Computers and Chemical Engineering 26 (2002) 1281–1293

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Multivariate process monitoring and fault diagnosis by multi-scale PCA

Manish Misra [a], H. Henry Yue [a,1], S. Joe Qin [a,*], Cheng Ling [b]

[a] *Department of Chemical Engineering, The University of Texas at Austin, Austin, TX 78712, USA*
[b] *Union Carbide Corporation, South Charleston, WV 25303, USA*

Received 5 January 2000; received in revised form 25 February 2002; accepted 25 March 2002

Abstract

Chemical process plant safety, production specifications, environmental regulations, operational constraints, and plant economics are some of the main reasons driving an upward interest in research and development of more robust methods for process monitoring and control. Principal component analysis (PCA) has long been used in fault detection by extracting relevant information from multivariate chemical data. The recent success of wavelets and multi-scale methods in chemical process monitoring and control has catalyzed an interest in the investigation of wavelets based methods for fault detection. In the present work, multi-scale principal component analysis (MSPCA) is used for fault detection and diagnosis. MSPCA simultaneously extracts both, cross correlation across the sensors (PCA approach) and auto-correlation within a sensor (wavelet approach). Using wavelets, the individual sensor signals are decomposed into approximations and details at different scales. Contributions from each scale are collected in separate matrices, and a PCA model is then constructed to extract correlation at each scale. The multi-scale nature of MSPCA formulation makes it suitable to work with process data that are typically non-stationary and represent the cumulative effect of many underlying process phenomena, each operating at a different scale. The proposed MSPCA approach is able to outperform the conventional PCA based approach in detecting and identifying real process faults in an industrial process, and yields minimum false alarms. Additionally, the advantage of MSPCA, over the traditional PCA approach for sensor validation, is also demonstrated on an industrial boiler data set.

Keywords: Fault detection; Sensor validation; Wavelets; Multi-scale; Principal component analysis

1. Introduction

Chemical process plant safety, production specifications, environmental regulations, operational constraints and plant economics are some of the main reasons driving an upward interest in research and development of more robust methods for process monitoring and control. In an era where process monitoring and control techniques are heavily dependent on the quality of data, it is imperative to have proper techniques to discriminate a normal plant operation from an abnormal situation. Though infrequent, these abnormal situations have caused a significant impact on the safety and economy of the process industry. It has been estimated that poor monitoring and control of such abnormal situations has caused the US petrochemical industry alone an annual loss of $20 billion (Nimmo, 1995). Efficient monitoring and early control of an abnormal situation have, therefore, become an area of great research interest.

Under normal operating conditions, sensor measurements are highly correlated. This correlation stems from the basic physical and chemical principles such as mass and energy balances that exist among the variables being measured. Various researchers have used techniques such as principal component analysis (PCA) to exploit the property of correlation among variables under normal operating conditions, in order to extract information from the data. Notable applications of PCA in chemical engineering have been in process monitoring (Nomikos & MacGregor, 1995; MacGregor

* Corresponding author. Tel.: +1-512-471-4417; fax: +1-512-471-7060.
E-mail address: qin@che.utexas.edu (S.J. Qin).
[1] Present address: Tokyo Electron, 2400 Grove Blvd., Austin, TX 78741, USA

0098-1354/02/$ - see front matter
PII: S0098-1354(02)00093-5

& Kourti, 1995; Kresta, MacGregor & Marlin, 1991; Wise, Ricker, Veltkamp & Kowalski, 1990), quality control (Piovoso, Kosanovich & Pearson, 1992), disturbance detection (Ku, Storer & Georgakis, 1995), sensor fault diagnosis (Dunia, Qin, Edgar & McAvoy, 1996; MacGregor, Jaeckle, Kiparissides & Koutoudi, 1994) and process fault diagnosis (Dunia & Qin, 1998a,b; Raich & Cinar, 1996). Using PCA, one is able to capture the variability in fewer principal components, thereby reducing data dimensionality.

Though PCA has been highly successful in statistical process monitoring, its best applications are restricted to analyzing steady state data containing linear relationships between the variables. As it is not always possible to meet such requirements, many modifications of PCA such as non-linear PCA (Dong & McAvoy, 1996; Tan & Mavrovouniotis, 1995; Kramer, 1991; Hastie & Stuetzle, 1989), multiblock PCA (Wold, Kettaneh & Tjessem, 1996; MacGregor et al., 1994), multiway PCA (Nomikos & MacGregor, 1994) and dynamic PCA (Li & Qin, 2001; Ku et al., 1995) have been developed. These techniques operate on data collected at a single scale, i.e. they present a convolved picture of events occurring at various time scales. Chemical processes, on the other hand, are known to operate at different scales, and have contributions from events occurring at different scales such as (Bakshi, 1998).

- Events occurring at different locations and with different localizations in time and frequency.
- Stochastic processes whose energy or power spectrum changes with time and/or frequency.
- Variables measured at different sampling rates or containing missing data.

Thus, a multi-scale PCA (MSPCA) formulation in which contributions from events occurring at different scales are captured by PCA models at the corresponding scale seems to be better suited for extracting information from process data. Wavelets, with their time-frequency localization and multi-resolution property, have offered a framework for multi-scale representation of data. In recent years, wavelet-based multi-scale methods have demonstrated superior performance over the conventional methods in various chemical engineering tasks such as process data compression (Misra, Kumar, Qin & Seemann, 2001, 2000; Bakshi & Stephanopoulos, 1996), sensor validation (Luo, Misra, Qin, Barton & Himmelblau, 1998; Luo, Misra & Himmelblau, 1999, 2002), data rectification (Nounou & Bakshi, 1998) and process monitoring and diagnosis (Vedam & Venkatasubramanian, 1997), and recently, Bakshi (1998) developed an MSPCA formulation and demonstrated its effectiveness for monitoring of multivariate statistical processes.

In essence, the MSPCA approach as proposed by Bakshi (1998) is the same as the one proposed in this paper. There are, however, some differences in the application and in extension of the MSPCA approach. Specifically:

- Our approach goes beyond the fault detection task. Once a fault is detected via MSPCA, we have proposed a multi-scale fault identification technique for identifying the type of fault. Bakshi (1998) does not perform multi-scale fault identification.
- We have introduced a comprehensive scheme for fault detection and identification right from the time a data point is generated. As per our approach, once a data point is generated, it is fed to a sensor validation scheme to validate the data for any bad/invalid data point, and also to serve as an early warning in case a fault of large magnitude is present.
- We have applied a recursive MSPCA scheme for fault detection and identification to real industrial data. Using our approach, we were able to successfully detect and identify real industrial faults.

The next section prepares the ground for MSPCA by presenting a brief overview of PCA for fault detection and of wavelets as a tool for multi-resolution analysis. Section 3 explains the MSPCA formulation. Section 4 describes an industrial gas phase tubular reactor system used in this work. Results with two sets of industrial data are presented in Section 5. Section 6 summarizes the paper with some key results.

2. Principal component analysis and wavelets

PCA has been used for various multivariate data analysis techniques such as process monitoring, quality control, sensor and process fault diagnosis. In this section, the general principles of using PCA for fault detection is presented. It is followed by a brief introduction to wavelets. This prepares the ground for MSPCA which is explained in Section 3.

2.1. Principal component analysis

In PCA, the correlation among sensors is used to transform the multivariate space into a subspace which preserves maximum variance of the original space in minimum number of dimensions. In other words, PCA rotates the original coordinate system along the direction of maximum variance. Consider a data matrix $\mathbf{X} \in \Re^{n \times m}$ consisting of n sample rows and m variable columns that are normalized to zero mean and unit variance. The matrix $\mathbf{X}$ can be decomposed into a score matrix $\mathbf{T}$ and a loading matrix $\mathbf{P}$ whose columns are the right singular vectors of $\mathbf{X}$ as follows,

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \tilde{\mathbf{X}} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \tilde{\mathbf{T}}\tilde{\mathbf{P}}^{\mathrm{T}} \tag{1}$$

where $\tilde{\mathbf{X}} = \tilde{\mathbf{T}}\tilde{\mathbf{P}}^{\mathrm{T}}$ is the residual matrix.

