Chemometrics and Intelligent Laboratory Systems xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space

José L. Godoy [a,b,*], Jorge R. Vega [a,c], Jacinto L. Marchetti [a]

[a] INTEC (CONICET and Universidad Nacional del Litoral), Güemes 3450, (3000) Santa Fe, Argentina
[b] FRP-UTN (Facultad Regional Paraná – Universidad Tecnológica Nacional), Almafuerte 1033 - (3100) Paraná, Argentina
[c] FRSF-UTN (Facultad Regional Santa Fe – Universidad Tecnológica Nacional), Lavaisse 610, (3000) Santa Fe, Argentina

ARTICLE INFO

Article history:
Received 2 November 2012
Received in revised form 5 July 2013
Accepted 12 July 2013
Available online xxxx

Keywords:
PLS
Process monitoring
Fault detection
Fault diagnosis
Fault isolation
Contribution analysis

ABSTRACT

A new statistical monitoring technique based on partial least squares (PLS) is proposed for fault detection and diagnosis in multivariate processes that exhibit collinear measurements. A typical PLS regression (PLSR) modeling strategy is first extended by adding the projections of the model outputs to the latent space. Then, a PLS-decomposition of the measurements into four terms that belongs to four different subspaces is derived. In order to online monitor the PLS-projections in each subspace, new specific statistics with non-overlapped domains are combined into a single index able to detect process anomalies. To reach a complete diagnosis, a further decomposition of each statistic was defined as a sum of variable contributions. By adequately processing all this information, the technique is able to: i) detect an anomaly through a single combined index, ii) diagnose the anomaly class from the observed pattern of the four component statistics with respect to their respective confidence intervals, and iii) identify the disturbed variables based on the analysis of the main variable contributions to each of the four subspaces. The effectiveness observed in the simulated examples suggests the potential application of this technique to real production systems.

1. Introduction

Statistical process monitoring (SPM) applies multivariate statistics and machine learning methods to product quality and production control, fault detection and diagnosis, and estimation of fault/fault-free magnitudes. The SPM approaches based on historical operating data are useful when applied to process having a large number of measured variables and when causal models are unavailable [1]. In particular, latent variable (LV) models are more adequate than causal models to deal with monitoring tasks. Along the last two decades, several strategies based on LV models have been proposed for multivariate SPM [2–7]. When a process operates under normal conditions (i.e., when only 'common-cause' variations are present), then the correlation structure underlying the measured data can be adequately described through LV models. Monitoring techniques based on LV are useful for detecting and diagnosing abnormal behaviors in complex multivariate processes; and therefore they are of great interest for their use in industrial applications [8]. For instance, partial least squares (PLS) is often applied in Process Analytical Technology (PAT) projects, where most applications have successfully been used for monitoring biotechnological processes of the pharmaceutical industry [9].

Several fault detection indexes based on LV models can be used to alert on the presence of possible anomalies during the process operation. An alarm signal typically appears when an index exceeds its predefined control limit. Once a fault is detected, the fault diagnosis is made by analyzing the contributions of each measured variable to the specific index that caused the alarm. An essential requirement for fault diagnosis is to avoid misdiagnosis. For example, in the case of a fault in a sensor or in an actuator, only one variable associated to the faulty device should ideally contribute to the index with a high positive value. Contribution charts basically calculate the contributions of the variables in a fault situation and select the variables with large contributions as indicators of the probable cause of the fault [10]. Thus, an analysis of well-defined contributions should have the following desirable properties: i) in absence of faults, all contributions would exhibit similar low average values, thus determining a normal baseline level; ii) in the presence of a single fault (i.e., a fault that is due to only one variable), the contribution corresponding to the faulty variable should be large, and iii) in the presence of a multiple fault (e.g., multiple sensor failures), the abnormally large contribution of each faulty variable must be greater than the rest [11].

In medium and large scale processes (such as chemical manufacturing plants, food industries, etc.), there are a large number of controlled, manipulated, and process variables. When a multivariate process operates under normal conditions, then the measurement space spanned by the measured variables exhibits a non-full or deficient rank. In such a case, a model of the correlation structure among all

* Corresponding author at: INTEC (CONICET and Universidad Nacional del Litoral), Güemes 3450, (3000) Santa Fe, Argentina. Tel.: +54 342 4559175; fax: +54 342 4550944.
E-mail address: jlgodoy@santafe-conicet.gov.ar (J.L. Godoy).

0169-7439/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.chemolab.2013.07.006

measured variables is required for reaching a proper process monitoring. Methods of projection to LVs can be used to transform noisy collinear data into well-conditioned and reduced-dimension data that preserve all the useful information about the process, thus allowing the capture of the 'common-cause' subspace [7]. These data-driven methods typically use historic data (collected during normal operating conditions) to develop a LV model able to effectively explain the 'common-cause' variability. Several of these monitoring techniques (such as PCA: principal component analysis, ICA: independent component analysis, Kernel PCA, and Kernel ICA) treat the data without differentiating outputs from inputs [12–15]. In contrast, a PLS regression (PLSR) model is closer to the intrinsic structure of multi-input multi-output process [16,17], because it allows the elimination of some undesired input variables from the original pre-sets (e.g., those interfering the regression model) [18,19].

In the last years, several tools based on PLSR have been proposed for monitoring industrial processes [18,20–23]. Simultaneously, some efforts have also been aimed at improving the fundamentals of PLSR [24,25]. The development of surveillance PLSR tools able to detect anomalies or poor process performances is undoubtedly an active field of current research.

This article aims at developing an online data processing technique useful for helping process engineers to reach a swift identification of the origin of a fault when the monitoring system trigger an alarm signal. The paper is organized as follows. Section 2 describes the fundamentals of an extended PLSR modeling and their geometric properties, which was needed in order to obtain a new online PLS-decomposition of the measurements with associated statistics and contribution analysis. Section 3 presents the main statistical monitoring tools and describes their roles in the measurement PLS-decomposition. Section 4 discusses some simulation tests based on both static and dynamic systems. Finally, the main conclusions are presented in Section 5.

2. Extended PLSR modeling

The PLSR model developed here is calculated by simultaneously deflating the data matrices with the classical PLS–NIPALS algorithm [26]. This procedure gives better results for multivariate prediction and process monitoring than other alternative PLS algorithms [25]. Besides, the simultaneous deflation on both data matrices allows the detection of predictor variables playing an interfering effect [18].

Consider a process with m measured input variables plus p measured output variables. Assume that N measurements of each variable are collected while the process is operating under normal conditions. In order to build a model, the N multivariate measurements are arranged into a predictor matrix $\mathbf{X} = [\mathbf{x}_1 \ldots \mathbf{x}_N]'$ $(N \times m)$ consisting of N samples of m variables per sample, and a response matrix $\mathbf{Y} = [\mathbf{y}_1 \ldots \mathbf{y}_N]'$ $(N \times p)$, with N samples of p variables per sample. Then, PLSR can be used to find a regression model between the measurement vectors $\mathbf{x} = [x_1 \ldots x_m]'$ and $\mathbf{y} = [y_1 \ldots y_p]'$, even when their correlation matrices ($\mathbf{R_x}$ and $\mathbf{R_y}$) are both positive semi-definite (i.e., $\mathbf{X}$ and $\mathbf{Y}$ have collinear variables). The method produces a projection of $\mathbf{X}$ and $\mathbf{Y}$ into low-dimension spaces defined by the number A of LV which are then regressed. At each iteration a, the implicit objective of the PLS–NIPALS algorithm (see Table 1) is to find a solution to the following optimization problem [25]:

$$\max_{\mathbf{w}_a,\mathbf{q}_a} (\mathbf{w}'_a \mathbf{X}'_a \mathbf{Y}_a \mathbf{q}_a) \quad \text{subject to:} \quad \|\mathbf{w}_a\| = 1, \|\mathbf{q}_a\| = 1. \tag{1}$$

In Eq. (1), $\mathbf{X}_a$ and $\mathbf{Y}_a$ stand for the a-times deflated versions of $\mathbf{X}_1 = \mathbf{X}$ and $\mathbf{Y}_1 = \mathbf{Y}$ respectively. The number A of deflations to be made is determined by checking the residual matrices until the leftover information can no longer be modeled [18]. In this way, the PLSR modeling strategy produces an external and an internal model. The external

Table 1
(**X**, **Y**)-deflated PLS-NIPALS algorithm. Outputs: **P**, **Q**, **B**, **T**, and **U**.

Center the columns of **X**, **Y** to zero mean and scale them to unit variance.
Set $a = 1$, $\mathbf{X}_1 = \mathbf{X}$, $\mathbf{Y}_1 = \mathbf{Y}$. (Initialization)
1. Set $\mathbf{t}_a^0$ equal to the maximum-variance column of $\mathbf{X}_a$.
   Set $\mathbf{u}_a$ equal to the maximum-variance column of $\mathbf{Y}_a$.
2. $\mathbf{w}_a = \mathbf{X}'_a\mathbf{u}_a / \|\mathbf{X}'_a\mathbf{u}_a\|$, ($\|\mathbf{w}_a\| = 1$).
3. $\mathbf{t}_a = \mathbf{X}'_a\mathbf{w}_a$.
4. $\mathbf{q}_a = \mathbf{Y}'_a\mathbf{t}_a / \|\mathbf{Y}'_a\mathbf{t}_a\|$, ($\|\mathbf{q}_a\| = 1$).
5. $\mathbf{u}_a = \mathbf{Y}'_a\mathbf{q}_a$.
   If $\|\mathbf{t}_a^0 - \mathbf{t}_a\| < \varepsilon$, go to step 6, else set $\mathbf{t}_a^0 = \mathbf{t}_a$ and return to step 2.
6. $\mathbf{p}_a^* = \mathbf{X}'_a\mathbf{t}_a / (\mathbf{t}'_a\mathbf{t}_a)$, $\mathbf{p}_a = \mathbf{p}_a^*/\|\mathbf{p}_a^*\|$, ($\|\mathbf{p}_a\| = 1$). #
7. $\mathbf{t}_a = \mathbf{t}_a\|\mathbf{p}_a^*\|$, $\mathbf{w}_a = \mathbf{w}_a\|\mathbf{p}_a^*\|$, ($\|\mathbf{w}_a\| \neq 1$). #
8. $b_a = \mathbf{u}'_a\mathbf{t}_a / (\mathbf{t}'_a\mathbf{t}_a)$, (inner regression).
9. $\mathbf{X}_{a+1} = \mathbf{X}_a - \mathbf{t}_a\mathbf{p}'_a$, $\mathbf{Y}_{a+1} = \mathbf{Y}_a - b_a\mathbf{t}_a\mathbf{q}'_a$, (deflations).
10. $\mathbf{P} \leftarrow \mathbf{p}_a$, $\mathbf{Q} \leftarrow \mathbf{q}_a$, $\mathbf{B} \leftarrow b_a$, $\mathbf{T} \leftarrow \mathbf{t}_a$, $\mathbf{U} \leftarrow \mathbf{u}_a$, (save data to matrices).
Set $a = a + 1$. If $a > A$, then stop; else return to step 1.

: Compensation of scales by previous normalization.

model decomposes $\mathbf{X}$ and $\mathbf{Y}$ into score vectors ($\mathbf{t}_a$ and $\mathbf{u}_a$), loading vectors ($\mathbf{p}_a$ and $\mathbf{q}_a$), and residual error matrices ($\widetilde{\mathbf{X}}$ and $\widetilde{\mathbf{Y}}_2$), as follows:

$$\mathbf{X} = \sum_{a=1}^{A} \mathbf{t}_a\mathbf{p}'_a + \widetilde{\mathbf{X}} = \mathbf{T}\mathbf{P}' + \widetilde{\mathbf{X}}, \qquad \mathbf{P} = [\mathbf{p}_1 \ldots \mathbf{p}_A], \tag{2}$$

$$\mathbf{Y} = \sum_{a=1}^{A} \mathbf{u}_a\mathbf{q}'_a + \widetilde{\mathbf{Y}}_2 = \mathbf{U}\mathbf{Q}' + \widetilde{\mathbf{Y}}_2, \quad \mathbf{Q} = [\mathbf{q}_1 \ldots \mathbf{q}_A], \tag{3}$$

where $\mathbf{T} = [\mathbf{t}_1 \ldots \mathbf{t}_A]$ and $\mathbf{U} = [\mathbf{u}_1 \ldots \mathbf{u}_A]$ are orthogonal by columns. In the internal model, $\mathbf{t}_a$ is linearly regressed against the y-score vector $\mathbf{u}_a$, i.e.,

$$\mathbf{U} = \mathbf{T}\mathbf{B} + \widetilde{\mathbf{U}}, \mathbf{B} = diag(b_1 \ldots b_A) \tag{4}$$

where $b_1 \ldots b_A$ are the regression coefficients determined by minimization of the residuals $\widetilde{\mathbf{U}}$.

