amcs

FAULT DETECTION AND ISOLATION WITH ROBUST PRINCIPAL COMPONENT ANALYSIS

YVON THARRAULT, GILLES MOUROT, JOSÉ RAGOT, DIDIER MAQUIN

Centre de Recherche en Automatique de Nancy (CRAN)
UMR 7039, Nancy Université, CNRS
2, Avenue de la forêt de Haye, F-54 516 Vandoeuvre-lès-Nancy, France
e-mail: {yvon.tharrault,gilles.mourot}@ensem.inpl-nancy.fr
{jose.ragot,didier.maquin}@ensem.inpl-nancy.fr

Principal component analysis (PCA) is a powerful fault detection and isolation method. However, the classical PCA, which is based on the estimation of the sample mean and covariance matrix of the data, is very sensitive to outliers in the training data set. Usually robust principal component analysis is applied to remove the effect of outliers on the PCA model. In this paper, a fast two-step algorithm is proposed. First, the objective was to find an accurate estimate of the covariance matrix of the data so that a PCA model might be developed that could then be used for fault detection and isolation. A very simple estimate derived from a one-step weighted variance-covariance estimate is used (Ruiz-Gazen, 1996). This is a "local" matrix of variance which tends to emphasize the contribution of close observations in comparison with distant observations (outliers). Second, structured residuals are used for multiple fault detection and isolation. These structured residuals are based on the reconstruction principle, and the existence condition of such residuals is used to determine the detectable faults and the isolable faults. The proposed scheme avoids the combinatorial explosion of faulty scenarios related to multiple faults to be considered. Then, this procedure for outliers detection and isolation is successfully applied to an example with multiple faults.

Keywords: principal component analysis, robustness, outliers, fault detection and isolation, structured residual vector, variable reconstruction.

1. Introduction

Principal component analysis (PCA) has been applied successfully in the monitoring of complex systems (Chiang and Colegrove, 2007; Kano and Nakagawa, 2008; Harkat, Mourot and Ragot, 2006). It is a widely used method for dimensionality reduction. Indeed, PCA transforms the data to a smaller set of variables which are linear combinations of the original variables while retaining as much information as possible. In the classical approach, the first principal component corresponds to the direction in which the projected observations have the largest variance. The second component is then orthogonal to the first one and again maximizes the variance of the data points projected on it. Continuing in this way, it produces all the principal components, which correspond to the eigenvectors of the empirical covariance matrix. From a regression point of view, PCA also constructs the optimal orthogonal linear projections (in terms of the mean square error) from the eigenvectors of the data covariance matrix.

By analysing the eigenstructure of the covariance matrix of data collected under normal operating conditions, linear relations among the variables are revealed. The PCA model so obtained describes the normal process behaviour and unusual events are then detected by referencing the observed behaviour against this model. The performance of the PCA model is then based on the accurate estimation of the covariance matrix from the data, which is very sensitive to abnormal observations.

In general, the majority of the training data set is associated with normal operating conditions. The remaining data (faulty data, data obtained during shutdown or startup periods or data issued from different operating modes) are referred to as "outliers". Often the outlying observations are not incorrect but they were made under exceptional circumstances. Consequently, they may disturb the correlation structure of the "normal data" and the result will be a model that does not accurately represent the process. The fact that multiple outliers can contaminate the model derived from a classical PCA has motivated the develop-

ment of robust methods that are less affected by outliers.

In practice, one often tries to detect outliers using diagnostic tools starting from a classical fitting method. However, classical methods can be affected by outliers so strongly that the resulting fitted model does not allow one to detect the deviating observations. This is called the masking effect. Additionally, some good data points might even appear to be outliers, which is known as swamping. To avoid these effects, the goal of robust PCA methods is to obtain principal components that are not influenced much by outliers. Large residuals from that robust fit indicate the presence of outliers.

Several ways of robustifying principal components have been proposed. They can be grouped as follows.

A first group of robust PCA methods is obtained by replacing the classical covariance matrix with a robust covariance estimator, such as the minimum covariance determinant (MCD) estimator (Rousseeuw, 1987). The MCD looks for those h observations in the data set whose classical covariance matrix has the lowest possible determinant. The user-defined parameter h is the number of fault-free data among all the data and determines the robustness as well as the efficiency of the resulting estimator. The computation of the MCD estimator is non-trivial and naively requires an exhaustive investigation of all h-subsets out of the N observations. This is no longer possible for large N or in a high dimension. Rousseeuw and Van Driessen (1999) constructed a much faster algorithm called the FAST-MCD, which avoids such a complete enumeration. It is obtained by combining a basic subsampling and an iterative scheme with an MCD estimator.

A second approach to robust PCA uses projection pursuit (PP) techniques. These methods maximize a robust measure of data spread to obtain consecutive directions on which the data points are projected (Hubert, Rousseeuw and Verboven, 2002; Li and Chen, 1985; Croux and Ruiz-Gazen, 2005; Croux, Filzmoser and Oliveira, 2007). The main step of these algorithms is then to search for the direction in which the projected observations have the largest robust spread to obtain the first component. The second component is then orthogonal to the first one and has the largest robust spread of the data points projected on it. Continuing in this way produces all the robust principal components. To make these algorithms computationally feasible, the collection of directions to be investigated are restricted to all directions that pass through the robust centre and a data point or through two data points. However, the robust directions obtained are approximations of the true ones. To improve the speed of algorithms, a PCA compression to the rank of the data is performed as a first step. According to the authors, these algorithms can deal with both low and high dimensional data.

Another approach to robust PCA was proposed by Hubert *et al.* (2005) and is called ROBPCA. This method combined the ideas of both projection pursuit and robust covariance estimation based on the FAST-MCD algorithm. It first applied projection pursuit techniques in the original data space. These results are then used to project the observations into a subspace of small to moderate dimensions. Within this subspace, robust covariance estimation is applied. According to the authors, this algorithm is a powerful tool for high dimensional data when the number of variables is greater than the number of observations. The authors also used a diagnostic plot to visualize and classify the outliers. It plots the squared Mahalanobis distance versus the orthogonal distance of each observation to the PCA subspace, these indices being known as T^2 and SPE statistics, respectively, in statistical process monitoring field (Qin, 2003).

Last proposals for robust PCA include the robust LTS-subspace estimator and its generalizations (Maronna, Martin and Yohai, 2006). The idea behind these approaches consists in minimizing a robust scale of the orthogonal distances of each observation to the PCA subspace, similar to the LTS estimator, S-estimators and many others in regression. These methods are based on iterative procedures for which the problem of starting values remains open. For example, for the LTS-subspace estimator, the classical PCA is performed on the h observations with the smallest orthogonal distance to the PCA subspace. Its drawbacks are the same as for the MCD-estimator: a high computational cost, the choice of the user-defined parameter h and the starting values. Like the MCD-estimator, a FAST-LTS algorithm has been proposed.

Our presentation is devoted to the problem of sensor fault detection and isolation in data. In this paper, a fast two-step algorithm is proposed. First, a very simple estimate derived from a one-step weighted variance-covariance estimate is used (Ruiz-Gazen, 1996). Second, structured residuals are employed for multiple fault detection and isolation. These structured residuals are based on the reconstruction principle. The variable reconstruction approach assumes that each set of variables, e.g., one, two, or n variables is unknown and suggests to reconstruct these variables using the PCA model from the remaining variables (Dunia and Qin, 1998). If the faulty variables are reconstructed, the fault effect is eliminated. This property is useful for fault isolation. Moreover, instead of considering the isolation of one up to all sensors, we determine the maximum number of faulty scenarios to be taken into account by evaluating the existence condition of structured residuals. Note that this number is usually much less than the total number of sensors. The proposed scheme avoids the combinatorial explosion of faulty scenarios related to multiple faults to consider. Section 2 is a short reminder, on the one hand, of the principal component analysis in the traditional case and, on the other hand, of the robust principal component analysis. A detection and isolation

procedure for outliers is proposed in Section 3. Then, in Section 4, it is applied to an example emphasizing the generation of fault signatures.

2. PCA fault detection and isolation

Let us consider a data matrix $X \in \mathbb{R}^{N \times n}$, with a row vector $x_i^T \in \mathbb{R}^n$, which gathers N measurements collected on n system variables.

