



**HAL**  
open science

# Principal Alarms in Multivariate Statistical Process Control using Independent Component Analysis

Isabel Gonzalez, Ismael Sanchez

► **To cite this version:**

Isabel Gonzalez, Ismael Sanchez. Principal Alarms in Multivariate Statistical Process Control using Independent Component Analysis. *International Journal of Production Research*, 2008, 46 (22), pp.6345-6366. 10.1080/00207540701361467. hal-00512981

**HAL Id: hal-00512981**

**<https://hal.science/hal-00512981>**

Submitted on 1 Sep 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



### Principal Alarms in Multivariate Statistical Process Control using Independent Component Analysis

Journal:	<i>International Journal of Production Research</i>
Manuscript ID:	TPRS-2006-IJPR-0570.R1
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	22-Dec-2006
Complete List of Authors:	Gonzalez, Isabel; Universidad Carlos III de Madrid, Mechanical Engineering Sanchez, Ismael; Universidad Carlos III de Madrid, Department of Statistics
Keywords:	SPC, MULTIVARIATE STATISTICS
Keywords (user):	ICA



# Principal Alarms in Multivariate Statistical Process Control using Independent Component Analysis

Isabel González\*  
Department of Mechanical Engineering  
Universidad Carlos III de Madrid

Ismael Sánchez†  
Department of Statistics  
Universidad Carlos III de Madrid

December 22, 2006

## Abstract

This article proposes a methodology that helps to predict the main mean shifts, denoted as principal alarms, in a non-normal multivariate process using the available in-control data. The analysis is based on the transformation of the observed correlated variables into independent factors using independent component analysis. These independent components allow us to simulate shifts preserving the covariance structure. The graphical representations of those simulated shifts are helpful to improve the design and the control of the process. Two real manufacturing processes are presented showing the advantage of the proposed methodology.

**Keywords:** Independent Component Analysis; Manufacturing processes; Multivariate process control; Principal alarms; **Principal components analysis.**

---

\*Correspondence author: Isabel González. Universidad Carlos III de Madrid. Avd. de la Universidad 30, 28911, Leganés, Madrid (Spain), [imgfarial@ing.uc3m.es](mailto:imgfarial@ing.uc3m.es), phone: 34-916245860, fax: 34-916249430

†Ismael Sánchez. Universidad Carlos III de Madrid. Avd. de la Universidad 30, 28911, Leganés, Madrid (Spain), [ismael@est-econ.uc3m.es](mailto:ismael@est-econ.uc3m.es), phone: 34-916249179, fax: 34-916249430

# Principal Alarms in Multivariate Statistical Process Control using Independent Component Analysis

December 22, 2006

## Abstract

This article proposes a methodology that helps to predict the main mean shifts, denoted as principal alarms, in a non-normal multivariate process using the available in-control data. The analysis is based on the transformation of the observed correlated variables into independent factors using independent component analysis. These independent components allow us to simulate shifts preserving the covariance structure. The graphical representations of those simulated shifts are helpful to improve the design and the control of the process. Two real manufacturing processes are presented showing the advantage of the proposed methodology.

**Keywords:** Independent Component Analysis; Manufacturing processes; Multivariate process control; Principal alarms; **Principal components analysis.**

# 1 Introduction

Monitoring a production process is important to detect and prevent faults or other special events. To that end, performing statistical process control (SPC) via control charts has proven to be an effective method for monitoring and also for improving manufacturing processes. Often the quality is measured by the joint level of several correlated variables. To monitor such processes, multivariate control charts have been proposed in the statistical literature. The Hotelling  $T^2$  chart is the most commonly used chart for correlated variables (Hotelling, 1947; Alt, 1985). Furthermore, the multivariate version of the CUSUM chart has been proposed by Woodall and Ncube (1985), and Croisier (1988). A multivariate extension of the exponentially weighted moving average (MEWMA) can be found in Lowry et al. (1992).

