Adaptive Multivariate Statistical Process Control for Monitoring Time-Varying Processes

Sang Wook Choi, Elaine B. Martin,* A. Julian Morris, and In-Beum Lee†

Centre for Process Analytics and Control Technology, School of Chemical Engineering and Advanced Materials, University of Newcastle upon Tyne, Newcastle NE1 7RU, U.K., and Department of Chemical Engineering, Pohang University of Science and Technology, San 31 Hyoja Dong, Pohang 790-784, Korea

An adaptive multivariate statistical process monitoring (MSPC) approach is described for the monitoring of a process with incurs operating condition changes. Samplewise and blockwise recursive formulas for updating a weighted mean and covariance matrix are derived. By utilizing these updated mean and covariance structures and the current model, a new model is derived recursively. On the basis of the updated principal component analysis (PCA) representation, two monitoring metrics, Hotelling's T^2 and the Q-statistic, are calculated and their control limits are updated. For more efficient model updating, forgetting factors, which change with time, for the updating of the mean and covariance are considered. Furthermore, the updating scheme proposed is robust in that it not only reduces the false alarm rate in the monitoring charts but also makes the model insensitive to outliers. The adaptive MSPC approach developed is applied to a multivariate static system and a continuous stirred tank reactor process, and the results are compared to static MSPC. The revised approach is shown to be effective for the monitoring of processes where changes are either fast or slow.

1. Introduction

Multivariate statistical projection methods such as principal component analysis (PCA) and partial least squares (PLS) combined with statistical monitoring charts have been widely applied for on-line continuous and batch process monitoring.[1−3] Typically, monitoring charts based on Hotelling T^2 and the Q-statistic[4,5] are used to detect faults with the contribution plot,[6] which shows the contribution of each measured variable to the monitoring statistic, being adopted to assist in the identification of an assignable cause of a fault.

Multivariate statistical process control (MSPC) schemes using PCA or PLS are formulated on the assumption that the process variables are independent, identically distributed, and linearly correlated. Furthermore, if the underlying distribution of the process variables is multivariate normal, then the process can be described totally by the mean and covariance of the variables. Thus, because of the limitations imposed by the underlying assumptions, MSPC is usually performed under steady-state conditions. Occasionally, the problems associated with the inability and inefficiency of a process-monitoring scheme can arise because the underlying assumptions of conventional MSPC are violated by one of several reasons, including nonlinear process behavior, time dependency (autocorrelation), process dynamics, and changes in operating conditions. To address these limitations, a number of alternative monitoring schemes have been proposed, such as dynamic, nonlinear, or adaptive MSPC.

In this paper, an adaptive PCA-based MSPC scheme is proposed to monitor processes where operating condition changes are common. Industrial processes usually exhibit time-varying behavior due to intentional disturbances, such as set-point changes, or unwanted disturbances, such as the aging of equipment, fouling, and catalyst degradation. There have been several papers presented on adaptive PCA and PLS. For example, Wold[7] proposed a scheme which utilized the exponentially weighted moving concept for the updating of both PCA and PLS models. However, one limitation with this algorithm is that it requires all the historical data to be used in the updating of the model every time a new sample becomes available. Dayal and MacGregor[8] described a recursive exponentially weighted PLS algorithm, using the PLS kernel algorithm, based on recursively updating the covariance matrices as opposed to using the total historical data set. Qin[9] developed an alternative samplewise and blockwise recursive PLS algorithm that only utilizes the loading matrix and the regression coefficients between the input and output score vector pairs in the PLS model, instead of the covariance matrices, materializing in more rapid updating of the model. More recently, Li et al.[10] proposed two efficient learning algorithms for recursive PCA for the monitoring of time-varying processes.

The study reported in this paper focuses on developing a new adaptive monitoring scheme for time-varying processes with a number of contributions being presented. First, samplewise and batchwise recursive formulas are derived for updating the sample mean, variance, and covariance prior to an efficient and fast learning algorithm being proposed for calculating a new PCA representation through either the samplewise or blockwise approach. Second, an overall model-updating and process-monitoring method is presented. More specifically, a new calculation procedure for the variable forgetting factors is developed for updating the mean and covariance. Finally, a robust model updating scheme is described to reduce the effect of outliers on the model-updating procedure.

The remainder of this paper is structured as follows. Section 2 presents the recursive updating forms of the sample mean and covariance, with the focus in Section 3 being on an alternative and efficient way by which to find the eigenvalue−eigenvector pairs in the updated covariance (or correlation) matrix. In Section 4, an overall adaptive statistical process-monitoring strategy is proposed. Two specific aspects are considered which impact on the monitoring statistics: outliers and the selection of the forgetting factors (one for the mean

* To whom all correspondence should be addressed. Tel.: +44 191 222 6231. Fax: +44 191 222 5748. E-mail: e.b.martin@newcastle.ac.uk.
† Pohang University of Science and Technology.

10.1021/ie050391w CCC: $33.50
Published on Web 03/24/2006

and the other for the covariance). Finally, in Section 5, the results of the proposed methodology are assessed through a number of simulated examples where different types of changes are considered including drift, set-point changes, a ramp, and a change in the underlying dimensionality of the process. Finally, a case study on a continuous stirred tank reactor (CSTR) process is considered.

2. Recursive Form of the Gaussian Density Function of a Random Vector

In MSPC, measurements obtained during normal operating conditions are regarded as being collected from a stationary Gaussian process, i.e., $\mathbf{x}_t \sim (\boldsymbol{\mu},\boldsymbol{\Sigma})$. The maximum likelihood estimate of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ is given by $\hat{\boldsymbol{\mu}} \equiv \mathbf{m} = t^{-1} \sum_{i=1}^{t} \mathbf{x}_i$ and $\hat{\boldsymbol{\Sigma}} \equiv \mathbf{S} = t^{-1} \sum_{i=1}^{t} (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^{\mathrm{T}}$, respectively, based on t samples.

2.1. Samplewise Updating. When the process operating conditions change either gradually or abruptly, the mean vector and covariance matrix will not be constant and will need to be updated. Each time a sample or a sample block becomes available, both the mean and covariance are updated with the degree of change in the model structure being dependent on the magnitude of the forgetting factors. Consequently, the PCA representation should accordingly be updated to allow the monitoring of such a time-varying process, $\mathbf{x}_t \sim (\boldsymbol{\mu}_t,\boldsymbol{\Sigma}_t)$. The estimated mean vector and covariance matrix at time point t are given in eqs 1 and 2, respectively:[11,12]

$$\mathbf{m}_t = \frac{\sum_{i=1}^{t} \alpha^{t-i}\mathbf{x}_i}{\sum_{i=1}^{t} \alpha^{t-i}} = \frac{\mathbf{x}_t + \alpha\mathbf{x}_{t-1} + \cdots + \alpha^{t-1}\mathbf{x}_1}{1 + \alpha + \cdots + \alpha^{t-1}} \quad (1)$$

and

$$\mathbf{S}_t = \frac{\sum_{i=1}^{t} \beta^{t-i}(\mathbf{x}_i - \mathbf{m}_i)(\mathbf{x}_i - \mathbf{m}_i)^{\mathrm{T}}}{\sum_{i=1}^{t} \beta^{t-i}} \quad (2)$$

The sample mean at time point t can be denoted more simply as a weighted sum of the sample mean at time point $t - 1$, $\mathbf{m}_{t-1}$, and the sample at time point t, $\mathbf{x}_t$:

$$\mathbf{m}_t = \frac{1-\alpha}{1-\alpha^t}\mathbf{x}_t + \alpha\frac{1-\alpha^{t-1}}{1-\alpha^t}\mathbf{m}_{t-1} \quad (3)$$

Likewise, an alternative form of the sample covariance is given by

$$\mathbf{S}_t = \frac{1-\beta}{1-\beta^t}\tilde{\mathbf{x}}_t\tilde{\mathbf{x}}_t^{\mathrm{T}} + \beta\frac{1-\beta^{t-1}}{1-\beta^t}\mathbf{S}_{t-1} \quad (4)$$

where $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \mathbf{m}_t$ is the mean-centered sample vector at time point t. As t becomes large, eqs 3 and 4 can be further simplified to the following forms:

$$\mathbf{m}_t = (1 - \alpha)\mathbf{x}_t + \alpha\mathbf{m}_{t-1} \quad (5)$$

and

$$\mathbf{S}_t = (1 - \beta)\tilde{\mathbf{x}}_t\tilde{\mathbf{x}}_t^{\mathrm{T}} + \beta\mathbf{S}_{t-1} \quad (6)$$

From eq 6, if only the sample variance is being considered, then $\mathbf{D}_t = (1 - \beta)\cdot\mathrm{diag}(\tilde{\mathbf{x}}_t\tilde{\mathbf{x}}_t^{\mathrm{T}}) + \beta\mathbf{D}_{t-1}$, where $\mathbf{D}_t$ is a diagonal matrix whose diagonal elements are identical to those of $\mathbf{S}_t$. Also, the correlation matrix is estimated as follows:

$$\mathbf{R}_t = \frac{\sum_{i=1}^{t} \beta^{t-i}\mathbf{D}_i^{-1/2}\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_i^{\mathrm{T}}\mathbf{D}_i^{-1/2}}{\sum_{i=1}^{t} \beta^{t-i}} \quad (7)$$

As t becomes large, eq 7 can be simplified:

$$\mathbf{R}_t = (1 - \beta)\mathbf{D}_t^{-1/2}\tilde{\mathbf{x}}_t\tilde{\mathbf{x}}_t^{\mathrm{T}}\mathbf{D}_t^{-1/2} + \beta\mathbf{R}_{t-1} \quad (8)$$

2.2. Blockwise Updating. When a process changes slowly and the sampling time is small compared with the process time constant, it is inefficient to update the model each time a new sample becomes available, as described in the previous section, since it is computationally intensive. Instead, the PCA representation can be updated after a block of samples has been collected. The blockwise updating approach has two attractive features: first, it has a low computational cost; and second, it reduces the risk of updating the model based on a false alarm, since it is possible to decide whether a detected alarm is false, prior to updating the model.

