Principal-Component Analysis of Multiscale Data for Process Monitoring and Fault Diagnosis

Seongkyu Yoon and John F. MacGregor
McMaster Advanced Control Consortium, Dept. of Chemical Engineering, McMaster University, Hamilton, Ontario L8S 4L7, Canada

DOI 10.1002/aic.10260
Published online in Wiley InterScience (www.interscience.wiley.com).

An approach is presented to multivariate statistical process control (MSPC) for process monitoring and fault diagnosis based on principal-component analysis (PCA) models of multiscale data. Process measurements, representing the cumulative effects of many underlying process phenomena, are decomposed by applying multiresolution analysis (MRA) by wavelet transformations. The decomposed process measurements are rearranged according to their scales, and PCA is applied to these multiscale data to capture process variable correlations occurring at different scales. Choosing an orthonormal mother wavelet allows each principal component to be a function of the process variables at only one scale level. The proposed method is discussed in the context of other multiscale approaches, and illustrated in detail using simulated data from a continuous stirred tank reactor (CSTR) system. A major contribution of the paper is to extend fault isolation methods based on contribution plots to multiscale approaches. In particular, once a fault is detected, the contributions of the variations at each scale to the fault are computed. These scale contributions can be very helpful in isolating faults that occur mainly at a single scale. For those scales having large contributions to the fault, one can further compute the variable contributions to those scales, thereby making fault diagnosis much easier. A comparison study is done through Monte Carlo simulation. The proposed method can enhance fault detection and isolation (FDI) performance when the frequency content of a fault effect is confined to a narrow-frequency band. However, when the fault frequency content is not localized, the multiscale approaches perform very comparably to the standard single-scale approaches, and offer no real advantage. © 2004 American Institute of Chemical Engineers *AIChE J,* 50: 2891–2903, 2004

Keywords: principal component analysis, process monitoring, fault detection, fault diagnosis, multiscale modeling

Introduction

A continuous dynamic process is usually characterized with autocorrelated measurements. This autocorrelation is attributed to the effects of process dynamics and disturbances. The common multivariate statistical process control (MSPC) approach to the continuous dynamic processes has been to use principal-component analysis (PCA) or projections to latent structures (PLS) to build a model that captures contemporaneous correlations among the variables but that ignores the serial correlation existing in the data during normal operation. The autocorrelated principal-component values or scores and the squared prediction error (SPE) are then plotted along with appropriate control limits.

As a more explicit method to account for the autocorrelated nature of process data in dynamic systems, one can expand the

Current address of S. Yoon: Umetrics, Inc., Kinnelon, NJ 07405 (seongkyu.yoon@umetrics.com).
Correspondence concerning this article should be addressed to J. F. MacGregor at macgreg@mcmaster.ca.

© 2004 American Institute of Chemical Engineers

X and **Y** matrices by including lagged values of all the variables (Ku et al., 1995; MacGregor et al., 1991; Nomikos and MacGregor, 1994, 1995; Wise et al., 1991; Wold et al., 1984). In this way PCA or PLS will explicitly model both contemporaneous and serial correlations among all the variables. The number of principal components required may now be larger, but the MSPC scheme will now be able to detect any changes in the serial correlation of the variables as well as changes in relationships among the variables. This is analogous to the use of time-series models to develop SPC charts on individual variables (Alwan and Roberts, 1988; Harris and Ross, 1991; Montgomery, 1985).

An alternative approach to account for the dynamic aspects of the data in MSPC is to use MRA (multiresolution analysis) by wavelet decomposition (Daubechies, 1990; Mallat, 1989). The individual signals are decomposed into different scales (or frequencies). Data in each decomposed scale are used for MSPC, which is an indirect way of handling process dynamics. As an application of this approach to fault detection, multiscale PCA (MSPCA) has been proposed by Bakshi (1998). MSPCA enables one to simultaneously extract process correlations across data and account for autocorrelation within sensor data. It captures correlations among the process variables made by various events occurring at different scales. Aradhye et al. (2000) presented the theoretical analysis of MSPCA and addressed several properties from a statistical process control perspective. It was shown that the existing statistical process control (SPC) methods would be best suited for situations where the scale of the signal features that represent abnormal operation is known in advance. If the nature of abnormal features cannot be predicted a priori, the multiscale approach was shown to provide better performance in general. Misra et al. (2002) also proposed a similar scheme. Several PCA models built with approximation and detail blocks of decomposed observations were used for FDI. Teppola and Minkkinen (2000) proposed wavelet-PLS regression models for data analysis and process monitoring. In their wavelet-PLS–based approach, periodic seasonal fluctuation and long-term drifting, shown as low-frequency variations, were filtered by removing the low-frequency blocks. The wavelet-PLS model was constructed based on the filtered measurements. The wavelet-based multiscale approach has also been used in other method fields such as data compression (Misra et al., 2000), sensor validation (Luo et al., 1999), data rectification (Nounou and Bakshi, 1999), and process monitoring and diagnosis (Bakshi, 1998; Kasashima et al., 1995; Vedam and Venkatasubramanian, 1997; Li, 1999).

Yoon and MacGregor (2001b) proposed a generalized or modified principal-component analysis (MSPCA) by which one can decompose the process signals and discriminate all the contributing events according to their scale (or frequency) contents. The MSPCA model is an extension of prior multiscale methods, and enables one to use additional scale information for fault diagnosis. In this article, fault detection and isolation (FDI) approaches based on MSPCA will be investigated. Both on-line and off-line implementation schemes of MSPCA for FDI are proposed. Based on the analogy between multiblock methods and multiscale methods, a new algorithm to calculate the MSPCA model is also presented. Fault isolation methods based on contribution plots are extended to multiscale methods. They are shown to facilitate the isolation of faults that occur in different scales. Performance of FDI using this MSPCA approach is assessed using a simulated continuous stirred tank reactor (CSTR) system. The FDI performance of the MSPCA based method is compared to that of the regular PCA-based method by Monte Carlo simulation.

An Improved MSPCA for Process Monitoring and Fault Diagnosis

Off-line PCA model of multiscale data

Let **X** be the observation matrix of dimension, $n \times m$, where n is the number of observations and m is the number of variables. Using the transformation matrix that is obtained by the boundary corrected wavelet filters, all the columns of **X** can be decomposed into the details at all levels and the approximation at the coarsest level (J). Further, the decomposed block matrices are reconstructed such that all the columns have the same number of rows as in the case of the original observation matrix

$$\mathbf{W}_S\mathbf{W}_A\mathbf{X} = \mathbf{G}_1^{(1)}\mathbf{H}_1^{(1)}\mathbf{X} + \mathbf{G}_1^{(2)}\mathbf{H}_1^{(2)}\mathbf{X} + \cdots + \mathbf{G}_1^{(j)}\mathbf{H}_1^{(j)}\mathbf{X} + \cdots + \mathbf{G}_1^{(J)}\mathbf{H}_1^{(J)}\mathbf{X} + \mathbf{G}_0^{(J)}\mathbf{H}_0^{(J)}\mathbf{X} \quad (1)$$

where each term represents the scale contribution of the original data matrix in terms of reconstructed values. $\mathbf{W}_S$ is the synthesis filter matrix and $\mathbf{W}_A$ is the analysis filter matrix. $\mathbf{H}_1^{(j)}$is the matrix containing wavelet filter coefficients corresponding to scale j and $\mathbf{H}_0^{(J)}$ is the matrix of scaling function filter coefficients at the coarsest scale. If an orthogonal mother wavelet is chosen, then $\mathbf{G}_0 = \mathbf{H}_0^*$, $\mathbf{G}_1 = \mathbf{H}_1^*$, where * denotes Hemitian transpose. Thus, $\mathbf{G}_0^{(J)} = \mathbf{H}_0^{(J)^T}$, $\mathbf{G}_1^{(j)} = \mathbf{H}_1^{(j)^T}$, and $\mathbf{W}_S\mathbf{W}_A = \mathbf{I}$. $\mathbf{H}_1^{(j)}$and $\mathbf{H}_0^{(J)}$ are expressed in terms of successive filters, $\mathbf{H}_0$ and $\mathbf{H}_1$, respectively. $\mathbf{H}_i$ is defined as

$$\mathbf{H}_i = \begin{pmatrix} \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ \cdots & h_i[L-1] & h_i[L-2] & h_i[L-3] & \cdots & h_i[0] & 0 & 0 & \cdots \\ \cdots & 0 & 0 & h_i[L-1] & \cdots & h_i[2] & h_i[1] & h_i[0] & \cdots \\ & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \end{pmatrix} \quad (2)$$

where each line is an even shift of the impulse response of $h_i[n]$ and L is the filter length. $\mathbf{H}_i$ has the effect of filtering the signal by $H_i(z)$ and subsampling by 2 (represented by the shift by 2 in $\mathbf{H}_i$). For details, refer to the literature (Yoon, 2001). All the products of

the analysis and the synthesis in Eq. 1 can then be simplified in terms of the analysis filter matrices, and reexpressed as

$$\begin{aligned}\mathbf{X} &= \mathbf{H}_1^{(1)T}\mathbf{H}_1^{(1)}\mathbf{X} + \mathbf{H}_1^{(2)T}\mathbf{H}_1^{(2)}\mathbf{X} + \cdots + \mathbf{H}_1^{(j)T}\mathbf{H}_1^{(j)}\mathbf{X} + \cdots \\ &\quad + \mathbf{H}_1^{(J)T}\mathbf{H}_1^{(J)}\mathbf{X} + \mathbf{H}_0^{(J)T}\mathbf{H}_0^{(J)}\mathbf{X} \\ &= \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_j + \cdots + \mathbf{X}_J + \mathbf{X}_{J+1} \end{aligned} \tag{3}$$

