17

Sensor fault detection and isolation by robust principal component analysis

Yvon THARRAULT, Gilles MOUROT, José RAGOT
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

Mohamed-Faouzi HARKAT
Université Badji Mokhtar – ANNABA Faculté des Sciences de l'Ingénieur, Département d'Electronique BP. 12, Annaba. 23000 Algérie

1. Introduction

Sensors are essential components of modern control systems. Any faults in sensors will affect the overall performance of a system because their effects can easily propagate to manipulative variables through feedback control loops and also disturb other process variables. The task for sensor validation is to detect and isolate faulty sensors and estimate fault magnitudes afterwards to provide fault-free values. Model-based methods constitute an important approach to sensor fault detection and isolation (FDI).
A model-based approach consists in generating residuals as the difference between the measurements and the estimates provided by the relationships existing between the various variables of the process. The analysis of these residuals may lead to detect and isolate the faulty sensors. Almost all conventional model-based methods presume the knowledge of an accurate model of the system, e.g. transfer function or system matrices in the state space representation. Principal component Analysis (PCA) is a data-driven method which is particularly well adapted to reveal linear relationships among the plant variables without formulating them explicitly and has also been employed for system identification. PCA has some other nice features. It can handle high dimensional and correlated process variables, provides a natural solution to the errors-in-variables problem and includes disturbance decoupling (Li & Qin, 2001). Moreover in the FDI field, Gertler & McAvoy (1997) have shown a close link between PCA and parity space method. Principal component analysis (PCA) has then been applied successfully in the monitoring of complex systems (Chiang & Colegrove, 2007; Harkat et al., 2006; Kano & Nakagawa, 2008).

PCA is used to model the normal process behavior from an empirical data set which is representative of a normal process operation. In general, the majority of the training data set is associated with such 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, these outlying observations are not incorrect but they were made under exceptional circumstances. Therefore, 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 development of robust methods that are less affected by outliers. Large residuals from that robust fit indicate the presence of outliers. Once a robust model is determined, the next step deals with multiple fault detection and isolation. Indeed, outliers corresponding to either multiple faulty sensors or a priori unknown operating conditions affect many process variables.

This chapter is devoted to the problem of multiple fault detection and isolation. Section 2 presents the classical PCA principle and summarizes the benefits of different indices generally used for fault detection. Section 3, after a definition of outliers, introduces the main robust methods generally used. Next, a new robust method called MMRPCA for MM-estimator Robust Principal Component Analysis is proposed. It extends to all kinds of outliers the robust subspace estimator of Maronna (2005). Section 4 deals with multiple fault isolation. After a brief state of the art on fault isolation, structured residuals are generated for multiple fault isolation. These structured residuals are based on the reconstruction principle of process variables (Dunia et al., 1996; Wang et al., 2004a;b). However, instead of considering all the subsets of faulty variables (one up to all sensors), we determine the isolable multiple fault by evaluating the existence condition of these structured residuals. The proposed scheme avoids the combinatorial explosion of faulty scenarios related to the multiple faults to consider. In the last section 5 this method is applied on a simulated example in order to illustrate the different steps of our method.

2. Background

PCA is a widely used method for dimensionality reduction. Indeed, PCA transforms the data to a smaller set of variables (scores) which are the linear combinations of the original variables while retaining as much information as possible. Data redundancy stems from linear relation between process variables. The PCA model appears in the form of loadings, scores and variances. The eigenvectors (loadings) associated to the eigenvalues of the data covariance matrix span the representation and residual subspaces. The representation subspace (respectively residual subspace) associated to the first (respectively last) principal components (scores) describes significant variations of the process (respectively the noise in the data). Different statistics are defined on these two subspaces in order to detect faults. Qin (2003) proposed a unified representation of these fault detection indices.

For dynamic systems, since the current and past values of variables are related, Ku et al. (1995) proposed to include time lagged variables into the data matrix and to perform PCA on this augmented matrix in order to reveal dynamic linear relationships among the process variables. However, this straightforward application of PCA, called Dynamic PCA (DPCA), has several limitations as a monitoring approach for dynamic systems. Firstly, Li & Qin (2001) have shown that the DPCA method is consistent only if all variables have identical noise variance. Secondly, it is not clear how many time-lagged variables should be included into the augmented data matrix. These dynamic modeling problems have been approached by Li &

Qin (2001), Wang & Qin (2002), Qin & Wang (2006) in a instrumental variable framework. Thirdly, Kruger et al. (2004) have shown that DPCA produces correlated scores which lead to an undesired impact upon the fault detection ability in the representation subspace (production of false alarms). Xie et al. (2006) have then proposed to filter the principal scores obtained by DPCA by a Kalman innovation filter to remove the correlation between them. These points being outside the scope of this chapter, for the sake of simplicity, in the following, only classic PCA is considered in the following.

2.1 PCA modelling of systems

Let us consider $\mathbf{x}(k) = [x_1(k)\ x_2(k)\ \dots\ x_m(k)]^{\mathrm{T}}$ the vector formed with m observed plant variables at time instant k. Define the data matrix $\mathbf{X} = [\mathbf{x}(1)\ \mathbf{x}(2)\ \dots\ \mathbf{x}(N)]^{\mathrm{T}} \in \Re^{N\times m}$ with N samples $\mathbf{x}(k)(k = 1,\dots,N)$ which is representative of a normal process operation. PCA determines an optimal linear transformation of the data matrix $\mathbf{X}$ in terms of capturing the variation in the data:

$$\mathbf{T} = \mathbf{X}\,\mathbf{P} \qquad \text{and} \qquad \mathbf{X} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}} \tag{1}$$

with $\mathbf{T} = [\mathbf{t}_1\ \mathbf{t}_2\ \dots\ \mathbf{t}_m] \in \Re^{N\times m}$, where the vectors $\mathbf{t}_i$ are called scores or principal components and the matrix $\mathbf{P} = [\mathbf{p}_1\ \mathbf{p}_2\ \dots\ \mathbf{p}_m] \in \Re^{m\times m}$, where the orthogonal vectors $\mathbf{p}_i$, called loading or principal vectors, are the eigenvectors associated to the eigenvalues λ_i of the covariance matrix (or correlation matrix) $\mathbf{\Sigma}$ of $\mathbf{X}$:

$$\mathbf{\Sigma} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{\mathrm{T}} \quad \text{with} \quad \mathbf{P}\,\mathbf{P}^{\mathrm{T}} = \mathbf{P}^{\mathrm{T}}\mathbf{P} = \mathbf{I}_m \tag{2}$$

where $\mathbf{\Lambda} = diag(\lambda_1 \dots \lambda_m)$ is a diagonal matrix with diagonal elements in decreasing magnitude order.

The relations (1) are useful when the dimension ℓ of the representation subspace is reduced ($\ell << m$). Once the component number ℓ to retain is determined, the data matrix $\mathbf{X}$ can be approximated. For that, the different matrices are partitioned into the form:

$$\mathbf{P} = \left[\ \hat{\mathbf{P}}\ \middle|\ \tilde{\mathbf{P}}\ \right], \quad \hat{\mathbf{P}} \in \Re^{m\times\ell} \quad \text{and} \quad \tilde{\mathbf{P}} \in \Re^{m\times(m-\ell)} \tag{3}$$

$$\mathbf{\Lambda} = \begin{bmatrix} \hat{\mathbf{\Lambda}} & 0 \\ 0 & \tilde{\mathbf{\Lambda}} \end{bmatrix}, \quad \hat{\mathbf{\Lambda}} \in \Re^{\ell\times\ell} \quad \text{and} \quad \tilde{\mathbf{\Lambda}} \in \Re^{(m-\ell)\times(m-\ell)} \tag{4}$$

$$\mathbf{T} = \left[\ \hat{\mathbf{T}}\ \middle|\ \tilde{\mathbf{T}}\ \right], \quad \hat{\mathbf{T}} \in \Re^{N\times\ell} \quad \text{and} \quad \tilde{\mathbf{T}} \in \Re^{N\times(m-\ell)} \tag{5}$$

Equation (1) can be rewritten as:

$$\mathbf{X} = \hat{\mathbf{T}}\hat{\mathbf{P}}^{\mathrm{T}} + \tilde{\mathbf{T}}\tilde{\mathbf{P}}^{\mathrm{T}} = \hat{\mathbf{X}} + \mathbf{E} \tag{6}$$

with

$$\hat{\mathbf{X}} = \mathbf{X}\,\hat{\mathbf{C}} \qquad \text{and} \qquad \mathbf{E} = \mathbf{X}\,\left(\mathbf{I}_m - \hat{\mathbf{C}}\right) \tag{7}$$

where the matrix $\hat{\mathbf{C}} = \hat{\mathbf{P}}\,\hat{\mathbf{P}}^{\mathrm{T}}$ constitutes the PCA model.

The matrices $\hat{\mathbf{X}}$ and $\mathbf{E}$ represent, respectively, the modeled variations and non modeled variations of $\mathbf{X}$ based on ℓ components ($\ell < m$). The first ℓ eigenvectors $\hat{\mathbf{P}}$ constitute the representation subspace which describes the significant data variations whereas the last $(m - \ell)$ eigenvectors $\tilde{\mathbf{P}}$ constitute the residual subspace which describes the noises affecting the data. The identification of the PCA model thus consists in estimating its parameters by an eigenvalue/eigenvector decomposition of the matrix $\mathbf{\Sigma}$ and determining the number of principal components ℓ to retain.