The blockwise updating procedure is identical to the samplewise approach except that a sample block is used to adapt the model instead of one sample. The recursive equations for the estimated mean and covariance are given by

$$\mathbf{m}_t = \frac{\sum_{i=1}^{t} \alpha^{t-i}\mathbf{X}_{\mathrm{B}(i)}^{\mathrm{T}}\mathbf{1}_{\mathrm{B}(i)}}{\sum_{i=1}^{t} \alpha^{t-i}} \approx (1 - \alpha)\mathbf{X}_{\mathrm{B}(t)}^{\mathrm{T}}\mathbf{1}_{\mathrm{B}(t)} + \alpha\mathbf{m}_{t-1} \quad (9)$$

and

$$\mathbf{S}_t = \frac{\sum_{i=1}^{t} \beta^{t-i}n_{\mathrm{B}(i)}^{-1}\,\tilde{\mathbf{X}}_{\mathrm{B}(i)}^{\mathrm{T}}\tilde{\mathbf{X}}_{\mathrm{B}(i)}}{\sum_{i=1}^{t} \beta^{t-i}} \approx (1 - \beta)\tilde{\mathbf{X}}_{\mathrm{B}(t)}^{\mathrm{T}}\tilde{\mathrm{X}}_{\mathrm{B}(t)} + \beta\mathbf{S}_{t-1} \quad (10)$$

where $\mathbf{X}_{\mathrm{B}(i)} \in \mathscr{R}^{N_i\times m}$ is the ith sample block (m is the number of measurements), $\mathbf{1}_{\mathrm{B}(i)} = [11 \cdots 1]^{\mathrm{T}} \in \mathscr{R}^{N_i}$, $n_{\mathrm{B}(i)}$ is the sample number in the ith block, and $\tilde{\mathbf{X}}_{\mathrm{B}(i)} = \mathbf{X}_{\mathrm{B}(i)} - \mathbf{1}_{\mathrm{B}(i)}\mathbf{m}_i^{\mathrm{T}}$. The recursive calculation for the sample correlation matrix is given by

$$\mathbf{R}_t = \frac{\sum\sum_{i=1}^{t} \beta^{t-i}n_{\mathrm{B}(i)}^{-1}\,\mathbf{D}_{\mathrm{B}(i)}^{-1/2}\,\tilde{\mathbf{X}}_{\mathrm{B}(i)}^{\mathrm{T}}\tilde{\mathbf{X}}_{\mathrm{B}(i)}\mathbf{D}_{\mathrm{B}(i)}^{-1/2}}{\sum_{i=1}^{t} \beta^{t-i}} \approx (1 - \beta)\mathbf{R}_{\mathrm{B}(t)} + \beta\mathbf{R}_{t-1} \quad (11)$$

where $\mathbf{R}_{B(t)} = \mathbf{D}_{B(t)}^{-1/2}\tilde{\mathbf{X}}_{B(t)}^{T}\tilde{X}_{B(t)}\mathbf{D}_{B(t)}^{-1/2}$ is the correlation matrix of the current sample block and $\mathbf{D}_{B(t)} = (1-\beta)\cdot\mathrm{diag}(\tilde{\mathbf{X}}_{B(t)}^{T}\tilde{X}_{B(t)}) + \beta\mathbf{D}_{B(t-1)}$. $\mathbf{D}_{B(t)}$ is a diagonal matrix whose elements are identical to those of $\mathbf{S}_{B(t)}$.

3. Principal Component Analysis (PCA) Model Updating

When a new sample or block becomes available, the eigenvalues and eigenvectors (loading vectors) of the newly updated correlation matrix are calculated to obtain a new PCA representation. A number of approaches have been proposed to calculate the eigenvalues and eigenvectors. One of the simplest approaches is to perform a singular value decomposition (SVD) on the current correlation matrix. More recently, Li et al.[10] proposed two recursive PCA algorithms based on rank-one modification and Lanczos tridiagonalization, which are computationally more efficient than SVD. The preceding methods require the storage and updating of the correlation matrix to construct an updated PCA representation. When the number of measured variables is fairly large, e.g., over 100, which is common in many industrial processes, the procedure is time-consuming. Consequently, a revised and more efficient algorithm[12] was proposed to update the eigenvalues and eigenvectors, and this idea is extended to the batchwise case.

The blockwise updating procedure is first considered, since samplewise updating is a special case of the blockwise case. On the basis of the mean vector at time point $t-1$, $\mathbf{m}_{t-1}$, the principal loading matrix $\mathbf{P}_{t-1} \in \mathcal{R}^{m\times a}$, and the diagonal matrix of eigenvalues $\mathbf{\Lambda}_{t-1} = \mathrm{diag}(\lambda_{t-1}^{1}, \lambda_{t-1}^{2}, \cdots, \lambda_{t-1}^{a}) \in \mathcal{R}^{a\times a}$ at time point $t-1$, the newly updated values are obtained by using these values and the new sample block $\mathbf{X}_{B(t)}$ at time point t, where a is the approximate rank of $\mathbf{R}_{t-1}$. First, the correlation matrix of $\mathbf{X}_{B(t)}$ is calculated, $\mathbf{R}_{B(t)} = n_{B(t)}^{-1}\mathbf{D}_{B(t)}^{-1/2}\tilde{\mathbf{X}}_{B(t)}^{T}\tilde{\mathbf{X}}_{B(t)}\mathbf{D}_{B(t)}^{-1/2}$, and SVD is applied to obtain an approximation of $\mathbf{R}_{B(t)}$,

$$\mathbf{R}_{B(t)} \approx \mathbf{P}_{B(t)}\mathbf{\Lambda}_{B(t)}\mathbf{P}_{B(t)}^{T} \tag{12}$$

where the columns of $\mathbf{P}_{B(t)} \in \mathcal{R}^{m\times b}$ are the principal eigenvectors, $\Lambda_{B(t)} = \mathrm{diag}(\lambda_{B(t)}^{1}, \lambda_{B(t)}^{2}, \cdots, \lambda_{B(t)}^{b}) \in \mathcal{R}^{b\times b}$ is a diagonal matrix whose elements are the significant eigenvalues of $\mathbf{R}_{B(t)}$, and b is the approximate rank of $\mathbf{R}_{B(t)}$. Likewise, the previous correlation matrix $\mathbf{R}_{t-1}$ can be approximated as

$$\mathbf{R}_{t-1} \approx \mathbf{P}_{t-1}\mathbf{\Lambda}_{t-1}\mathbf{P}_{t-1}^{T} \tag{13}$$

Substituting eqs 12 and 13 into eq 11, $\mathbf{R}_t$ may be approximated as follows:

$$\hat{\mathbf{R}}_t = (1-\beta)\mathbf{P}_{B(t)}\mathbf{\Lambda}_{B(t)}\mathbf{P}_{B(t)}^{T} + \beta\mathbf{P}_{t-1}\mathbf{\Lambda}_{t-1}\mathbf{P}_{t-1}^{T} \tag{14}$$

Combining the two terms on the right-hand side of eq 14, the equation can be rewritten as

$$\hat{\mathbf{R}}_t = [\mathbf{P}_{B(t)}\mathbf{P}_{t-1}]\begin{bmatrix}(1-\beta)\mathbf{\Lambda}_{B(t)} & \mathbf{0}\\ \mathbf{0} & \beta\mathbf{\Lambda}_{t-1}\end{bmatrix}\begin{bmatrix}\mathbf{P}_{B(t)}^{T}\\ \mathbf{P}_{t-1}^{T}\end{bmatrix} = \mathbf{K}_t\mathbf{K}_t^{T} \tag{15}$$

where

$$\mathbf{K}_t = [\sqrt{(1-\beta)\lambda_{B(t)}^{1}}\mathbf{p}_{B(t)}^{1} \cdots \sqrt{(1-\beta)\lambda_{B(t)}^{a}}\mathbf{p}_{B(t)}^{a}\sqrt{\beta\lambda_{t-1}^{1}}\mathbf{p}_{t-1}^{1} \cdots \sqrt{\beta\lambda_{t-1}^{b}}\mathbf{p}_{t-1}^{b}] \in \mathcal{R}^{m\times(a+b)}$$

The updated loading vectors and eigenvalues can be obtained by performing SVD on $\mathbf{K}_t\mathbf{K}_t^{T}$:

$$(\mathbf{K}_t\mathbf{K}_t^{T})\mathbf{p}_t^{i} = \lambda_t^{i}\,\mathbf{p}_t^{i} \tag{16}$$

Alternatively, the loading vectors and eigenvalues can be calculated using $\mathbf{K}_t^{T}\mathbf{K}_t \in \mathcal{R}^{(a+b)\times(a+b)}$ instead of $\mathbf{K}_t\mathbf{K}_t^{T} \in \mathcal{R}^{m\times m}$, i.e., $(\mathbf{K}_t^{T}\mathbf{K}_t)\mathbf{q}_t^{i} = \gamma_t^{i}\,\mathbf{q}_t^{i}$. Multiplying both sides of this relationship by $\delta\mathbf{K}_t$ gives

$$\mathbf{K}_t\mathbf{K}_t^{T}(\delta\mathbf{K}_t\mathbf{q}_t^{i}) = \gamma_t^{i}(\delta\mathbf{K}_t\mathbf{q}_t^{i}) \tag{17}$$