2.1. Classical approach.
In the classical PCA case, data are supposed to be collected on a system being in a normal process operation. PCA determines an optimal linear transformation of the data matrix X in terms of capturing the variation in the data:

$$T = XP \quad \text{and} \quad X = TP^T, \tag{1}$$

with $T \in \mathbb{R}^{N \times n}$ being the principal component matrix, and the matrix $P \in \mathbb{R}^{n \times n}$ contains the principal vectors which are the eigenvectors associated with the eigenvalues λ_i of the covariance matrix (or correlation matrix) Σ of X:

$$\Sigma = P\Lambda P^T \qquad \text{with} \qquad PP^T = P^T P = I_n, \tag{2}$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix with diagonal elements in decreasing magnitude order.

The relations (1) are useful when the dimension of the representation space is reduced. Once the component number ℓ to retain is determined, the data matrix X can be approximated. For that purpose, the eigenvector matrix is partitioned as follows:

$$P = \begin{pmatrix} \hat{P} & \tilde{P} \end{pmatrix}, \qquad \hat{P} \in \mathbb{R}^{n \times \ell}. \tag{3}$$

From the decomposition (1), $\hat{X}$ is the principal part of the data explained by the first ℓ eigenvectors and the residual part $\tilde{X}$ is explained by the remaining components:

$$\hat{X} = X\hat{P}\hat{P}^T = XC_\ell, \tag{4}$$

$$\tilde{X} = X - \hat{X} = X(I - C_\ell), \tag{5}$$

where the matrix $C_\ell = \hat{P}\hat{P}^T$ is not equal to the identity matrix, except for the case $\ell = n$.

Therefore, the PCA model partitions the measurement space into two orthogonal spaces:

- The principal component space, which includes data variations according to the principal component model.
- The residual space, which includes data variation not explained by the model. Such variations are due to noise and model errors in the data.

When the process variables are noise-free or corrupted with multivariate zero mean independent and identically distributed measurement noise (Li and Qin, 2001),

the eigenvectors $\tilde{P}$ of Σ corresponding to its $n - \ell$ smallest eigenvalues are such that

$$X\tilde{P} = 0. \tag{6}$$

2.2. Robust approach.
A major difficulty in PCA comes from its sensitivity to outliers. In order to reduce this sensitivity, various techniques can be applied and, in particular, that which consists in carrying out PCA directly on the data possibly contaminated by outliers. One approach is to replace the covariance matrix by its robust variant which leads to robust PCA. This seems to be a straightforward way since the principal components are the eigenvectors of the covariance matrix.

Ruiz-Gazen (1996) define a "local" matrix of variance in the sense that the suggested form tends to emphasize the contribution of close observations in comparison with distant observations (outliers). The matrix is defined in the following way:

$$T = \frac{\displaystyle\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} w_{i,j}(x_i - x_j)(x_i - x_j)^T}{\displaystyle\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} w_{i,j}}, \tag{7}$$

where the weights $w_{i,j}$ themselves are defined by

$$w_{i,j} = \exp\left(-\frac{\beta}{2}(x_i - x_j)^T \Sigma^{-1}(x_i - x_j)\right), \tag{8}$$

β being a tuning parameter to reduce the influence of the observations far away (the authors recommend a value close to 2). It is shown in the sequel that the results are not very sensitive to this parameter.

Thanks to the presence of adapted weights $w_{i,j}$, PCA can then be carried out on this "new" matrix of covariance considered robust with respect to outliers. From this new model, the detection and isolation of outliers are carried out using the reconstruction principle. Then a weight of zeros is applied on the outliers, and a classical covariance matrix is calculated. The PCA is then constructed from fault-free data. To ensure the detection of all outliers, a fault detection procedure can then be applied again.

3. Fault detection and isolation

The variable reconstruction approach assumes that each variable may be faulty (in the case of a single fault) and suggests to reconstruct the assumed faulty variable using the PCA model from the remaining variables (Dunia and Qin, 1998). This reconstructed variable is then used to detect and isolate the faults. Moreover, this principle allows us to determine replacement values for the faulty variables.

3.1. Data reconstruction. The PCA model being known according to (4) and (5), a new measurement vector x can be decomposed as

$$x = \hat{x} + \tilde{x}, \quad \hat{x} = C_\ell\, x, \quad \tilde{x} = (I - C_\ell)\, x, \quad (9)$$

where $\hat{x}$ and $\tilde{x}$ are respectively the projections of x onto the principal space and residual space.

From (9), it is possible to estimate a part of the vector x, for example, the subset R containing the indices of r reconstructed variables. However, the presence of outliers in the observation vector x returns the estimated $\hat{x}$ sensitive to these values. It is then preferable to express this estimated $\hat{x}$ by using only the fault-free part of the observation vector x.

The reconstruction of variables consists in estimating the reconstructed vector $\hat{x}_R$ by eliminating the effect of the faults. Matrix Ξ_R indicates the reconstruction directions. This matrix is orthonormal with dimension $(n \times r)$ and is built with 0 and 1, where 1 indicates the reconstructed variables from the other variables (with 0). For example, to reconstruct variables $R = \{2, 4\}$ among five variables, matrix Ξ_R is formed as follows:

$$\Xi_R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}^T.$$

The expression for the reconstruction $\hat{x}_R$ of the vector x is given by

$$\hat{x}_R = G_R\, x, \quad (10)$$

where

$$\tilde{\Xi}_R = (I - C_\ell)\, \Xi_R, \quad (11)$$
$$G_R = I - \Xi_R (\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1} \tilde{\Xi}_R^T. \quad (12)$$

Reconstruction condition. Let us note that if $\tilde{\Xi}_R$ has full column rank, then $(\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1}$ exists and the variables of the subset R are completely reconstructible. This condition implies that the number of reconstructed variables r must satisfy

$$r \le n - \ell \quad (13)$$

and that the columns of matrix $\tilde{\Xi}_R$ are neither null nor collinear.

If we write $\hat{x}_R$ in the case where the matrix of the reconstruction directions is reorganized as follows:

$$\Xi_R = \begin{bmatrix} I_1 \\ (r \times r) \end{bmatrix} \begin{matrix} 0 \\ ((n-r) \times r) \end{matrix} \Big]^T \in \mathbb{R}^{n \times r} \quad (14)$$

with $I_1 \in \mathbb{R}^{r \times r}$ being the identity matrix, then C_ℓ is splitted in four parts:

$$C_\ell = \begin{bmatrix} \underset{(r \times r)}{c_1} & \underset{(r \times (n-r))}{c_2} \\ \underset{((n-r) \times r)}{c_2^T} & \underset{((n-r) \times (n-r))}{c_4} \end{bmatrix} \in \mathbb{R}^{n \times n}. \quad (15)$$

The reconstruction $\hat{x}_R$ of the vector x is written as follows:

$$\hat{x}_R = \begin{bmatrix} 0 & (I_1 - c_1)^{-1} c_2 \\ 0 & I_2 \end{bmatrix} x, \quad (16)$$

with $I_2 \in \mathbb{R}^{n-r \times n-r}$ being the identity matrix.

This form highlights two characteristics. First, the reconstructed vector $\hat{x}_R$ is formed of the r reconstructed variables and a copy of the remaining $n - r$ variables. Second, the reconstructed variables are estimated without using their own measurement.