A key issue to develop a PCA model is to choose the adequate number of principal components. Indeed, in most practical cases (noisy measurements), the small eigenvalues indicate the existence of linear or quasilinear relations among the process variables. However, the distinction between significant or insignificant eigenvalues may not be obvious due to modelling errors (disturbances and nonlinearities) and noise. Most methods to determine the number of principal components are rather subjective in the general practice of PCA (Qin & Dunia, 2000). Other methods are based on criteria actually used in system identification (Aikaike information criterion, minimum description length, ...) to determine the system order and emphasize the approximation of the data matrix $\mathbf{X}$ (see Valle et al. (1999) for a survey of these methods). However, the number of principal components has a significant impact on each step of the sensor fault detection and isolation scheme. Tamura & Tsujita (2007) proposed to choose ℓ which maximizes the sensitivity of fault detection indices to each sensor fault. This method can lead to monitor in parallel until m PCA models with various numbers of principal components. Qin & Dunia (2000) proposed to determine ℓ by minimization of the variance of the reconstruction error in the residual subspace. This variable reconstruction consists in estimating a variable from other plant variables using the PCA model, i.e. using the redundancy relations between this variable and the others. The reconstruction accuracy is thus related to the capacity of the PCA model to reveal the redundancy relations among the variables, i.e. to the number of principal components. This method will be retained here to determine ℓ.

2.2 Residual properties

After the PCA model has been built, we now examine its use for sensor fault detection and isolation. Let us consider now the fault propagation on the two signals obtained by the projection of the measurement vector $\mathbf{x}(k)$ onto the representation and the residual subspaces. In the presence of f faulty variables whose indices belong to the subset F, the measurement vector $\mathbf{x}(k)$ can be expressed as:

$$\mathbf{x}(k) = \mathbf{x}^{\mathrm{o}}(k) + \boldsymbol{\epsilon}(k) + \boldsymbol{\Xi}_F \, \mathbf{f}(k) \tag{8}$$

where $\mathbf{x}^{\mathrm{o}}(k)$ is true value vector, $\boldsymbol{\epsilon}(k)$ is the zero mean i.i.d. measurement noise vector, $\mathbf{f}(k)$ is the fault magnitude vector (unknown) and $\boldsymbol{\Xi}_F$ is the matrix of the fault directions. This orthonormal matrix with dimension ($m \times$ f) is built with 0 and 1, where 1 indicates the faulty variables from the other variables (with 0). For example, for the subset of faulty variables $F = \{2, 4\}$ among 5 variables, matrix $\boldsymbol{\Xi}_F$ is formed as follows:

$$\boldsymbol{\Xi}_F = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}^{\mathrm{T}}$$

Following decomposition (6), the measurement vector can also be represented as:

$$\mathbf{x}(k) = \hat{\mathbf{x}}(k) + \mathbf{e}(k) \tag{9}$$

where $\hat{\mathbf{x}}(k) = \hat{\mathbf{C}} \, \mathbf{x}(k)$ is the estimation vector, $\mathbf{e}(k) = \left(\mathbf{I_m} - \hat{\mathbf{C}}\right) \mathbf{x}(k)$ is the vector of estimation errors or residual vector.
The principal component vector is given by:

$$\mathbf{t}(k) = \mathbf{P}^{\mathrm{T}} \, \mathbf{x}(k) = \left[\hat{\mathbf{t}}(k) \;\; \tilde{\mathbf{t}}(k)\right] \tag{10}$$

where:

$$\hat{\mathbf{t}}(k) = \hat{\mathbf{P}}^{\mathrm{T}} \, \mathbf{x}(k), \tag{11}$$

$$\tilde{\mathbf{t}}(k) = \tilde{\mathbf{P}}^{\mathrm{T}} \, \mathbf{x}(k) \tag{12}$$

2.2.1 In the residual subspace

There is an equivalence between the residual vector and the last principal component vector $\tilde{\mathbf{t}}$:

$$\mathbf{e}(k) = \tilde{\mathbf{P}}\,\tilde{\mathbf{t}}(k) \quad \in \Re^{m\times 1} \tag{13}$$

So, it becomes simpler to work with this new residual vector $\tilde{\mathbf{t}}$ with dimension $(m-\ell)$. From (8) and (12), the residual vector is given by:

$$\tilde{\mathbf{t}}(k) \quad = \quad \tilde{\mathbf{P}}^{\mathrm{T}}\,\mathbf{x}^{\mathrm{o}}(k) + \tilde{\mathbf{P}}^{\mathrm{T}}\,\boldsymbol{\epsilon}(k) + \tilde{\mathbf{P}}^{\mathrm{T}}\,\boldsymbol{\Xi}_F\,\mathbf{f}(k) \tag{14}$$

Since $\tilde{\mathbf{P}}^{\mathrm{T}}\,\mathbf{x}^{\mathrm{o}}(k) = 0$ then (14) becomes:

$$\tilde{\mathbf{t}}(k) \quad = \quad \tilde{\mathbf{P}}^{\mathrm{T}}\,\boldsymbol{\epsilon}(k) + \tilde{\mathbf{P}}^{\mathrm{T}}\,\boldsymbol{\Xi}_F\,\mathbf{f}(k) \tag{15}$$

In the fault-free case, the expectation of the residual vector is zero. In the presence of faults, the expectation of the residual vector is no longer zero and the fault affects all the components of the residual vector. However if a variable is not correlated with others, its projection onto the residual subspace will be very small and then it will be very difficult to detect if this variable is faulty in the residual subspace. To conclude with this residual, it is possible to detect a fault (provided that residual is sufficiently sensitive to the fault) but it is difficult to isolate the faulty sensors.

Typically two main fault detection indices are used to monitor these residuals. The SPE (squared prediction error) is a statistic which measures the lack of fit of the PCA model to the data. At time k, the detection index SPE is given by:

$$SPE(k) = \tilde{\mathbf{t}}(k)^{\mathrm{T}}\tilde{\mathbf{t}}(k) \tag{16}$$

This quantity suggests the existence of an abnormal situation in the data when:

$$SPE(k) > \delta_\alpha^2 \tag{17}$$

where δ_α^2 is a control limit for SPE with a significance level α (Nomikos & MacGregor, 1995) or estimated using the historical data.

The SPE is formed by summing the square of residuals obtained from PCA model. However, modelling errors could be projected onto the residual subspace which results in residuals with a higher variance than the others. Then the SPE will be heavily biased in favour of those residuals with the largest residual variance whereas the residuals with the smallest residual variances are most useful for sensor fault diagnosis because they are associated with linear relationships. In this case using the indicator SWE (squared weighted error) can be more judicious because it takes into account the variances of the residual (Westerhuis et al., 2000), the problem inversion of near-zero singular eigenvalues seldom occuring in practice. At time k, the detection index SWE is given by:

$$SWE(k) = \tilde{\mathbf{t}}(k)^{\mathrm{T}}\tilde{\boldsymbol{\Lambda}}^{-1}\,\tilde{\mathbf{t}}(k) \tag{18}$$

This quantity suggests the existence of an abnormal situation in the data when:

$$SWE(k) > \chi^2_{m-\ell,\alpha} \tag{19}$$

where $\chi^2_{m-\ell,\alpha}$ is a Chi-2 upper control limit with $m-\ell$ degrees of freedom for a significance level α.

2.2.2 In the representation subspace

From (8) and (12), the principal score vector is given by:

$$\hat{\mathbf{t}}(k) \quad = \quad \hat{\mathbf{P}}^{\mathrm{T}}\,\mathbf{x}^{\mathrm{o}}(k) + \hat{\mathbf{P}}^{\mathrm{T}}\,\boldsymbol{\epsilon}(k) + \hat{\mathbf{P}}^{\mathrm{T}}\,\Xi_F\,\mathbf{f}(k) \tag{20}$$

The faults affect all the components of $\hat{\mathbf{t}}(k)$. Moreover, the term $\hat{\mathbf{P}}^{\mathrm{T}}\,\mathbf{x}^{\mathrm{o}}(k)$ describing the significant variations of the data, can mask the fault influence on the principal score vector except if the fault amplitude is high relative to this term which depends on the operating point $\mathbf{x}^{\mathrm{o}}(k)$. The T^2 statistic is the main fault detection index to monitor the principal scores. At time k, the detection index T^2 is given by:

$$T^2(k) = \hat{\mathbf{t}}(k)^{\mathrm{T}}\,\hat{\mathbf{\Lambda}}^{-1}\,\hat{\mathbf{t}}(k) \tag{21}$$

This quantity suggests the existence of an abnormal situation in the data when:

$$T^2(k) > \chi^2_{\ell,\alpha} \tag{22}$$

where $\chi^2_{\ell,\alpha}$ is a Chi-2 upper control limit with ℓ degrees of freedom for a significance level α.

2.2.3 In the entire space

However, according to the magnitude and the kind of faults, faults can be only projected onto the representation or the residual subspace and then affect only one of the previous indices. Rather than monitoring different fault detection indices, combined indices are preferred. The Mahalanobis distance can be used as a measure of variability in the entire space. At time k, the Mahalanobis distance D^2 is given by:

$$D^2(k) = \mathbf{x}(k)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\,\mathbf{x}(k) \tag{23}$$

This quantity suggests the existence of an abnormal situation in the data when:

$$D^2(k) > \chi^2_{m,\alpha} \tag{24}$$

where $\chi^2_{m,\alpha}$ is a Chi-2 upper control limit with m degrees of freedom for a significance level α. Previously we have seen that faults affect all the components of a residual or principal score vector. Therefore it is rather difficult to isolate the faulty sensors by using these vectors. The problem of isolation enhancement will be addressed in Section 4.

