15th Triennial World Congress, Barcelona, Spain

DETECTION OF PROCESS MODEL CHANGE IN PLS BASED PERFORMANCE MONITORING

Sukhbinder Kumar, Elaine B. Martin and Julian Morris

Centre for Process Analytics & Control Technology
Merz Court, University of Newcastle
Newcastle Upon Tyne, NE1 7RU, England

Abstract: The detection of process changes using a partial least squares (PLS) based monitoring scheme can be achieved through the interrogation of two metrics, Hotelling's T^2 and the Q-statistic. The Q-statistic has been shown to be insensitive to small changes in the process model parameters. In this paper, a modified statistic based on the local approach is proposed to detect changes in the model parameters in a PLS based monitoring scheme. The performance of the Q-statistic is compared with the modified statistic through their application to fault detection in a continuous stirred tank reactor. *Copyright 2002 IFAC.*

Keywords: - Fault detection; Process monitoring; Local Approach; Statistical analysis

1. INTRODUCTION

The problem of abnormal change detection in a process has received considerable attention from the research community (e.g. Willsky, 1976; Basseville, 1988). Two main approaches have been proposed to address the problem (i) model based and (ii) knowledge-based. In the model based approach, a mathematical model of the process is established from data collected when it was working under normal operating conditions (NOC) and any abnormal change in the process is detected by comparing the behavior of the process with that predicted by the model. In the knowledge-based approach, artificial intelligence techniques (neural networks, fuzzy logic or a combination of both) are used to classify the data into different groups that correspond to different operating conditions of the process. In this paper, interest is in the model based approach.

In a complex process, hundreds of process variables can be measured that exhibit (cross) correlation. Thus the number of independent (latent) events driving the process is less than the number of measured variables. The multivariate statistical projection techniques of Principal Component Analysis (PCA) and Partial Least Squares (PLS) are empirical techniques that are suitable in such situations. These methods have been applied widely for the monitoring of processes and the detection of abnormal changes in process operation (Kresta et al. 1991; Martin et al. 1996). In this paper attention focuses on PLS for the monitoring of process performance.

In a process, it is possible to distinguish between two classes of change. The first is associated with changes in process operation that result in greater variation occurring in some process variables than that captured under NOC, but the relationship between the variables is unaffected. Statistically, these changes result in a shift in the mean value of one or more of the process variables, but the (cross) correlation structure is unaffected and hence the model is still valid for such changes. The metric used to detect these changes is Hotelling's T^2:

$$T^2 = (\hat{\mathbf{x}} - \mathbf{m_0})^T \mathbf{S}^{-1} (\hat{\mathbf{x}} - \mathbf{m}_0) \qquad (1)$$

where $\mathbf{m}_0$ is the mean and $\mathbf{S}$ is the covariance matrix calculated under NOC.

The second class of change is associated with a change in the relationship or correlation structure between the process variables. For these changes, the model determined under NOC is no longer valid.

These changes in PLS based monitoring scheme are detected by monitoring the sum of squares of the residuals (also known as the Q-statistic) either in the process variable or the quality variable space if the quality variables are available as frequently as the process variables. In the process variable space the Q-statistic is given by:

$$Q_x = (\mathbf{x} - \hat{\mathbf{x}})^T (\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{e}_x{}^T \mathbf{e}_x \qquad (2)$$

and in the quality variable space:

$$Q_y = (\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}}) = \mathbf{e}_y{}^T \mathbf{e}_y \qquad (3)$$

where $\mathbf{x}$ and $\mathbf{y}$ are the process and quality variable vectors, $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ are the PLS predictions of $\mathbf{x}$ and $\mathbf{y}$ and $\mathbf{e}_x, \mathbf{e}_y$ are the residuals.

Basseville (1993) showed that the residuals, in general, are not a sufficient statistic for detecting changes in the parameters of a system. That is, the residuals do not capture the complete information as far as changes in the parameters of a system are concerned. Therefore a statistic based on the residuals may not detect such changes or will be less sensitive. The objective of this paper is to propose a statistic based on the local approach (Benveniste et al. 1987; Zhang et al. 1994) for the detection of parameter changes in a PLS model.