The use of multivariate control charts has, however, an important problem related with the interpretation of the out-of-control signals in terms of the original quality variables and, consequently, with the corrective decisions that are to be taken. This problem occurs primarily because multivariate control charts attempt to reduce a  $k$ -dimensional data vector into a one-dimensional statistic where the information of the original variables is mixed. The identification of the cause of the alarm in a multivariate process can then be a complex issue. **An initial approach to deal with this problem was based on the principal components of the multivariate data. Principal components analysis (PCA) is an appropriate tool to project a set of correlated variables onto a set of new uncorrelated variables (principal components). This projection allows a univariate analysis of each principal component. Jackson (1980, 1985) proposes the plotting of not only the  $T^2$  chart but all the univariate charts of the principal components. However, since each principal component is a combination of all the original variables, these plots do not usually lead to a clear interpretation.**

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
Alternative techniques for the interpretation of out-of-control signals are based on variations of the  $T^2$  chart with the aim of extracting as much information of the out-of-control signal as possible. The aim of these procedures is to account for the contribution of the original variables in the  $T^2$  value. For instance, Murphy (1987) proposes a searching procedure where the overall  $T^2$  value is compared to the  $T^2$  value based on different subsets of variables. Doganaksoy et al. (1991) propose a set of univariate t tests that help to rank the variables according to their participation in the out-of-control signal. Mason et al. (1995) propose an orthogonal decomposition of the  $T^2$  value each of which reflects the contribution of an individual variable and can be used to test the significance of each individual variable or a set of them. Alternative procedures to decompose the  $T^2$  values can also be found in Hawkins (1993), and Wade and Woodall (1993). Some procedures like those proposed by Mason et al.(1997), can give the quality characteristics responsible for the shift.

35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
All these procedures are helpful in identifying the cause of a specific observed out-of-control signals. However, they do not supply information about potential alarms before these occur. In González and Sánchez (2006), a complementary analysis to find the causes of the alarms is proposed. The approach, however, is very different to multivariate control charts in that the analysis is not aimed at monitoring the process but at identifying the main expected alarms. These main expected alarms are denoted as principal alarms (PA). The analysis of PA can be made with the same data used for building the control charts, and can be seen as a descriptive and a preventive tool, as opposed to control charts, which are used as an on-line control tool . The analysis of PA has several advantages. First, the PA help to identify weak points in the process, which is very useful in the initial stages of the design of a product or

1  
2  
3 a process. Second, since the PA point out the weakness of the process, they can be  
4 used to select the most convenient control chart, as done in González and Sánchez  
5 (2006). And third, when a multivariate control chart triggers an alarm, we can start  
6 the identification of the causes with the PA, since they are the most likely shifts. The  
7 PA are, therefore, useful for the prevention of alarms, for the design of efficient charts,  
8 and for allowing a faster response to an out-of-control signal.  
9

10  
11  
12  
13  
14  
15  
16 The main argument in a PA analysis is that a mean shift in a multivariate process  
17 involving correlated variables,  $\mathbb{X} = (X_1, X_2, \dots, X_K)$ , can not be considered as a set of  
18 independent shifts on each variable  $X_i$ . The measured process variables are not inde-  
19 pendent, but may be combinations of independent factors that might not be directly  
20 measurable. In multivariate analysis, these independent factors are usually referred  
21 to as latent factors. González and Sánchez (2006) use the PA analysis to simulate a  
22 realistic mean shift by using a set of independent latent factors  $\mathbb{Y} = (Y_1, Y_2, \dots, Y_J)$ ,  
23 where  $J \leq K$ . This vector  $\mathbb{Y}$  is obtained using an appropriate transformation of the  
24 original variables  $\mathbb{Y} = f(\mathbb{X})$  that can be interpreted as the primary independent sources  
25 of variability of the process. Since the variables  $\mathbb{Y}$  are independent, it is possible to  
26 simulate a mean shift in  $\mathbb{Y}$  by only modifying each component  $Y_j$  independently of  
27 each other. This mean shift in  $\mathbb{Y}$  can be translated in terms of the original variables  
28  $\mathbb{X}$  by applying the inverse of the transformation,  $f^{-1}(\mathbb{Y})$ . The shifts in terms of  $\mathbb{X}$ ,  
29 provoked by shifting each column of  $\mathbb{Y}$  are the so called principal alarms of the process.  
30 In order to obtain this vector of independent variables  $\mathbb{Y}$ , González and Sánchez (2006)  
31 use principal components analysis (PCA) based on an assumption of normality. As is  
32 well known, under the normality assumption, independence is equivalent to uncorre-  
33 latedness. However, if normality can not be assumed, uncorrelated variables can still  
34 have non linear dependence.  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54