3. Robust PCA

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 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 mean squared error) from the eigenvectors of the data covariance matrix. The performance of PCA model is then based on the accurate estimation of the covariance matrix from the data which is very sensitive to abnormal observations.

In practice one often tries to detect outliers by 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 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.
After having described outliers using PCA in the following section, the robust PCA methods are introduced.

3.1 Characteristics of outliers using PCA

The different types of outliers are classified according to their effects on the construction of the PCA model. They are explained in figure 1 considering a system with 3 variables and 2 principal components. Thus, 4 types of Observations can be distinguished (Hubert et al., 2005):

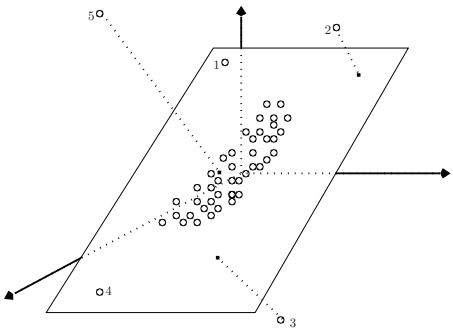

Fig. 1. Illustration of different types of outliers for an example with 3 variables and 2 principal components

- *Regular observations:* observation which belong to a homogeneous group that is close to the representation subspace.
- *Good leverage points:* Outliers close to the representation subspace (small projection onto the residual subspace) but far from the regular observations such as the observations 1 and 4 of figure 1.
- *Orthogonal outliers:* Outliers whose orthogonal distance to representation subspace is large (large projection onto the residual subspace) but not visible with its projection onto the representation subspace, like the observation 5 of figure 1.
- *Bad leverage points:* Outliers that have a large orthogonal distance (large projection onto the residual subspace) and whose projection onto the representation subspace is remote from the typical projections, such as the observations 2 and 3 of figure 1.

To construct a robust model with respect to outliers, the influences of outliers in the construction of the PCA model have to be minimised.

3.2 Classical robust PCA methods

Several ways of robustifying principal components have been proposed (Daszykowski et al., 2007; Filzmoser et al., 2008; Rousseeuw et al., 2006). To enable the comparison of different robust methods, measures of performance are necessary. One such performance measure for

robust methods is the *breakdown point* (Donoho & Huber, 1983) which can be defined as the maximal fraction of outlying objects in the data that the estimator can handle in order to yield acceptable estimates. For instance, the breakdown point of the mean estimator with 0% being the smallest possible. For such breakdown point a single outlier can completely corrupt the estimate. Conceptually, it is impossible to distinguish between the good and the bad parts in the data if the fraction of outliers becomes larger than 50%. In the following, we focus on robust methods with the highest breakdown point, i.e. close to 50%.
To robustify PCA with respect to outliers, M-estimators (Huber, 1964) can be used. These estimators minimise a more general objective function than the classical criterion on Mean Square Error. However, the breakdown point of M-estimators is limited by $\frac{1}{m+1}$, with m the number of variables. Then this method is not very robust when considering large systems (large m). The robust methods can be grouped as follows :

- A first group of robust PCA methods is obtained by replacing the classical covariance matrix by 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 but also 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 high dimension. Rousseeuw & Van Driessen (1999) constructed a much faster algorithm called FAST-MCD which avoids such a complete enumeration. It is obtained by combining a basic subsampling and iterative scheme with MCD estimator. Another example is Caussinus et al. (2003) who 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). As for the MCD method, this method has a parameter setting β, which depends on the number of outliers a priori unknown.
- 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 (Croux et al., 2007; Croux & Ruiz-Gazen, 2005; Huber, 1964; Li & Chen, 1985). 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 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 center of the data 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 robust PCA method is to combine the two previous robust approaches like the method proposed in Hubert et al. (2005) and called ROBPCA. This method combined ideas of both projection pursuit and robust covariance estimation based on FAST-MCD algorithm. It first applied projection pursuit techniques in the original data space. These results are then used to project the observations onto a subspace of small to moderate

dimension. 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 representation subspace.

- The last proposals for robust PCA include the robust LTS-subspace estimator and its generalizations (Maronna et al., 2006). The idea behind these approaches consists in minimizing a robust scale of the orthogonal distances of each observation to the representation subspace, similar to the LTS estimator, S-estimators and many others in regression. These methods are based on iterative procedures for which there remains the problem of starting values. 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 the MCD-estimator: a high computational cost, the choice of the user-defined parameter h and the starting values. Like MCD-estimator, a FAST-LTS algorithm has been proposed.

All these methods have a tuning parameter which changes with the fraction of outliers in the data. However, this information is unknown. To fill this gap, we propose a new robust method called MMRPCA for MM-estimator Robust Principal Component Analysis. Thus, a MM-estimator is used to determine a robust model. This estimator is a combination of two M-estimators: a M-estimator to estimate the model and a second M-estimator to estimate the robust scale of the residual. This estimator, calculated with an iterative algorithm is initialized with a robust estimator of the covariance matrix (Caussinus et al., 2003). Then fault detection tools are used, from the robust model for finding outliers in the data. The influences of outliers can then be eliminated and the resulting PCA model becomes unbiased. Moreover, a robust procedure is proposed to determine the number of principal components.

3.3 MMRPCA Method (MM-estimator Robust Principal Component Analysis)

Our approach consists in carrying out PCA directly on the data possibly contaminated by outliers. For that, a simple robust estimator, called MM-estimator, is used. However, this estimator is computed by an iterative procedure. Thus a good initialization parameter procedure is needed to avoid local minima. To initialize this MM-estimator a robust covariance matrix is first calculated.

3.3.1 Robust covariance matrix

Caussinus et al. (2003) define a "local" matrix of covariance 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:

$$\boldsymbol{Q} = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \boldsymbol{\Omega}(i,j)(\boldsymbol{x}(i) - \boldsymbol{x}(j))(\boldsymbol{x}(i) - \boldsymbol{x}(j))^{\mathrm{T}}}{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} w(i,j)} \tag{25}$$

where the weights $\mathbf{\Omega}(i,j)$ themselves are defined by:

$$\mathbf{\Omega}(i,j) = \exp\left(-\frac{\beta}{2}(\boldsymbol{x}(i)-\boldsymbol{x}(j))^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\boldsymbol{x}(i)-\boldsymbol{x}(j))\right) \tag{26}$$

β being a tuning parameter to reduce the influence of the observations faraway, as recommended by the authors, β value is equal to 2. For $\beta = 0$, the robust covariance matrix $\boldsymbol{Q}$ is equal to $2\,\mathbf{\Sigma}$. And for a high value of β, only the closest observations are taken into account in the robust covariance matrix $\boldsymbol{Q}$.

3.3.2 MM-estimator

Two M-estimators are used, one to estimate the minimum of the objective function and another one for the estimation of the robust residual scale. The general MM-estimator minimizes the following objective function with the constraint $\tilde{\boldsymbol{P}}^{\mathrm{T}}\tilde{\boldsymbol{P}} = \boldsymbol{I}_{m-\ell}$ (Maronna, 2005):

$$\frac{1}{N}\sum_{k=1}^{N}\rho\left(\frac{\boldsymbol{r}(k)}{\hat{\sigma}}\right) \tag{27}$$

with $\boldsymbol{r}(k) = ||\tilde{\boldsymbol{P}}\,\boldsymbol{x}(k) - \tilde{\boldsymbol{P}}\,\boldsymbol{\mu}||^2$ is the residual, $\boldsymbol{\mu} = \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{x}(k)$ is the mean, $\hat{\sigma}$ is the robust scale of the residual $\boldsymbol{r}(k)$ and the function $\rho{:}\Re^{+} \rightarrow [0,1]$ is nondecreasing, with $\rho(0) = 0$, $\rho(\infty) = 1$, and differentiable. $\tilde{\boldsymbol{P}}$ is the eigenvector matrix of the robust covariance matrix $\boldsymbol{S}$ (29) corresponding to its $m-\ell$ smallest eigenvalues. Then the weighted mean $\boldsymbol{\mu}$ and the covariance $\boldsymbol{S}$ are defined as follows:

$$\boldsymbol{\mu} = \frac{\sum_{k=1}^{N}\boldsymbol{w}(k)\boldsymbol{x}(k)}{\sum_{k=1}^{N}\boldsymbol{w}(k)} \qquad \text{with} \quad \boldsymbol{w}(k) = \dot{\rho}\left(\frac{\boldsymbol{r}(k)}{\hat{\sigma}}\right) \tag{28}$$

$$\boldsymbol{S} = \sum_{k=1}^{N}\boldsymbol{w}(k)(\boldsymbol{x}(k)-\boldsymbol{\mu})(\boldsymbol{x}(k)-\boldsymbol{\mu})^{\mathrm{T}} \tag{29}$$

where $\dot{\rho}(x) = \frac{\partial\rho(x)}{\partial x}$.
Then the scale factor $\hat{\sigma}$ is defined as the solution to:

$$\frac{1}{N}\sum_{k=1}^{N}\rho\left(\frac{\boldsymbol{r}(k)}{\hat{\sigma}}\right) = \delta \tag{30}$$

with $\delta \in [0,1]$. This parameter δ is directly related to the fraction of outliers in the data *i.e.* the breakdown point. To maximize the breakdown point, Maronna et al. (2006) define δ as follows:

$$\delta = \frac{N-m+\ell-1}{2N} \qquad \text{if } N >> m \text{ then } \delta \approx 0.5 \tag{31}$$