Comparing eq 17 with eq 16, it can be deduced that $\gamma_t^{i} = \lambda_t^{i}$ and $\mathbf{p}_t^{i} = \delta\mathbf{K}_t\mathbf{q}_t^{i}$, where $\delta = (\lambda_t^{i})^{-1/2}$ to ensure $||\mathbf{p}_t^{i}||^2 = ||\delta\mathbf{K}_t\mathbf{q}_t^{i}||^2 = 1$. Samplewise updating is a special case, and hence, $\mathbf{K}_t$ becomes

$$\mathbf{K}_t = [\sqrt{(1-\beta)\lambda_{B(t)}^{1}}\mathbf{p}_{B(t)}^{1} \cdots \sqrt{(1-\beta)\lambda_{B(t)}^{p}}\mathbf{p}_{B(t)}^{a}\sqrt{\beta}\mathbf{D}_t^{-0.5}\tilde{\mathbf{x}}_t] \in \mathcal{R}^{m\times(a+1)} \tag{18}$$

If $m > a+1$ in the samplewise updating procedure or if $m > a+b$ in the blockwise updating procedure, then it is computationally efficient to use $\mathbf{K}_t^{T}\mathbf{K}_t$ instead of $\mathbf{K}_t\mathbf{K}_t^{T}$ to calculate $\mathbf{P}_t$ and Λ_t, because $\mathbf{K}_t^{T}\mathbf{K}_t$ makes the eigenvalue problem more simple and rapid to solve. It should, however, be noted that there are two sources of truncation error in calculating $\mathbf{P}_t$ and Λ_t in the blockwise updating procedure. These materialize from discarding the insignificant eigenvector and eigenvalue pairs in eqs 12 and 13. In the case of the samplewise updating procedure, truncation error may occur when approximating the correlation matrix as defined in eq 13. In the above updating algorithm, the eigenvectors and eigenvalues, which are considered to be insignificant, are discarded to approximate $\mathbf{R}_t$ and $\mathbf{R}_{B(t)}$.

Therefore, the dimensionalities of the two matrices $\mathbf{R}_t$ and $\mathbf{R}_{B(t)}$, a and b, are required to be determined. In each recursive model-updating step, there are several ways to select a and b,[13] but the approaches are dependent on the practical application. Three methods can be considered: (1) fixing the parameter (a or b) as a constant values; (2) retaining those eigenpairs whose eigenvalues are larger than a predefined threshold; or (3) retaining those eigenpairs such that a specified fraction of the total sum of eigenvalues, e.g., 0.99, is included.

4. Adaptive Process Performance Monitoring

4.1. Adaptive Monitoring Statistics. In general, a PCA-based MSPC scheme utilizes two monitoring statistics, Hotelling's T^2,

$$T^2 = \mathbf{t}^{T}\mathbf{\Lambda}_p^{-1}\mathbf{t} = \mathbf{x}^{T}\mathbf{P}^{T}\mathbf{\Lambda}_p^{-1}\mathbf{P}\mathbf{x} \tag{19}$$

where $\Lambda_p \in \mathcal{R}^{p\times p}$ is the leading principal matrix of Λ, and the Q-statistic,

$$Q = \mathbf{x}^{T}(\mathbf{I} - \mathbf{P}\mathbf{P}^{T})\mathbf{x} \tag{20}$$

which are expressed in terms of the Mahalanobis and Euclidean distances, respectively. Hotelling's T^2 represents the distance in the PCA model space, whereas the Q-statistic indicates a distance from the model space. Under the assumption of multivariate normality of the observations and temporal independence, the $100(1-\delta)\%$ control limit for Hotelling's T^2 is calculated by means of a χ^2 approximation, i.e., $T_{\lim}^2 = \chi_\delta^2(p)$, which is generally used in quality control because of its simplicity.[10] In this case, δ denotes the significance level. This method is favorable since it requires only a small computational time in each model-updating step. Also, the $100(1 - \delta)$ % control limit for the Q-statistic is given by

$$Q_{\lim} = \theta_1[z_\delta\sqrt{2\theta_2 h_0^2}/\theta_1 + 1 + \theta_2 h_0(h_0 - 1)/{\theta_1}^2]^{1/h_0} \quad (21)$$

where $\theta_j = \sum_{k=p+1}^{a} \Lambda_{kk}^{j}$ for $j = 1, 2, 3$; $h_0 = 1 - 2\theta_1\theta_3/3\theta_2^2$; and z_δ is the normal deviate.[14] In adaptive PCA monitoring, the two control limits are required to be calculated every time the PCA model is updated, since the number of principal components, p, and the diagonal matrix whose elements are the eigenvalues of the covariance matrix, $\mathbf{\Lambda}$, are time-varying.

4.2. Variable Forgetting Factor. As presented in eqs 1 and 2, calculation of the weighted mean and covariance requires the weighting parameters (termed forgetting factors) α and β to be determined. If both forgetting factors are unity, the mean and covariance are the maximum-likelihood estimates calculated from all the data. By using a forgetting factor that is <1, previous samples are automatically weighted out without deleting data from the process model. As the value gets close to unity, the process has a long-term memory, that is, the number of previous samples that have an effect on the current model increases. To date, most model-updating approaches have used an empirical constant forgetting factor. However, the optimal value of the forgetting factor varies significantly depending on the rate of process change. When the process changes rapidly, the updating rate should be high, whereas when the change is slow, and thus the essential process information is valid for a long period, the updating rate should be small in magnitude. Also, the lack of persistent excitation in the process contributes to the newly recorded samples having insufficient process information;[8] thus, the magnitude of the forgetting factor should be large to prevent the covariance matrix from being inaccurate or unreliable. However, it is likely that the rate of process change or variation in a real process will vary with time. If this is the case, the forgetting factor should be determined according to the underlying objective of the process-monitoring scheme.

To handle this situation, that is, where the rate of process change is not constant, instead of using a fixed forgetting factor, a factor that can be adjusted according to the current process conditions is adopted. Dayal and MacGregor[8] used the algorithm proposed by Fortescue at al. [15] to adjust the forgetting factor in their recursive PLS algorithm. The same concept of the variable forgetting factor has been applied in recursive PCA.[16] In this study, a new algorithm for adapting the forgetting factor is proposed. Fundamentally it differs from the previous algorithm proposed by Fortescue at al.[15] in two aspects. First, the variable forgetting factor used for updating the sample mean and the covariance matrix can take different values, enabling greater flexibility and generality. Second, the two forgetting factors directly depend on the change in the mean and covariance structures, unlike the previous approaches in which both depended on Hotelling's T^2 and the Q-statistic.

In the proposed updating algorithm, two forgetting parameters are used to update the sample mean vector and covariance (or correlation) matrix, respectively. The forgetting factor for updating the mean vector is calculated as

$$\alpha_t = \alpha_{\max} - (\alpha_{\max} - \alpha_{\min})[1 - \exp\{-k(||\Delta\mathbf{m}_{t-1}||/||\Delta\mathbf{m}_{\text{nor}}||)^n\}] \quad (22)$$

where $\alpha_{\max}$ and $\alpha_{\min}$ are the maximum and minimum forgetting values, respectively, k and n are function parameters (defined below), and $||\Delta\mathbf{m}||$ is the Euclidean vector norm of the difference between two consecutive mean vectors. Here, $||\Delta\mathbf{m}_{\text{nor}}||$ is the averaged $||\Delta\mathbf{m}||$ obtained using historical data, as will be explained in Section 4.4. Similarly, the forgetting factor β for updating the covariance (or correlation) matrix is given by

$$\beta_t = \beta_{\max} - (\beta_{\max} - \beta_{\min})[1 - \exp\{-k(||\Delta\mathbf{R}_{t-1}||/||\Delta\mathbf{R}_{\text{nor}}||)^n\}] \quad (23)$$

where $\beta_{\max}$ and $\beta_{\min}$ are the maximum and minimum forgetting values, respectively, and $||\Delta\mathbf{R}||$ is the Euclidean matrix norm of the difference between two consecutive correlation matrices.

There are four function parameters that are required to be determined: $\alpha_{\max}$ (or $\beta_{\max}$), $\alpha_{\min}$ (or $\beta_{\min}$), k, and n. The influence of these four parameters on the forgetting factor value is shown in Figure 1. The change in the patterns of the forgetting factor were investigated according to the magnitude of $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}||$, where $\mathbf{x} = \mathbf{m}$ or $\mathbf{R}$, by varying one of the four parameters and fixing the other three parameters. This procedure was repeated iteratively. The default values of the four function parameters were taken as $\alpha_{\max} = 0.99$, $\alpha_{\min} = 0.90$, $k = 0.6931$, and $n = 1$. Here, the default k value, 0.6931, corresponds to the case where $\alpha = (\alpha_{\max} + \alpha_{\min})/2$ when $n = 1$ and $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}|| = 1$.

Parts a and b of Figure 1 show the effect of the maximum and minimum values, respectively, of α on the forgetting factor. Both $\alpha_{\min}$ and $\alpha_{\max}$ affect the range of the change, i.e., the forgetting factor varies within the range $\alpha_{\min}$ and $\alpha_{\max}$. Also, $\alpha_{\min}$ and $\alpha_{\max}$ regulate the maximum and minimum rate of model adaptation, respectively. The parameters k and n control the sensitivity of the change in α. The larger the value of k, the faster α decreases as $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}||$ increases (Figure 1c), indicating that the model can be updated more rapidly by a change in the mean or covariance if k is large. In Figure 1d, the traces converge at (1, 0.945) since k is set such that $\alpha = (0.99 + 0.90)/2 = 0.945$ when $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}|| = 1$. As n increases, α changes rapidly about the point (1, 0.945). In other words, model adaptation is more rapid resulting in greater sensitivity to $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}||$ about unity. As a limiting case, α becomes a binary variable, $\alpha_{\max}$ and $\alpha_{\min}$, as n becomes sufficiently large. Then, if $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}|| > 1$, α is equal to $\alpha_{\max}$; otherwise, α is equal to $\alpha_{\min}$.

Even though these parameters vary depending on the process characteristics, there needs to be practical guidance to select them; thus, it is proposed that an empirical parameter selection procedure is adopted, as follows:

(1) Select $\alpha_{\max}$ and $\alpha_{\min}$ (typically, $\alpha_{\max} = 0.999-0.99$ and $\alpha_{\min} = 0.95-0.90$).