Once a PCA model is built and a new data sample, $\mathbf{x}$, is to be tested for fault detection, it is first scaled and then decomposed as follows,

$$\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}} \tag{2}$$

where,

$$\hat{\mathbf{x}} = \mathbf{P}\mathbf{P}^{\mathrm{T}}\mathbf{x} \in S_{\mathrm{p}} \tag{3}$$

is the projection on the principal component subspace (PCS), S_{p}, and

$$\tilde{\mathbf{x}} = (\mathbf{I} - \mathbf{P}\mathbf{P}^{\mathrm{T}})\mathbf{x} \in S_{\mathrm{r}} \tag{4}$$

is the projection on the residual subspace (RS), S_{r}.

For fault detection in the new sample $\mathbf{x}$, a deviation in $\mathbf{x}$ from the normal correlation would change the projections onto the subspaces, either S_p or S_r. Consequently, the magnitude of either $\bar{\mathbf{x}}$ or $\hat{\mathbf{x}}$ would increase over the values obtained with normal data.

The square prediction error (SPE), also known as Q, is a statistic that measures lack of fit of a model to data. The SPE statistic indicates the difference, or residual, between a sample and its projection into the k components retained in the model. The exact description of the distribution of SPE is given by Jackson (1991). Mathematically,

$$\mathrm{SPE} \equiv \|\bar{\mathbf{x}}\|^2 = \|(\mathbf{I} - \mathbf{P}\mathbf{P}^{\mathrm{T}})\mathbf{x}\|^2 \tag{5}$$

The process is considered normal if

$$\mathrm{SPE} \leq \delta^2 \tag{6}$$

where δ^2 is a confidence limit for SPE. A confidence limit expression for SPE, when x follows a normal distribution, is developed by Jackson and Mudholkar (1979). Since then, it has often been used as a limit for fault detection (Wise & Ricker, 1991; Dunia et al., 1996; Qin, Yue & Dunia, 1997). In this article, the SPE test is used as the main criterion for fault detection.

2.2. Wavelets

The theory of wavelet transforms is based on multiresolution analysis (Strang & Nguyen, 1996) which implies that the space of finite energy square integrable functions $L^2(R)$ can be decomposed into nested subspaces at multiple resolutions. The subspaces are defined by a series of orthonormal basis functions $\{\phi_{j,n}(x)\}$. Any square integrable function in this space can be expressed as a linear combination of these bases. Consider a signal $x(t)$ as the Jth level signal $a_J(t) = x(t)$. This signal can be decomposed into two components, one of which is lying in the *coarse approximation* subspace, denoted as $a_{j-1}(t)$, and the other in the $(J-1)$th level detail, $b_{J-1}(t)$. The wavelet decomposition then calculates the $(J-2)$th level approximation and detail from $a_{J-1}(t)$.

The simplest wavelet transform is the Haar wavelet with filter length of 2. The process of obtaining wavelet coefficients $\hat{a}_{j-1,n}$ and $\hat{b}_{j-1,n}$ from the coefficients at higher resolution $\hat{a}_{j,m}$'s, such that $\hat{a}_{J,k}$ represents the discrete raw signal $x(k\Delta t)$ obtained at the highest resolution J, is called analysis. For Haar wavelets, the analysis equations are,

$$\hat{a}_{j-1,n} = \frac{1}{\sqrt{2}}(\hat{a}_{j,2n} + \hat{a}_{j,2n+1}) \tag{7}$$

$$\hat{b}_{j-1,n} = \frac{1}{\sqrt{2}}(\hat{a}_{j,2n} + \hat{a}_{j,2n+1}) \tag{8}$$

The process of reconstructing the signal $x(k\Delta t)$ back from the coefficients at lower resolutions $\hat{a}_{j,n}$ and $\hat{b}_{j,n}$ is called synthesis. The Haar wavelet synthesis equations for obtaining coefficient $\hat{a}_{j,2n}$ and $\hat{a}_{j,2n+1}$ from coefficients $\hat{a}_{j-1,n}$ and $\hat{b}_{j-1,n}$ are as follows:

$$\hat{a}_{j,2n} = \frac{1}{\sqrt{2}}(\hat{a}_{j-1,n} + \hat{b}_{j-1,n}) \tag{9}$$

$$\hat{a}_{j,2n+1} = \frac{1}{\sqrt{2}}(\hat{a}_{j-1,n} - \hat{b}_{j-1,n}) \tag{10}$$

The wavelet transform can be used to decompose multivariate signals $\mathbf{X}$ into approximations $\mathbf{A}_1$ and details $\mathbf{D}_1$ coefficients at the first level. Application of the same transform on the approximations $\mathbf{A}_1$ causes them to be decomposed further into approximations $\mathbf{A}_2$ and details $\mathbf{D}_2$ coefficients at the second level. The decomposition process can continue to a level L as long as the length of approximation coefficients in $\mathbf{A}_L$ is more than the length of coefficients in the wavelet filter. At this stage, there are details coefficients, $\mathbf{D}_1$, $\mathbf{D}_2$, ..., $\mathbf{D}_L$, and approximation coefficients, $\mathbf{A}_L$. The MSPCA approach presented in Section 3 uses this property of wavelet decomposition in the development of a multi-scale framework for analysis of process data.

3. MSPCA formulation

The idea of jointly using PCA and wavelets was first reported by Wickerhauser (1994) who proposed an approximation to PCA using a joint-best basis algorithm over wavelet packet decomposition of the multivariate data. In Wickerhauser's algorithm, first the complete wavelet packet tree for all the sensors are generated. Then the entropy of the square of wavelet packet coefficients is calculated across all trees for each level and time index. The basis which minimizes this overall entropy is chosen as the best basis for the multivariate data. Saito (1994), Coifman and Saito (1996) proposed a local PCA approach by applying local cosine transform on the data and selecting those bases that capture the maximum variance. Instead of applying PCA on time-domain data, Kosanovich and Piovoso (1997) applied PCA on the wavelet coefficients for process monitoring. Bakshi (1998) developed an MSPCA formulation and demonstrated its effectiveness for monitoring multivariate processes.

In this paper, the primary motivation for jointly using PCA and wavelets emanated from the idea that, in PCA, the correlation among sensors is used to transform the multivariate space into a subspace which preserves maximum variance of the original space. However, PCA fails to make use of correlation within the sensor along the time line. In other words, it does not utilize the information pertaining to the frequency or scale characteristics of the individual sensors. Wavelets, on the other hand, capture correlation *within* a sensor whereas PCA correlates *across* sensors. Thus, wavelets and PCA based analysis of multivariate data represent two extremes, one, making use of only the signal trends, and the other, using only correlation. In this article, these two techniques are effectively combined for their complementary strengths and to extract maximum information from multivariate sensor data. This combined approach is applied on real plant data for process fault detection and it demonstrates superior results over conventional PCA methods.

Fig. 1 shows the complete flow diagram for the proposed MSPCA system. In any chemical or processing unit, there are many sensors, measuring variables such as temperature, pressure and flow rates. The presence of many sensors provides valuable redundancy, and the measurements are highly correlated. This property is exploited in the first step of the proposed work where sets of highly correlated data, go into a regular PCA module. The objective of this is to first, validate the sensor data for any bad/invalid data point, and also, to serve as an early warning system in the case that a fault of large magnitude is present. Some variables, that may not be highly correlated, are fed directly into the sensor validation module without being clustered with other variable measurements. A priori knowledge of the physical and chemical principles governing the process operation could help in deciding which variables may exhibit a high correlation with others. As an example, information about mass and energy balances of the variables in the process is generally available. This information can be used to determine which variables are correlated, and to cluster together the correlated variables.

