Regression-Based Analysis of Multivariate Non-Gaussian Datasets for Diagnosing Abnormal Situations in Chemical Processes

Jiusun Zeng
Institute of Thermal Detection and Control, College of Metrology and Measurement Engineering, China Jiliang University, Hangzhou 310018, P.R. China

Dep. of Control, State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, P.R. China

Lei Xie
Dep. of Control, State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, P.R. China

Uwe Kruger
Dept. of Mechanical and Industrial Engineering, Sultan Qaboos University, Muscat, Oman

Chuanhou Gao
Dept. of Mathematics, Zhejiang University, Hangzhou 310027, P.R. China

DOI 10.1002/aic.14230
Published online October 21, 2013 in Wiley Online Library (wileyonlinelibrary.com)

This article presents a regression-based monitoring approach for diagnosing abnormal conditions in complex chemical process systems. Such systems typically yield process variables that may be both Gaussian and non-Gaussian distributed. The proposed approach utilizes the statistical local approach to monitor parametric changes of the latent variable model that is identified by a revised non-Gaussian regression algorithm. Based on a numerical example and recorded data from a fluidized bed reactor, the article shows that the proposed approach is more sensitive when compared to existing work in this area. A detailed analysis of both application studies highlights that the introduced non-Gaussian monitoring scheme extracts latent components that provide a better approximation of non-Gaussian source signal and/or is more sensitive in detecting process abnormities. © 2013 American Institute of Chemical Engineers *AIChE J*, 60: 148–159, 2014
Keywords: non-Gaussian regression, mutual information, non-Gaussian source signals, statistical local approach

Introduction

The requirement for improving reliability, safety, and efficiency in the chemical process industry has led to an increased research interest in the area of fault detection and diagnosis (FDD) over the past few decades. This culminated in the introduction of different approaches,[1–5] which can be divided into three main categories: (1) quantitative model-based methods, (2) qualitative model representations, and (3) methods based on feature extraction.

A large number of FDD concepts for industrial systems are model-based, for example, methods based on parameter estimation, parity equations, or state observers.[5] In practice, however, it is often very difficult to obtain an accurate mechanistic process model for complex chemical systems. Resulting from advances in sensor and computer technology, automatic data acquisition systems provide the basis for recording larger numbers of variables. This, in turn, allows identifying statistical-based process models and analyzing them to determine the state of the process operation.[6,7]

Multivariate statistical models rely on latent variable (LV) techniques, such as principal component analysis (PCA)[8,9] and partial least squares (PLS).[10,11] Statistically, however, both PCA and PLS only consider second order information under the assumption that the process variables follow a Gaussian distribution. Recent extensions that address this deficiency include the use of independent component analysis (ICA),[12–14] which is a non-Gaussian extension of PCA.

Based on ICA, this article introduces a regression-based approach to monitor processes that produce non-Gaussian data records and involves the recently proposed non-Gaussian regression (NGR) algorithm.[15] The objective function of the NGR algorithm is designed to extract non-Gaussian components based on mutual information and negentropy. The first contribution of this article is to design an NGR algorithm that maximally extract $\mathcal{M}$ LV sets that is equal to the number of input or output variables, whichever is larger. This is accomplished by introducing different orthogonality constraints compared to original NGR algorithm.

Motivated by recent work involving standard PLS,[16] the extracted non-Gaussian LVs are then embedded into the statistical local approach to construct monitoring functions that are, asymptotically, Gaussian distributed. For process monitoring, this, in turn, allows a statistical inference using standard multivariate statistical techniques, for example, the Hotelling's T^2 statistic.

The benefits of utilizing the statistical local approach as a monitoring tool is its increased sensitivity in detecting incipient fault conditions compared to score-based monitoring statistics.[16–18] Consequently, the proposed work combines the advantages of (1) a NGR technique, (2) a simplified statistical inference using low dimensional LV sets, and (3) the increased

Correspondence concerning this article should be addressed to L. Xie at leix@iipc.zju.edu.cn and U. Kruger at uwekruger@squ.edu.om.

sensitivity associated with the statistical local approach. The application studies to a simulation example and recorded data from an industrial process demonstrate these benefits over the use of ICA and the PLS-based monitoring approach.[16]

The remainder of the article is organized as follows. The next section gives preliminaries of the NGR algorithm and the statistical local approach. The section of Defining Primary Residual Functions then develops the proposed regression-based monitoring approach, followed by constructing monitoring statistics for diagnosing process abnormalities in the section of Fault Diagnosis Using Primary Residuals. Application studies involving a simulation example and an industrial fluidized bed reactor are presented in the following two sections, respectively. The last section provides a Concluding Summary.

Preliminaries

This section presents a summary of NGR and the statistical local approach. Without restriction of generality, the following discussion assumes that the input and output vectors, $\mathbf{x} \in \mathbb{R}^M$ and $\mathbf{y} \in \mathbb{R}^N$ have zero mean.

NGR algorithm

The core NGR algorithm relies on the construction of a regression model by extracting latent components that maximize a multiobjective function.[15] Defining the sample projections of the input and output variables onto weight vectors of unit length, $\mathbf{w} \in \mathbb{R}^M$ and $\mathbf{v} \in \mathbb{R}^N$, by $t=\mathbf{x}^T\mathbf{w}$ and $u=\mathbf{y}^T\mathbf{v}$, this multiobjective function includes weighted contributions of:

- the negentropy of t;
- mutual information between t and u; and
- the negentropy of u.

Prior to the determination of $\mathbf{w}$ and $\mathbf{v}$, the variable sets $\mathbf{x}$ and $\mathbf{y}$ are prewhitened, which is a common practice for extracting independent components[12]

$$\tilde{\mathbf{x}}=\mathrm{diag}\left(\lambda_1^{-\frac{1}{2}} \quad \lambda_2^{-\frac{1}{2}} \quad \cdots \quad \lambda_a^{-\frac{1}{2}}\right)\mathbf{P}_x\mathbf{x} \tag{1a}$$

$$\tilde{\mathbf{y}}=\mathrm{diag}\left(\mu_1^{-\frac{1}{2}} \quad \mu_2^{-\frac{1}{2}} \quad \cdots \quad \mu_b^{-\frac{1}{2}}\right)\mathbf{P}_y\mathbf{y} \tag{1b}$$

Here, $\mathbf{P}_x$ and $\mathbf{P}_y$ store the orthonormal eigenvectors, diag $(\cdot)$ is a diagonal matrix, and λ_i and μ_j are the ith and jth largest eigenvalue of $\mathbf{S}_{xx}=E\{\mathbf{x}\mathbf{x}^T\}$ and $\mathbf{S}_{yy}=E\{\mathbf{y}\mathbf{y}^T\}$, respectively. For simplicity, the remainder of this article refers to the prewhitened input and output vectors as $\mathbf{x}$ and $\mathbf{y}$. The integer values a and b denote the number of nonzero eigenvalues of $\mathbf{S}_{xx}$ and $\mathbf{S}_{yy}$, respectively.

Using the scaled input and output vectors, the multiple objective function for extracting the ith pair of weight vectors is defined as follows

$$\begin{aligned}\begin{pmatrix}\hat{\mathbf{w}}_i\\ \hat{\mathbf{v}}_i\end{pmatrix}=\arg\max_{\mathbf{w}_i\mathbf{v}_i} \; & \underbrace{\alpha\left[E\{G(\mathbf{w}_i^T\mathbf{x})\}-E\{G(\upsilon)\}\right]^2}_{\text{Negentropy of } t}+\\ & \underbrace{\frac{\beta}{12}\left(E\{\mathbf{w}_i^T\mathbf{x}\}^3\right)^2+\frac{\beta}{12}\left(E\{\mathbf{v}_i^T\mathbf{y}\}^3\right)^2+\frac{\beta}{4}\left(E\left\{(\mathbf{w}_i^T\mathbf{x})^2\mathbf{v}_i^T\mathbf{y}\right\}\right)^2}_{\text{Negentropy between } t \text{ and } u}+\\ & \underbrace{\frac{\beta}{4}\left(E\left\{\mathbf{w}_i^T\mathbf{x}(\mathbf{v}_i^T\mathbf{y})^2\right\}\right)^2-\frac{\beta}{2}\log\left(1-E\{\mathbf{w}_i^T\mathbf{x}\mathbf{y}^T\mathbf{v}_i\}^2\right)}_{\text{Negentropy between } t \text{ and } u}+\\ & \underbrace{\gamma\left[E\{G(\mathbf{v}_i^T\mathbf{y})\}-E\{G(\nu)\}\right]^2}_{\text{Negentropy of } u}\end{aligned} \tag{2}$$

Here, υ and ν are zero mean Gaussian distributed variables of the same variance as t and u, respectively, and α, β, and $\gamma, \alpha+\beta+\gamma=1$, are parameters. The multiobjective function is subject to the following constraints

$$\mathbf{w}_i^T\mathbf{w}_i-1=0 \quad \mathbf{v}_i^T\mathbf{v}_i-1=0 \tag{3}$$

Appendix A provides a detailed derivative of Eq. 2. Ref. 15 shows how to obtain an optimal solution for the ith pair of weight vectors and how to determine α, β, and γ. To determine the $(i+1)$th set of LVs, the NGR algorithm requires the incorporation of orthogonality constraints for both sets of weight vectors. For process monitoring applications, however, this leads to the following two problems:

(P1) if $M \geq N$, that is, there are at least as many input as output variables; and

(P2) if $M < N$, there are more output than input variables.

To address problems (P1) and (P2), orthogonality constraints are only imposed on the w-weight vectors and the v-weight vectors, respectively. This guarantees that the number of LV sets is $\mathcal{M}=\max\{M,N\}$.

Statistical local approach

The formulation of the statistical local approach, commonly used for change detection and fault diagnosis, relies on the central limit theorem.[17–19] Denoting $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}$ as parameter vectors representing normal and abnormal process behavior, respectively, the statistical local approach is based on defining $\boldsymbol{\theta}$ as follows

$$\boldsymbol{\theta}=\boldsymbol{\theta}_0+\frac{\Delta\boldsymbol{\theta}}{\sqrt{K}} \tag{4}$$

where $\Delta\boldsymbol{\theta}$ is an unknown but otherwise fixed vector describing the impact of the fault upon the fault-free parameter vector $\boldsymbol{\theta}_0$ and K is the number of recorded samples describing the fault condition. Next, defining $\mathbf{z}_k$ as a data vector storing the recorded values for each variable at the kth sampling instance, fault detection is carried out by monitoring the vector-valued primary residual function $\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_k)$. Appendix B shows the requirements for defining primary residual functions.