2. PARTIAL LEAST SQUARES

Linear PLS is a multivariate regression method that projects process variables ($\mathbf{x} \in R^m$) and the quality variables ($\mathbf{y} \in R^d$),onto a number of latent variables, say t_j and u_j and develop a regression model between them (Geladi and Kowalski, 1986):

$$u_j = b_j t_j + e_j \quad \text{where } (j = 1, \ldots, k) \qquad (4)$$

where k is the number of latent variables and is determined by cross validation. The latent variables u_j and t_j are chosen in such a way that the correlation between them is maximised. The matrices **X** and **Y** containing L samples are decomposed as the sum of the outer products of latent variables, $\mathbf{t}_j$, and the loadings, $\mathbf{p}_j$, and the prediction $\hat{\mathbf{u}}_j$ of $\mathbf{u}_j$ and the loadings $\mathbf{q}_j$, respectively:

$$\mathbf{X} = \sum_{j=1}^{k} \mathbf{t}_j \mathbf{p}_j^T + \mathbf{E} \qquad (5)$$

$$\mathbf{Y} = \sum_{j=1}^{k} \hat{\mathbf{u}}_j \mathbf{q}_j^T + \mathbf{F} \qquad (6)$$

where **E** and **F** are the residual matrices for the matrix decomposition of **X** and **Y**.

Although partial least squares was originally used to model steady state relationships, recently its use has been extended to modelling dynamic relationships. There are two approaches. One is to augment the matrices **X** and **Y** with lagged values of the input and output data and use these matrices in the PLS algorithm (Baffi et al. 2000). An alternative approach is to fit a dynamic relationship between the latent variables $\mathbf{u}_j$ and $\mathbf{t}_j$ in equation (4) instead of a steady state relationship (Lakshminarayanan, 1997). For example, if an ARX (p, q) model is chosen to fit the relationship between the latent variables then:

$$u_j(n) = a_{j1} u_j(n-1) + \ldots + a_{jp} u_j(n-p) + b_{j1} t_j(n) + \ldots + b_{jq} t_j(n-q+1) \qquad (7)$$

where $j = 1,2\ldots k$. The PLS computation engine for dynamic relationships remains the same as for steady state PLS except that the expression for $\hat{u}_j$ is now computed from a dynamic relationship.

3. LOCAL APPROACH FOR CHANGE DETECTION

The local approach for change detection is briefly described to provide the building blocks for the multivariate extension in section 4. It is assumed that a 'true' monitored system is represented by:

$$y = f(\mathbf{\Theta}, \varphi) + \eta \qquad (8)$$

where $\mathbf{\Theta}$ is the system parameter vector, φ is a regression vector and η is noise. The parameter vector $\mathbf{\Theta}$ determines the behavior of the process. Suppose that in the NOC mode, $\mathbf{\Theta} = \mathbf{\Theta}_0$. When $\mathbf{\Theta}$ takes values other than $\mathbf{\Theta}_0$, abnormal behavior of the system is indicated. The problem of change or abnormality detection in the system can be formulated in the framework of a hypothesis test.

Given a set of observations $y_1, y_2 \ldots\ldots y_n$ from a system with parameter vector $\mathbf{\Theta}$, it is necessary to decide between the hypotheses:

$$H_0: \ \mathbf{\Theta} = \mathbf{\Theta}_0 \ \text{ for } t = 1,2,\ldots n \qquad (9)$$

$$H_1: \ \exists \text{ an instance } r \ (1 \le r \le n) \text{ such that } \begin{cases} \mathbf{\Theta} = \mathbf{\Theta}_0 & \text{for } t = 1,2,\ldots r\text{-}1 \\ \mathbf{\Theta} \ne \mathbf{\Theta}_0 & \text{for } t = r, r+1,\ldots n \end{cases}$$