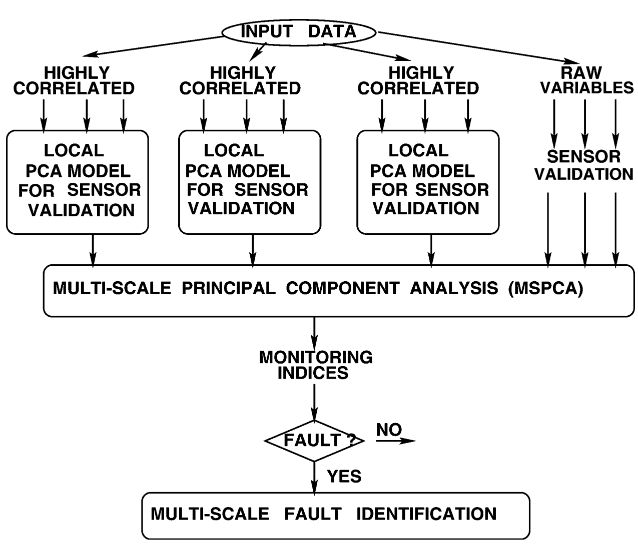

Fig. 1. Flow diagram for the recursive MSPCA formulation.

For sensor validation, it is unusual to have multiple faults occurring simultaneously. A faulty sensor would break the normal correlation and the resultant SPE would be significantly different from the others. Qin et al. (1997) proposed a scheme for sensor validation by reconstructing the normal part of the faulty data in all possible fault directions. A set of faults were assumed, and for each assumed fault, a fault magnitude was estimated by reconstruction from other sensors using the approach discussed in Qin et al. (1997). Although reconstruction along all directions tend to reduce SPE, the reconstruction with the largest reduction in SPE corresponds to the faulty sensor. A sensor validity index (SVI) is defined as

$$\eta_j = \frac{\mathrm{SPE}(\mathbf{x}_j)}{\mathrm{SPE}(\mathbf{x})} \tag{11}$$

where $\mathrm{SPE}(\mathbf{x}_j)$ is the SPE when the jth sensor is assumed faulty and reconstructed. SVI is fault sensitive and $0 \le \eta_j \le 1$. The details of this scheme are described in Dunia et al. (1996). After being validated, the sample data is then fed into the MSPCA module for process monitoring and fault detection.

For the MSPCA formulation, consider an $n \times m$ data matrix $\mathbf{X}$ having m variables and n samples. Each of the m columns are first decomposed individually by applying a discrete wavelet transform (DWT) as explained in Section 2. It may be noted that the same wavelet transform with the same level of decomposition, L, is applied to each of the m variables. The wavelet approximations A_L from each of the m decompositions are collected in one matrix of size $m \times n/2^L$ (as shown by the $\mathbf{A}_L$ solid line matrix in Fig. 2). Similarly, the wavelet details ($\mathbf{D}_1$ to $\mathbf{D}_L$ from each of the L levels) from each of the m decompositions are collected in L corresponding matrices (as shown by the $\mathbf{D}_L$ dashed and $\mathbf{D}_1$ dotted line matrices in Fig. 2), with matrix size varying $m \times n/2^i$, $i = 1, 2, \ldots, L$. Thus a total of $L+1$ matrices are formed, each being represented at a different scale, and captured trends within the sensors at the corresponding scale. As explained in Section 2, PCA is then applied to each of the $L+1$ matrices, the objective being to extract the correlation across the sensors. Depending on the process objective, the performance indices such as SPE can then be monitored. Fig. 2 is a pictorial representation of our MSPCA formulation.

Process monitoring is started by selecting a normal (fault-free) data set. The variables from the selected data set are separated into clusters that are highly

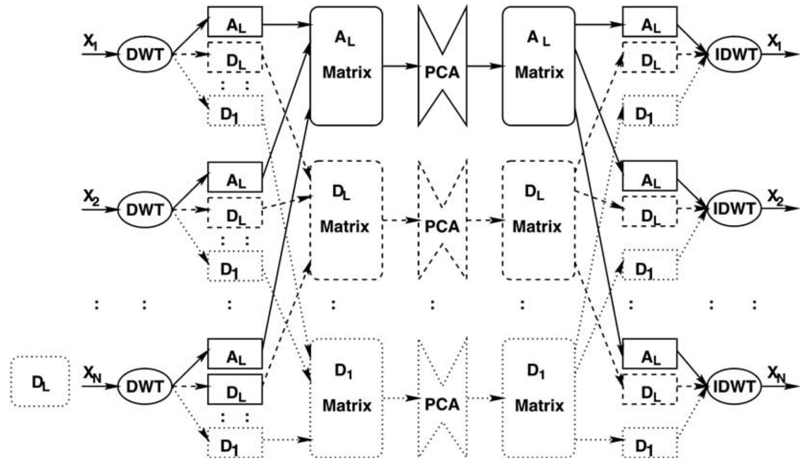

Fig. 2. Multi-Scale PCA. Solid lines, wavelet approximations model; dashed and dotted lines, wavelet details models.

correlated among themselves. A basic understanding of the physical and chemical principles governing the process operation helps in clustering the highly correlated variables together. Sets of clustered variables are then fed into local PCA modules for sensor validation, and a normal PCA model is constructed. This data set is also used to build the normal MSPCA model.

For on-line monitoring, a moving window approach is adopted. A window of incoming sample data matrix is fed into the system. Local PCA modules help to detect any bad/invalid data points, and send an advanced warning if a fault of large magnitude is present. After being validated, the data set is then fed into the MSPCA framework, which calculates the SPE for each level of wavelet decomposition. If the SPE is within limits, as described in Section 2, then the data is considered normal, and the normal model is updated by including this sample data also. If the SPE at a certain scale violates the limits, then the sample data contains a fault corresponding to that scale. For example, if the SPE at the approximation scale is violated, it implies the presence of a slowly drifting fault. Experimental results are shown in Section 5.

Once a fault at a certain scale is detected, multi-scale fault identification is conducted to identify the source of the fault. Contribution plots of the variables, at the time when the SPE limit is violated, are generated and the variables with abnormally high contribution to SPE are identified. This step is performed only at the scales where the SPE violates the limit. Multi-scale fault identification results with the industrial reactor system data are shown in Section 5.2.

In testing the data from an industrial reactor system, the proposed MSPCA approach is successful in early detection and identification of real process faults. In another example, data from an industrial boiler is taken, and faults of different scales and magnitudes are injected. The proposed MSPCA approach is able to detect and identify all the faults, whereas the conventional PCA approach is unable to detect the fault early and reliably. Four different types of faults are tested: drift, bias, precision degradation and spike. The results are shown in Section 5.3.

4. Tubular reactor process description

An important part of the proposed work is to apply the MSPCA formulation to an industrial data set about which no a priori assumptions could be made. Data, from an industrial gas phase tubular reactor system, are analyzed for real process faults. The reactor system is commonly found in petrochemical industry and is used in the manufacturing of several products. No assumptions are made about the quality of the data, and it is possible that the data has some gross errors or faults. The objective is early prediction of an abnormal situation that will eventually lead to sudden pressure surge in the reactor tubes.

4.1. General process description

The reactor system consists of a number of parallel tubular reactors, residing in and sharing a common hot box. The reactors draw feed from a common header. The product is combined at a manifold before being chilled by an exit heat exchanger. Multiple small-bore reactors facilitate better heat transfer as compared with a single large bore reactor, and allow for better process control, resulting in a better overall product quality. However, due to the physical connections at the common header and manifold, significant process interactions are also present. Fig. 3 shows the schematic of the tubular reactor system. A general description of individual variables listed in Fig. 3 is given in Section 4.3.

The primary reaction in the reactor is endothermic,

$$A + \text{impurities} \rightarrow B + \text{byproducts} \quad (12)$$

The least desirable byproduct deposits on the inside of the reactor tubes and forms a solid layer. This deposition impedes the flow of process gases and reduces heat transfer efficiency. Various abnormal events often lead to accelerated solid layer deposition, thereby resulting in early system shutdown. Hence the desire for early prediction of abnormal situation.

4.2. Basic control strategy

In this endothermic reaction system, the feed rate is manipulated to regulate the reaction temperature. Feed also serves as a heat sink; heat entering the reactors in excess of that required for reaction is carried off. The process time constant is in the order of sub to low minutes. Each reactor in the hot box is controlled independent of others. The product composition is analyzed. This is used to determine the degree of reaction, which in turn, is used to set the reaction temperature setpoint.