Call $\mathbf{R}$ and $\mathbf{S}$ the pseudo-inverses of $\mathbf{P}'$ and $\mathbf{Q}'$, respectively (i.e., $\mathbf{P}'\mathbf{R} = \mathbf{I}$ and $\mathbf{Q}'\mathbf{S} = \mathbf{I}$). Then, for new $\mathbf{X}$ and $\mathbf{Y}$, the predictions of $\mathbf{T}$ and $\mathbf{U}$ are directly obtained from Eqs. (2) and (3), as:

$$\mathbf{T} = \mathbf{X}\mathbf{R}, \quad \mathbf{R} = [\mathbf{r}_1 \ldots \mathbf{r}_A], \tag{5}$$

$$\mathbf{U} = \mathbf{Y}\mathbf{S}, \quad \mathbf{S} = [\mathbf{s}_1 \ldots \mathbf{s}_A]. \tag{6}$$

Since the row space of $\widetilde{\mathbf{X}}$ belongs to the null-space of $\mathbf{R}$, then $\widetilde{\mathbf{X}}\mathbf{R} = 0$ [27]. Similarly for $\widetilde{\mathbf{Y}}_2$, that belongs to the null space $\mathbf{S}$, and consequently $\widetilde{\mathbf{Y}}_2\mathbf{S} = 0$. Hence, by combining Eqs. (3)–(5), the following prediction model is obtained:

$$\mathbf{Y} = \mathbf{X}\mathbf{R}\mathbf{B}\mathbf{Q}' + \widetilde{\mathbf{U}}\mathbf{Q}' + \widetilde{\mathbf{Y}}_2 = \widehat{\mathbf{Y}} + \widetilde{\mathbf{Y}}_1 + \widetilde{\mathbf{Y}}_2, \tag{7}$$

where $\widetilde{\mathbf{Y}}_2 = \mathbf{Y} - \mathbf{Y}\mathbf{S}\mathbf{Q}'$ and $\widetilde{\mathbf{Y}}_1 = \mathbf{Y}\mathbf{S}\mathbf{Q}' - \widehat{\mathbf{Y}}$ are the projection and transformation error matrices, respectively. Hence, we have extended the PLSR model description by adding the projection of $\mathbf{Y}$ to $\mathbf{U}$ (Eq. (6)), which allows the decomposition of the prediction error $\widetilde{\mathbf{Y}}$ into two terms: $\widetilde{\mathbf{Y}}_1$ and $\widetilde{\mathbf{Y}}_2$ (Eq. (7)).

2.1. PLS-decomposition of the input and output spaces

After synthesizing an "in-control" PLSR model, the measurement vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^p$ can be decomposed as described below.

Lemma 1. *Call* $\mathbf{\Pi}_{\mathbf{P}|\mathbf{R}^\perp}$ ($\mathbf{\Pi}_{\mathbf{Q}|\mathbf{S}^\perp}$) *the projector on the model subspace* $S_{MX} \equiv Span\{\mathbf{P}\} \subseteq \mathbb{R}^m$ ($S_{MY} \equiv Span\{\mathbf{Q}\} \subseteq \mathbb{R}^p$), *along the residual subspace* $S_{RX} \equiv Span\{\mathbf{R}\}^\perp$ ($S_{RY} \equiv Span\{\mathbf{S}\}^\perp$). *Then:*

$$\mathbf{\Pi}_{\mathbf{P}|\mathbf{R}^\perp} = \mathbf{P}\mathbf{R}', \quad \mathbf{\Pi}_{\mathbf{R}^\perp|\mathbf{P}} = \mathbf{I} - \mathbf{P}\mathbf{R}', \tag{8}$$

$$\mathbf{\Pi}_{\mathbf{Q}|\mathbf{S}^\perp} = \mathbf{Q}\mathbf{S}', \quad \mathbf{\Pi}_{\mathbf{S}^\perp|\mathbf{Q}} = \mathbf{I} - \mathbf{Q}\mathbf{S}', \tag{9}$$

where $^{\perp}$ *denotes the orthogonal complement of the subspace* (*see Proof 1 in Appendix A*).

From Lemma 1, we propose the following theorem on the PLS-decomposition.

Theorem 1. *The predictor and response vectors can be decomposed (by PLSR) in complementary oblique projections as*:

$$\begin{aligned}
&\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}} \in \mathbb{R}^m,\\
&\hat{\mathbf{x}} = \mathbf{P}\mathbf{R}'\mathbf{x} \in S_{MX} \equiv Span\{\mathbf{P}\},\\
&\tilde{\mathbf{x}} = \left(\mathbf{I}-\mathbf{P}\mathbf{R}'\right)\mathbf{x} \in S_{RX} \equiv Span\{\mathbf{R}\}^{\perp}
\end{aligned} \tag{10}$$

$$\begin{aligned}
&\mathbf{y} = \hat{\mathbf{y}}^{*} + \tilde{\mathbf{y}}_2 \in \mathbb{R}^p,\\
&\hat{\mathbf{y}}^{*} = \mathbf{Q}\mathbf{S}'\mathbf{y} \in S_{MY} \equiv Span\{\mathbf{Q}\},\\
&\tilde{\mathbf{y}}_2 = (\mathbf{I}-\mathbf{Q}\mathbf{S}')\mathbf{y} \in S_{RY} \equiv Span\{\mathbf{S}\}^{\perp},
\end{aligned} \tag{11}$$

with the projections on the model subspaces S_{MX} *and* S_{MY} *interrelated according to*:

$$\begin{aligned}
&\hat{\mathbf{y}}^{*} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}_1 \in S_{MY},\\
&\hat{\mathbf{y}} = \mathbf{Q}\mathbf{B}\mathbf{R}'\hat{\mathbf{x}} \in S_{MY},\\
&\tilde{\mathbf{y}}_1 = \mathbf{Q}\mathbf{S}'\mathbf{y} - \mathbf{Q}\mathbf{B}\mathbf{R}'\mathbf{x} \in S_{MY},
\end{aligned} \tag{12}$$

where $\tilde{\mathbf{y}}_1$ *is the error of the linear transformation* $\hat{\mathbf{x}} \rightarrow \hat{\mathbf{y}}^{*}$, *and* $\hat{\mathbf{y}}$ *denotes the predictable part of* $\hat{\mathbf{y}}^{*}$ *from* $\hat{\mathbf{x}}$ (*see Proof 2 in Appendix A*).

The minimum angle θ_X between the two complementary subspaces S_{MX} and S_{RX} is defined as the number $0 \le \theta_X \le \pi/2$ that satisfies: $\cos\theta_X = \max(\hat{\mathbf{x}}'\tilde{\mathbf{x}}/\|\hat{\mathbf{x}}\|\|\tilde{\mathbf{x}}\|)$. A similar definition applies for the minimum angle θ_Y between S_{MY} and S_{RY}. In turn, θ_X (θ_Y) can be estimated through [28]:

$$\sin\theta_X = 1/\|\mathbf{P}\mathbf{R}'\|_2, \quad \sin\theta_Y = 1/\|\mathbf{Q}\mathbf{S}'\|_2 \tag{13}$$

The coordinates of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}^{*}$ on the model subspaces S_{MX} and S_{MY}, are given by:

$$\mathbf{t} = \mathbf{R}'\mathbf{x}, \quad \mathbf{u} = \mathbf{S}'\mathbf{y}, \tag{14a}$$

where $\mathbf{t} = [t_1 \cdots t_A]'$ and $\mathbf{u} = [u_1 \cdots u_A]'$. The latent vectors are related by:

$$\mathbf{u} = \mathbf{B}\mathbf{t} + \tilde{\mathbf{u}}, \tag{14b}$$

where $\hat{\mathbf{u}} = \mathbf{B}\mathbf{t} = \mathbf{S}'\hat{\mathbf{y}}$ is the prediction of $\mathbf{u}$. Then, from Eqs. (10)–(12), (14a) it results:

$$\hat{\mathbf{x}} = \mathbf{P}\mathbf{t}, \quad \hat{\mathbf{y}} = \mathbf{Q}\hat{\mathbf{u}}. \tag{14c}$$

Therefore, the corresponding correlation matrices are related as follows:

$$\mathbf{\Lambda} = \mathbf{R}'\mathbf{R}_{\mathbf{x}}\mathbf{R} = (N-1)^{-1}\mathbf{T}'\mathbf{T} = diag(\lambda_1 \ldots \lambda_A), \tag{15a}$$

$$\mathbf{\Delta} = (N-1)^{-1}\hat{\mathbf{U}}'\hat{\mathbf{U}} = \mathbf{B}\mathbf{\Lambda}\mathbf{B} = diag\left(\hat{\delta}_1 \ldots \hat{\delta}_A\right), \tag{15b}$$

$$\mathbf{R}_{\hat{\mathbf{x}}} = (N-1)^{-1}\hat{\mathbf{X}}'\hat{\mathbf{X}} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \sum_{a=1}^{A}\lambda_a\mathbf{p}_a\mathbf{p}_a', \tag{15c}$$

$$\mathbf{R}_{\hat{\mathbf{y}}} = (N-1)^{-1}\hat{\mathbf{Y}}'\hat{\mathbf{Y}} = \mathbf{Q}\mathbf{\Delta}\mathbf{Q}' = \sum_{a=1}^{A}\lambda_a b_a^2\mathbf{q}_a\mathbf{q}_a', \tag{15d}$$

where λ_i and $\hat{\delta}_i$ are the estimated variances of t_i and $\hat{u}_i$, respectively.

The main geometric properties of the PLS-decomposition are represented in Fig. 1 for an idealized low-order hypothetical system. Both measurement vectors ($\mathbf{x}$ and $\mathbf{y}$) are decomposed into their respective projections; and the corresponding control regions (intervals for S_{RX} and S_{RY}, and ellipses for S_{MX} and S_{MY}), are also indicated. The oblique minimum angles θ_X (θ_Y) between the model subspace S_{MX} (S_{MY}) and the complementary residual subspace S_{RX} (S_{RY}) are easily visualized.

3. Process monitoring based on PLS-decomposition

The SPM includes three main activities: 1) the detection of an out-of-control condition or occurrence of an anomaly; 2) the diagnostic achieved by classifying the type of anomaly that has generated the abnormal behavior, and 3) the isolation of the disturbed variables and – ideally – of the disturbing sources. In the following, every one of these points is analyzed.

3.1. PLS-based fault detection

Once an in-control PLSR model is developed using process data under normal operating conditions, the process state can be followed by projecting the actual measurements $\mathbf{x}$ and $\mathbf{y}$ on subspaces S_{MX}, S_{RX}, S_{MY}, and S_{RY}, and by using the proper statistics with its control limits. For example, a signal of an unexpected change in the input variables $\mathbf{x}$ can be detected by using the $\mathbf{t}$-based Hotelling's T^2 statistic in S_{MX},

$$T_{\mathbf{t}}^2 = \left\|\mathbf{\Lambda}^{-1/2}\mathbf{t}\right\|^2 = \left\|\left(\mathbf{\Lambda}^{-1/2}\mathbf{R}'\right)\hat{\mathbf{x}}\right\|^2 = \left\|\left(\mathbf{\Lambda}^{-1/2}\mathbf{R}'\right)\mathbf{x}\right\|^2, \tag{16}$$

which can be compared against a control limit. This statistic measures a meaningful change on the model subspace S_{MX}, and serves for discriminating excessive changes from normal variations.

However, when a special event that was not originally considered by the PLSR model occurs, the new observations $\mathbf{x}$ will partially move

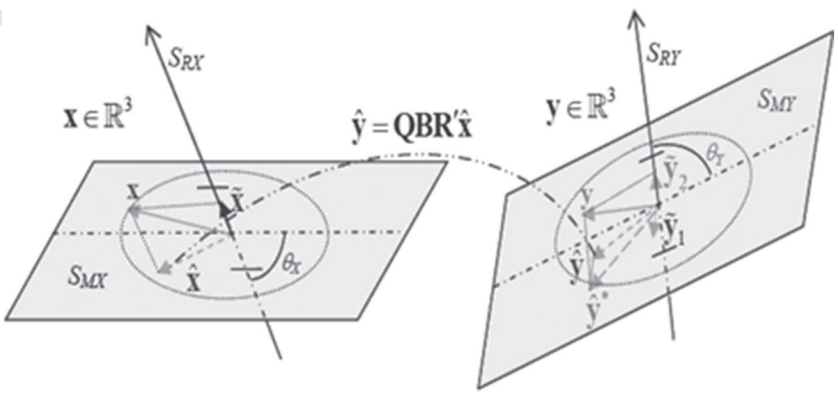

Fig. 1. Schematic representation of the PLS-decomposition of the measurement vectors $\mathbf{x}$ and $\mathbf{y}$ in their projections on the model and residue subspaces. The in-control regions are also indicated.

outside S_{MX} towards S_{RX} , showing an increment in the squared of the predicted error $SPE_{\mathbf{x}}$,

$$SPE_{\mathbf{x}} = \|\tilde{\mathbf{x}}\|^2 = \|(\mathbf{I}-\mathbf{P}\mathbf{R}')\mathbf{x}\|^2. \quad (17)$$

This last statistic measures the distance to the model lying in S_{RX}; in other words, it shows a departure of $\mathbf{x}$ from the regular expected behavior by going to the subspace S_{RX}. Similarly, the $\hat{\mathbf{u}}$-based Hotelling's T^2 statistic can be written as

$$T^2_{\hat{\mathbf{u}}} = \left\|\Delta^{-1/2}\hat{\mathbf{u}}\right\|^2 = \left\|\left(\Delta^{-1/2}\mathbf{S}'\right)\hat{\mathbf{y}}\right\|^2, \quad (18)$$

and used for detecting significant amplitude changes lying in the model subspace S_{MY}. Also, the distance to the regression model in S_{MY} can be written as [see Eqs. (12, 14)]:

$$SPE_{\mathbf{y}1} = \|\tilde{\mathbf{y}}_1\|^2 = \left\| \begin{bmatrix} \mathbf{Q}\mathbf{S}' & -\mathbf{Q}\mathbf{B}\mathbf{R}' \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} \right\|^2 = \|\mathbf{Q}(\mathbf{S}'\mathbf{y}-\mathbf{B}\mathbf{R}'\mathbf{x})\|^2. \quad (19)$$

Similarly, the distance to the model lying in S_{RY} is defined as:

$$SPE_{\mathbf{y}2} = \|\tilde{\mathbf{y}}_2\|^2 = \|(\mathbf{I}-\mathbf{Q}\mathbf{S}')\mathbf{y}\|^2. \quad (20)$$

Matrices $\mathbf{R}_{\hat{\mathbf{x}}}$ and $\mathbf{R}_{\hat{\mathbf{y}}}$ (Eqs. 15c, 15d) are often singular because $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ typically have collinear variables, as it is inferred from Eq. (14c). Consequently, the generalized Mahalanobis distance of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ can be used to measure the projections as follows:

$$D_{\hat{\mathbf{x}}} = \hat{\mathbf{x}}'\mathbf{R}^-_{\hat{\mathbf{x}}}\hat{\mathbf{x}}, \quad (21a)$$

$$D_{\hat{\mathbf{y}}} = \hat{\mathbf{y}}'\mathbf{R}^-_{\hat{\mathbf{y}}}\hat{\mathbf{y}}. \quad (21b)$$

The following theorem states that not all these statistics are independent.