In a complex system, the function, f, in Equation (8) is either not known because of a lack of understanding of the physics and chemistry of the system, or is too complex to be of any use in practical applications such as in monitoring and control. Therefore, data driven system identification techniques are commonly used to build simpler representations (models) of the system. Let the model structure selected for the system in Equation (8) be:

$$y = g(\boldsymbol{\theta}, \boldsymbol{\psi}) + e \tag{10}$$

where $\boldsymbol{\theta}$ is the parameter vector of the model, $\boldsymbol{\psi}$ is the regression vector and e is the residual. In general, $f \neq g$ and $\boldsymbol{\Theta}$ and $\boldsymbol{\theta}$ belong to vector spaces of different dimensions. When the system in Equation (8) is operating in the NOC mode, $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. From now on, a distinction is made between the system in Equation (8) and its model in Equation (10). The former is termed the '*true system*' and the latter is defined simply as the '*model*' and it is assumed that the parameter vector $\boldsymbol{\theta}_0$ is known.

To detect changes in the behavior of the *true system*, it is necessary to define a function $K(\boldsymbol{\theta}_0, y)$ that satisfies the following properties:

$$E_{\boldsymbol{\Theta}}\left[K(\boldsymbol{\theta}_0, y)\right] = 0 \quad \text{when} \quad \boldsymbol{\Theta} = \boldsymbol{\Theta}_0 \tag{11}$$

$$E_{\boldsymbol{\Theta}}\left[K(\boldsymbol{\theta}_0, y)\right] \neq 0 \quad \text{when} \quad \boldsymbol{\Theta} \neq \boldsymbol{\Theta}_0 \tag{12}$$

where $E_{\boldsymbol{\Theta}}$ denotes the expectation when the parameter of the *true system* is $\boldsymbol{\Theta}$. The function $K(\boldsymbol{\theta}_0, y)$ is known as the monitoring statistic or primary residual (Bassseville, 1998). Benveniste et al. (1987) and Zhang et al. (1994) developed local asymptotic approach in which the monitoring statistic can be associated with system identification procedures. To understand the approach of Zhang et al. (1994), it is necessary to understand how the parameter vector $\boldsymbol{\theta}_0$ of the model corresponding to the NOC mode of the *true system* is identified. The parameter $\boldsymbol{\theta}_0$ is determined by minimizing some criterion function $V(\boldsymbol{\theta})$:

$$\boldsymbol{\theta}_0 = \arg \min_{\boldsymbol{\theta}} \; V(\boldsymbol{\theta}) \tag{13}$$

The most commonly used function is the expectation of the square of the residual:

$$V(\boldsymbol{\theta}) = E_{\boldsymbol{\Theta}0}[(y - g(\boldsymbol{\theta}, \boldsymbol{\psi})^2] = E_{\boldsymbol{\Theta}0}[e^2] \tag{14}$$

$$\boldsymbol{\theta}_0 = \arg . \frac{\partial}{\partial \boldsymbol{\theta}} E_{\boldsymbol{\Theta}0}[e^2] = 0 \tag{15}$$

Exchanging the expectation and differential operators, and choosing $K(\boldsymbol{\theta}, y) = \frac{\partial}{\partial \boldsymbol{\theta}} e^2$ gives:

$$\boldsymbol{\theta}_0 = \arg . \, E_{\boldsymbol{\Theta}0}[K(\boldsymbol{\theta}, y)] = 0 \tag{16}$$

Thus $E_{\boldsymbol{\Theta}0}[K(\boldsymbol{\theta}_0, y)] = 0$ and

$$E_{\boldsymbol{\Theta}}[K(\boldsymbol{\theta}_0, y)] \neq 0 \quad \text{if} \quad \boldsymbol{\Theta} \neq \boldsymbol{\Theta}_0 \tag{17}$$

