AIChE

Diagnosis of Process Faults in Chemical Systems Using a Local Partial Least Squares Approach

Uwe Kruger
Dept. of Electrical Engineering, The Petroleum Institute, Abu Dhabi, United Arab Emirates

Grigorios Dimitriadis
Dept. of Aerospace and Mechanical Engineering, Université de Liège, Chemin des chevreuils, 1, B-4000, Liège 1, Belgium

DOI 10.1002/aic.11576
Published online July 29, 2008 in Wiley InterScience (www.interscience.wiley.com).

This article discusses the application of partial least squares (PLS) for monitoring complex chemical systems. In relation to existing work, this article proposes the integration of the statistical local approach into the PLS framework to monitor changes in the underlying model rather than analyzing the recorded input/output data directly. As discussed in the literature, monitoring changes in model parameters addresses the problems of nonstationary behavior and presents an analogy to model-based approaches. The benefits of the proposed technique are that (i) a detailed mechanistic plant model is not required, (ii) nonstationary process behavior does not produce false alarms, (iii) parameter changes can be non-Gaussian, (iv) Gaussian monitoring statistics can be established to simplify the monitoring task, and (v) fault magnitude and signatures can be estimated. This is demonstrated by a simulation example and the analysis of recorded data from two chemical processes. © 2008 American Institute of Chemical Engineers *AIChE J,* 54: 2581–2596, 2008
Keywords: condition monitoring, fault detection, fault diagnosis, partial least squares, local approach

Introduction

The condition monitoring of industrial processes is standard practice in many industries, particularly the chemical industry. Based on the large body of work in the area of fault detection and diagnosis (FDD), the goals of condition monitoring are threefold: (i) to detect anomalous behavior in a process from the measured process variables, (ii) to locate and identify the source of such behavior, and (iii) to identify the magnitude of the fault. An ideal condition monitoring procedure can prevent a major malfunction in processing units by identifying incipient faults before they actually cause significant damage. Detailed discussions of condition monitoring and fault detection aspects can be found in many survey papers, for example, Refs. 1–4, or research texts.[5–7]

Techniques for process monitoring and FDD are based on a variety of paradigms including signal-based techniques[8–11] that are mainly applied to mechanical systems, model-based techniques[4,12–14] which address a wider spectrum of applications, rule-based techniques,[15–18] and more recently, knowledge-based techniques[19–22] that blend heuristic knowledge into monitoring application. Such methodologies have shown their potential whenever cost-benefit economics have justified the required effort in developing applications. A good recent overview of various FDD techniques was given in Refs. 23–26.

Correspondence concerning this article should be addressed to U. Kruger at ukruger@pi.ac.ae.

With advances in instrumentation and computer technology, the number of recorded process variables has increased to the point that some or many of the measured variables can be highly correlated.[27] In such cases, the creation of condition monitoring models is hindered by multicollinearity and suitably derived methods must be used. The most popular of these methods are statistical-based techniques, such as principal component analysis (PCA), partial least squares (PLS), and their extensions.

In a previous paper,[28] a statistical theory was noted which can be readily used for the early detection of incipient faults, known as the local approach. The theory can be used to transform the problem of detecting a fault in a stochastic process to that of monitoring the mean of a Gaussian vector that is constructed from the process observations. More precisely, rather than performing a direct analysis of the recorded variables, the local approach examines changes in model parameters.

Most fault detection concepts for technical systems are model-based, for example, observer-based or related to eigenstructure assignments and parity relations.[6] However, models for complex chemical systems are often unavailable, mainly because of the complexity associated with determining a detailed mechanistic model. The work presented in this article makes use of a model-based approach, describing input/output relationships of the recorded data. Instead of relying on an available model, the introduced FDD scheme is based on identified models using reference data from the process.

The modeling task of identifying parametric relationships between the input and output variables is performed using PLS, a technique that, in analogy to model-based approaches, identifies a linear regression model between a predictor and a response variable set. The local approach is then integrated as the process-monitoring engine. This, in turn, translates changes of the parameters of the identified PLS model into univariate monitoring statistics. The strength of this integration is that (i) nonstationary process behavior can be accommodated, (ii) simple univariate statistics are used for process monitoring, and (iii) an identification scheme can be derived for estimating fault magnitudes and signatures.

Nonstationary behavior presents a considerable challenge for data-driven techniques.[29] Existing work on dealing with nonstationary behavior relates to the use of ARIMA filters.[30] For fault detection, however, in refs. 31 and 32 it was demonstrated that step-type faults, for example, are short-lived after they are detected. This follows from the differentiation of the analyzed sequences. The theory of integrating the local approach into PCA, in comparison to PLS a method applied to a single variable set, has been discussed in Kruger et al.[33]

In a similar fashion to model-based approaches, the contribution of this work is (i) to construct input/output models from recorded data using PLS, (ii) to formulate monitoring statistics for the on-line monitoring of parameter changes in the PLS models, and (iii) to introduce a fault diagnosis scheme for estimating the fault magnitude and signature. The utility of this scheme is shown by a simulation example and application studies to recorded data from two chemical processes.

The article is divided into the following sections. Preliminaries of PLS are summarized first. Then, the statistical local approach is briefly revisited, prior to the introduction of the proposed PLS-based fault detection scheme. This is followed by the development of a fault diagnosis scheme to estimate fault magnitude and signature. Finally, the utility of the proposed monitoring scheme is demonstrated using a simulation example and the two industrial application studies.

Preliminaries

This section provides a brief overview of PLS along with data preprocessing. PLS was first developed for use in economics by Wold.[34,35] Since then PLS has been applied to many disciplines,[36] including the condition monitoring of industrial processes.[37,38]

PLS is designed to analyze relationships between and interrelationships within predictor and response variable sets. The observations for each variable set are stored in the predictor matrix $\mathbf{X} \in \mathbb{R}^{K \times N}$ and the response matrix $\mathbf{Y} \in \mathbb{R}^{K \times M}$. This matrix notation entails that each column and row refers to a particular variable and a particular observation, respectively. Typically, $\mathbf{X}$ and $\mathbf{Y}$ are mean centered and appropriately scaled. PLS determines the ith pair of weight vectors, $\mathbf{v}_i \in \mathbb{R}^M$ and $\mathbf{w}_i \in \mathbb{R}^N$, and score vectors, $\mathbf{t}_i \in \mathbb{R}^K$ and $\mathbf{u}_i \in \mathbb{R}^K$, by maximizing J_i:

$$J_i = E\{t_i u_i\} = E\{\mathbf{w}_i^T \mathbf{x}_i \mathbf{y}_i^T \mathbf{v}_i\} \approx \frac{\mathbf{t}_i^T \mathbf{u}_i}{K-1} = \frac{\mathbf{w}_i^T \mathbf{X}_i^T \mathbf{Y}_i \mathbf{v}_i}{K-1} \approx \mathbf{w}_i^T \mathbf{S}_{XY_i} \mathbf{v}_i. \quad (1)$$

Here, $\mathbf{S}_{XY_i}$ is a covariance matrix, $E\{\cdot\}$ is the expectation operator and t_i and u_i represent the ith pair of score variables of the predictor and response variables $\mathbf{x}_i$ and $\mathbf{y}_i$, respectively. Equation 1 is subject to the following constraints:

$$G_i^{(\mathbf{w}_i)} = \|\mathbf{w}_i\|_2^2 - 1 = 0, \quad G_i^{(\mathbf{v}_i)} = \|\mathbf{v}_i\|_2^2 - 1 = 0. \quad (2)$$

This produces $\mathbf{w}_i$ and $\mathbf{v}_i$ to be the most dominant eigenvectors of $\widehat{\mathbf{S}}_{XY_i} \widehat{\mathbf{S}}_{XY_i}^T$ and $\widehat{\mathbf{S}}_{XY_i}^T \widehat{\mathbf{S}}_{XY_i}$, respectively, where $\widehat{\mathbf{S}}_{XY_i} = \frac{\mathbf{X}_i^T \mathbf{Y}_i}{K-1}$ is the estimated covariance matrix. The solutions for $\mathbf{w}_i$, $\mathbf{v}_i$, $\mathbf{t}_i$, and $\mathbf{u}_i$ are given by the power method.[39] Upon convergence, the loading vectors $\mathbf{p}_i$ and $\mathbf{q}_i$ are calculated as

$$\mathbf{p}_i = \frac{\mathbf{X}^T \mathbf{t}_i}{\mathbf{t}_i^T \mathbf{t}_i} \quad \text{and} \quad \mathbf{q}_i = \frac{\mathbf{Y}^T \mathbf{t}_i}{\mathbf{t}_i^T \mathbf{t}_i}. \quad (3)$$

For determining subsequent factors, PLS uses a deflation procedure, which involves the contributions of $\mathbf{t}_i$ being subtracted, or deflated, from $\mathbf{X}_i$ and $\mathbf{Y}_i$ to produce $\mathbf{X}_{i+1}$ and $\mathbf{Y}_{i+1}$, respectively:

$$\mathbf{X}_{i+1} = \mathbf{X}_i - \mathbf{t}_i \mathbf{p}_i^T \quad \text{and} \quad \mathbf{Y}_{i+1} = \mathbf{Y}_i - \mathbf{t}_i \mathbf{q}_i^T. \quad (4)$$

When n loading vectors are obtained, the final outer model, $\mathbf{B}_n$, can be obtained by Eq. 5:

$$\mathbf{B}_n = \mathbf{W}_n [\mathbf{P}_n^T \mathbf{W}_n]^{-1} \mathbf{Q}_n^T. \quad (5)$$

The NIPALS[39] and SIMPLS[40] algorithms are the most popular PLS implementations.

Data preprocessing

To remove noise and to reveal the underlying linear relationships between the predictor and response variable sets, a singular value decomposition is applied prior to the identification of a PLS model:

$$[\mathbf{X} \quad \mathbf{Y}] = [\mathcal{U}_1 \quad \mathcal{U}_2]\begin{bmatrix} \mathcal{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathcal{S}_2 \end{bmatrix}\begin{bmatrix} \mathcal{V}_{\mathbf{X1}}^{\mathbf{T}} & \mathcal{V}_{\mathbf{Y1}}^{\mathbf{T}} \\ \mathcal{V}_{\mathbf{X2}}^{\mathbf{T}} & \mathcal{V}_{\mathbf{Y2}}^{\mathbf{T}} \end{bmatrix} \tag{6}$$

where $\mathcal{U}_1 \in \mathbb{R}^{K\times m}$ and $\mathcal{U}_2 \in \mathbb{R}^{K\times(M+N-m)}$ are the left singular vectors, $\mathcal{V}_1^{\mathbf{T}} = [\mathcal{V}_{\mathbf{X1}}^{\mathbf{T}}\ \mathcal{V}_{\mathbf{Y1}}^{\mathbf{T}}] \in \mathbb{R}^{m\times(M+N)}$ and $\mathcal{V}_2^{\mathbf{T}} = [\mathcal{V}_{\mathbf{X2}}^{\mathbf{T}}\ \mathcal{V}_{\mathbf{Y2}}^{\mathbf{T}}] \in \mathbb{R}^{(M+N-n)\times(M+N)}$ are the right singular vectors, and $\mathcal{S}_1 \in \mathbb{R}^{m\times m}$ and $\mathcal{S}_2 \in \mathbb{R}^{(M+N-m)\times(M+N-m)}$ are diagonal matrices storing the singular values. The separation of Eq. 6 relies on the singular values, where the m ones stored in $\mathcal{S}_1$ are considerably larger than those in $\mathcal{S}_2$. Reformulating Eq. 6 gives rise to

$$[\mathbf{X} \quad \mathbf{Y}] = \mathcal{U}_1\mathcal{S}_1[\mathcal{V}_{\mathbf{X1}}^{\mathbf{T}} \quad \mathcal{V}_{\mathbf{Y1}}^{\mathbf{T}}] + \mathcal{U}_2\mathcal{S}_2[\mathcal{V}_{\mathbf{X2}}^{\mathbf{T}} \quad \mathcal{V}_{\mathbf{Y2}}^{\mathbf{T}}] \tag{7}$$

which can be rewritten as follows:

$$\mathbf{X} = \mathcal{X} + \mathcal{E} \quad \mathbf{Y} = \mathcal{Y} + \mathcal{F} \tag{8}$$

where $\mathcal{X} = \mathcal{U}_1\mathcal{S}_1\mathcal{V}_{\mathbf{X1}}^{\mathbf{T}}$, $\mathcal{E} = \mathcal{U}_2\mathcal{S}_2\mathcal{V}_{\mathbf{X2}}^{\mathbf{T}}$, $\mathcal{Y} = \mathcal{U}_1\mathcal{S}_1\mathcal{V}_{\mathbf{Y1}}^{\mathbf{T}}$, and $\mathcal{F} = \mathcal{U}_2\mathcal{S}_2\mathcal{V}_{\mathbf{Y2}}^{\mathbf{T}}$.

The matrices $\mathcal{E}$ and $\mathcal{F}$ are residual matrices of the predictor and response matrices that represent sensor noise and $\mathcal{Y} = \mathcal{X}\mathbf{B}$ and $\mathcal{X} = \mathcal{Y}\mathbf{C}$ are the linear relationships between the predictor and response variables with $\mathbf{B} = [\mathcal{V}_{\mathbf{X1}}^{\mathbf{T}}]^{\dagger}\ \mathcal{V}_{\mathbf{Y1}}^{\mathbf{T}}$ and $\mathbf{C} = [\mathcal{V}_{\mathbf{Y1}}^{\mathbf{T}}]^{\dagger}\ \mathcal{V}_{\mathbf{X1}}^{\mathbf{T}}$ and † being the generalized inverse.

For the subsequent PLS analysis, the matrices $\mathbf{X}$ and $\mathbf{Y}$ are used to denote $\mathcal{X}$ and $\mathcal{Y}$, respectively. The use of $\mathcal{X}$ and $\mathcal{Y}$ is useful for producing a dual model approach, detailed below. The next section summarizes the principles of the local approach, which is to be blended into the conventional PLS-based monitoring scheme. For the remainder of this article, Eqs. 1–3 are referred to as characteristic PLS equations.