Using the primary residual function $\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_k)$, an improved residual function can be defined as follows

$$\boldsymbol{\zeta}(\boldsymbol{\theta}_0)=\frac{1}{\sqrt{K}}\sum_{k=1}^{K}\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_k) \tag{5}$$

In the fault-free case, $\boldsymbol{\zeta}(\boldsymbol{\theta}_0)$ follows a multivariate Gaussian distribution of zero mean and covariance

$$\mathbf{S}_{\phi\phi}=\lim_{K\to\infty}\frac{1}{K}\sum_{i=1}^{K}\sum_{j=1}^{K}\boldsymbol{\phi}(\boldsymbol{\theta}_0, \mathbf{z}_i)\boldsymbol{\phi}^T(\boldsymbol{\theta}_0, \mathbf{z}_j) \tag{6}$$

It is important to note that the covariance matrix for $\boldsymbol{\zeta}(\boldsymbol{\theta}_0+\frac{\Delta\boldsymbol{\theta}}{\sqrt{K}})$ is identical to that of $\zeta(\boldsymbol{\theta}_0)$ but the mean vector $E\{\boldsymbol{\zeta}(\boldsymbol{\theta}_0+\frac{\Delta\boldsymbol{\theta}}{\sqrt{K}})\}$ departs from zero. Thus, a Hotelling's T^2 statistics can be utilized to detect a change in mean.[16] The next section details the determination of the primary residual functions for the NGR algorithm.

Defining Primary Residual Functions

The primary residual vector $\boldsymbol{\theta}$ can be obtained on the basis of the objective function in Eq. 2. Commencing by reformulating Eq. 2

$$\begin{aligned}\mathcal{J}_i=\max_{\mathbf{w}_i\mathbf{v}_i} \; & \alpha\left[E\{G(\mathbf{w}_i^T\mathbf{x})\}-E\{G(\upsilon)\}\right]^2+\frac{\beta}{12}\left(E\{\mathbf{w}_i^T\mathbf{x}\}^3\right)^2\\ & +\frac{\beta}{12}\left(E\{\mathbf{v}_i^T\mathbf{y}\}^3\right)^2+\frac{\beta}{4}\left(E\left\{(\mathbf{w}_i^T\mathbf{x})^2\mathbf{v}_i^T\mathbf{y}\right\}\right)^2\\ & +\frac{\beta}{4}\left(E\left\{\mathbf{w}_i^T\mathbf{x}(\mathbf{v}_i^T\mathbf{y})^2\right\}\right)^2-\frac{\beta}{2}\log\left(1-\left(E\{\mathbf{w}_i^T\mathbf{x}\mathbf{y}^T\mathbf{v}_i\}\right)^2\right)\\ & +\gamma\left[E\{G(\mathbf{v}_i^T\mathbf{y})\}-E\{G(\nu)\}\right]^2-\lambda_{1,i}(\mathbf{w}_i^T\mathbf{w}_i-1)-\lambda_{2,i}(\mathbf{v}_i^T\mathbf{v}_i-1),\end{aligned} \tag{7}$$

where $\lambda_{1,i}$ and $\lambda_{2,i}$ are Lagrangian multipliers, the optimal solution satisfies

$$\hat{\mathbf{w}}_i=\arg\max_{\mathbf{w}_i}\left(J_i-\lambda_{1,i}\left(\mathbf{w}_i^T\mathbf{w}_i-1\right)\right)\quad\Rightarrow\quad\frac{\partial J_i}{\partial\mathbf{w}_i}\Big|_{\hat{\mathbf{w}}_i}-2\lambda_{1,i}\hat{\mathbf{w}}_i=0 \tag{8a}$$

$$\hat{\mathbf{v}}_i=\arg\max_{\mathbf{v}_i}\left(J_i-\lambda_{2,i}\left(\mathbf{v}_i^T\mathbf{v}_i-1\right)\right)\quad\Rightarrow\quad\frac{\partial J_i}{\partial\mathbf{v}_i}\Big|_{\hat{\mathbf{v}}_i}-2\lambda_{2,i}\hat{\mathbf{v}}_i=0 \tag{8b}$$

In Eqs. 8, J_i is the value of the objective function in Eq. 2. Working out the partial derivative of $\mathcal{J}_i$ with respect to $\mathbf{w}_i$ yields

$$\begin{aligned}\frac{\partial\mathcal{J}_i}{\partial\mathbf{w}_i}&=2\alpha\left[E\left\{G\left(\mathbf{w}_i^T\mathbf{x}\right)\right\}-E\{G(\upsilon)\}\right]E\left\{\mathbf{x}\frac{\partial G\left(\mathbf{w}_i^T\mathbf{x}\right)}{\partial\left(\mathbf{w}_i^T\mathbf{x}\right)}\right\}\\&+\frac{\beta E\{\mathbf{x}\mathbf{y}^T\mathbf{v}_i\}E\left\{\mathbf{w}_i^T\mathbf{x}\mathbf{y}^T\mathbf{v}_i\right\}}{1-\left(E\left\{\mathbf{w}_i^T\mathbf{x}\mathbf{y}^T\mathbf{v}_i\right\}\right)^2}+\frac{\beta}{2}\left(E\left\{\mathbf{w}_i^T\mathbf{x}\right\}\right)3E\left\{\mathbf{x}\left(\mathbf{w}_i^T\mathbf{x}\right)^2\right\}\\&+\beta E\left\{\left(\mathbf{w}_i^T\mathbf{x}\right)^2\mathbf{y}^T\mathbf{v}_i\right\}E\left\{\mathbf{x}\left(\mathbf{w}_i^T\mathbf{x}\right)\mathbf{y}^T\mathbf{v}_i\right\}\\&+\frac{\beta}{2}E\left\{\mathbf{w}_i^T\mathbf{x}\left(\mathbf{y}^T\mathbf{v}_i\right)^2\right\}E\left\{\mathbf{x}\left(\mathbf{y}^T\mathbf{v}_i\right)^2\right\}-2\lambda_{1,i}\mathbf{w}_i\end{aligned} \tag{9}$$

Defining $\hat{s}_{1i}=\hat{\mathbf{w}}_i^T\mathbf{x},\hat{s}_{2i}=\hat{\mathbf{v}}_i^T\mathbf{y}$ and $g(\hat{s}_{1i})=[E\{G(\hat{s}_{1i})\}-E\{\upsilon\}]\frac{\partial G(\hat{s}_{1i})}{\partial\hat{s}_{1i}}$ allows simplifying Eq. 9 to become

$$\begin{aligned}\frac{\partial\mathcal{J}_i}{\partial\mathbf{w}_i}\Big|_{\mathbf{w}_i=\hat{\mathbf{w}}_i}&=\alpha E\{\mathbf{x}g(\hat{s}_{1i})\}+\frac{\beta E\left\{\mathbf{x}\hat{s}_{2i}^T\right\}E\{\hat{s}_{1i}\hat{s}_{2i}\}}{1-\left(E\left\{\hat{s}_{1i}\hat{s}_{2i}^T\right\}\right)^2}\\&+\frac{\beta}{2}E\left\{\hat{s}_{1i}^3\right\}E\left\{\mathbf{x}\hat{s}_{1i}^2\right\}+\beta E\left\{\hat{s}_{1i}^2\hat{s}_{2i}\right\}E\left\{\mathbf{x}\hat{s}_{1i}\hat{s}_{2i}\right\}\\&+\frac{\beta}{2}E\left\{\hat{s}_{1i}\hat{s}_{2i}^2\right\}E\left\{\mathbf{x}\hat{s}_{2i}^2\right\}-2\lambda_{1,i}\hat{\mathbf{w}}_i\end{aligned} \tag{10}$$

Setting the above derivative zero and multiply both sides by $\hat{\mathbf{w}}_i^T$ now allows to determine $\lambda_{1,i}$

$$\begin{aligned}\lambda_{1,i}&=\alpha E\{\hat{s}_{1i}g(\hat{s}_{1i})\}+\frac{\beta\left(E\{\hat{s}_{1i}\hat{s}_{2i}\}\right)^2}{2-2\left(E\{\hat{s}_{1i}\hat{s}_{2i}\}\right)^2}+\frac{\beta}{4}\left(E\left\{\hat{s}_{1i}^3\right\}\right)^2\\&+\frac{\beta}{2}\left(E\left\{\hat{s}_{1i}^2\hat{s}_{2i}\right\}\right)^2+\frac{\beta}{4}\left(E\left\{\hat{s}_{1i}\hat{s}_{2i}^2\right\}\right)^2\end{aligned} \tag{11}$$

Similarly, the partial derivative of $\mathcal{J}_i$ with respect to $\mathbf{v}_i$ produces

$$\begin{aligned}\frac{\partial\mathcal{J}_i}{\partial\mathbf{v}_i}&=2\gamma\left[E\left\{G\left(\mathbf{v}_i^T\mathbf{y}\right)\right\}-E\{G(v)\}\right]E\left\{\mathbf{y}\frac{\partial G\left(\mathbf{v}_i^T\mathbf{y}\right)}{\partial\left(\mathbf{v}_i^T\mathbf{y}\right)}\right\}\\&+\frac{\beta E\{\mathbf{y}\mathbf{x}^T\mathbf{w}_i\}E\left\{\mathbf{w}_i^T\mathbf{x}\mathbf{y}^T\mathbf{v}_i\right\}}{1-\left(E\left\{\mathbf{w}_i^T\mathbf{x}\mathbf{y}^T\mathbf{v}_i\right\}\right)^2}+\frac{\beta}{2}E\left\{\mathbf{v}_i^T\mathbf{y}\right\}3E\left\{\mathbf{y}\left(\mathbf{v}_i^T\mathbf{y}\right)2\right\}\\&+\beta E\left\{\left(\mathbf{v}_i^T\mathbf{y}\right)^2\mathbf{x}^T\mathbf{w}_i\right\}E\left\{\mathbf{y}\left(\mathbf{v}_i^T\mathbf{y}\right)\mathbf{x}^T\mathbf{w}_i\right\}\\&+\frac{\beta}{2}E\left\{\mathbf{v}_i^T\mathbf{y}\left(\mathbf{x}^T\mathbf{w}_i\right)^2\right\}E\left\{\mathbf{y}\left(\mathbf{x}^T\mathbf{w}_i\right)^2\right\}-2\lambda_{2,i}\mathbf{v}_i\end{aligned} \tag{12}$$

Again, denoting $g(\hat{s}_{2i})=[E\{G(\hat{s}_{2i})\}-E\{v\}]\frac{\partial G(\hat{s}_{2i})}{\partial\hat{s}_{2i}}$, setting the derivative to zero and multiply both sides with $\hat{\mathbf{v}}_i^T$ allows determining $\lambda_{2,i}$

$$\begin{aligned}\lambda_{2,i}&=\gamma E\{\hat{s}_{2i}g(\hat{s}_{2i})\}+\frac{\beta\left(E\{\hat{s}_{1i}\hat{s}_{2i}\}\right)^2}{2-2\left(E\{\hat{s}_{1i}\hat{s}_{2i}\}\right)^2}+\frac{\beta}{4}\left(E\left\{\hat{s}_{2i}^3\right\}\right)^2\\&+\frac{\beta}{2}\left(E\left\{\hat{s}_{1i}\hat{s}_{2i}^2\right\}\right)^2+\frac{\beta}{4}\left(E\left\{\hat{s}_{1i}^2\hat{s}_{2i}\right\}\right)^2\end{aligned} \tag{13}$$