(2) Determine k such that $\alpha = \mu(\alpha_{\max} - \alpha_{\min}) + \alpha_{\min}$ when $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}|| = 1$, indicating that, when the current mean change is the same as the normal mean change (i.e., as per nominal data), α is set to a value between $\alpha_{\min}$ and $\alpha_{\max}$. For example, $\alpha = 0.5(\alpha_{\max} + \alpha_{\min})$ when $\mu = 0.5$ and $\alpha = 0.67\alpha_{\max} + 0.33\alpha_{\min}$ when $\mu = 0.67$.

(3) Select n from between 1 and 3 by considering the sensitivity of to $||\Delta\mathbf{x}_{t-1}||/||\Delta\mathbf{x}_{\text{nor}}||$. This guidance is equally applicable for the determination of β.

4.3. Robust Adaptation using Outlying Samples. Outlying observations can occur during the normal operation of a process. There are two strategies when dealing with outliers in adaptive modeling: first, outliers can be ignored and a model is held without updating until the next sample or block is considered;

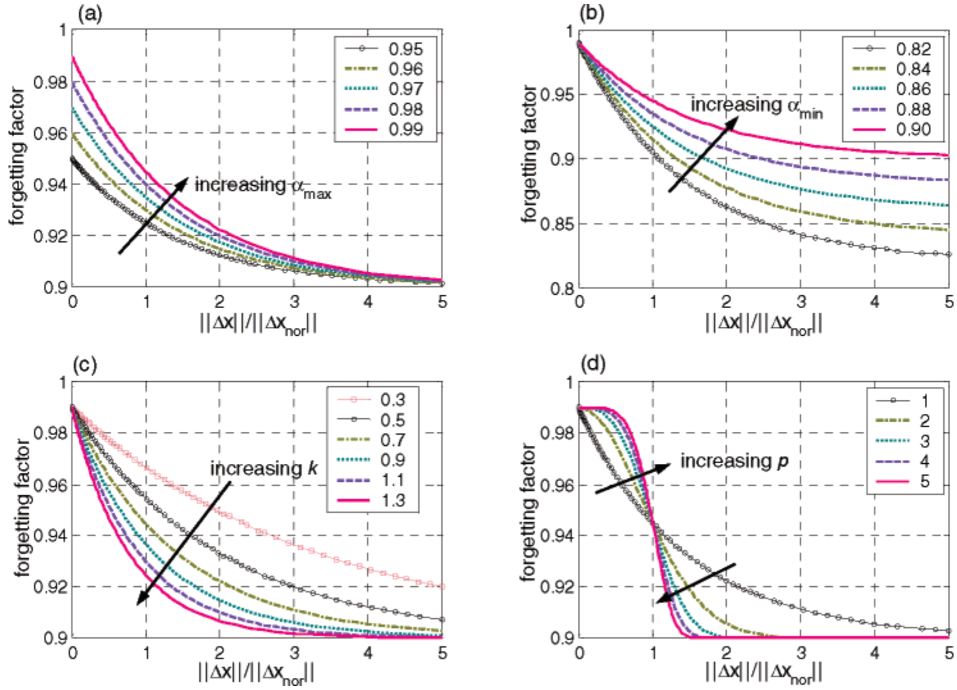

Figure 1. Effect of function parameters on variable forgetting factor: (a) effect of α_{max}, (b) effect of α_{min}, (c) effect of k, and (d) effect of p on α. The default values of the four parameters are $\alpha_{max} = 0.99$, $\alpha_{min} = 0.90$, $k = 0.6931$, and $n = 1$.

and alternatively, a robust parameter estimator can be used to update the model using the outlying samples so that the effect of outlying samples on model updating can be suppressed. One common robust estimator is the M-estimator proposed by Huber.[17] This approach has been used for regression and dimensionality reduction problems.[18,19]

Most robust methods are used to identify a reliable model where a data set contains outliers. Consequently, modeling and outlier identification are performed iteratively until a robust model is identified, since the presence of outliers is unknown a priori. However, in the adaptive modeling approach proposed, a reliable PCA model will have been developed in the previous step and, thus, it may be possible to determine if a new sample is an outlier by examining the monitoring charts. Therefore, the aim is to identify unusual measured variables in the new sample and correct for them. Once an outlying sample, **x**, is detected by using the monitoring charts, its value is replaced with a robust estimate, **z**. Thus, instead of **x**, its robust estimate **z** is used in the calculation of the loading vectors and the eigenvalues to update the PCA model. Here, **z** is represented as a weighted **x**,

$$\mathbf{z} = \mathbf{W}\mathbf{x} \tag{24}$$

where $\mathbf{W} = \mathrm{diag}(w_1, w_2, \cdots, w_m)$ is a weight matrix. The weight of each variable is calculated according to its reliability, which is a function of the variable residual obtained using the current model,

$$w_k = \left(1 + \frac{(x_k - \hat{x}_k)^2}{c_k^2}\right)^{-1} \tag{25}$$

where c_k is a tuning parameter, which controls the sharpness of the weight function w_k. Considering the covariance of the residual $\mathbf{e}_t$ at time point t, i.e., $\Sigma_{\mathbf{e}_t} = (\mathbf{I} - \mathbf{P}_t\mathbf{P}_t^{\mathrm{T}})\Sigma_{\mathbf{x}}(\mathbf{I} - \mathbf{P}_t\mathbf{P}_t^{\mathrm{T}}) = (\mathbf{I} - \mathbf{P}_t\mathbf{P}_t^{\mathrm{T}})$, and assuming $\Sigma_{\mathbf{x}} = \mathbf{I}$, c_k can be calculated as $c_k = z_\delta(1 - \sum_{i=1}^{p} p_{ki}^2)$, where z_δ is a normal deviate. As the variable residuals e_k increase, the weight w_k decreases and, thus, the difference between the original and corrected values increases.

Outlying samples may trigger a warning signal in the monitoring chart in a manner similar to that of a fault. Therefore, there is a risk that a current model adapts to a fault or a disturbance as well as an outlier when using a robust updating method. Hence, it is important to discriminate between outliers and process faults. Outliers are unlikely to be generated consecutively, whereas faults and disturbances last for a certain period. In this respect, if a monitoring chart shows out-of-control signals for several consecutive samples, e.g., three samples, a process is considered to be abnormal. Otherwise, the current model is updated using the robust updating method.

4.4. Overview of the Monitoring Procedure. The overall strategy for time-varying monitoring using the proposed method is as follows:

Off-line Learning.

(1) Obtain the initial values of the sample mean $\mathbf{m}_0$, the variance $\mathbf{D}_0$, the loading matrix $\mathbf{P}_0$, and the eigenvalue matrix $\mathbf{\Lambda}_0$ based on the training data block.

(2) Obtain the two control limits, $T_{\mathrm{lim},0}^2$ and $Q_{\mathrm{lim},0}$, and the number of principal components, p_0.

(3) Select α_{max}, α_{min}, β_{max}, and β_{max}.

(4) Calculate $||\Delta\mathbf{m}_{nor}||$ and $||\Delta\mathbf{R}_{nor}||$. Divide the training data into two sections and calculate the initial **m**, **D**, and **R** using the first training data set. Then, with a fixed α and β (e.g., set both to 0.96, which corresponds to a value of $\alpha = 0.67\alpha_{max} + 0.33\alpha_{min}$, when $\alpha_{max} = 0.99$ and $\alpha_{min} = 0.90$), update **m** (eq 5), **D** (eq 6), and **R** (eq 7) and store $||\Delta\mathbf{m}||$ and $||\Delta\mathbf{R}||$ using the second training data set. Calculate the average values $||\Delta\mathbf{m}_{nor}||$ and $||\Delta\mathbf{R}_{nor}||$.

(5) Determine the model parameters: n and k.

(6) Determine the initial values of the two forgetting factors α_0 and β_0 (Here, both are set to 0.99). In effect, they are deemed not to be critical in on-line model adaptation.

On-line Learning and Monitoring. At time point t, using the previous values of $\mathbf{m}_{t-1}$, $\mathbf{D}_{t-1}$, $\mathbf{P}_{t-1}$, Λ_{t-1}, $T_{\mathrm{lim},t-1}^2$, $Q_{\mathrm{lim},t-1}$, and p_{t-1}:

Table 1. Four Types of Process Changes Introduced into the Simulation Study

	patterns		scenario
P1	$x_1(t) = x_1^*(t) + 0.01(t - 300)$	$t > 300$	change in the mean until the end of simulation time
	$x_{11}(t) = x_{11}^*(t) - 0.015(t - 500)$	$t > 500$	
	$x_{21}(t) = x_{21}^*(t) + 0.02(t - 700)$	$t > 700$	
P2	$x_1(t) = x_1^*(t) + 2\exp(-(t - 100)/100)$	$t > 300$	change in the mean
	$x_{11}(t) = x_{11}^*(t) - 2\exp(-(t - 700)/30)$	$t > 700$	
P3	$\Xi'_{11} = \Xi_{11} + 0.01(t - 100)$	$t > 100$	change in the covariance
P4	$\Xi \in R^{30\times 10} \rightarrow \Xi'' \in R^{30\times 14}$	$t > 500$	change in the number of PCs.

(1) Calculate T_t^2 and Q_t for a new sample $\mathbf{x}_t$ (or block) after mean centering and autoscaling using $\mathbf{m}_{t-1}$ and $\mathbf{D}_{t-1}$, eqs 19 and 20, respectively.

(2) If $T_t^2 > T_{\lim,t-1}^2$ or $Q_t > Q_{\lim,t-1}$, go to step 3; otherwise, go to step 6.