Proof. Let us note that the matrix C_ℓ is an idempotent and symetric matrix. From (14) and (15), (11) becomes $\tilde{\Xi}_R^T = \begin{bmatrix} I_1 - c_1 & -c_2 \end{bmatrix}$. From (11) we get

$$\tilde{\Xi}_R^T \tilde{\Xi}_R = \Xi_R^T (I - C_\ell)\, \Xi_R = I_1 - \Xi_R^T C_\ell \Xi_R.$$

As $\Xi_R^T C_\ell \Xi_R = c_1$, we have

$$\left(\tilde{\Xi}_R^T \tilde{\Xi}_R \right)^{-1} = (I_1 - c_1)^{-1}.$$

These different terms are replaced in (12):

$$G_R = \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} - \begin{bmatrix} I_1 \\ 0 \end{bmatrix} \begin{bmatrix} I_1 & -(I - c_1)^{-1} c_2 \end{bmatrix}$$
$$= \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} - \begin{bmatrix} I_1 & -(I - c_1)^{-1} c_2 \\ 0 & 0 \end{bmatrix}.$$

■

3.2. Structured residual generation. With a diagnosis objective in mind, residuals are generated for fault detection and isolation. The residuals are obtained by projecting the reconstructed variables onto the residual space. Residuals are defined by $\tilde{x}_R$, the projection of $\hat{x}_R$ onto the residual space:

$$\tilde{x}_R = (I - C_\ell)\, \hat{x}_R$$
$$= (I - C_\ell)\, G_R\, x = P_R^{(\ell)} x, \quad (17)$$

where

$$P_R^{(\ell)} = (I - C_\ell)\, G_R \quad (18)$$
$$= (I - C_\ell) - \tilde{\Xi}_R (\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1} \tilde{\Xi}_R^T. \quad (19)$$

Property 1. *Matrix $P_R^{(\ell)}$ has the following property:*

$$P_R^{(\ell)} \Xi_R = 0. \quad (20)$$

This means that the components of $\tilde{x}_R$ are not sensitive to the components of x belonging to the subset R. This property can be used to identify which components of x are disturbed by faults.

Proof. From the matrix partition adopted above, (18) becomes

$$P_R^{(\ell)} = \begin{bmatrix} I_1 - c_1 & -c_2 \\ -c_2^T & I_2 - c_4 \end{bmatrix} \begin{bmatrix} 0 & (I_1 - c_1)^{-1} c_2 \\ 0 & I_2 \end{bmatrix}$$

$$= \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & -c_2^T (I_1 - c_1)^{-1} c_2 + I_2 - c_4 \end{array} \right]. \quad (21)$$

■

Example 1. Consider a measurement x composed of the true value x^*, a noise ϵ with zero mean and one fault of amplitude d and direction Ξ_F, where F is a subset containing the indices of the fault directions:

$$x = x^* + \epsilon + \Xi_F d. \quad (22)$$

Then the residual is

$$\tilde{x}_R = P_R^{(\ell)}(x^* + \epsilon + \Xi_F d) = P_R^{(\ell)}(\epsilon + \Xi_F d), \quad (23)$$

with $P_r^{(\ell)} x^* = 0$, cf. (6), and its expected value is

$$E(\tilde{x}_R) = P_R^{(\ell)} \Xi_F d. \quad (24)$$

The following observations can be made:

- If the reconstruction directions Ξ_R are the same as the fault directions, i.e., if $R = F$, then all components of the vector $P_R^{(\ell)} \Xi_F$ are zero and $E(\tilde{x}_R) = 0$.

- If the reconstruction directions Ξ_R are different from the fault directions, then all components of the vector $P_R^{(\ell)} \Xi_F$ are *a priori* nonzero except the components belonging to the subset R.

The analysis of the residual amplitudes $\tilde{x}_R$ for all possible combinations shows the presence of faults and makes it possible to determine the components of the measurement vector affected by this fault. ♦

Property 2. *If $\tilde{\Xi}_F$ has full column rank, there is no loss in sensitivity for some fault component.*

Proof. If some components of $\tilde{\Xi}_F$ and $\tilde{\Xi}_R$ are orthogonal, then

$$\tilde{\Xi}_R^T \tilde{\Xi}_F = \left[\begin{array}{c|c} 0 & \times \\ (r \times k) & (r \times r-k) \end{array} \right], \quad (25)$$

with $\times$ being a matrix without null data.

From (24) it follows that

$$E(\tilde{x}_R) = ((I - C_\ell) - \tilde{\Xi}_R(\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1}\tilde{\Xi}_R^T)\Xi_F d \quad (26)$$

$$= \left(\tilde{\Xi}_F - \left[\begin{array}{c|c} 0 & \times \\ (n \times k) & (n \times r-k) \end{array} \right] \right) d. \quad (27)$$

To suppress some directions of the fault, the first k columns of the matrix $\tilde{\Xi}_F$ have to be zero. But in this case, $\tilde{\Xi}_F$ is rank-deficient and the faults are not detectable. Therefore if the directions $\tilde{\Xi}_F$ and $\tilde{\Xi}_R$ are orthogonal, this does not disturb the localization process. ■

3.3. Fault isolation. The structural condition for fault isolation is as follows:

$$r \leq n - \ell - 1. \quad (28)$$

All the directions of reconstruction have to be explored for fault detection and isolation. Solutions for which the faults associated with the reconstruction directions are not detectable are useless. The number of possible reconstructions can then be reduced, and the detectable faults are defined.

The maximum reconstruction number can be calculated as follows:

$$\sum_{r=1}^{n-\ell} \mathbb{C}_n^r, \quad (29)$$

with $\mathbb{C}_n^r$ denoting the number of combinations of n elements taken r at a time. This number takes into account only the number of reconstructions and not the amplitude of the projection of the reconstructed directions onto the residual space. It can be reduced when the matrix of projected fault directions is rank-deficient or near rank-deficient. To detect these cases, the condition number ('$Rcond$'), defined as the ratio between the smallest singular value and the greatest singular value of the matrix $\tilde{\Xi}_R$, is used:

$$Rcond = \frac{\min\left(\sigma(\tilde{\Xi}_R)\right)}{\max\left(\sigma(\tilde{\Xi}_R)\right)}. \quad (30)$$

For the near rank-deficient case, fault detection and localization are possible only if its amplitude is huge. In the following, faults with huge amplitude are not considered as realistic.

The process to detect useful directions of reconstruction can be summarized as follows:

1. $r = 1$ (single-fault): compute all available directions $\tilde{\Xi}_R$. If $\tilde{\Xi}_R^T \tilde{\Xi}_R$ is close to zero, this means that the fault is not projected onto the residual space and then not detectable. To detect and localize this fault, its projection onto the principal space can be used.

2. $r = r + 1$: for all available directions $\tilde{\Xi}_R$ compute the values of the condition number $Rcond$.

   - If $Rcond$ is close to zero, then the r faulty variables of the subset R are not detectable. Therefore, all combinations which take into account these r variables will not be detectable. So they are useless. Moreover, all the combinations of $r - 1$ variables among the variables of the subset R are only detectable because their

fault signatures are identical. Then, it is useful to reconstruct only one combination of these $r - 1$ variables selected from these r variables. Moreover, all the combinations of $r - 2$ variables among the r variables of the subset R are isolable.

- If $Rcond$ is not close to zero, then all the combinations of $r - 1$ variables selected from the r variables of the subset R are isolable.

3. While $r \leq n - \ell$, go to Step 2.

This analysis of the model structure allows us to determine the detectable and isolable faults. The number of useful reconstructions can then be reduced.

4. Numerical example

4.1. Single fault case. A simple example based on four variables (x_1, x_2, x_3 and x_4) and two models is used as a device to illustrate the above ideas.

4.1.1. Data generation. The data matrix X includes $N = 240$ measurements defined in the following way:

$$x_{i,1} = v_i^2 + 1 + \sin(0.1i), \quad v_i \sim \mathcal{N}(0,1), \quad (31)$$
$$x_{i,2} = x_{i,1}, \quad x_{i,3} = -2x_{i,1}, \quad x_{i,4} \sim \mathcal{N}(0,1).$$

Realizations of centred normal distributions with the same standard deviation equal to 0.02 are added to the first three variables. A constant bias of the amplitude equal to 30% of the amplitudes of the variables simulates the presence of outliers δx_1, δx_2, δx_3 affecting the variables x_1, x_2 and x_3, from 24 to 44 for the variable x_1, from 80 to 100 for the variable x_2, from 140 to 160 for the variable x_3, respectively. This bias amplitude is important to highlight the robust PCA. It is important to notice that 60 observations contained abnormal values, and hence 25 percent of the data are contaminated by these values. The objective is to detect and especially isolate them.

Using raw data (Fig. 1), we established a robust PCA model by applying the propositions of Section 2.2. The analysis of the decrease in the standardized eigenvalues of the covariance matrix (92.93, 6.96, 0.06, 0.04) allows us to retain two principal components ($\ell = 2$).