55  
56 In this article, a methodology is proposed to obtain the PA when  $\mathbb{X}$  is non-normally  
57 distributed. This methodology does not use any transformation of  $\mathbb{X}$  to obtain normal-  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

ity. The vector of independent factors  $\mathbb{Y}$  will be obtained using independent component analysis (ICA) instead of PCA. ICA is a recently developed method in which the goal is to decompose non-normal multivariate data into a linear combination of statistically independent components (Lee, 1998; Hyvärinen, 1999; Hyvärinen and Oja, 2000; Hyvärinen et al.2001). **ICA has recently been used in SPC to build multivariate control charts (Chen et al. 2004; Lee et al. 2004; Kano et al. 2004).** The idea of multivariate control charts using ICA can be seen as the extension of the control charts based on PCA proposed by Jackson (1980, 1985) and Jackson and Mudholkar (1979) to non-normal data. By using ICA, the original multivariate variables are decomposed into a set of independent variables (independent components). Then, univariate and multivariate control charts of these independent components are built to monitor the process. In this article, the independent components of the process will no be used to monitor the process but to identify the main alarms of a non-normal process. As in the normal case based on PCA, the independent components are estimated with historical data collected during normal operation. These independent components will be used as a descriptive tool that helps to simulate the main out-of-control situations that will take place if a shift in the independent components occur. A graphical representation of these shifts will help quality engineers to improve the understanding of the process.

The structure of the article is as follows: Section 2 introduces the technique of ICA, Section 3 describes the proposed methodology of estimating the principal alarms of a non-normal process using ICA, Section 4 applies the methodology to two real cases, and Section 5 gives the conclusions.

## 2 Independent component analysis

### 2.1 The independent component transformation

The most popular methods for transforming a set of related variables into **uncorrelated** ones are PCA and factor analysis. These methods can be labeled as second order methods since they only use the information contained in the covariance matrix of the observed data set. Hence, these techniques are to be understood under normality, because under such assumption all the information of zero mean normal variables is contained in the covariance matrix. ICA, however, is a higher-order method which uses information that is not contained in the covariance matrix. ICA is a signal processing technique aimed at transforming a set of multivariate observations into statistically independent components. Therefore, the ICA model can be seen as a generative model that describes how the observed variables are generated by a process made by linearly combining these independent components.

The ICA model can be written as  $\mathbb{X} = \mathbf{A}\mathbb{Y} + \mathbf{E}$ , where  $\mathbb{X} = (X_1, X_2, \dots, X_K)'$  is the vector of  $K$  measured variables of zero mean,  $\mathbb{Y} = (Y_1, Y_2, \dots, Y_J)'$  with  $J \leq K$  is the zero mean vector of independent components,  $\mathbf{A}$  is a  $K \times J$  matrix denoted as mixing matrix, and  $\mathbf{E}$  is the residual vector. This model is known as the noise ICA model. However, the estimation of this noise model is difficult and, therefore, the great majority of ICA research and applications have focused on the following noise-free ICA model:

$$\mathbb{X} = \mathbf{A}\mathbb{Y}. \quad (1)$$

If  $J = K$ , and since  $\mathbf{A}$  is invertible, we can write the ICA model as:

$$\mathbb{Y} = \mathbf{W}\mathbb{X}, \quad (2)$$

where  $\mathbf{W} = \mathbf{A}^{-1}$  is called the demixing matrix. Using (2), the objective of ICA is to

1  
2  
3 calculate a demixing matrix  $\mathbf{W}$  so that the components  $\mathbb{Y}$  become **as independent**  
4 **as possible**. In this article, we will apply this noiseless ICA model.  
5  
6  
7

## 2.2 The estimation of the ICA transformation

8  
9  
10  
11  
12 The intuition of the estimation of the independent components in the ICA model  
13 is the following. If an observed variable  $\mathbb{X}$  is generated by a linear combination of  
14 some independent components  $\mathbb{Y}$ , it will be closer to normality than these independent  
15 components, according to the central limit theorem. Then,  $\mathbb{Y}$  will be obtained in such  
16 a way that they departure from normality as much as possible. Therefore, using the  
17 model (2), we can estimate the  $j$ -th independent component  $Y_j$  finding a vector  $\mathbf{w}_j$ , in  
18  $Y_j = \mathbf{w}_j' \mathbb{X}$ , that maximizes the non-normality of  $Y_j$ .  
19  
20  
21  
22  
23  
24  
25  
26