Theorem 2. *The Mahalanobis distances computed from* $\hat{\mathbf{x}}$, $\mathbf{t}$, $\hat{\mathbf{u}}$, *and* $\hat{\mathbf{y}}$ *are equivalent*, i.e., $D_{\hat{\mathbf{x}}} = T^2_{\mathbf{t}} = T^2_{\hat{\mathbf{u}}} = D_{\hat{\mathbf{y}}}$ (*see Proof* 3 *in Appendix* A).

The identity in Theorem 2 suggests that the behavior of the response variables $\mathbf{y}$ can be monitored through the predictor variables $\mathbf{x}$. Therefore, the monitoring of the complete measurement space can be implemented through the following independent statistics: $T_{\mathbf{t}}^2$, $SPE_{\mathbf{x}}$, $SPE_{\mathbf{y}1}$, and $SPE_{\mathbf{y}2}$, each one actuating in a different subspace: S_{MX}, S_{RX}, S_{MY}, and S_{RY}, respectively (see Fig. 1). Consequently, a combined detection index I_{TC} can be defined to attain the whole measurement space without signal superposition,

$$I_{TC} = \frac{T^2_{\mathbf{t}}}{\tau^2_\alpha} + \frac{SPE_{\mathbf{x}}}{\delta^2_{\mathbf{x},\alpha}} + \frac{SPE_{\mathbf{y}1}}{\delta^2_{\mathbf{y}1,\alpha}} + \frac{SPE_{\mathbf{y}2}}{\delta^2_{\mathbf{y}2,\alpha}} = \begin{bmatrix} \mathbf{y}' & \mathbf{x}' \end{bmatrix} \Phi \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix}, \quad (22)$$

where τ^2_α, $\delta^2_{\mathbf{x},\alpha}$, $\delta^2_{\mathbf{y}1,\alpha}$, $\delta^2_{\mathbf{y}2,\alpha}$ are the respective control limits with a confidence level α, and Φ is a symmetric positive-definite matrix ($\Phi > 0$)). Note that in this way the significance level of each normalized statistic is 1. The vector arrangement on the right of Eq. (22) is derived from Eqs. (16, 17, 19, 20), and it shows that the resulting quadratic index depends on the extended vector $[\mathbf{y}'\ \mathbf{x}']'$. As indicated in Appendix B, a control limit I_α (with a confidence level α) can also be defined for I_{TC}; and this index is useful for simultaneous monitoring of product quality, process changes, and sensor problems. This combined detection index defines a multidimensional elliptic region that is compatible with the assumption of multi-normal data. Therefore, the amount of false alarms and undetected faults are significantly reduced with respect to the performance typically yielded by separated indexes [12]. The main difference among the statistics composing the combined index I_{TC} are the different scaling factor and the subspace where they are relevant. Typically, the scale τ^2_α for the statistic $T_{\mathbf{t}}^2$ is much larger than the scale factors $\delta^2_{\mathbf{x},\alpha}$, $\delta^2_{\mathbf{y}1,\alpha}$, or $\delta^2_{\mathbf{y}2,\alpha}$ for the quadratic errors. This happens because the noise-like variability of the errors are much smaller than the signal variability of $T_{\mathbf{t}}^2$ responding to the correlation structure. Thus, a multivariate change in $\mathbf{x}$ (from its normal values) that lies on S_{MX} produces a contribution to I_{TC} that is lower than a change of similar magnitude on S_{RX}, S_{MY}, or S_{RY}, measured by $SPE_{\mathbf{x}}$, $SPE_{\mathbf{y}1}$, or $SPE_{\mathbf{y}2}$, respectively.

3.2. Diagnosis of process anomalies

In order to determine the discriminating capacity of the four statistics composing the index I_{TC}, an artificial process system – here identified with the subscript 0 – is created for generating ideal data obeying to a predetermined correlation structure. This artificial process is defined by first setting values to the PLSR matrices $\mathbf{\Lambda}_0$, $\mathbf{P}_0$, $\mathbf{B}_0$, and $\mathbf{Q}_0$, that stand for a model of the system behavior under normal condition. In this procedure, the random score vector $\mathbf{t}_0 \in \mathbb{R}^A \sim N(\mathbf{0}, \mathbf{\Lambda}_0)$ works as an independent variable to generate the associate vectors of "input" and "output" responses by

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{P}_0\,\mathbf{t}_0 \quad \in S_{MX} \equiv Span\{\mathbf{P}_0\} \subseteq \mathbb{R}^m, \\ \mathbf{y}_0 &= \mathbf{Q}_0\mathbf{B}_0\mathbf{t}_0 \quad \in S_{MY} \equiv Span\{\mathbf{Q}_0\} \subseteq \mathbb{R}^p, \end{aligned} \quad (23)$$

Since Eq. (23) are linear combinations of random variables, the resulting input and output vectors respectively follow: $\mathbf{x}_0 \sim N(\mathbf{0}, \mathbf{P}_0\mathbf{\Lambda}_0\mathbf{P}'_0)$ and $\mathbf{y}_0 \sim N(\mathbf{0}, \mathbf{Q}_0\mathbf{B}_0\mathbf{\Lambda}_0\mathbf{B}_0\mathbf{Q}'_0)$, thus featuring 'common-cause' variations only. Since these data stand for an ideal perfect model, the residuals $\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1$ and $\tilde{\mathbf{y}}_2$ are null, and there are no differences between model predictions and ideal data, i.e., $\hat{\mathbf{x}} = \mathbf{x}_0$ and $\hat{\mathbf{y}} = \hat{\mathbf{y}}^* = \mathbf{y}_0$.

An anomalous event is detected by a significant change in the measurements, which triggers the alarm condition when $I_{TC} \geq I_\alpha$. The above measurement decomposition allows the discrimination among six different classes of anomaly sources, facilitating in this way the diagnostic task. Each one of these six alternatives is analyzed by assuming that data are generated under localized pure disturbances and then observing the detection track followed with the available model ($\mathbf{\Lambda}_0, \mathbf{P}_0, \mathbf{B}_0, \mathbf{Q}_0$). The sketch in Fig. 2 helps to visualize how a warning signal (at $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}_1$, $\tilde{\mathbf{y}}_2$ or $\mathbf{t}$) is generated as the disturbed measurements ($\mathbf{x}$ or $\mathbf{y}$) go through the PLSR model.

Hence, under the above framework, the following classes of anomalies are discriminated:

Class 1. Sensor faults associated to $\mathbf{x}$, represented by additional signals $\boldsymbol{\xi}$, that take the $\mathbf{x}$ vector out of the pattern cast by the PLSR model. In this case, the input vector can be written as,

$$\mathbf{x} = \mathbf{x}_0 + \boldsymbol{\xi}, \quad \text{(Data generation)}$$

where $\mathbf{x}_0$ is the "free-of-change" part of the input-vector measurement. Let us assume that the disturbance $\boldsymbol{\xi}$ does not follow the correlation pattern of $\mathbf{x}_0$ at all (see Eq. (23)), in other words, $\boldsymbol{\xi} \in S_{RX}$. Then, $\mathbf{t} = \mathbf{R}'_0\mathbf{x} = \mathbf{R}'_0\mathbf{x}_0 + \mathbf{R}'_0\boldsymbol{\xi} = \mathbf{t}_0 + 0 = \mathbf{t}_0$, which shows that the disturbance goes completely to the residual variable

$$\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}} = \mathbf{x} - \mathbf{P}_0\mathbf{t} = \mathbf{x} - \mathbf{x}_0 = \boldsymbol{\xi} \neq 0 \ \in S_{RX} \text{(Disturbance detection)}$$

Class 2. Sensors faults associated to the output variable $\mathbf{y}$,

$$\mathbf{y} = \mathbf{y}_0 + \boldsymbol{\eta}, \quad \text{(Data generation)}$$

where $\mathbf{y}_0$ is the "free-of-change" part with a population structure determined by the normal conditions only. This disturbance is analyzed by assuming that $\boldsymbol{\eta} \in S_{RY}$. Hence, the disturbance track from generation to detection is as follows:

$$\hat{\mathbf{y}}^* = \mathbf{Q}_0\mathbf{S}'_0\mathbf{y} = \mathbf{Q}_0\mathbf{S}'_0\mathbf{y}_0 + \mathbf{Q}_0\mathbf{S}'_0\boldsymbol{\eta} = \mathbf{y}_0 + 0 = \mathbf{y}_0,$$

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006

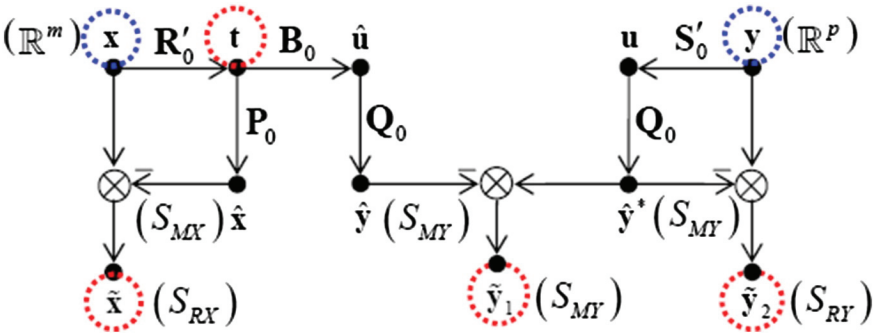

Fig. 2. Measurements decomposition based on projections onto the subspaces created by a PLSR model.

$$\tilde{\mathbf{y}}_2 = \mathbf{y} - \hat{\mathbf{y}}^* = \boldsymbol{\eta} \neq 0 \in S_{RY}, \text{(Disturbance detection)}$$

and again, a measurement disturbance not associable with the normal behavior is sent to a residual subspace, in this case, S_{RY}.

Class 3. Changes in the correlation structure of $\mathbf{x}$. An $\mathbf{x}$ vector population affected by this kind of change can also be generated by the same score $\mathbf{t}_0$, but using an "unknown" change $\Delta\mathbf{P} = \mathbf{P} - \mathbf{P}_0$, i.e.,

$$\mathbf{x} = \mathbf{P}\mathbf{t}_0 = (\mathbf{P}_0 + \Delta\mathbf{P})\mathbf{t}_0 = \mathbf{x}_0 + \Delta\mathbf{P}\mathbf{t}_0. \quad \text{(Data generation)}$$

It should be noted, however, that when using the measurement $\mathbf{x}$ with the available model $\mathbf{P}_0$, the scores change to $\mathbf{t} = \mathbf{R}_0'\mathbf{x} = \mathbf{R}_0'(\mathbf{P}_0 + \Delta\mathbf{P})\mathbf{t}_0 = \mathbf{t}_0 + \mathbf{R}'_0\Delta\mathbf{P}\mathbf{t}_0$, thus generating the following no-null residual values

$$\begin{aligned}\tilde{\mathbf{x}} &= \mathbf{x} - \hat{\mathbf{x}} = (\mathbf{x}_0 + \Delta\mathbf{P}\mathbf{t}_0) - \mathbf{P}_0\mathbf{t} = (\mathbf{x}_0 + \Delta\mathbf{P}\mathbf{t}_0) - (\mathbf{x}_0 + \mathbf{P}_0\mathbf{R}'_0\Delta\mathbf{P}\mathbf{t}_0)\\ &= (\mathbf{I} - \mathbf{P}_0\mathbf{R}'_0)\Delta\mathbf{P}\mathbf{t}_0 \neq 0 \in S_{RX}, \tilde{\mathbf{y}}_1 = \hat{\mathbf{y}}^* - \hat{\mathbf{y}} = \mathbf{y}_0 - \mathbf{Q}_0\mathbf{B}_0\mathbf{t}\\ &= \mathbf{y}_0 - \mathbf{Q}_0\mathbf{B}_0(\mathbf{t}_0 + \mathbf{R}'_0\Delta\mathbf{P}\mathbf{t}_0) = -\mathbf{Q}_0\mathbf{B}_0\mathbf{R}'_0\Delta\mathbf{P}\mathbf{t}_0 \neq 0 \in S_{MY}.\end{aligned}$$

where the residues $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}_1$ are used to detect the disturbance. Also, it is important to note that $\left\|\boldsymbol{\Lambda}_0^{-1/2}\mathbf{R}'_0\tilde{\mathbf{x}}\right\|^2/\tau_\alpha^2 \ll \|\tilde{\mathbf{x}}\|^2/\delta_{\mathbf{x},\alpha}^2$ when $\|\hat{\mathbf{x}}\|^2 = \|\tilde{\mathbf{x}}\|^2$.