4.1.2. Useful reconstruction. From (29), the maximum number of reconstructions is 10. Now, the direction projections onto the residual space are studied (Table 1). From this table, let us note that the last variable x_4 is not reconstructible. Then, it is not possible to detect a fault on x_4 in the residual space. A solution is to work in the principal space to detect the fault. Then for $r = 2$, the indicator $Rcond$ is computed for all useful directions ($R = \{1, 2\}$, $R = \{1, 3\}$ and $R = \{2, 3\}$).

Table 1. Existence condition of residuals.

$\tilde{\Xi}_1^T \tilde{\Xi}_1$	$\tilde{\Xi}_2^T \tilde{\Xi}_2$	$\tilde{\Xi}_3^T \tilde{\Xi}_3$	$\tilde{\Xi}_4^T \tilde{\Xi}_4$
0.83	0.83	0.33	**0.00**

Table 2. Existence condition of residuals.

R	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$
$Rcond$	0.82	0.41	0.41

The result is shown in Table 2. All the $Rcond$ values are not close to zero. This means that a fault on x_1, x_2 or x_3 is isolable. From 10 reconstruction possibilities, only 6 are really reconstructible. Among these reconstructible directions, only 3 combinations are useful to isolate the faulty variables. To conclude, in the following, only the first three variables are reconstructed to ensure localization.

Table 3. Fault occurrences.

	$R = \{1\}$				$R = \{2\}$				$R = \{3\}$			
	$\tilde{x}_{11}$	$\tilde{x}_{12}$	$\tilde{x}_{13}$	$\tilde{x}_{14}$	$\tilde{x}_{21}$	$\tilde{x}_{22}$	$\tilde{x}_{23}$	$\tilde{x}_{24}$	$\tilde{x}_{31}$	$\tilde{x}_{32}$	$\tilde{x}_{33}$	$\tilde{x}_{34}$
δx_1	0	0	0	0	$\times$	0	$\times$	$\times$	$\times$	$\times$	0	$\times$
δx_2	0	$\times$	$\times$	$\times$	0	0	0	0	$\times$	$\times$	0	$\times$
δx_3	0	$\times$	$\times$	$\times$	$\times$	0	$\times$	$\times$	0	0	0	0

4.1.3. Sensitivity analysis. The data in Table 3 summarize the relationship between the residual sensitivity $\tilde{x}_R$ (17) and the outliers or faults δx_1, δx_2 and δx_3 (the fault δx_4 on the variable x_4 is not considered). This table was constructed while taking into account the property (20) of the matrix $P_R^{(2)}$. For example, the first four residuals, $\tilde{x}_{11}$ to $\tilde{x}_{14}$ (relative to the variables x_1, x_2, x_3 and x_4), were obtained by projecting onto the residual space the reconstructed variables without using the variable x_1. As the first row and the first column of $P_1^{(2)}$ are zero, according to (20), the residual $\tilde{x}_{11}$ is not sensitive to the variables x_1, x_2 and x_3 and, consequently, to the potential faults δx_1, δx_2 or δx_3 affecting these variables. Moreover, the residuals $\tilde{x}_{12}$, $\tilde{x}_{13}$ and $\tilde{x}_{14}$ are not sensitive to the variable x_1 and thus to the fault δx_1 which can affect them. To summarize these different situations, the symbols $\times$ and 0 denote the fault influence on the residuals. The other parts of the table were constructed with this same principle by considering the different projection matrices $P_2^{(2)}$ and $P_3^{(2)}$. By analysing the dependence of the columns of the signature matrix, one can establish necessary conditions enabling fault detection and isolation.

Let us note that only two projection matrices and two residuals are necessary for fault detection and isolation. For example, the matrices $P_1^{(2)}$ and $P_2^{(2)}$ (19) allow us to build the residuals $\tilde{x}_{12}$ (relative to x_2), $\tilde{x}_{21}$ (relative to
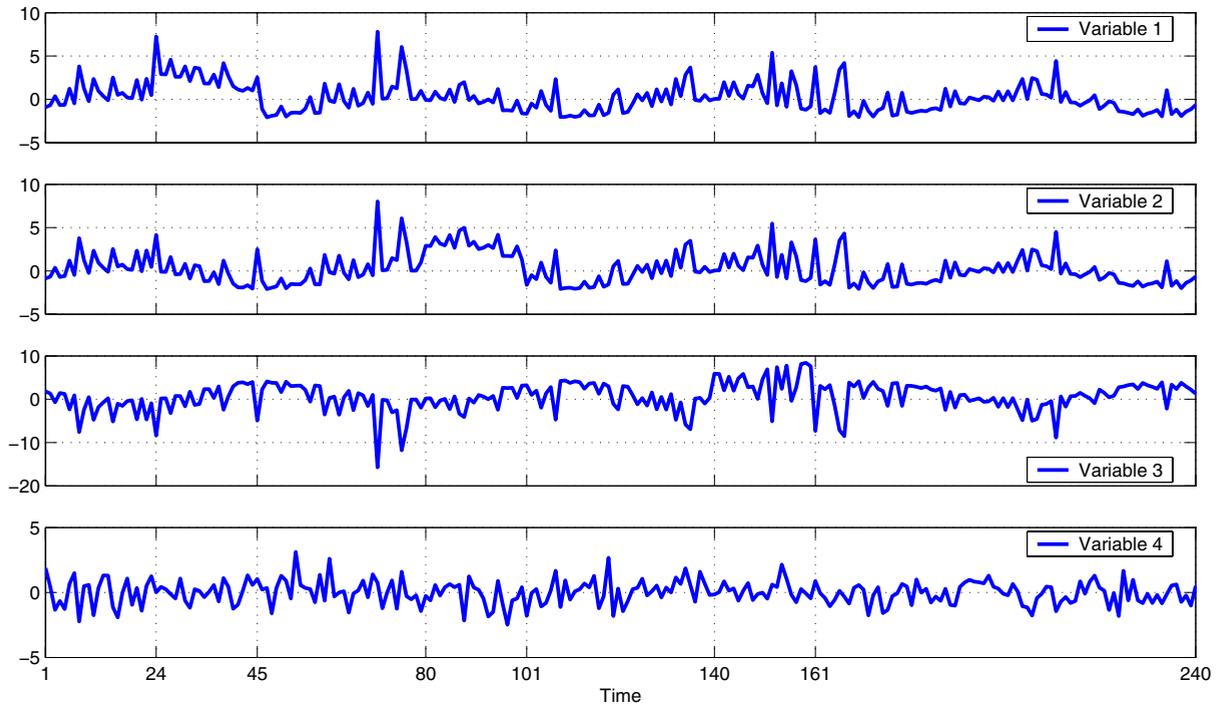
Fig. 1. Raw data.
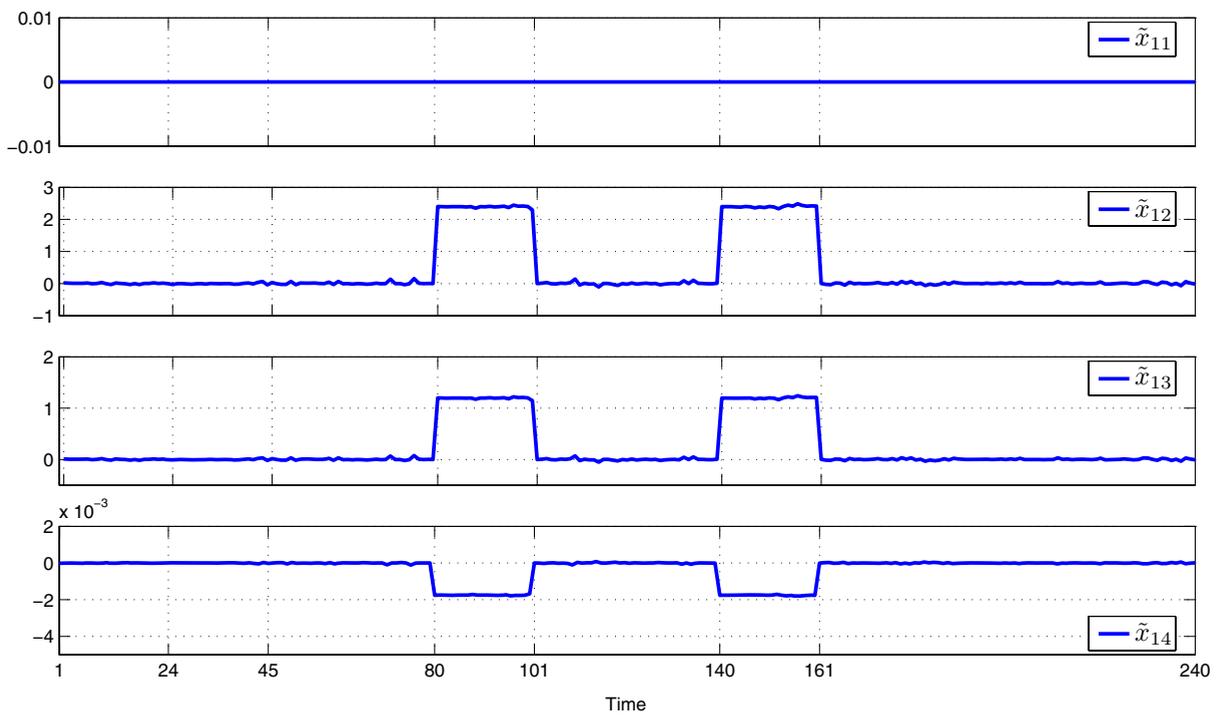
Fig. 2. Residuals $\tilde{x}_{11}$, $\tilde{x}_{12}$, $\tilde{x}_{13}$ and $\tilde{x}_{14}$.