27 In the literature, many algorithms have been suggested in order to solve the ICA  
28 model; that is, to compute the independent components  $\mathbb{Y}$  and the demixing matrix  
29  $\mathbf{W}$  (see Bell and Sejnowsky, 1995; Amari et al, 1996; Hyvärinen, 1999, which includes  
30 a survey on the topic; Hyvärinen and Oja, 2000; Cheung and Xu, 2001; Hyvärinen et  
31 al. 2001, among others). Two of the more common measures of non-normality are  
32 kurtosis and negentropy. Kurtosis is the classical measure of non-normality but it is  
33 sensitive to outliers. Negentropy is a measure of non-normality based on the entropy.  
34 It is zero for normal variables and always nonnegative. **A simple approximation of**  
35 **the negentropy of a random variable  $y$ , denoted as  $J(y)$ , is introduced by**  
36 **Hyvärinen (1998), as  $J(y) \approx [E \{G(y)\} - E \{G(\nu)\}]^2$ , where  $\nu$  is normal with**  
37 **zero mean and unit variance and  $G$  is a non-quadratic function. By choosing**  
38  **$G$  wisely, it is possible to obtain approximations of negentropy that make**  
39 **a good compromise between kurtosis and negentropy. The functions sug-**  
40 **gested for  $G$  are  $G(y) = a_1^{-1} \log \cosh(a_1 y)$ ,  $1 \leq a_1 \leq 2$ , and  $G(y) = -\exp(-y^2/2)$ .**  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60 Based on this approximation, Hyvärinen and Oja (2000) and Hyvärinen et

1  
2  
3 al.(2001) have developed an algorithm called FastICA that is available at  
4  
5 [www.cis.hut.fi/projects/ica/fastica/](http://www.cis.hut.fi/projects/ica/fastica/).  
6

7 The FastICA algorithm uses a whitened vector  $\check{\mathbf{X}}$ ; that is, a vector with compo-  
8 nents that are uncorrelated and with unit variances. The model (2) can be written  
9 as  $\mathbf{Y} = \mathbf{B}\check{\mathbf{X}}$ , where  $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_J)'$  is now an orthogonal matrix whose compo-  
10 nents  $\mathbf{b}_i$  are vectors of unit norm. FastICA allows us to estimate the independent  
11 components one by one (deflationary orthogonalization) or all vectors  $\mathbf{b}_i$  in parallel  
12 (symmetric orthogonalization). The deflationary method has the drawback that the  
13 estimation errors of the first vectors are cumulated in the subsequent ones by orthog-  
14 onalization. Hence, symmetric orthogonalization is preferred. A description of the  
15 symmetric orthogonalization algorithm can be found below.  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

26 An important drawback of ICA estimation is that we cannot determine an ordering  
27 of the independent components, in contrast to other techniques such as PCA, where the  
28 ordering is determined by the eigenvalues. A number of methods have been suggested  
29 to determine some ordering in the independent components. For example, Hyvärinen  
30 (1999) suggests sorting the independent components according to their non-normality.  
31 He also proposes using the norm of the columns of the mixing matrix  $\mathbf{A}$  in (1) to order  
32 the independent components. Lee et al.(2004) propose, in the context of SPC, the use  
33 of the Euclidean norms of the rows of the demixing matrix  $\mathbf{W}$ . These norms give the  
34 contribution of the independent components to the variance of  $\mathbf{X}$ . Other criteria related  
35 with data reconstruction criterion have also been proposed in Cheung and Xu (2001).  
36 In this article we use the Euclidean norm of  $\mathbf{W}$  to order the independent components.  
37 This method is a simple measure and gives an ordering of independent components  
38 reminiscent of PCA.  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### 3 Principal alarms in a non-normal process

#### 3.1 General considerations

Let  $\mathbb{X} = (X_1, X_2, \dots, X_K)'$  be the vector of  $K$  non-normal variables that can be measured and used for statistical control of the process. Let us assume that  $\boldsymbol{\mu} = E(\mathbb{X}) = (\mu_1, \mu_2, \dots, \mu_K)'$  and  $\text{Cov}(\mathbb{X}) = \boldsymbol{\Sigma}$ . One of the purposes of multivariate SPC is to detect a multivariate mean shift  $\boldsymbol{\mu}^* = \boldsymbol{\mu} + \boldsymbol{\delta}$  assuming that the covariance  $\boldsymbol{\Sigma}$  remains constant. The vector  $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_K)'$  represents the shift in each original variable. Once a deviation is detected, the next goal is to determine which elements of the vector  $\boldsymbol{\delta}$  hold  $\delta_k \neq 0$ , and interpret it. The goal in the analysis of the PA is different from the latter in that the interest is not in the analysis of observed mean shifts, but in the identification of potential mean shifts  $\boldsymbol{\delta}$  that have not been observed yet. To that end, we will simulate mean shifts that, although not observed yet, are very likely to happen when the process is out of control. In this article, we will consider out-of-control situations where the mean changes, but  $\boldsymbol{\Sigma}$  remains constant. González and Sánchez (2006) also consider the case of PA with changes in the elements of  $\boldsymbol{\Sigma}$ .