This statistic:

$$K(\boldsymbol{\theta}_0, y) = \frac{\partial}{\partial \boldsymbol{\theta}} e^2 \bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}0} \tag{18}$$

therefore satisfies the properties of the primary residuals defined in Equations (11) and (12). Now let $h(\boldsymbol{\Theta}_0, \boldsymbol{\theta}) = E_{\boldsymbol{\Theta}0}[K(\boldsymbol{\theta}, y)]$, then from Equation (16):

$$h(\boldsymbol{\Theta}_0, \boldsymbol{\theta}_0) = 0 \tag{19}$$

If $\boldsymbol{\Theta}$ is any other parameter of the system (that is the process is operating under fault conditions) and $\boldsymbol{\theta}$ is the corresponding identified model parameter then:

$$h(\boldsymbol{\Theta}, \boldsymbol{\theta}) = 0 \qquad \forall \; \boldsymbol{\Theta} \text{ and } \boldsymbol{\theta}. \tag{20}$$

Now if the function $h(\boldsymbol{\Theta}, \boldsymbol{\theta}) = E_{\boldsymbol{\Theta}}[K(\boldsymbol{\theta}, y)]$ is continuously differentiable, then $h(\boldsymbol{\Theta}, \boldsymbol{\theta}) = 0$ defines a unique mapping between $\boldsymbol{\Theta}$ and $\boldsymbol{\theta}$. That is:

$$\boldsymbol{\theta} = \mathbf{b}(\boldsymbol{\Theta}). \tag{21}$$

where the function $\mathbf{b}$ is known as the bias function. The change detection problem, Equation (9), for the *true system*, can therefore be formulated as:

$$\begin{aligned} &H_0 : \; \mathbf{b}(\boldsymbol{\Theta}) = \boldsymbol{\theta}_0 \text{ for } t = 1,2.......n \\ &H_1 : \; \exists \text{ an instance r } (1 \leq r \leq n) \text{ such that} \\ &\quad \begin{cases} \mathbf{b}(\boldsymbol{\Theta}) = \boldsymbol{\theta}_0 & \text{for } t = 1,2,\ldots r\text{-}1 \\ \mathbf{b}(\boldsymbol{\Theta}) \neq \boldsymbol{\theta}_0 & \text{for } t = r, r+1, \ldots n \end{cases} \end{aligned} \tag{22}$$

For the design and analysis of a change detection algorithm based on the statistic $K(\boldsymbol{\theta}_{0,} y)$, the underlying distribution function is required. Unfortunately, the distribution function of $K(\boldsymbol{\theta}_0, y)$ is difficult to compute. To overcome this problem, the hypotheses in Equation (22) are replaced by the so called local hypotheses. In the local approach of hypotheses testing it is assumed that the null and alternative hypotheses come closer as the sample size increases. That is, under the local approach the null and alternative hypotheses are:

$$\begin{aligned} &H_0 : \; \mathbf{b}(\boldsymbol{\Theta}) = \boldsymbol{\theta}_0 \text{ for } t = 1,2.......n \\ &H_1 : \exists \text{ an instance r } (1 \leq r \leq n) \text{ such that} \\ &\begin{cases} \mathbf{b}(\boldsymbol{\Theta}) = \boldsymbol{\theta}_0 & \text{for } t = 1,2,\ldots r\text{-}1 \\ \mathbf{b}(\boldsymbol{\Theta}) = \boldsymbol{\theta}_0 + (v/n)\boldsymbol{\gamma} & \text{for } t = r, r+1, \ldots n \end{cases} \end{aligned} \tag{23}$$

where v is an unknown scalar and γ is a vector with $\|\gamma\|=1$. It should be noted that the local hypotheses of change detection approach is especially suitable for detecting small changes in the process parameters. This approach has also been used to design and analyze likelihood ratio based testing procedures (Basseville, 1993). It was shown in Zhang et al. (1994) that under the local approach of hypotheses testing, the cumulative sum:

$$S_n(\boldsymbol{\theta}_0)=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}K(\boldsymbol{\theta}_0,y_t) \tag{24}$$

converges weakly to Brownian motion as $n\to\infty$ under both the hypotheses H_0 and H_1 with drift equal to zero under H_0 and a non-zero time varying drift under H_1. This is equivalent to assuming that $K(\boldsymbol{\theta}_0,y_t)$ is an independent, identically distributed (i.i.d.) Gaussian process with zero mean under the null hypothesis, H_0, and a non-zero mean under the hypotheses, H_1, with the same covariance matrix under both the hypotheses. The problem of change detection in the parameters of the system therefore reduces to the detection of changes in the mean of an i.i.d. Gaussian process, $K(\boldsymbol{\theta}_0,y_t)$. This can be solved using the generalized likelihood ratio (GLR) test (Basseville, 1993).