The hot box has a single fuel control system. It is normally set to maintain a constant heat input to the hot box. A number of factors, as described in Section 4.3, cause the fuel control program to reduce the fuel rate.

4.3. Key process variables

4.3.1. Feed rate

Measured independently for each reactor, the feed rate is the primary indication of throughput. Ideally, this should be maximized for better economic returns. Due to mechanical considerations, the total feed rates for all reactors are limited to a maximum amount.

4.3.2. Process gas (reaction) temperature

This temperature is measured at the reactor outlet. The reaction temperature is regulated by manipulating the feed rate (1).

4.3.3. Product composition

The product composition is analyzed to determine the degree of reaction. The degree of reaction is regulated by manipulating the reaction temperature setpoint.

4.3.4. Heat value of fuel

Due to variations in fuel quality, the heat value is measured. This allows regulation of heat input to the hot box. Unless otherwise noted, this is normally maintained at a constant per unit feed.

4.3.5. Heat exchanger inlet pressure

This measurement is used as the primary indicator of reactor system health. Higher pressure represents excessive solid deposit, which restricts feed flow and affects economic bottom line. The system cannot be operated economically with excessive pressure build up. Typically, the heat input is reduced to retard reaction.

4.3.6. Heat exchanger outlet temperature

This temperature is the primary indicator of the heat exchanger health. Abnormally high temperature indicates a decrease in heat transfer efficiency through fouling. The heat input is cut to lower heat carried off by the process gas to the heat exchanger.

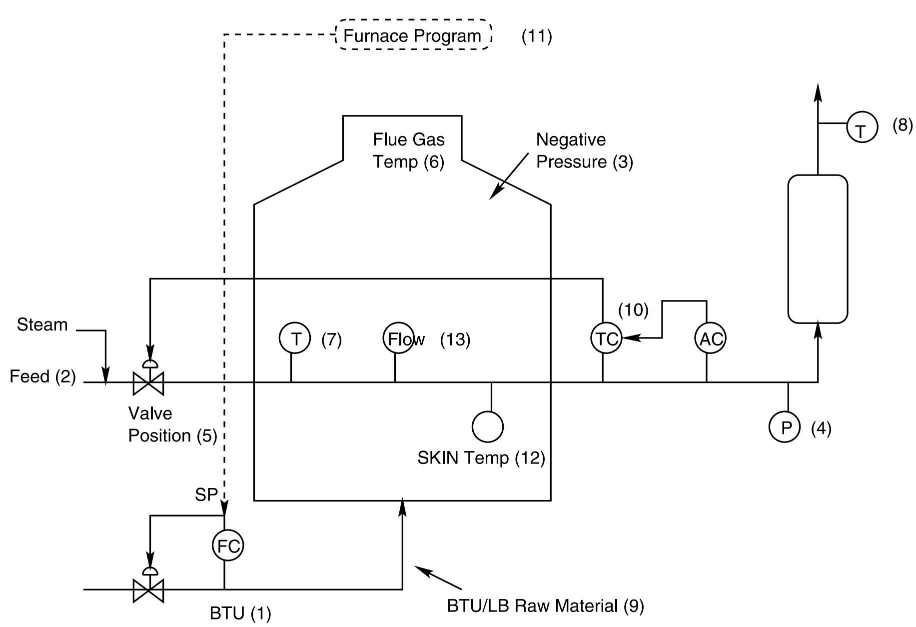

Fig. 3. Industrial tubular reactor system.

4.3.7. External reactor tube temperature

This is a direct indicator of solid deposit within tubes. A high temperature indicates poor heat transfer caused by solid layer. High temperatures may lead to mechanical damage to the tubes. Thus, the heat input is reduced to alleviate it.

4.3.8. Flue gas temperature

The flue gas temperature is one indicator of reactor health. A high temperature indicates poor heat transfer into the tubes, which indicates severe solid deposit. In this case, the heat input is cut down in order to prevent overheating of the tubes.

4.3.9. Feed flow valve position

The reaction temperature controller manipulates the feed valve directly. Under normal operations, the valve position correlates with feed rate. A lower than expected feed rate at corresponding valve position is indicative of excessive solid deposit. The fuel rate is cut to lower the chance of overheating the reactor tubes.

4.3.10. Fuel rate

Proper monitoring and control of the fuel flow rate is important, as sufficient fuel is required to sustain the reaction. But the reactors may not be able to effectively remove excess heat input to hot box if total fuel flow input is too high. A high total heat input causes overheating of tubes, resulting in (a) damage to tubes, and/or (b) accelerated solid deposit, both of which are undesirable. Fuel flow rate directly affects the tube temperature, and that in turn, affects feed rate.

4.3.11. Operator constraint

This is a mechanism to allow the operator to override fuel control program determined fuel rate. This is done when there is an indication of excessive solid deposit not reflected by above described variables, which causes automated fuel rate reduction.

The reaction system selected for this work suffered from the problem of a pressure build-up in the reactor tubes. After the pressure reaches a certain level, it is no longer economical and safe to operate the system and it has to be shut down for overhauling. A slow pressure build-up is expected as the solid continues to be deposited within the tubes. What is of concern is the rate of solid layer formation deposition that, in turn, determines the duration of a run, and the sudden increase in pressure.

The solid formation reduces heat transfer efficiency and results in heating up of the tubes, which may create hot spots within the tubes. In order to achieve desired heat transfer, it may be required to raise the reactor external temperature to the point of causing equipment damage. Thus, a typical action is to reduce the heat input to an acceptable level. Clearly, solid deposit is not desirable because it reduces heat transfer that, in turn, increases operating cost as throughput goes down and per unit product fuel cost goes up.

A common problem is a sudden jump in pressure before the heat exchanger ((5) in Fig. 3). This is frequently caused by deposited solid falling off the tube surface, and resulting in a drop in flow. The piece of solid is then blown off and accumulates at the entry of the heat exchanger. In such a situation, since feed also serves as a coolant, the temperature will go up. In order to counter the temperature rise, the controller increases feed flow rate. This would be indicated by the controller output. One goal of this project is to identify abnormal events, which cause excessive thermal contraction and expansion in the tubes. This is suspected to cause solid chips flaking off.

5. Experimental results

To demonstrate the effectiveness of MSPCA approach over the PCA approach, two sets of experiments are conducted. In the first set, data from an industrial reactor system are taken and analyzed for real process faults. The objective is early detection and identification of an abnormal situation. In the second set of experiments, industrial data set from a boiler system are taken and faults introduced into the data. The objective here is to detect sensor faults of different scales at the corresponding scale.

5.1. Multi-scale process fault detection in the reactor system

For the recursive MSPCA formulation, a moving window of data collected over 12 hour (approximately 750 data points) are decomposed into wavelet coefficients and an MSPCA model built as described in Section 3. Monitoring the process in windows of 12 hour worth of data was considered optimal in the given process conditions. Incoming data samples from the moving window are then fed into the MSPCA model and the SPE at each scale is computed and compared with the 95% SPE limit as calculated by the MSPCA model. If the SPE is below the 95% SPE limit, the sample is assumed to be normal and the MSPCA model is updated to include this data window also. Fig. 4 is a part of the reactor system data showing some of the variables for which the SPE violated the SPE limit. As discussed in the last section, one of the criteria of normal operation is having minimum fluctuations in feed rate and temperatures profiles. It can be seen from Fig. 4 that there are no sudden variations in feed rate and temperature till about time 1380, when there is a sudden drop followed by a sharp rise in system feed rate. It is also reflected in temperature and pressure profiles. The ob-