x_1), which permit to detect and isolate one of the three faults. Indeed, Table 3 indicates that with these two residuals, the signature faults δx_1, δx_2 and δx_3 are respectively $(0 \quad \times)$, $(\times \quad 0)$ and $(\times \quad \times)$; these three signa-

tures are independent and thus the faults are isolable from each other. However, this condition is only structural and does not take into account the sensitivity of the residuals to the fault. Indeed, the residuals do not have the same sen-

sitivity to different faults. Therefore, if only two residuals are used, it is not sure if the faulty variable is detectable and localizable.

4.1.4. Fault detection.

Figure 2 shows the residuals $\tilde{x}_{11}$, $\tilde{x}_{12}$, $\tilde{x}_{13}$ and $\tilde{x}_{14}$ (relative to x_1, x_2, x_3 and x_4), defined with $\ell = 2$ by (17), and obtained by the projection of all reconstructed variables without using the variable x_1. As indicated by Property 1, in the case of the residual generated without using the variable x_1, only faults affecting the variables x_2 and x_3 are detectable on the residuals $\tilde{x}_{12}$, $\tilde{x}_{13}$ and $\tilde{x}_{14}$. Panels of Fig. 3 are relative to a global indicator SPE_R (norm of the projection vector) calculated for each observation from time 1 to time 240:

$$SPE_R = \|\tilde{x}_R\|^2. \tag{32}$$

The faulty observations are determined as follows:

$$SPE_R > \delta_\alpha^2, \tag{33}$$

with δ_α^2 being the detection threshold of SPE_R (Jackson and Mudholkar, 1979).

Table 4. Table of fault signatures.

	SPE_1	SPE_2	SPE_3
δx_1	0	×	×
δx_2	×	0	×
δx_3	×	×	0

A global indicator is computed for each direction ($R = \{1\}$, $R = \{2\}$ and $R = \{3\}$). It takes into account their corresponding residuals, e.g., SPE_1 is computed from $\tilde{x}_{11}$, $\tilde{x}_{12}$, $\tilde{x}_{13}$ and $\tilde{x}_{14}$. Then the global indicators use all the sensitivity of the residuals to the fault. Table 4 shows the fault signatures with the global indicators and is constructed in the same way as Table 3.

Fault detection and isolation are realized without ambiguity and are in accordance with the theoretical results of the isolation procedure (Table 4).

4.1.5. Choice of β.

For different values of β, Fig. 4 represents the indicator SPE_1. For $\beta = 0$, we have the same case as that of the classical PCA. We notice that it is not possible to detect or isolate the fault. For $\beta = 0.5$ to 5, we can detect and isolate the fault with a similar result. So there is a large range of values for β where the detection ability is identical. Then, we recommend to choose $\beta = 2$.

4.2. Multiple fault case.

We consider here the situation in which several faults affect variables at the same time.

4.2.1. Data generation.

The matrix X includes $N = 108$ observations of a vector x with 8 components generated in the following way:

$$x_{i,1} = v_i^2 + \sin(0.1i), \tag{34}$$
$$x_{i,2} = 2\sin(i/6)\cos(i/4)\exp(-i/N), \quad v_i \sim \mathcal{N}(0,1),$$
$$x_{i,3} = \log(x_{i,2}^2), \quad x_{i,4} = x_{i,1} + x_{i,2},$$
$$x_{i,5} = x_{i,1} - x_{i,2}, \quad x_{i,6} = 2x_{i,1} + x_{i,2},$$
$$x_{i,7} = x_{i,1} + x_{i,3}, \quad x_{i,8} \sim \mathcal{N}(0,1).$$

To the data thus generated were added realizations of random variables with centred normal distributions and standard deviations equal to 0.02 as well as the faults δx_1, δx_2, δx_3, δx_4, δx_5, δx_6 represented by a bias of the amplitude equal to 10% of the amplitudes of the variables and defined in the following way: observations from 10 to 24 for the variable x_1, observations from 35 to 49 for the variables x_2 and x_3, observations from 60 to 74 for the variables x_4 and x_5, observations from 85 to 99 for the variable x_1, x_4 and x_6. In the following, these four intervals are indicated by I_1, I_2, I_3, I_4.

From the contaminated data, the robust *PCA* model, with four principal axes ($\ell = 4$), was chosen.

Table 5. Existence condition of residuals.

$\tilde{\Xi}_1^T\tilde{\Xi}_1$	$\tilde{\Xi}_2^T\tilde{\Xi}_2$	$\tilde{\Xi}_3^T\tilde{\Xi}_3$	$\tilde{\Xi}_4^T\tilde{\Xi}_4$	$\tilde{\Xi}_5^T\tilde{\Xi}_5$	$\tilde{\Xi}_6^T\tilde{\Xi}_6$	$\tilde{\Xi}_7^T\tilde{\Xi}_7$	$\tilde{\Xi}_8^T\tilde{\Xi}_8$
0.84	0.72	0.46	0.71	0.41	0.40	0.46	**0.00**

4.2.2. Useful reconstruction.

From the size of the residual space, we cannot reconstruct more than four variables simultaneously. The maximum number of reconstructions is then equal to 162 (cf. (29)). Now, the projections of fault directions onto the residual space are studied (Table 5). From this table, let us note that the last variable is not reconstructible. Indeed, the variable x_8 is uncorrelated with the other variables. To detect and isolate a fault on this variable, it is better to work in the principal space. To analyze multiple fault directions, the indicator $Rcond$ is calculated for all available directions (for $r = 2$ to 4). Results for the reconstruction of two variables are shown in Table 6. For example, for $R = \{1, 2\}$, the value of $Rcond$ is the intersection of the first row and the second column of Table 6.