Class 4. Changes in the intrinsic gains. The diagonal matrix $\mathbf{B}$ is the core place where the PLSR model ties up input with output variables. Let us assume an unknown change occurred in this matrix, i.e.,

$$\hat{\mathbf{y}}^* = \mathbf{Q}_0\mathbf{B}\mathbf{t}_0 = \mathbf{Q}_0(\mathbf{B}_0 + \Delta\mathbf{B})\mathbf{t}_0 = \mathbf{y}_0 + \mathbf{Q}_0\Delta\mathbf{B}\mathbf{t}_0, \text{(Data generation)}$$

and analyze the effects on the statistics being considered for monitoring,

$$\tilde{\mathbf{y}}_1 = \hat{\mathbf{y}}^* - \hat{\mathbf{y}} = (\mathbf{y}_0 + \mathbf{Q}_0\Delta\mathbf{B}\mathbf{t}_0) - \mathbf{Q}_0\mathbf{B}_0\mathbf{t}_0 = \mathbf{Q}_0\Delta\mathbf{B}\mathbf{t}_0 \neq 0 \in S_{MY}$$

which tells that the "detecting signal" should be observed by monitoring the residue $\tilde{\mathbf{y}}_1$ on S_{MY}, while the remaining statistics remain unaffected (see Fig. 2).

Class 5. Changes in the correlation structure of $\mathbf{y}$. Let us assume an unknown change in the matrix $\mathbf{Q}$, i.e.

$$\mathbf{y} = \mathbf{Q}\mathbf{B}_0\mathbf{t}_0 = (\mathbf{Q}_0 + \Delta\mathbf{Q})\mathbf{B}_0\mathbf{t}_0 = \mathbf{y}_0 + \Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0 \quad \text{(Data generation)}$$

and the detection track characterized by

$$\begin{aligned}\hat{\mathbf{y}}^* &= \mathbf{Q}_0\mathbf{S}'_0\mathbf{y} = \mathbf{Q}_0\mathbf{S}'_0(\mathbf{y}_0 + \Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0) = \mathbf{y}_0 + \mathbf{Q}_0\mathbf{S}'_0\Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0 \tilde{\mathbf{y}}_2 = \mathbf{y} - \hat{\mathbf{y}}^*\\ &= (\mathbf{y}_0 + \Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0) - (\mathbf{y}_0 + \mathbf{Q}_0\mathbf{S}'_0\Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0)\\ &= (\mathbf{I} - \mathbf{Q}_0\mathbf{S}'_0)\Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0 \neq 0 \in S_{RY}\end{aligned}$$

and also

$$\tilde{\mathbf{y}}_1 = \hat{\mathbf{y}}^* - \hat{\mathbf{y}} = \left(\mathbf{y}_0 + \mathbf{Q}_0\mathbf{S}'_0\Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0\right) - \mathbf{y}_0 = \mathbf{Q}_0\mathbf{S}'_0\Delta\mathbf{Q}\mathbf{B}_0\mathbf{t}_0 \neq 0 \in S_{MY}.$$

where the residues $\tilde{\mathbf{y}}_1$ and $\tilde{\mathbf{y}}_2$ are used to "detect the disturbance".

Class 6. A significant change in the process condition produces measurements that follow the correlation structure captured by the undisturbed PLSR model. This type of disturbance can be represented by changes in the population parameters of $\mathbf{t}_0 \sim N(\mathbf{0}, \boldsymbol{\Lambda}_0)$, i.e., we have to assume a displacement of $E\{\mathbf{t}_0\}$ from $\mathbf{0}$ to a $\boldsymbol{\mu}_\mathbf{t} \neq \mathbf{0}$, or a significant change in the variability from $\boldsymbol{\Lambda}_0$ to $\boldsymbol{\Lambda}_\mathbf{t} \neq \boldsymbol{\Lambda}_0$. Hence, the anomaly in this case produces

$$\mathbf{t} = \mathbf{t}_0 + \Delta\mathbf{t} \sim N(\boldsymbol{\mu}_\mathbf{t}, \boldsymbol{\Lambda}_\mathbf{t}) \quad \text{(Data generation)}$$

with a $\Delta\mathbf{t}$ of magnitude such that: $T_\mathbf{t}^2 = \|\boldsymbol{\Lambda}_0^{-1/2}\mathbf{t}\|^2 = \|\boldsymbol{\Lambda}_0^{-1/2}(\mathbf{t}_0 + \Delta\mathbf{t})\|^2 \geq \tau_\alpha^2$ (disturbance detection).

The above results are summarized in Table 2. The highlighted discrimination patterns indicate the statistics that are activated as soon as the measurements $(\mathbf{x}, \mathbf{y})$ bring information about a localized model mismatch. The six analyzed anomalies can qualitatively be grouped into the following three categories: sensor fault (classes 1 and 2), process change (classes 3, 4 and 5), and excessive (quality) operation change (class 6).

Some final comments about the effect of the dimension A on the discriminating sensitivity seem convenient at this point. Recall that the predictor variables $\mathbf{x}$ are assumed as belonging to an m dimensional space, i.e., $\mathbf{x} \in \mathbb{R}^m$. This space is then projected onto the subspaces S_{MX} and S_{RX} where $dim(S_{MX}) = A \leq m$ and $dim(S_{RX}) = m\text{-}A$. Hence, when $A \to m$, the subspace S_{RX} (useful for detecting sensor faults in the predictor variables $\mathbf{x}$) reduces, leaving information of possible anomalous measurements in S_{MX}. It is then understandable that when A increases beyond a certain point, the ability for failure detection on S_{RX} tends to disappear. At the same time, too small values of A might move the correlated behavior to S_{RX}, playing an interfering roll in the residue index $SPE_\mathbf{x}$. A similar reasoning can be done with the output variables $\mathbf{y}$, originally in a space of dimension p, i.e., $\mathbf{y} \in \mathbb{R}^p$.

Table 2
Expected statistics pattern for each anomaly class.

Anomaly class	Statistics composing I_{TC} (associated subspaces)			
	$T_\mathbf{t}^2/\tau_\alpha^2$ (S_{MX})	$SPE_\mathbf{x}/\delta_{\mathbf{x},\alpha}^2$ (S_{RX})	$SPE_{\mathbf{y}1}/\delta_{\mathbf{y}1,\alpha}^2$ (S_{MY})	$SPE_{\mathbf{y}2}/\delta_{\mathbf{y}2,\alpha}^2$ (S_{RY})
Class 1: fault of sensors in $\mathbf{x}$ $(\boldsymbol{\xi})$	−	++	−	−
Class 2: fault of sensors in $\mathbf{y}$ $(\boldsymbol{\eta})$	−	−	−	++
Class 3: changes in correlations of $\mathbf{x}$ $(\Delta\mathbf{P})$	−	++	++	−
Class 4: changes in the intrinsic gains $(\Delta\mathbf{B})$	−	−	++	−
Class 5: changes in correlations of $\mathbf{y}$ $(\Delta\mathbf{Q})$	−	−	++	++
Class 6: excessive change in $(\boldsymbol{\mu}_\mathbf{t}, \boldsymbol{\Lambda}_\mathbf{t})$	++	−	−	−

++: significant value. −: negligible value.

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006

3.3. Isolation of disturbed variables or disturbing sources

When trying to localize a faulty sensor or detect an abnormal correlation, the identification of the involved variables becomes very helpful. The use of contribution analysis to characterize an anomaly has been previously proposed [11]. The preliminary classification of the anomalous event according to Table 2, enable us to restrict the search problem to one or two statistics.

A simple strategy for decomposing the four statistics in Eq. (22) is also proposed here. The arrangement can be applied to any quadratic form $I^2 = \mathbf{z}'\mathbf{A}\mathbf{z}$ (like T_t^2, $SPE_{\mathbf{x}}$, $SPE_{\mathbf{y}1}$, or $SPE_{\mathbf{y}2}$), to reveal the contribution intensity of each variable z_i. This decomposition is performed as follows:

$$I^2 = \mathbf{z}'\mathbf{A}\sum_{i=1}^{n}\mathbf{z}_i = \sum_{i=1}^{n}\mathbf{z}'\mathbf{a}_i z_i = \sum_{i=1}^{n} cI^2(z_i) \tag{24}$$

where $\mathbf{a}_i = [a_{i1} \cdots a_{in}]'$ is the i-th column of the matrix $\mathbf{A}$, $\mathbf{z}_i \triangleq [0 \cdots z_i \cdots 0]'$ $(n \times 1)$ is a vector with only a nonzero value in the i-th position, and $cI^2(z_i)$ stands for the contribution of z_i to I^2. This contribution captures the gradient of the statistic in the z_i direction, i.e.,

$$\frac{1}{2}\nabla_{\mathbf{z}}\left(I^2\right)\cdot\mathbf{z}_i = \frac{1}{2}\nabla_{\mathbf{z}}(\mathbf{z}'\mathbf{A}\mathbf{z})\cdot\mathbf{z}_i = \mathbf{z}'\mathbf{a}_i z_i = cI^2(z_i). \tag{25}$$

Note that it can also be written in terms of the projections by substituting the relationships $\mathbf{R}'\widehat{\mathbf{x}} = \mathbf{R}'\mathbf{x}$, $(\mathbf{I}-\mathbf{P}\mathbf{R}')\widetilde{\mathbf{x}} = (\mathbf{I}-\mathbf{P}\mathbf{R}')\mathbf{x}$, $\mathbf{S}'\widehat{\mathbf{y}}^* = \mathbf{S}'\mathbf{y}$ and $(\mathbf{I}-\mathbf{Q}\mathbf{S}')\widetilde{\mathbf{y}}_2 = (\mathbf{I}-\mathbf{Q}\mathbf{S}')\mathbf{y}$ in Eqs. (16, 17, 19, 20), yielding

$$cT_{\mathbf{t}}^2(x_i) = \mathbf{x}'\mathbf{a}_i x_i \equiv \widehat{\mathbf{x}}'\mathbf{a}_i\widehat{x}_i, \quad \text{with } \mathbf{A} = \mathbf{R}\boldsymbol{\Lambda}^{-1}\mathbf{R}' = [\mathbf{a}_1 \cdots \mathbf{a}_m]_{m\times m} \tag{26}$$

$$cSPE_{\mathbf{x}}(x_i) = \widetilde{\mathbf{x}}'\mathbf{a}_i\widetilde{x}_i, \text{with } \mathbf{A} = (\mathbf{I}-\mathbf{P}\mathbf{R}')'(\mathbf{I}-\mathbf{P}\mathbf{R}') = [\mathbf{a}_1 \cdots \mathbf{a}_m]_{m\times m} \tag{27}$$

$$cSPE_{\mathbf{y}1}(x_i) = [\widehat{\mathbf{y}}^{*\prime} \;\; \widehat{\mathbf{x}}']\mathbf{a}_j\widehat{x}_i, \quad cSPE_{\mathbf{y}1}\left(y_j\right) = [\widehat{\mathbf{y}}^{*\prime} \;\; \widehat{\mathbf{x}}']\mathbf{a}_j\widehat{y}_j^*$$
$$\text{with } \mathbf{A} = \begin{bmatrix} \mathbf{S}\mathbf{S}' & -\mathbf{S}\mathbf{B}\mathbf{R}' \\ -\mathbf{R}\mathbf{B}\mathbf{S}' & \mathbf{R}\mathbf{B}^2\mathbf{R}' \end{bmatrix} = \left[\mathbf{a}_1 \cdots \mathbf{a}_p \mathbf{a}_{p+1} \cdots \mathbf{a}_{p+m}\right]_{(p+m)\times(p+m)} \tag{28}$$

$$cSPE_{\mathbf{y}2}\left(y_j\right) = \widetilde{\mathbf{y}}'_2\,\mathbf{a}_j\widetilde{y}_{2,j}, \text{with } \mathbf{A} = (\mathbf{I}-\mathbf{Q}\mathbf{S}')'(\mathbf{I}-\mathbf{Q}\mathbf{S}') = \left[\mathbf{a}_1 \cdots \mathbf{a}_p\right]_{p\times p}. \tag{29}$$

Then, by combining Eqs. (22), (24), (26)–(29) gives

$$I_{TC} = I_{MX}(\mathbf{x}) + I_{RX}(\mathbf{x}) + I_{RY1}(\mathbf{x},\mathbf{y}) + I_{RY2}(\mathbf{y})$$
$$I_{TC} = \sum_{i=1}^{m} cI_{MX}(x_i) + \sum_{i=1}^{m} cI_{RX}(x_i) + \left(\sum_{i=1}^{m} cI_{RY1}(x_i) + \sum_{j=1}^{p} cI_{RY1}\left(y_j\right)\right) + \sum_{j=1}^{p} cI_{RY2}\left(y_j\right) \tag{30}$$

where e.g., the term $cI_{MX}(x_i) = cT_t^2(x_i)/\tau_\alpha^2$ in Eq. (30) is the normalized contribution of x_i to I_{TC} through the projection on S_{MX} (see Eq. (26)). Furthermore, since I_{MX} has a significance level of 1, then the significance level of their contributions cI_{MX} is also adopted equal to 1. In particular, $cT_{\mathbf{t}}^2(x_i) = x_i\mathbf{a}_i'\mathbf{x}$ (Eq. (26)) is equivalent to the contribution defined by Westerhuis et al. [29]; and it tends to the contribution reported by Alvarez et al. [30], when $S_{MX}\to\mathbb{R}^m$, because $D_{\widehat{\mathbf{x}}}\to T_{\mathbf{x}}^2 = \mathbf{x}'\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{x}$ with $\mathbf{R}_{\mathbf{x}} > 0$.