4. LOCAL APPROACH TO DETECT CHANGE IN A PLS MODEL

The above approach is used to detect changes in the parameters of the PLS model. Two cases are considered (i) the quality variables are available as frequently as the process variables and, (ii) the sampling frequency of the quality variables is less than for the process variables.

Case 1: Assume that an ARX (p, q) model is fitted between each pair of variables u_j and t_j. From Equation (7):

$$\mathbf{u}=\boldsymbol{\Phi}_{\mathbf{xy}}\boldsymbol{\theta}+\mathbf{e} \tag{25}$$

where $\boldsymbol{\Phi}_{\mathbf{xy}}$ is a $k\times(p+q)k$ block diagonal matrix with the j^{th} diagonal element equal to $[u_j(n-1),\ldots u_j(n-p),t_j(n),t_j(n-1),\ldots t_j(n-q+1)]$, $\boldsymbol{\theta}=[a_{11},..a_{1p}\ b_{11}..b_{1q}\ \ldots a_{k1}\ldots a_{kp}\ b_{k1}\ldots b_{kq}]^T$ is a parameter vector, $\mathbf{u}=[u_1,u_2\ldots u_k]^T$ and $\mathbf{e}=[e_1,e_2\ldots e_k]^T$ is a vector of residuals. When the (cross) correlation structure of the process changes, it will be reflected in the parameter vector $\boldsymbol{\theta}$. The local approach can be used to detect this change. The multivariate version of the statistic in equation (18), given the measurements of the process and quality variable vectors, $x(t)$ and $y(t)$ are given by:

$$K(\boldsymbol{\theta}_0,\mathbf{x}(t),\mathbf{y}(t))=\frac{\partial}{\partial\boldsymbol{\theta}}(\mathbf{e}^T\mathbf{e})=\frac{\partial}{\partial\boldsymbol{\theta}}(\mathbf{u}-\hat{\mathbf{u}})^T(\mathbf{u}-\hat{\mathbf{u}})\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

where $\hat{\mathbf{u}}=\boldsymbol{\Phi}_{\mathbf{xy}}\boldsymbol{\theta}$, and $\boldsymbol{\theta}_0$ is the parameter vector identified under NOC. By using differential vector calculus, the primary residual at time t is given by:

$$\mathbf{K}(\boldsymbol{\theta}_0,\mathbf{x}(t),\mathbf{y}(t))=\boldsymbol{\Phi}_{\mathbf{xy}}{}^T(t)\mathbf{e}(t) \tag{26}$$

It is noted that $\mathbf{K}(\boldsymbol{\theta}_0,\mathbf{x}(t),\mathbf{y}(t))$ is a vector of dimension $(p+q)\times k$. The GLR test can now be applied to $\mathbf{K}(\boldsymbol{\theta}_0,\mathbf{x}(t),\mathbf{y}(t))$ to detect the change.

Case 2: Now consider the situation where the quality variables are not measured as frequently as the process variables. This situation is more common in practice. The prediction of the observed process variable vector $\mathbf{x}$ is given by:

$$\hat{\mathbf{x}}=t_1\mathbf{p_1}+t_2\mathbf{p_2}+\ldots.t_k\mathbf{p}_k \tag{27}$$

where $\mathbf{p_i}\in\mathbf{R}^m$, $t_i\in\mathbf{R}$; $i=1,2....k$. The prediction error vector $\mathbf{e}\in\mathbf{R}^m$ is then given by:

$$\mathbf{e}=\mathbf{x}-\hat{\mathbf{x}} \tag{28}$$

Writing (27) in component form:

$$\hat{\mathbf{x}}=\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\theta} \tag{29}$$