Let us notice that for $R = \{3, 7\}$, $Rcond$ is close to zero. This means that the fault signatures for a fault on x_3 or on x_7 are identical ($SPE_3 = SPE_7$). So we can only detect the fault and conclude that the variables x_3 or x_7 or x_3 and x_7 are faulty. To illustrate this case, Fig. 5 shows the values of the global indicator SPE_R for $R = \{3\}$, $R = \{7\}$, $R = \{2,3\}$ and $R = \{2,7\}$. This figure shows
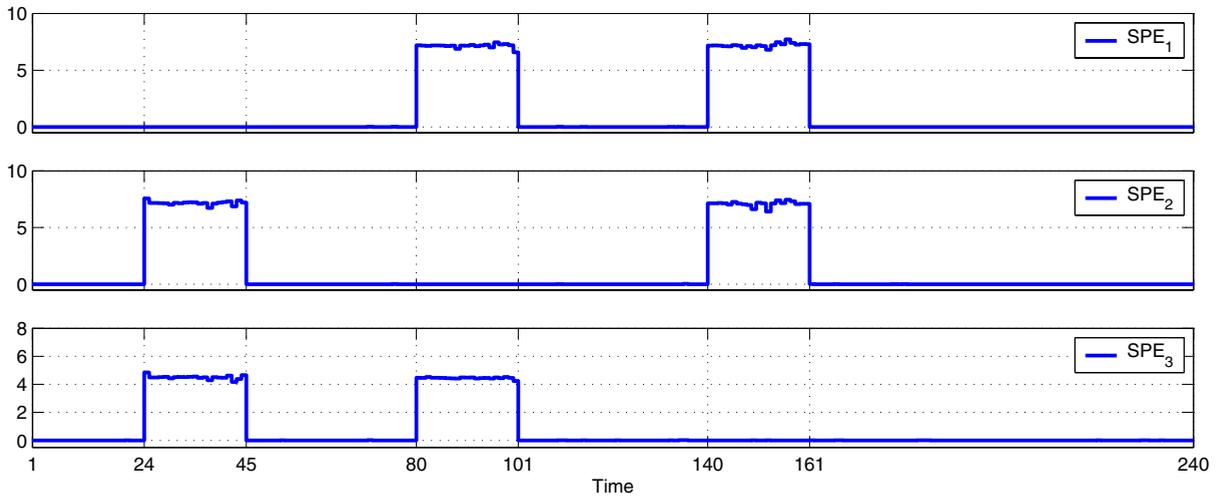
Fig. 3. SPE without using x_1, x_2 and x_3.
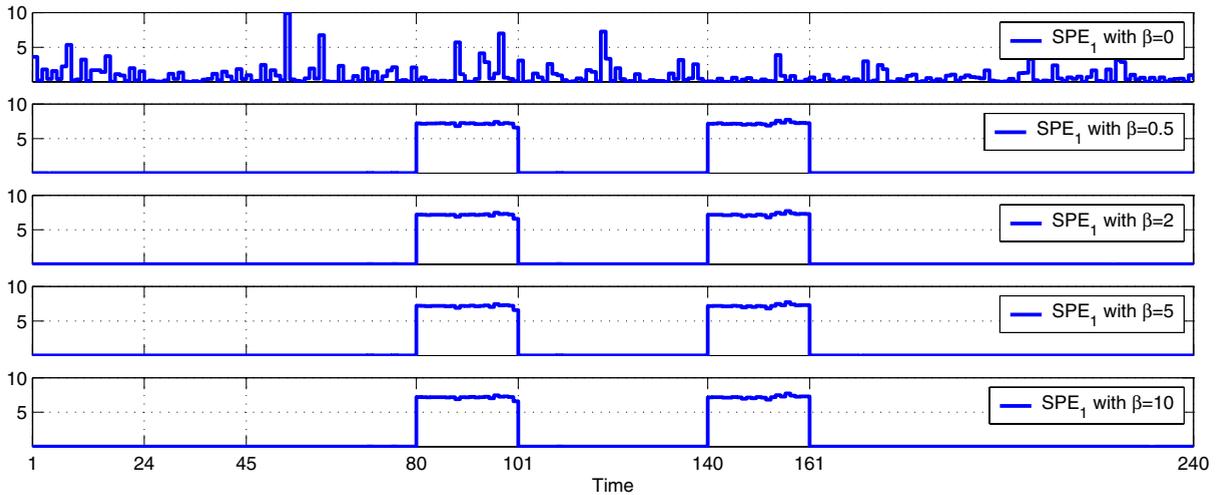
Fig. 4. SPE_1 for different values of β.

Table 7. Table of fault signatures.

	SPE_1	SPE_2	SPE_3	SPE_4	SPE_5	SPE_6	SPE_{12}	SPE_{14}	SPE_{15}	SPE_{16}	SPE_{24}	SPE_{25}	SPE_{26}	SPE_{45}	SPE_{46}	SPE_{56}	SPE_{124}	SPE_{125}	SPE_{126}	SPE_{145}	SPE_{146}	SPE_{156}
δ_1	0	×	×	×	×	×	0	0	0	0	×	×	×	×	×	×	0	0	0	0	0	0
δ_2	×	0	×	×	×	×	0	×	×	×	0	0	0	×	×	×	0	0	0	×	×	×
δ_3	×	×	0	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
δ_4	×	×	×	0	×	×	×	0	×	×	0	×	×	0	0	×	0	×	×	0	0	×
δ_{12}	×	×	×	×	×	×	0	×	×	×	×	×	×	×	×	×	0	0	0	×	×	×
δ_{14}	×	×	×	×	×	×	×	0	×	×	×	×	×	×	×	×	0	×	×	0	0	×
δ_{24}	×	×	×	×	×	×	×	×	×	×	0	×	×	×	×	×	0	×	×	×	×	×
δ_{124}	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	0	×	×	×	×	×
δ_{125}	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	0	×	×	×	×

that $SPE_3 = SPE_7$ and $SPE_{23} = SPE_{27}$, and then only one SPE of each combination is useful to detect a fault. In the following, only the combinations with 3 are consid-

ered (not with 7). On the interval I_2, SPE_3 is nonzero, so there is another fault at the same time. Moreover, this figure shows that for SPE_{23}, on the interval I_2, of SPE is
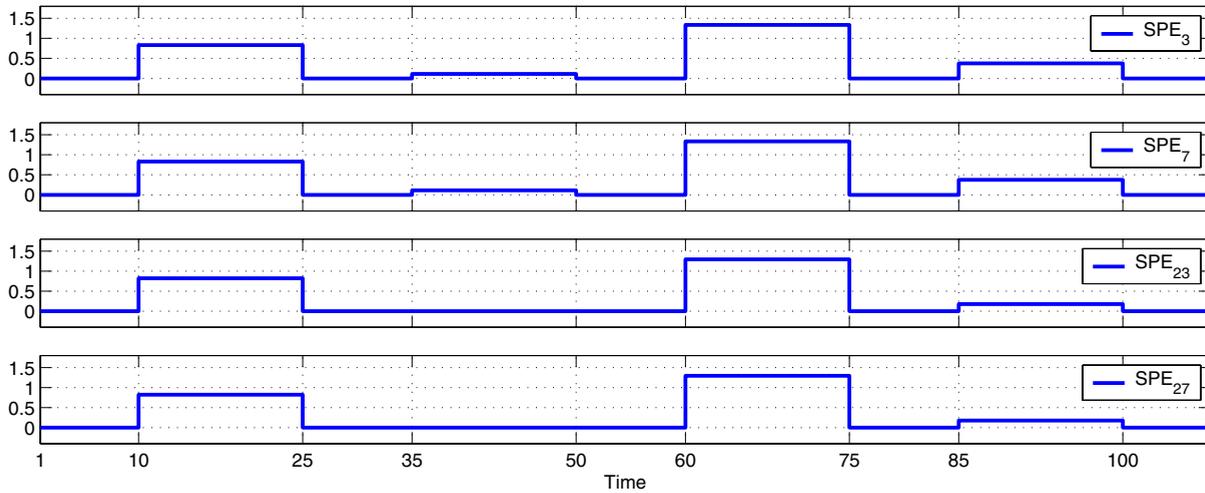
Fig. 5. SPE for $R = 2, 3$ and $R = 2, 7$.

Table 6. Existence condition of residuals.

$Rcond$	1	2	3	4	5	6	7
1		0.88	0.72	0.88	0.57	0.57	0.72
2			0.79	0.73	0.42	0.68	0.80
3				0.80	0.75	0.76	**0.01**
4					0.68	0.41	0.80
5						0.79	0.75
6							0.75

zero. We conclude that on the interval I_2, the variables x_2 and x_3 or/and x_7 are faulty.

For all the directions of reconstruction these indicators are calculated. Another case where $Rcond$ is close to zero with $R = \{2, 4, 5, 6\}$ is detected. Then all the combinations of three variables selected from among the variables of the subset R are only detectable and their fault signatures are identical ($SPE_{245} = SPE_{246} = SPE_{256} = SPE_{456}$). Therefore only one SPE is useful to detect a fault, e.g., SPE_{245}. Thus, a fault can be detected but not isolated. The faulty variables are among the variables x_2, x_4, x_5 and x_6.