The variable contribution defined by Eq. (24) guarantees a correct diagnosis when $\mathbf{A}$ is positive semidefinite [11], which always holds for the four statistics of Eqs. (26)–(29). A reliable isolation of faulty sensors must correctly indicate the disturbed variable. For example, a fault in the k-th sensor of $\mathbf{x}$ can be represented as $\mathbf{x} = \mathbf{x}_0 + f_k\boldsymbol{\xi}_k$ [10], where f_k is the fault magnitude and $\boldsymbol{\xi}_k = [0\,0 \cdots 1 \cdots 0]'$ is the fault direction. If the faulty measurement projection is mostly included in S_{RX} (see Class 1), then it can be expressed as: $\widetilde{\mathbf{x}} \cong f_k\boldsymbol{\xi}_k$ (i.e. $\widetilde{x}_k \cong f_k$ and $\widetilde{x}_i \cong 0 \forall i \neq k$). Thus, by including such value into the contributions to I_{TC} (depending on the subspace), it results:

$$I_{MX}, I_{RY1}, I_{RY2} \cong 0,$$
$$cI_{RX}(x_i) = \widetilde{\mathbf{x}}'\mathbf{a}_i\widetilde{x}_i/\delta_{\mathbf{x},\alpha}^2 = \begin{cases} f_k\;\boldsymbol{\xi}'_k\mathbf{a}_i\;0\;/\;\delta_{\mathbf{x},\alpha}^2 = 0 & \text{for } i\neq k \\ f_k\boldsymbol{\xi}'_k\mathbf{a}_i f_k/\delta_{\mathbf{x},\alpha}^2 = f_k^2 a_{ii}/\delta_{\mathbf{x},\alpha}^2 & \text{for } i = k \end{cases} \tag{31}$$

A correct diagnosis occurs when the contribution of the non-faulty variable is less than or equal to the contribution of the faulty variable, i.e. when $cI_{RX}(x_k) \geq 0$ (see Eq. (31)). Therefore, a correct isolation is guaranteed when $a_{ii} \geq 0$, which is always true since $\mathbf{A} \geq 0$ (Eq. (27)).

4. Simulation test

The described monitoring procedure is now tested with random measurements obtained from the numerical simulation of two systems of different complexity. The first system is a synthetic example representing a hypothetical static process, where the internal data structure was arbitrary chosen. The second application is a more realistic problem where the variables of a dynamic process under normal conditions are empirically modeled by PLSR.

4.1. A four-state static process

The synthetic example is first simulated to better interpret the proposed methodology as a monitoring tool. The normal operation of the chosen system includes four internal states, which are represented by the following four points of its latent space ($\mathbf{t}$-scores): $\{(t_1^*,t_2^*)\}_{1\ldots4} = \{(1,1),(1,3),(3,3),(3,1)\}$. The "measurements" of the external variables, $\mathbf{x}$ and $\mathbf{y}$, are generated by adding zero-mean Gaussian random noises ($\boldsymbol{\varepsilon}_i$, $i = 1\ldots4$) to the PLSR correlation structure characterized by the arbitrary-selected process matrices $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{B}$, as follows:

$$\textit{Internal part}\begin{cases} \mathbf{t}_0 = \mathbf{t}^* + \boldsymbol{\varepsilon}_1, & \boldsymbol{\varepsilon}_1 \sim N\left(\mathbf{0}, 0.1^2\mathbf{I}_2\right), \\ \mathbf{u} = \mathbf{B}\mathbf{t}_0 + \boldsymbol{\varepsilon}_2, \quad \mathbf{B} = diag(2, 0.5), & \boldsymbol{\varepsilon}_2 \sim N\left(\mathbf{0}, 0.05^2\mathbf{I}_2\right), \end{cases}$$
$$\textit{External part}\begin{cases} \mathbf{x} = \mathbf{P}\mathbf{t}_0 + \boldsymbol{\varepsilon}_3, \quad \mathbf{P} = [\mathbf{p}_1 \;\; \mathbf{p}_2], & \boldsymbol{\varepsilon}_3 \sim N\left(\mathbf{0}, 0.05^2\mathbf{I}_7\right), \\ \mathbf{y} = \mathbf{Q}\mathbf{u} + \boldsymbol{\varepsilon}_4, \quad \mathbf{Q} = [\mathbf{q}_1 \;\; \mathbf{q}_2], & \boldsymbol{\varepsilon}_4 \sim N\left(\mathbf{0}, 0.05^2\mathbf{I}_5\right), \end{cases}$$
$$\mathbf{p}_1 = [0.4045, 0, 0.5394, 0.2697, 0.1348, 0, 0.6742]',$$
$$\mathbf{p}_2 = [0, 0.7906, 0.1581, -0.1581, -0.3162, 0.4743, 0]',$$
$$\mathbf{q}_1 = [0.5883, 0, -0.1961, 0, 0.7845]',$$
$$\mathbf{q}_2 = [0.0081, 0.7054, 0.0678, -0.7054, 0.0109]', \tag{32}$$

This model is used to simulate 40 multivariate observations under 'normal conditions' and the generated dataset is used to fit the PLSR model. The selection of $A = 2$ is determined by monitoring the simultaneous deflation of $\mathbf{X}$ and $\mathbf{Y}$ [18]; in this way the errors regarding the "true" matrices $\mathbf{Q}$, $\mathbf{B}$, and $\mathbf{P}$ are negligible. The control limits τ_α^2, $\delta_{\mathbf{x},\alpha}^2$, $\delta_{\mathbf{y}1,\alpha}^2$ and $\delta_{\mathbf{y}2,\alpha}^2$ are calculated as explained in Appendix B.

Table 3 shows six simulated anomalies (one for each class in Table 2): a) anomalies of classes 1 and 2 correspond to sensor biases that are simulated by disturbing the measurements $\mathbf{x}$ and $\mathbf{y}$; b)

Table 3
Simulated scenarios of anomalies.

Location	Anomaly class	Magnitude of the change/fault	I_{TC}/I_a ($I_a = 2.266$)
$k = 11$	1	$\Delta\mathbf{x} = [0.3\;0\;0\;0\;0\;0.25\;0]'$ (Multiple fault)	1.980
$k = 19$	2	$\Delta\mathbf{y} = [0.35\;0\;0\;0\;0]'$ (Single fault)	2.381
$k = 27$	3	$\Delta\mathbf{p}_2 = [0\;0.28\;0\;0\;-0.07\;0.14\;-0.14]'$	5.799
$k = 35$	4	$\Delta\mathbf{B}_{22} = 0.25$	3.291
$k = 43$	5	$\Delta\mathbf{q}_1 = [-0.05\;0.025\;0.05\;0.025\;-0.1]'$	3.700
$k = 51$	6	$\Delta\mathbf{t} = [0\;6]'$	2.715

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006

anomalies of classes 3, 4, and 5 are implemented by altering the process matrices; and c) anomaly of class 6 consists in adding up a change $\Delta\mathbf{t}$ to $\mathbf{t}_0$, such that $I_{TC} > I_\alpha$. Each fault is simulated by affecting only one sample point at a discrete time, k; and immediately the anomaly is canceled from $k + 1$ onwards. These anomalies represent a hard test for evaluating the ability of the method.

Fig. 3a shows the time evolution of the I_{TC} detection index and its component statistics. The alarm condition is triggered at a given sample k, when the I_{TC} index overpasses the control limit. The index I_{TC} proved to be effective for detecting all simulated anomalies. Table 3 shows the proportion of the variation of I_{TC} with respect to control limit for each simulated fault-scenario. The patterns of the statistics recorded in

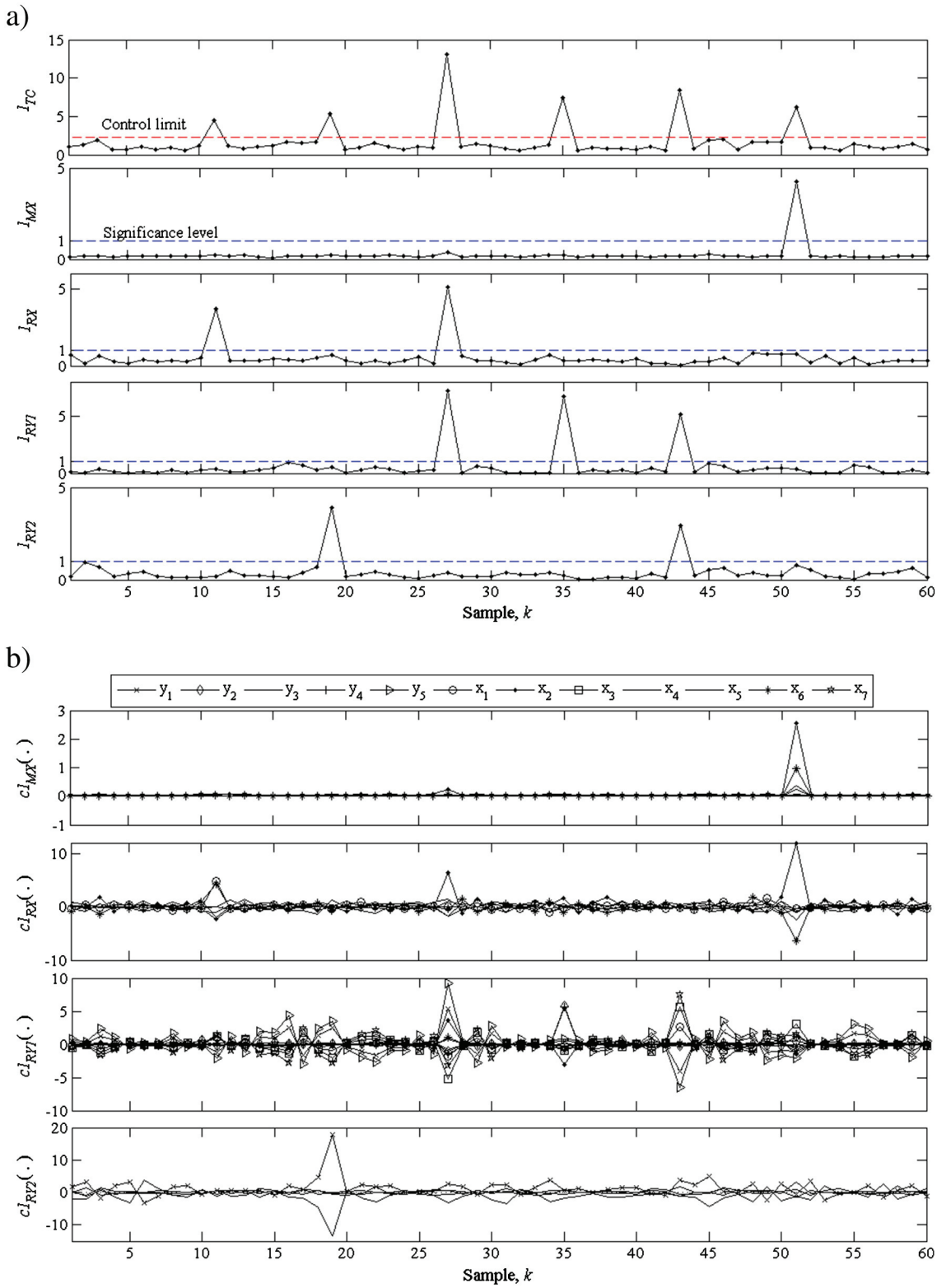

Fig. 3. Simulation example based on a statistic process. (a) Temporal evolution of the combined index I_{TC} and of its component statistics. (b) Variable contributions to each component statistics of I_{TC}.

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006

Fig. 3a along with the information given in Table 2 allow an unambiguous classification of each anomaly class (see Table 4). Fig. 3b shows the instantaneous variable contributions to each component statistic of I_{TC}. For example, at $k = 11$, the main contributions to I_{TC} are $cI_{RX}(x_1)$ and $cI_{RX}(x_6)$; and therefore faults in sensors x_1 and x_6 are diagnosed (first row of Table 4). Similarly, at $k = 19$, $cI_{RY2}(y_1)$ diagnoses a fault in sensor y_1 (second row of Table 4). The remaining results in Table 4 are summarized as follows: a) at $k = 27$ and $k = 43$, correlation changes in $\mathbf{x}$ and $\mathbf{y}$ are respectively diagnosed, which indicate major changes in the original external correlations (i.e., $\mathbf{P}$ and $\mathbf{Q}$ in Eq. (32)); b) at $k = 35$, $cI_{RY1}(y_2)$ and $cI_{RY1}(y_4)$ diagnose an upset of the t_2-u_2 inner relation, because the contributions of y_2 and y_4 are more correlated with u_2 (see the main components of $\mathbf{q}_2$ in Eq. (32)); and c) at $k = 51$, $cI_{MX}(x_2)$ and $cI_{MX}(x_6)$ diagnose an excessive change in the normal operation point (though still following the correlation structure), because the contributions of x_2 and x_6 are more correlated with t_2 (see the main components of $\mathbf{p}_2$ in Eq. (32)).