where $\boldsymbol{\Phi}_{\mathbf{x}}$ is a $m\times mk$ block diagonal matrix, with the diagonal element equal to $[t_1,t_2,\ldots t_k]$. $\boldsymbol{\theta}=[p_{11}..p_{k1},p_{12}..p_{k2},\ldots p_{1m}..p_{km}]^T$ is an $mk\times\mathbf{1}$ vector of the components of loading vector and p_{ij} denotes the j^{th} component of the i^{th} loading vector. Again using the multivariate version of the statistic in equation (18), the primary residual at time *t* given the measurement of process variable vector $x(t)$ is given by:

$$\mathbf{K}(\boldsymbol{\theta}_0,\mathbf{x}(t))=\frac{\partial}{\partial\boldsymbol{\theta}}(\mathbf{e}^T\mathbf{e})=\frac{\partial}{\partial\boldsymbol{\theta}}(\mathbf{x}-\hat{\mathbf{x}})^T(\mathbf{x}-\hat{\mathbf{x}})\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \tag{30}$$

Here $\boldsymbol{\theta}_0$ denotes the vector of loading vector components in NOC mode. Note that $\mathbf{K}(\boldsymbol{\theta}_0,\mathbf{x}(t))$ is a vector of dimension $m\times k$. The statistic at time *t* is then given by:

$$\mathbf{K}(\boldsymbol{\theta}_0,\mathbf{x}(t))=\boldsymbol{\Phi}_{\mathbf{x}}{}^T(t)\mathbf{e}(t) \tag{31}$$

The Generalized Likelihood Ratio (GLR) test (Basseville, 1993) can now be applied to $\mathbf{K}(\boldsymbol{\theta}_0, \mathbf{x}(t))$ to detect the change. A summary of the GLR test is given below. At time n, for $r \le n$:

$$\mathbf{D}_r^{\,n} = (n-r+1)^{-1/2} \sum_{t=r}^{n} \mathbf{K}(\boldsymbol{\theta}_0, .) \tag{32}$$

$$S_r^{\,n} = (\mathbf{D}_r^{\,n})^T \mathbf{R}^{-1} \mathbf{D}_r^{\,n} \tag{33}$$

are computed where **R** is the covariance matrix of $\mathbf{K}(\boldsymbol{\theta}_0, .)$ under the NOC mode of system. The decision rule is then given by:

$$\text{Decide in favor of } H_1 \text{ if } \max_{r \in W} S_r^{\,n} \ge \lambda$$

$$\text{Decide in favor of } H_0 \text{ if } \max_{r \in W} S_r^{\,n} < \lambda$$

where λ is a suitable threshold and W is a window function. The window W is a set of points within which the occurrence of a fault is sought up to the current time point n. The covariance matrix **R** is generally unknown and can be estimated from N observations of the statistics under NOC mode:

$$\mathbf{R} = \frac{1}{N} \sum_{t=1}^{N} \mathbf{K}(\boldsymbol{\theta}_0, .)\, \mathbf{K}(\boldsymbol{\theta}_0, .)^T \tag{34}$$

If the covariance matrix of the monitoring statistic is rank deficient, the covariance matrix can be decomposed as:

$$\mathbf{R} = \mathbf{A}\mathbf{D}\mathbf{A}^T \tag{35}$$

so that

$$\mathbf{R}^{-1} = \mathbf{A}\mathbf{D}^{-1}\mathbf{A}^T \tag{36}$$

where **D** is a diagonal matrix of non-zero eigenvalues and **A** is the matrix of corresponding eigenvectors. The threshold λ is calculated by observing that $S_r^{\,n}$ is distributed as a χ^2 with degrees of freedom equal to the dimension of $\mathbf{K}(\boldsymbol{\theta}_0, .)$ as $n \to \infty$, where n is time. The window, W is a moving window, where $W = \{(n-n_1)\ldots(n-n_0)\}$ and $(n_0 < n_1)$. In practice the size of the window is selected as a tradeoff between the delay in detecting the change and the false alarm rate, Zhang et al. (1994) by tuning the values of n_0 and n_1.