Usually, when the number of observations (N) is larger than the number of variables (m) then parameter δ is chosen equal to 0.5. By defining a weight function $\boldsymbol{w}_{\sigma}$ in the following way:

$$\boldsymbol{w}_{\sigma}(\boldsymbol{r}) = \begin{cases} \rho(\boldsymbol{r})/\boldsymbol{r} & \text{if } \boldsymbol{r} \neq 0 \\ \ddot{\rho}(0) & \text{if } \boldsymbol{r} = 0 \end{cases} \tag{32}$$

where $\ddot{\rho}(x) = \frac{\partial^2 \rho(x)}{\partial x^2}$.
The solution of equation (30) is expressed by the following equation:

$$\hat{\boldsymbol{\sigma}} = \frac{1}{N\delta} \sum_{k=1}^{N} \boldsymbol{w}_{\boldsymbol{\sigma}} \left(\frac{\boldsymbol{r}(k)}{\hat{\boldsymbol{\sigma}}} \right) \boldsymbol{r}(k) \tag{33}$$

One notices that an iterative resolution of equation (33) is needed to determine the scale factor $\hat{\boldsymbol{\sigma}}$.
We chose the function ρ (27) as the Bisquare function ($\boldsymbol{r}$ represents the square of residuals) because it allows the cancellation of the influence of outliers.

$$\rho(\boldsymbol{r}) = \min\{1, 1 - (1 - \boldsymbol{r})^3\} \tag{34}$$

and

$$\boldsymbol{w}_{\boldsymbol{\sigma}}(\boldsymbol{r}) = \min\{3 - 3\boldsymbol{r} + \boldsymbol{r}^2, 1/\boldsymbol{r}\} \tag{35}$$

However, this method is only robust to fault with a projection onto the residual subspace. Then to be robust to all kinds of faults, a similar approach in the representation subspace is used. In that case the MM-estimator maximizes the following objective function with the constraint $\hat{\boldsymbol{P}}^{\mathrm{T}} \hat{\boldsymbol{P}} = \boldsymbol{I}_{\ell}$:

$$\frac{1}{N} \sum_{k=1}^{N} \rho \left(\frac{||\hat{\boldsymbol{P}}^{\mathrm{T}} \boldsymbol{x}(k) - \hat{\boldsymbol{P}}^{\mathrm{T}} \boldsymbol{\mu}||^2}{\hat{\sigma}} \right) \tag{36}$$

To ensure the elimination of any fault disturbing the residual subspace when the MM-estimator in the representation subspace is used, the minimum between the weight determined with the first MM-estimator (robust to fault with a projection onto the residual subspace) and the weight obtained with the second MM-estimator (robust to fault with a projection onto the representation subspace) is used (line 26 of table 1). The algorithm of the MMRPCA method is described in table 1.

Finally, to improve the estimation of the covariance matrix, and thus the PCA model, a last weighting step is implemented by using the Mahalanobis distance to eliminate outliers.

$$\begin{cases} \boldsymbol{w}(k) = 1 & \text{if} \quad D^2(k) \leq \chi^2_{m,\alpha} \\ \quad = 0 & \text{else} \end{cases} \tag{37}$$

where $D^2(k)$ is the Mahalanobis distance
The robust mean $\boldsymbol{\mu}$ and the variance matrix $\boldsymbol{S}$ are then defined as follows:

$$\boldsymbol{\mu} = \frac{\sum_{k=1}^{N} \boldsymbol{w}(k) \boldsymbol{x}(k)}{\sum_{k=1}^{N} \boldsymbol{w}(k)} \tag{38}$$

$$\boldsymbol{S} = \left(\sum_{k=1}^{N} \boldsymbol{w}(k) (\boldsymbol{x}(k) - \boldsymbol{\mu})(\boldsymbol{x}(k) - \boldsymbol{\mu})^{\mathrm{T}} \right) \Big/ \left(\sum_{k=1}^{N} \boldsymbol{w}(k) - 1 \right) \tag{39}$$

However, this algorithm requires the number of principal components to be known. Hence, a robust method to find the number of principal components is introduced.

1.	$it = 1$ **and** $\sigma_0 = \infty$**;**
2.	**Compute** $\tilde{\boldsymbol{P}}$ **the eigenvector matrix of the robust covariance matrix** $\boldsymbol{Q}$ **corresponding to its** $m - \ell$ **smallest eigenvalues;**
3.	**Compute** $\boldsymbol{a} = median(\boldsymbol{X}\,\tilde{\boldsymbol{P}})$**;**
4.	**Do until convergence of** Δ**;**
5.	**Compute** $\boldsymbol{r}(k) = \|\|\tilde{\boldsymbol{P}}\boldsymbol{x}(k) - \boldsymbol{a}\|\|^2$ **for** $k = 1...N$**;**
6.	**Compute** $\hat{\sigma}$ **from (33);**
7.	**If** $it > 1$**, set** $\Delta = 1 - \hat{\sigma}/\sigma_0$**;**
8.	**Set** $\sigma_0 = \hat{\sigma}$**;**
9.	**Compute** $w(k) = \dot{\rho}\left(\boldsymbol{r}(k)/\hat{\sigma}\right)$ **for** $k = 1...N$**;**
10.	**Compute** $\boldsymbol{\mu}$ **from (28);**
11.	**Compute** $\boldsymbol{S}$ **from (29);**
12.	**Compute** $\tilde{\boldsymbol{P}}$ **the eigenvector matrix of the covariance matrix** $\boldsymbol{S}$**;**
13.	**corresponding to its** $m - \ell$ **smallest eigenvalues;**
14.	**Compute** $\boldsymbol{a} = \tilde{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{\mu}$**;**
15.	**Set** $it = it + 1$**;**
16.	**End do;**
17.	**Set** $\boldsymbol{w}_{res} = \boldsymbol{w}$**,** $it = 1$ **and** $\sigma_0 = \infty$**;**
18.	**Compute** $\hat{\boldsymbol{P}}$ **the eigenvector matrix of the robust covariance matrix** $\boldsymbol{S}$ **corresponding to its** ℓ **largest eigenvalues;**
19.	**Compute** $\boldsymbol{a} = median(\boldsymbol{X}\,\hat{\boldsymbol{P}})$**;**
20.	**Do until convergence of** Δ**;**
21.	**Compute** $\boldsymbol{r}(k) = \|\|\hat{\boldsymbol{P}}\boldsymbol{x}(k) - \boldsymbol{a}\|\|^2$ **for** $k = 1...N$**;**
22.	**Compute** $\hat{\sigma}$ **from (33);**
23.	**If** $it > 1$**, set** $\Delta = 1 - \hat{\sigma}/\sigma_0$**;**
24.	**Set** $\sigma_0 = \hat{\sigma}$**;**
25.	**Compute the** $\boldsymbol{w}(k) = \dot{\rho}\left(\boldsymbol{r}(k)/\hat{\sigma}\right)$ **for** $k = 1...N$**;**
26.	**Set** $\boldsymbol{w} = min(\boldsymbol{w}, \boldsymbol{w}_{res})$ **;**
27.	**Compute** $\boldsymbol{\mu}$ **from (28);**
28.	**Compute** $\boldsymbol{S}$ **from (29);**
29.	**Compute** $\hat{\boldsymbol{P}}$ **the eigenvector matrix of the covariance matrix** $\boldsymbol{S}$**;**
30.	**corresponding to its** ℓ **largest eigenvalues;**
31.	**Compute** $\boldsymbol{a} = \hat{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{\mu}$**;**
32.	**Set** $it = it + 1$**;**
33.	**End do.**

Table 1. Algortithm of the MMRPCA method

3.4 Robust determination of the principal component number

The number ℓ of principal components to choose is obtained by minimizing the normalized VRE (variance of reconstruction error) with respect to the number ℓ (Qin & Dunia, 2000), the criterion is then :

$$J(\ell) = \sum_{j=1}^{m} \frac{\boldsymbol{\xi}_j^{\mathrm{T}}\left(\boldsymbol{I}_m - \hat{\boldsymbol{C}}\right)\boldsymbol{S}\left(\boldsymbol{I}_m - \hat{\boldsymbol{C}}\right)\boldsymbol{\xi}_j}{\left(\boldsymbol{\xi}_j^{\mathrm{T}}(\boldsymbol{I}_m - \hat{\boldsymbol{C}})\boldsymbol{\xi}_j\right)^2} \tag{40}$$

with $\ell = 1, \ldots, m-1$, $\hat{\boldsymbol{C}}$ the robust PCA model, $\boldsymbol{S}$ the robust covariance matrix and $\boldsymbol{\xi}_j$ the reconstruction direction ($\boldsymbol{\xi}_j = [0\ ...\ 1\ ...\ 0]^{\mathrm{T}}$ where value 1 is at the j^{th} position). Qin & Dunia (2000) show that this criterion may present a minimum in the interval $[1, m]$. The algorithm to determine both the number of principal components ℓ and the robust model is summarized in table 2.