4.2. Application to a chemical process with feedback control

The chemical process described by Yoon and MacGregor [31] is here adopted to evaluate the proposed method of process monitoring with fault detection and isolation. The process consists of a first-order chemical reaction carried out in a non-isothermal continuously-stirred tank reactor, where both the solute (A) and the solvent (S) are continuously fed into the reactor (Fig. 4). This example has widely been used in the literature to test other methods of fault detection and diagnosis; e.g. methods based on both linear [31] and nonlinear [14,32,33] LV models. The mathematical model of the process is given by [31]:

$$
\begin{aligned}
&F = F_A + F_S,\\
&C_{A0} = (C_{AA}F_A + C_{AS}F_S)/F,\\
&\frac{dC_A}{dt} = \frac{F}{V}C_{A0} - \frac{F}{V}C_A - k_0 e^{-E/RT}C_A,\\
&\frac{dT}{dt} = \frac{F}{V}(T_0 - T) - \frac{aF_C^{3/2}}{\left(F_C + aF_C^{1/2}/2\rho_C C_{PC}\right)V\rho C_P}(T - T_C)\\
&\qquad + \frac{(-\Delta H_r)}{\rho C_P}k_0 e^{-E/RT}C_A,\\
&Q = (-\Delta H_r)Vk_0 e^{-E/RT}C_A,
\end{aligned}
\tag{33}
$$

with the following parameters: $V = 1$ (m^3), $\rho = 10^6$ (g/m^3), $\rho_C = 10^6$ (g/m^3), $E/R = 8330.1$ (K), $C_P = 1$ (cal/gK), $C_{PC} = 1$ (cal/gK), $k_0 = 10^{10}$ (m^3/kmol × min), $a = 1.678 \times 10^6$ (cal/min × K), and $\Delta H_r = -1.3 \times 10^7$ (cal/kmol).

The process is monitored from the following measurements: the inlet A flow rate (F_A), the inlet A concentration (C_{AA}), the inlet solvent flow rate (F_S), the A concentration in the solvent (C_{AS}), the reactant mixture temperature (T_0), the cooling water temperature (T_C), the coolant flow rate (F_C), the outlet concentration (C_A), the outlet temperature (T), and the reaction heat rate (Q). These measurements are arranged as follows: $\mathbf{x} = [F_A\ C_{AA}\ F_S\ C_{AS}\ T_0\ T_C\ F_C]'$ (process inputs), and $\mathbf{y} = [C_A T Q]'$ (process outputs). Additionally, the following measured process disturbances were selected: C_{AA}, F_S, C_{AS}, T_0, and T_C, which were simulated (under both normal and fault conditions) through a first-order autoregressive (AR) model, i.e.: $x_i(k) = \phi_i x_i(k-1) + \sigma_{e,i}e(k)$, where $e(k) \sim N(0,1)$ and $\sigma^2_{e,i}$ is the process variance, and ϕ_i is the AR parameter. Random measurement noises are added to all process variables according to: $x_{i,meas}(k) = x_i(k) + \sigma_{m,i}m(k)$ and $y_{i,meas} = y_i(k) + \sigma_{m,i}m(k)$, where $m(k) \sim N(0,1)$ and $\sigma^2_{m,i}$ is the noise variance. The parameters of the disturbances and measurement noises are shown in Table 5.

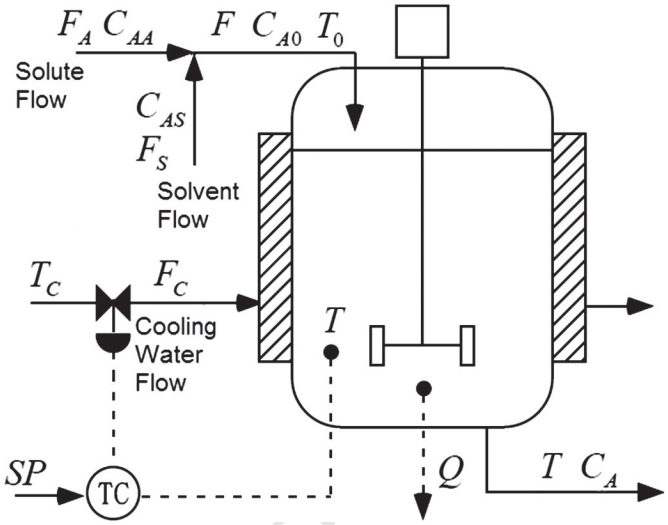

Fig. 4. Simulation example based on a dynamic process. Flow diagram of the process with 7 input- and 3 output measured variables.

The reaction temperature is regulated through a PI controller, implemented as: $dF_C/dt = K_C d(SP_T - T)/dt + K_C(SP_T - T)/T_I$, with the following parameters: $K_C = -1.5$ and $T_I = 5.0$. The adopted initial conditions are: $T_0 = 370.0$ K, and $C_A = 0.8$ kmol/m^3. The remaining variables are kept constant at the following values: $T_C = 365.0$ K, $F_C = 15$ m^3/min, $T = 368.25$ K, $F_S = 0.9$ m^3/min, $F_A = 0.1$ m^3/min, $C_A = 0.8$ kmole/m^3, $C_{AS} = 0.1$ kmol/m^3, and $C_{AA} = 19.1$ kmol/m^3. All variables were sampled at regular intervals of one minute. Under normal conditions, 200 samples were used to define the model calibration dataset. The following PLSR model was identified in this work:

$$\mathbf{y}(k) = \mathbf{QBR}'\mathbf{x}(k-1), \tag{34}$$

where $\mathbf{y}(k)$ represents the response vector at the discrete time k, and $\mathbf{x}(k\text{-}1)$ is the predictor vector at k-1. Eq. (34) represents a static model with a one-point delayed input. Based on criteria given by Godoy et al. [18], three LVs were required to fit the PLSR model. Since the number of LVs equals the number of process outputs, then the response space

Table 4
Anomaly classification and identification of disturbed variables.

Sample point	$cI_{MX}(.)$	$cI_{RX}(.)$	$cI_{RY1}(.)$	$cI_{RY2}(.)$	Diagnosis of the abnormal event	Diagnosed class
$k = 11$	–	x_1, x_6	–	–	*Faults in sensors of x_1 and x_6*	1
$k = 19$	–	–	–	y_1	*Fault in the sensor of y_1*	2
$k = 27$	–	x_2	y_5, y_1, x_2	–	*Correlation changes in* $\mathbf{x}$[a]	3
$k = 35$	–	–	y_2, y_4	–	*Change in the intrinsic gain*	4
$k = 43$	–	–	x_7, x_3	y_1	*Correlation changes in* $\mathbf{y}$[b]	5
$k = 51$	x_2, x_6	–	–	–	*Excessive operation change*	6

[a] Main changes occur in x_2.
[b] Main changes occur in y_1.

Table 5
Parameters for simulating process disturbances and measurement noises.

	Process disturbances		Measurement noises
	ϕ_i	$\sigma^2_{e,i}$	$\sigma^2_{m,i}$
F_A (x_1)			4.0×10^{-6}
C_{AA} (x_2)	0.9	0.475×10^{-1}	1.0×10^{-2}
F_S (x_3)	0.9	0.190×10^{-2}	4.0×10^{-6}
C_{AS} (x_4)	0.5	1.875×10^{-3}	2.5×10^{-5}
T_0 (x_5)	0.9	0.475×10^{-1}	2.5×10^{-3}
T_C (x_6)	0.9	0.475×10^{-1}	2.5×10^{-3}
F_C (x_7)			1.0×10^{-2}
C_A (y_1)			2.5×10^{-5}
T (y_2)			4.0×10^{-4}
Q (y_3)			1.0×10^{3}

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006

cannot be separated into two subspaces. Anyway, the monitoring technique can still be applied.

To evaluate the performance of the proposed technique, four types of faults are simulated as follows: F.1) a (non-propagating) sensor fault, simulated as a bias of +2.4 K in T_0; F.2) a sensor fault (propagated by the PI control loop), simulated as a bias of +1.5 K in T; F.3) a process change, simulated as a slow drift in the reaction kinetics, given by $k_0(k) = 0.999\ k_0(k-1)$; and F.4) a process disturbance, simulated as an unmeasured increment in the reactant mixture temperature, given by $T_{0,in} = T_0 + 0.03k$. According to the classification given by Yoon and MacGregor [31], F.1 is a simple fault while F.2, F.3, and F.4 are complex faults. The proposed cases are qualitatively similar to fault scenarios investigated in previous works [14,31–33]. Also, a quite similar process has been used to test a diagnosis technique based on a linear LV model, with the fault scenarios F.1 and F.2 [34].

For simplicity, each fault was introduced at $k = 50$, and was kept until the end of the simulation. All cases were simulated along 400 min, and therefore 400 samples were collected. The simulation results are depicted in Figs. 5–8. The first row of each figure represents the fault detection pattern, where the ordinates are normalized to I_{TC}/I_α and therefore the control limit is 1. The three last rows of each figure constitute the basis of the diagnosis strategy based on the contribution analysis.

For the fault F.1), Fig. 5 shows a significant change in $cI_{RX}(x_5)$ for $k > 50$, while both $cI_{MX}(.)$ and $cI_{RY1}(.)$ remain close or slightly below 1. The variable x_5 ($=T_0$) was identified as dominant, with a neat prevalence on I_{RX} and smaller effects on the other subspaces.

In the case of F.2), Fig. 6 shows a rapid alarm in I_{RY1} at $k = 51$, while negligible contributions to I_{MX} and I_{RX} are observed. Around 30 min after the beginning of the fault ($k = 51$–80), the variable contributions to I_{RY1} properly indicate the root source (i.e., the bias in T). Since T is a controlled variable, the bias fault is gradually removed by effect of the closed-loop action, and the anomaly is then transferred to other variables. From $k = 80$ onwards, the PI controller propagates the anomaly to the variable F_C (x_7), which modified T and C_A, as it is notified by cI_{RY1}. Additionally, an alert in I_{MX} diagnoses a change in the operating point together with an alert in I_{RX} that indicates a correlation change in F_C (x_7) with respect to the remaining variables (i.e., $x_1 \ldots x_6$). The disturbed x_7 is correctly diagnosed by the contributions to I_{MX} and I_{RX}. The variable contributions to each component statistic exhibit transient responses until the process is settled to the new steady state.

The simulated fault F.3) corresponds to a gradual degradation of the reaction rate, as typically caused by catalyst poisoning. The process change was detected at $k \approx 170$, after collection of around 120 samples. The high levels of $cI_{RY1}(y_1)$ and $cI_{RY1}(y_3)$ in Fig. 7 indicate that both C_A and Q cannot be adequately predicted. Since no predictor variable is attributed to such changes, then it can be inferred that the anomaly is due to the modification of some parameter of the plant (k_0 in this case). Note that k_0 is related to T, C_A, and Q (see Eq. (33)), but T is regulated to SP_T by the PI controller.

Fault F.4) consists in introducing a slow drift of +0.03 K/min in T_0. This continuous temperature rise indirectly causes an increase in the inlet coolant flow rate F_C (x_7), as required to keep the reactor temperature constant at its desired value. According to Fig. 8, a correlation loss between F_C and the remaining input variables is first diagnosed by a high $cI_{RX}(x_7)$; then, a change in the operating point of F_C is diagnosed by a high $cI_{MX}(x_7)$.

The four analyzed example faults allow a direct comparison with some detection and diagnosis studies previously-reported [14,31–34]. Each of these works investigated a different (linear or nonlinear) LV technique that was applied to the evaluation of only a limited number of the described fault scenarios F.1 to F.4. More specifically, a Kernel PCA technique was applied to F.1, F.2, and F.3 [14], a steady-state fault signature technique was applied to F.1 and F.2 [31], a nonlinear multiscale modeling technique was applied to F.1, F.3, and F.4 [33], and a contribution plot without smearing effect was applied to F.1 and F.2 [34]. Besides, a technique based on Gaussian process LV models was evaluated onto the fault scenario F.2, but with purposes of fault detection only [32]. In contrast, the technique presented in this work is linear and successfully detects and diagnoses the four simulated fault scenarios. In general, it was observed that the detection sensitivity (i.e., the minimal disturbance that triggers an out-of-control of the index I_{TC}) is equivalent to the sensitivity achieved by the individual approaches of [14,31–34] in the cases of the fault scenarios F.1, F.2, and F.4. However, the present approach is more sensitive than that used in [14,33] for the fault scenario F.3, and allows adequate detection and diagnosis even under lower levels of the reaction degradation. Additionally, it should be noticed that the

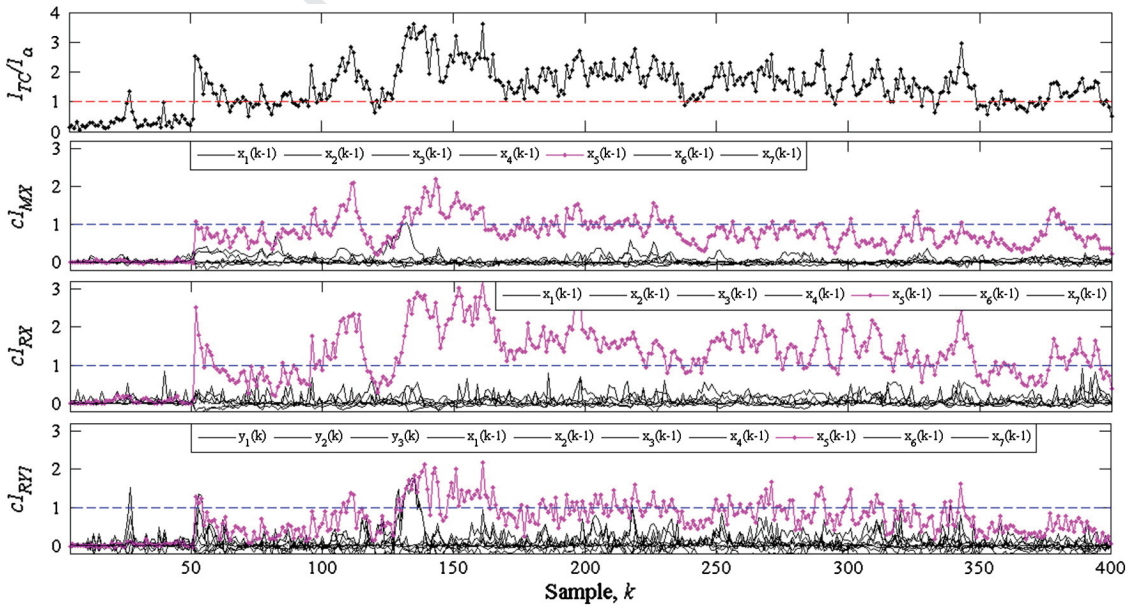

Fig. 5. Simulation of fault F.1): bias in the measurement of T_0. The contributions to the statistic I_{RX} identify the variable x_5 ($=T_0$) as the source of the fault.