Next, writing Eqs. 10 and 12 in a compact form yields the following vector valued function

$$\mathbf{f}_i=\begin{pmatrix}\mathcal{F}_i^{(x)}-2\lambda_{1,i}\mathbf{w}_i\\\mathcal{F}_i^{(x)}-2\lambda_{2,i}\mathbf{v}_i\end{pmatrix} \tag{14}$$

where $\mathcal{F}_i^{(x)}$ and $\mathcal{F}_i^{(y)}$ are derivatives of J_i with respect to $\mathbf{w}_i$ and $\mathbf{v}_i$, respectively. It follows from Eq. 14 that

$$E\{\mathbf{f}_i\}\Big|_{\begin{pmatrix}\mathbf{w}_i=\hat{\mathbf{w}}_i\\\mathbf{v}_i=\hat{\mathbf{v}}_i\end{pmatrix}}=0 \tag{15}$$

The vector-valued function $\mathbf{f}_i$ satisfies Assumption 1 in Appendix B. It can also be shown easily that the remaining three assumptions are also satisfied.[19] Therefore, $\mathbf{f}_i$ can be used as a primary residual vector for process monitoring by evaluating $E\{\mathbf{f}_i\}$. The modeling procedure yields a total of n significant LVs that describe non-Gaussian signal components. It important to note that the proposed NGR algorithm reduces to canonical correlation regression (CCR) if the signal components are Gaussian distributed.[15] This algorithm, therefore, allows extracting the relevant components for producing a regression model between the input and output variable sets by retaining the n significant components. The integer n can be determined, for example, using cross validation or a bootstrapping method.[15] The separation between retained and discarded components gives rise to the construction of two different sets of vector-valued primary residual functions that are constructed by stacking the $\mathcal{M}=\max\{M,N\}$ components as follows

$$\boldsymbol{\phi}_R^T=\begin{pmatrix}\mathbf{f}_1^T&\mathbf{f}_2^T&\cdots&\mathbf{f}_n^T\end{pmatrix}\in\mathbb{R}^{(M+N)\cdot n} \tag{16a}$$

$$\boldsymbol{\phi}_D^T=\begin{pmatrix}\mathbf{f}_{n+1}^T&\mathbf{f}_{n+2}^T&\cdots&\mathbf{f}_N^T\end{pmatrix}\in\mathbb{R}^{(M+N)\cdot(\mathcal{M}-n)} \tag{16b}$$

Based on the analysis in Ref. 17 for PCA, a second set of primary residuals can be obtained by premultiplying Eq. 14 with the augmented vector $\begin{pmatrix}\mathbf{w}_i^T&\mathbf{v}_i^T\end{pmatrix}$, which produces

$$f_i=\mathbf{w}_i^T\mathcal{F}_i^{(x)}+\mathbf{v}_i^T\mathcal{F}_i^{(y)}-\left(2\lambda_{1,i}+2\lambda_{2,i}\right). \tag{17}$$

Similar to Eq. 16, dividing the $\mathcal{M}$ primary residuals f_i into retained and discarded components yields

$$\phi_R^T=\begin{pmatrix}f_1&f_2&\cdots&f_n\end{pmatrix}\in\mathbb{R}^n \tag{18a}$$

$$\phi_D^T=\begin{pmatrix}f_{n+1}&f_{n+2}&\cdots&f_{\mathcal{M}}\end{pmatrix}\in\mathbb{R}^{\mathcal{M}-n} \tag{18b}$$

Fault Diagnosis Using Primary Residuals

To utilize the primary residuals for process monitoring requires the definition of improved residuals. Refs. 16 and 17 proposed the use of a moving window approach by summing the primary residuals that are inside the sliding window. This sum of primary residuals then yields improved residuals that, asymptotically, converge to a Gaussian distribution function

$$\boldsymbol{\zeta}_{R_k}=\frac{1}{\sqrt{k_0}}\sum_{j=k-k_0+1}^{k}\boldsymbol{\phi}_{R_j}\qquad\boldsymbol{\zeta}_{D_k}=\frac{1}{\sqrt{k_0}}\sum_{j=k-k_0+1}^{k}\boldsymbol{\phi}_{D_j} \tag{19a}$$

$$\zeta_{R_k}=\frac{1}{\sqrt{k_0}}\sum_{j=k-k_0+1}^{k}\phi_{R_j}\qquad \zeta_{D_k}=\frac{1}{\sqrt{k_0}}\sum_{j=k-k_0+1}^{k}\phi_{D_j} \tag{19b}$$

where k_0 is the window length, k denotes the current sample, and the integer j refers to the samples inside the sliding window. Equation 19 allows defining Hotelling's T^2 statistics from the improved residual vectors

$$T^2_{\zeta_R}=\boldsymbol{\zeta}_R^T\mathbf{S}^{-1}_{\boldsymbol{\phi}_R\boldsymbol{\phi}_R}\boldsymbol{\zeta}_R\qquad T^2_{\zeta_D}=\boldsymbol{\zeta}_D^T\mathbf{S}^{-1}_{\boldsymbol{\phi}_D\boldsymbol{\phi}_D}\boldsymbol{\zeta}_D \tag{20a}$$

$$T^2_{\zeta_R}=\zeta_R^T\mathbf{S}^{-1}_{\phi_R\phi_R}\zeta_R\qquad T^2_{\zeta_D}=\zeta_D^T\mathbf{S}^{-1}_{\phi_D\phi_D}\zeta_D \tag{20b}$$

Here, $\mathbf{S}_{zz}$ denotes the covariance matrix of the variable set $\mathbf{z}$. Under the assumption that the improved residuals asymptotically converge to a multivariate Gaussian distribution, the above Hotelling's T^2 statistics converge, asymptotically, to a χ^2 distribution if the covariance matrices are known *a priori* or an F distribution if they need to be estimated.[20] Confidence or control limits can then be obtained for a significance of 0.01 or 0.05.

The next subsection analyzes the sensitivity of the primary, and hence improved, residuals if a fault condition affects the input, the output, or both variable sets. The detection of an abnormal operating condition is usually followed by diagnosing the root cause of this anomalous behavior. Subsection shows how to diagnose a fault condition, which assists process operators in narrowing down potential root causes responsible for the detected process abnormality.

Sensitivity analysis

This subsection discusses (1) whether the introduced primary residual vectors are capable of detecting all possible fault scenarios and (2) whether the higher dimensional primary residual vectors $\boldsymbol{\phi}_R$ and $\boldsymbol{\phi}_D$ are required or whether the lower dimensional primary residual vectors ϕ_R and ϕ_D are sufficient.

To conduct the sensitivity analysis, we assume that a fault condition can impact the output variables only (Case 1), impacts the input variables only (Case 2), and impacts both variable sets (Case 3). This gives rise to formulate the measured vectors of the input and output variables as follows

$$\mathbf{x}=\mathbf{x}_0+\Delta\mathbf{x}\qquad \mathbf{y}=\mathbf{y}_0+\Delta\mathbf{y}, \tag{21}$$

where $\mathbf{x}_0$ and $\mathbf{y}_0$ are data vectors of the input and output variables that describe normal process behavior and $\Delta\mathbf{x}\in\mathbb{R}^M$ and $\Delta\mathbf{y}\in\mathbb{R}^N$ describe the impact of the fault condition, where

- $||\Delta\mathbf{x}||=0$ and $||\Delta\mathbf{y}||>0$ describes Case 1;
- $||\Delta\mathbf{x}||>0$ and $||\Delta\mathbf{y}||=0$ represents Case 2; and
- $||\Delta\mathbf{x}||>0$ and $||\Delta\mathbf{y}||>0$ refers to Case 3.

The operator $||\cdot||$ returns the length of a vector. Given that the primary residuals are, in fact, based on the derivatives of Eq. 7, the sensitivity analysis can be carried out by assuming that $||\mathbf{x}||\gg||\Delta\mathbf{x}||$ and $||\mathbf{y}||\gg||\Delta\mathbf{y}||$. This allows approximating Eq. 14 using a first-order Taylor expansion

$$\frac{\partial\mathcal{J}_i(\mathbf{x}_0+\Delta\mathbf{x},\mathbf{y}_0+\Delta\mathbf{y})}{\partial\mathbf{w}_i}\Big|_{\mathbf{w}_i=\hat{\mathbf{w}}_i}\approx\underbrace{\frac{\partial\mathcal{J}_i(\mathbf{x}_0,\mathbf{y}_0)}{\partial\mathbf{w}_i}\Big|_{\mathbf{w}=\hat{\mathbf{w}}_i}}_{0}+\boldsymbol{f}_1(\Delta\mathbf{x},\Delta\mathbf{y}) \tag{22a}$$

$$\frac{\partial\mathcal{J}_i(\mathbf{x}_0+\Delta\mathbf{x},\mathbf{y}_0+\Delta\mathbf{y})}{\partial\mathbf{v}_i}\Big|_{\mathbf{v}_i=\hat{\mathbf{v}}_i}\approx\underbrace{\frac{\partial\mathcal{J}_i(\mathbf{x}_0,\mathbf{y}_0)}{\partial\mathbf{v}_i}\Big|_{\mathbf{v}_i=\hat{\mathbf{v}}_i}}_{0}+\boldsymbol{f}_2(\Delta\mathbf{x},\Delta\mathbf{y}). \tag{22b}$$

The terms $\boldsymbol{f}_1(\Delta\mathbf{x},\Delta\mathbf{y})$ and $\boldsymbol{f}_2(\Delta\mathbf{x},\Delta\mathbf{y})$ correspond to the change in mean of $\mathbf{f}_i$. Abbreviating Case 1 to 3 by C1 to C3, the analysis of each individual case using Eqs. 22 yields