1.	**Compute the robust covariance matrix $\boldsymbol{Q}$ (25);**
2.	**Set $\ell_{ini} = 1$, ℓ_{ini} is the number of principal component used to construct the robust model with the MMRPCA method;**
3.	**Do until $\ell_{ini} < m$;**
4.	**Compute the robust covariance with the MM-estimator $\boldsymbol{S}$ (39);**
5.	**Find the minimum of the normalized VRE (ℓ varying from 1 to $m - \ell$);**
6.	**Set ℓ_{ini}=ℓ_{ini}+1;**
7.	**End do;**
8.	**The smallest value obtain by minimizing the normalized VRE is associated with the number of principal component to use.**

Table 2. Algorithm to determine the number of principal components

4. Fault isolation

After the presence of faults has been detected, it is important to identify these faults and to apply the necessary corrective actions to eliminate the abnormal data. In the PCA framework, the well known isolation approaches are residual enhancement, contribution plot and variable reconstruction methods.

In structured design, each residual responds to a specific subset of faults, and thus a specific subset of the residuals responds to each fault, resulting in a unique fault code. Structured residuals may be obtained from a full PC model by algebraic transformation. But they can also be generated by first specifying residual structures and then obtaining subsystem models (partial PCA models), each corresponding to a residual (Huang & Gertler, 1999; Qin & Weihua, 1999). However, for a high dimensionality process, it is not always possible to find the residual structuration that enables to obtain the desired isolation properties because these properties are only defined according to the occurrence of the faults in the residuals without taking into account the sensitivities of the residuals to the faults.

Contribution plots are well known diagnostic tools for fault isolation. The most common indices used for fault diagnosis with contribution plots are SPE and T^2. Contribution plots on SPE indicate the significance of the effect of each variable on this index at different sampling times. If a sample vector $\mathbf{x}$ has an abnormal SPE, the components of this vector that appear to have a significant contribution are investigated. A contribution plot on principal scores indicates the significance of the effect of each variable on the T^2 index. The variables with the largest contribution are considered as major contributors to the fault. The contribution plots are very easy to calculate, with no prior knowledge required to generate the plots. Prior knowledge, however, is often used and required to interpret the plots. As explained by MacGregor & Kourti (1995), the contribution plots may not explicitly identify the cause of an abnormal condition. The reason is that when there is no fault, the contributions are uneven across variables. Therefore, a fault in a normally small contribution variable may not have the largest contribution unless the fault magnitude is very large. This can be a source of misdiagnosis.

An alternative approach for fault isolation is the variable reconstruction method proposed by Dunia et al. (1996). The reconstruction based approach eliminates the fault when the actual fault direction is used for reconstruction. In the case of arbitrary process fault directions the reconstruction-based approach brings a fault indicator within the normal control limit (completely removes the effect of the fault when the faulty variables are reconstructed).

4.1 Reconstruction approach

In this section the variable reconstruction approach is presented. The reconstruction $\mathbf{x}_R(k)$ of an observation $\mathbf{x}(k)$ is defined as follows:

$$\mathbf{x}_R(k) = \mathbf{x}(k) - \Xi_R \mathbf{f}_R \tag{41}$$

with $\mathbf{f}_R$ the fault magnitude (unknown) and the matrix Ξ_R indicating the reconstruction directions. This matrix is orthonormal with dimension ($m \times$ r), with r the number of component to reconstruct, and is built with 0 and 1, where 1 indicates the reconstructed variables from the other variables (with 0). For example, to reconstruct the set of variables $R = \{2, 4\}$ among 5 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 reconstruction can be formulated by an optimization problem of the detection indicator with respect to fault $\mathbf{f}_R$ according to the direction Ξ_R. So, according to the indicator used for detection SPE, SWE, T^2 or D^2, several types of reconstruction can take place.

The estimation of the fault magnitude $\mathbf{f}_R$ is obtained by solving the following optimization problem, according to the detection indicator:

$$\hat{\mathbf{f}}_R = \arg\min_{\mathbf{f}_R} \left\{ \mathbf{x}_R^{\mathrm{T}}(k) \mathbf{\Phi} \mathbf{x}_R(k) \right\} \tag{42}$$

where

- In reconstruction according to SPE indicator, we have: $\mathbf{\Phi} = \tilde{\mathbf{P}}\tilde{\mathbf{P}}^T$,
- In reconstruction according to SWE indicator, we have: $\mathbf{\Phi} = \tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}^{-1}\tilde{\mathbf{P}}^T$,
- In reconstruction according to T^2 indicator, we have: $\mathbf{\Phi} = \hat{\mathbf{P}}\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{P}}^T$
- In reconstruction according to D^2 indicator, we have: $\mathbf{\Phi} = \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}^T$

Given a subset R, the expression for the reconstruction $\mathbf{x}_R(k)$ of the vector $\mathbf{x}(k)$ is then expressed by:

$$\mathbf{x}_R(k) = \mathbf{G}_R \, \mathbf{x}(k) \tag{43}$$

with $\mathbf{G}_R = \left(\mathbf{I} - \Xi_R (\Xi_R^T \mathbf{\Phi} \Xi_R)^{-1} \Xi_R^T \mathbf{\Phi} \right)$

Faults can only be projected onto the representation or the residual subspace. To isolate all kinds of faults, reconstruction according to D^2 (Mahalanobis distance) indicator is considered ($\mathbf{\Phi} = \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}^T$). We define $D_R^2(k)$ as the Mahalanobis distance, calculated after the reconstruction of the subset R of variables which is given by:

$$D_R^2(k) = \mathbf{x}_R^{\mathrm{T}}(k) \mathbf{P} \mathbf{\Lambda}^{-1} \mathbf{P}^T \mathbf{x}_R(k) \tag{44}$$

The system is considered normal if:

$$D_R^2(k) \leq \chi^2_{m-r,\alpha} \tag{45}$$

where $D_R^2(k)$ is the fault detection indicator (Mahalanobis distance), calculated after the reconstruction of a subset R of variables, and $\chi^2_{m-r,\alpha}$ is a chi-2 control limit with $m-r$ degrees of freedom for a significance level α.

If we write $\mathbf{x}_R$ (43) in the case where the matrix of the reconstruction directions is reorganized as follows (using column permutations):

$$\Xi_R = \left[\begin{array}{c|c} \underset{(r\times r)}{\mathbf{I}_1} & \underset{((m-r)\times r)}{\mathbf{0}} \end{array} \right]^{\mathrm{T}} \in \Re^{m\times r} \tag{46}$$

with $\mathbf{I}_1 \in \Re^{r\times r}$ an identity matrix. Then $\mathbf{P}$ and $\mathbf{\Lambda}$ are split in four parts:

$$\mathbf{P}^{\mathrm{T}} = \left[\begin{array}{c|c} \underset{(r\times r)}{\mathbf{P}_{11}^{\mathrm{T}}} & \underset{(r\times(m-r))}{\mathbf{P}_{12}^{\mathrm{T}}} \\ \hline \underset{((m-r)\times r)}{\mathbf{P}_{21}^{\mathrm{T}}} & \underset{((m-r)\times(m-r))}{\mathbf{P}_{22}^{\mathrm{T}}} \end{array} \right] \in \Re^{m\times m} \tag{47}$$

$$\mathbf{\Lambda} = \left[\begin{array}{c|c} \underset{(r\times r)}{\mathbf{\Lambda}_1} & \underset{(r\times(m-r))}{\mathbf{0}} \\ \hline \underset{((m-r)\times r)}{\mathbf{0}} & \underset{((m-r)\times(m-r))}{\mathbf{\Lambda}_2} \end{array} \right] \in \Re^{m\times m} \tag{48}$$

Using definition (43), the reconstruction $\mathbf{x}_R$ of the vector $\mathbf{x}$ is written as follows:

$$\mathbf{x}_R = \left[\begin{array}{c|c} \mathbf{0} & \left(\Xi_R^{\mathrm{T}}\Xi_R\right)^{-1}\left(\mathbf{P}_{11}\mathbf{\Lambda}_1^{-1}\mathbf{P}_{12}^{\mathrm{T}} + \mathbf{P}_{21}\mathbf{\Lambda}_2^{-1}\mathbf{P}_{22}^{\mathrm{T}}\right) \\ \hline \mathbf{0} & \mathbf{I}_2 \end{array} \right] \mathbf{x} \tag{49}$$

with $\mathbf{I}_2 \in \Re^{m-r\times m-r}$ an identity matrix.

This form highlights two characteristics. First, the reconstructed vector $\mathbf{x}_R$ is constituted by the r reconstructed variables and a copy of the $m-r$ remaining variables. Secondly, the reconstructed variables are estimated without using their own measurement.