The Statistical Local Approach

The basis of the statistical local approach is the central limit theorem, which can be stated as follows: The sum of a series of K samples, $\boldsymbol{\zeta} = \Sigma_{k=1}^{K}\boldsymbol{\zeta}_k$ where each sample $\boldsymbol{\zeta}_k$ has the same distribution function, asymptotically follows a multinormal distribution. To see how this relates to condition monitoring consider a model of a process described by the parameter vector $\boldsymbol{\theta}$. Denote by $\boldsymbol{\theta}_0$ the parameters describing normal operation and by $\mathbf{z}_k$ the data vector at the kth instance. Fault detection is performed by monitoring the primary residual function $\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_k)$ under the conditions that $E\{\boldsymbol{\phi}(\boldsymbol{\theta}, \mathbf{z}_k)\} = \mathbf{0}$ if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, $\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_k)$ exists and $E\{\boldsymbol{\phi}(\boldsymbol{\theta}, \mathbf{z}_k)\} \neq \mathbf{0}$ if $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ but $\boldsymbol{\theta}$ is in the neighborhood of $\boldsymbol{\theta}_0$, and finally, $\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_k)$ is differentiable with respect to $\boldsymbol{\theta}$.

In Ref. 28 the local approach was applied to the primary residual function by assuming that any deviations from $\boldsymbol{\theta}$ are local to $\boldsymbol{\theta}_0$, i.e. by describing an incipient fault condition by

$$\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \frac{\Delta\boldsymbol{\theta}}{\sqrt{K}} \tag{9}$$

Using the central limit theorem, an improved residual can be defined as

$$\mathcal{Z}_K(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{K}}\sum_{k=1}^{K}\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_k) \tag{10}$$

so that $\mathcal{Z}_K(\boldsymbol{\theta}_0)$ follows a multinormal distribution of zero mean and covariance

$$\mathbf{S}_{\phi\phi} = \lim_{K\to\infty}\frac{1}{K}\sum_{i=1}^{K}\sum_{j=1}^{K}\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_i)\boldsymbol{\phi}^T(\boldsymbol{\theta}_0, \mathbf{z}_j) \tag{11}$$

if no fault condition is present. It should be noted that the covariance matrix of $\mathcal{Z}_K(\boldsymbol{\theta}_0 + \Delta\boldsymbol{\theta}/\sqrt{K})$ is equal to that of $\mathcal{Z}_K(\boldsymbol{\theta}_0)$ but, as stated above, the mean value $E\left\{\mathcal{Z}_K\left(\boldsymbol{\theta}_0 + \frac{\Delta\boldsymbol{\theta}}{\sqrt{K}}\right)\right\}$ departs from zero. Using this fact, a Hotelling's T^2 criterion can be defined such that

$$T_K^2 = \mathcal{Z}_K^T\mathbf{S}_{\phi\phi}^{-1}\mathcal{Z}_K \tag{12}$$

Now, normal operation is described by $T_K^2 \leq T_0^2$, where T_0 is the 95 or 99% confidence limit for a χ^2 distribution for which the number of degrees of freedom is equal to the dimension of $\mathcal{Z}_K$. Conversely, faulty operation is described by $T_K^2 > T_0^2$.

Integrating the Local Approach into PLS Monitoring

This section shows how to detect fault conditions in chemical systems by incorporating the statistical local approach into PLS-based process monitoring. We first introduce PLS-based functions to be considered as primary residual functions, required for the statistical local approach. Based on these functions, the section then establishes improved residual functions and subsequently shows how to construct univariate monitoring statistics.

Introduction of primary residual functions

The first step involves the identification of a PLS model using reference data from the process to be monitored. This model includes the weight and loading matrices $\mathbf{W}_N$, $\mathbf{V}_N$, $\mathbf{P}_N$, $\mathbf{Q}_N$, $\mathbf{B}_N$ and the vector $\mathbf{J}_N$, whose elements are the maximum of J_i. The objective of the primary residuals is to utilize the characteristic PLS equations to identify incipient faults as early as possible.

The discussion in the previous section shows that the primary residuals must have an expected value of zero under normal operation. To obtain such primary residuals from the characteristic PLS equations the cost function for the ith $\mathbf{w}$- and $\mathbf{v}$-vectors is utilized here:

$$J_i = \arg\max_{\mathbf{w},\mathbf{v}}\{E\{\mathbf{w}_i^T\mathbf{x}_i\mathbf{y}_i^T\mathbf{v}_i\}\}. \tag{13}$$

With $G_i^{(\mathbf{w}_i)} = \|\mathbf{w}_i\|_2^2 - 1$ and $G_i^{(\mathbf{v}_i)} = \|\mathbf{v}_i\|_2^2 - 1$, reformulating Eq. 13 gives rise to

$$\mathbf{w}_i = \arg\max_{\mathbf{w}}\left\{J_i - \lambda_{1,i}G_i^{(\mathbf{w}_i)}\right\} \Leftrightarrow \arg\left\{\frac{\partial J_i}{\partial \mathbf{w}_i} - \lambda_{1,i}\frac{\partial G_i^{(\mathbf{w}_i)}}{\partial \mathbf{w}_i}\right\} = 0 \tag{14}$$

$$\mathbf{v}_i = \arg\max_{\mathbf{v}} \left\{ J_i - \lambda_{2,i} G_i^{(\mathbf{v}_i)} \right\} \Leftrightarrow \arg \left\{ \frac{\partial J_i}{\partial \mathbf{v}_i} - \lambda_{2,i} \frac{\partial G_i^{(\mathbf{v}_i)}}{\partial \mathbf{v}_i} \right\} = 0 \tag{15}$$

where $\lambda_{1,i}$ and $\lambda_{2,i}$ are Lagrangian multipliers. The relationships between $\lambda_{1,i}$, $\lambda_{2,i}$ and the cost function J_i are governed by the following lemma.

Lemma 1. The two Lagrangian multipliers possess an identical value, that is, $\lambda_{1,i} = \lambda_{2,i} \triangleq \lambda_i$, and the cost function J_i is twice the value of λ_i.

A proof of Lemma 1 is given in Appendix A. Utilizing Lemma 1, and working out the partial derivatives of Eqs. 14 and 15 produces

$$\begin{pmatrix} \mathbf{w}_i \\ \mathbf{v}_i \end{pmatrix} = \arg \begin{Bmatrix} E\{\mathbf{x}_i^T \mathbf{y}_i\}\mathbf{v}_i - J_i \mathbf{w}_i \\ E\{\mathbf{y}_i^T \mathbf{x}_i\}\mathbf{w}_i - J_i \mathbf{v}_i \end{Bmatrix} = 0. \tag{16}$$

Finally, defining

$$\mathbf{g}_i \triangleq \begin{pmatrix} \mathbf{x}_i \mathbf{y}_i^T \mathbf{v}_i - \frac{K-1}{K} J_i \mathbf{w}_i \\ \mathbf{y}_i \mathbf{x}_i^T \mathbf{w}_i - \frac{K-1}{K} J_i \mathbf{v}_i \end{pmatrix} \tag{17}$$

allows Eq. 16 to be reformulated as

$$\begin{pmatrix} \mathbf{w}_i \\ \mathbf{v}_i \end{pmatrix} = \arg \{ E\{\mathbf{g}_i\} \} = 0 \tag{18}$$

The functions $\mathbf{g}_i$ are suitable for the generation of primary residuals, since it is clear that under nominal operating conditions $E\{\mathbf{g}_i\} = 0$. Conversely, for faulty operating conditions in the neighborhood of the nominal conditions, it must be concluded that $E\{\mathbf{g}_i\} \neq 0$.

There is a total of $N \times (N + M)$ functions. In most practical PLS monitoring applications, however, some of the latent variables contribute very little to the covariance structure of the data[27,41,42] and can therefore be discarded. Retaining n latent variables leads to the definition of the following primary residuals:

$$\mathcal{G}_{R0}^T = \begin{pmatrix} \mathbf{g}_1^T & \mathbf{g}_2^T & \cdots & \mathbf{g}_n^T \end{pmatrix} \in \mathbb{R}^{(N+M)\times n} \tag{19}$$

$$\mathcal{G}_{D0}^T = \begin{pmatrix} \mathbf{g}_{n+1}^T & \mathbf{g}_{n+2}^T & \cdots & \mathbf{g}_N^T \end{pmatrix} \in \mathbb{R}^{(N+M)\times(N-n)} \tag{20}$$

where the subscript R denotes primary residuals constructed from retained latent variables and D denotes those constructed from discarded latent variables. The subscript 0 refers to primary residuals whose statistical properties represent normal operating conditions.

A further set of monitoring functions can be obtained by pre-multiplying the first line of the definition of $\mathbf{g}_i$, Eq. 17, by $\mathbf{w}_i^T$, i.e.

$$\mathbf{w}_i^T \mathbf{x}_i \mathbf{y}_i^T \mathbf{v}_i - \frac{K-1}{K} J_i \mathbf{w}_i^T \mathbf{w}_i = t_i u_i - \frac{K-1}{K} J_i \triangleq h_i \tag{21}$$

It should be noted that pre-multiplying the second line in Eq. 17 by $\mathbf{v}_i^T$ produces an identical expression to that of Eq. 21. It can be seen that $E\{h_i\} = 0$ under normal operating conditions while $E\{h_i\} \neq 0$ under faulty conditions. Therefore, the functions h_i can alternatively be selected for the construction of primary residuals:

$$\mathcal{H}_{R0}^T = \begin{pmatrix} h_1 & h_2 & \cdots & h_n \end{pmatrix} \in \mathbb{R}^n \tag{22}$$

$$\mathcal{H}_{D0}^T = \begin{pmatrix} h_{n+1} & h_{n+2} & \cdots & h_N \end{pmatrix} \in \mathbb{R}^{N-n} \tag{23}$$

A third set of monitoring functions can be extracted from the cost function for p:

$$J_{\mathbf{p}_i} = \arg\min_{\mathbf{p}_i} \operatorname{trace} \left\{ (\mathbf{X}_i - \boldsymbol{t}_i \mathbf{p}_i^T)^T (\mathbf{X}_i - \boldsymbol{t}_i \mathbf{p}_i^T) \right\} \tag{24}$$

This equation can be reformulated as

$$\mathbf{p}_i = \arg\min_{\mathbf{p}_i} \left(E\left\{ \operatorname{trace} \left(\mathbf{x}_i \mathbf{x}_i^T - 2\mathbf{x}_i t_i \mathbf{p}_i^T + \mathbf{p}_i t_i^2 \mathbf{p}_i^T \right) \right\} \right) \tag{25}$$

The solution can be obtained, as earlier, by taking the partial derivative and setting it to zero, i.e.

$$\mathbf{p}_i = \arg \left(\frac{\partial}{\partial \mathbf{p}_i} E\left\{ \operatorname{trace} \left(\mathbf{x}_i \mathbf{x}_i^T - 2\mathbf{x}_i t_i \mathbf{p}_i^T + \mathbf{p}_i t_i^2 \mathbf{p}_i^T \right) \right\} \right) = 0 \tag{26}$$

so that

$$\mathbf{p}_i = \arg \left(E\left\{ 2t_i^2 \mathbf{p}_i - 2t_i \mathbf{x}_i \right\} \right) = 0 \tag{27}$$

Now define

$$\mathbf{l}_i = 2t_i^2 \mathbf{p}_i - 2t_i \mathbf{x}_i \tag{28}$$

such that $E\{\mathbf{l}_i\} = \mathbf{0}$ for normal operating conditions and $E\{\mathbf{l}_i\} \neq \mathbf{0}$ under faulty conditions. The functions $\mathbf{l}_i$ can be selected as a new set of primary residuals:

$$\mathcal{L}_{R0}^T = \begin{pmatrix} \mathbf{l}_1^T & \mathbf{l}_2^T & \cdots & \mathbf{l}_n^T \end{pmatrix} \in \mathbb{R}^{N\times n} \tag{29}$$

$$\mathcal{L}_{D0}^T = \begin{pmatrix} \mathbf{l}_{n+1}^T & \mathbf{l}_{n+2}^T & \cdots & \mathbf{l}_N^T \end{pmatrix} \in \mathbb{R}^{N\times(N-n)} \tag{30}$$

Finally, a fourth set of monitoring functions can be derived for $\mathbf{q}$:

$$J_{\mathbf{q}_i} = \arg\min_{\mathbf{q}_i} \operatorname{trace} \{ (\mathbf{Y}_i - \mathbf{t}_i \mathbf{q}_i)^T (\mathbf{Y}_i - \mathbf{t}_i \mathbf{q}_i) \}. \tag{31}$$

Following an approach identical to that used for $\mathbf{p}$, the $\mathbf{q}$-based residuals can be defined as

$$\mathbf{m}_i = 2t_i^2 \mathbf{q}_i - 2t_i \mathbf{y}_i \tag{32}$$

so that their expected value $E\{\mathbf{m}_i\} = \mathbf{0}$ for normal operating conditions and $E\{\mathbf{m}_i\} \neq 0$ under faulty conditions. As earlier, the functions $\mathbf{m}_i$ can be selected as a fourth set of primary residuals:

$$\mathcal{M}_{R0}^T = \begin{pmatrix} \mathbf{m}_1^T & \mathbf{m}_2^T & \cdots & \mathbf{m}_n^T \end{pmatrix} \in \mathbb{R}^{M\times n} \tag{33}$$

$$\mathcal{M}_{D0}^T = \begin{pmatrix} \mathbf{m}_{n+1}^T & \mathbf{m}_{n+2}^T & \cdots & \mathbf{m}_N^T \end{pmatrix} \in \mathbb{R}^{M\times(N-n)} \tag{34}$$