5. CASE STUDY
FAULT DETECTION IN A CSTR

The proposed statistic for PLS model change detection is now applied to detect faults in a continuous stirred tank reactor (CSTR) in which an irreversible heterogeneous catalytic exothermic reaction $A \to B$ takes place. A schematic of the CSTR is shown in Fig. 1.

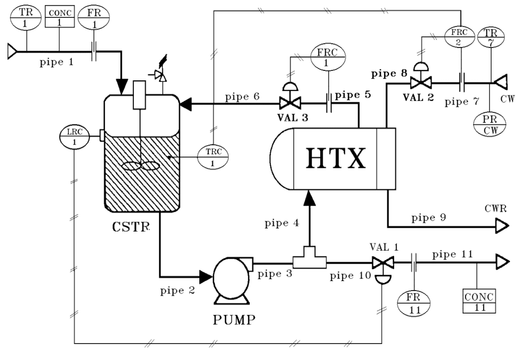

Fig. 1. Continuous Stirred Tank Reactor

The objective is to maintain the product concentration at a desired level by controlling the temperature of the reactor, the level in the reactor and the reactor mixing conditions. Temperature in the reactor is controlled by manipulating the flow rate of the coolant to the heat exchanger via a cascade control system. Manipulating the product flow rate controls level in the reactor. The mixing conditions are controlled by manipulating the recycle flow rate.

The system has eight measured process variables namely temperature of feed, concentration of feed, flow rate of feed, flow rate of product, temperature of cold water, pressure of cold water, coolant flow rate and the control signal to valve 3. The three response variables considered are temperature of the reactor, level in the reactor and recycle flow rate. To identify the model, a PRBS signal was superimposed and 1000 data points generated with a sampling time of one second. A dynamic PLS model was then built by fitting an ARX (1,1) model to each pair of latent variables. The order of the model was selected arbitrarily, although a more technical approach like Akaike Information Criterion (AIC) can be used. Three latent variables were selected using cross validation, with a total of 93 % of the variance being explained for **Y** block and 80 % for **X** block. The fault investigated was the fouling of the heat exchanger and is simulated by changing (reducing) the heat transfer coefficient of the heat exchanger from its nominal value. This fault is selected because the occurrence of this fault changes the correlation structure and hence the PLS model.

Three simulations were run. In each of the simulations the heat transfer coefficient was reduced by a different percentage at time $t = 450$, namely (i) 5 % (ii) 10 % (iii) 20 % of its normal value. The plots of the decision statistics determined from $\mathbf{K}(\boldsymbol{\theta}_0, \mathbf{y}(t), \mathbf{x}(t))$ and $\mathbf{K}(\boldsymbol{\theta}_0, \mathbf{x}(t))$ based on the local approach corresponding to the above changes in the heat transfer coefficient with 99.9 % confidence limits are shown in Figs. 2 and 3, respectively. The parameters n_0 and n_1 for the window, W , for the above confidence limit are given as 50 and 550 respectively.

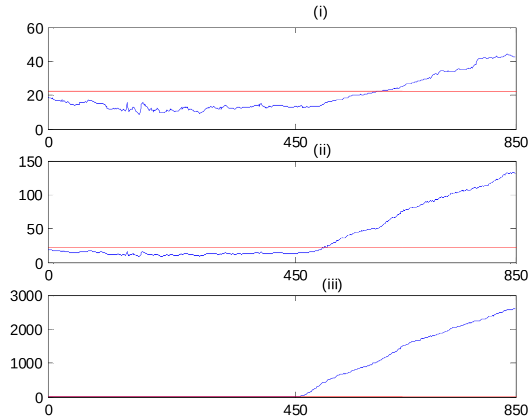

Fig. 2. Plot of decision statistic based on $\mathbf{K}(\boldsymbol{\theta}_0, \mathbf{x}(t), \mathbf{y}(t))$ versus time.