From among 162 reconstruction possibilities, only 81 are really reconstructible. Among these reconstructible directions, only 21 combinations are useful to isolate the faulty variables. For the others, a set of variables is considered as faulty but it is not possible to determine the faulty variables in the set.

4.2.3. Sensitivity analysis. All the useful reconstruction directions and the corresponding SPE are computed. Concerning the *a priori* analysis of fault isolation, a reduced table of signatures established from the properties (20) is given (Table 7). It reveals only some possible faults, as noted by δ in the first row, those affecting

the variables 1, 2, 3, 4 and those affecting the couples of variables $\{1, 2\}, \{1, 4\}, \{2, 4\}, \{1, 2, 4\}, \{1, 2, 5\}$. The columns pertain to the norm SPE_R of the residual vectors obtained by the reconstruction-projection of the variables by using all the components of x except those with the indices belonging to R. The residuals are defined by (17).

This table, which the reader will be able to extend, provides a correspondence between the symptoms SPE_R and the faults δ_R. For example, the fault δ_2 affects all projections except those established without components 2, $\{1, 2\}, \{2, 4\}, \{2, 5\}, \{2, 6\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 2, 6\}$.

4.2.4. Fault detection. The reconstruction is carried out from all useful directions. Figure 6 visualizes the reconstruction of variables without using Variable 1. This figure shows the reconstruction of the first seven variables which are associated with the column SPE_1 of Table 7 specifying the isolable faults. The N reconstructed data were then projected onto the residual space. For each observation, fault indicators SPE_R were calculated.

Let us analyze Fig. 6. Variable 1, biased for the observations of the interval I_1, is not used for the reconstruction and the other variables which are used for the reconstruction do not present any bias. For these observations, the reconstructions are thus correct, emphasizing the first graph (starting from the top of the figure), which shows the superposition of the reconstructed variables (the symbol '○') with the true variables (in practice, the latter are unknown, but at the stage where the data are generated the comparison is possible). The measurement of the variable is also indicative (continuous line) in order to compare it with the reconstruction.

This result is confirmed by the first graph of Fig. 7, where the norm of the residual vector (32) is depicted. For the observations of the interval I_1, this norm close to zero
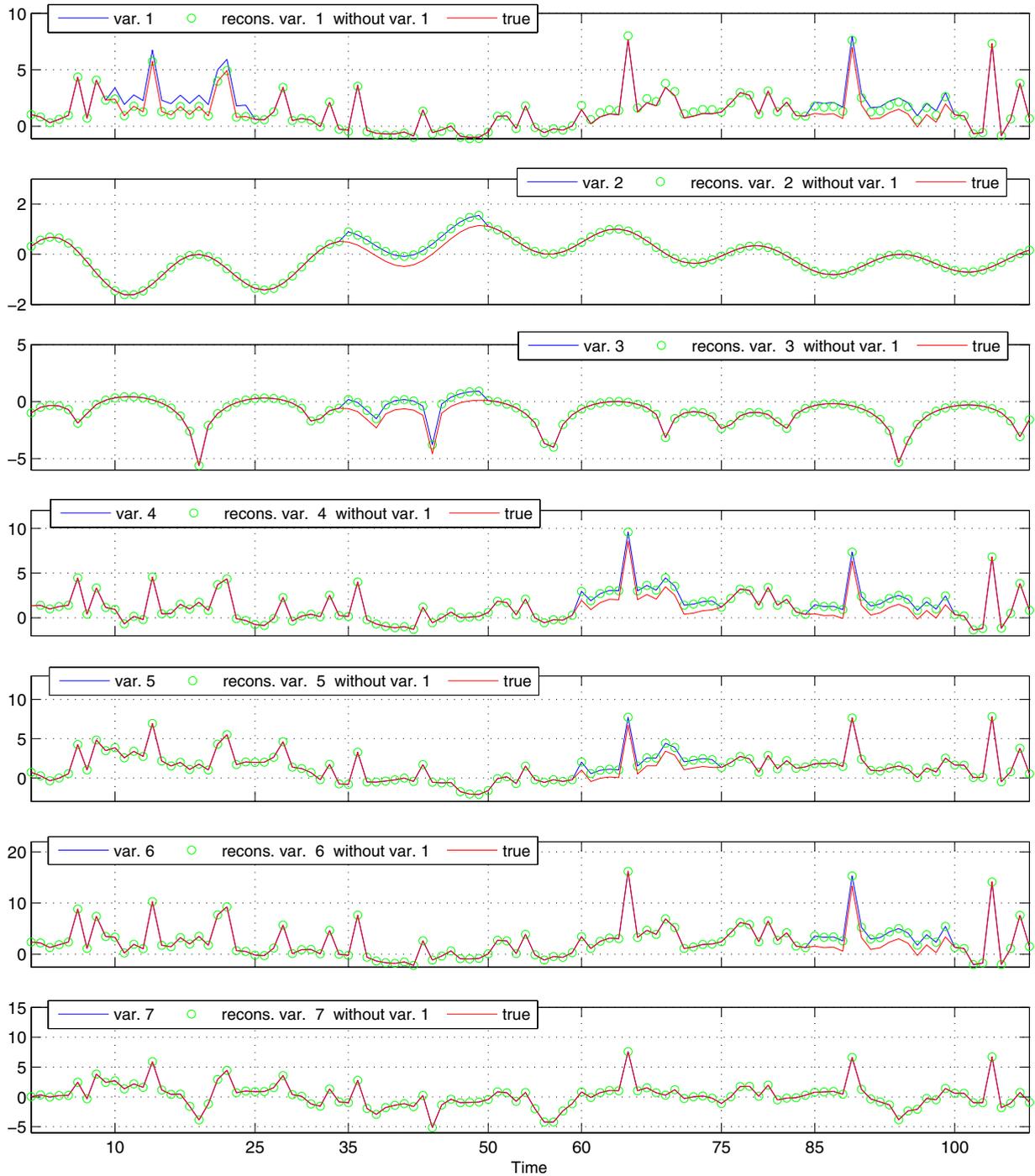
Fig. 6. Variable reconstruction without using Variable 1.

thus shows the absence of outliers in the variables used for the reconstruction and projection, i.e., all the variables except x_1. Let us note that the three other groups of observa-