(C1) Inspecting Eq. 22 for $\Delta\mathbf{x}=0,\Delta\mathbf{y}\neq0$ gives rise to

$$\begin{aligned}E\{\boldsymbol{f}_1(\Delta\mathbf{x},\Delta\mathbf{y})\}=&\beta E\Big\{\big(\hat{s}^*_{1i}\big)^2\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)\Big\}E\big\{\mathbf{x}_0\hat{s}^*_{1i}\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)\big\}\\&-\beta E\Big\{\big(\hat{s}^*_{1i}\big)^2\hat{s}^*_{2i}\Big\}E\big\{\mathbf{x}_0\hat{s}^*_{1i}\hat{s}^*_{2i}\big\}\\&+\frac{\beta}{2}E\Big\{\hat{s}^*_{1i}\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)^2\Big\}E\Big\{\mathbf{x}_0\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)^2\Big\}\\&-\frac{\beta}{2}E\Big\{\hat{s}^*_{1i}\big(\hat{s}^*_{2i}\big)^2\Big\}E\Big\{\mathbf{x}_0\big(\hat{s}^*_{2i}\big)^2\Big\}-\frac{\beta E\{\mathbf{x}_0\hat{s}^*_{2i}\}E\{\hat{s}^*_{1i}(\hat{s}^*_{2i})\}}{1-\Big(E\Big\{\hat{s}^*_{1i}\big(\hat{s}^*_{2i}\big)^T\Big\}\Big)^2}\\&+\frac{\beta E\{\mathbf{x}_0(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i})\}E\{\hat{s}^*_{1i}(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i})\}}{1-\Big(E\Big\{\hat{s}^*_{1i}\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)^2\Big\}\Big)^2}\end{aligned} \tag{23}$$

In the above derivative $\hat{s}^*_{1i}=\mathbf{w}_i^T\mathbf{x}_0,\hat{s}^*_{2i}=\mathbf{v}_i^T\mathbf{y}_0$ and $\Delta\hat{s}^*_{2i}=\mathbf{v}_i^T\Delta\mathbf{y}_0$. The analysis of Eq. 22 yields

$$\begin{aligned}E\{\boldsymbol{f}_2(\Delta\mathbf{x},\Delta\mathbf{y})\}=&2\gamma g^*\big(\Delta\hat{s}^*_{2i}\big)-2\gamma g^*\big(s^*_{2i}\big)\\&+\frac{\beta E\{(\mathbf{y}_0+\Delta\mathbf{y})\hat{s}^*_{1i}\}E\{\hat{s}^*_{1i}(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i})\}}{1-\big(E\{\hat{s}^*_{1i}(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i})\}\big)^2}\\&-\frac{\beta E\{\mathbf{y}_0\hat{s}^*_{1i}\}E\{\hat{s}^*_{1i}(\hat{s}^*_{2i})\}}{1-\big(E\{\hat{s}^*_{1i}(\hat{s}^*_{2i})\}\big)^2}-\frac{\beta}{2}E\Big\{\big(\hat{s}^*_{2i}\big)^3\Big\}E\Big\{\mathbf{y}_0\big(\hat{s}^*_{2i}\big)^2\Big\}\\&+\frac{\beta}{2}E\Big\{\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)^3\hat{s}_{1i}\Big\}E\Big\{(\mathbf{y}_0+\Delta\mathbf{y})\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)^2\Big\}\\&+\beta E\Big\{\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)^2\hat{s}_{1i}\Big\}E\big\{(\mathbf{y}_0+\Delta\mathbf{y})\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)\hat{s}_{1i}\big\}\\&-\beta E\Big\{\big(\hat{s}^*_{2i}\big)^2\hat{s}^*_{1i}\Big\}E\{\mathbf{y}_0\hat{s}^*_{2i}\hat{s}^*_{1i}\}-\frac{\beta}{2}E\Big\{\hat{s}^*_{2i}\big(\hat{s}^*_{1i}\big)^2\Big\}E\Big\{\mathbf{y}_0\big(\hat{s}^*_{1i}\big)^2\Big\}\\&+\frac{\beta}{2}E\Big\{\big(\hat{s}^*_{2i}+\Delta\hat{s}^*_{2i}\big)\big(\hat{s}^*_{1i}\big)^2\Big\}E\Big\{(\mathbf{y}_0+\Delta\mathbf{y})\big(\hat{s}^*_{1i}\big)^2\Big\}\end{aligned} \tag{24}$$

Given that $E\{\mathbf{x}_0\}=0,E\{\mathbf{y}_0\}=0$, it follows that $E\{\boldsymbol{f}_1(\Delta\mathbf{x},\Delta\mathbf{y})\}=0$. This also yields that all terms in Eq. 24 from the second one onwards are equal to zero. However, the first term in Eq. 24 departs from zero, which implies that $E\{\boldsymbol{f}_2(\Delta\mathbf{x},\Delta\mathbf{y})\}\neq0$. Consequently, the fault condition described by Case 1 is detectable by monitoring the primary residual vectors $\boldsymbol{\phi}_R$ and $\boldsymbol{\phi}_D$.

(C2) Following the same analysis, examining the case for $\Delta\mathbf{x}\neq0,\Delta\mathbf{y}=0$ using Eqs. 22a and 22b yields that $E\{\boldsymbol{f}_1(\Delta\mathbf{x},\Delta\mathbf{y})\}\neq0$ and $E\{\boldsymbol{f}_2(\Delta\mathbf{x},\Delta\mathbf{y})\}=0$. The scenario of Case 2 is, therefore, detectable.

(C3) For Case 3, $\Delta\mathbf{x}\neq0$ and $\Delta\mathbf{y}\neq0$, the analysis of Eqs. 22a and 22b gives rise to $E\{\boldsymbol{f}_1(\Delta\mathbf{x},\Delta\mathbf{y})\}\neq0$ and $E\{\boldsymbol{f}_2(\Delta\mathbf{x},\Delta\mathbf{y})\}\neq0$, which implies that also Case 3 is detectable.

It is easy to show that the mean vector of the primary residual vectors ϕ_R and ϕ_D also departs from zero for each of the three cases, as they are linear combinations of $\boldsymbol{\phi}_R$ and $\boldsymbol{\phi}_D$, respectively. Similar to the work in Ref. 17 for PCA, this, in turn, implies that it is sufficient to utilize the Hotelling's T^2 statistics for the improved residuals ζ_R and ζ_D in Eq. 20b.

Table 1. Comparison of Modeling Accuracy Between NGR, ICA, and PLS

Method	ICA	PLS	NGR ($\alpha = 0$)	NGR ($\alpha = 0.2$)	NGR ($\alpha = 0.4$)	NGR ($\alpha = 0.6$)	NGR ($\alpha = 0.8$)
NE(t_1)	0.5970	0.1756	0.1969	0.2040	0.2124	0.2450	0.5969
NE(t_2)	0.4290	0.0552	0.5909	0.5741	0.5953	0.5967	0.2300
NE(t_3)	0.2736	0.0528	0.3057	0.4183	0.4166	0.4166	0.4273
NE(t_4)	0.2438	0.0377	0.3041	0.2788	0.2795	0.2777	0.2821
NE(t_5)	0.0887	0.0327	0.0435	0.0888	0.0920	0.0873	0.0895
NE(t_6)	0.0875	0.0463	0.0427	0.0823	0.0847	0.0860	0.0847
R_1^2	0.1570	0.7698	0.7961	0.7959	0.7952	0.7930	0.1582
R_2^2	0.1581	0.7825	0.9536	0.9536	0.9536	0.9536	0.9485
R_3^2	0.1584	0.8054	0.9536	0.9536	0.9536	0.9536	0.9487
R_4^2	0.9378	0.8485	0.9537	0.9537	0.9537	0.9537	0.9488
R_5^2	0.9400	0.8664	0.9537	0.9537	0.9537	0.9537	0.9488
R_6^2	0.9400	0.9537	0.9537	0.9537	0.9537	0.9537	0.9507

Fault diagnosis

On the basis of Eq. 17b, fault diagnosis can be carried out under the assumption that the fault condition is described by a fault direction of unity length and a fault magnitude.[14] Next, assuming that there are a total of J fault conditions that represent individual fault directions, denoted here by $\mathfrak{F}_j$, $j = 1, \ldots, J$, the impact of the fault condition upon the input and output variables is

$$\begin{pmatrix} \hat{x}_0 \\ \hat{\mathbf{y}}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} - \mathfrak{F}_j \hat{e}_j. \tag{25}$$

Here, $\hat{\mathbf{x}}_0$ and $\hat{\mathbf{y}}_0$ are estimates of $\mathbf{x}_0$ and $\mathbf{y}_0$, respectively, and, according to Eq. 21, $\mathfrak{F}_j^T \hat{e}_j$ is the estimation of $(\Delta \mathbf{x}^T \Delta \mathbf{y}^T)$. For the jth fault condition, defined by the fault direction $\mathfrak{F}_j$, the fault magnitude, e_j, can be estimated by minimizing the following objective function

$$\hat{e}_j = \arg\min_{e_j} \| \zeta_R^T \mathbf{S}_{\phi_R \phi_R}^{-\frac{1}{2}} + \zeta_D^T \mathbf{S}_{\phi_D \phi_D}^{-\frac{1}{2}} \|^2 \tag{26}$$

Analyzing each of the j fault conditions using Eq. 27 then allows to determine the most likely condition that has caused the abnormal process behavior. If the jth fault condition is correct, that is, if the jth fault condition describes the correct fault subspace, both of the Hotelling's T^2 statistics do not respond to an out-of-statistical-control situation when the corrected data vectors $\hat{\mathbf{x}}_0$ and $\hat{\mathbf{y}}_0$ are utilized to determine Eq. 20b. Reference 14 showed that the assessment of which of the j conditions describes the most likely fault condition can be assisted by dividing $T_{\zeta_R}^2$ and $T_{\zeta_D}^2$ by their associated confidence or control limits.

Simulation Example

This section presents the first comparison between the proposed monitoring scheme, the local PLS approach[16] and conventional ICA.[12] The choice of ICA and local PLS relate to the fact that:

- ICA is a modeling tool for extracting non-Gaussian components; and
- local PLS is a monitoring tool that incorporate the statistical local approach into a PLS-based monitoring scheme.

Utilizing ICA, the proposed monitoring strategy is to apply ICA to the input variable set. The extracted non-Gaussian components are then regressed on the output variables. This yields a total of four monitoring statistics, which are

1. the I_d^2 statistic for the extracted non-Gaussian components;
2. the I_e^2 statistic for the Gaussian components;
3. the squared prediction error (SPE) statistic for the residual space; and
4. the T_{Res}^2 for the regression residuals.