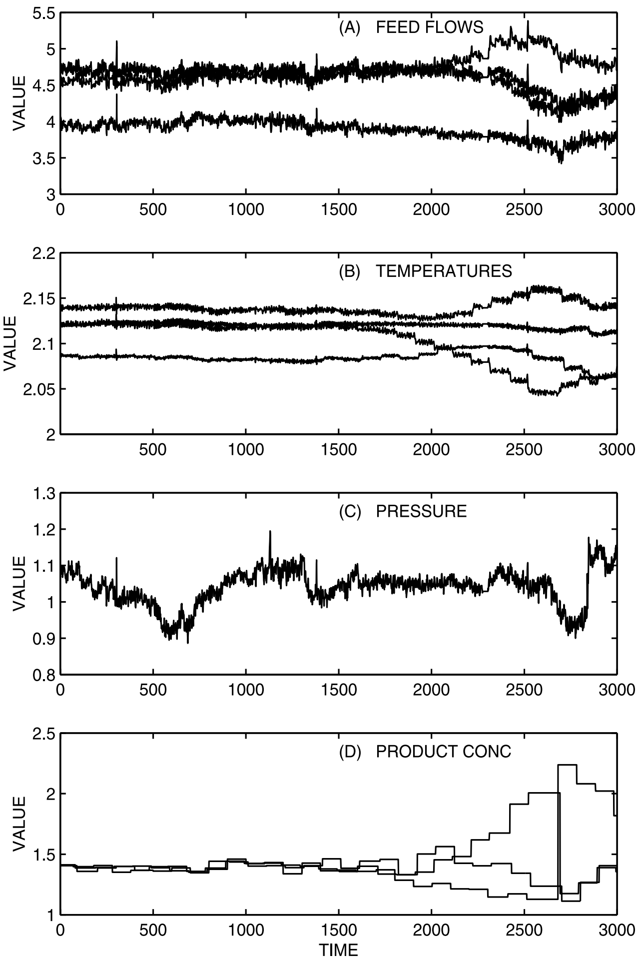

Fig. 4. Reactor system date with process fault.

served pressure variation is also indicative of feed header pressure changes. This sudden fluctuation causes a change in the feed rate that thereby significantly changes the residence time of the feed in the reactor and affects the reaction chemistry in the reactor. The change in the product composition is detected by the controller that, in turn, modifies the tube temperature setpoint, causing the temperature controller to further change the flow rate. As the reactors are physically connected, a change in the flow rate of one tube affects the flow rates of other tubes, thereby affecting the reaction chemistry and consequently the product composition in the other tubes. This triggers off a series of changes in other reactors in a fashion similar to the one described above, which leads to an abnormal situation that is developing after time 1500 and more clearly after time 2000. For efficient operation of the reactor system, this situation is not desirable and the goal is to minimize these fluctuations by early detection and identification of such situations.

Fig. 5 shows the multi-scale SPE results. Fig. 5(A) is the result of the MSPCA model built on the wavelet approximations (shown as the solid line in Fig. 2). It shows that the SPE violates the 95% SPE limit after time 1400 and the SPE keeps on increasing, signaling that a slow drifting fault at that scale had started around this time. Fig. 5(B–D) are the results of the MSPCA model built on the wavelet details (shown as the dashed and dotted lines in Fig. 2). The SPE in Fig. 5(B) is constantly below the 95% limit signaling no abnormality at that scale. Fig. 5(C) has spikes violating the 95% confidence limit at around time 1200, 1400 and 2000, whereas Fig. 5(D) has numerous spikes violating the 95% SPE limit. A sudden and abrupt fluctuation is a high frequency change and is reflected as a spike in the high scale MSPCA models. For reasons explained above, the attention is focused on the spikes around time 1400. The other spikes in Fig. 5(D) are due to local disturbances in some of the variables that are not strong enough to have an impact at other scales, and are thereby ignored.

Fig. 6 compares the recursive MSPCA formulation with the conventional PCA and the recursive PCA

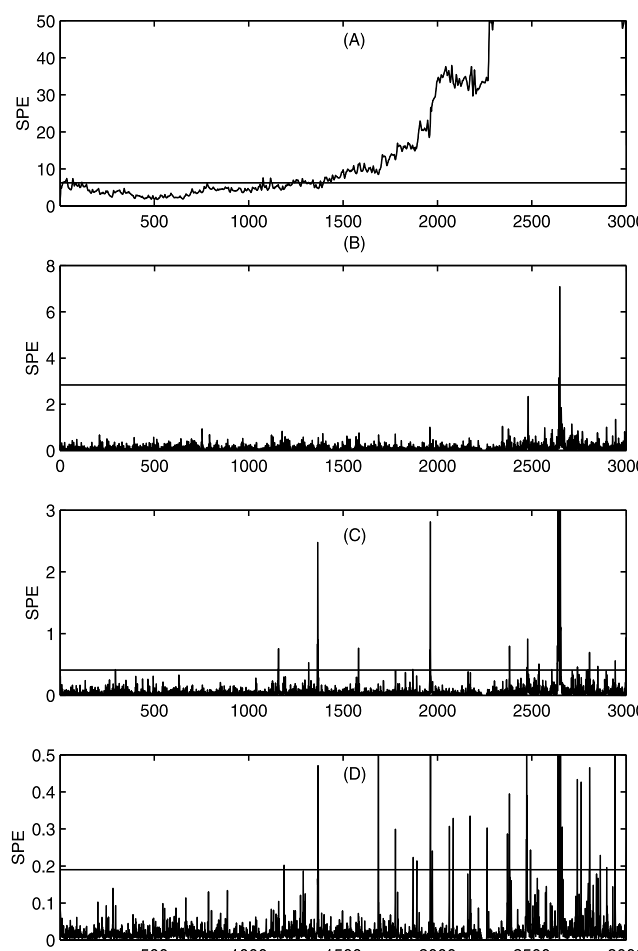

Fig. 5. Multi-scale fault detection: (A) MSPCA model of level 3 wavelet approximations, A3, (B) MSPCA model of level 3 wavelet details, D3, (C) MSPCA model of level 2 wavelet details D2, (D) MSPCA model of level 1 wavelet details, D1.

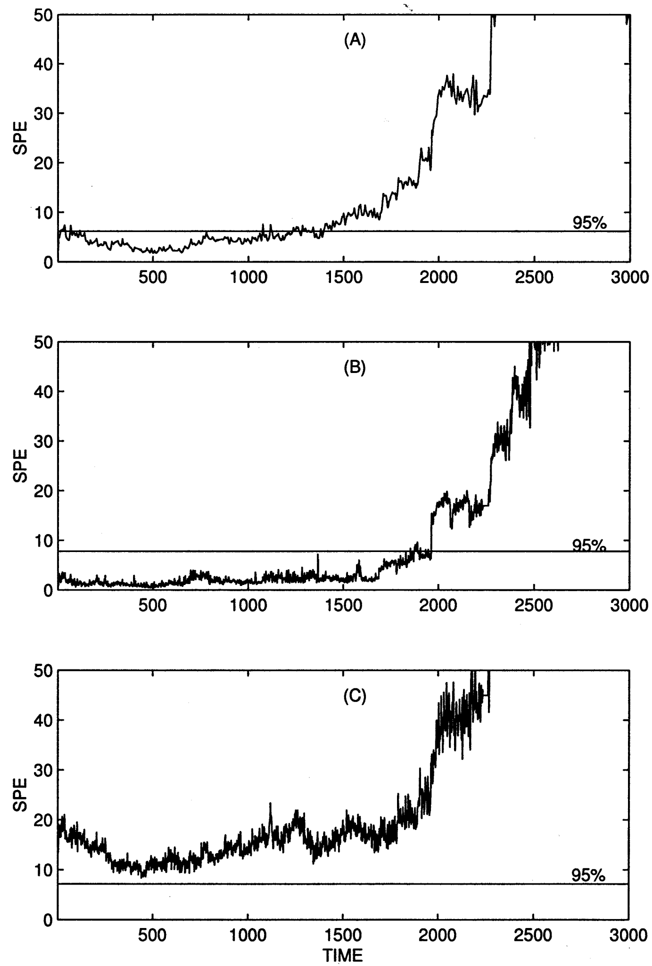

Fig. 6. Comparison of recursive MSPCA with conventional recursive PCA and conventional PCA approaches. Results with reactor system data. (A) Recursive MSPCA of wavelet approximations at level 3. (B) Recursive PCA. (C) Conventional PCA.

approaches. Fig. 6(A) is the SPE of the recursive MSPCA approach with the lowest scale approximations. Fig. 6(B) is the SPE of the recursive PCA approach while Fig. 6(C) is the SPE of the conventional PCA approach. As mentioned in Section 1, conventional PCA is best suited for analyzing steady state data and this is clearly evident from Fig. 6(C) where, although the SPE seems to be steady within a certain limit, yet it is constantly above the 95% confidence limit as determined by the conventional PCA model, thus giving a false alarm. This result demonstrates the limitation of conventional PCA that, for gradually changing processes, better results can be obtained by updating the model using recursive PCA. This can be seen from Fig. 6(B) where the SPE of the recursive PCA approach violates the 95% confidence limit only around time 2000. Thus the recursive MSPCA approach is able to detect the fault almost 600 intervals earlier than the recursive PCA.