On the basis of the analysis in Ref. [42] the calculation of the $\mathbf{g}_i$, h_i, $\mathbf{l}_i$, and $\mathbf{m}_i$ functions can be simplified by deflating only the input data, by means of

$$\mathbf{t}_i = \mathbf{X}_i \mathbf{w}_i \triangleq \mathbf{X}\mathbf{r}_i \tag{35}$$

where $\mathbf{r}_i$ can be iteratively computed as follows:

$$\mathbf{r}_i = \mathbf{w}_i - \sum_{j=1}^{i=1} \mathbf{p}_j^T \mathbf{w}_i \mathbf{r}_j. \tag{36}$$

For completeness, a proof of Eq. 36 is given in Appendix B. This appendix also shows that

$$\mathbf{X}_i = \mathbf{X}\mathbf{R}_i \tag{37}$$

where $\mathbf{R}_i = \mathbf{I} - \Sigma_{j=1}^{i-1} \mathbf{r}_j \mathbf{p}_j^T$ for $i = 2, \ldots, N$ and $\mathbf{R}_1 = \mathbf{I}$. Utilizing Eqs. 36 and 37 leads to the construction of compact equations that incorporate the deflation procedure and hence, enables to compute the monitoring functions directly from the recorded data, i.e. the predictor and response variable sets:

$$h_i = \mathbf{r}_i^T \mathbf{x} \mathbf{y}^T \mathbf{v}_i - \frac{K-1}{K} J_i \tag{38}$$

$$\mathbf{g}_i = \begin{pmatrix} \mathbf{R}_i^T \mathbf{x} (\mathbf{y}^T \mathbf{v}_i) - \frac{K-1}{K} J_i \mathbf{w}_i \\ \mathbf{y} \mathbf{x}^T \mathbf{r}_i - \frac{K-1}{K} J_i \mathbf{v}_i \end{pmatrix} \tag{39}$$

$$\mathbf{l}_i = 2\mathbf{x}^T \mathbf{r}_i \big((\mathbf{x}^T \mathbf{r}_i) \mathbf{p}_i - \mathbf{R}_i^T \mathbf{x} \big) \tag{40}$$

$$\mathbf{m}_i = 2\mathbf{x}^T \mathbf{r}_i \big((\mathbf{x}^T \mathbf{r}_i) \mathbf{q}_i - \mathbf{y} \big) \tag{41}$$

Introduction of improved residuals

Improved residuals can now be calculated based on the four primary residual functions, detailed in the previous subsection. Denoting the time instance by j, the improved residuals are given by

$$\begin{aligned}
\mathbb{G}_{R0_K} &= \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{G}_{R0_j}, \quad \mathbb{G}_{D0_K} = \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{G}_{D0_j} \\
\mathbb{H}_{R0_K} &= \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{H}_{R0_j}, \quad \mathbb{H}_{D0_K} = \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{H}_{D0_j} \\
\mathbb{L}_{R0_K} &= \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{L}_{R0_j}, \quad \mathbb{L}_{D0_K} = \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{L}_{D0_j} \\
\mathbb{M}_{R0_K} &= \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{M}_{R0_j}, \quad \mathbb{M}_{D0_K} = \frac{1}{\sqrt{K}} \sum\nolimits_{j=1}^{K} \mathcal{M}_{D0_j}
\end{aligned} \tag{42}$$

For large K, however, the values of $\mathbb{H}_{R0}$ and $\mathbb{H}_{D0}$ etc become stabilized and are less sensitive to incipient changes in the behavior of the plant. This can be overcome by performing the averaging not over the entire data record but over a shorter window of k_0 data-points.[43] Incorporating such a moving window approach renders the above equations to become

$$\begin{aligned}
\mathbb{G}_{R0_k} &= \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{G}_{R0_j}, \quad \mathbb{G}_{D0k} = \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{G}_{D0_j} \\
\mathbb{H}_{R0_k} &= \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{H}_{R0_j}, \quad \mathbb{H}_{D0_k} = \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{H}_{D0_j} \\
\mathbb{L}_{R0_k} &= \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{L}_{R0_j}, \quad \mathbb{L}_{D0k} = \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{L}_{D0_j} \\
\mathbb{M}_{R0_k} &= \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{M}_{R0_j}, \quad \mathbb{M}_{D0_k} = \frac{1}{\sqrt{k_0}} \sum\nolimits_{j=k-k_0+1}^{k} \mathcal{M}_{D0_j}
\end{aligned} \tag{43}$$

where k denotes the current sampling instance. Next, it is shown how to use the improved residuals in order to construct univariate monitoring statistics.

Univariate monitoring statistics

As shown in Eq. 12, a Hotelling's T^2 monitoring statistic can be defined for each of the improved residual vectors:

$$\begin{aligned}
T^2_{\mathbb{G}_{R0}} &= \mathbb{G}_{R0}^T \mathbf{S}^{-1}_{\mathcal{G}_{R0}\mathcal{G}_{R0}} \mathbb{G}_{R0}, \quad T^2_{\mathbb{G}_{D0}} = \mathbb{G}_{D0}^T \mathbf{S}^{-1}_{\mathcal{G}_{D0}\mathcal{G}_{D0}} \mathbb{G}_{D0} \\
T^2_{\mathbb{H}_{R0}} &= \mathbb{H}_{R0}^T \mathbf{S}^{-1}_{\mathcal{H}_{R0}\mathcal{H}_{R0}} \mathbb{H}_{R0}, \quad T^2_{\mathbb{H}_{D0}} = \mathbb{H}_{D0}^T \mathbf{S}^{-1}_{\mathcal{H}_{D0}\mathcal{H}_{D0}} \mathbb{H}_{D0} \\
T^2_{\mathbb{L}_{R0}} &= \mathbb{L}_{R0}^T \mathbf{S}^{-1}_{\mathcal{L}_{R0}\mathcal{L}_{R0}} \mathbb{L}_{R0}, \quad T^2_{\mathbb{L}_{D0}} = \mathbb{L}_{D0}^T \mathbf{S}^{-1}_{\mathcal{L}_{D0}\mathcal{L}_{D0}} \mathbb{L}_{D0} \\
T^2_{\mathbb{M}_{R0}} &= \mathbb{M}_{R0}^T \mathbf{S}^{-1}_{\mathcal{M}_{R0}\mathcal{M}_{R0}} \mathbb{M}_{R0}, \quad T^2_{\mathbb{M}_{D0}} = \mathbb{M}_{D0}^T \mathbf{S}^{-1}_{\mathcal{M}_{D0}\mathcal{M}_{D0}} \mathbb{M}_{D0}
\end{aligned} \tag{44}$$

where $\mathbf{S}_{zz}$ denotes the covariance matrix of the variable set $\mathbf{z}$. Applying the central limit theorem, the statistics $\mathbb{G}_{R0_k}$, $\mathbb{G}_{D0_k}$, $\mathbb{H}_{R0_k}$, $\mathbb{H}_{D0_k}$, etc., each follow a multinormal distribution of dimension n $(N + M)$, $(N - n)$ $(N + M)$, n, $N - n$, etc., respectively. If the covariance matrices are estimated from process reference data, each of the above statistics follow an F-distribution,[44] e.g.

$$\begin{aligned}
T^2_{\mathbb{G}_{R0}} &\sim \frac{\tilde{n}_1 (K^2 - 1)}{K(K - \tilde{n}_1)} \mathcal{F}(\tilde{n}_1, K - \tilde{n}_1) \\
T^2_{\mathbb{G}_{D0}} &\sim \frac{\tilde{n}_2 (K^2 - 1)}{K(K - \tilde{n}_2)} \mathcal{F}(\tilde{n}_2, K - \tilde{n}_2) \\
T^2_{\mathbb{H}_{R0}} &\sim \frac{\tilde{n}_1 (K^2 - 1)}{K(K - \tilde{n})} \mathcal{F}(\tilde{n}, K - \tilde{n}) \\
T^2_{\mathbb{H}_{D0}} &\sim \frac{(N - \tilde{n})(K^2 - 1)}{K(K - N + \tilde{n})} \mathcal{F}(\tilde{n}_1, K - \tilde{n}_1),
\end{aligned} \tag{45}$$

etc., where $\tilde{n}_1 = n(N + M)$, $\tilde{n}_2 = (N - n)(N + M)$, and $\mathcal{F}(m_1, m_2)$ is an F-distribution with m_1 and m_2 degrees of freedom.

Sensitivity analysis

In this subsection, the sensitivity of each primary (and consequently the improved) residuals to deterministic additive faults is examined. Under these two assumptions, the presence of a step-type fault manifests itself in the predictor and response variables as follows:

$$\mathbf{x} = \mathbf{x}_0 + \Delta\mathbf{x} \tag{46}$$

$$\mathbf{y} = \mathbf{y}_0 + \Delta\mathbf{y} \tag{47}$$

where $\mathbf{x}_0 \in \mathbb{R}^N$ and $\mathbf{y}_0 \in \mathbb{R}^M$ are the data that would have been measured if there was no fault and $\Delta\mathbf{x} \in \mathbb{R}^N$ and $\Delta\mathbf{y} \in \mathbb{R}^M$ are vectors containing the fault magnitude.

The sensitivity of the residuals derived earlier in this section to such faults can be determined by considering their expected values. Substituting Eqs. 46 and 47 into Eqs. 38–41 and taking the expected values of the four sets of primary residuals yields

$$E\{h_i\} = (\mathbf{r}_i \otimes \mathbf{v}_i)^T (\Delta\mathbf{x} \otimes \Delta\mathbf{y}) \tag{48}$$

$$E\{\mathbf{g}_i\} = \begin{pmatrix} (\mathbf{R}_i \otimes \mathbf{v}_i)^T \\ \mathbf{r}_i^T \otimes \mathbf{I}_M \end{pmatrix} (\Delta\mathbf{x} \otimes \Delta\mathbf{y}) \tag{49}$$

$$E\{\mathbf{l}_i\} = 2\left(\mathbf{p}_i(\mathbf{r}_i \otimes \mathbf{r}_i)^T - (\mathbf{R}_i \otimes \mathbf{r}_i)^T\right)(\Delta\mathbf{x} \otimes \Delta\mathbf{x}) \tag{50}$$

$$E\{\mathbf{m}_i\} = 2\mathbf{q}_i(\mathbf{r}_i \otimes \mathbf{r}_i)^T - (\Delta\mathbf{x} \otimes \Delta\mathbf{x}) - 2(\mathbf{r}_i \otimes \mathbf{I}_M)^T(\Delta\mathbf{x} \otimes \Delta\mathbf{y}) \tag{51}$$

where $\otimes$ is the Kronecker product and $\mathbf{I}_M$ is a unity matrix of size M. These results were obtained using the facts that $E\{\mathbf{x}_0\} = \mathbf{0}$ and $E\{\mathbf{y}_0\} = 0$.

Equations 48–51 demonstrate the following possibilities:

- $\Delta\mathbf{x} = \mathbf{0}$, $\Delta\mathbf{y} = \mathbf{0}$: there is no fault and the expected values of all the monitoring functions are zero.
- $\Delta\mathbf{x} \neq \mathbf{0}$, $\Delta\mathbf{y} \neq \mathbf{0}$: the expected values of all the monitoring functions are nonzero, and hence, the fault is detected
- $\Delta\mathbf{x} \neq \mathbf{0}$, $\Delta\mathbf{y} = \mathbf{0}$: the expected values of the h_i and g_i monitoring functions are zero and those of the $\mathbf{l}_i$ and $\mathbf{m}_i$ functions are nonzero. Therefore, the fault is detected only by the $\mathbf{l}_i$ and $\mathbf{m}_i$ monitoring functions.
- $\Delta\mathbf{x} = \mathbf{0}$, $\Delta\mathbf{y} \neq \mathbf{0}$: the expected values of all the monitoring functions are zero and the fault goes undetected.

Therefore, the case $\Delta\mathbf{x} = \mathbf{0}$, $\Delta\mathbf{y} \neq \mathbf{0}$ is problematic because the monitoring strategy detailed here fails. However, as outlined earlier, $\mathbf{X}$ and $\mathbf{Y}$ are, in fact, $\mathcal{X}$ and $\mathcal{Y}$. Two PLS models can be created, a *direct* model based on $\mathcal{Y} = \mathcal{X}\mathbf{B}$ and an *inverted* model based on $\mathcal{X} = \mathcal{Y}\mathbf{C}$. More precisely, the first model treats the plant inputs as model inputs and the plant outputs as model outputs while the second model treats the plant inputs as model outputs and the plant outputs as model inputs.

Using this dual model approach two sets of monitoring functions can be selected. For example, one denoted by $\mathbf{m}_i$, as given in Eq. 41, which detects faults in the form $\Delta\mathcal{X} \otimes \Delta\mathcal{X}$, $\Delta\mathcal{X} \otimes \Delta\mathcal{Y}$, and one denoted by $\mathbf{m}_i^*$, also described by Eq. 41 but obtained from the inverted model and treating the plant inputs as model outputs and vice versa, which detects faults in the form $\Delta\mathcal{Y} \otimes \Delta\mathcal{Y}$ and $\Delta\mathcal{Y} \otimes \Delta\mathcal{X}$. Primary residuals $\mathcal{M}_{R0}^*$ and $\mathcal{M}_{D0}^*$ can be defined as in Eqs. 33 and 34.

$$\mathcal{M}_{R0}^{*T} = \begin{pmatrix} \mathbf{m}_1^{*T} & \mathbf{m}_2^{*T} & \cdots & \mathbf{m}_n^{*T} \end{pmatrix} \in \mathbb{R}^n \tag{52}$$

$$\mathcal{M}_{D0}^{*T} = \begin{pmatrix} \mathbf{m}_{n+1}^{*T} & \mathbf{m}_{n+2}^{*T} & \cdots & \mathbf{m}_N^{*T} \end{pmatrix} \in \mathbb{R}^{N-n} \tag{53}$$

Improved residuals can be defined according to Eqs. 43 as

$$\mathbb{M}_{R0_k}^* = \frac{1}{\sqrt{k_0}} \sum_{j=k-k_0+1}^{k} \mathcal{M}_{R0_j}^*, \quad \mathbb{M}_{D0_k}^* = \frac{1}{\sqrt{k_0}} \sum_{j=k-k_0+1}^{k} \mathcal{M}_{D0_j}^* \tag{54}$$

Finally, univariate monitoring statistics can be defined as

$$T_{\mathbb{M}_{R0}^*}^2 = \mathbb{M}_{R0}^{*T}\mathbf{S}_{\mathcal{M}_{R0}^*\mathcal{M}_{R0}^*}^{-1}\mathbb{M}_{R0}^*, \quad T_{\mathbb{M}_{D0}^*}^2 = \mathbb{M}_{D0}^{*T}\mathbf{S}_{\mathcal{M}_{D0}^*\mathcal{M}_{D0}^*}^{-1}\mathbb{M}_{R0}^* \tag{55}$$

in the manner of Eqs. 44. The univariate statistics $T_{\mathbb{M}_{R0}}^2$, $T_{\mathbb{M}_{D0}}^2$, $T_{\mathbb{M}_{R0}^*}^2$, $T_{\mathbb{M}_{D0}^*}^2$ are capable of detecting any additive faults in the inputs and/or outputs of the plant.