(3) Check if the new sample is an outlier either automatically or manually. For instance, if $T_{t-s}^2 > T_{\lim,t-s-1}^2$ or $Q_{t-s} > Q_{\lim,t-s-1}$ ($s = 1, 2$) (i.e., if three consecutive out-of control signals have been generated), the new sample is not an outlier. Otherwise, it is an outlying sample.

(4) If it is an outlier, go to step 5. Otherwise, consider the current condition to be abnormal. Typically, the current process condition cannot be determined by examining one sample. In this case, the model is retained without updating and the current sample is stored. Then, if the process condition is proven to be normal subsequently, the model can be updated in a blockwise manner.

(5) Calculate the robust estimated value, $\mathbf{z}_t$, of the new sample, eqs 24 and 25.

(6) Recalculate T_t^2 and Q_t based on $\mathbf{z}_t$, eqs 19 and 20, respectively.

(7) Calculate $\mathbf{m}_t$ and $\mathbf{D}_t$, eqs 5 and 6 (or eqs 9 and 10), respectively.

(8) Calculate $\mathbf{P}_t$ and Λ_t by performing SVD on $\mathbf{K}_t^{\mathrm{T}}\mathbf{K}_t$ or $\mathbf{K}_t\mathbf{K}_t^{\mathrm{T}}$ (eq 18).

(9) Identify the number of principal components to retain, p_t, and calculate the control limits, $T_{\lim,t}^2$ and $Q_{\lim,t}$.

5. Application Studies

5.1. Illustrative Examples. Four examples are considered to illustrate the updating procedures proposed in the paper. It should be noted that the following scenarios are assumed to be reflective of normal process changes.

The following multivariate static process is considered: $\mathbf{x} = \Xi$t, where $\mathbf{t} = [t_1,t_2, \cdots,t_{10}]^{\mathrm{T}} \in \mathscr{R}^{10}$ is a Gaussian random vector and $\Xi \in \mathscr{R}^{30\times 10}$ is a transformation matrix. Here, $t_i \sim (0, 0.1^2)$ and $\Xi_{ij} \sim (0, 1)$. Therefore, the measurement vector $\mathbf{x} \in \mathscr{R}^{30}$ follows a normal distribution with $\mu_{\mathbf{x}} = \mathbf{0}$ and $\Sigma_{\mathbf{x}} = 0.01\cdot\Xi\Xi^{\mathrm{T}}$. Two-hundred samples were generated that were reflective of normal process operation. Four types of process changes were simulated, each comprising 1000 samples. Table 1 summarizes these process changes and the conditions implemented to simulate these changes.

In this study, a blockwise adaptive PCA modeling approach was adopted with a block size of five samples. That is, the PCA model was updated every five time points. For the forgetting factors to be adjustable, the following parameters, $\alpha_{\max} = \beta_{\max} = 0.99$, $\alpha_{\min} = \beta_{\min} = 0.9$, $k = 0.4055$, and $n = 1$, were selected. By setting $k = 0.4055$, if the mean change is equal to the average values obtained from the training data set, then the forgetting factor takes the value 0.96.

The first step was to develop the baseline model using the off-line learning procedure described previously. The 200 training samples were utilized to identify the initial PCA model. The first 100 samples were used to identify the initial mean and covariance structures, and the remaining 100 samples were used for the determination of the parameters for the forgetting factor updating (eqs 22 and 23). After identifying the model and the parameters, the procedure for on-line monitoring described in Section 4.4 was applied to the samples taken from the slowly drifting process.

In the first case study (P1), three measured variables, x_1, x_{11}, and x_{21}, slowly drift away from the normal operating conditions, thereby simulating a slow process change such as catalyst deactivation. The slow drifts of magnitude 0.01, -0.015, and 0.02 are introduced to the three variables at time points $t =$ 301, 501, and 701, respectively, and are continued until the end of the simulation. As shown in Figure 2 parts a−c, the change in the mean of variables x_1, x_{11}, and x_{21} is tracked closely by the model-updating procedure with some time delay. During the simulation, the number of principal components is fixed at eight, since the rank of the covariance matrix does not change over time. The forgetting factor α for the sample mean $\mathbf{m}$ continues to decrease from time point $t = 301$, with it taking a constant low value of 0.92 from time point $t = 700$ until the end of the simulation; hence, the model is updated rapidly according to the mean changes in x_1, x_{11}, and x_{21}. On the other hand, the forgetting factor β fluctuates around 0.96 with no distinct changes being evident, since the sample covariance matrix does not change significantly during this period (Figure 2d).

Hotelling's T^2 and the Q-statistic for adaptive PCA and static PCA are compared in Figure 3. The PCA model without updating is no longer valid after the process begins to change at time point $t = 301$, and thus, it results in the process appearing to be out of control in both monitoring charts. In contrast, the monitoring charts for the adaptive PCA model show that the process remains in control after the changes have been introduced.

The second case study (P2) reflects two types of changes in two of the measured variables. Both describe set-point changes in the process by assuming that the two variables x_1 and x_{11} are measured controlled outputs. The set point of x_1 changes from 0 to +2 with first-order dynamics and a time constant of 100 sampling time points at time point $t = 101$, while that of x_{11} changes from 0 to -2 with first-order dynamics and a time constant of 30 sampling time points at time point $t = 701$. Parts a and b of Figure 4 show the actual values and updated means of variables x_1 and x_{11}, respectively, with Figure 4c capturing the difference between the theoretical and estimated means. As time progresses after a process change has occurred, the updated sample mean approaches its theoretical value.

Similar to the first case, the number of principal components does not change, since the underlying dimensionality of the process is constant. As illustrated in Figure 4d, the value of the forgetting factor α decreases sharply at around time points $t =$ 100 and $t = 500$ in order to update the model according to the

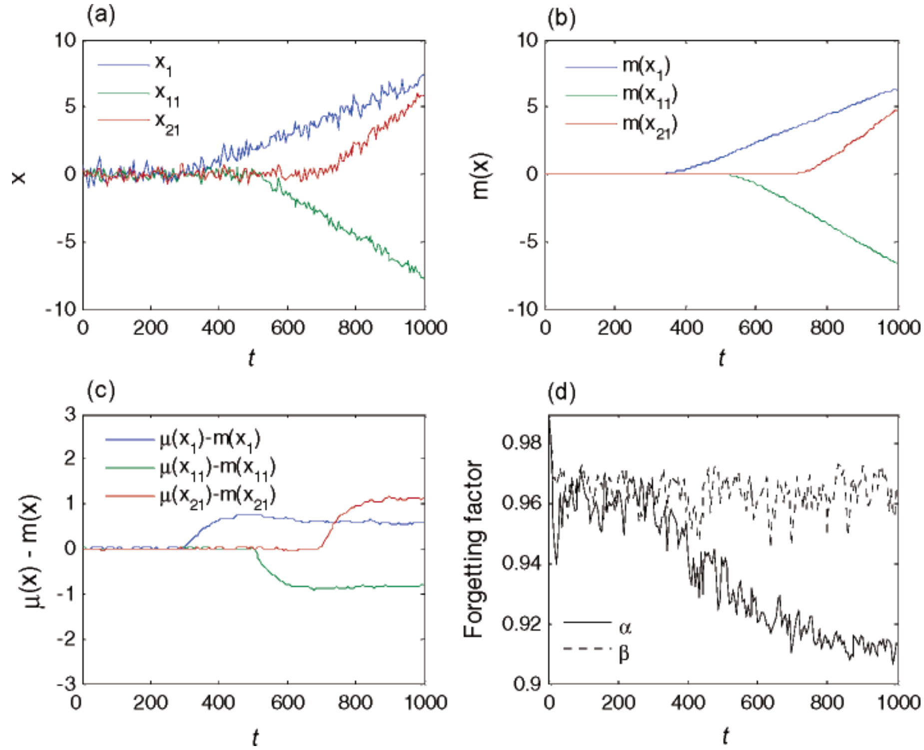

Figure 2. Time-series plots of variable mean and model parameters for case P1: (a) measured variables x_1, x_{11}, and x_{21}; (b) estimated mean $m(x)$; (c) difference between theoretical and estimated mean values $\mu(x) - m(x)$; and (d) forgetting factors α and β.

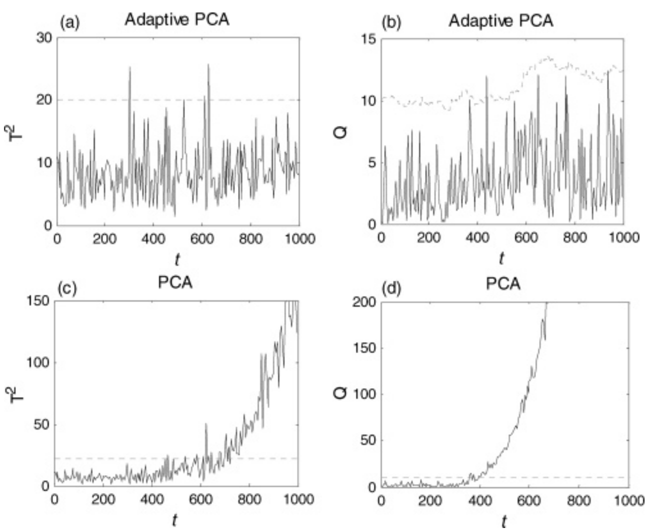

Figure 3. Process-monitoring charts for case P1: (a) Hotelling's T^2 for adaptive PCA, (b) Q-statistic for adaptive PCA, (c) Hotelling's T^2 for PCA, and (d) Q-statistic for PCA.