4.1.1 Reconstruction condition

To reconstruct a fault, it must be at least projected into the representation subspace ($r \leq \ell$) or onto the residual subspace ($r \leq m-\ell$). This implies that the number of reconstructed variables r must respect the following inequality:

$$r \leq \max(m-\ell, \ell) \tag{50}$$

4.1.2 Structured residual generation

In a diagnosis objective, residuals are generated for fault detection and isolation. Considering equations (8) and (43), then $\mathbf{x}_R$ can be expressed as:

$$\mathbf{x}_R(k) = \left(\mathbf{I} - \Xi_R(\Xi_R^T\mathbf{\Phi}\Xi_R)^{-1}\Xi_R^T\mathbf{\Phi}\right)\ (\mathbf{x}^{\mathrm{o}}(k) + \boldsymbol{\epsilon}(k) + \Xi_F\mathbf{f}) \tag{51}$$

As the fault influence is generally unknown, we have to consider all possible reconstruction directions Ξ_R

- If the reconstruction directions Ξ_R (8) are the same as the fault directions, i.e. if $R = F$, then $D_R^2(k)$ is under the detection threshold ($D_R^2 < \chi^2_{m-r,\alpha}$), indeed, the fault influence is removed:

$$\left(\mathbf{I} - \Xi_R(\Xi_R^T\mathbf{\Phi}\Xi_R)^{-1}\Xi_R^T\mathbf{\Phi}\right)\ \Xi_R = \mathbf{0} \tag{52}$$

  and the reconstruction is expressed as:

$$\mathbf{x}_R(k) = \left(\mathbf{I} - \Xi_R(\Xi_R^T\mathbf{\Phi}\Xi_R)^{-1}\Xi_R^T\mathbf{\Phi}\right)\ (\mathbf{x}^{\mathrm{o}}(k) + \boldsymbol{\epsilon}(k)) \tag{53}$$

- If the reconstruction directions Ξ_R are different from the fault directions, then $D_R^2(k)$ is higher than the detection threshold $\chi^2_{m-r,\alpha}$ if the projection of the reconstruction directions are not collinear to the fault projection onto the residual subspace and onto the representation subspace.

For the faulty observation k, the faulty variables subset $\hat{R}$ is determined as follows:

$$\hat{R} = \underset{R\in\Im}{\arg}\ D_R(k) < \chi^2_{m-r,\alpha} \tag{54}$$

with $\Im$ the set of possible reconstruction directions.

4.2 Fault isolation

The proposed fault isolation procedure consists in two steps: offline and online steps. The offline step consists of a priori analysis of all reconstruction directions which allows to determine the isolable faults (useful reconstructions). The second online step consists in using the useful reconstruction directions (corresponding to isolable faults) for multiple fault isolation.

4.2.1 Offline step

All the directions of reconstruction Ξ_R have to be explored for fault isolation. The maximum reconstruction number can be calculated as follows:

$$\sum_{r=1}^{\max(m-\ell,\ell)-1} \mathbb{C}_m^r \tag{55}$$

where $\mathbb{C}_m^r$ denotes the combination of r from m.
This number takes only into account the number of reconstructions in the different subspaces. However, collinear projections have the same fault signature. Then we will analyze the angles between the different projections of reconstruction directions. The largest primary angle θ between two subspaces of the same size is linked to the concept of distance between these two subspaces (Golub & Van Loan, 1996).
This distance is defined in the representation subspace $d(R_i, R_j)$ and in the residual subspace $\tilde{d}(R_i, R_j)$ as follows:

$$d(R_i, R_j) = ||\hat{\Xi}_{R_i}(\hat{\Xi}_{R_i}^{\mathrm{T}}\hat{\Xi}_{R_i})^{-1}\hat{\Xi}_{R_i}^{\mathrm{T}} - \hat{\Xi}_{R_j}(\hat{\Xi}_{R_j}^{\mathrm{T}}\hat{\Xi}_{R_j})^{-1}\hat{\Xi}_{R_j}^{\mathrm{T}}||_2 \tag{56}$$

$$\tilde{d}(R_i, R_j) = ||\tilde{\Xi}_{R_i}(\tilde{\Xi}_{R_i}^{\mathrm{T}}\tilde{\Xi}_{R_i})^{-1}\tilde{\Xi}_{R_i}^{\mathrm{T}} - \tilde{\Xi}_{R_j}(\tilde{\Xi}_{R_j}^{\mathrm{T}}\tilde{\Xi}_{R_j})^{-1}\tilde{\Xi}_{R_j}^{\mathrm{T}}||_2 \tag{57}$$

with $\hat{\Xi}_{R_i} = \hat{\Lambda}^{-1/2}\hat{\mathbf{P}}^{\mathrm{T}}\Xi_{R_i}$, $\tilde{\Xi}_{R_i} = \tilde{\Lambda}^{-1/2}\tilde{\mathbf{P}}^{\mathrm{T}}\Xi_{R_i}$ and R_i and R_j correspond to two sets of variable reconstructions.

Analyzing these distances, then the isolable fault can be determined a priori. Hence, a global indicator $\mathcal{K}$ is built.

$$\mathcal{K}(R_i, R_j) = \max\{(d(R_i, R_j), \tilde{d}(R_i, R_j)\} \tag{58}$$

Thus, if $\mathcal{K}(R_i, R_j)$ is close to zero, it means that the projections of the set of reconstructed variables R_i and R_j are collinear in the residual subspace and in the representation subspace. It means that a fault for the sets of reconstructed variables R_i or R_j are not isolable. The process to detect useful directions of reconstruction can be summarized as follows:

1. $r = 1$
2. Calculate for all available directions ($R_i \in \Im$ and $R_j \in \Im$) the indicator $\mathcal{K}(R_i, R_j)$ (58). The smaller the value of this indicator, the higher the magnitude of the fault has to be to ensure fault isolation. And if this indicator is equal to zero, then only a set of potentially faulty variables may be determined, i.e. the faulty variables are associated to the indices R_i or R_j or R_i and R_j. Thus, it is only required to study one single subset of directions, for example R_i.
3. $r = r + 1$
4. While $r \leq \max(\ell, m - \ell) - 1$ do to the step 2

This analysis of the structure of the model allows to determine a priori the isolable faults. The number of useful reconstructions can then be greatly reduced.

4.2.2 Online step

If a fault is detected on the $D^2(k)$ indicator:

1. For $r = 1$, (number of reconstructed variables)
2. Compute $D^2_R(k)$, where $R \in \Im$ is a subset (of r variables) of useful reconstruction selected in the offline step:
   - if for a particular reconstruction direction Ξ_R, $D^2_R(k) \leq \chi^2_{m-r,\alpha}$, variables in this subset R are the faulty variables. Isolation procedure is stopped.
   - Otherwise, if there are more than the r faulty variables, go to step 3
3. r=r+1
4. While $r \leq max(m - \ell, \ell) - 1$, go to step 2
5. The fault is not isolable (more than $r = max(m - \ell, \ell) - 1$ variables can be faulty)

5. Numerical example

5.1 Data generation

We consider here the situation in which several faults affect different variables at the same time. The matrix $\mathbf{X}$ includes $N = 450$ observations of a vector x with $m = 9$ components generated in the following way:

$$\begin{aligned} &x_{i,1} = 1 + v_i^2 + \sin(i/3), \quad x_{i,2} = 2\sin(i/6)\cos(i/4)\exp(-i/N), \quad v_i \sim \mathcal{N}(0,1) \\ &x_{i,3} = \log(x_{i,2}^2), \quad x_{i,4} = x_{i,1} + x_{i,2}, \quad x_{i,5} = x_{i,1} - x_{i,2} \\ &x_{i,6} = 2x_{i,1} + x_{i,2}, \quad x_{i,7} = x_{i,1} + x_{i,3}, \quad x_{i,8} \sim \mathcal{N}(0,1), \quad x_{i,9} \sim \mathcal{N}(0,1) \end{aligned} \tag{59}$$

On the data thus generated were added realizations of random variables with centred normal distribution and standard deviations equal to 0.02 as well as faults δx_1 represented by a bias of amplitude equal to 20% of the amplitude of the variable, δx_2, δx_3 represented by a bias of amplitude equal to 10% of the amplitudes of the variables, δx_8 represented by a bias of amplitude equal to 150% of the amplitude of the variable. Faults are defined on specific time intervals: observations from 50 to 100 (interval I_1) for the variable x_1, observations from 150 to 200 (interval I_2) for the variables x_2 and x_3, observations from 250 to 300 (interval I_3) for the variables x_8.

5.2 Robust PCA

To determine the number of principal components, the robust approach using the VRE, proposed in the section 3.4, is used. Five principal components are selected ($\ell = 5$). The robust model is then built. Figure 2 shows the measure of the first variable x_1 with its estimation obtained with the classic PCA model and the robust PCA model and the associated residual (measure - estimate).

This figure shows that in the fault-free case (for example interval from 300 to 450) the robust residuals are centred on zero while, with the classical PCA, the residuals are not centred on zero. It means that a smaller fault magnitude can be detected much better using the robust method than using the classical method. This shows the advantage of using a robust approach.