Fault Diagnosis

Up to this point, it has been shown that the application of the PLS method within the remit of the statistical local approach produces monitoring functions can detect the occurrence of additive faults in industrial processes. This relates to the identification of a statistically significant change in the mean value of the primary and improved residuals. However, the basic approach detailed above cannot identify where exactly the fault has occurred nor can it estimate its magnitude.

This section shows that the primary residuals can be utilized to extract the fault magnitude and signature. More precisely, under the assumption that the fault signature is deterministic in nature, this signature can be exactly reconstructed for step-type faults. For general deterministic fault conditions, this section shows that an estimate of the fault magnitude and signature can be extracted. The estimated magnitudes and signatures can be attributed to each variable and thus provide an experienced plant operator with valuable information to narrow down potential root causes for anomalous behavior.

The objective for fault diagnosis here is to estimate the additive faults $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ as defined in Eqs. 46 and 47 and thus obtain information on both fault magnitude and location. This can be accomplished by investigating the monitoring functions, $\mathbf{g}_i$, $\mathbf{l}_i$, and $\mathbf{m}_i$ which have zero expected values only when they are calculated using $\mathbf{x}_0$ and $\mathbf{y}_0$. Considering the functions $\mathbf{g}_i$ it is clear that

$$E\{\mathbf{g}_i(\mathbf{x}_0, \mathbf{x}_0)\} = E\left\{ \begin{matrix} \mathbf{R}_i^T\mathbf{x}_0\mathbf{y}_0^T\mathbf{v}_i - \frac{K-1}{K}J_i\mathbf{w}_i \\ \mathbf{y}_0\mathbf{x}_0^T\mathbf{r}_i - \frac{K-1}{K}J_i\mathbf{v}_i \end{matrix} \right\} = \mathbf{0} \tag{56}$$

Substituting from Eqs. 46 and 47 yields

$$\begin{aligned} E\{\mathbf{g}_i\} &= E\left\{ \begin{matrix} \mathbf{R}_i^T(\mathbf{x}-\Delta\mathbf{x})(\mathbf{y}-\Delta\mathbf{y})^T\mathbf{v}_i - \frac{K-1}{K}J_i\mathbf{w}_i \\ (\mathbf{y}-\Delta\mathbf{y})(\mathbf{x}-\Delta\mathbf{x})^T\mathbf{r}_i - \frac{K-1}{K}J_i\mathbf{v}_i \end{matrix} \right\} \\ &= E\left\{ \begin{matrix} \mathbf{R}_i^T\mathbf{x}\mathbf{y}^T\mathbf{v}_i - \frac{K-1}{K}J_i\mathbf{w}_i \\ \mathbf{y}\mathbf{x}^T\mathbf{r}_i - \frac{K-1}{K}J_i\mathbf{v}_i \end{matrix} \right\} - \begin{pmatrix} (\mathbf{R}_i \otimes \mathbf{v}_i)^T \\ \mathbf{r}_i^T \otimes \mathbf{I}_M \end{pmatrix}(\Delta\mathbf{x} \otimes \Delta\mathbf{y}) \end{aligned} \tag{57}$$

Note that this result was obtained using the fact that $E\{\mathbf{x}\} = \Delta\mathbf{x}$ and $E\{\mathbf{y}\} = \Delta\mathbf{y}$, which follows from Eqs. 46 and 47. This translates to

$$\sum_{j=k-k_0+1}^{k} \left\{ \begin{matrix} \mathbf{R}_i^T\mathbf{x}_j\mathbf{y}_j^T\mathbf{v}_i - \frac{K-1}{K}J_i\mathbf{w}_i \\ \mathbf{y}_j\mathbf{x}_j^T\mathbf{r}_i - \frac{K-1}{K}J_i\mathbf{v}_i \end{matrix} \right\} = \begin{pmatrix} (\mathbf{R}_i \otimes \mathbf{v}_i)^T \\ \mathbf{r}_i^T \otimes \mathbf{I}_M \end{pmatrix}(\Delta\mathbf{x} \otimes \Delta\mathbf{y}) \tag{58}$$

for $i = 1, \ldots, N$. The number of equations is $N(N + M)$ and the number of unknowns is $N \times M$. Obviously, Eq. 58 can be solved, as there are more equations than unknowns. The solution can be obtained in a least squares sense. Notice that a new value of $\Delta\mathbf{x}_k \otimes \Delta\mathbf{y}_k$ can be obtained at each data window, $k - k_0 + 1$ to k, for which the $\mathbf{g}_i$ monitoring functions are calculated. Therefore, both step-type and general deterministic faults can be diagnosed.

The evaluation of $\Delta\mathbf{x} \otimes \Delta\mathbf{y}$ is only part of the fault identification procedure. The individual faults $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ cannot be evaluated from $\Delta\mathbf{x} \otimes \Delta\mathbf{y}$. In order to achieve this, a separate set of equations must be developed, based on the $\mathbf{l}_i$ primary residuals. Equation 40 is evaluated at $\mathbf{x}_0$. Then, substituting from Eqs. 46 and 47 and taking the expected value yields

$$E\{\mathbf{l}_i\} = E\{2\mathbf{x}^T\mathbf{r}_i\mathbf{x}^T\mathbf{r}_i\mathbf{p}_i - 2\mathbf{x}^T\mathbf{r}_i\mathbf{x}\} - 2\Delta\mathbf{x}^T\mathbf{r}_i\Delta\mathbf{x}^T\mathbf{r}_i\mathbf{p}_i + 2\Delta\mathbf{x}^T\mathbf{r}_i\Delta\mathbf{x} = 0 \quad (59)$$

Finally, the $\Delta\mathbf{x}$ faults can be obtained from

$$\sum_{j=k-k_0+1}^{k} (2\mathbf{x}_j^T\mathbf{r}_i\mathbf{x}_j^T\mathbf{r}_i\mathbf{p}_i - 2\mathbf{x}_j^T\mathbf{r}_i\mathbf{x}_j) = 2\Big(\mathbf{p}_i(\mathbf{r}_i\otimes\mathbf{r}_i)^T - (\mathbf{R}_i\otimes\mathbf{r}_i)^T\Big)(\Delta\mathbf{x}\otimes\Delta\mathbf{x}) \quad (60)$$

for $i = 1, \ldots, N$. There are $\Sigma_{j=1}^{N} j$ unknowns (because some of the terms in $\Delta\mathbf{x}\otimes\Delta\mathbf{x}$ are identical) with N^2 equations. Additionally, it can be shown that the rank of the data matrix in Eq. 60 is equal to N. Therefore, Eq. 60 is not sufficient to solve for the unknown $\Delta\mathbf{x}\otimes\Delta\mathbf{x}$. A third set of equations is required, based on the $\mathbf{m}_i$ residuals. For the latter, it is obvious that

$$E\{\mathbf{m}_i\} = 2\mathbf{x}_0^T\mathbf{r}_i\mathbf{x}_0^T\mathbf{r}_i\mathbf{q}_i - 2\mathbf{x}_0^T\mathbf{r}_i\mathbf{y}_0 = 0 \quad (61)$$

Substituting from Eqs. 46 and 47 and carrying out all the necessary algebra yields

$$E\{\mathbf{m}_i\} = E\{2\mathbf{x}^T\mathbf{r}_i\mathbf{x}^T\mathbf{r}_i\mathbf{q}_i - 2\mathbf{x}^T\mathbf{r}_i\mathbf{y}\} - 2\Delta\mathbf{x}^T\mathbf{r}_i\Delta\mathbf{x}^T\mathbf{r}_i\mathbf{q}_i + 2\Delta\mathbf{x}^T\mathbf{r}_i\Delta\mathbf{y} = 0 \quad (62)$$

Defining primary and improved residuals in the usual manner leads to an equation for the $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ faults of the form

$$\sum_{j=k-k_0+1}^{k} (2\mathbf{x}_j^T\mathbf{r}_i\mathbf{x}_j^T\mathbf{r}_i\mathbf{p}_i - 2\mathbf{x}_j^T\mathbf{r}_i\mathbf{y}_j) + 2(\mathbf{r}_i\otimes\mathbf{I}_M)^T\Delta\mathbf{x}\otimes\Delta\mathbf{y} = 2\mathbf{q}_i(\mathbf{r}_i\otimes\mathbf{r}_i)^T\Delta\mathbf{x}\otimes\Delta\mathbf{x} \quad (63)$$

where $\Delta\mathbf{x}\otimes\Delta\mathbf{y}$ has already been evaluated from Eq. 58. Finally, Eqs. 60 and 63 are combined to yield a complete set of equations:

$$\begin{pmatrix} 2(\mathbf{p}_i(\mathbf{r}_i\otimes\mathbf{r}_i)^T - (\mathbf{R}_i\otimes\mathbf{r}_i)^T) \\ 2b_i\mathbf{q}_i(\mathbf{r}_i\otimes\mathbf{r}_i)^T \end{pmatrix}\Delta\mathbf{x}\otimes\Delta\mathbf{x} = \begin{pmatrix} \sum_{j=k-k_0+1}^{k}(2\mathbf{x}_j^T\mathbf{r}_i\mathbf{x}_j^T\mathbf{r}_i\mathbf{p}_i - 2\mathbf{x}_j^T\mathbf{r}_i\mathbf{x}_j) \\ \sum_{j=k-k_0+1}^{k}(2b_i\mathbf{x}_j^T\mathbf{r}_i\mathbf{x}_j^T\mathbf{r}_i\mathbf{q}_i - 2b_i\mathbf{x}_j^T\mathbf{r}_i\mathbf{y}_j) + 2b_i(\mathbf{r}_i\otimes\mathbf{I}_M)^T\Delta\mathbf{x}\otimes\Delta\mathbf{y} \end{pmatrix} \quad (64)$$

which is of the form

$$\mathbf{\Phi\Delta} = \mathbf{\Psi} \quad (65)$$

and of rank $2N$. Notice that $\mathbf{\Delta}$ contains only the elements of $\Delta\mathbf{x}\otimes\Delta\mathbf{x}$ that are not repeated and $\mathbf{\Phi}$ has been compressed accordingly. Solving Eq. 65 yields $\Delta\mathbf{x}\otimes\Delta\mathbf{x}$ from which $\Delta\mathbf{x}$, and hence, $\Delta\mathbf{y}$ can be extracted.

Case 1: Additive faults in the inputs, $\Delta\mathbf{x} \neq \mathbf{0}$

If there are additive faults in the inputs, then the fault magnitude identification procedure begins with the solution of Eq. 58 for $\Delta\mathbf{x}\otimes\Delta\mathbf{y}$. Then, Eq. 65 is solved to yield $\Delta\mathbf{x}\otimes\Delta\mathbf{x}$. Finally, the individual values of $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ are obtained from $\Delta\mathbf{x}\otimes\Delta\mathbf{y}$ and $\Delta\mathbf{x}\otimes\Delta\mathbf{x}$.

Case 2: No additive faults in the inputs, $\Delta\mathbf{x} = \mathbf{0}$

If $\Delta\mathbf{x} = \mathbf{0}$ and all the additive faults are in the outputs then Eq. 58 will fail as $\Delta\mathbf{x}\otimes\Delta\mathbf{y} = 0$. Nevertheless, the fault magnitude can still be identified in such cases from the inverted PLS model. In this case, the monitoring functions $\mathbf{g}_i^*$, $\mathbf{l}_i^*$, and $\mathbf{m}_i^*$ are used to derive equations identical in form to Eqs. 58 and 65 but in which $\mathbf{X}$ and $\mathbf{Y}$ denote $\mathcal{Y}$ and $\mathcal{X}$, respectively, and where $\mathbf{v}_i$, $\mathbf{w}_i$, $\mathbf{r}_i$, $\mathbf{p}_i$, $\mathbf{q}_i$, are related to the inverted model.

FDD for a Simulated Example

In this section the proposed FDD approach is applied to a simulated plant. The plant has three inputs and two outputs and a $\mathbf{B}$ matrix equal to

$$\mathbf{B} = \begin{bmatrix} 0.3412 & 0.5341 & 0.7271 \\ 0.3093 & 0.8385 & 0.5681 \end{bmatrix} \quad (66)$$

The inputs are white noise with zero mean and unit variance. The linear relationships between the input and output variables are calculated from $\mathbf{y}^* = \mathbf{B}\mathbf{x}^*$ and the recorded variables are $\mathbf{y} = \mathbf{y}^* + \mathbf{e}$ and $\mathbf{x} = \mathbf{x}^* + \mathbf{f}$, where $\mathbf{g}^T = (\mathbf{e}^T\ \mathbf{f}^T)$ are normally distributed noise sequences, which are independently and identically distributed, and of zero mean and a variance of 0.05, that is $\mathbf{g} \sim \mathcal{N}\{\mathbf{0}, 0.05\mathbf{I}\}$. Furthermore, the sequences of $\mathbf{x} \sim \mathcal{N}\left\{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{bmatrix}1 & 0.2 & 0.1\\ 0.2 & 1 & 0.3\\ 0.1 & 0.3 & 1\end{bmatrix}\right\}$. Two sets of data are used, one training set of 1000 records and one operational set of 4000 records. The last 1000 records of the operational data contain additive faults.