The confidence or control limits for the I_d^2, I_e^2, and I_{Res}^2 statistics are obtained by the kernel density estimator similar to that discussed in Lee et al.[12] The control limit for the SPE statistic is computed based on Theorem 3.1 in Box.[21] The significance of each confidence or control limit level is 0.01. The local PLS approach is applied here according to the discussion in Kruger and Dimitriadis.[16]

This first comparison here is based on a simulation example that involves a total of 10 input and 6 output variables. For the kth sample, the input variables are linear combinations of four non-Gaussian and six Gaussian distributed variables $\mathbf{s}$ and $\mathbf{r}$, respectively

$$\mathbf{x}_k = \mathbf{A}\mathbf{s}_k + \mathbf{B}\mathbf{r}_k \tag{27}$$

The construction of the non-Gaussian and Gaussian signals are as follows

$$\mathbf{s}_k = \begin{pmatrix} \sin(0.06k) \\ \sin(0.3k) \\ 0 \\ 0.45\sin(0.4k) \end{pmatrix} + \begin{pmatrix} 5\cos(0.08k) \\ 0 \\ \cos(0.1k) \\ 3\cos(0.2k) \end{pmatrix}, \mathbf{r}_k \sim \mathcal{N}\{0, \mathbf{I}\} \tag{28}$$

The parameter matrices $\mathbf{A}$ and $\mathbf{B}$ are given by

$$\mathbf{A} = \begin{bmatrix} -0.5419 & -0.1983 & 0.1845 & 0.3533 \\ -0.2787 & 0.3243 & -0.1811 & 0.8849 \\ 0.4553 & 0.9312 & 0.4070 & -0.1825 \\ -0.7747 & -0.0382 & -0.9221 & 0.1616 \\ 0.2155 & -0.4718 & -0.7931 & 0.4139 \\ 0.6315 & 0.1748 & -0.4355 & -0.7824 \\ 0.9625 & -0.8828 & -0.6938 & -0.2056 \\ -0.8725 & 0.1914 & -0.3999 & -0.3106 \\ 0.5999 & -0.1150 & -0.3623 & 0.2600 \\ -0.8061 & 0.0668 & 0.3931 & -0.2609 \end{bmatrix} \tag{29a}$$

$$\mathbf{B}=\begin{bmatrix} -0.6869 & 0.8298 & -0.1508 & -0.4984 & 0.3143 & 0.7724 \\ -0.8423 & 0.7969 & -0.1689 & 0.2102 & -0.2139 & 0.5477 \\ 0.7176 & 0.2929 & 0.4165 & 0.9035 & 0.7804 & 0.0092 \\ 0.8224 & -0.2213 & 0.4027 & -0.7570 & 0.8539 & -0.4402 \\ 0.6650 & 0.7151 & 0.8084 & 0.0004 & -0.9106 & 0.8172 \\ 0.2453 & -0.5186 & -0.3004 & 0.2640 & 0.0784 & -0.0615 \\ -0.4253 & -0.0637 & -0.1987 & -0.2310 & -0.0263 & 0.0458 \\ -0.1975 & 0.7287 & -0.2168 & -0.9401 & 0.4707 & -0.7617 \\ 0.2371 & 0.1976 & -0.3884 & -0.5078 & -0.1991 & -0.2510 \\ 0.5635 & 0.9360 & -0.5784 & 0.5489 & -0.5961 & -0.2499 \end{bmatrix} \tag{29b}$$

The output set is a linear combination of the first two non-Gaussian signals and corrupted by an error vector

$$\mathbf{y}=\mathbf{C}\mathbf{s}_{\text{red}}+\boldsymbol{e} \quad \mathbf{s}_{\text{red}}=\begin{pmatrix} \sin(0.06k)+5\cos(0.08k) \\ \sin(0.3k) \end{pmatrix} \tag{30}$$

$$\mathbf{C}=\begin{bmatrix} 0.4550 & -0.7526 & -0.6914 & 0.0269 & 0.3209 & 0.8825 \\ 0.6112 & -0.5094 & 0.7772 & 0.4026 & -0.6274 & -0.2937 \end{bmatrix} \tag{31}$$

The noise sequences are zero mean Gaussian distributed i.d. sequences

$$\mathbf{e}\sim\mathcal{N}\{0,\mathbf{S}_{ee}\} \qquad \mathbf{S}_{ee}=\text{diag}\,(\,0.0330 \quad 0.0454 \quad 0.0480 \quad 0.0261 \quad 0.0334 \quad 0.0443\,) \tag{32}$$

With this setup, the process input variables contain both, Gaussian and non-Gaussian components, whilst the output variables are only non-Gaussian distributed. From the simulated process, described in Eqs. 27–32, a total of 2000 samples are generated to contrast the three competitive methods.

The selection of the weighting parameters for the proposed NGR technique is a tradeoff between the non-Gaussianity of the extracted components and the regression accuracy. As a rule of thumb, the larger $\alpha+\gamma\leq 1$ is selected the more LVs need to be retained to achieve satisfactory regression accuracy. This follows from the fact that the NGR algorithm focuses on determining components that maximize the non-Gaussianity of t and u. Conversely, if $\beta\leq 1$ is selected to be large, the NGR algorithm emphasizes of extracting latent components that produce a significantly better prediction accuracy of the output variables.[15]

Selecting a number of different combinations for $\alpha+\beta+\gamma=1$, a tradeoff between the accuracy of extracting the non-Gaussian signals that are encapsulated in the output variables and the accuracy of predicting the system outputs can be achieved. The argument for selecting $\gamma=0$ is that the source signals that predict the output variables are present in both variable sets. Table 1 presents the comparison of models obtained from the NGR algorithm using different combinations of the weighting parameters, ICA and PLS.

In Table 1, NE(t_i) is the negentropy of the ith LV t_i, R_i^2 is the R^2 statistic describing the accuracy for predicting the output variables, the first i LVs. The computation of the NE(t_i) and R^2 statistic is well known, details may be obtained in Ref. 15. The negentropy is used to measure the non-Gaussianity of a variable, large values imply significant departures from a Gaussian distribution. The R^2 statistic measures the prediction accuracy of a model with values close to 1 indicating a small prediction error. The results summarized in Table 1 suggest that selecting $\alpha=0.6, \beta=0.4$, and $\gamma=0$ is a good tradeoff between modeling accuracy and non-Gaussianity for the extracted score variables. It is interesting to note that the negentropy values for the first four extracted components are 0.2450, 0.5967, 0.4166, and 0.2797. This implies that the NGR algorithm, different from ICA, does not extract components to maximize negentropy only but also balances negentropy with model accuracy, which follows form the inclusion of the mutual information in the NGR objective function.

Figure 1(a) show the four non-Gaussian signals, Figure 1(b) the first six extracted t-score variables using the NGR algorithm, and Figure 1(c) the first six estimated t-scores produced by the PLS algorithm. Comparing these Figures 1a–1c, it follows that the selection of the parameters allows the NGR algorithm to accurately recover the four non-Gaussian source signals. Moreover, the NGR algorithm extracts the first two components as the most important ones for predicting the output variables. In sharp contrast, PLS can only produce an approximation of the first source signal. For process monitoring, the more accurate extraction of the non-Gaussian source signals by the NGR algorithm is useful, as it increases the sensitivity in detecting abnormal behavior, which is demonstrated next.

Simulating a further 1000 samples that describe a bias in the second non-Gaussian source signal

$$s_{2_k}=\sin(0.3k)+0.3 \tag{33}$$

Figures 2–4 summarize the performance of the proposed NGR monitoring scheme, local PLS, and the ICA approach to the 1000 samples generated from Eq. (33), respectively. The application of the bootstrapping procedure[15] suggested, as expected, the retention of $n=2$ set of LV sets. This give rise to the construction of $\zeta_R\in\mathbb{R}^2$ and $\zeta_D\in\mathbb{R}^8$ for a

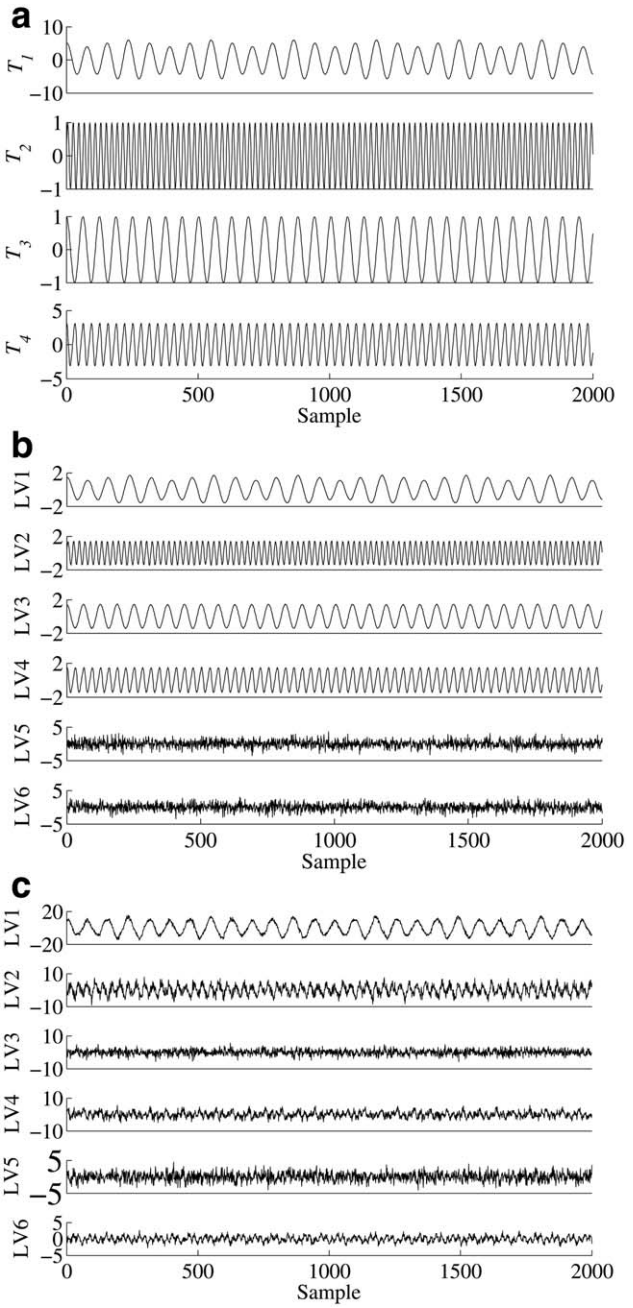

Figure 1. Comparison of signals extracted by NGR and PLS.

selected window length of $k_0 = 100$ and subsequently to the two Hotelling's T^2 statistics $T^2_{\zeta_R}$ and $T^2_{\zeta_D}$. In addition, a third monitoring statistic that relate to the prediction residuals for the output variables can be constructed, which is defined here as T^2_{Res}. Figure 2 highlights that the abnormal condition can be detected by $T^2_{\zeta_R}$, whilst the number of violations of the $T^2_{\zeta_D}$ and T^2_{Res} does not exceed the significance of 0.01. This is expected, as the added offset to the source signal s_2 does not affect the prediction accuracy of the NGR model, nor the discarded eight components from the output variable set.