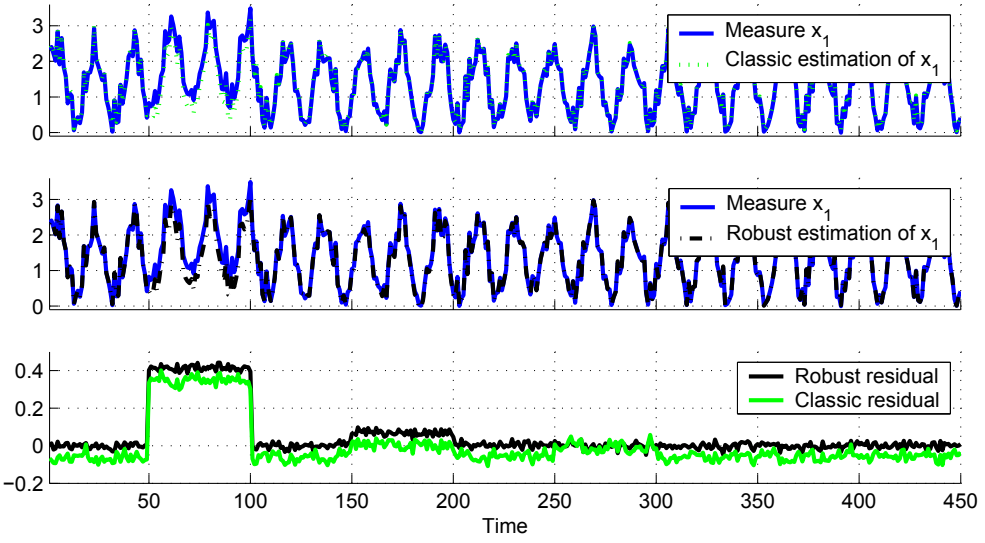

Fig. 2. Measure and estimation of x_1

5.3 Fault detection and isolation

5.3.1 Offline step

The Mahalanobis distance is used to detect and to isolate the faults. Considering the dimension of the residual subspace and of the representation subspace, we cannot reconstruct simultaneously more than five variables ($\max(m - \ell, \ell)$). The maximum number of reconstructions is then equal to 255 (55). Table 3 shows the values of the global indicator $\mathcal{K}$ (58) with $r = 1$,

i.e. only one variable is reconstructed. The two sets R_1 and R_2 contain the indices of the reconstructed variables. The smaller the value of this indicator $\mathcal{K}$, the higher the magnitude of the fault has to be to ensure fault isolation. All the values of $\mathcal{K}$ are not null, so all the faults on one variable are isolable.
For all the directions of reconstruction ($r = 2, 3, 4$) this indicator is calculated. A case where $\mathcal{K}$ is close to zero is detected between $D^2_{1,3}$ and $D^2_{1,7}$. Then the fault signatures of these two directions are identical ($D^2_{1,3} = D^2_{1,7}$). Therefore, only one indicator is useful to detect this fault, for example $D^2_{1,3}$. Moreover, we concluded that the signatures of reconstruction directions taking into account these sets are identical ($D^2_{1,3,4} = D^2_{1,4,7}$, $D^2_{1,3,6} = D^2_{1,6,7}$, ...). The number of useful reconstructions can be reduced to 168.

K		R_1								
		1	2	3	4	5	6	7	8	9
R_2	1	0	1.00	0.99	0.99	0.89	0.94	0.99	1.00	1.00
	2		0	1.00	0.95	0.72	0.97	1.00	1.00	0.98
	3			0	1.00	0.98	0.95	0.23	1.00	1.00
	4				0	0.98	0.75	1.00	1.00	0.99
	5					0	0.99	0.99	1.00	0.99
	6						0	0.96	1.00	1.00
	7							0	1.00	1.00
	8								0	1.00

Table 3. Indicator $\mathcal{K}$ for $r = 1$

5.3.2 Online step

Figure 3 shows the Mahalanobis distance divided by its detection threshold obtained for a significance level of $\alpha = 99\%$, i.e. a fault is detected if the normalised Mahalanobis distance is greater than one. Then, the faults on intervals I_1, I_2 and I_3 are detected only by using the robust Mahalanobis distance (constructed with the robust model). Let us remark that the classical Mahalanobis distance is not able to detect the faults.

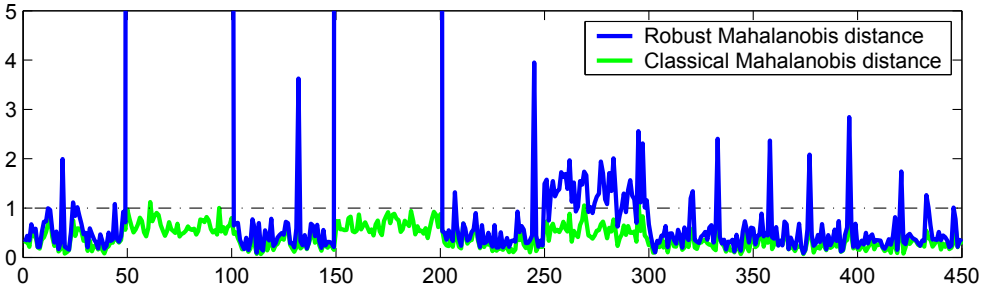

Fig. 3. Fault detection with Mahalanobis distance

Once the faults are detected, we will try to isolate them. First, all useful reconstruction directions for the reconstruction of one variable (r=1) are calculated. The first graph of the figure 4 shows the global indicator D^2_1 (44). For the observations of the interval I_1 this distance is close to the value 0 and thus shows the absence of outliers in the variables used for the reconstruction, i.e. all the variables except x_1. Let us note that the two other groups of observations (I_2,

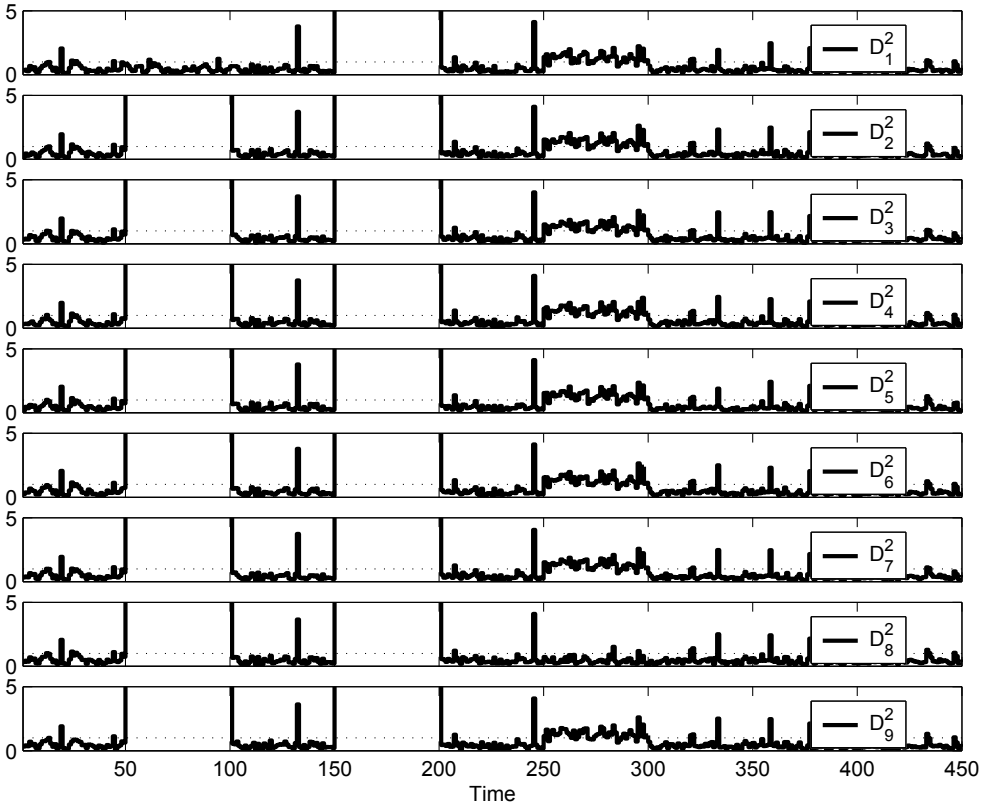

Fig. 4. Fault isolation with Mahalanobis distance

I_3) are affected by faults, but we don't know exactly which components of the measurement vector are faulty. Finally, by taking into account the fault presence in the three intervals, the examination of the first graph of the figure 4 helps us to conclude that:

- in each interval I_2, I_3 other variables than x_1 are faulty.

Other reconstructions are built and are interpreted in a similar way. Figure 4 shows all values of the indicator D_R^2 for useful reconstruction directions always in the case $r = 1$. From the reconstruction of one variable, the diagnosis is as follows:

- in the interval I_1, x_1 is faulty,
- in the interval I_2, more than one variable is faulty,
- in the interval I_3, x_8 is faulty.