The analysis of the training set, $\mathbf{Z} = [\mathbf{Y}\ \mathbf{X}]$, yielded the following singular values: 64.8282, 30.9326, 27.2072, and as expected, two times 7.0628. A subsequent PLS analysis of $\mathcal{Y}$ and $\mathcal{X}$ produced J values of 1.9805, 0.1454, and 0.0021, and for the inverse model values of 1.9805, 0.1225, suggesting that for both models the first two pairs of latent variables have a significant contribution. Next, these models were applied to detect and diagnose a total of two different fault conditions.

Fault detection

Case 1: $\Delta\mathbf{x} \neq \mathbf{0}$, $\Delta\mathbf{y} \neq \mathbf{0}$. The following additive faults were seeded to the last 1000 records of the operational data:

$$\Delta\mathbf{x} = [2.0 \quad -1.1 \quad -2.0]^T, \quad \Delta\mathbf{y} = [2.5 \quad -1.8]^T \quad (67)$$

The upper four plots in Figure 1 show the $T^2_{\mathbb{M}_{R0}}$, $T^2_{\mathbb{M}_{D0}}$ and $T^2_{\mathbb{M}^*_{R0}}$, $T^2_{\mathbb{M}^*_{D0}}$ criteria, respectively, evaluated from the operational data. The criteria were calculated using a window length $k_0 = 100$ and are plotted along with the 99% confidence limits. The $T^2_{\mathbb{M}^*_{R0}}$ plot is blank because all latent variables were retained in the inverse application. It is clear that the injected fault condition has been correctly detected.

Case 2: $\Delta\mathbf{x} = \mathbf{0}$, $\Delta\mathbf{y} \neq \mathbf{0}$. In this case the seeded additive fault is of the form

$$\Delta\mathbf{x} = [0 \quad 0 \quad 0]^T, \quad \Delta\mathbf{y} = [2.5 \quad -1.8]^T \quad (68)$$

occurring again after the 3000th record. As discussed in the "Sensitivity Analysis" section, this case is sensitive because

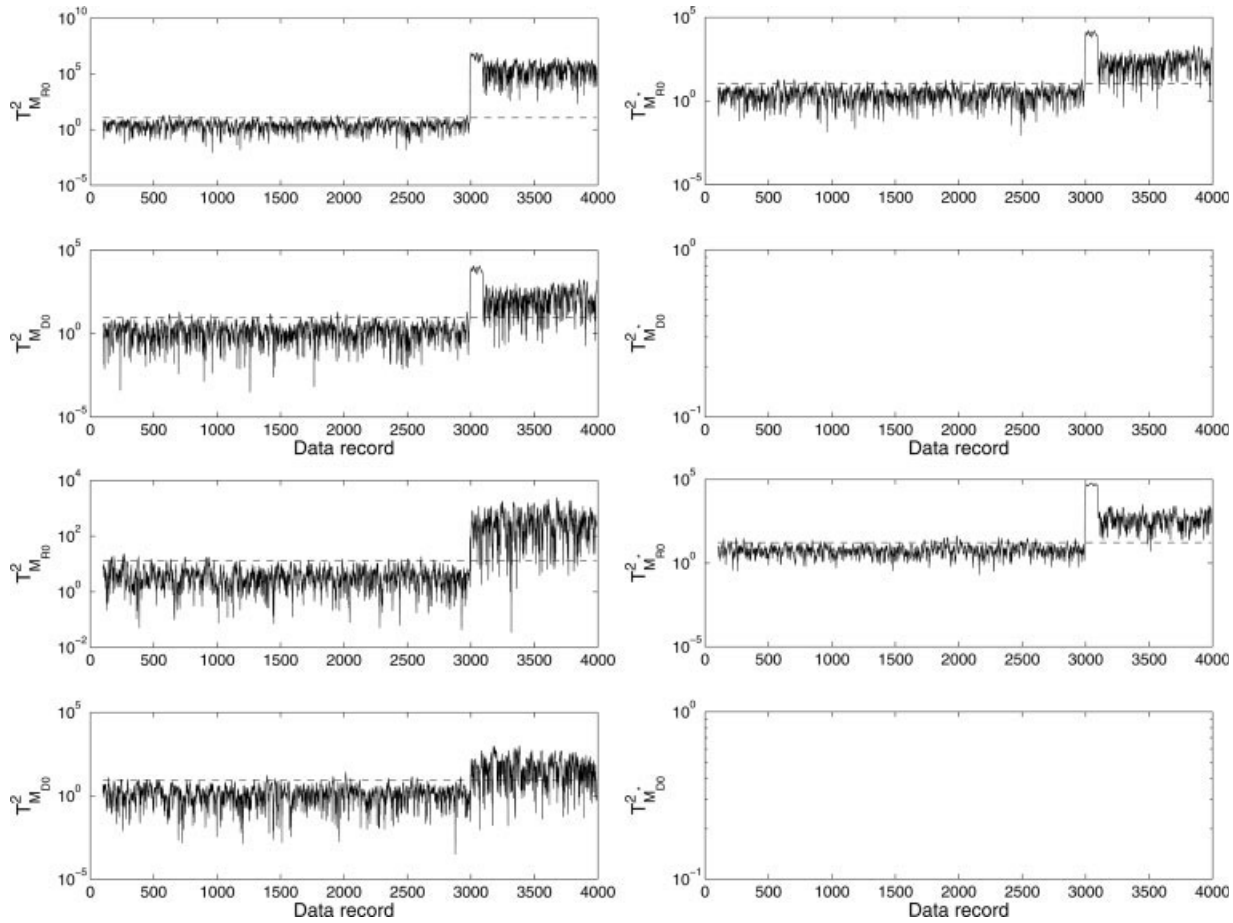

Figure 1. Upper four plots: Four T^2 statistics for $\Delta \mathbf{y} \neq \mathbf{0}$ and $\Delta \mathbf{x} \neq \mathbf{0}$; Lower four plots: four T^2 statistics for the $\Delta \mathbf{x} = \mathbf{0}$ and $\Delta \mathbf{y} \neq \mathbf{0}$.

the direct application of the monitoring procedure yields monitoring functions whose expected values are not affected by the fault. The lower four plots in Figure 1 show the $T^2_{\mathbb{M}_{R0}}$, $T^2_{\mathbb{M}_{D0}}$ and $T^2_{\mathbb{M}^*_{R0}}$, $T^2_{\mathbb{M}^*_{D0}}$ statistics are calculated from the direct and inverse applications, respectively. It can be seen that the fault is correctly monitored by all statistics except $T^2_{\mathbb{M}^*_{D0}}$ which, as before, is not available because there are no discarded latent variables in the inverse application. The detection of the fault by the $T^2_{\mathbb{M}_{R0}}$ and $T^2_{\mathbb{M}_{D0}}$ statistics is interesting because the sensitivity analysis suggests that the fault should be undetectable. The reason for the detection is that, even though the fault does not affect the expected values of the m_i functions, it affects their variance. The local approach picks up this change, and as a result, the fault is detected. Nevertheless, such detection is less sensitive than detection based on changes in the expected value, and therefore it is useful to monitor also the $T^2_{\mathbb{M}^*_{R0}}$ and $T^2_{\mathbb{M}^*_{D0}}$ statistics.

Fault magnitude identification

Case $\Delta \mathbf{x} \neq \mathbf{0}$, $\Delta \mathbf{y} \neq \mathbf{0}$. The fault magnitude identification methodology is applied to the same simulated example. Operational data have been created using the system matrix of Eq. 66. The data are again 4000 records long and the last 3000 records have been corrupted with the additive fault of Eq. 67.

Equations 58 and 64 are applied in a sliding window of $k_0 = 100$ data records and a new set of faults $\Delta \mathbf{x}_k$ and $\Delta \mathbf{y}_k$ are obtained at each data record. The resulting fault magnitudes are plotted in Figure 2, where the dotted lines denote the true faults. It can be seen that the identified faults are very close to their true magnitudes.

Alternatively, Eqs. 58 and 64 can be applied in a single window over the entire range of faulty data. The identified fault magnitudes from this calculation are shown in Eq. 69. It can be seen that they are very accurate when compared with the true fault magnitudes of Eq. 67.

$$\Delta \mathbf{x} = [\,1.99 \quad -1.08 \quad -2.0\,], \quad \Delta \mathbf{y} = [\,2.49 \quad -1.79\,] \tag{69}$$

Case $\Delta x = 0$, $\Delta y \neq 0$. A second example was created by contaminating the data from the system matrix of Eq. 66 with the additive fault of Eq. 68.

In this case $\Delta \mathbf{x} = \mathbf{0}$ and therefore the direct application of Eqs. 58–64 fails, as discussed earlier. Nevertheless, the inverted application of the methodology can yield a highly accurate estimate of the seeded additive fault in the outputs. Figure 2 shows the window-by-window estimated fault obtained from the inverted application. It can be seen that

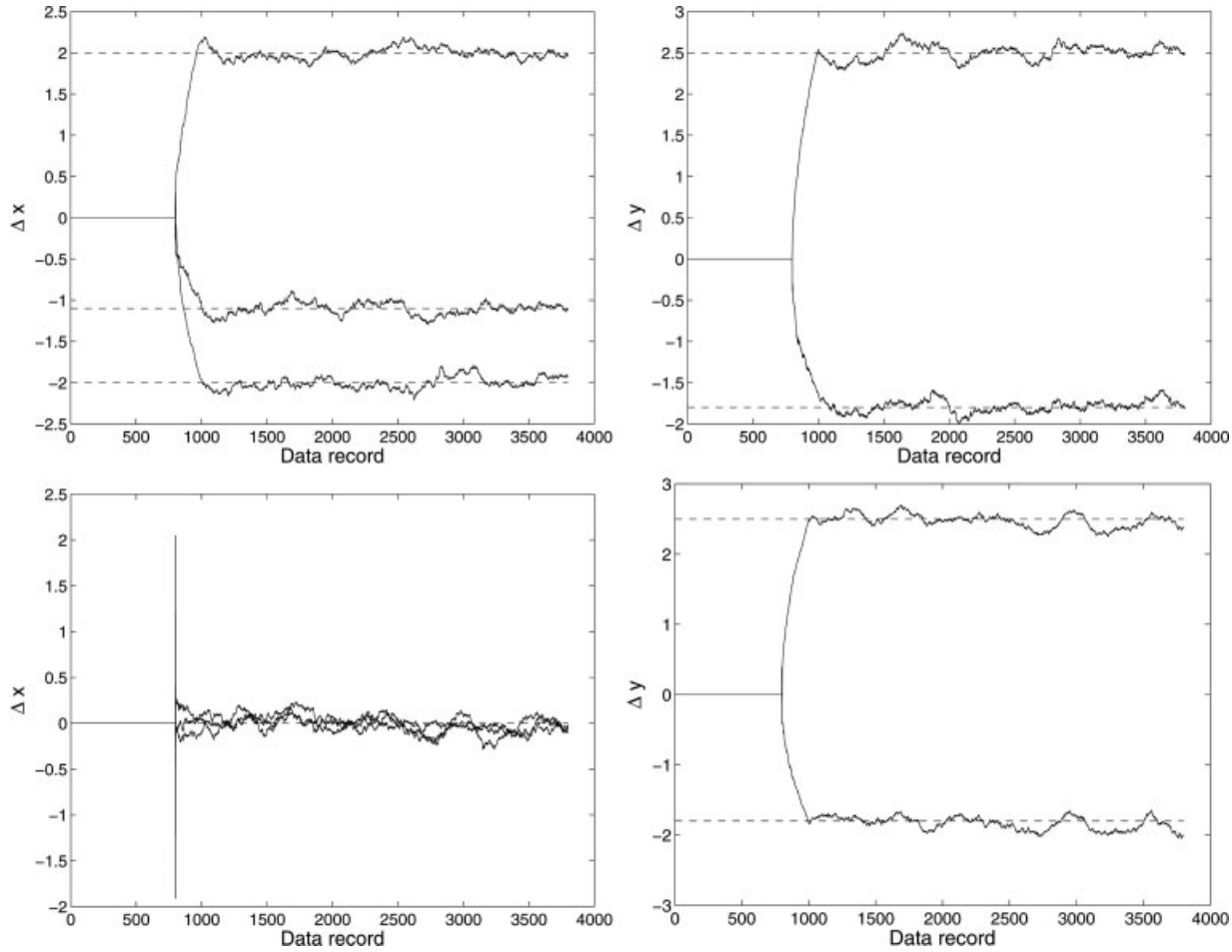

Figure 2. Upper two plots: Identified magnitude of faults in inputs and outputs; Lower two plots: Identified magnitude of faults in inputs and outputs, no input fault.

the identified faults for the inputs are random with zero mean and small standard deviation while the estimated faults for the outputs are centered around the true fault values. Single window application of the inverted approach yields the following estimated faults:

$$\Delta\mathbf{x} = [\,0.02 \quad 0.00 \quad -0.01\,], \quad \Delta\mathbf{y} = [\,2.49 \quad -1.79\,] \tag{70}$$

FDD for Industrial Data

In this section the proposed FDD approach is applied to data from actual industrial processes. As the data recorded from these processes, a distillation unit for purifying a mixture of hydrocarbons and fluidized bed reactor which carries out a complex chemical reaction, have arbitrary mean values and arbitrary scaling, it is necessary to center them to zero and scale them to unity variance before applying the PLS modeling procedure.

Distillation unit

This process has 4 inputs and 8 output variables and is designed to purify a specific compound from a mixture of hydrocarbons. The input variables are refluxflow (1), reboiler steam flow (2), fresh feed flow (3), and fresh feed temperature (4). The distillation unit has a total of 33 trays and the output variables are the tray 14 temperature (1), tray 2 temperature (2), reflux vessel level (3), product flow (4), tray 31 temperature (5), bottom draw (6), reboiler outlet temperature (7), and overhead column pressure (8). From this process, a total of four sets of data were recorded; the first data-set is from a period of healthy operation and was used here as a reference set for identifying a monitoring model. The other three data-sets were operational data and featured each a significant drop in the fresh feed flow.