To investigate the variable correlations over frequency, all the scale contributions in Eq. 2 are rearranged side by side to produce a two-dimensional matrix of size $[n \times (J+1)m]$. The two-dimensional matrix is defined as follows

$$\mathbf{X}_G = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \cdots \quad \mathbf{X}_j \quad \cdots \quad \mathbf{X}_J \quad \mathbf{X}_{J+1}] \tag{4}$$

where $\mathbf{X}_G$ has the dimension of $[n \times (J+1)m]$. One can express $\mathbf{X}_G$ in terms of $\mathbf{V}$ and $\text{Diag}(\mathbf{X})_{J+1}$ as follows

$$\mathbf{X}_G = \mathbf{V}\,\text{Diag}(\mathbf{X})_{J+1} \tag{5}$$

where $\text{Diag}(\mathbf{X})_{J+1}$ is a diagonal matrix whose diagonal component is $\mathbf{X}$ and off-diagonal terms are zero matrices with the same dimension as $\mathbf{X}$. The dimension of $\text{Diag}(\mathbf{X})_{J+1}$ becomes $[(J+1)n, (J+1)m]$. $\mathbf{V}$ is defined as

$$\begin{aligned}\mathbf{V} = [&\mathbf{H}_1^{(1)T}\mathbf{H}_1^{(1)} \quad \mathbf{H}_1^{(2)T}\mathbf{H}_1^{(2)} \quad \cdots \quad \mathbf{H}_1^{(j)T}\mathbf{H}_1^{(j)} \quad \cdots \\ &\mathbf{H}_1^{(J)T}\mathbf{H}_1^{(J)} \quad \mathbf{H}_0^{(J)T}\mathbf{H}_0^{(J)}] \end{aligned} \tag{6}$$

The data matrix is then mean-centered and scaled with the variances. An off-line MSPCA model is obtained by applying the regular PCA on $\mathbf{X}_G$. One can capture the variable correlations over several scales (Yoon, 2001).

On-line MSPCA model

To implement the MSPCA model online, the observations have to be decomposed online. Consider a sequence of variable observations $x[n]$, expressed as

$$\begin{aligned}x[n] = \sum_{j=1}^{J}\sum_{k\in Z} X^{(j)}[2k+1]\,g_1^{(j)}[n-2^jk] \\ + \sum_{k\in Z} X^{(J)}[2k]\,g_0^{(J)}[n-2^Jk] \end{aligned} \tag{7}$$

where

$$X^{(j)}[2k+1] = \langle h_1^{(j)}[2^jk-l],\, x[l]\rangle \qquad j = 1, \ldots, J$$

$$X^{(J)}[2k] = \langle h_0^{(J)}[2^Jk-l],\, x[l]\rangle$$

These are the convolutions of the input with $h_0[n]$ and $h_1[n]$ evaluated at evenly indexed 2^jk. In these equations, $h_i^{(1)}[n] = h_i[n]$and $g_i^{(1)}[n] = g_i[n]$. If $h_i[n]$ is an orthogonal filter, $h_i[-n] = g_i[n]$ and $h_i^{(j)}[n] = g_i^{(j)}[-n]$. This structure implements an orthogonal discrete time wavelet series expansion. The basis functions $g_i^{(j)}[n]$ are defined as the time domain versions of $G_0^{(j)}(z)$ and $G_0^{(j)}(z)$, given by

$$G_0^{(J)}(z) = G_0^{(J-1)}(z)G_0(z^{2^{J-1}}) = \prod_{K=0}^{J-1} G_0(z^{2^K}) \tag{8}$$

$$G_1^{(j)}(z) = G_0^{(j-1)}(z)G_1(z^{2^{j-1}}) = G_1(z^{2^{j-1}})\prod_{K=0}^{j-2} G_0(z^{2^K}) \tag{9}$$

where $G_0^{(1)}(z) = G_0(z)$, $G_1^{(1)}(z) = G_1(z)$, and $j = 1, \ldots, J$. The expressions for analysis filters, $H_0^{(J)}(z)$ and $H_1^{(j)}(z)$ for $j = 1, \ldots, J$ can be similarly obtained. Using these expressions and the multirate identity for filtering, the octave-band filter banks that perform the multiresolution analysis can be expressed in terms of $J+1$ filter banks that are the combination of analysis and synthesis filters, downsamplers, and upsamplers. The filter banks at coarser scales need more samples.

Because wavelet filters are in general noncausal in nature and require future measured data for calculating the current wavelet coefficient, special wavelets at the edges that eliminate boundary errors, while being orthogonal to the other wavelets, are required. These boundary corrected filters are causal and require no information about the future observations to compute wavelet coefficients at the signal ends (Cohen et al., 1993; Nounou and Bakshi, 1999).

Thus, the on-line implementation scheme requires that one decomposes the measured data within a window of dyadic length using a causal boundary–corrected wavelet filter and retains only the last data point of the reconstructed signal for on-line use. When new measured data are available, one moves the window in time to include the most recent measurement while maintaining the maximum dynamic window length. Nounou and Bakshi (1999) used a similar scheme for on-line data rectification. It is noted that the data transformation is done on a row as the new measured observation is obtained. Thus on-line transformation does not guarantee orthogonality of the scale block. However, one can obtain the block orthogonal PCA model by applying multiblock principal-component analysis (MBPCA) (Cheng and McAvoy, 1997) or consensus principal-component analysis (CPCA) (Wold et al., 1987) algorithm on the data transformed by the on-line scheme. The algorithm is summarized below.

Modified MSPCA algorithm

(1) Perform regular PCA on each scale block $\mathbf{X}_j$ to obtain loadings, $\mathbf{P}_j$ and eigenvalues, λ_j.

(2) Select the largest eigenvalue among all the existing values and the corresponding loading.

(3) Obtain a MSPCA loading $\mathbf{P}_{G,k}$ with the selected loading of a scale block, j. For example, if the selected eigenvalue is the ith eigenvalue of the jth scale block, then the corresponding MSPCA loading $\mathbf{P}_{G,k}$ is

$$\mathbf{P}_{G,k}^T = \Big[\underbrace{\mathbf{0}(1,m) \quad \cdots \quad \mathbf{0}(1,m)}_{(j-1)m} \quad \mathbf{P}_{ij}^T \quad \underbrace{\mathbf{0}(1,m) \quad \cdots \quad \mathbf{0}(1,m)}_{(J+1-j)m}\Big]$$

(4) The scores can be obtained in two ways. One can use the block score according to the loading selected, or use $\mathbf{X}_G$ and $\mathbf{P}_{G,k}$, the MSPCA loading

$$\mathbf{T}_{G,k} = \mathbf{X}_G \mathbf{P}_{G,k}$$

(5) Deflate residuals

$$\mathbf{E}_G = \mathbf{X}_G - \mathbf{T}_{G,k}\mathbf{P}_{G,k}^T$$

(6) Return to step 2 and repeat the procedure until one obtains all the MSPCA loadings with all block loadings. In step 2, the eigenvalue is selected among the remaining eigenvalues except for the one used in the previous step.

Process Monitoring and Fault Diagnosis Using Scale and Variable Contributions

With the notion of MSPCA, one can estimate a scale contribution to Hotelling's T^2 for fault diagnosis

$$T^2 = \sum_{i=1}^{A} \frac{t_i^2}{s_{t_i}^2} \approx \sum_{i=1}^{A} \frac{1}{s_{t_i}^2} \left[\sum_{k=1}^{m(J+1)} (p_{ki} x_{G,k})^2 \right] = \sum_{k=1}^{m(J+1)} x_{G,k}^2 \left[\sum_{i=1}^{A} \left(\frac{p_{ki}}{s_{t_i}} \right)^2 \right] = \sum_{k=1}^{m(J+1)} x_{G,k}^2 w_k \quad (10)$$

where p_{ki} is kth variable weight in the ith latent variable, m is the number of variables, J is the number of scales, A is the number of PCA components, $x_{G,k}$ is the kth variable of the decomposed current signal x_G, and s_{t_i} is the standard deviation of the ith score. The overall T^2 is partitioned into a contribution from each scale as follows

$$T_j^2 = \sum_{i=1}^{m} x_{G,(j-1)m+i}^2 w_{(j-1)m+i} \qquad j = 1, \ldots, J+1 \quad (11)$$

This is the expression for the scale contribution of T^2 to a fault. Once the contributing scale for a fault is identified, one can compute the variable contributions of the contributing scale. It can be expressed as follows

$$T_{jk}^2 = x_{G,(j-1)m+k}^2 w_{(j-1)m+k} \quad (12)$$

The above equation is the scale variable contribution to T^2 relative to zero. The expression for contribution to a change in T^2 over some time interval is

$$\Delta T_{jk}^2 = T_{jk}^2(t) - T_{jk}^2(t-1) = \Delta x_{G,(j-1)m+k}^2 w_{(j-1)m+k} \quad (13)$$

where $j = 1, \ldots, J+1$, $k = 1, \ldots, m$. The upper control limits for interval for T^2, T_j^2, and T_{jk}^2 for a given significance level α are given by

$$T_{UCL}^2 = \frac{(n-1)(n+1)q}{n(n-q)} F_\alpha(q, n-q) \quad (14)$$

where $F_\alpha(q, n-q)$ is the upper $100\alpha\%$ critical point of the F-distribution with q and $n-q$ degrees of freedom (MacGregor and Kourti, 1995; Nomikos and MacGregor, 1995; Tracy et al., 1992). In the case of T^2, T_j^2, and T_{jk}^2, the degrees of freedom of the numerator terms (q) are $A(J+1)$, A, and 1, respectively. One needs to carefully determine the significance level, α of T_j^2 because the monitoring task using the multiple scale T_j^2 plots becomes a multiple statistical test and then the actual Type I error will be $1 - (1-\alpha)^{J+1}$. It has an effect of lowering the upper critical point and increasing the sensitivity of the fault detection (Bakshi, 1998). Thus, the α-value of the test using multiple T_j^2 has to be modified such that the overall Type I error would be the same as the desired Type I error. Otherwise, comparison results between the regular PCA/PLS and the multiway, or multiscale methods may lead to an incorrect assessment of the sensitivity of the fault detection (Misra et al., 2002; Qin et al., 2001). Further detail of the Bonferroni adjustment is given in the next section.