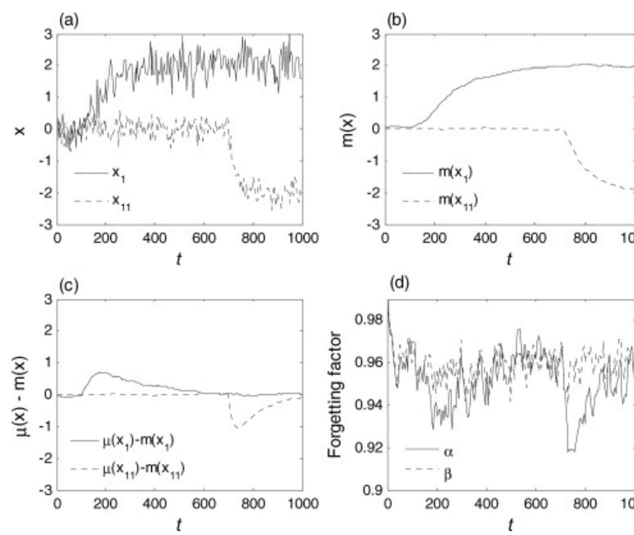

Figure 4. Time-series plots of variable mean and model parameters for case P2: (a) measured variables x_1 and x_{11}; (b) estimated mean $m(x)$; (c) difference between theoretical and estimated mean values $\mu(x) - m(x)$; and (d) forgetting factors α and β.

change in the mean values, whereas the value of the forgetting factor, β, fluctuates around 0.96 without any significant change being incurred. As shown in Figure 5, the Q-statistic for the adaptive PCA model does not show any abnormal patterns indicating out-of-control behavior, unlike that for the static PCA model. It can be observed that Hotelling's T^2 does not pick up the change, since the direction of this change is nearly orthogonal to the model subspace, and thus, the process change does not significantly increase Hotelling's T^2 (see Figure 5c).

In the third case study (P3), a ramp signal is introduced with magnitude 0.01 to the (1, 1) elements of the transformation matrix from time point $t = 101$ until the end of the simulation, i.e., $\Xi'_{11} = \Xi_{11} + 0.01(t - 100)$. As expected, the covariance of $\mathbf{x}$ changes with time after the change in the transformation matrix has been introduced. Parts a and b of Figure 6 represent the theoretical and updated variances of x_1, respectively, and Figure 6c represents the difference between them. As for the first two case studies, the number of principal components remains constant during the simulation. The value of the forgetting factor, α, does not show any significant change, since the parameter change in Ξ does not affect the mean of the variables. In contrast, the forgetting factor, β, is slightly lower than 0.96; thus, the covariance matrix is continuously updated after time point $t = 100$ (Figure 6d). Even though Hotelling's T^2 for adaptive PCA shows several false alarms, whose frequency is statistically acceptable, the adaptive PCA model gives reliable adaptive monitoring charts for both Hotelling's T^2 and the

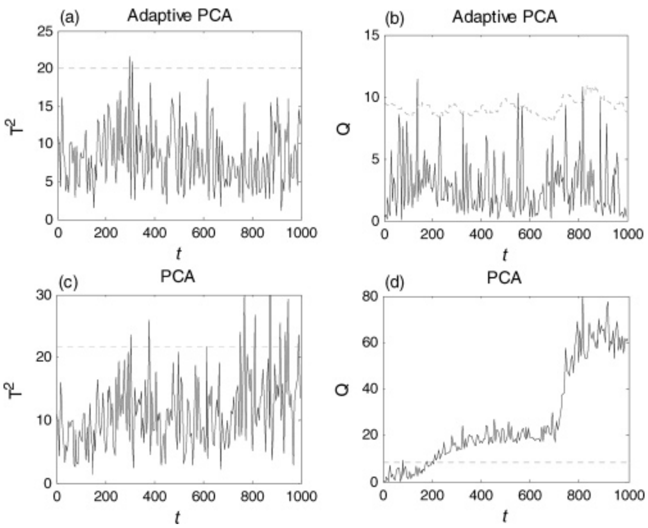

Figure 5. Process-monitoring charts for case P2: (a) Hotelling's T^2 for adaptive PCA, (b) Q-statistic for adaptive PCA, (c) Hotelling's T^2 for PCA, and (d) Q-statistic for PCA.

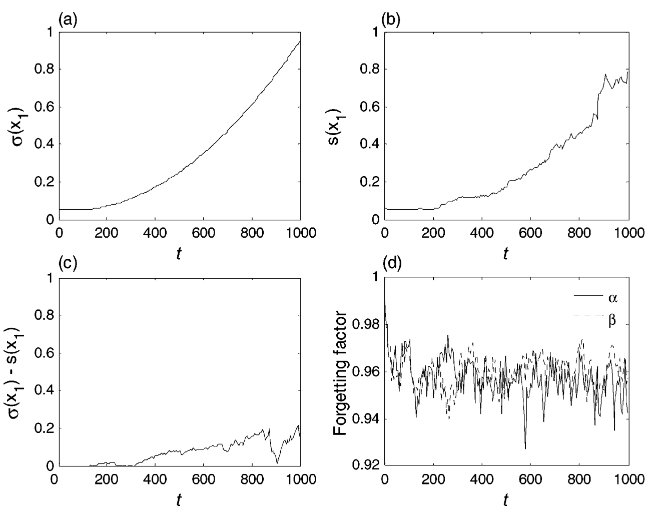

Figure 6. Time-series plots of variable variance and model parameters for case P3: (a) Theoretical standard deviation of variables x_1 and $\sigma(x_1)$; (b) estimated standard deviation $s(x_1)$; (c) difference between theoretical and estimated standard deviation $\sigma(x_1) - s(x_1)$; and (d) forgetting factors α and β.

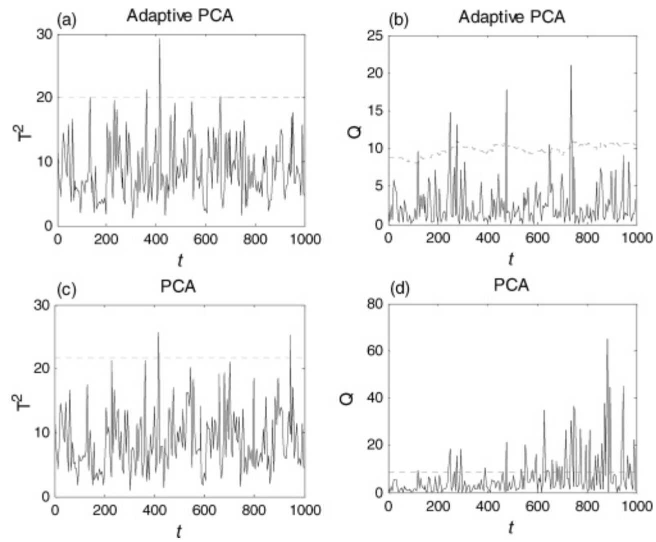

Figure 7. Process-monitoring charts for case P3: (a) Hotelling's T^2 for adaptive PCA, (b) Q-statistic for adaptive PCA, (c) Hotelling's T^2 for PCA, and (d) Q-statistic for PCA.

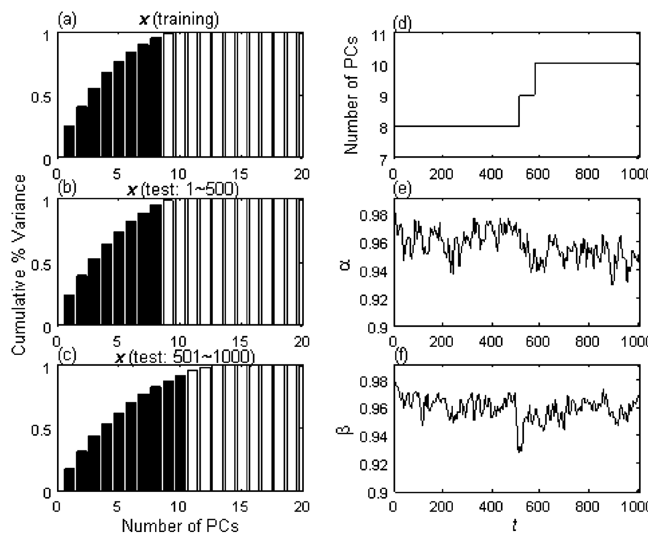

Figure 8. Cumulative percent variance (CPV) captured by PCA model and the model parameters for case P4: (a) CPV captured by the PCA model for the training data, (b) CPV captured by the PCA model for the test data for time points $t = 1-500$, (c) CPV captured by the PCA model for the test data for time points $t = 501-1000$, (d) number of PCs retained for adaptive PCA, (e) forgetting factor α, and (f) forgetting factor β.

Q-statistic, for the process showing the slow covariance change compared with the results for the static PCA model (Figure 7).

The last study (P4) simulates a change in the underlying dimensionality of the process, which gives rise to a change in the number of principal components retained in the PCA model. At time point $t = 501$, the dimensionality of **t**, the principal components, changes from 10 to 14, and thus, the rank of Ξ also increases by the order of four. Parts a−c of Figure 8 show the cumulative percentage variance captured by the PCA model constructed based on the training data set, the first 500 test data points collected before the process change, and the final 500 test data points taken after the process change, respectively. In each figure, the number of black bars indicates the number of principal components retained, which is selected as that when the cumulative percentage variance attains a value larger than 90%. Before the process change, the number of principal components for the adaptive PCA model is the same as that for the PCA model obtained using the training data. However, after the change in the underlying dimensionality of the process, the number of principal components retained increases by two and, thus, becomes 10, as observed in Figure 8d. As indicated in Figure 8f, β rapidly decreases to 0.93 at time point $t = 501$ to update the model as a result of this abrupt process change.

The monitoring results with adaptive PCA are illustrated in Figure 9 parts a and b, and those with static PCA are displayed in Figure 9 parts c and d. The chart for Hotelling's T^2 for adaptive PCA has a smaller false alarm rate than that for static PCA. It can be observed that the control limit in this chart increases as the number of principal components retained in the PCA model increases. The Q-statistic without model updating produces out-of-control signals just after time point $t = 501$, whereas that with model updating is still valid after that time, except that there are some false alarms just after the process changes. This is because the model-updating procedure requires time until sufficient samples are collected from the newly changed process to describe the new condition.