An application of a singular value decomposition to the augmented data matrix, storing the process reference data (Eqs. 7 and 8), by including four components and a subsequent application of PLS yielded the following values for the cost function J_i, Eq. 1: 1.6668, 1.5713, 0.8801, and 0.0250. Of these, the first three sets of latent variables were retained and the last set discarded. The local approach is then applied to the operational data using a window k_0 of 100 samples. The inverse application yielded the following 8 values for J_i^*: 1.6668, 1.4604, 0.8095, 0.4953, 0.0000, 0.0000, 0.0000 and 0.0000. As earlier, the first three sets of latent variables were

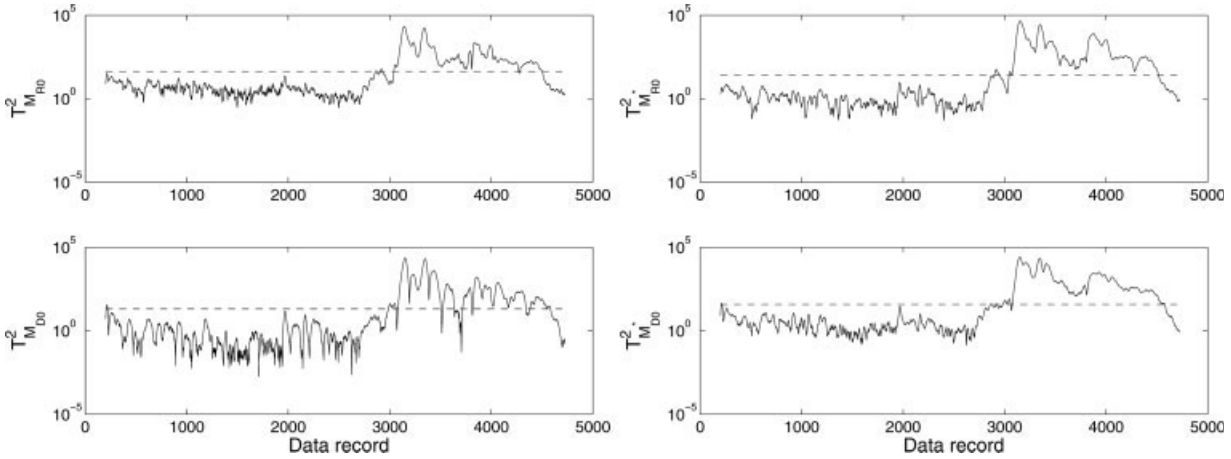

Figure 3. Four T^2 criteria for process fault 1.

retained, whilst the fourth set was discarded. The resulting T^2 criteria are plotted in Figures 3–5, for the three fault cases respectively. It can be seen that for all three sets of operational data the PLS-based monitoring technique has detected an abnormal operating condition caused by the drop in fresh feed.

Figures 6–8 show the fault magnitudes identified for these three faults. In all cases $\Delta\bar{\mathbf{x}}$ and $\Delta\bar{\mathbf{y}}$ are plotted, i.e. values of $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ re-scaled and re-centered by the standard deviations and means of the raw measured data. Referring to the three instances of severe and prolonged drops in fresh feed to faults 1–3, respectively, the next three subsections offer a detailed analysis of the information that can be obtained from Figures 6 to 8.

As the unit does not operate under closed-loop control, any alteration in the fresh feed flow has an immediate affect on the material balance. With some delay, the changes in the material balance translate in changes within the enthalpy balance of the unit. The impact of each drop in fresh feed is further studied in the following three subsections.

Analysis of fault 1. Inspecting first the identified fault signatures for the input variables suggests that variables 1, 2, and 4 are constant and hence, did not contribute to the identified fault. Variable 3, however, showed some sporadic and short-lived drops around 1000 samples into the data record but a substantial drop from around 25 to 20 over a prolonged period of time (from around 2800th sample to the end of the data record. Variable 3 is the fresh feed flow and indeed the root cause of this event.

Analyzing the identified fault signatures for the output variables, no or negligible affects could be noticed for variable 3, 5, 7, and 8, while variables 1, 2, 4, and 6 showed a significant impact. Variables 4 and 6 are the product flow (top draw) and the bottom draw and showed a decrease in flow that coincided with the drop in fresh feed. Moreover, variables 1 and 2 are the tray 14 and the tray 2 temperature. Particularly the tray 14 temperature responded significantly, which relates to the fact this temperature is measured in the vicinity of the fresh feed entering the column.

The isolation of temperatures and flow rates in the output variable set and the fresh feed flow in the input variable set therefore suggested to trace the root cause of this fault condition back to an undesired alteration in the input feed flow, and therefore demonstrate the usefulness of the proposed technique for extracting the fault signature from the recorded date.

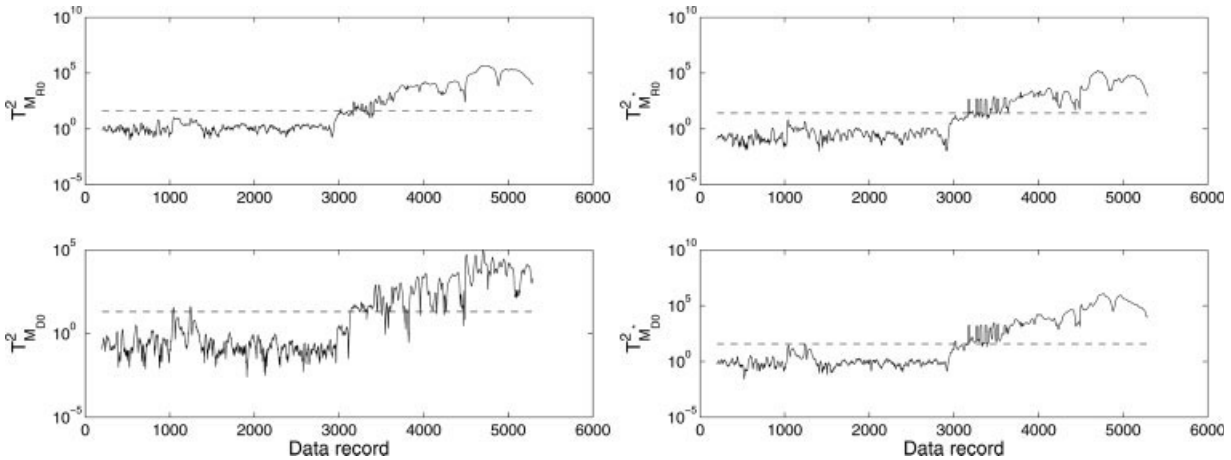

Figure 4. Four T^2 criteria for process fault 2.

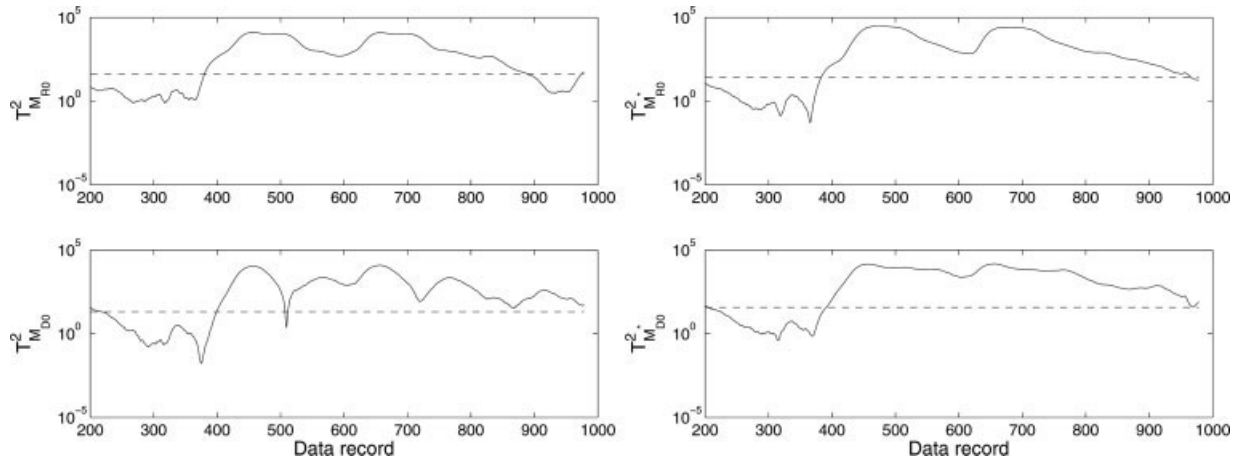

Figure 5. Four T^2 criteria for process fault 3.

Analysis of fault 2. By inspecting the identified fault signatures for the input variable set variables 1 and 3 can be isolated as significant contributors to this event. Variable 3 (fresh feed flow) showed a decrease in value from around 25 to below 20 over a prolonged period of time. After detecting this severe drop, the process operators reduced the reflux flow (variable 1) to reduce the impact of the feed drop (which went unnoticed in the previous example).

The analysis of the fault signatures for the output variables highlighted that variables 1 and 2 (tray 31 and 2 temperatures) and variables 4 and 6 (top and bottom draw) had to most significant response to the drop in feed and the alteration of the reflux flow.

The additional impact of changing the reflux flow produced a stronger affect upon the enthalpy balance within the unit. This can be noticed by the contribution of variables 5 and 7 (increase in tray 31 temperature and reboiler outlet temperature). This reduction in the material balance produced a decrease in the amount of liquid at the bottom of the column and by a constant reboiler steam flow slowly increased the temperature of this liquid (tray 31 temperature and reboiler outlet temperature). Again, the analysis above suggested that the initial problem can be traced to a reduction in the feed flow entering the unit.

Analysis of fault 3. The third data set which described the occurrence of a significant and prolonged drop in fresh feed was shorter than the previous sets. The inspection of the identified fault signatures for this data set yielded the fresh feed flow as the main contributor of the input variable set and the tray 14 and 2 temperatures (variables 1 and 2) and the top and bottom flow (variables 4 and 6) and to a minor extend the reboiler outlet temperature (variables 7) as the most affected variables of the output set. As earlier, this picture would lead and experienced plant operator to attribute this event to the drop in fresh feed.

Fluidized bed reactor

The second application study was based on the analysis of recorded data from an industrial fluidized bed reactor. This process has six input and 35 output variables. A total of 5 input feeds enter the reactor to produce two chemical compounds that are processed in downstream units. The sixth input variable is a stream required to reduce the pressure in

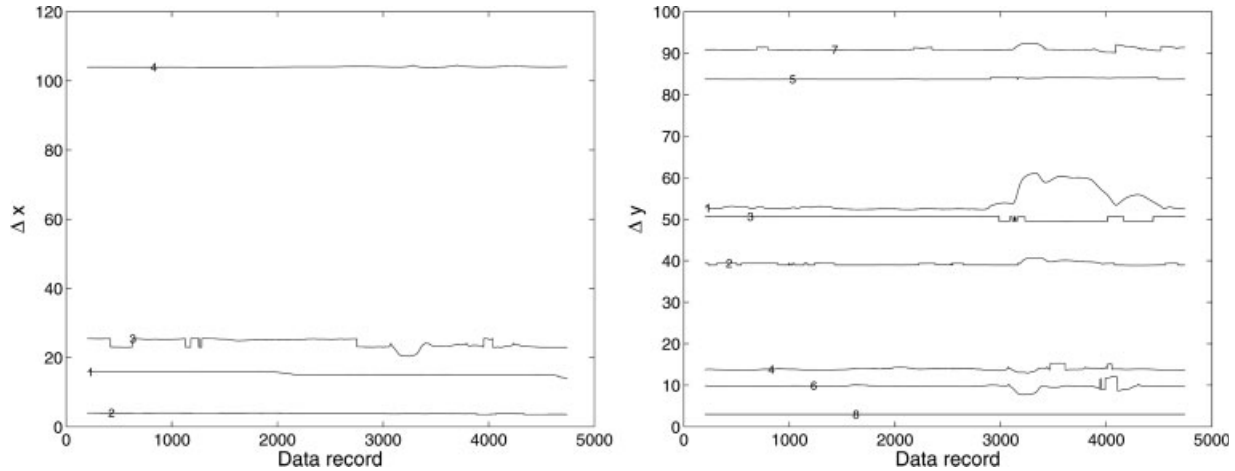

Figure 6. Identified magnitude of faults in inputs and outputs for process fault 1.

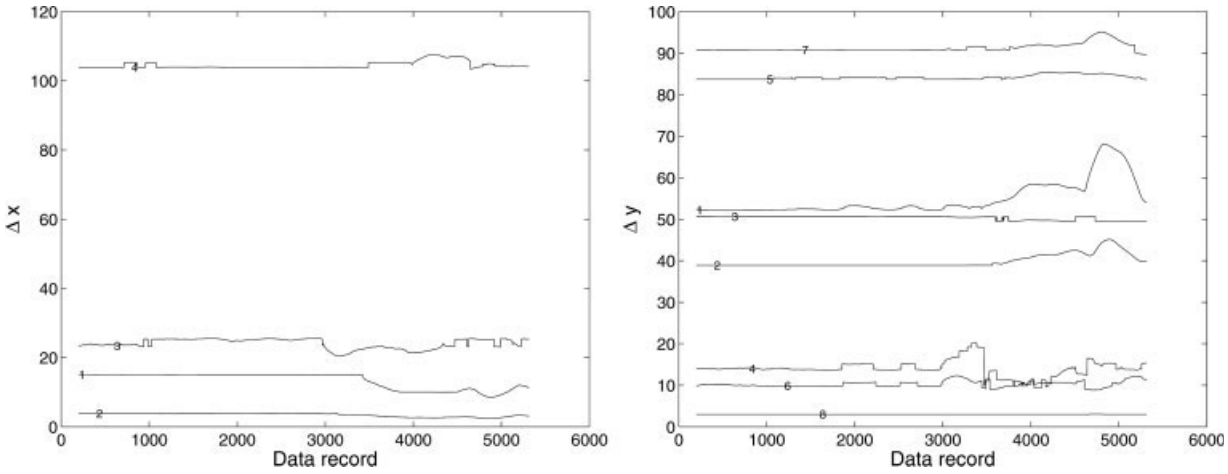

Figure 7. Identified magnitude of faults in inputs and outputs for process fault 2.

an adjacent vaporizer unit which vaporizes two of the input feeds. The output variables are the temperatures of a total of 35 vertically oriented tubes in which the chemical reaction is carried out, supported by fluidized catalyst. From this process, three data sets were recorded, one data set containing normal operating data of the reactor and two operational sets which describe the impact of a drop in steam pressure supplied to the vaporizer and a fluidization problem in one of the tubes.