5.2. Multi-scale fault identification

From Fig. 5 it is clear that a high frequency abnormal situation is detected at the high scales, $D3$ at time 1381 and $D2$ at time 1382. By time 1417, a drifting fault is detected at the approximation scale, $A3$. In order to identify the sensors that caused this abnormal situation, a multi-scale identification scheme is adopted. For multi-scale fault identification, contribution plots of the variables contributing to the SPE at the time when the SPE violated the limits, are generated at the corresponding scales. Fig. 7 is the contribution plot for the reactor system variables, generated at the following time and scales: Fig. 7(A) is for scale $A3$ at time 1417, Fig. 7(B) is for scale $D2$ at time 1382 and Fig. 7(C) is for scale $D1$ at time 1381. Variables 14, 18 and 19 seem to be the major contributors in the SPE at scale $D1$ violating the limit. These variables correspond to the feed rate to one of the tubes in the reactor system, implying thereby that the sudden change is first de-

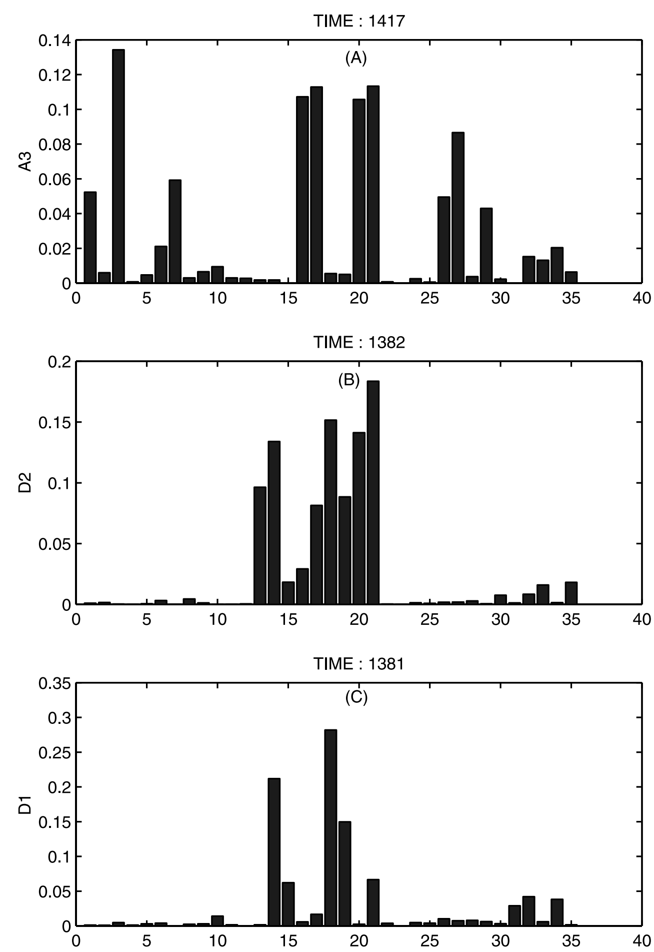

Fig. 7. Multi-scale fault identification, results with data. (A) Fault identification at wavelet approximation at level 3 at time 1417. (B) Fault identification at wavelet detail at level 2 at time 1382. (C) Fault identification at wavelet detail at level 1 at time 1381.

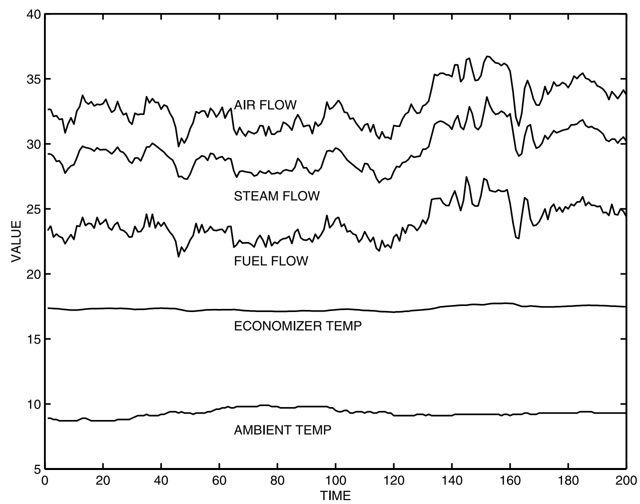

Fig. 8. Industrial boiler system data.

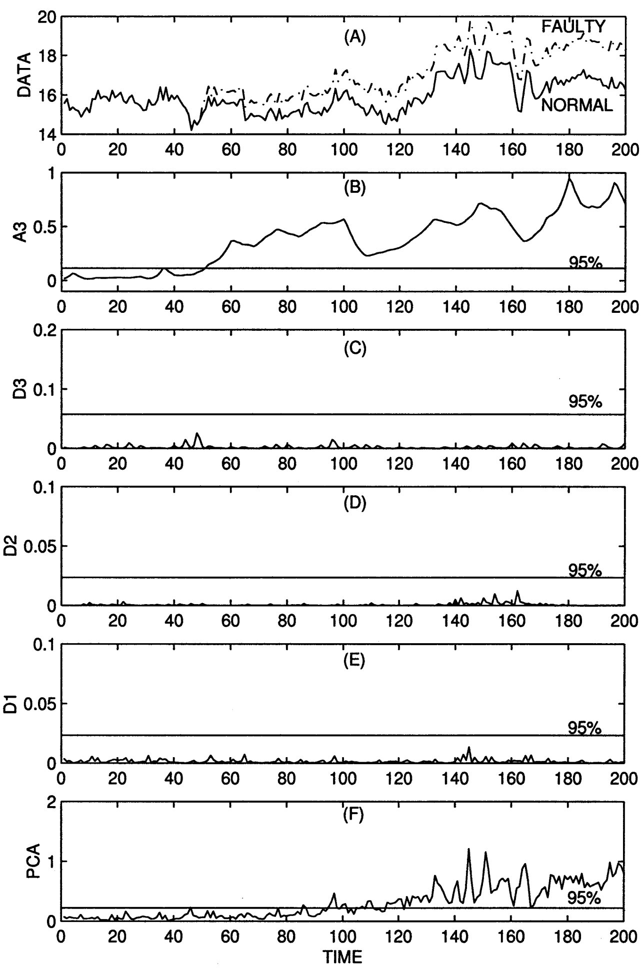

Fig. 9. Sensor fault: drift.

tected in the feed rate measurements. This result falls exactly in line with the observations made from the analysis of data set in Fig. 4. Once the flow to one set of tubes is disturbed, a corresponding change is made to the flow to the other tubes to compensate for the fact that the total flow in all the reactor tubes is normally at a constant maximum. This is validated from Fig. 7(B) where all the feed rate variables (13, 14 and 17 to 21) have a change that is detected at scale $D2$. It may also be noted from Fig. 7(B and C) that at this time, no other variable, apart from the feed rate variables, is detected as having a sharp change.

The changes in the feed rate caused corresponding changes in the residence times that affected the reaction chemistry in the reactors. Once the analyzer controller detects this change in the product compositions, corresponding changes are made to the temperature set points to counter the effect of this change. This again causes some of the feed rates to change so as to facilitate the changes in the temperature setpoints (as the feed serves as a coolant to the system). These changes in the temperature (variables 26, 27 and 29), feed rate (variables 16, 17, 20 and 21), and product composition (variables 2, 4 and 7) are detected in Fig. 7(A). These changes are secondary responses to the sudden fluctuations in the feed rate. They are caused by closed-loop control and process interactions, and are detected in the wavelet approximation scale. Furthermore, they occur at around time 1417, about 26 time intervals after the first sharp change is detected. This time lag is partly due to the analyzer controller that measures the product composition at a slower sampling rate, and partly due to the slow drift rate at the approximations scale, which is caused by various interactions.