5.2. Simulated CSTR Process. A nonisothermal CSTR process[20,21] is considered for the application of the adaptive

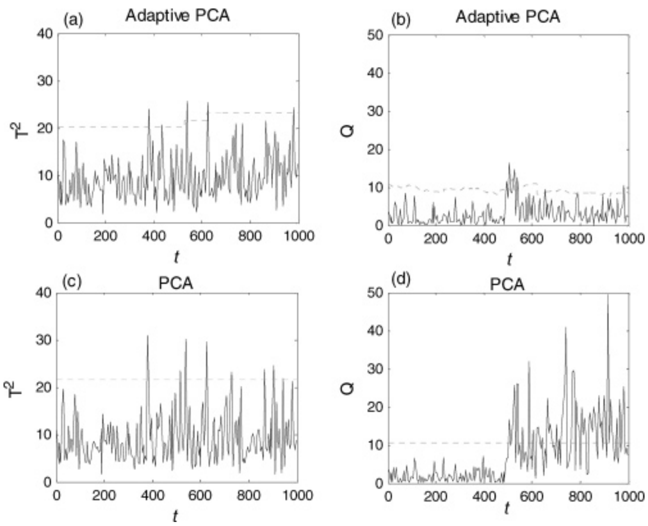

Figure 9. Process-monitoring charts for case P4: (a) Hotelling's T^2 for adaptive PCA, (b) Q-statistic for adaptive PCA, (c) Hotelling's T^2 for PCA, and (d) Q-statistic for PCA.

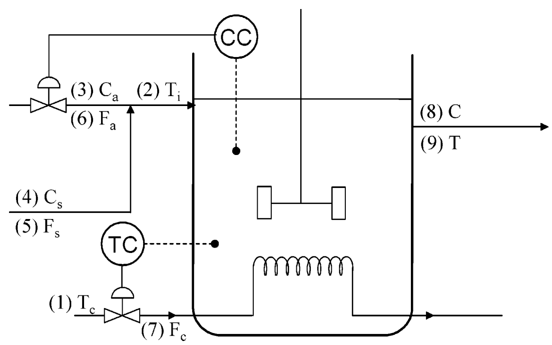

Figure 10. Nonisothermal CSTR process and nine measured variables: (1) coolant temperature, (2) reactant mixture temperature, (3) reactant A concentration, (4) solvent concentration, (5) solvent flow rate, (6) solute flow rate, (7) coolant flow rate, (8) outlet concentration, and (9) outlet temperature.

process-monitoring algorithm. A schematic of the process is given in Figure 10. In the reactor, reactant A was premixed with a solvent, and then it was converted into product B with rate $r = \beta_r k_0\ e^{-E/RT}C$. The dynamic behavior of the process is described by the mass balance of reactant A and the total energy balance for the reacting system,

$$V\frac{dC}{dt} = F(C - C_i) - Vr \tag{26}$$

$$V\rho C_p\frac{dT}{dt} = \rho C_p F(T_i - T) - \frac{UA}{1 + UA/2F_c\rho_c C_{pc}}(T - T_c) + (-\Delta H_r)Vr \tag{27}$$

where the heat transfer coefficient UA is empirically represented as $UA = \beta_{UA} a F_c^b$. The concentration of the reactant mixture is calculated from $C_i = (F_a C_a + F_s C_s)/(F_a + F_s)$. The outlet temperature (T) and the outlet concentration (C) are controlled by manipulating the inlet coolant flow rate (F_c) and the inlet reactant flow rate (F_a), respectively. The model parameters used in the CSTR modeling and the controller parameters of the two PI controllers are given in Table 2. All process inputs and disturbances are generated using first-order autoregressive models or sinusoidal signals.[21] In addition, all measured variables are contaminated with white Gaussian noise to describe measurement noise (Table 3).

Table 2. . Model Parameters Used in the CSTR Modeling and the PI Control Parameters

notation	parameters and constants	value
V	volume of reaction mixture in the tank	1 m^3
ρ	density of reaction mixture	10^6 g/m^3
ρ_c	density of coolant	10^6 g/m^3
C_p	specific heat capacity of the rxn mixture	1 cal/(g K)
C_{pc}	specific heat capacity of the coolant	1 cal/(g K)
ΔH_r	heat of reaction	-1.3×10^7 cal/kmol
k_0	preexponential kinetic constant	10^{10} min^{-1}
E/R	activation energy/ideal gas constant	8330 K
$K_c(T)$	temperature controller gain	-1.5
$\tau_I(T)$	integral time	5
$K_c(C)$	concentration controller gain	0.4825
$\tau_I(C)$	integral time	2

Table 3. Patterns of Process Inputs, Disturbances, and Measurement Noise[a]

input and disturbance				measurements		
variable	mean	ϕ	σ_v^2	variable		σ_e^2
F_s	0.9 m^3/min	0.1	1.90×10^{-2}	1	T_c	2.5×10^{-3}
T_i	370 K	0.2	4.75×10^{-2}	2	T_i	2.5×10^{-3}
T_c	365 K	0.2	4.75×10^{-2}	3	C_a	1.0×10^{-2}
β_r	1	0.9	1.90×10^{-3}	4	C_s	2.5×10^{-5}
β_{UA}	1	0.95	9.75×10^{-4}	5	F_s	4.0×10^{-6}
				6	F_a	4.0×10^{-6}
				7	F_c	1.0×10^{-2}
				8	C	2.5×10^{-5}
				9	T	4.0×10^{-4}

variable	mean	a	b	σ_v^2
C_a	19.1 kmol/m^3	2	0.25	4.75×10^{-2}
C_s	0.3 kmol/m^3	0.2	0.2	1.875×10^{-3}

[a] C_a and C_s are modeled as $x(t) = a \sin(bt) + v(t)$, and the remaining inputs and disturbances are modeled as $x(t) = \phi x(t - 1) + v(t)$, where the process noise $v(t) \approx (0, \sigma_v^2)$. All measured variables are contaminated with white gaussian noise $e(t) \approx (0, \sigma_e^2)$.

To evaluate the performance of the new adaptive method, two kinds of process change are considered: (1) a slow decrease in the reaction rate and (2) a set-point change in the reactor temperature. For each simulation, the CSTR process was run for 500 min, with the process change being introduced at time point $t = 301$ min. Nine process variables were measured every minute (Table 3). The samples obtained for time points $t = 100-300$ min were used to build an initial PCA model, and the subsequent 200 samples ($t = 301-500$ min) were used to update the model on-line.

In the first case, a slow drift in β_r of magnitude 10^{-3} min^{-1} was introduced at time point $t = 301$ min and continued until the end of the simulation. The decrease in the reaction rate results in a decrease in the reactant flow rate to maintain the required outlet concentration, and the decrease in the coolant flow rate will maintain the outlet temperature. Parts a and b of Figure 11 represent the actual values and their estimated mean values, through the adaptive method, of F_a and F_c, respectively. The proposed adaptive scheme was able to closely track the time trajectory of the variables with a time delay of 25 min. During the drift, the forgetting factor α decreases from 0.96 to ~0.93 to take account of the mean change in the process (Figure 11d). The process-monitoring methods based on static PCA and two adaptive PCA models with and without outlier correction are compared in Figure 12. The static PCA model becomes invalid just after the process change occurs (Figure 12 parts a and d), whereas the adaptive PCA model adapts the Hotelling's T^2 and Q-statistics and their control limits, making the current model valid. There are several outliers caused by the measurement error, as depicted in Figure 12 parts b and e. The presence

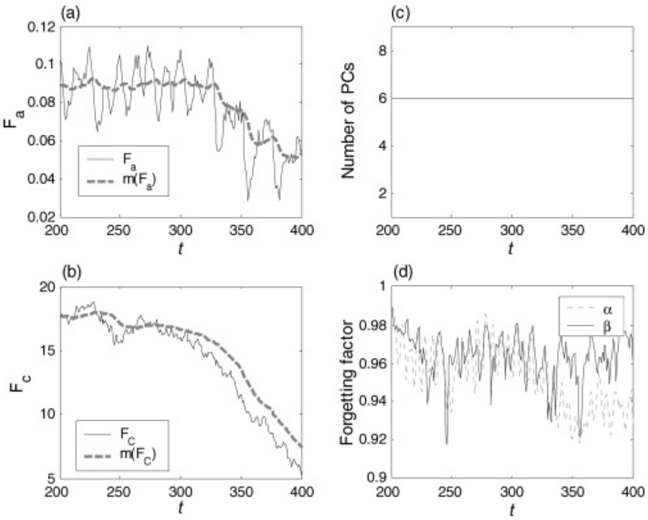

Figure 11. Time-series plots of variable mean and model parameters for the first case study (reaction rate decrease) in the CSTR process: (a) measured value of F_a and its estimated mean, (b) measured value of F_c and its estimated mean, (c) number of PCs for adaptive PCA, and (d) forgetting factors.

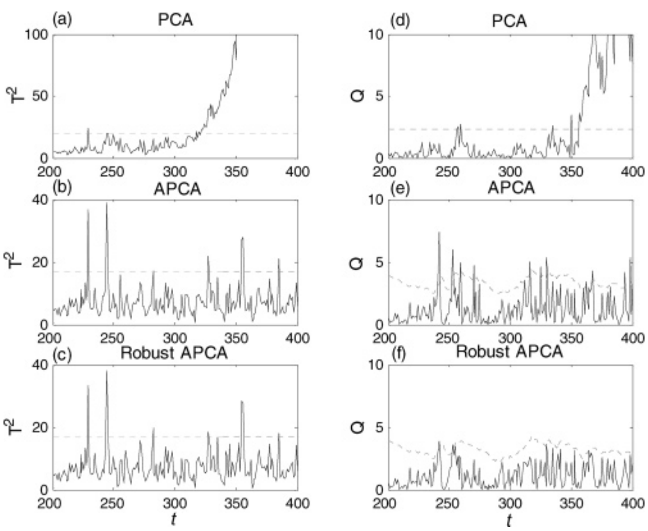

Figure 12. Process-monitoring charts for PCA, adaptive PCA, and robust adaptive PCA. Hotelling's T^2 for (a) PCA, (b) adaptive PCA, and (c) robust adaptive PCA; and Q-statistic for (d) PCA, (e) adaptive PCA, and (f) robust adaptive PCA.

of outliers deteriorates the performance of the model-updating procedure. By using the proposed robust updating methodology when an outlier occurs, the false alarm rate in the monitoring charts could be significantly reduced, especially with respect to the Q-statistic (Figure 12 parts c and f).