Applying a singular value decomposition to the augmented data matrix, storing the data of the reference data, and a subsequent PLS analysis yielded the following values for J_i: 3.5201, 2.9173, 2.6273, 1.1723, 0.1615, and 0.0159. This suggested the retention of the first four sets of latent variables. The first six corresponding values of J_i^* for the inverse model were: 3.5201, 2.3882, 0.8511, 0.6951, 0.3470, 0.0926, while the remaining were equal to zero. Since the sixth value is negligible, a total of five sets of latent variables were retained for the inverse model.

Figures 9 and 10 show the T^2 statistics from the two operational data sets. Again, a window k_0 of 100 records is used. In the case of Figure 9, the effect of the drop in steam pressure, fault number 1, was picked just after record index 1000 by $T^2_{\mathbb{M}_{R0}}$. The other statistics detect the fault slightly later, at around the 1300th samples were recorded. The fluidization problem, fault number 2, as shown in Figure 10, is detected after around 1250 samples into the data record by each statistic. The next two subsections discuss the result of estimating the fault signature for each of the process input and output variables.

Analysis of fault 1. Figure 11 plots the identified fault signatures for the first fault condition, the drop in steam pressure. As the steam pressure is not available in the recorded data set, none of the input variables could respond to this event. It was, however, noticeable that input feeds 1, 3 and the input stream to the vaporizer unit to reduce pressure have been altered by plant personnel after the drop in steam pressure had been noticed. These alterations were carried out at around 1900 samples into the data record. It should be noted that the change in first input feed did not have a profound effect upon the tube temperatures.

Inspecting the identified fault signatures of the 35 tube temperatures, it is noticeable that some of the temperatures'

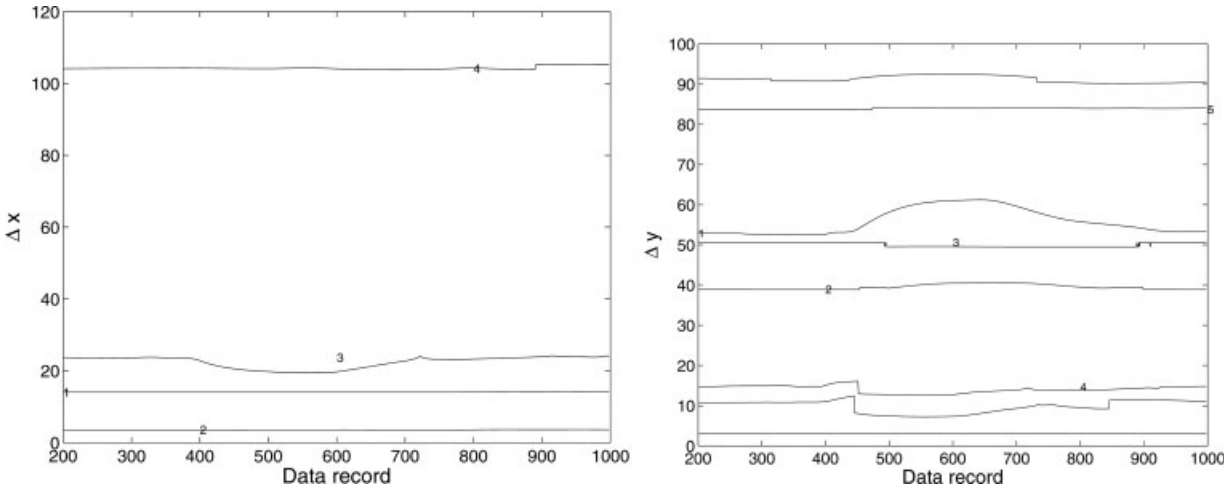

Figure 8. Identified magnitude of faults in inputs and outputs for process fault 3.

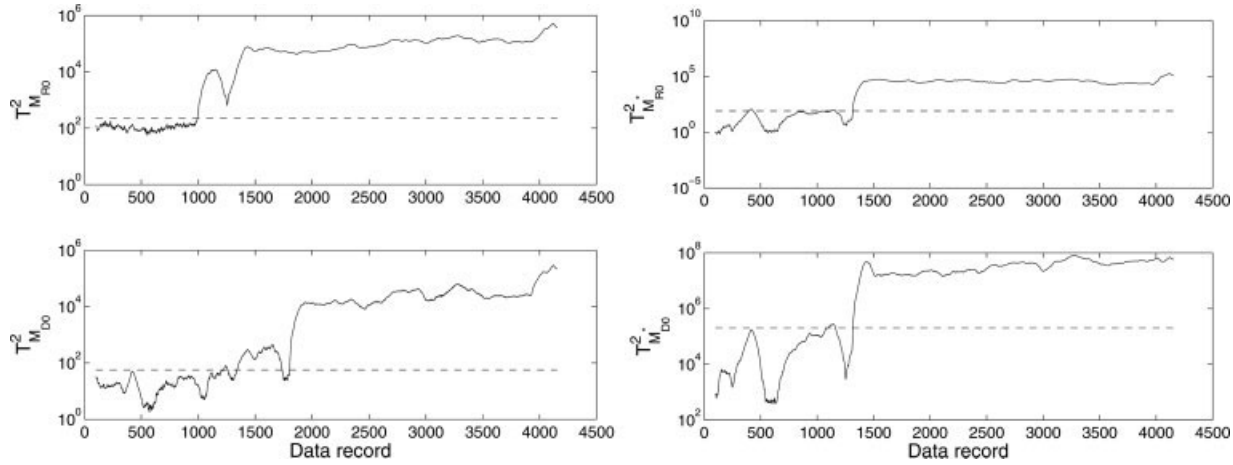

Figure 9. Four T^2 criteria for process fault 1.

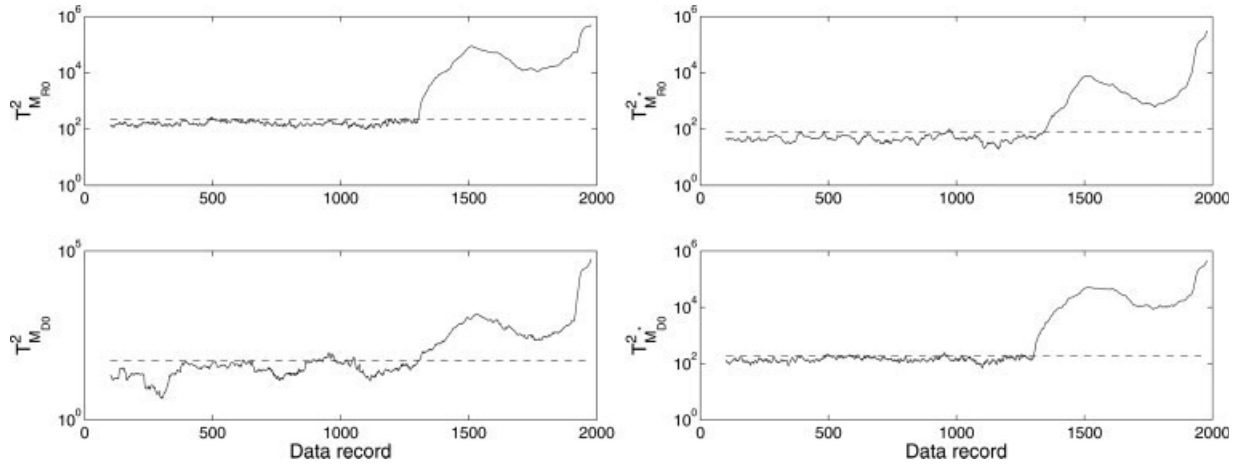

Figure 10. Four T^2 criteria for process fault 2.

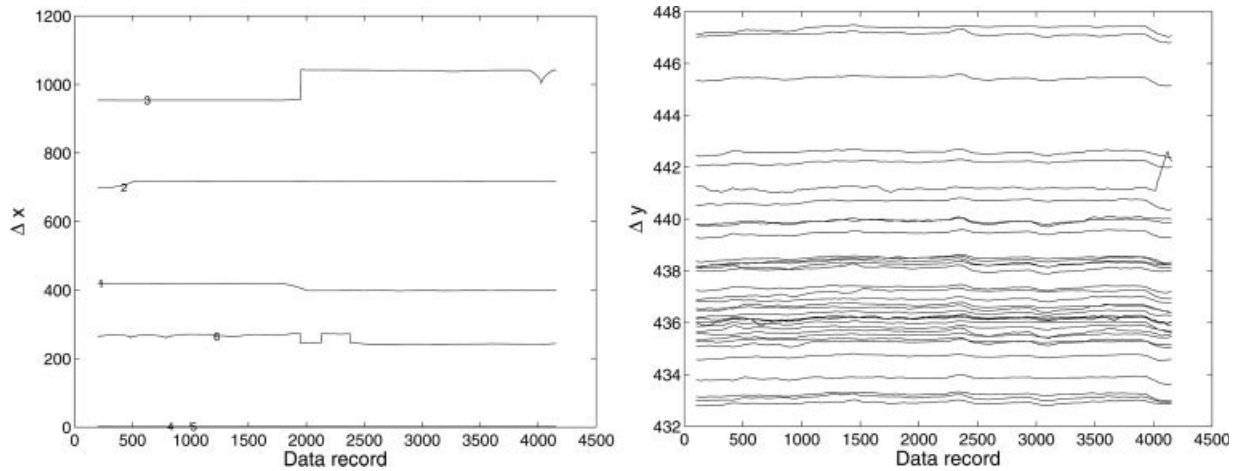

Figure 11. Identified magnitude of faults in inputs and outputs for process fault 1.

increase after the first 1000 samples was recorded. These minor increases, however, changed the relationships between the input and output variables described by the PLS model. The drop in steam pressure reduced the enthalpy with which the steam entered the vaporizer which, in turn, reduced the enthalpy with which two of the input feeds entered the reactor. The reduced enthalpy of these input streams produced the alterations in around half of the tube temperatures. Compared with the increase in tube temperatures that was detected by the plant personnel after 1900 samples were recorded, the PLS based monitoring scheme detected and correctly diagnosed the impact of the drop in steam pressure after about 1000 samples.

Analysis of fault 2. Figure 12 shows the identified fault signatures of each variable for the fluidization problem in one of the tubes. It can be seen that small sporadic peaks in the outputs at 950 records, which could be detected by the $T^2_{\mathbb{M}_{D0}}$ statistic. These minor violations were caused by changes in the second input feed and the stream for reducing the pressure in the vaporizer (input variable 6). These were operator interventions and not the result of the fluidization problem, which arose after about 1250 samples into the data record.

As expected, none of the inputs responded to this event, since fluidization problems are not caused by changes in the inputs. More precisely, such problems emerge if the catalyst density is considerably greater at the bottom of the tube, which enhances the reaction conditions there and consequently increase the temperature recorded at the bottom of each tube. If the fluidization problem is maintained over a longer period of time, it has an undesired effect that negatively affects product quality. In such a case, the tube must be shut down.

By inspecting the identified fault signature of each tube temperature, one particular tube shows a substantial increase in temperature. A closer inspection of Figure 12, and particularly the plot at the bottom of that figure showing the behavior of the tube in question, yields a substantial increase in temperature from the normal level of around 440 degrees (for this specific tube) to around 450 degrees over a prolonged period of time (between the 1250th and 1430th data record). The fluidization conditions in this tube improved after this period and went back to a normal temperature level of 440 degrees. From sample 1970 onwards, the fluidization problem emerged again but this time with a greater impact, which could be noticed from the increase in temperature to a level of over 470 degrees.

Although the first occurrence of the temperature increase in this tube went unnoticed, the second time corrective action was taken by the plant operators to bring the tube temperature back to its normal temperature level. This can be noticed by the sharp decrease in feed level of the sixth input and the temperature dips of each tube. After this intervention failed to inverse this runaway condition, noticeable by a further increase of the temperature to over 540 degrees, the tube was eventually shut down.

The application of the proposed FDD scheme to this chemical reaction process has therefore demonstrated that both faults, although minor if referring to the initial first increase in tube temperature, could be success fully detected and diagnosed. The extracted fault signatures gave a correct representation of both events, which would have enabled an experienced plant operator to narrow down potential root causes for each event.