5.3. Multi-scale sensor fault detection in boiler data

The boiler system data consists of 600 data points from eight sensor measurements sampled at a 5-min interval. Some of the variables are shown in Fig. 8. This set of experiments are basically conducted to demonstrate that sensor faults of different scales and magnitudes are detected at their corresponding scales. Four different types of sensor faults are tested: drift, bias, precision degradation and spike.

Fig. 9(A) shows a slow drift starting from time 50 in the fuel rate to the boiler. A slow drift is a low frequency change which is expected to appear in the wavelet approximation model of MSPCA. Fig. 9(B–E) are the SPE plots for the MSPCA approach. It can be seen from Fig. 9(B) that this fault is immediately detected in the approximation scale, and does not appear in the other MSPCA models at other scales. Conventional PCA, on the other hand, detects the fault at time 117. Considering the fact that the data is sampled once every 5 min, PCA detects the fault after 335 min. Thus, the MSPCA approach is not only successful in early detection but also in identification of the type of fault.

In Fig. 10(A), a constant bias is injected in the steam rate to the boiler at time 50. A bias is a sudden deviation from the normal value followed by a constant difference. The sudden deviation, being a high frequency change, is detected at the $D3$ scale of the SPE plot of the MSPCA approach in Fig. 10(C) as a sharp spike, and then as a slow change in the low frequency MSPCA model in Fig. 10(B). Here again the conventional PCA approach is late in detecting the fault. Similarly, Fig. 11(A) shows the ambient temperature readings of the boiler data corrupted with precision degradation, starting at time 50. In precision degradation, the normal data starts getting corrupted with high frequency noise that is expected to appear constantly in the high scales of an MSPCA model. Fig. 11(B) through Fig. 11(E) are results with the MSPCA approach, and the high frequency fault is clearly detected in the higher scales, $D1$ and $D2$. The conventional PCA approach fails to detect the fault. Fig. 12(A) shows a spike in the air flow rate reading at time 50. A spike is a high frequency change and is expected to show once in the high frequency scales. Fig. 12(B) through Fig. 12(E) are results with the MSPCA approach and the

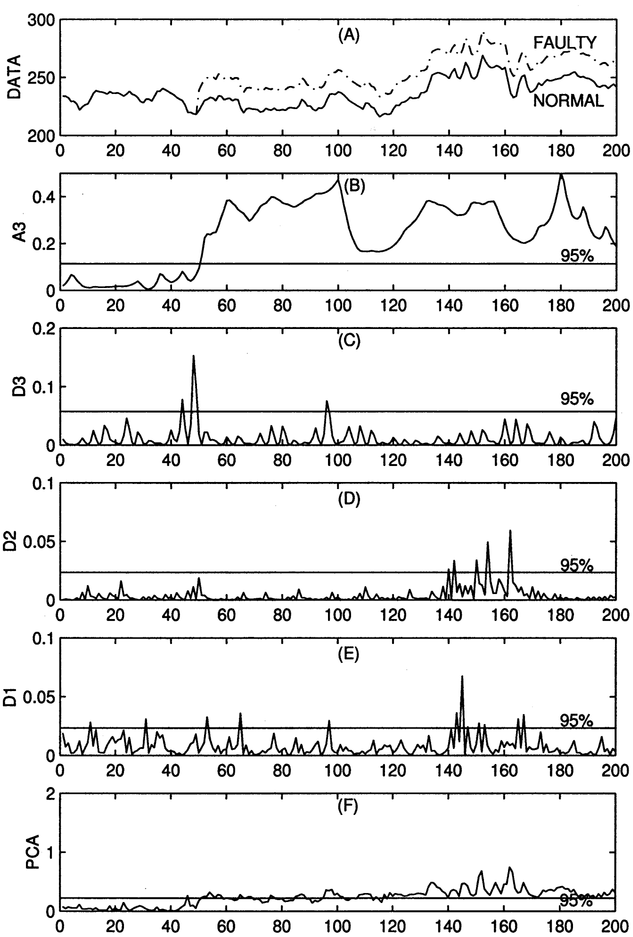

Fig. 10. Sensor fault: bias.

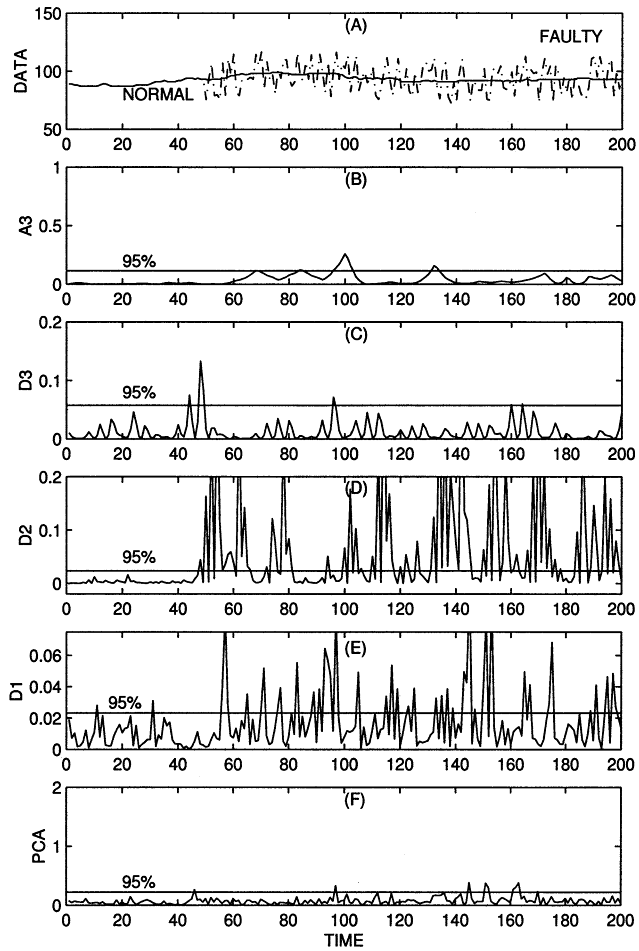

Fig. 11. Sensor fault: precision degradation.

spike is detected in all the wavelet details scales. The conventional PCA approach is also able to detect the spike.

From the results of Figs. 9–12 it is concluded that the recursive MSPCA approach outperforms the conventional PCA approach in timely detection and identification of the sensor faults.

6. Conclusions and future directions

The properties of PCA to capture correlation across variables and of wavelets to capture correlation within a variable are effectively combined in this article to formulate MSPCA. In the MSPCA formulation, the individual variables are decomposed into wavelet approximations and details at different scales. Contributions from each scale are collected in separate matrices, and a PCA model is then constructed to extract correlation at each scale. Once a significant event is detected at a scale, a multi-scale fault identification approach is used to identify the variables contributing most to the events at that particular scale. The MSPCA approach is applied to two industrial data sets for process fault diagnosis and sensor fault detection. It is able to detect and identify faults and abnormal events earlier than the conventional PCA approach.

Future directions in this work include the modification of MSPCA for use in strictly on-line applications. Multivariate process data compression is another candidate for potential application of MSPCA. Since MSPCA is known to jointly extract correlations both across and within sensors, it can effectively be used for feature extraction and multivariate process data compression. MSPCA may also be used in tasks in which PCA has demonstrated its effectiveness over other methods. Study of the modification of MSPCA for application in gross-error detection, estimation of missing data, etc., would be of interest.

Acknowledgements

Financial support from National Science Foundation (CTS-9814340), Texas Higher Education Co-ordinating Board, and Union Carbide for this work is gratefully acknowledged.