Similar expressions can be obtained for the squared prediction error (SPE), that is

$$\text{SPE} = \sum_{i=1}^{m(J+1)} (x_{G,i} - \hat{x}_{G,i})^2 \quad (15)$$

From this expression, the scale SPE (that is, the scale contribution to the overall SPE) is given by

$$\text{SPE}_j = \sum_{i=1}^{m} [x_{G,(j-1)m+i} - \hat{x}_{G,(j-1)m+i}]^2 \qquad j = 1, \ldots, J+1 \quad (16)$$

where $\hat{x}_{G,k} = (\mathbf{P}_A \mathbf{P}_A^T)_k x_{G,k}$ and the j variable contribution to the SPE at that scale SPE_{jk} is

$$\text{SPE}_{jk} = [x_{G,(j-1)m+k} - \hat{x}_{G,(j-1)m+k}]^2 \qquad j = 1, \ldots, J+1, k = 1, \ldots, m \quad (17)$$

The upper control limits for SPE, SPE_j, and SPE_{jk} for a given confidence $1-\alpha$ can be approximated by a weighted chi-squared distribution (Box, 1954) as $\delta_{\text{SPE},\alpha} \equiv g\chi_\alpha^2(h)$, where the weight ($g$) and the degree of freedom (h) are both functions of the eigenvalues of the covariance matrix. The mean and variance of the $g\chi_\alpha^2(h)$ distribution ($\mu = gh$, $\sigma^2 = 2g^2h$) are equated to the sample mean (m) and variance (ν) of the SPE sample (Nomikos and MacGregor, 1995). This is a quick and reliable way to estimate g and h provided that the number of SPE observations is sufficiently large. Thus the upper control limit on the SPE at the given significance level α is given by

$$\delta_{\text{SPE},\alpha} \equiv \left(\frac{\nu}{2m} \right) \chi_\alpha^2 \left(\frac{2m^2}{\nu} \right) \quad (18)$$

This is the upper $100\alpha\%$ critical point of the chi-squared variable weighted with $\nu/2m$ and with $2m^2/\nu$ degrees of freedom at the significance level α (MacGregor and Kourti, 1995; Nomikos and MacGregor, 1995; Tracy et al., 1992). Based on this expression, the upper control limit for the scale SPE and the variable contribution to the scale SPE can be obtained. On

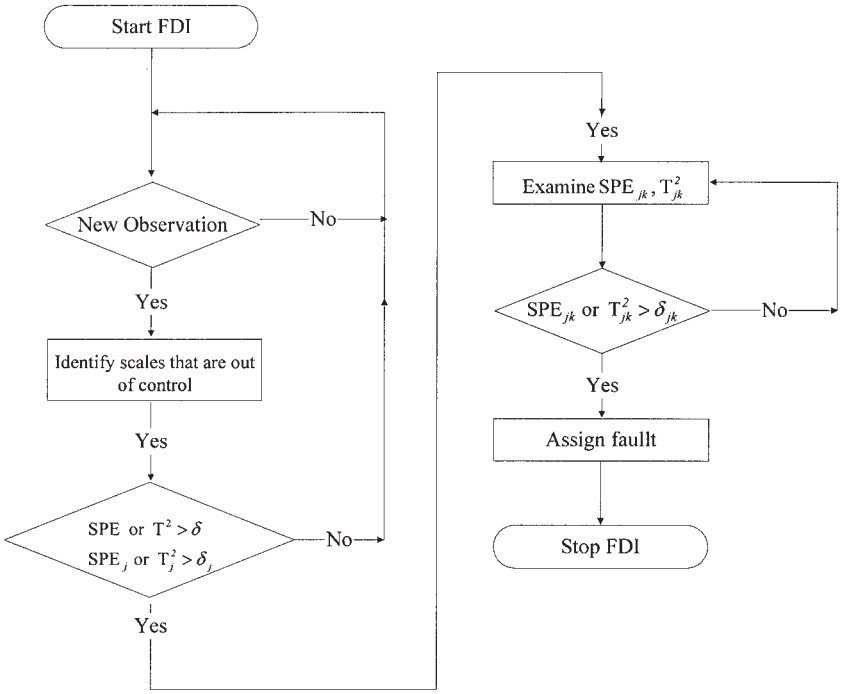

Figure 1. Fault detection and isolation using MSPCA.

the other hand, the upper control limit can also be computed from historical data using the approximate distribution for quadratic forms proposed by Jackson and Mudholkar (1979). Nomikos and MacGregor proved that both methods give almost the same result when one uses the Wilson–Hilferty equation (Nomikos and MacGregor, 1995). In the case of the SPE statistics, one also has to modify the significance level α according to the Bonferroni adjustment attributed to the same reason explained above.

The process monitoring and fault diagnosis procedure based on the MSPCA can be summarized as follows:

(1) Build an MSPCA model with the training data set and calculate the control limits for T^2, T_j^2, SPE, and SPE_j based on the training data set.

(2) For a new sample, calculate, T^2, T_j^2, SPE, and SPE_j.

(3) If either T^2 or SPE or both of them are out of their limits, a fault is detected.

(4) The scale j whose T_j^2 or SPE_j violates its control limit is deemed to be related to the fault.

(5) Any of the variable contributions T_{jk}^2 and SPE_{jk} for $k = 1, 2, \ldots, m$, that are large compared to the others in the responsible scale blocks are then diagnosed to be related to the fault.

It is noted that steps 3 and 4 would be done together assuming that the control limits for the T^2, T_j^2, SPE, and SPE_j are adjusted by Bonferroni corrections. Depending on the fault characteristics, only a few values of T_j^2 and SPE_j in step 4, T_{jk}^2 and SPE_{jk} in step 5 should be affected. In general, one might be able to recognize a fault type with the scale contribution plots, T_j^2 and SPE_j. When the fault induces variations in only a narrow-frequency band, one can further investigate a root cause by examining the variable contributions to T_j^2 and SPE_j, that is T_{jk}^2 and SPE_{jk} ($k = 1, 2, \cdots$). It is recommended to use the overall MSPCA statistics (T^2 and SPE) at first as overall monitoring tools, and the scale MSPCA statistics (T_j^2 and SPE_j) for further diagnosis. Because MSPCA allows one to use information on scale localization in addition to time localization, the fault investigation thus may provide a suitable framework for the fault isolation. Figure 1 summarizes the FDI procedure using the MSPCA model. The proposed scheme provides a systematic approach for diagnosing abnormalities in processes.

Performance measures

Fault detectability is measured with the SPE and T^2 statistics with respect to various types of faults. With the following null and alternative hypotheses

$$H_0\text{: No fault} \qquad H_1\text{: fault} \tag{19}$$

Type I error is defined as a probability of rejecting the null hypothesis that is true (Mongomery and Runger, 1994), and presented by the α risk

$$\alpha = p(\text{reject } H_0 \mid H_0 \text{ is true}) \tag{20}$$

By calculating the false alarm rate during normal operating conditions for the testing set and comparing it against the level of significance upon which the threshold is based, one can measure the robustness of a fault detection method. The discrepancy between the calculated Type I error from data and the theoretical significance level indicates deviation from the assumed distributions on the statistics of T^2 and SPE. One can adjust the control limit such that the calculated Type I values are consistent with the specified significance level (Russel et al., 2000). The adjustment of thresholds to provide the same

Table 1. Comparison between Multiblock and Multiscale Approaches

	Multiblock	Multiscale
Decomposition methods	Process units or variables CPCA (Wold et al., 1987); MBPCA (Cheng and McAvoy, 1997);	Scales (Frequencies) MSPCA (Bakshi et al., 1998; Misra et al., 1998); MSPCA (Yoon and MacGregor, 2001b)
Tools for FDI	T^2/SPE, T_j^2/SPE$_j$, T_{jk}^2/SPE$_{jk}$	T^2/SPE, T_j^2/SPE$_j$, T_{jk}^2/SPE$_{jk}$
Requirements	Location information of process variations and faults	Frequency characteristics of process variations and faults
Applications	Large processes	Small but complicated processes

Type I error, or α, provides a fairer basis for the comparison of the sensitivities of the statistics.