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006

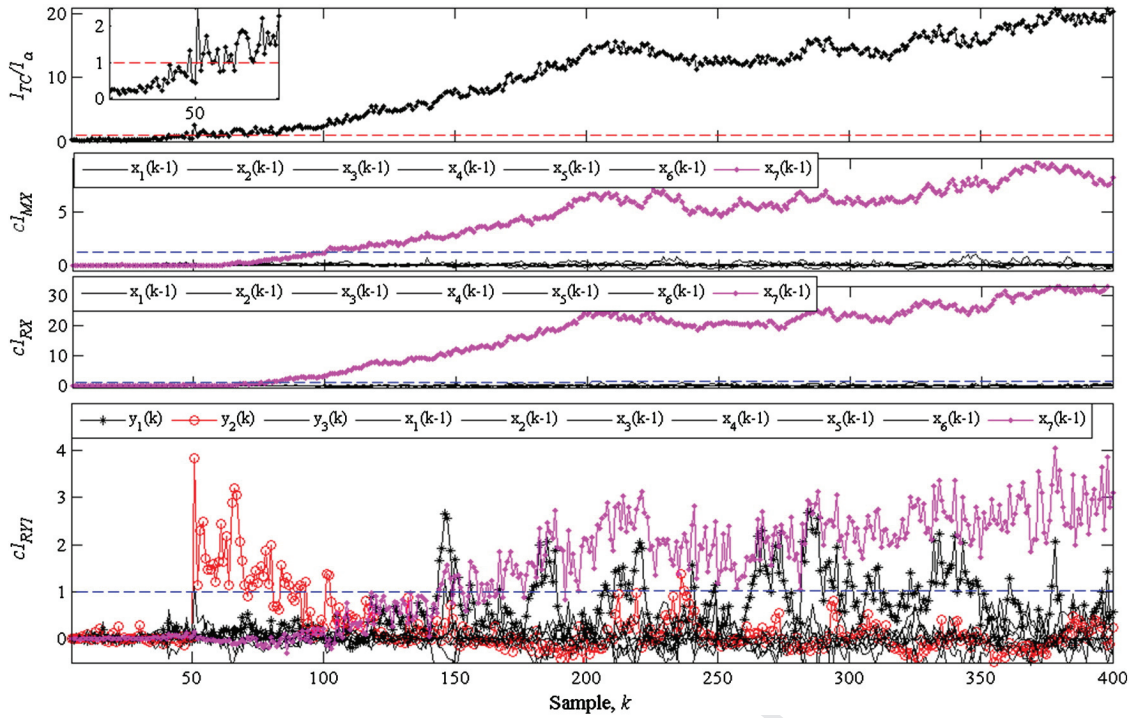

Fig. 6. Simulation of fault F.2): bias in the measurement of T. At $50 < k < 80$, the contributions to the statistic I_{RY1} identify the variable y_2 ($=T$) as the source of the fault.

present detection and diagnosis procedure, implemented through an anomaly classification step followed by a further variable identification, is only based on theoretical considerations that enabled a proper diagnosis of several classes of faults (Table 4).

5. Conclusions

The proposed monitoring technique is based on a PLSR model initially designed under 'in-control' conditions, and is especially useful for monitoring processes that exhibit collinear measurements. The fundamentals of the method are based on detecting meaningful deviations of the measurements from their expected behaviors, which in turn serve for diagnosing abnormal events. The extension of the PLSR modeling strategy consisted in a new PLS-decomposition that included more specific statistics, such as the distance to the regression model. By means of several numerical simulations, it was verified that the proposed method was effective for detecting and diagnosing different anomalies.

The proposed detection index, I_{TC}, involves a balanced merging of several scaled metrics; and it represents a statistical distance that considers the correlation structure of the process as well as three Euclidean distances to the model. The analytical expression of I_{TC} [Eq. (22)] contains an optimized combination of statistics. In fact, if a new statistics were added (e.g., $T^2_{\hat{u}}$), then the new combined index would be: $I^*_{TC} = 2I_{MX} + I_{RX} + I_{RY1} + I_{RY2} = I_{TC} + I_{MX}$. Thus, the

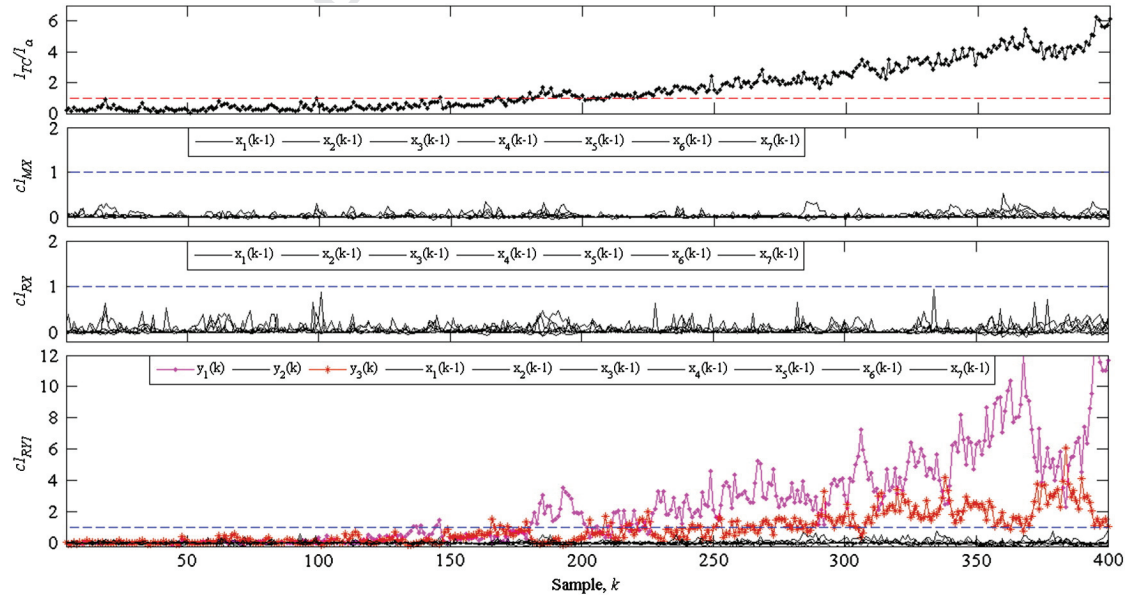

Fig. 7. Simulation of fault F.3): degradation of the reaction rate. At $k > 170$, the contributions to the statistic I_{RY1} indicate that predictions of the variables y_1 ($=C_A$) and y_3 ($=Q$) could be highly erroneous.

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006

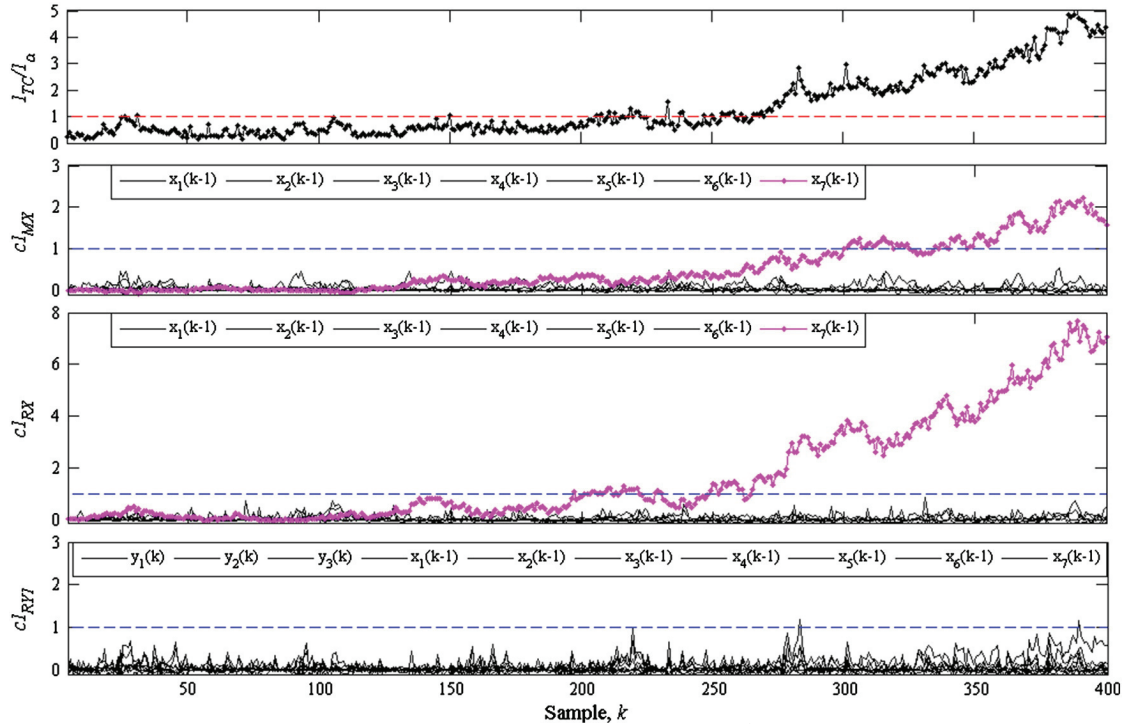

Fig. 8. Simulation of fault F.4): disturbance in T_0. At $k > 200$, the contributions to the statistic I_{RX} identify a correlation loss of the variable x_7 ($=F_C$) followed by a change in its operating point (as identified in I_{MX} from $k = 300$ onwards).

magnitude of I_{MX} would be overweighed, and therefore the information contained in the remaining three statistics will be partially hidden.

In the application examples, only small deviations with respect to the nominal values were adopted, in order to evaluate the ability of the control volume associated to I_{TC} for detecting the simulated errors. When an anomaly was present in the process, the combination of the alerts observed in the isolated statistics conforming I_{TC} was efficiently used for classifying the perturbation source. Such preliminary diagnostic was then completed through the analysis of contributions to the alarmed statistics. Isolation of the disturbed variables was feasible through a decomposition of each component statistic in their contributions. The results obtained with two simulation examples show that the proposed technique is effective and accurate enough to be exploited with more emphasis in the future.

In summary, the proposed technique exhibits the following advantages: 1) it is an effective tool to simultaneously cope with quality monitoring, detection, and diagnosis of process faults, and isolation of multiple sensor faults; 2) it is based on a linear LV model and its calibration is relatively simple because only requires data taken from the process operating under normal conditions; and 3) the detection and diagnosis method is mainly based on the theoretical considerations exposed in the first part of this work and compiled in Table 2; instead of being based on recorded patterns of historical faults, such as in the fault signature methods. A limitation of the proposed technique is that it might not be advisable when the data collected under normal operating conditions are multimodal and/or when strong nonlinear correlations are present. However, it could be extended to dynamic models (e.g., by using dynamic PLS), and also to nonlinear models (e.g., by using kernel PLS).

Acknowledgments

The authors are grateful for the financial support received from CONICET, MinCyT, Universidad Nacional del Litoral, and Universidad Tecnológica Nacional (Argentina).

Appendix A. Proofs of Lemma 1 and Theorems 1 and 2

Proof 1. The oblique projector onto *Span*{**A**} along *Span*{**B**} can be obtained through the following equation [27]:

$$\mathbf{\Pi}_{\mathbf{A}|\mathbf{B}} = \mathbf{A}\left(\mathbf{A}'\mathbf{\Pi}_{\mathbf{B}}^{\perp}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{\Pi}_{\mathbf{B}}^{\perp} \tag{A.1}$$

where $\mathbf{\Pi}_{\mathbf{B}}^{\perp}$ is the orthogonal projector onto *Span*{**B**}$^{\perp}$.

Since **R** and **S** are full column rank, then:

$$\mathbf{\Pi}_{\mathbf{R}^{\perp}}^{\perp} = \mathbf{\Pi}_{\mathbf{R}} = \mathbf{R}(\mathbf{R}'\mathbf{R})^{-1}\mathbf{R}', \tag{A.2}$$

$$\mathbf{\Pi}_{\mathbf{S}^{\perp}}^{\perp} = \mathbf{\Pi}_{\mathbf{S}} = \mathbf{S}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'. \tag{A.3}$$

Since $\mathbf{P}'\mathbf{R} = \mathbf{R}'\mathbf{P} = \mathbf{I}$ (or $\mathbf{Q}'\mathbf{S} = \mathbf{S}'\mathbf{Q} = \mathbf{I}$), then substituting Eq. (A.2) [or Eq. (A.3)] into Eq. (A.1) yield: $\mathbf{\Pi}_{\mathbf{P}|\mathbf{R}^{\perp}} = \mathbf{P}\mathbf{R}'$ (or $\mathbf{\Pi}_{\mathbf{Q}|\mathbf{S}^{\perp}} = \mathbf{Q}\mathbf{S}'$). Similarly, we have $\mathbf{\Pi}_{\mathbf{R}^{\perp}|\mathbf{P}} = \mathbf{I}-\mathbf{P}\mathbf{R}'$ (or $\mathbf{\Pi}_{\mathbf{S}^{\perp}|\mathbf{Q}} = \mathbf{I}-\mathbf{Q}\mathbf{S}'$). □

Partially, Lemma 1 has also been proved by Gang et al. [25].