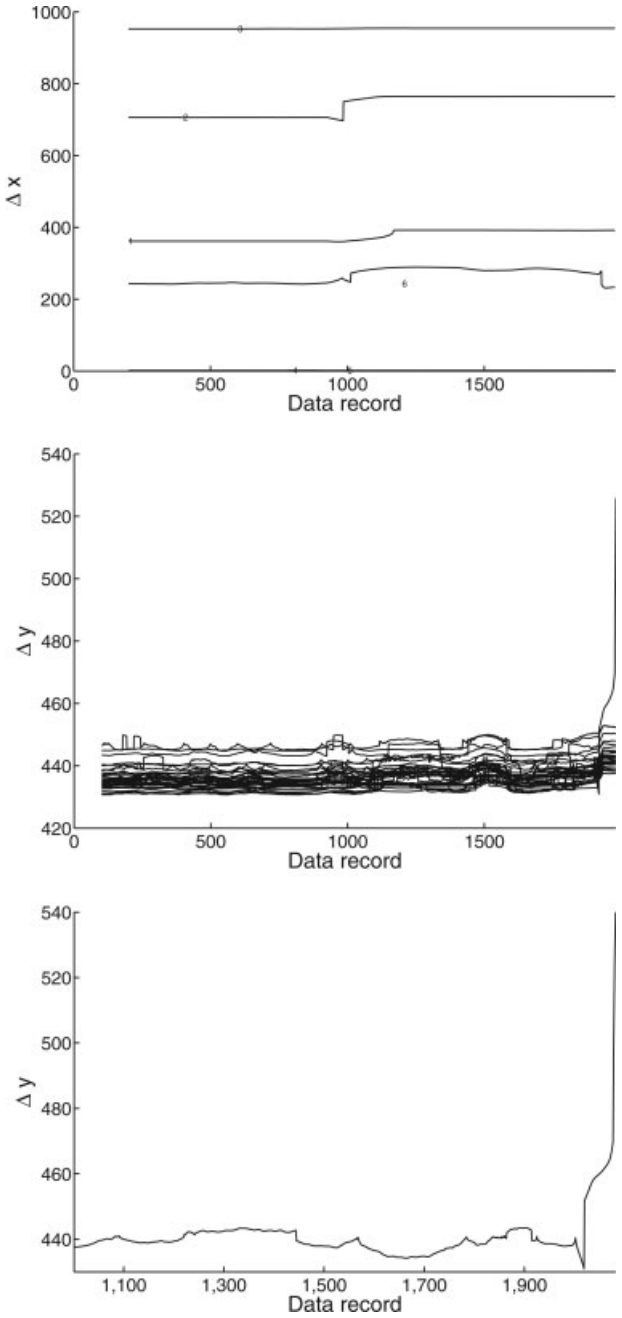

Figure 12. Identified magnitude of faults in inputs and outputs for process fault 2 (upper, middle, and lower plots showing the identified fault signals for predictor variables, the response variables, and the misbehaving tube).

Conclusions

In this article a local approach for the monitoring of industrial processing using a PLS model has been presented. It is shown that the method yields four primary statistics that can

be used for process monitoring. The success of these statistics is demonstrated on a simulated process and two industrial processes with several faults. The methodology is extended by the derivation of additional monitoring functions that can be used in order to identify the magnitude of additive faults in processes. Fault magnitude identification is successfully demonstrated on both simulated and industrial process.

The main contribution in this article is the development of a fault diagnosis scheme that allows to extract the fault signature from the primary residuals of the local PLS model. The choice of PLS within the context of this work relates to its close link to model-based FDD, where both approaches rely on the analysis of input and output variable sets. Further work that involves the proposed local PLS approach will concentrate on multivariate dynamic models, for example, identified state space models, to take advantage of the relationship between local PLS and model-based FDD.

We note that this fault diagnosis scheme can be readily applied to the primary residuals of a PCA model.[33] A potential but not significant advantage of selecting PLS over PCA is that its improved residuals separate between process input and output variables. This, in turn, enables to localize a fault, for example a sensor or actuator fault, within the input or output variable sets by examining the associated monitoring statistics.

Literature Cited

1. Isermann R. Process fault detection based on modeling and estimation methods—a survey. *Automatica*. 1984;20:387–404.
2. Gertler JJ. Survey of model-based failure detection and isolation in complex plants. *IEEE Control Syst Mag*. 1988;8:3–11.
3. Basseville M. Detecting changes in signals and systems—a survey. *Automatica*. 1988;24:309–326.
4. Frank PM. Analytical and qualitative model-based fault diagnosis—a survey and some new results. *Eur J Control*. 1996;2:6–28.
5. Patton R, Frank P, Clark R. *Fault Diagnosis in Dynamic Systems—Theory and Applications*. London: Prentice Hall, 1989.
6. Chen J, Patton RJ. *Robust Model-Based Fault Diagnosis for Dynamic Systems. Asian Studies in Computer and Information Science*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999.
7. Russel EL, Chiang LH, Braatz RD. *Data-Driven Techniques for Fault Detection and Diagnosis in Chemical Processes (Advances in Industrial Control)*. London: Springer, 2000.
8. Bardon O, Sidahmed M. Early detection of leakages in the exhaust and discharge systems of reciprocating machines by vibration analysis. *Mech Syst Signal Process*. 1996;8:551–570.
9. Chen YD, Du R, Qu LS. Fault features of large rotating machinery and diagnosis using sensor fusion. *J Sound Vibration*. 1995;188: 227–242.
10. Kim K, Parlos AG. Reducing the impact of false alarms in induction motor fault diagnosis. *J Dyn Syst Meas Control Trans ASME*. 2003; 125:80–95.
11. Hu N, Chen M, Wen X. The application of stochastic resonance theory for early detecting rub-impact fault of rotor system. *Mech Syst Signal Process*. 2003;17:883–895.
12. Leyval L, Gentil S, Feray-Beaumont S. Model based causal reasoning for process supervision. *Automatica*. 1994;30:1295–1306.
13. Mylaraswamy D, Kavuri S, Venkatasubramanian V. A framework for automated development of causal models for fault diagnosis. In: *Proceedings of the Annual AIChE Meeting*, San Francisco, CA, Nov. 13–18, 1994.
14. Rajaraman S, Hahn J, Mannan MS. A methodology for fault detection, isolation and identification for nonlinear processes with parametric uncertainties. *Ind Eng Chem Res*. 2004;43:6774–6786.
15. Kramer MA, Palowitch BL. A rule-based approach to fault diagnosis using the signed directed graph. *AIChE J*. 1987;33:1067–1078.
16. Iserman R. Fault diagnosis of machines via parameter estimation and knowledge processing: tutorial paper: fault detection, supervision and safety for technical processes. *Automatica*. 1993;29:815–835.
17. Shin R, Lee LS. Use of fuzzy cause-effect digraph for resolution fault diagnosis of process plants. I. Fuzzy cause-effect digraph. *Ind Eng Chem Res*. 1995;34:1688–1702.
18. Upadhyaya BR, Zhao K, Lu B. Fault monitoring of nuclear power plant sensors and field devices. *Prog Nucl Energy*. 2003;43:337–342.
19. Lehane M, Dube F, Halasz M, Orchard R, Wylie R, Zaluski M. Integrated diagnosis system (ids) for aircraft fleet maintenance. In: Mostow J, Rich C, editors. *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI 98)*, Menlo Park, CA, July 26–30, 1998. Madison, WI: AAAI Press.
20. Ming R, Haibin Y, Heming Y. Integrated distribution intelligent system architecture for incidents monitoring and diagnosis. *Comput Ind*. 1998;37:143–151.
21. Qing Z, Zhihan X. Design of a novel knowledge-based fault detection and isolation scheme. *IEEE Trans Syst Man Cybernet Part B (Cybernetics)*. 2004;34:1089–1095.
22. Shing CT, Chee PL. Application of an adaptive neural network with symbolic rule extraction to fault detection and diagnosis in a power generation plant. *IEEE Trans Energy Conversion*. 2004;19:369–377.
23. Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis. Part i: quantitative model-based methods. *Comput Chem Eng*. 2003;27:293–311.
24. Venkatasubramanian V, Rengaswamy R, Kavuri SN. A review of process fault detection and diagnosis. Part ii: qualitative models and search strategies. *Comput Chem Eng*. 2003;27:313–326.
25. Venkatasubramanian V, Rengaswamy R, Kavuri SN, Yin K. A review of process fault detection and diagnosis. Part iii: process history based methods. *Comput Chem Eng*. 2003;27:327–346.
26. Cinar A, Palazoglu A, Kayihan F. *Chemical Process Performance Evaluation*. Boca Raton, FL: CRC Press, 2007.
27. MacGregor JF, Marlin TE, Kresta JV, Skagerberg B. Multivariate statistical methods in process analysis and control. In: *AIChE Symposium Proceedings of the Fourth International Conference on Chemical Process Control*, New York, 1991:79–99. AIChE Publication No. P-67.
28. Basseville M. On-board component fault detection and isolation using the statistical local approach. *Automatica*. 1998;34:1391–1415.
29. Wang X, Kruger U, Lennox B. Recursive partial least squares algorithms for monitoring complex industrial processes. *Control Eng Pract*. 2003;11:613–632.
30. G. Box EP, Jenkins GM, Reinsel GC. *Time Series Analysis*. Englewood-Cliffs, NJ: Prentice-Hall, 1994.
31. Superville CR, Adams BM. An evaluation of forecast-based quality control schemes. *Commun Stat: Simul Comput*. 1994;23:645–661.
32. Apley DW, Jianjun S. The glrt for statistical process control of autocorrelated processes. *IIE Trans*. 1999;31:1123–1134.
33. Kruger U, Kumar S, Littler T. Improved principal component modelling using the local approach. *Automatica*. 2007;43:1532–1542.
34. Wold H. Estimation of principal components and related models by iterative least squares. In: Krishnaiah PR, editor. *Multivariate Analysis*. New York: Academic Press, 1966:391–420.
35. Wold H. Non-linear estimation by iterative least squares procedures. In: David F, editor. *Research Papers in Statistics*. New York: Wiley, 1966.
36. Geladi P. Pls model building: theory and application. *Chemometrics Intel Lab Syst*. 1988;5:231–246.
37. Morud T. Multivariate statistical process control; example from the chemical process industry. *J Chemometrics*. 1996;10:669–675.
38. Kruger U, Chen Q, Sandoz DJ, McFarlane RC. Extended pls approach for enhanced condition monitoring of industrial processes. *AIChE J*. 2001;47:2076–2091.
39. Geladi P, Kowalski BR. Partial least squares regression: A tutorial. *Anal Chim Acta*. 1986;185:231–246.

40. de Jong S. Simpls, an alternative approach to partial least squares regression. *Chemometrics Intel Lab Syst*. 1993;18:251–263.
41. Piovoso MJ, Kosanovich KA. Process data chemometric. *IEEE Trans Instrum Meas*. 1992;41:262–268.
42. Ḣoskuldsson A. Pls regression models. *J Chemometrics*. 1988;2: 211–228.
43. Wang X, Kruger U, Irwin GW. Process monitoring approach using fast moving window pca. *Ind Eng Chem Res*. 2005;44:5691–5702.
44. Tracey ND, Young JC, Mason RL. Multivariate control charts for individual observations. *J Quality Technol*. 1992;24:88–95.

Appendix A: Proof of Lemma 1

The proof commences by reformulating Eq. 1:

$$J_i \stackrel{\Delta}{=} \arg\max_{\mathbf{w}_i,\mathbf{v}_i}\{E\{t_i u_i\}\} = \arg\max_{\mathbf{w}_i,\mathbf{v}_i}\{E\{\mathbf{w}_i^T\mathbf{x}_i\mathbf{y}_i^T\mathbf{v}_i\}\} \tag{A1}$$

by incorporating Eq. 2 to produce

$$\begin{bmatrix} E\{\mathbf{x}_i\mathbf{y}_i^T\}^T & -2\lambda_{1,i}\mathbf{I} \\ -2\lambda_{2,i}\mathbf{I} & E\{\mathbf{x}_i\mathbf{y}_i^T\} \end{bmatrix}\begin{pmatrix}\mathbf{w}_i \\ \mathbf{v}_i\end{pmatrix} = 0. \tag{A2}$$

Simplifying the above derivative terms with respect to $\lambda_{1,i}$ and $\lambda_{2,i}$ gives rise to

$$2\lambda_{1,a}\mathbf{w}_i = E\{\mathbf{x}_i\mathbf{y}_i\}\mathbf{v}_i, \quad 2\lambda_{2,i}\mathbf{v}_i = E\{\mathbf{x}_i\mathbf{y}_i^T\}^T\mathbf{w}_i \tag{A3}$$

Pre-multiplying the above equations by $\mathbf{w}_i^T$ and $\mathbf{v}_i^T$ yields

$$\lambda_{1,a} = \frac{\mathbf{w}_i^T E\{\mathbf{x}_i\mathbf{y}_i^T\}\mathbf{v}_i}{2} = \frac{J_i}{2}, \quad \lambda_{2,a} = \frac{\mathbf{v}_i^T E\{\mathbf{x}_i\mathbf{y}_i^T\}^T\mathbf{w}_i}{2} = \frac{J_i}{2} \tag{A4}$$

It follows from Eqs. A4 that

$$\lambda_{1,a} = \lambda_{2,a} \stackrel{\Delta}{=} \lambda_i, \quad J_i = 2\lambda_i \tag{A5}$$

Appendix B: Proof of Eq. 36

Revising the deflation procedure for the predictor matrix to compute the *i*th pair of weight and score vectors gives rise to

$$\begin{aligned} \mathbf{t}_i &= \mathbf{X}_i\mathbf{w}_i = [\mathbf{X}_{i-1} - \mathbf{t}_{i-1}\mathbf{p}_{i-1}^T]\mathbf{w}_i \\ \mathbf{t}_i &= [\mathbf{X} - \mathbf{t}_1\mathbf{p}_1^T - \mathbf{t}_2\mathbf{p}_2^T - \cdots - \mathbf{t}_{i-2}\mathbf{p}_{i-2}^T - \mathbf{t}_{i-1}\mathbf{p}_{i-1}^T]\mathbf{w}_i \\ \mathbf{t}_i &= \mathbf{X}[\mathbf{I} - r_1\mathbf{p}_1^T - r_2\mathbf{p}_2^T - \cdots - r_{i-2}\mathbf{p}_{i-2}^T - r_{i-1}\mathbf{p}_{i-1}^T]\mathbf{w}_i \end{aligned} \tag{B1}$$

which yields the following relationship between the r-vectors and the w-weight vectors:

$$\mathbf{r}_i = \mathbf{w}_i - \sum_{j=1}^{i-1}\mathbf{r}_j\mathbf{p}_j^T\mathbf{w}_i = \mathbf{w}_i - \sum_{j=1}^{i-1}\mathbf{p}_j^T\mathbf{w}_i\mathbf{r}_j. \tag{B2}$$

Manuscript received Sept. 5, 2007, revision received Feb. 21, 2008, and final revision received May 20, 2008.