The average run length (ARL) is the average number of samples required to detect a shift (Mongomery and Runger, 1994). When the magnitude of shift is zero, the corresponding ARL value is a reciprocal of the probability of false alarms, and is referred to as the in-control run lengths ($ARL_0 = 1/\alpha$). It is desirable to have the same in-control run length for calculating the consistent ARL values for nonzero mean shifts. This approach provides a standard way of comparing the relative performance of different SPC techniques. For example, when plotted against the magnitude of the shift, the ARL curve is expected to be nonincreasing and typically converges to the location of the mean shift as the magnitude of shift tends to infinity (Aradhye et al., 2003).

Herein, SPE and T^2 statistics and their scale contributions, SPE$_j$ and T_j^2, with respect to various values of the fault-to-signal ratio (FNR $= \mu/\sigma$, where μ is fault magnitude and σ is standard deviation of signal variation), are used for the performance evaluation of the fault detection. The T^2 and SPE statistics are used together without any smoothing effect to calculate Type I and II errors, and the ARL.

One can address a fault isolability of a method using the contributions to the T^2 and SPE statistics. The scale contributions to the SPE and T^2 (SPE$_j$ and T_j^2) and the variable contributions to SPE$_j$ and T_j^2 statistics (SPE$_{jk}$ and T_{jk}^2) are the main tools for the fault isolation under the assumption that a fault is detected. However, they will be limited to the simple type of fault attributed to the fault propagation caused by process dynamics and feedback controls.

Bonferroni adjustment

The significance level (α) is the chance taken to incorrectly declare a fault because of the Type I error. In the case of more than one statistical test, the chance of finding at least one test statistically significant attributed to chance fluctuation increases. Thus, multiple tests increase the chance to incorrectly declare a fault as significant. Because the overall significance level using all tests should result in the same value of the original significance level intended, the significance level of each test should be adjusted downward when there is more than one test in a particular study. For a set of independent tests, the Bonferroni adjustment (Alt et al., 1988) can be used. When α is set at 0.05, the test will show "something" in one out of 20 statistical tests, whereas in fact there is nothing. With five independent tests, the chance of finding at least one fault significant arising from chance fluctuation equals 0.23 ($=1 - 0.95^5$). In ten tests, this chance becomes 0.40 ($=1 - 0.95^{10}$), which is about one in two. Using the Bonferroni method, the α value of each individual test is adjusted downward to ensure that the overall risk for a number of tests remains 0.05. For example, α in the case of five tests will be 0.01 ($=1 - 0.95^{1/5}$). Then the risk of incorrectly finding a fault continues to be 0.05. Thus, in the case of MSPCA with J levels of decomposition, the significance levels of the scale statistics, T_j^2 and SPE$_j$, are adjusted as $\alpha_j = 1 - (1 - \alpha_{\text{given}})^{1/(J+1)}$. Then, the risk of incorrectly finding a fault continues to be 0.05, providing each of the test statistics is statistically independent.

The Type I error of T_j^2/SPE$_j$ is based on the threshold value adjusted with the Bonferroni rule. This adjustment is based on the assumption that one uses the scale monitoring plots, T_j^2/SPE$_j$ as the main fault detection tools (Bakshi, 1998; Misra et al., 2002). On the other hand, adjusting the threshold with the Bonferroni rule is not appropriate in the case that the scale monitoring plots are used only for the purpose of fault diagnosis because it is assumed that a fault is already detected.

Analogy between multiblock and multiscale methods

In very large processes involving many processing units with many variables in each unit, the number of potential errors or faults can be very large, making the diagnosis more difficult. Schemes for process monitoring using multivariate statistical projection methods such as PCA and PLS can be extended to situations where the processes can be naturally blocked into several subsections (MacGregor et al., 1994). The multiblock projection methods allow one to establish monitoring charts for the individual process subsections as well as for the entire process. Multiblock data analysis methods have their origins in path analysis and path modeling in the fields of sociology and econometrics (Wagen and Kowalski, 1988; Wold, 1987). Westerhuis et al. (1998) presented a unifying view of all the multiblock PCA and PLS algorithms. (See Table 1.)

The main advantage of such blocking is to allow for easier interpretation of the data by looking at smaller meaningful blocks and the relationship between blocks. The choice of blocking depends on engineering judgments. The block should correspond to distinct units of the process where all the variables within a block or process unit may be highly coupled, but where there is minimal coupling among variables in different blocks. Variables associated with streams that leave one block and enter another (feed or recycle streams) should generally be included in both blocks.

If a special event or fault occurs in a certain section of the process, then the multiblock methods can generally detect the event earlier and reveal the subsection within which the event has occurred (MacGregor et al., 1994). This is because a process is monitored by the deviations of the block T^2/SPE statistics relative to normal variations in the variables of the corresponding process subsection, but not with respect to variations in all the process variables. However, if a fault occurs

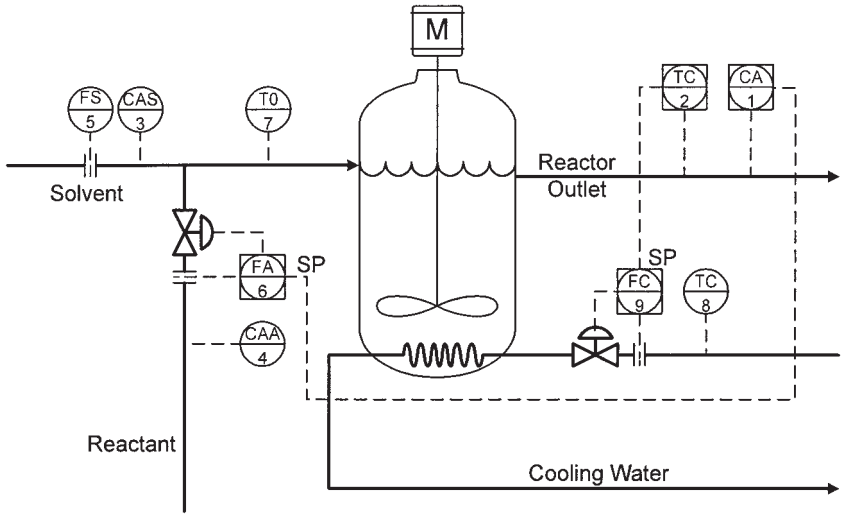

Figure 2. CSTR system with sensors indicated.

that affects all sections of the plant almost simultaneously, then the superblock (or consensus) scores and SPE should detect the event earlier because it is simultaneously using information on the fault gathered from all blocks of the process. Note that the superscores level in MBPCA/MBPLS is equivalent to the conventional (single block) PCA/PLS.

The analogy between the multiblock methods and multiscale methods comes from the fact that both methods decompose the overall monitoring statistics and provide detailed frameworks for diagnosing process abnormalities. The multiblock methods block the information according to variables, whereas the multiscale methods block the data with respect to frequency char-

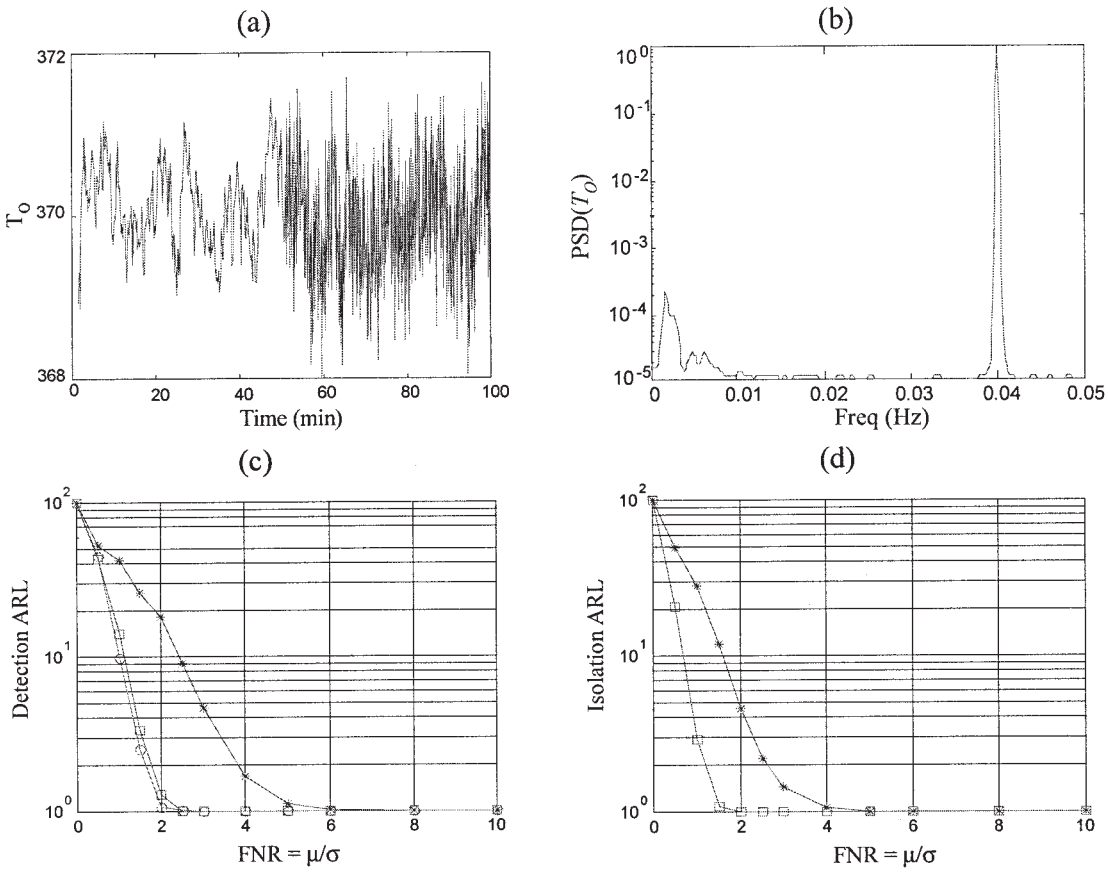

Figure 3. Monte Carlo simulation of an oscillatory fault in the T_o sensor for various FNR ratios.