Since the fault on interval I_2 is not isolate, then the useful reconstructions with two variables are performed. Figure 5 shows some values of the indicator D_R^2 for useful reconstruction directions with $r = 2$. From the reconstruction of two variables, the diagnosis is the following:

- in the interval I_2, x_2 and x_3 are faulty,

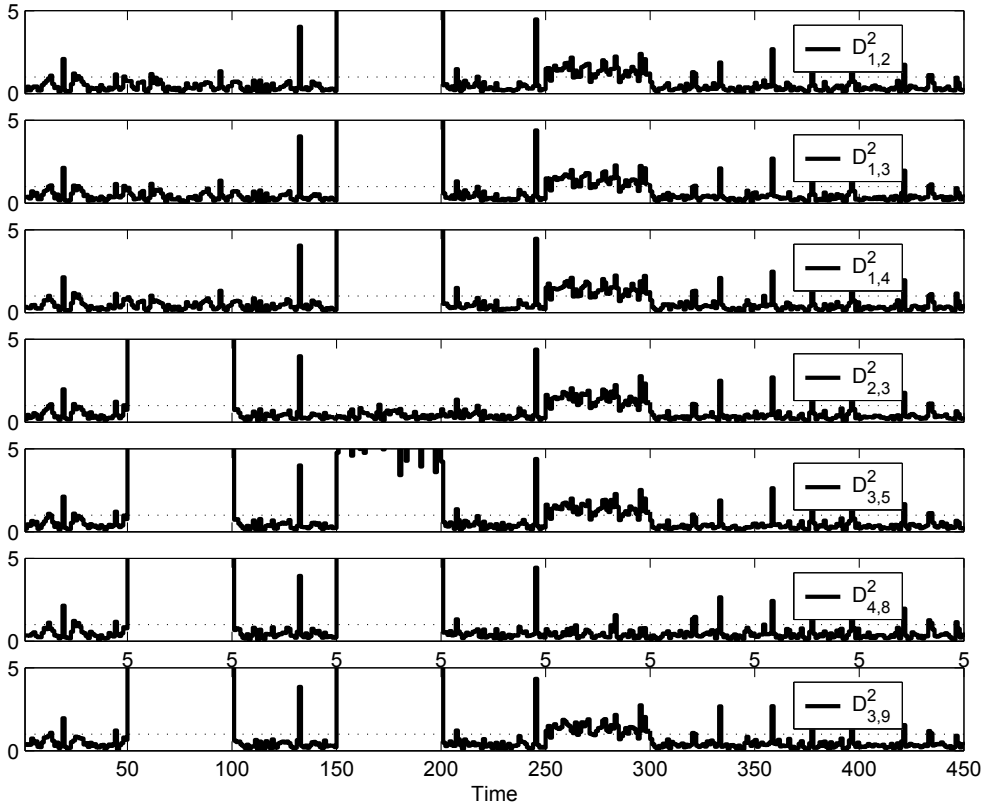

Fig. 5. Fault isolation with Mahalanobis distance

- in the intervals I_1 and I_3 the previous conclusions are confirmed.

If all the faults are isolated, then it is not useful to reconstruct more variables. Table 4 summarizes the conclusions resulting from the D_R^2 analysis (figures 4 and 5). Symbol 0 denotes the fault absence and symbol $\times$ denotes the fault presence in the considered interval.

	I_1	I_2	I_3
D_1^2	0	$\times$	$\times$
$D_{2,3}^2$	$\times$	0	$\times$
D_8^2	$\times$	$\times$	0

Table 4. Fault signatures

6. Conclusion

Principal components analysis reduces the data representation subspace and enables the determination of the redundancy relationships (linear relations among the variables). The re-

dundancy relations are then used to detect and isolate the faulty data. PCA is constructed with fault-free data from a decomposition in eigenvalues and eigenvectors of a covariance matrix.
However, real data sets are usually not fault-free then the covariance matrix is disturbed by outliers. In order to reduce the sensitivity of the model to all kinds of outliers (with a projection onto the representation or the residual subspace), a fast two-step algorithm is proposed. First, a MM-estimator is used to determine a robust model. This estimator is computed by using an iterative re-weighted least squares (IRWLS) procedure. This algorithm is initialized from a very simple estimate derived from a one-step weighted covariance estimate. Therefore, a model robust with respect to outliers is constructed. Secondly, structured residuals are generated for multiple fault detection and isolation. These structured residuals are based on the reconstruction-projection 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 consider by evaluating the existence condition of structured residuals. Therefore, the detectable faults and the isolable faults are determined as well as the different faulty scenarios for which it is not possible to distinguish the faulty variables. This procedure has been applied on one example, with single and multiple faults. The presence of approximately 30 percent of outliers authorizes a correct estimation of the principal components, then the estimation is not very sensitive to outliers. The method is efficient for fault detection and isolation.

7. References

Caussinus, H., Hakam, S. & Ruiz-Gazen, A. (2003). Projections révélatrices contrôlées : Groupements et structures diverses, *Revue de statistique appliquée* **51**(1): 37–58.

Chiang, L. H. & 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. & Oliveira, M. (2007). Algorithms for projection-pursuit robust principal component analysis, *Chemometrics and Intelligent Laboratory Systems* **87**(2): 218–225.

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

Daszykowski, M., Kaczmarek, K., Vander Heyden, Y. & Walczak, B. (2007). Robust statistics in data analysis - a review: basic concepts, *Chemometrics and Intelligent Laboratory Systems* **85**(2): 203–219.

Donoho, D. L. & Huber, P. J. (1983). *A Festschrift for Eric Lehmann*, Wadsworth, Belmont, CA, chapter The notion of breakdown point.

Dunia, R., Qin, S. & Edgar, T. F. (1996). Identification of faulty sensors using principal component analysis, *American Institute Of Chemical Engineers Journal* **42**(10): 2797–2812.

Filzmoser, P., Maronna, R. & Werner, M. (2008). Outlier identification in high dimensions, *Computational Statistics & Data Analysis* **52**(3): 1694–1711.

Gertler, J. & McAvoy, T. (1997). Principal component analysis and parity relations - a strong duality, *IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, SAFEPROCESS'97, Hull,UK*. 837-842.

Golub, G. H. & Van Loan, C. F. (1996). *Matrix computations*, The Johns Hopkins University Press, Baltimore and London.

Harkat, M.-F., Mourot, G. & 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.

Huang, Y. & Gertler, J. (1999). Fault isolation by partial PCA and partial NLPCA, *14th triennial world congress, IFAC'99, Beijing, P. R. China*.

Huber, P. (1964). Robust estimation of a location parameter, *Annals of Mathematical Statistics* **35**: 73–101.

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

Kano, M. & 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.

Kruger, U., Zhou, Y. & Irwin, G. W. (2004). Improved principal component monitoring of large-scale processes, *Journal of Process Control* **14**: 879–888.

Ku, W., Storer, R. & Georgakis, G. (1995). Disturbance detection and isolation by dynamic principal component analysis, *Chemometrics and Intelligent Laboratory Systems* **30**: 179–196.

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

MacGregor, J. F. & Kourti, T. (1995). Statistical process control of multivariate processes, *Control Engineering Practice* **3**(3): 403–414.

Maronna, R. A. (2005). Principal components and orthogonal regression based on robust scales, *Technometrics* **47**(3): 264–273.

Maronna, R. A., Martin, R. & Yohai, V. J. (2006). *Robust Statistics: Theory and Methods*, John Wiley & Sons, Ltd.

Nomikos, P. & MacGregor, J. F. (1995). Multivariate SPC charts for monitoring batch processes, *Technometrics* **37**: 41–59.

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

Qin, S. J. & Dunia, R. (2000). Determining the number of principal components for best reconstruction, *Journal of Process Control* **10**: 245–250.

Qin, S. J. & Weihua, L. (1999). Detection, identification and reconstruction of faulty sensors with maximized sensitivity, *American Institute Of Chemical Engineers Journal* **45**(9): 1963–1976.

Qin, S. & Wang, J. (2006). Closed-loop subspace identification using the parity space, *Automatica* **42**: 315–320.

Rousseeuw, P. (1987). *Robust regression and outliers detection*, John Wiley & Sons Inc.

Rousseeuw, P., Debruyne, M., Engelen, S. & Hubert, M. (2006). Robustness and outlier detection in chemometrics, *Analytical Chemistry,* **36**: 221–242.

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

Tamura, M. & Tsujita, S. (2007). A study on the number of principal components and sensitivity of fault detection using PCA, *Computers and Chemical Engineering* **31**: 1035–1046.

Valle, S., Li, W. & Qin, S. J. (1999). Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods, *Industrial Engineering Chemical Research* **38**: 4389–4401.

Wang, H., Jiang, N. & Yang, D. (2004a). Process monitoring in principal component subspace: part 1 - fault reconstruction study, *IEEE international conference on systems, man and cybernetics, The Hague, Netherlands, 10-13 october*.

Wang, H., Jiang, N. & Yang, D. (2004b). Process monitoring in principal component subspace: part 2. fault identification and isolation study, *IEEE international conference on systems, man and cybernetics, The Hague, Netherlands, 10-13 october*.

Wang, J. & Qin, S. (2002). A new subspace identification approach based on principal component analysis, *Journal of Process Control* **12**: 841–855.

Westerhuis, J., Gurden, S. & Smilde, A. (2000). Standardized q-statistic for improved sensitivity in the monitoring of residuals in MSPC, *Journal of Chemometrics* **14**(4): 335–349.

Xie, L., Kruger, U., Lieftucht, D., Littler, T., Chen, Q. & Wang, S. (2006). Statistical monitoring of dynamic multivariate process - part 1. modeling autocorrelation and cross-correlation, *Industrial Engineering Chemical Research* **45**: 1659–1676.

Fault Detection

Edited by Wei Zhang

ISBN 978-953-307-037-7
Hard cover, 504 pages
Publisher InTech
Published online 01, March, 2010
Published in print edition March, 2010

In this book, a number of innovative fault diagnosis algorithms in recently years are introduced. These methods can detect failures of various types of system effectively, and with a relatively high significance.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yvon Tharrault, Gilles Mourot, Jose Ragot and Mohamed-Faouzi Harkat (2010). Sensor Fault Detection and Isolation by Robust Principal Component Analysis, Fault Detection, Wei Zhang (Ed.), ISBN: 978-953-307-037-7, InTech, Available from: http://www.intechopen.com/books/fault-detection/sensor-fault-detection-and-isolation-by-robust-principal-component-analysis

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821