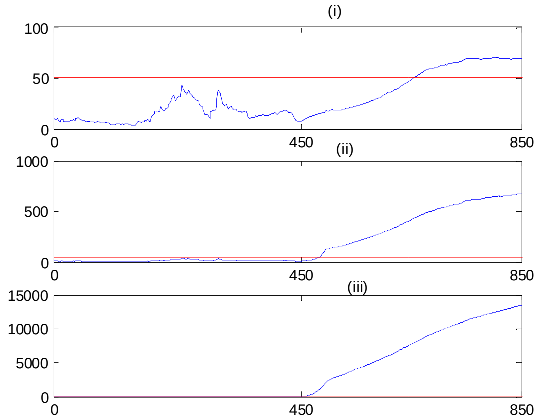

Fig. 3: Plot of decision statistics based on $\mathbf{K}(\boldsymbol{\theta}_0, \mathbf{x}(t))$ versus time.

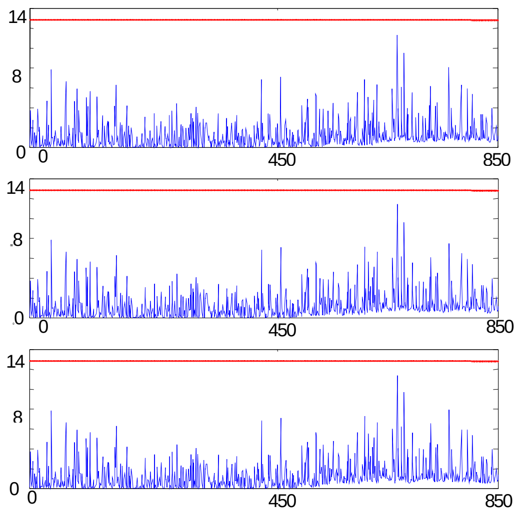

Fig. 4. $Q_{\mathbf{x}}$ statistic versus time.

The plot of the $Q_{\mathbf{x}}$ statistic that is traditionally used in MSPC is shown in Fig (4). It is clear from the Fig. 4 that the $Q_{\mathbf{x}}$ statistic is not able to detect the fault corresponding to the above changes unlike the statistic based on the local approach, Figs. 2 and 3. Although not shown, Hotelling's T^2 was unable to detect any of these faults.

6. DISCUSSION AND CONCLUSIONS

The Q-statistic has been shown to be less sensitive in terms of the detection of changes in model parameters. In this paper the local approach of Zhang et al. (1994) was extended to PLS based performance monitoring for the detection of model parameter changes. The methodology was applied to the detection of the onset of faults in a CSTR and was shown to be more sensitive to the subtle faults which are not detected by the more traditional Q-statistic. Only the step changes in the parameters of the system has been considered. The case for slow drift in parameters of the system however, needs to be investigated further.

7. ACKNOWLEDGEMENTS

S. Kumar acknowledges the financial support of the EPSRC and CPACT.

8. REFERENCES

Baffi, G., Martin, E.B., Morris, A.J., 2000. Non-linear dynamic projection to latent structures modelling:. *Chemom. Intell. Lab. Syst.* 52, 5-22.

Basseville, M. and Nikiforov, I. (1993). Detection of abrupt changes – Theory and applications. Prentice Hall Information and System Sciences Series, Englewood Cliffs, NJ.

Basseville, M (1988), Detecting Changes in Signals and Systems-A Survey, *Automatica*, 24, 309-326,

Benveniste, A. Basseville, M. and Moustakides, G. (1987). The asymptotic local approach to change detection and model validation'. *IEEE Trans. on Automatic Control*, AC-32(7), 583-592.

Geladi . P. and Kowalski. B. R (1986) Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, 18, 1-17.

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

Lakshminarayanan, S. (1997). Process characterization and control using multivariate statistical techniques, PhD Thesis, Edmonton, Alberta.

Martin, E.B., Morris, A. J. and Zhang, J. (1996). Process Performance Monitoring using Multivariate SPC, *Proceedings of IEE Control Theory and Applications Pt. D*, 143(2), 132-144

Willsky, A. S. and Jones, H. L. (1976). A generalized likelihood ratio approach to detection and estimation of jumps in linear systems. IEEE Trans. Automatic Control AC(21), 1, 108-112.

Zhang, Q., Basseville, M. and Benvenite (1994). Early warning of slight changes in systems. *Automatica*, 30(1), 95-113.