tions (I_2, I_3, I_4) are affected by faults, without knowing exactly which components of the measurement vector are faulty. Finally, by taking into account the fault presence
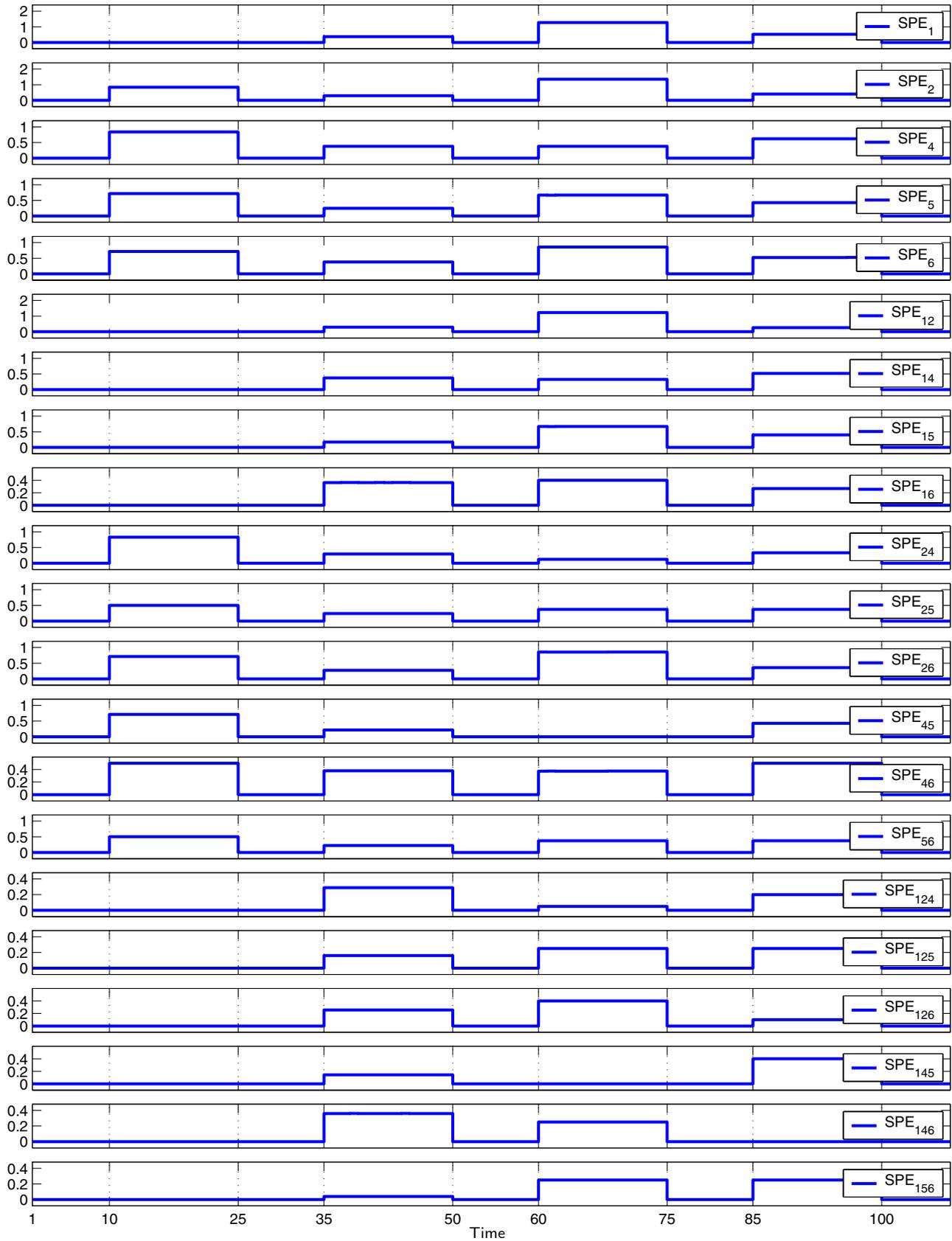
Fig. 7. *SPE* for isolable directions.

Table 8. Fault signatures.

	I_1	I_2	I_3	I_4
SPE_1	0	×	×	×
SPE_{45}	×	×	0	×
SPE_{146}	0	×	×	0

in the four intervals, the examination of the first graph of Fig. 7 leads to the conclusion that, in each interval I_2, I_3, I_4, a variable other than x_1 is faulty or more than one variable is faulty.

Other projections are built and are interpreted in a similar way. Figure 7 shows the global indicator for reconstruction directions which ensure isolation. Table 8 summarizes the conclusions resulting from the projection analysis (Fig. 7). SPE_1 pertains to the reconstructed residuals without using the first variable. The symbol '0' denotes the fault absence in the interval considered. The diagnosis is as follows:

- in the interval I_1, x_1 is faulty,

- in the interval I_2, x_2 and x_3 or/and x_7 are faulty, the fault is not isolable,

- in the interval I_3, x_4, x_5 are faulty,

- in the interval I_4, x_1, x_4, x_6 are faulty.

5. Conclusion

Principal components analysis reduces the data representation space and enables the determination of the redundancy relationships (linear relations among the variables). The redundancy relations are then used to detect and isolate the faults. PCA is constructed with fault-free data from a decomposition into the eigenvalues and eigenvectors of a covariance matrix. However, real data sets are not usually fault-free. Then the covariance matrix is disturbed by outliers. In order to reduce the sensitivity of the model to outliers, a fast two-step algorithm is proposed. First, the covariance matrix is replaced by its robust variant which leads to robust PCA. This one-step weighted estimate tends to emphasize the contribution of close observations in comparison with distant observations (outliers). Moreover, the results are not very sensitive to the tuning parameter β. Therefore, a model robust with respect to outliers was constructed. Second, structured residuals are generated for multiple fault detection and isolation. These structured residuals are based on the reconstruction principle. For fault isolation, the proposed scheme avoids the combinatorial explosion of faulty scenarios related to multiple faults. Indeed, instead of considering all combinations of one up to all sensors, we limit the maximum number of faulty scenarios to be considered by evaluating

the existence condition of structured residuals. Therefore, the detectable and isolable faults are determined, as well as the different faulty scenarios for which it is not possible to distinguish the faulty variables. This procedure was applied to two examples, the first with a single fault and the second with multiple faults. The presence of approximately 25 percent of outliers authorizes a correct estimation of the principal directions. Then the estimation is not very sensitive to these values. In both examples, the method is efficient for fault detection and isolation.

A PCA model can thus be built directly from the available data containing potential faults. One advantage of the suggested method is that it is not iterative unlike a majority of robust methods. The most important result concerns the diagnosis of the systems, applied here to the detection and isolation of outliers. For that purpose, we showed how to build fault indicators and to determine the isolable faults. The use of the principle of the reconstruction and projection of the reconstructed data together made it possible to detect and isolate outliers in an effective way.

In a future work, a comparison, in terms of fault detection and localization, of this robust approach with others like the robust LTS-subspace estimator and its generalizations will be performed. Another way is to improve the detection indicator. The detection is carried out only in the residual space and not in the principal space. For example, a combined fault indicator (Yue and Qin, 2001) using the residual and the principal space can be used instead.

References

Chiang L. H. and Colegrove L. F. (2007). Industrial implementation of on-line multivariate quality control, *Chemometrics and Intelligent Laboratory Systems* **88**(2): 143–153.

Croux C., Filzmoser P. and Oliveira M. (2007). Algorithms for projection-pursuit robust principal component analysis, *Chemometrics and Intelligent Laboratory Systems* **87**(2): 218–225.

Croux C. and Ruiz-Gazen A. (2005). High breakdown estimators for principal components: The projection-pursuit approach revisited, *Journal of Multivariate Analysis* **95**(1): 206–226.

Dunia R. and Qin S. (1998). A subspace approach to multidimensional fault identification and reconstruction, *American Institute of Chemical Engineers Journal* **44 (8)**: 1813–1831.

Harkat M.-F., Mourot G. and Ragot J. (2006). An improved PCA scheme for sensor FDI: Application to an air quality monitoring network, *Journal of Process Control* **16**(6): 625–634.

Hubert M., Rousseeuw P. and Van den Branden, K. (2005). RobPCA: A new approach to robust principal component analysis, *Technometrics* **47 (1)**: 64–79.

Hubert M., Rousseeuw P. and Verboven S. (2002). A fast method for robust principal components with applications

to chemometrics, *Chemometrics and Intelligent Laboratory Systems* **60**(1-2): 101–111.

Jackson J. and Mudholkar G. S. (1979). Control procedures for residuals associated with principal component analysis, *Technometrics* **21**(3): 341–349.

Kano M. and Nakagawa Y. (2008). Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry, *Computers & Chemical Engineering* **32**(1-2): 12–24.

Li G. and Chen Z. (1985). Projection-pursuit approach to robust dispersion matrices and principal components: Primary theory and Monte Carlo, *Journal of the American Statistical Association* **80**(391): 759–766.

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

Maronna R. A., Martin R. and Yohai V. J. (2006). *Robust Statistics: Theory and Methods*, Wiley, New York, NY .

Qin S. J. (2003). Statistical process monitoring: Basics and beyond, *Journal of Chemometrics* **17**(8-9): 480–502.

Rousseeuw P. (1987). *Robust Regression and Outliers Detection*, Wiley, New York, NY.

Rousseeuw P. and Van Driessen K. (1999). Fast algorithm for the minimum covariance determinant estimator, *Technometrics* **41**(3): 212–223.

Ruiz-Gazen, A. (1996). A very simple robust estimator of a dispersion matrix, *Computational Statistics and Data Analysis* **21**(2): 149–162.

Yue, H. and Qin, S. (2001). Reconstruction-based fault identification using a combined index, *Industrial and Engineering Chemistry Research* **40**(20): 4403–4414.