New York, NY: Oxford University Press, 1995.
50. Duda RO, Hart PE, Stork DG. *Pattern Classification*, *2nd ed*. New York, NY: Wiley, 2001.
51. Paalanen P, Kamarainen JK, Ilonen J, Kälviäinen H. Feature representation and discrimination based on Gaussian mixture model probability densities - Practices and algorithms. *Pattern Recogn*. 2006;39:1346–1358.
52. Mao J, Jain AK. A self-organizing network for hyperellipsoidal clustering (HEC). *IEEE Trans Neural Networks*. 1996;7:16–29.
53. Archambeau C, Vrins F, Verleysen M. Flexible and robust Bayesian classification by finite mixture models. In: *European Symposium on Artificial Neural Networks*, Bruges, Belgium; 2004:75–80.
54. Box GEP. Some theorems on quadratic forms applied in the study of analysis of variance problems. I. Effect of inequality of variance in the one-way classification. *Ann Math Stat*. 1954;25:290–302.
55. Thornhill NF, Patwardhan SC, Shah SL. A continuous stirred tank heater simulation model with applications. *J Process Control*. 2008;18:347–360.
56. Thornhill NF, Patwardhan SC, Shah SL. The CSTH simulation website 2007. Available at: http://personal-pages.ps.ic.ac.uk/~nina/CSTHSimulation/index.htm. Accessed on November 22nd, 2007.
57. Downs JJ, Vogel EF. Plant-wide industrial process control problem. *Comput Chem Eng*. 1993;17:245–255.
58. McAvoy TJ, Ye N. Base control for the Tennessee Eastman problem. *Comput Chem Eng*. 1994;18:383–413.
59. Lyman PR, Georgakis C. Plant-wide control of the Tennessee Eastman problem. *Comput Chem Eng*. 1995;19:321–331.
60. Ricker NL, Lee JH. Nonlinear model predictive control of the Tennessee Eastman challenge process. *Comput Chem Eng*. 1995;19:961–981.
61. Ricker NL. Decentralized control of the Tennessee Eastman challenge process. *J Process Control*. 1996;6:205–221.
62. Larsson T, Hestetun K, Hovland E, Skogestad S. Self-optimizing control of a large scale plant: the Tennessee Eastman process. *Ind Eng Chem Res*. 2001;40:4889–4901.
63. Ricker NL. Tennessee Eastman challenge archive 1999. Available at: http://depts.washington.edu/control/LARRY/TE/download.html. Accessed on November 22nd, 2007.

Manuscript received Dec. 2, 2007, and revision received Mar. 24, 2008.