Unlike the first simulation study, the second process change caused by the set-point change is very fast and predictable. Since the time of the process change is known, it may be straightforward to discriminate between the alarms from the process change and those from outliers or faults. In this case, the set-point of the outlet temperature changed from 368 to 370 at time point $t = 301$ min. The sample mean values of F_c and T calculated using the recursive scheme are illustrated in parts a and b of Figure 13, respectively. The number of principal components for the adaptive model is six, except for several time points where it increases to seven (Figure 13c). As shown in Figure 13d, the forgetting factor, α, drops rapidly from 0.96 to 0.91 just after the set-point change and then returns to its normal value (0.96) after time point $t = 350$ because the estimated mean values have converged to their new values.

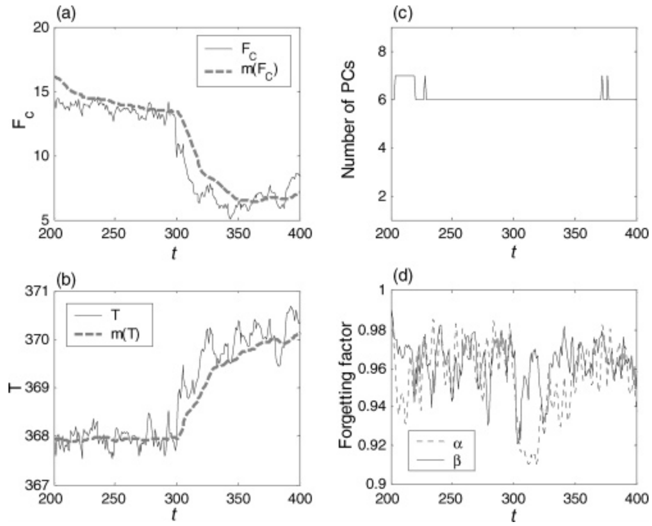

Figure 13. Time-series plots of variable mean and model parameters in the second case study (set-point change) in the CSTR process: (a) measured value of F_c and its estimated mean, (b) measured value of T and its estimated mean, (c) number of PCs for adaptive PCA, and (d) forgetting factors.

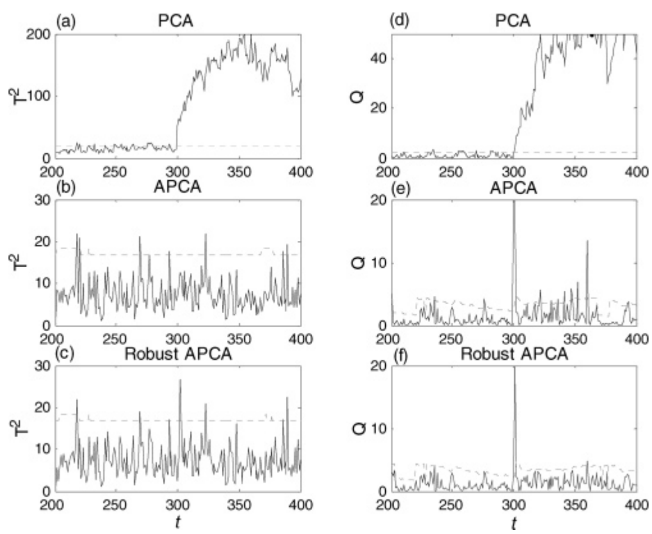

Figure 14. Process-monitoring charts for PCA, adaptive PCA, and robust adaptive PCA. Hotelling's T^2 for (a) PCA, (b) adaptive PCA, and (c) robust adaptive PCA; and Q-statistic for (d) PCA, (e) adaptive PCA, and (f) robust adaptive PCA.

The monitoring results are shown in Figure 14. For time points $t = 301-305$ min, the Q-statistic for all the methods shows out-of-control signals due to the set-point change. With the exception of this time period, the adaptive PCA model combined with the outlier correction method gave the most reliable monitoring results (Figure 14f), making it possible to suppress the effect of outliers on the updating of the model and the control limits of the monitoring statistics.

6. Conclusions

A new adaptive MSPC method for time-varying process monitoring has been proposed. On the basis of the newly defined weighted mean and covariance, recursive forms of the mean and covariance are defined in a samplewise and blockwise manner. Then, by using these recursive formulas, an efficient method for finding a PCA model recursively is developed. In this approach, a loading matrix is stored instead of a full covariance matrix for the next model update, which means that this method has lower storage requirements compared with the previous approaches described in the literature. Furthermore, several important issues for adaptive process monitoring are addressed. First, two different variable forgetting factors for mean and covariance updating are considered, and their calculation based on the recent mean or covariance change is proposed. Second, outlying samples, which are recorded quite often in real processes, are pretreated by means of a robust estimator such that the model is insensitive to the outliers and, thus, is correctly updated.

On the basis of the overall process-monitoring strategy presented in this study, two application examples were considered: a simple static multivariate system and a CSTR process. The ability of the adaptive model to keep track of the changes in the variable mean and covariance was investigated. Also, a change in the number of principal components and the forgetting factors according to process changes were considered. Charts of Hotelling's T^2 and the Q-statistic based on the adaptive monitoring statistics and their control limits were observed to be reliable for the monitoring of both slowly and abruptly changing processes. In particular, the correction for outliers using the robust estimation method results in not only reducing the number of false alarms significantly but also preventing the current model from being updated by outliers.

Acknowledgment

S.W.C. acknowledges the financial contribution of the Centre for Process Analytics and Control Technology.

Literature Cited

(1) Kresta, J.; MacGregor, J. F.; Marlin, T. E. Multivariate statistical monitoring of process operating performance. *Can. J. Chem. Eng.* **1991**, *69*, 35−47.

(2) Wise, B. M.; Gallagher, N. B. The process chemometrics approach to process monitoring and fault detection. *J. Process Control* **1996**, *6*, 329−348.

(3) MacGregor, J. F.; Kourti, T. Statistical process control of multivariate processes. *Control Eng. Pract.* **1995**, *3* (4), 403−414.

(4) Tracy, N. D.; Young, J. C.; Mason, R. L. Multivariate control charts for individual observations. *J. Qual. Technol.* **1992**, *24* (2), 88−95.

(5) Jackson, J. E. *A user's guide to principal components*; John Wiley & Sons: New York, 1991.

(6) Miller, P.; Swanson, R.; Heckler, C. Contribution Plots: The missing link in multivariate quality control. Presented at Fall Conference of the ASQC and ASA, Milwaukee, WI, 1993.

(7) Wold, S. Exponentially weighted moving principal components analysis and projections to latent structures. *Chemom. Intell. Lab. Syst.* **1994**, *23*, 149−161.

(8) Dayal, B. S.; MacGregor, J. F. Recursive exponentially weighted PLS and its applications to adaptive control and prediction. *J. Process Control* **1997**, *7* (3), 169−179.

(9) Qin, S. J. Recursive PLS algorithms for adaptive data monitoring. *Comput. Chem. Eng.* **1998**, *22*, 503−514.

(10) Li, W.; Yue, H. H.; Valle-Cervantes, S.; Qin, S. J. Recursive PCA for adaptive process monitoring. *J. Process Control* **2000**, *10*, 471−486.

(11) Feltz, C. J.; Shiau, J.-J. H. Statistical process monitoring using an empirical bayes multivariate process control chart. *Qual. Reliab. Eng. Int.* **2001**, *17*, 119−124.

(12) Liu, X.; Chen, T.; Thornton, S. M. Eigenspace updating for nonstationary process and its application to face recognition. *Pattern Recognit.* **2003**, *36*, 1945−1959.

(13) Hall, P.; Marshall, D.; Martin, R. Merging and splitting Eigenspace models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22* (9), 1042−1049.

(14) Jackson, J. E.; Mudholkar, G. Control procedures for residuals associated with principal component analysis. *Technometrics* **1979**, *21*, 341−349.

(15) Fortescue, T. R.; Kershenbaum, L. S.; Ydstie, B. E. Implementation of self-tuning regulators with variable forgetting factors. *Automatica* **1981**, *17*, 831−835.

(16) Lane, S.; Martin, E. B.; Morris, A. J.; Gower, P. Application of exponentially weighted principal component analysis for the monitoring of a polymer film manufacturing process. *Trans. Inst. Meas. Control* **2003**, *25* (1), 17−35.

(17) Huber, P. J. A robust estimation of a location parameter. *Ann. Math. Stat.* **1964**, *35*, 135−139.

(18) Huber, P. J. *Robust statistics*; John Wiley & Sons: New York, 1981.

(19) Liang, Y.-Z.; Kvalheim, O. M. Robust methods for multivariate analysis−A tutorial review. *Chemom. Intell. Lab. Syst.* **1996**, *32*, 1−10.

(20) Yoon, S.; MacGregor, J. F. Fault diagnosis with multivariate statistical models. Part I: Using steady-state fault signatures. *J. Process Control* **2001**, *11*, 387−400.

(21) Choi, S. W.; Martin, E. B.; Morris, A. J.; Lee, I.-B. Fault detection based on a maximum likelihood PCA mixture. *Ind. Eng. Chem. Res.* **2005**, *44*, 2316−2327.

Received for review March 29, 2005
Revised manuscript received January 3, 2006
Accepted January 11, 2006

IE050391W