(a) Time series plot of T_0; (b) power spectral density function of T_0 of fault data (FNR = 2 at 51 min); (c) detection ARL; (d) isolation ARL (# of simulations at one FNR: 500; $\alpha = 0.01$; scaling on $\mathbf{X}_C$; $*$: T^2/SPE(PCA); □: T^2/SPE(MSPCA); ○: T_j^2/SPE$_j$).

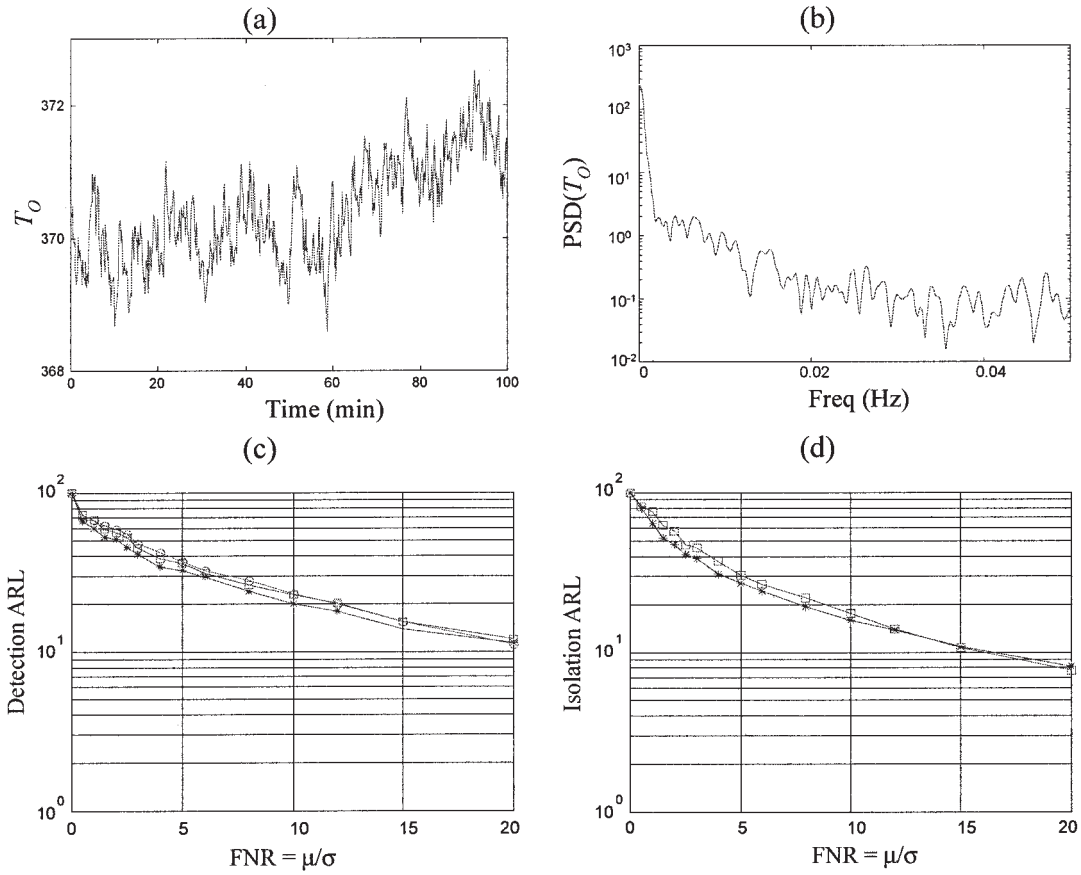

Figure 4. FDI of T_0 sensor drift.

(a) Time series plot of T_0 with FNR = 2.8; (b) power spectral density function of T_0 sensor after the fault; (c) detection ARL based on Monte Carlo simulation; (d) isolation ARL based on Monte Carlo simulation (# of simulations at one FNR: 500; $\alpha = 0.01$; scaling on $\mathbf{X}_G$; $*$: T^2/SPE(PCA); □: T^2/SPE(MSPCA); ○: T_j^2/SPE$_j$).

acteristics. When a fault occurs in one subsection of the process, or in one frequency band of the whole frequency range of the process variations, either the multiblock or multiscale method would be better than the regular methods in detecting and isolating faults. However, it is questionable if the multiblock and multiscale methods would result in better performance when the fault effect is spread over more than one subsection of the process or one frequency range of the process variations. Thus, a suitable monitoring method to give the best detection and isolation of the faults must be based on the fault characteristics. Such information on the faults is very crucial if one is to optimize an FDI scheme.

Numerical Examples

In this section, we illustrate how the MSPCA-based FDI performs for several types of faults with different frequency characteristics. The CSTR system illustrated in Figure 2 (taken from Marlin, 1995, pp. 90–92) is used in this simulation study. The reaction rate, the overall heat transfer coefficient, and all the input variables (C_{AS}, C_{AA}, T_0, and T_C) were simulated to have first-order autoregressive variations in them. These variations, together with additional white noise measurement errors on the input variables, made up the normal or "common-cause" variations in the process when it is operating normally (see Yoon and MacGregor, 2001a for details). All faults were then introduced on top of these normal variations.

Fault with a localized frequency content

This example is used to illustrate a situation where the MSPCA approach should show an advantage over regular PCA. The fault consist of the appearance of an additional oscillatory variation in the measured value of the inlet temperature (T_0) of the CSTR, whereas the real value of T_0 continues to behave in its normal manner (autoregressive plus white noise variation). The fault is simulated with a sinusoidal function whose oscillation period is 25 s and magnitude is M

$$\mu = M \sin\left(\frac{2\pi t}{25}\right) \tag{21}$$

Using the CSTR model, a set of observations was collected for 200 min under normal operation. The sampling interval was 10 s. Testing data were generated with the fault introduced at 51 min into the run as shown in Figure 3a. The PCA and MeSPCA models were estimated with the training data. The scaling for the MSPCA model was done on $\mathbf{X}_G$. Note from the power spectral density (PSD) of the inlet temperature measure-

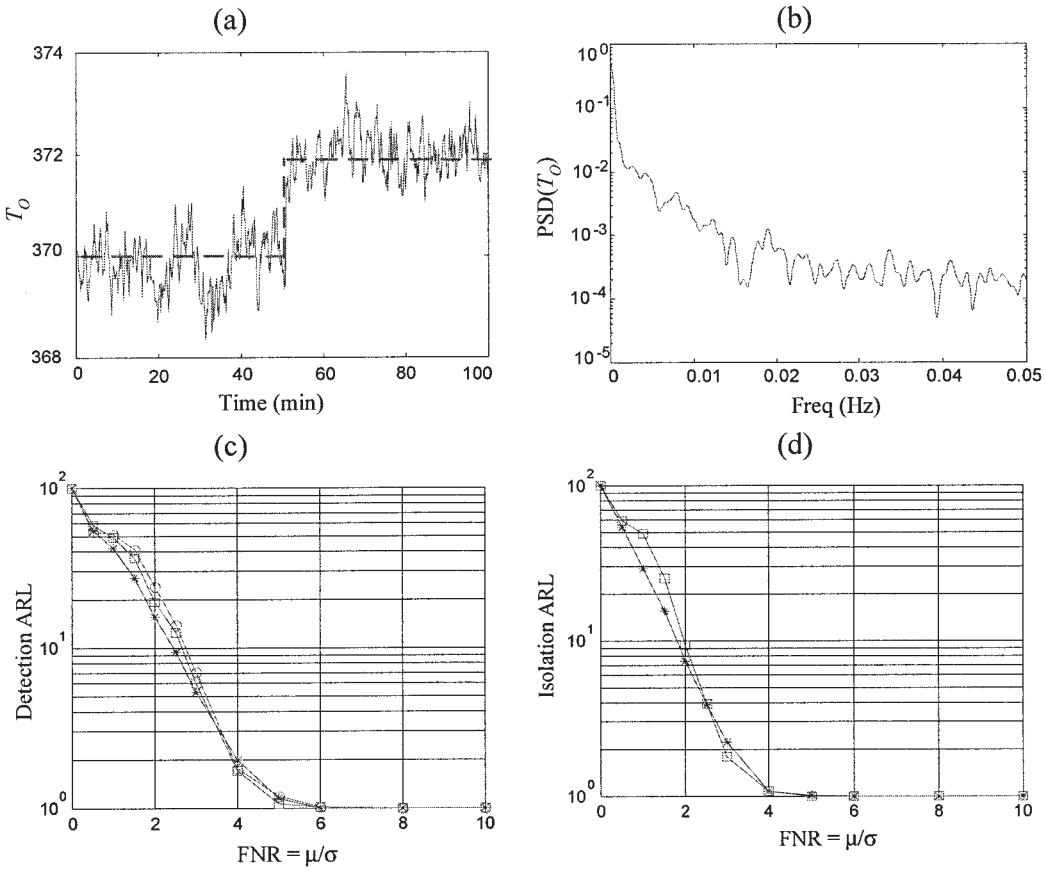

Figure 5. FDI of T_0 sensor bias.