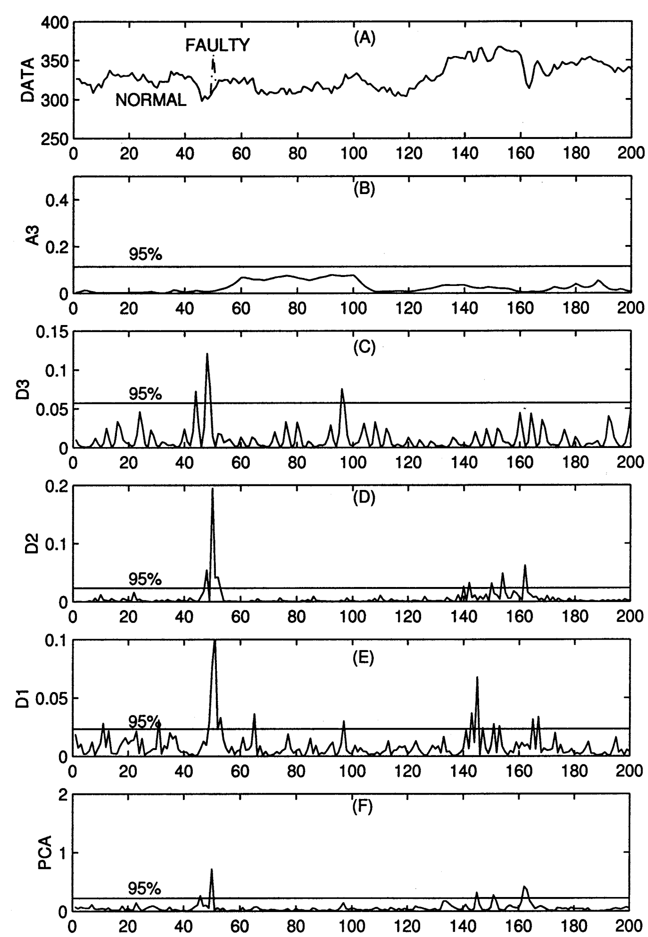

Fig. 12. Sensor fault: spike.

References

Bakshi, B. R. (1998). Multiscale PCA with applications to multivariate statistical process monitoring. *The American Institute of Chemical Engineering Journal*, *44*(7), 1596–1610.

Bakshi, B. R., & Stephanopoulos, G. (1996). Compression of chemical process data by functional approximation and feature extraction. *The American Institute of Chemical Engineering Journal*, *42*(2), 477–492.

Coifman R., & Saito, N. (1996). The local Karhunen–Loeve basis. *Proc. IEEE International Symposium on Time–Frequency and Time–Scale Analysis*, pp. 129–132, IEEE Signal Processing Society, Paris.

Dong, D., & McAvoy, T. J. (1996). Nonlinear principal component analysis-based on principal curves and neural networks. *Computers and Chemical Engineering*, *20*, 65–78.

Dunia, R., & Qin, S. J. (1998a). Subspace approach to multidimensional fault identification and reconstruction. *The American Institute of Chemical Engineering Journal*, *44*(8), 1813–1831.

Dunia, R., & Qin, S. J. (1998b). A unified geometric approach to process and sensor fault identification. *Computers and Chemical Engineering*, *22*, 927–943.

Dunia, R., Qin, S. J., Edgar, T. F., & McAvoy, T. J. (1996). Identification of faulty sensors using principal component analysis. *The American Institute of Chemical Engineering Journal*, *42*(10), 2797–2812.

Hastie, T. J., & Stuetzle, W. (1989). Principal curves. *Journal of American Statistical Association*, *84*, 502–516.

Jackson, J. E. (1991). *A user's guide to principal components*. New York: Wiley-Interscience.

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

Kosanovich, K. A., & Piovoso, M. J. (1997). PCA of wavelet transformed process data for monitoring. *Intellectual Data Analysis*, *1*(2), 1997.

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *The American Institute of Chemical Engineering Journal*, *37*, 233–243.

Kresta, J., MacGregor, J. F., & Marlin, T. E. (1991). Multivariate statistical monitoring of process operating performance. *Canadian Journal of Chemical Engineering*, *69*(1), 35–47.

Ku, W., Storer, R. H., & Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemical Intellectual Laboratory System*, *30*, 179–196.

Li, W., & Qin, S. J. (2001). Consistent dynamic PCA based on errors-in-variables subspace identification. *Journal of Process Control*, *11*(6), 661–678.

Luo, R., Misra, M., & Himmelblau, D. M. (2002). Sensor fault detection via multiscale analysis of prediction model residuals. *Industrial Engineering and Chemical Research*, revised for publication.

Luo, R., Misra, M., & Himmelblau, D. M. (1999). Sensor fault detection via multiscale analysis and dynamic PCA. *Industrial Engineering and Chemical Research*, *38*(4), 1489–1495.

Luo, R., Misra, M., Qin, S. J., Barton, R., & Himmelblau, D. M. (1998). Sensor fault detection via multiscale analysis and nonparametric statistical inference. *Industrial Engineering and Chemical Research*, *37*(3), 1024–1032.

MacGregor, J. F., & Kourti, T. (1995). Statistical process control of multivariate processes. *Controlled Engineering Practice*, *3*(3), 403–414.

MacGregor, J. F., Jaeckle, C., Kiparissides, C., & Koutoudi, M. (1994). Process monitoring and diagnosis by multiblock PLS methods. *The American Institute of Chemical Engineering Journal*, *40*, 826–838.

Misra, M., Kumar, S., Qin, S. J., & Seemann, D. (2000). Recursive on-line data compression and error analysis using wavelet technology. *The American Institute of Chemical Engineering Journal*, *46*(1), 119–132.

Misra, M., Kumar, S., Qin, S. J., & Seemann, D. (2001). Error based criterion for on-line wavelet data compression. *Journal of Process Control*, *11*(6), 717–731.

Nimmo, I. (1995). Adequately address abnormal situation operations. *Chemical Engineering Programme*, *91*(9), 36–45.

Nomikos, P., & MacGregor, J. F. (1994). Monitoring batch processes using multiway principal component analysis. *The American Institute of Chemical Engineering Journal*, *40*(8), 1361–1375.

Nomikos, P., & MacGregor, J. F. (1995). Multivariate SPC charts for monitoring batch processes. *Technometrics*, *37*(1), 41–59.

Nounou, M., & Bakshi, B. R. (1998). On-line multiscale filtering of random and gross errors without process models. *The American Institute of Chemical Engineering Journal*, *45*(5), 1041–1058.

Piovoso, M. J., Kosanovich, K. A., & Pearson, R. K. (1992). Monitoring process performance in real time. *Proceedings of ACC*, *3*, 2359–2363.

Qin, S. J., Yue, H., & Dunia, R. (1997). Self-validating inferential sensors with application to air emission monitoring. *Industrial Engineering and Chemical Research*, *36*, 1675–1685.

Raich, A., & Cinar, A. (1996). Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. *The American Institute of Chemical Engineering Journal*, *42*(4), 995–1009.

Saito, N. (1994). *Local Feature Extraction and Its Applications Using a Library of Bases*. Ph.D. thesis, Yale University.

Strang, G., & Nguyen, T. (1996). *Wavelets and filter banks*. Wellesley-Cambridge Press.

Tan, S., & Mavrovouniotis, M. L. (1995). Reducing data dimensionality through optimizing neural network inputs. *The American Institute of Chemical Engineering Journal*, *41*, 1471–1479.

Vedam, H., & Venkatasubramanian, V. (1997). Signed digraph based multiple fault diagnosis. *Computer and Chemical Engineering*, *21*(12), S655–S660.

Wickerhauser, M. V. (1994). *Adapted wavelet analysis from theory to software*. Wellesley, MA: A.K. Peters.

Wise, B. M., & Ricker, N. L. (1991). Recent advances in multivariate statistical process control: improving robustness and sensitivity. In *Proceedings of the IFAC. ADCHEM Symposium* (pp. 125–130).

Wise, B. M., Ricker, N. L., Veltkamp, D. F., & Kowalski, B. R. (1990). A theoretical basis for the use of principal component models for monitoring multivariate processes. *Process Quality and Control*, *1*, 41–51.

Wold, S., Kettaneh, N., & Tjessem, K. (1996). Hierarchical multiblock PLS and PC model for easier model interpretation and as an alternative to variable selection. *Journal Chemometrics*, *10*, 463.