Proof 2. Equations (10, 11) can be proved by taking into account that: (i) *Span*{$\mathbf{I} - \mathbf{P}\mathbf{R}'$} = *Span*{**R**}$^{\perp}$ and *Span*{$\mathbf{I} - \mathbf{Q}\mathbf{S}'$} = *Span*{**S**}$^{\perp}$ (see Lemma 1); and (ii) the projections belong to complementary subspaces, because $rank(\mathbf{P}\mathbf{R}'(\mathbf{I} - \mathbf{P}\mathbf{R}')) = \dim(S_{MX}) + \dim(S_{RX}) = m$ and $rank(\mathbf{Q}\mathbf{S}'(\mathbf{I} - \mathbf{Q}\mathbf{S}')) = \dim(S_{MY}) + \dim(S_{RY}) = p$. Eq. (12) is directly derived from Eq. (7). □

Proof 3. Substituting $\hat{\mathbf{y}} = \mathbf{Q}\hat{\mathbf{u}}$ into Eq. (21b), one obtains $D_{\hat{\mathbf{y}}} = \hat{\mathbf{u}}'\mathbf{Q}'(\mathbf{Q}\mathbf{\Delta}\mathbf{Q}')^{-}\mathbf{Q}\hat{\mathbf{u}}$. The singular value decomposition of the (full-column rank) matrix **Q** ($p \times A$) is $\mathbf{Q}' = \mathbf{V}[\mathbf{\Sigma}\ \mathbf{0}]\mathbf{W}'$, where $\mathbf{\Sigma}$ ($A \times A$) is a non-singular diagonal matrix, and **W** ($p \times p$) and **V** ($A \times A$) are orthonormal matrices. Then,

$$\begin{aligned}\mathbf{Q}'(\mathbf{Q}\mathbf{\Delta}\mathbf{Q}')^{-}\mathbf{Q} &= \mathbf{V}[\mathbf{\Sigma}\ \ \mathbf{0}]\mathbf{W}'\left(\mathbf{W}\begin{bmatrix}\mathbf{\Sigma}\\ \mathbf{0}\end{bmatrix}\mathbf{V}'\mathbf{\Delta}\mathbf{V}[\mathbf{\Sigma}\ \ \mathbf{0}]\mathbf{W}'\right)^{-}\mathbf{W}\begin{bmatrix}\mathbf{\Sigma}\\ \mathbf{0}\end{bmatrix}\mathbf{V}',\\ &= \mathbf{V}\mathbf{\Sigma}(\mathbf{\Sigma}\mathbf{V}'\mathbf{\Delta}\mathbf{V}\mathbf{\Sigma})^{-1}\mathbf{\Sigma}\mathbf{V}' = \mathbf{\Delta}^{-1}.\end{aligned} \tag{A.4}$$

Therefore, $D_{\hat{\mathbf{y}}} = \hat{\mathbf{u}}'\boldsymbol{\Delta}^{-1}\hat{\mathbf{u}} = T^2_{\hat{\mathbf{u}}}$. By combining Eqs. (18) and (15b) with $\hat{\mathbf{u}} = \mathbf{B}\mathbf{t}S$, we obtain $T^2_{\hat{\mathbf{u}}} = \mathbf{t}'\mathbf{B}\left(\mathbf{B}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{B}^{-1}\right)\mathbf{B}\mathbf{t} = T^2_{\mathbf{t}}$. By replacing $\hat{\mathbf{x}} = \mathbf{P}\mathbf{t}$ into Eq. (21a): $D_{\hat{\mathbf{x}}} = \mathbf{t}'\mathbf{P}'\left(\mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'\right)^{-}\mathbf{P}\mathbf{t}$. Similarly to the derivation of Eq. (A.4), it is obtained: $\mathbf{P}'(\mathbf{P}\boldsymbol{\Lambda}\mathbf{P}')^{-}\mathbf{P} = \boldsymbol{\Lambda}^{-1}$; hence $D_{\hat{\mathbf{x}}} = \mathbf{t}'\boldsymbol{\Lambda}^{-1}\mathbf{t} = T^2_{\mathbf{t}}$. From all these equalities, Theorem 2 is proven. □

Appendix B. Control limits

To calculate the confidence limits we assume that $\mathbf{x}$ and $\mathbf{y}$ approximately follow multivariate normal behaviors. The T^2 control limit at the α significance level is given by [5] $\tau^2_\alpha = [(AN^2 - A)/(N^2 + NA)]F_{A,N-A}(\alpha)$, where $F_{A,N\text{-}A}(.)$ is the F-distribution with A and N-A degrees of freedom. On the other hand, the α-control limit of $SPE_{\mathbf{x}}$ is given by [7] $\delta^2_{\mathbf{x},\alpha} = (\sigma^2/2\nu)\,\chi^2_{2\nu^2/\sigma^2}(\alpha)$ where ν and σ^2 are the sample average and variance estimated with N observations of $SPE_{\mathbf{x}}$, and $\chi^2_{2\nu^2/\sigma}(.)$ is the χ^2-distribution with $2\nu^2/\sigma^2$ degrees of freedom. Similarly, the control limits of $SPE_{\mathbf{y}1}$ and $SPE_{\mathbf{y}2}$ are calculated.

The 100(1-α)% control limit of the I_{TC} is given by [11]

$$I_\alpha = g\,\chi^2_h(\alpha) \tag{B.1}$$

with g and h given by:

$$g = \frac{tr\left[\left(\mathbf{R}_{[\mathbf{y}'\,\mathbf{x}']}\boldsymbol{\Phi}\right)^2\right]}{tr\left(\mathbf{R}_{[\mathbf{y}'\,\mathbf{x}']}\boldsymbol{\Phi}\right)},\quad h = \frac{\left[tr\left(\mathbf{R}_{[\mathbf{y}'\,\mathbf{x}']}\boldsymbol{\Phi}\right)\right]^2}{tr\left[\left(\mathbf{R}_{[\mathbf{y}'\,\mathbf{x}']}\boldsymbol{\Phi}\right)^2\right]}, \tag{B.2}$$

where $\mathbf{R}_{[\mathbf{y}'\,\mathbf{x}']}$ is the correlation matrix of the extended vector $[\mathbf{y}'\ \mathbf{x}']'$ and $\boldsymbol{\Phi}$ is equal to

$$\boldsymbol{\Phi} = \begin{bmatrix} \mathbf{S}\mathbf{Q}'\mathbf{Q}\mathbf{S}'/\delta^2_{\mathbf{y}1,\alpha} + (\mathbf{I}-\mathbf{Q}\mathbf{S}')'(\mathbf{I}-\mathbf{Q}\mathbf{S}')/\delta^2_{\mathbf{y}2,\alpha} & -\mathbf{S}\mathbf{Q}'\mathbf{Q}\mathbf{B}\mathbf{R}'/\delta^2_{\mathbf{y}1,\alpha} \\ -\mathbf{R}\mathbf{B}\mathbf{Q}'\mathbf{Q}\mathbf{S}'/\delta^2_{\mathbf{y}1,\alpha} & \boldsymbol{\Phi}_{22} \end{bmatrix},$$
$$\boldsymbol{\Phi}_{22} = \mathbf{R}\boldsymbol{\Lambda}^{-1}\mathbf{R}'/\tau^2_\alpha + (\mathbf{I}-\mathbf{P}\mathbf{R}')'(\mathbf{I}-\mathbf{P}\mathbf{R}')/\delta^2_{\mathbf{x},\alpha} + \mathbf{R}\mathbf{B}\mathbf{Q}'\mathbf{Q}\mathbf{B}\mathbf{R}'/\delta^2_{\mathbf{y}1,\alpha}. \tag{B.3}$$

In all cases, the control limits were set at the 99 percentile ($\alpha = 0.01$).

References

[1] S.J. Qin, Survey on data-driven industrial process monitoring and diagnosis, Annual Reviews in Control 36 (2012) 220–234.
[2] J. Kresta, J.F. MacGregor, T.E. Marlin, Multivariate statistical monitoring of process operating performance, Canadian Journal of Chemical Engineering 69 (1991) 35–39.
[3] P. Nomikos, J.F. MacGregor, Monitoring batch process using multiway principal component analysis, AIChE Journal 40 (1994) 1361–1375.
[4] P. Nomikos, J.F. MacGregor, Multivariate SPC charts for monitoring batch processes, Technometrics 37 (1995) 41–59.
[5] T. Kourti, J.F. MacGregor, Process analysis, monitoring and diagnosis using multivariate projection methods, Chemometrics and Intelligent Laboratory Systems 28 (1995) 3–21.
[6] T. Kourti, P. Nomikos, J.F. MacGregor, Analysis, monitoring and fault diagnosis of batch process using multiblock and multiway PLS, Journal of Process Control 5 (1995) 277–284.
[7] T. Kourti, Application of latent variable methods to process control and multivariate statistical process control in industry, International Journal of Adaptive Control and Signal Processing 19 (2005) 213–246.
[8] V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri, K. Yin, A review of process fault detection and diagnosis Part III: process history based methods, Computers and Chemical Engineering 27 (2003) 327–346.
[9] A.L. Pomerantsev, O.Y. Rodionova, Process analytical technology: a critical view of the chemometricians, Journal of Chemometrics 26 (2012) 299–310.
[10] C.F. Alcala, S.J. Qin, Reconstruction-based contribution for monitoring, Automatica 45 (2009) 1593–1600.
[11] C.F. Alcala, S.J. Qin, Analysis and generalization of fault diagnosis methods for process monitoring, Journal of Process Control 21 (2011) 322–330.
[12] H.H. Yue, S.J. Qin, Reconstruction-based fault identification using a combined index, Industrial and Engineering Chemistry Research 40 (2001) 4403–4414.
[13] J.M. Lee, C.K. Yoo, I.B. Lee, Statistical process monitoring with independent component analysis, Journal of Process Control 14 (2004) 467–485.
[14] C.F. Alcala, S.J. Qin, Reconstruction-based contribution for process monitoring with kernel principal component analysis, Industrial and Engineering Chemistry Research 49 (2010) 7849–7857.
[15] L. Wang, H. Shi, Multivariate statistical process monitoring using an improved independent component analysis, Chemical Engineering Research and Design 88 (2010) 403–414.
[16] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems 58 (2001) 109–130.
[17] H.A.L. Kiers, A.K. Smilde, A comparison of various methods for multivariate regression with highly collinear variables, Statistical Methods and Applications 16 (2007) 193–228.
[18] J.L. Godoy, R.J. Minari, J.R. Vega, J.L. Marchetti, Multivariate statistical monitoring of an industrial SBR Process. Soft-sensor for production and rubber quality, Chemometrics and Intelligent Laboratory Systems 107 (2011) 258–268.
[19] T. Mehmood, K.H. Liland, L. Snipen, S. Sæbø, A review of variable selection methods in Partial Least Squares Regression, Chemometrics and Intelligent Laboratory Systems 118 (2012) 62–69.
[20] J. Gabrielsson, N.-O. Lindberg, T. Lundstedt, Review: multivariate methods in pharmaceutical applications, Journal of Chemometrics 16 (2002) 141–160.
[21] J.A. Lopes, P.F. Costa, T.P. Alves, J.C. Menezes, Chemometrics in bioprocess engineering: process analytical technology (PAT) applications, Chemometrics and Intelligent Laboratory Systems 74 (2004) 269–275.
[22] B. Junker, H. Wang, Bioprocess monitoring and computer control: key roots of the current PAT initiative, Biotechnology and Bioengineering 95 (2006) 226–261.
[23] A. AlGazzawi, B. Lennox, Model predictive control monitoring using statistics, Journal of Process Control 19 (2009) 314–327.
[24] H.W. Lee, M.W. Lee, J.M. Park, Multi-scale extension of PLS algorithm for advanced on-line process monitoring, Chemometrics and Intelligent Laboratory Systems 98 (2009) 201–212.
[25] L.G. Gang, S.J. Qin, D. Zhou, Geometric properties of partial least squares for process monitoring, Automatica 46 (2010) 204–210.
[26] P. Geladi, B. Kowalski, Partial least-squares regression: a tutorial, Analytica Chimica Acta 185 (1986) 1–17.
[27] D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM, USA, 2000.
[28] I.C.F. Ipsen, C.D. Meyer, The angle between complementary subspaces, American Mathematical Monthly 102 (1995) 904–911.
[29] J.A. Westerhuis, S.P. Gurden, A.K. Smilde, Generalized contribution plots in multivariate statistical process monitoring, Chemometrics and Intelligent Laboratory Systems 51 (2000) 95–114.
[30] C.R. Alvarez, A. Brandolin, M.C. Sánchez, On the variable contributions to the D-statistic, Chemometrics and Intelligent Laboratory Systems 88 (2007) 189–196.
[31] S. Yoon, J.F. MacGregor, Fault diagnosis with multivariate statistical models part I: using steady state fault signatures, Journal of Process Control 11 (2001) 387–400.
[32] J. Serradilla, J.Q. Shi, A.J. Morris, Fault detection based on Gaussian process latent variable models, Chemometrics and Intelligent Laboratory Systems 109 (2011) 9–21.
[33] S.W. Choi, J. Morris, I. Lee, Nonlinear multiscale modelling for fault detection and identification, Chemical Engineering Science 63 (2008) 2252–2266.
[34] J. Liu, Fault diagnosis using contribution plots without smearing effect on non-faulty variables, Journal of Process Control 22 (2012) 1609–1623.

Please cite this article as: J.L. Godoy, et al., A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space, Chemometrics and Intelligent Laboratory Systems (2013), http://dx.doi.org/10.1016/j.chemolab.2013.07.006