(a) Time series plot of T_0 with $M = 2$; (b) power spectral density function of T_0 signal; (c) detection ARL based on Monte Carlo simulation; (d) isolation ARL based on Monte Carlo simulation (# of simulations at one FNR: 500; $\alpha = 0.01$; scaling on $\mathbf{X}_G$; ∗: T^2/SPE(PCA); □: T^2/SPE(MSPCA); ○: T_j^2/SPE$_j$).

ment for the testing data (Figure 3b), the fault in T_0 shows up as a peak at 0.04 Hz (in the high-frequency–low-scale band) and it does not propagate into other frequencies nor into the other variables. The significance levels of the T_j^2 and SPE$_j$ statistics were compensated with the Bonferroni adjustments The adjusted significance level is 0.0020 [$=1 - (1 - \alpha)^{1/5}$]. The upper control limits of MSPCA T^2/SPE statistics are thus higher than those of the PCA statistics.

To fairly assess the performance of the proposed method, Monte Carlo simulations (Sobol, 1974) were implemented. A total of 500 sets of training and testing data were generated at each value of various fault to normal noise ratios (FNR) on the inlet temperature measurement. The FNR is defined as the ratio of the magnitude of the fault to the standard deviation of the T_0 measurement under normal operation {FNR = [$M/\sigma(T_0)$]}. With each data set, the ARL values for the fault detection and isolation were calculated. The Type I error of T_j^2/SPE$_j$ is based on the threshold value adjusted with the Bonferroni rule. With these monitoring statistics, the overall ARL based on the T^2 and SPE plots is calculated. The individual control limits for the T^2 and SPE plots are then slightly adjusted such that the overall type I error is 0.01 in all cases. Average ARL values are calculated with 500 simulations for each FNR condition.

Figure 3c shows the average ARL values of several statistics for fault detection of the oscillatory fault in the T_0 sensor at several FNR levels. Those of the MSPCA based on T^2/SPE, and the scale contribution based on T_j^2/SPE$_j$ are shorter than those of the regular PCA statistics. This is attributed to the characteristic of the oscillatory sensor fault whose effect is concentrated in a single-frequency range. However, one may not have the same result for other types of faults if their frequency characteristics are different. The ARL for isolation measure is also calculated with the testing data under the assumption that the fault is detected. In the case of MSPCA, the variable contributions to the scale statistics T_j^2 and SPE$_j$, that is, T_{jk}^2 and SPE$_{jk}$ are used, whereas the variable contributions T_k^2 and SPE$_k$ are used for PCA. From Figure 3d, it is obvious that one can diagnose the oscillatory sensor fault better with the statistics of the MSPCA than with those of the standard PCA approach.

On the other hand, this enhancement may result from the scaling effect. Unit scaling of each scale block would put equal weights on all the process correlations of the decomposed scales regardless of the scale contributions to the overall process variation. If the scale contribution of the high-frequency block is small, the unit scaling might unreasonably increase the small contribution. The issue on the scaling will be discussed in a later example.

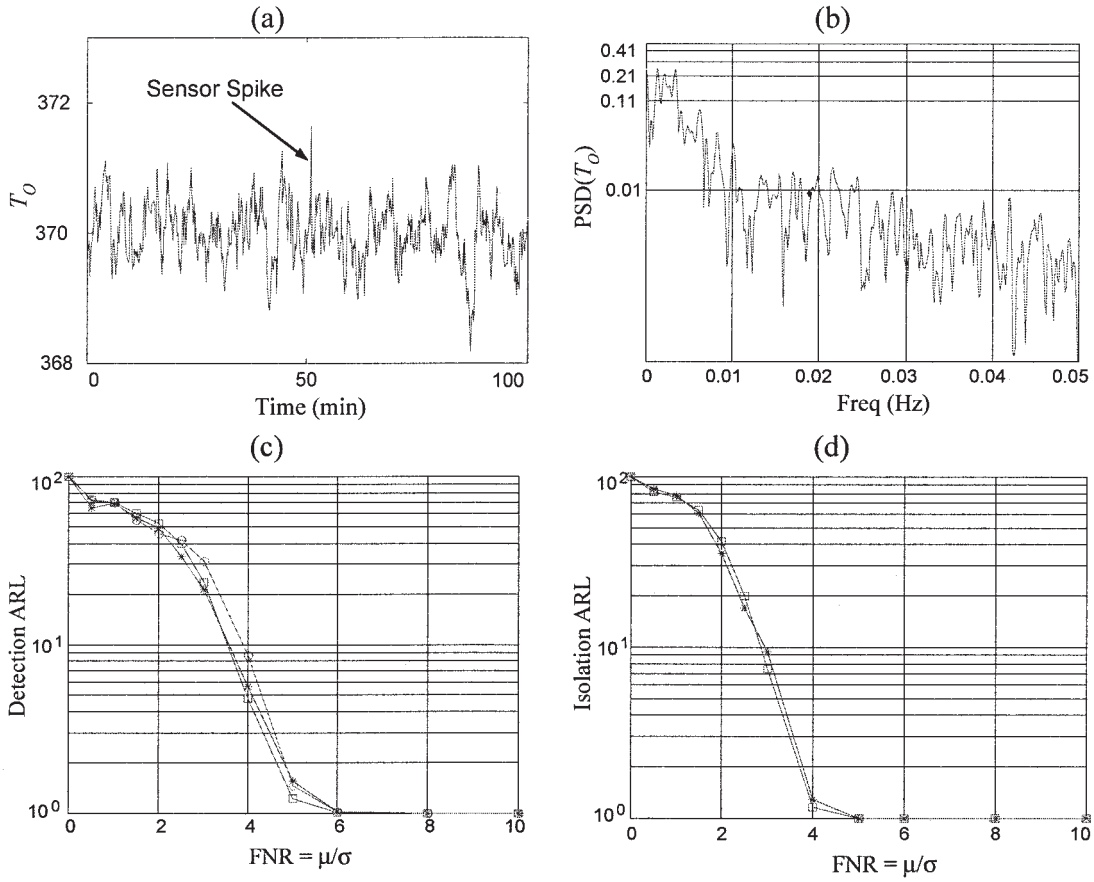

Figure 6. FDI of T_0 sensor spike.

(a) Time series plot of T_0; (b) power spectral density function of T_0 signal; (c) detection ARL based on Monte Carlo simulation; (d) isolation ARL based on Monte Carlo simulation (# of simulations at one FNR: 500; $\alpha = 0.01$; scaling on $\mathbf{X}_G$; ∗: T^2/SPE(PCA); □: T^2/SPE(MSPCA); ○: T_j^2/SPE_j).

Sensor drift (incipient fault)

The fault effect of the sensor drift is also confined in one scale band but at the low-frequency–high-scale bandwidth. A seasonal variation or long-term ageing of reactor catalyst may show this type of characteristics. Figure 4a shows a drift on the inlet temperature measurement. The drifting sensor is simulated to start at 51 min and continue until the end of the testing simulation. The eventual fault magnitude is M. The frequency characteristic of the sensor drift is shown in Figure 4b to be located mainly at the low frequency.

Monte Carlo simulations were done for the sensor drift. All the simulation conditions were the same as those for the previous example except for the type of fault. A total of 500 simulations were made for each FNR $[M/\sigma(T_0)]$ of the drift fault. Figures 4c and d summarize the detection and isolation ARL for the PCA, MSPCA, and the scale contribution with respect to various magnitudes of the fault-to-signal ratio. Both plots show that the detection and isolation performances of the three methods are almost identical. This phenomenon is explained by the fact that the low-frequency scale contains most of the fault variance, and all these approaches are therefore based on essentially the same frequency information. The FDI, using the MSPCA and its scale statistics, does not provide an enhanced performance if the fault effect is confined to the low-frequency region.

Sensor bias

A sensor bias of magnitude M in the inlet temperature measurement (T_0) is simulated as in Figure 5a. All the other simulation conditions are the same as in the above case. Figure 5a shows the process trend of the inlet temperature. Even though this type of fault is a high-frequency event, the high-frequency feature is easily missed and instead other effects such as measurement noise and measurement dynamic characterize the fault signature. Figure 5b confirms this fact.

A total of 500 sets of training and testing data were generated with different magnitudes of the bias on the inlet temperature measurement {FNR = $[M/\sigma(T_0)]$}. Figures 5c and d show the fault detection and isolation ARL for the sensor bias for different values of FNR. The detection and isolation ARL of the MSPCA based on T^2/SPE, and the scale contribution based on T_j^2/SPE_j, show almost the same results as those of the regular PCA. This is attributed to the characteristic of the sensor bias, which consists of a short-duration and high-frequency event, but mainly a low-frequency shift. The simulation result confirms that the multiscale model does not significantly outperform the regular PCA in such a case.