The application of cross validation suggested the retention of six sets of LVs for the PLS model. Based on this model, Figure 3 shows that local PLS[16] does not clearly detect this abnormal event. More precisely, the number of violations of the control limits that exceed the significance of 0.01 can be

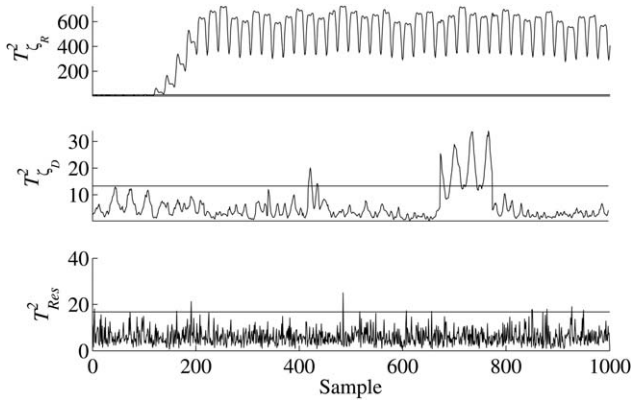

Figure 2. Monitoring results of sensor fault for local NGR.

noted from the conventional T^2_R statistic and the residual T^2_{Res} statistic. Figure 1(b) shows that the second source signal is not clearly extracted by the t-score variables t_2 to t_6. It is, therefore, not surprising that the T^2_R does not strongly respond to the small alteration of the second source signal. In contrast to the NGR monitoring approach, it is also not possible, to clearly identify that this fault condition is injected 100 samples into the dataset.

The application of ICA extracts the four non-Gaussian signals. However, different from the NGR algorithm, the order of the signals is 2, 3, 4, and 1. This implies that the second and third component does, in fact, not contribute to the prediction of the output variables. This emphasizes the main advantage of the NGR algorithm over ICA or independent component regression (ICR), as the extracted components are designed to balance between negentropy and prediction accuracy. This is similar to the well-known problem of PCR, which may discard components that are highly predictive and may include components that are not strongly correlated to the outputs.[22]

Figure 4 shows the performance of the four monitoring statistics, which are defined in the preceding discussion of this section. In fact, only the I^2_d statistic responds to this event, as the number of violations exceed the significance level of 0.01. However, as for the local PLS approach, it is not possible to clearly detect that this fault condition was injected in 100 samples into the dataset. The poor performance of the ICA-based approach follows from the fact that the statistical local

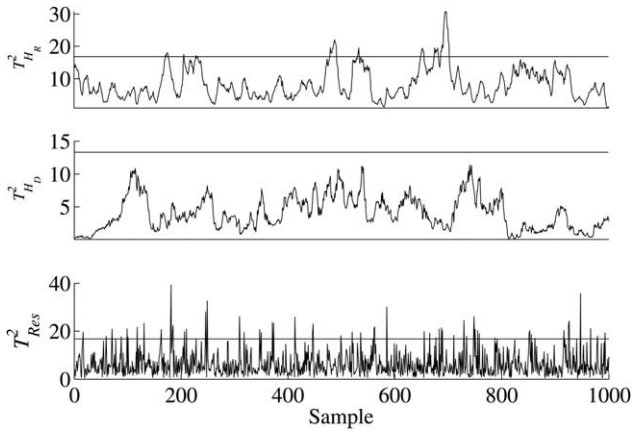

Figure 3. Monitoring results of sensor fault for local PLS.

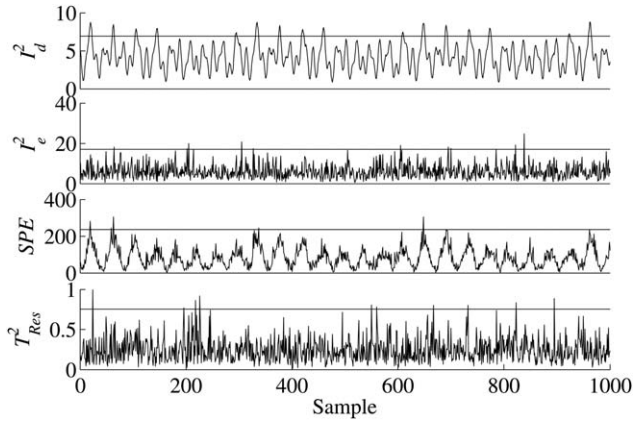

Figure 4. Monitoring results of sensor fault for ICA-based approach.

approach has an increased sensitivity to incipient fault conditions compared to score-based statistics.[17,18]

Application to Fluidized Bed Reactor

This section contrasts the performance of the proposed monitoring strategy with ICA and local PCA on the basis of recorded data from a fluidized bed reactor. This process involves the production of two solvent chemicals by complex chemical reactions. The core elements of this plant are five parallel fluidized bed reactors, each producing the desired chemicals by complex exothermic chemical reactions. This study analyzes recorded data from one reactor that receives a total of five different reactants, which are input variables. The sixth input variables is a stream required to reduce the pressure in an adjacent vaporizer unit which vaporizes two of the input feeds.

Each reactor consists of a large shell and a number of vertically oriented tubes in which the chemical reaction is carried out, supported by fluidized catalyst. A thermocouple is attached to the bottom of each tube to measure the temperature of the fluidized bed. There are a total of 35 tubes and the temperature readings are regarded here as the output variables. The process, therefore, consists of 6 input and 35 output variables.

For the purpose of fault detection, a total of two abnormal operating conditions are studied here. The first condition describes the presence of a sudden an unmeasured disturbance that resulted in abnormal pressure variations in the supply of steam to the vaporizer unit. The second fault condition is an abnormal behavior in a particular tube, which is caused by an increased catalyst density at the bottom of this tube.

A reference set containing 2338 samples that describes normal operation condition is utilized to identify the three monitoring models. Figure 5 shows time-based plots of the recorded variables.

In Figure 5, the first and second figures show the plots of the 35 tube temperatures, with the unit of °C, which are measured using the thermocouple at the bottom of each tube; the third figure plots the six input variables. The first through fifth input variables are the flow rates of 5 input feeds, with the unit of m^3/h. The sixth input variable is the flow rate of the additional steam, with the unit of m^3/h. These plots highlight (1) that the data cannot be assumed to follow a Gaussian distribution and (2) that there are strong collinearities between the output variables. Selecting $\alpha=0.3, \beta=0.7, \gamma=0$ for the NGR algorithm yielded a good tradeoff between model accuracy and the non-Gaussianity of the extracted score variables. Moreover, the retention of 10 sets of LVs provided an accurate prediction model. Figure 6 presents the 10 retained non-Gaussian source signals. By inspection, the non-Gaussian nature of the common cause variation,

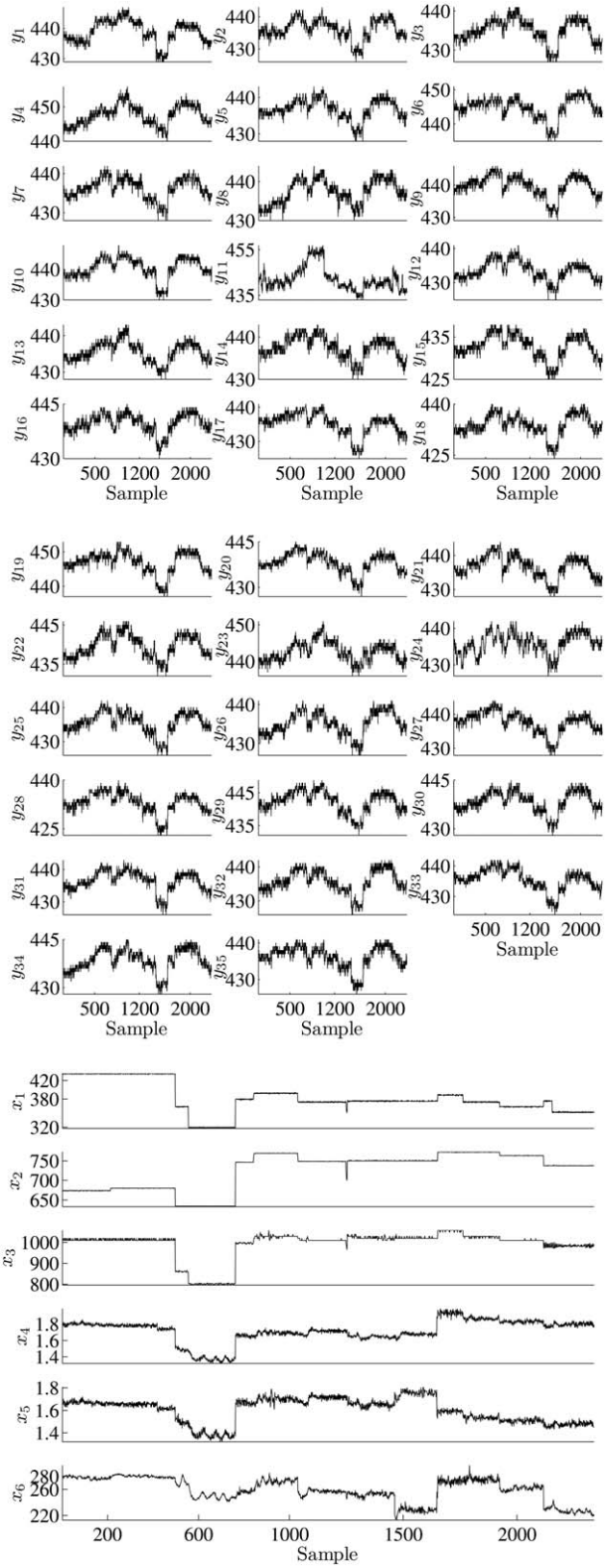

Figure 5. Original data of the fluidized bed reactor.

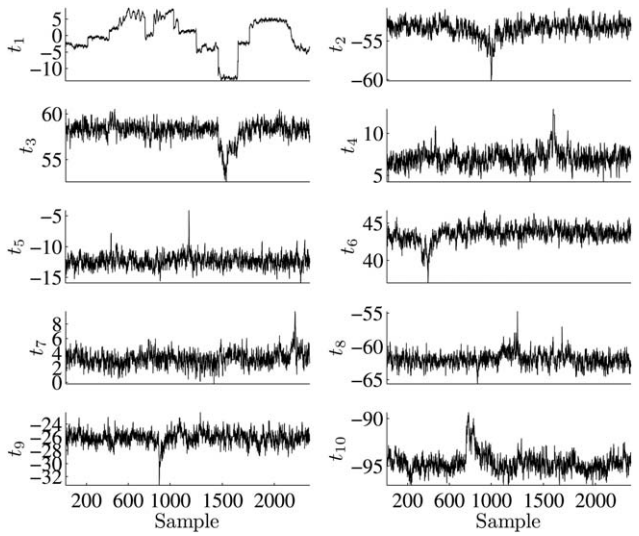

Figure 6. The 10 retained non-Gaussian components by NGR.