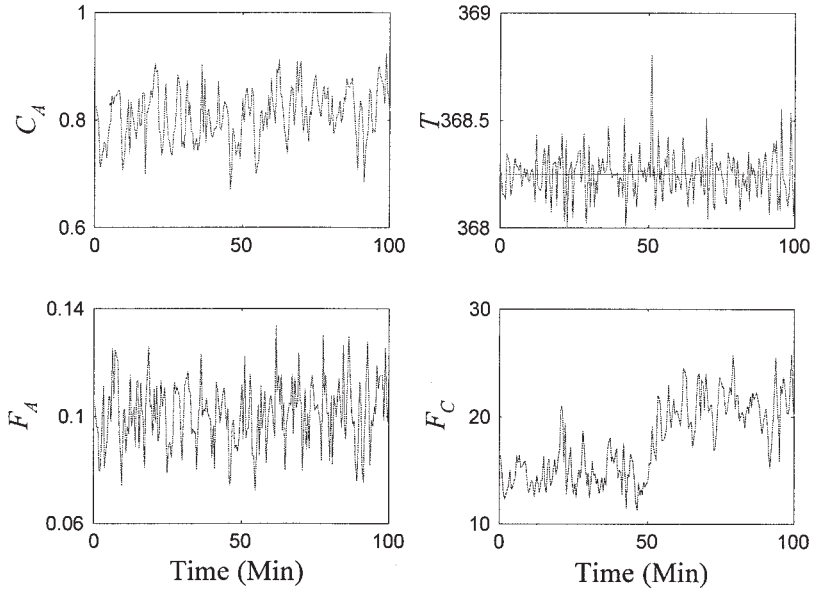

Figure 7. Complex fault with a measurement bias of $\Delta T = +1°$ K in the reactor at 51 min.

Sensor spike

Figure 6a shows a spike on the inlet temperature measurement. Such a fault may arise simply from a temporary sensor error or from a real short-term variation (the latter usually being more serious). The fault, simulated to occur at the 51st minute, occurs in a higher-frequency band. Because it is an instantaneous action, it is not clearly shown in the PSD plot of the sensor spike in Figure 6b.

Monte Carlo simulations were done for the sensor spike case to generalize the result obtained. All the simulation conditions were the same as those for the previous simulation except for the type of fault. A total of 500 simulations were made for each magnitude (M) of the fault {FNR = $[M/\sigma(T_0)]$}. Figures 6c and d show the ARL values of several statistics for fault detection and isolation of the T_0 sensor spike at different FNR values. Both plots show that the detection and isolation performances are almost identical in the three cases, which is explained by the fact that the effect of the sensor spike is easily missed when its magnitude is small. The FDI, using the MSPCA and its scale statistics, does not provide an enhanced performance. It is also noted that the overall performance of the FDI with respect to the sensor spike is worse than that with respect to the sensor bias.

Complex fault

In contrast with the simple faults simulated above, a complex fault is the one whose effects are propagated into many measurements. Complex faults are generally the most important type of fault to detect. They arise from real changes occurring in the process, such as might arise from reaction rate changes attributed to the presence of impurities or contaminants, or from heat transfer changes as a result of fouling, or from any sensor fault that occurs within a feedback loop. In all these cases the effect of the fault will appear in many process variables. The isolation of complex faults becomes complicated and problematic because process dimension and complexity increase the fault propagation. The complex fault has an initial fault signature, a time-varying trajectory, and a steady-state fault signature. The initial fault signature may provide a reliable and prompt source for fault isolation because it is not affected by fault propagation. However, the initial fault signature can be easily missed by any fault-detection scheme. The transient behavior and final steady-state vector of the measurements are thus generally used as an alternative to the initial fault signature (Yoon and MacGregor, 2001a).

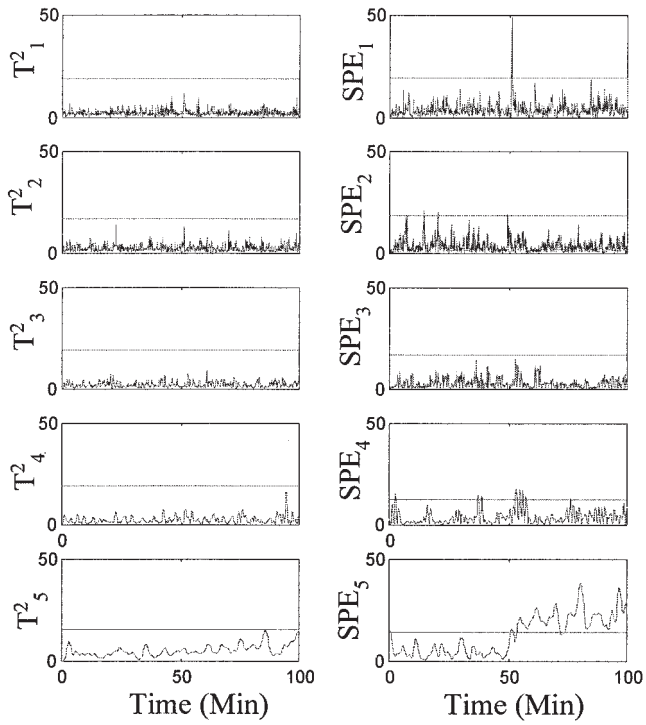

Figure 8. Detection and isolation of a complex fault with $\Delta T = +1$ at 51 min using T^2_j/SPE_j.

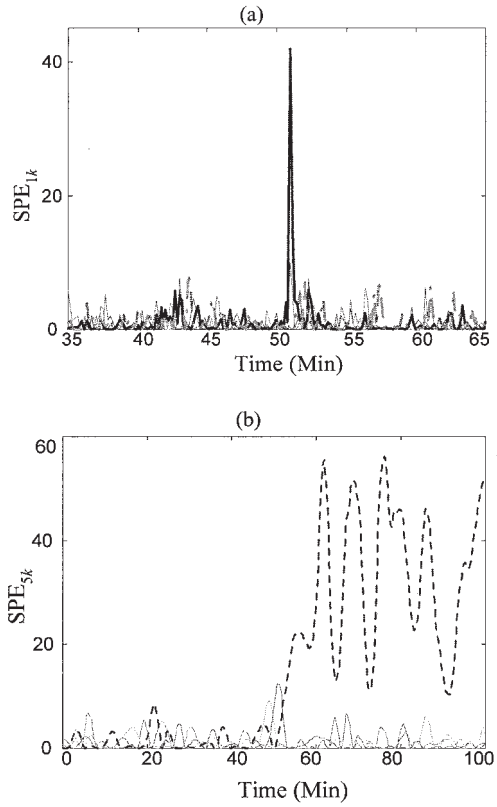

Figure 9. Isolation of a complex fault with $\Delta T = +1$ at 51 min.

(a) $SPE_{1[infi]k}$; (b) $SPE_{5[infi]k}$ (T: —; C_{A0}: · · ·; F_C: – – –; T_0: —; F_A: – · –).

The fault is introduced as a measurement bias of +1°C that suddenly appears on the reactor outlet temperature at time 51, as shown in Figure 7. Because the reactor outlet temperature is the controlled variable, the effect of measurement bias is propagated into other variables. The coolant flow rate, which is the manipulated variable, is changed to remove the deviation of the reactor outlet temperature from its setpoint, and this causes a change in the reactor outlet concentration.

The change of the coolant flow rate is an action at a low-frequency level compared to the random variation in a normal case. On the other hand, the fault—a sudden change in the temperature measurement—occurs at a higher-frequency band. In the conventional PCA, the T^2 and SPE statistics include all the events occurring at all scale bandwidths. With T^2/SPE statistics, one can easily detect the fault. With the variable contribution plot, one can diagnose that the outlet temperature and the coolant flow rate are associated with the fault. Similarly, one can detect the fault successfully with the MSPCA-based T^2/SPE monitoring plots (Yoon, 2001), and diagnose the faults with the scale contributions to T^2 and SPE. Figure 8 indicates that both the lowest and highest scale blocks are responsible for the fault alarm. In the variable contribution plot to the SPE_1 of the D_1 scale in Figure 9a, the outlet temperature measurement is clearly associated with the fault. The variable contribution plot to the SPE_5 of the A_4 scale, the highest approximation block in Figure 9b indicates that the measurement of the cooling water flow rate is also associated with the fault. Based on these two plots, one might conclude that the measurement of the outlet temperature may be a root source of the fault and the cooling water flow rate as an affected variable. This example illustrates how MSPCA does help in the fault isolation through contribution plots in specific scales, even though MSPCA does not improve the detection of the fault.

Scaling effect

To examine the scaling effect on the FDI performance, the Monte Carlo simulation has been done for the sensor precision degradation with the scaling on $\mathbf{X}$ instead of $\mathbf{X}_G$. The simulation conditions used in the previous examples were used.

Figure 10a thus shows the detection ARL values for the oscillatory fault in the T_0 sensor. The ARL has been calculated with T^2 and SPE statistics of the regular PCA, the MSPCA, and the scale contribution of the MSPCA model (MSPCA). Compared to Figure 3a the detection ARL of MSPCA in this example is degraded. This occurs because $\mathbf{X}_G$ scaling puts much greater weight on the faults occurring in high frequency, whereas $\mathbf{X}$ scaling weights the less-dominant low frequencies more. On the other hand, the isolation ARL values in Figure 10b show the same result as given in Figure 3b, regardless of the scaling, because the same statistics T^2_{jk}/SPE_{jk} are used in both cases.

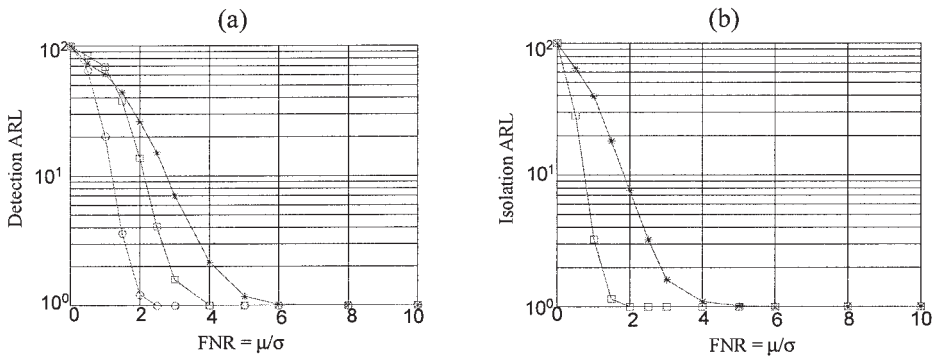

Figure 10. Scaling effect: Monte Carlo simulation of T_0 sinusoidal fault.

(a) Detection ARL; (b) isolation ARL based on Monte Carlo simulation (# of simulations at one FNR: 500; $\alpha = 0.01$; scaling on $\mathbf{X}$; ∗: T^2/SPE(PCA); □: T^2/SPE(MSPCA); ○: T^2_j/SPE_j).

Conclusion

A modified MSPCA algorithm is presented for use in process monitoring. Procedures for both fault detection and isolation are presented. Contributions of the various scales to the overall T^2 and SPE, and contributions of the variables to each scale, have been proposed as additional tools for fault isolation. Because of the usage of scale information, FDI performance sometimes can be enhanced.

FDI performance using the modified MSPCA was assessed through Monte Carlo simulation with a CSTR system for several types of both simple and complex faults. When an abnormal event occurs within one scale bandwidth, the MSPCA-based FDI method is shown to be more effective than that based on the regular PCA model for detecting and isolating faults. However, the multiscale approaches appear to provide little improvement if the fault effect is spread over more than one frequency band, or the fault effect occurs mainly in the scale with the dominant variance. Thus, a monitoring method that gives the best detection and isolation of faults will depend on the fault characteristics; however, unless one has faults that are expected to be very localized in frequency or that appear in scales that normally have small variance, there is little to be gained by using MSPCA models.

Literature CIted

Alt, F. B., and N. D. Smith, "Multivariate Process Control," *Handbook of Statistics*, Vol. **7**, P. R. Krishnaish and C. R. Rao, eds., North-Holland, Amsterdam, pp. 333–351 (1988).

Alwan, L. C., and H. V. Roberts, "Time Series Modeling for Statistical Process Control," *J. Econ. Bus. Stat.*, **6**, 87 (1988).

Aradhye, H. B., B. R. Bakshi, R. A. Strauss, and J. F. David, "Multiscale SPC Using Wavelets: Theoretical Analysis and Properties," *AICHE J.*, **49**(4), 939 (2003).

Bakshi, B., "Multiscale PCA with Application to Multivariate Statistical Process Monitoring," *AIChE J.*, **44**(7), 1596 (1998).

Box, G. E. P., "Some Theorems on Quadratic Forms Applied on the Study of Analysis of Variance Problems: Effect of Inequality of Variance in One-Way Classification," *Ann. Math. Stat.*, **25**, 290 (1954).

Cheng, G., and T. J. McAvoy, "Multi-Block Predictive Monitoring of Continuous Processes," Proc. of IFAC ADCHEM 1997, Banff, Canada (1997).

Cohen, A., I. Daubechies, and V. Pierre, "Wavelets on the Interval and Fast Wavelet Transforms," *Appl. Comput. Harmonic Anal.*, **1**, 54 (1993).

Daubechies, I., "The Wavelet Transform, Time–Frequency Localization and Signal Analysis," *IEEE Trans. Inform. Theory*, **36**(5), 961 (1990).

Harris, T. J., and W. H. Ross, "Statistical Process Control Procedures for Correlated Observations," *Can. J. Chem. Eng.*, **69**, 48 (1991).

Jackson, J. E., and G. S. Mudholkar, "Control Procedure for Residuals Associated with Principal Component Analysis," *Technometrics*, **21**(3), 341 (1979).

Kasashima, N., K. Mori, R. G. Herrera, and N. Taniguchi, "On-Line Failure Detection in Face Milling Using Discrete Wavelet Transform," *CIRP Ann. Manuf. Technol.*, **44**(1), 483 (1995).

Ku, W., R. H. Storer, and C. Georgakis, "Disturbance Detection and Isolation by Dynamic Principal Component Analysis," *Chemom. Intell. Lab. Syst.*, **30**, 179 (1995).

Li, X., "On-Line Detection of the Breakage of Small Diameter Drills Using Current Signature Wavelet Transform," *Int. J. Machine Tools & Manuf.*, **39**(1), 157 (1999).

Luo, R., M. Misra, and D. M. Himmelblau, "Sensor Fault Detection via Multiscale Analysis and Dynamic PCA," *Ind. Eng. Chem. Res.*, **38**, 1489 (1999).

MacGregor, J. F., C. Jaeckle, C. Kiparissides, and M. Koutoudi, "Process Monitoring and Diagnosis by Multiblock PLS Methods," *AIChE J.*, **40**(5), 826 (1994).

MacGregor, J. F., and T. Kourti, "Statistical Process Control of Multivariable Processes," *Control Eng. Pract.*, **3**(3), 403 (1995).

MacGregor, J. F., T. Kourti, and J. V. Kresta, "Multivariate Identification: A Study of Several Methods," Proc. of IFAC Symposium ADCHEM-91, Toulouse, France (1991).

Mallat, S. G., "Multiresolution Approximations and Wavelet Orthonormal Bases of $L_2(R)$," *Trans. Am. Math. Soc.*, **315**, 69 (1989).

Marlin, T. E., *Process Control: Designing Processes and Control Systems for Dynamic Performances*, McGraw-Hill, New York (1995).

Misra, M., S. J. Qin, S. Kumar, and D. Seemann, "On-Line Data Compression and Error Analysis Using Wavelet Technology," *AIChE J.*, **46**(1), 119 (2000).

Misra, M., H. Yue, S. J. Qin, and C. Ling, "Multivariate Process Monitoring and Fault Diagnosis by Multi-Scale PCA," *Comput. Chem. Eng.*, **26**(9), 1281 (2002).

Montgomery, D. C., *Introduction to Statistical Quality Control*, Wiley, New York. (1985).

Montgomery, D. C., and G. C. Runger, *Applied Statistics and Probability for Engineers*, Wiley, New York. (1994).

Nomikos, P., and J. F. MacGregor, "Monitoring Batch Processes Using Multiway Principal Component Analysis," *AIChE J.*, **40**(8), 1361 (1994).

Nomikos, P., and J. F. MacGregor, "Multivariate SPC Charts for Monitoring Batch Processes," *Technometrics*, **37**(1), 41 (1995).

Nounou, M. N., and B. R. Bakshi, "On-Line Multiscale Filtering of Random and Gross Errors without Process Models," *AIChE J.*, **45**(5), 1041 (1999).

Qin, S. J., S. Valle, and M. J. Piovoso, "On Unifying Multi-Block Analysis with Application to Decentralized Process Monitoring," *J. Chemom.*, **15**, 715 (2001).

Russel, E. L., L. H. Chiang, and R. D. Braatz, "Fault Detection in Industrial Processes Using Canonical Variate Analysis and Dynamic Principal Component Analysis," *Chemom. Intell. Lab. Syst.*, **51**, 81 (2000).

Sobol, I. M., *The Monte Carlo Method*, The University of Chicago Press, Chicago and London (1974).

Teppola, P., and P. Minkkinen, "Wavelet-PLS Regression Models for Both Exploratory Data Analysis and Process Monitoring," *J. Chemom.*, **14**, 383 (2000).

Tracy, N. D., J. C. Young, and R. L. Mason, "Multivariate Control Charts for Individual Observations," *J. Qual. Technol.*, **24**, 88 (1992).

Vedam, H., and V. Venkatasubramanian, "Signed Diagraph Based Multiple Fault Diagnosis," *Comput. Chem. Eng.*, **21**(12), S655 (1997).

Wagon, L. E., and B. R. Kowalski, "A Multiblock Partial Least Squares Algorithm for Investigating Complex Chemical Systems," *J. Chemom.*, **3**, 3 (1988).

Westerhuis, J. A., T. Kourti, and J. F. MacGregor, "Analysis of Multiblock and Hierachical PCA and PLS Models," *J. Chemom.*, 12, 301 (1998).

Wise, B. M., D. J. Veltkamp, N. L. Ricker, B. R. Kowalski, S. M. Barnes, and V. Arakali, "Application of Multivariate Statistical Process Control (MSPC) to the West Valley Slurry-Red Ceramic Melter Process," Proc. Waste Management, Tuscon, AZ. (1991).

Wold, S., C. Albano, W. J. Dunn III, O. Edlund, K. Esbensen, P. Geladi, S. Hellberg, E. Johansson, W. Lindberg, and M. Sjostrom, "Multivariate Data Analysis in Chemistry," *Chemometrics, Mathematics and Statistics in Chemistry*, B. R. Kowalski, ed., Reidel, Dordrecht, The Netherlands, pp. 17–95 (1984)

Wold, S., S. Hellberg, T. Lundstedt, M. Sjostrom, and H. Wold, Proc. of Symposium on PLS Model Building Theory and Application, Frankfurt, Germany (1987).

Yoon, S., "Using External Information for Statistical Process Control," PhD Thesis, McMaster University, Hamilton, Canada (2001).

Yoon, S., and J. F. MacGregor, "Fault Diagnosis with Multivariate Statistical Models, Part I: Using Steady-State Fault Signatures," *J. Process Control*, **11**, 387 (2001a).

Yoon, S., and J. F. MacGregor, "Unifying PCA and Multiscale Approaches to Fault Detection and Isolation," Proc. of 6th IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS-6), June, Chejudo, Korea (2001b).

Manuscript received Jan. 6, 2003, and revision received Feb. 27, 2004.