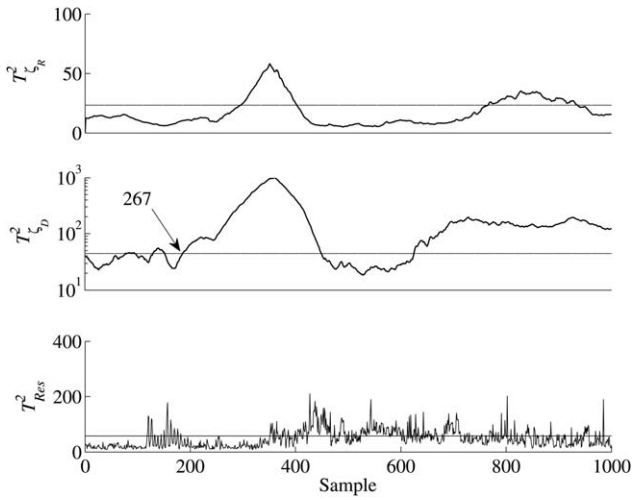

Figure 7. T^2 statistics of local NGR for the first fault condition.

particularly for the first few score variables can be noticed. Table 2 confirms this by summarizing the results of applying the well-known Jarque-Bera test[23] to the 10 retained components.

In Table 2, an h value of 1 implies rejecting the null hypothesis that the score variable follows a Gaussian distribution. Furthermore, the p value of 0.001 is the significance level for this test.

The next step is to determine the improved residuals $\zeta_R \in \mathbb{R}^{10}$ and $\zeta_D \in \mathbb{R}^{10}$ as well as the Hotelling's T^2 statistics $T^2_{\zeta_R}, T^2_{\zeta_D}$, and T^2_{Res} together with their control limits for a significance of 0.01. Figure 7 shows the monitoring statistics of the proposed NGR approach. As there are more outputs than inputs in this case, ICA is applied to the output variables and the extracted score variables are regressed on the input variables. The same approach is also applied to the local PLS algorithm,[16] that is, the outputs are regressed on the input variables. The application of cross validation suggested the retention of 10 LVs and discarding the remaining 25 sets.

The monitoring results of local PLS- and ICA-based approach are shown in Figures 8 and 9, respectively. Figure 9 outlines that the ICA-based approach is unable to detect the sudden disturbance. Comparing Figures 7 and 8, however, the NGR- and PLS-based monitoring statistics that relate to the discarded components and the model residuals. The NGR-based approach, however, is more sensitive to the earlier stages of this abnormal condition. The $T^2_{\zeta_D}$ statistic of the NGR approach shows excessive violation of its control limit after around 150 samples into the dataset and the T^2_{Res} statistic is sensitive after 267 samples into the dataset with sporadic violations of its control limit shortly before. In contrast, PLS-based approach is sensitive to this abnormal event only after 290 samples into the dataset.

The monitoring results for second abnormal process behavior present a similar picture. Figures 10, 11, and 12 show the monitoring results for the NGR-, the PLS- and ICA-based monitoring techniques. Different from the first abnormal case, this time, each technique detected the abnormal temperature measurement in one of the tubes. The sensitive statistics here are again based on the discarded components although the NGR-based approach is more sensitive and detects this event after 335 samples, whilst the PLS-based technique only after 354 samples into the dataset. In contrast, the ICA-based technique requires more than 400 samples to detect the abnormal tube behavior.

The increased sensitivity of the NGR- and PLS-based techniques over the ICA-based method relates to the incorporation of the statistical local approach. For PCA, this is discussed in detail in Refs. 16–18. By virtue of its construction, the better performance of the NGR-based monitoring technique follows from the fact that the NGR algorithm aims to extract components that are a tradeoff between negentropy and model accuracy. If the process variables are highly non-Gaussian, which is the case for the application studies considered here, the NGR algorithm, consequently, extracts components that describe the non-Gaussian source signals more accurately than to PLS.

The ICA technique, although also capable of extracting non-Gaussian components, suffers from the following two problems: (1) the extracted components may not be good regressors and (2) it does not rely on the conceptually more sensitive statistical local approach. Both application studies have, therefore, confirmed that the NGR-based algorithm provides a better modeling capability compared to ICA and produces a more accurate estimation of the source signals if some or all of them are non-Gaussian.

It is also important to note that the utilization of the statistical local approach allows to handle non-Gaussian as well as Gaussian distributed source signals, which follows from the central limit theorem. Hence, the application of a time consuming estimation of a non-Gaussian probability density function can be circumvented.

Table 2. The Results of Jarque-Bera Test on the 10 Retained Components

Components	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}
h value	1	1	1	1	1	1	1	1	1	1
p value	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

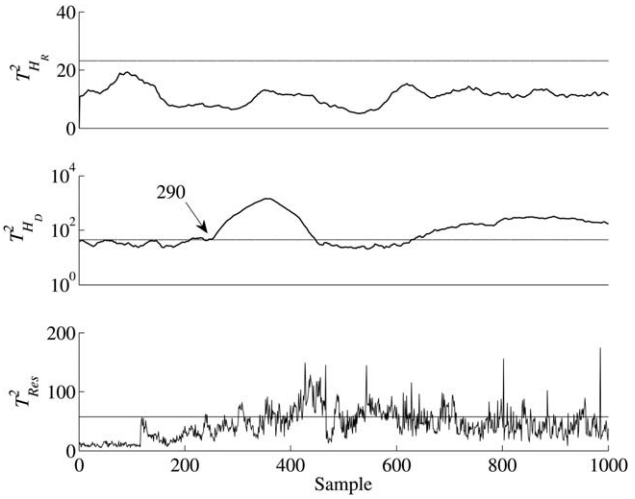

Figure 8. T^2 statistics of local PLS for the first fault condition.

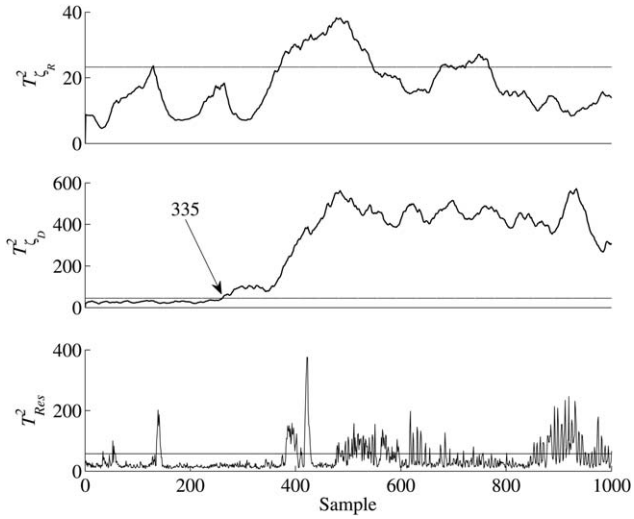

Figure 10. T^2 statistics of local NGR for the second fault condition.

Concluding Summary

This article has developed the recently proposed NGR algorithm into a modeling tool to conduct a regression-based monitoring of complex chemical processing systems that produce non-Gaussian and Gaussian distributed data readings. More precisely, the NGR algorithm is a modeling technique that considers the maximization of negentropy and mutual information between the extracted LVs and the response variables. The extracted LVs are non-Gaussian components that are related to the response variables by the mutual information criterion. Similar to PLS, the proposed NGR algorithm models that may provide a better modeling accuracy compared to PCR. Moreover, different to PLS, the NGR algorithm produces a more accurate estimation of non-Gaussian components. Different from the original NGR algorithm, the proposed monitoring scheme utilizes an algorithm that relies on a different set of orthogonality constraints to allow the number of components to be determined to be maximum of the number of input and output variables.

To construct monitoring statistics that, irrespective of the distribution function of the extracted LVs, follow a Gaussian distribution, the article has proposed the use of the statistical local approach. This allows online process monitoring using standard multivariate monitoring statistics. The developed monitoring statistics are sensitive to process faults as the NGR algorithm extracts score variables that provide a good approximation of non-Gaussian source signals. Through the incorporation of mutual information in its objective function, the NGR algorithm determines score variables that are good predictors for the system response variables. Faults occurring in the source signals that influence both the input and output variables, which is common in process industry, can, therefore, be more easily detected. To diagnose anomalous process behavior, a fault reconstruction approach has been discussed that relates to existing work.

The article has then contrasted the proposed NGR-based monitoring scheme with existing work that includes the use of ICA and work that combines PLS with the application of the statistical local approach. This comparison has been based on a simulation example and recorded data from a chemical reaction process and demonstrated the proposed NGR-based technique can outperform both competitive techniques. A detailed analysis has yielded the NGR model can extract score

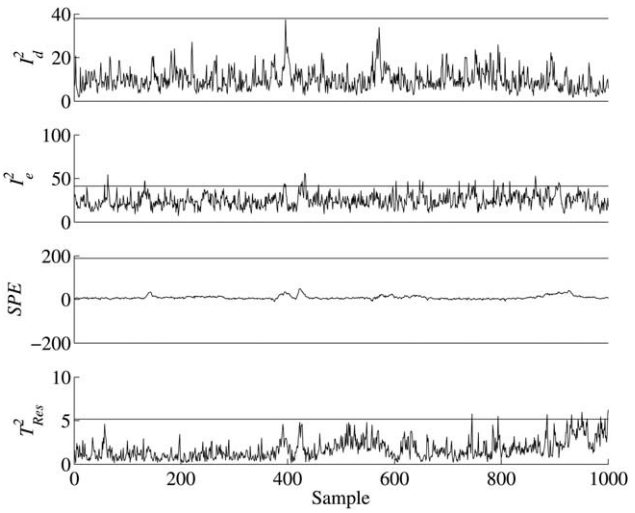

Figure 9. Monitoring statistics of ICA-based approach for the first fault condition.

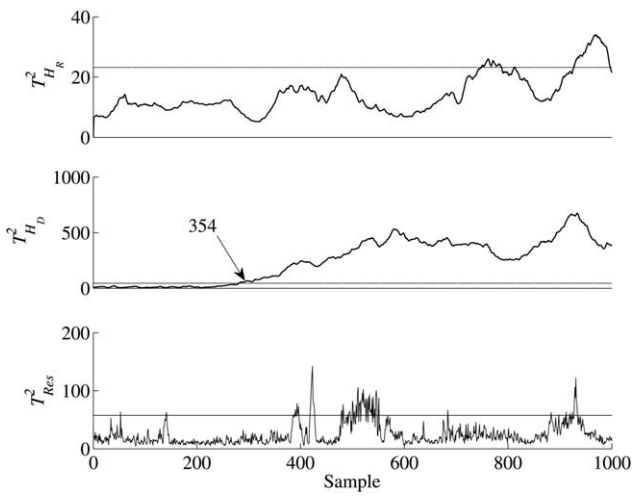

Figure 11. T^2 statistics of local PLS for the second fault condition.

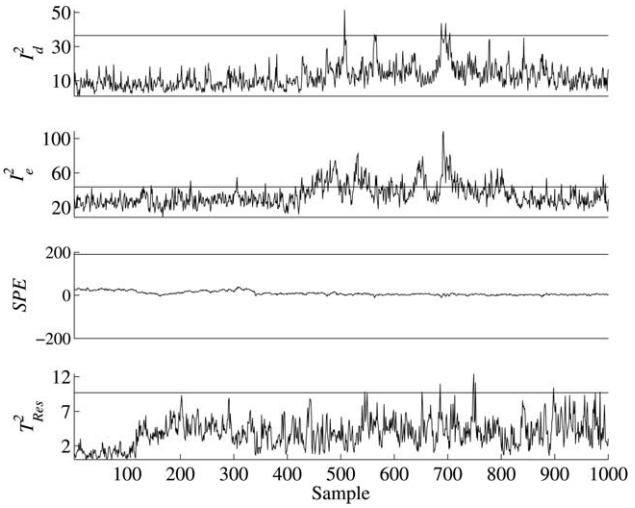

Figure 12. Monitoring statistics of ICA-based approach for the second fault condition.

variables that describe the source signals more accurately than PLS and it can produce a more accurate regression model than ICA/ICR. The proposed approach allows to construct monitoring statistics that are more sensitive than the contrasted methods.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No. 61203088, 61134007, 61374121), Science and Technology Project of Zhejiang Province (Grant No. 2012C01026-1), Natural Science Foundation of Zhejiang Province (Grant No. LQ12F03015), and the 111 Project (Grant No. B07031).

Literature Cited

1. 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.
2. Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis part ii: qualitative models and search strategies. *Comput Chem Eng*. 2003;27:313–326.
3. Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis part iii: process history based methods. *Comput Chem Eng*. 2003;27:327–346.
4. Ding SX. Model-Based Fault Diagnosis Techniques. Berlin: Springer, 2008.
5. Isermann R. Model-based fault-detection and diagnosis-status and applications. *Annu Rev Control*. 2005;29:71–85.
6. MacGregor JF, Marlin TE, Kresta JV, Skagerberg B. Multivariate statistical methods in process analysis and control. Proceedings of the Fourth International Conference on Chemical Process Control. AIChE, Padre Island, TX, 1991:79–99.
7. Yoon S, MacGregor JF. Statistical and causal model-based approaches to fault detection and isolation. *AIChE J*. 2000;46(9): 1813–1824.
8. Tamura M, Tsujita S. A study on the number of principal components and sensitivity of fault detection using pca. *Comput Chem Eng*. 2007;31:1035–1046.
9. Zhao CH, Wang FL, Lu NY, Jia MX. Stage-based soft-transition multiple PCA modeling and online monitoring strategy for batch processes. *J Process Control*. 2007;17:728–741.
10. Zhou DG, Li G, Qin SJ. Total projection to latent structures for process monitoring. *AIChE J*. 2009;55:168–178.
11. Li G, Qin SJ, Zhou DH. Geometric properties of partial least squares for process monitoring. *Automatica*. 2010;46:204–210.
12. Lee JM, Yoo CK, Lee IB. Statistical process monitoring with independent component analysis. *J Process Control*. 2004;14:467–485.
13. Lee JM, Qin SJ, Lee IB. Fault detection and diagnosis based on modified independent component analysis. *AIChE J*. 2006;52:3501–3514.
14. Ge ZQ, Xie L, Kruger U, Song ZH. Local ica for multivariate statistical fault diagnosis in systems with unknown signal and error distributions. *AIChE J*. 2012;58:2357–2371.
15. Zeng JS, Xie L, Kruger U, Gao CH. A non-Gaussian regression algorithm based on mutual information maximization. *Chemometr Intell Lab Syst*. 2012;111:1–19.
16. Kruger U, Dimitriadis G. Diagnosis of process faults in chemical systems using a local partial least squares approach. *AIChE J*. 2008; 54:2581–2598.
17. Kruger U, Kumar V, Littler T. Improved principal component modelling using the local approach. *Automatica*. 2007;43:1532–1542.
18. Kruger U, Xie L. Statistical Monitoring of Complex Multivariate Processes. Chichester: Wiley, 2012.
19. Basseville M. On-board component fault detection and isolation using the statistical local approach. *Automatica*. 1998;34:309–326.
20. Tracey ND, Young JC, Mason RL. Multivariate control charts for individual observations. *J Qual Technol*. 1992;24:88–95.
21. Box GEP. Some theorems on quadratic forms applied in the study of analysis of variance problems: effects of inequality of variance in one-way classification. *Ann Math Stat*. 1954;15:290–302.
22. Jackson JE. A Users Guide to Principal Components. New York: Wiley, 2003.
23. Jarque CM, Bera AK. A test for normality of observations and regression residuals. *Int Stat Rev*. 1987;55(2):1–10.
24. Hyvärinen A, Oja E. Fast fixed-point algorithm for independent component analysis. *Neural Comput*. 1997;9:1483–1492.
25. Barndorff-Nielsen OE, Cox DR. Inference and Asymptotics. London: Chapman and Hall, 1989.
26. Hulle MMV. Edgeworth approximation of multivariate differential entropy. *Neural Comput*. 2005;17:1903–1910.

Appendix A: Derivation of Eq. 2

Different from PLS, the weight vectors for the NGR algorithm produces $t_i=\mathbf{w}^T\mathbf{x}$ and $u_i=\mathbf{v}^T\mathbf{y}$ by using an objective function that includes the following negentropy criteria

$$I_{1_i}=H\{\upsilon\}-H\{t_i\} \tag{A1a}$$

$$I_{3_i}=H\{v\}-H\{u_i\} \tag{A1b}$$

Here, υ and v are a Gaussian distributed variables that have the same variance as t_i and u_i, respectively, and $H\{\upsilon\}$, $H\{t_i\}$, $H\{v\}$, and $H\{u_i\}$are the entropies of υ, t_i, v, and u_i, respectively. The entropy of a random variable z, which has a probability density function $p(z)$ is defined as follows

$$H\{z\}=-\int p(z)\log p(z)dz \tag{A2}$$

Ref. 24 showed that the computation of entropy can be circumvented by approximating Eqs. A1a and A1b as follows

$$I_{i_1} \approx [E\{G(t_i)\}-E\{G(\upsilon)\}]^2 \tag{A3a}$$

$$I_{i_3} \approx [E\{G(u_i)\}-E\{G(v)\}]^2. \tag{A3b}$$

There are different selections for $G(\cdot)$[24] of which this article uses

$$G(z)=\frac{1}{a_1}\log\cosh(a_1 z), \tag{A4}$$

where$1 \le a_1 \le 2$. Maximizing Eqs. A3a and A3b allows extracting non-Gaussian from $\mathbf{x}$ and $\mathbf{y}$, **that is**, $t_i=\mathbf{w}_i^T\mathbf{x}$ and $u_i=\mathbf{v}_i^T\mathbf{y}$. Next, the objective function of the NGR algorithm also includes mutual information to maximize the dependence between the t_i, and u_i.

$$I_{i_2}(t_i,u_i)=H\{t_i\}+H\{u_i\}-H\{u_i,t_i\} \tag{A5}$$

where $H\{u_i,t_i\}$ is the joint entropy between t_i and u_i**.** The complete NGR objective function presents a trade-off between I_{i1}, I_{i_2}, and I_{i_3}, which is achieved by introducing the weight parameters $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$

$$\begin{pmatrix}\widehat{\mathbf{w}}_i\\ \\ \hat{\mathbf{v}}_i\end{pmatrix}=\arg\max_{\mathbf{w}_i\mathbf{v}_i}\tilde{\alpha}(H\{\upsilon\}-H\{t_i\})$$
$$+\tilde{\beta}\left(H\{t_i\}+H\{u_i\}-H\left\{t_i,u_i^T\right\}\right)+\tilde{\gamma}(H\{v\}-H\{u_i\}) \tag{A6}$$
$$\begin{pmatrix}\widehat{\mathbf{w}}_i\\ \\ \hat{\mathbf{v}}_i\end{pmatrix}=\arg\max_{\mathbf{w}_i\mathbf{v}_i}\left(\tilde{\alpha}-\tilde{\beta}\right)(H\{\upsilon\}-H\{t_i\})$$
$$-\tilde{\beta}H\{t_i,u_i\}\left(\tilde{\gamma}-\tilde{\beta}\right)(H\{v\}-H\{u_i\})+\tilde{\beta}H\{\upsilon\}+\tilde{\gamma}H\{v\}$$

Next defining $\alpha=\tilde{\alpha}-\tilde{\beta}, \beta=\tilde{\beta}$ and $\gamma=\tilde{\gamma}-\tilde{\beta}$ and omitting the constant terms $H\{\upsilon\}$ and $H\{v\}$ allows reducing Eq. A6 to become

$$\begin{pmatrix}\widehat{\mathbf{w}}_i\\ \\ \hat{\mathbf{v}}_i\end{pmatrix}=\arg\max_{\mathbf{w}_i\mathbf{v}_i}\alpha[E\{G(t_i)\}-E\{G(\upsilon)\}]^2 \tag{A7}$$
$$-\beta H\{t_i,u_i\}+\gamma[E\{G(u_i)\}-E\{G(v)\}]^2$$

Finally, the joint entropy $H\{t_i,u_i\}$ can be approximated by the Edgeworth expansion[25,26]

$$H\{t_i,u_i\}\approx-\frac{1}{12}\left(E\{t_i\}^3\right)^2-\frac{1}{12}E\left\{(u_i)^3\right\}^2-\frac{1}{4}\left(E\{t_i\}^2u_i\right)^2$$
$$-\frac{1}{4}E\left\{t_i(u_i)^2\right\}^2+\frac{1}{2}\log\left(1-E(t_iu_i)^2\right)+\log(2\pi e) \tag{A8}$$

Combining Eqs. A7 and A8 and omitting the constant log (27re) yields Eq. 2.

Appendix B: Assumptions for Constructing Primary Residuals

With respect to the parameter vectors describing normal and abnormal process behavior, that is θ_0 and θ, the hypothesis that the process is in-statistical or out-of-statistical-control relies on the following test

$$H_0:\theta=\theta_0\quad H_1:\theta=\theta_0+\frac{\Delta\theta}{\sqrt{K}}. \tag{B1}$$

In Eq. B1, H_0 and H_1 are the null hypothesis representing in-statistical-control behavior and the alternative hypothesis corresponding to out-of-statistical-control behavior. Using the statistical local approach, the null hypothesis is tested using the primary residuals $\phi(\theta_0,\mathbf{z})$, where $\mathbf{z}$ is a data vector. The constructing of a primary residual function $\phi(\cdot)$ must meet the following four assumptions

1. $E\{\phi(\theta_0,\mathbf{z})\}=0$ if $\theta=\theta_0$;
2. $E\{\phi(\theta_0,\mathbf{z})\}\neq 0$ if $\theta\neq\theta_0$ but θ is in the neighborhood of θ_0;
3. $\phi(\theta_0,\mathbf{z})$ is differentiate with respect to θ; and
4. $\phi(\theta_0,\mathbf{z})$ exists in the vicinity of θ_0.

Manuscript received May 2, 2012, revision received July 1, 2013, and final revision received Sept. 3, 2013.