As mentioned before, since the variables  $\mathbb{X}$  are dependent, we will simulate a mean shift in a set of independent variables  $\mathbb{Y} = f(\mathbb{X})$  by simply adding a vector of constants  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_J)'$  to  $\mathbb{Y}$ ; where the individual  $\gamma_j$ ,  $j = 1, \dots, J$  do not need to be related. This simulated change is translated in terms of the original variables  $\mathbb{X}$  by applying the inverse of the transformation,  $f^{-1}(\mathbb{Y} + \boldsymbol{\gamma})$ . Hence,  $\boldsymbol{\delta} = E[f^{-1}(\mathbb{Y} + \boldsymbol{\gamma})] - \boldsymbol{\mu}$ . The ICA model will be used to obtain the independent variables  $\mathbb{Y}$ .

#### 3.2 Computation of principal alarms using ICA

Let  $\mathbf{X}$  be a  $n \times K$  matrix of observations of  $\mathbb{X}$ . It should be noted that in ICA literature the data are defined as  $\mathbf{X}'$  instead of  $\mathbf{X}$ , as is customary in the SPC literature and also

here. Assuming that the process has been in control during these observations, they can be used to obtain the estimates  $\hat{\boldsymbol{\mu}}$  and  $\hat{\boldsymbol{\Sigma}}$  as  $\hat{\boldsymbol{\mu}}' = \mathbf{1}'\mathbf{X}/n$  and  $\hat{\boldsymbol{\Sigma}} = \tilde{\mathbf{X}}'\tilde{\mathbf{X}}/n$ , where  $\mathbf{1}$  is a vector of ones of appropriate dimension and:

$$\tilde{\mathbf{X}} = (\mathbf{X} - \mathbf{1}\hat{\boldsymbol{\mu}}'). \quad (3)$$

In the context of SPC, it is important to detect the most relevant mean shifts of  $\mathbb{X}$ . With this purpose in mind, the first step in the estimation of the PA will be to reduce the dimension  $K$  of the original non-normal variables  $\mathbf{X}$  to a new dimension  $J$  with  $J < K$  that captures most of the information and facilitates the analysis. PCA can be used to obtain this dimension reduction. PCA projects the variables  $\mathbf{X}$  onto a new set of variables  $\mathbf{Z}$  of dimension  $J < K$  retaining most of the original variance of  $\mathbf{X}$ . **This dimension reduction is in the very nature of SPC, since the interest is in the effect of the so-called special causes of variability, whereas the remaining variability (due to the so-called common causes) is ignored. The special causes are usually referred to as the main set of causes which provoke large variations in the process. The idea of dimension reduction in SPC has also been used in Jackson and Mudholkar (1979) and Kano et al. (2004), among others, as a way to improve the control by focusing on essential variables that drive the process.**

After this dimension reduction, it is possible to apply the ICA model (2) to obtain the  $J < K$  independent components of the main set of variables  $\mathbf{Z}$ . **These independent components will be interpreted as the main latent independent sources of variability of  $X$ .** The steps to obtain the PA are the follows:

STEP 1 Reduce the dimension  $K$  to dimension  $J$  with PCA. This step can be carried out by using the singular value decomposition of the covariance matrix  $\hat{\boldsymbol{\Sigma}}$ , or the

1  
2  
3 correlation matrix  $\hat{\mathbf{R}}$  if the variables have different magnitude. Using  $\hat{\Sigma}$ , this  
4  
5 singular value decomposition can be written as  $\hat{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{U}'$ , where  $\mathbf{U}$  is the  $K \times K$   
6  
7 matrix of eigenvectors of  $\hat{\Sigma}$ , and  $\mathbf{D}$  is the diagonal matrix with its  $K$  eigenvalues  
8  
9  $\lambda_1, \lambda_2, \dots, \lambda_K$ . Using the first  $J$  columns of this matrix  $\mathbf{U}$ , it is possible to obtain  
10  
11 the new set of variables  $\mathbf{Z}$  of dimension  $J$  as:

$$12 \quad \mathbf{Z} = \tilde{\mathbf{X}}\mathbf{U}_{(J)}, \quad (4)$$

13  
14  
15 where  $\mathbf{U}_{(J)}$  is the  $K \times J$  matrix containing the first  $J$  eigenvectors. Note that  
16  
17 since the variables  $\mathbf{X}$  are non-normal then the variables  $\mathbf{Z}$  are uncorrelated but  
18  
19 still not independent. The eigenvalue  $\lambda_j$  represents the variance of  $Z_j$ . The value  
20  
21  $J$  can be chosen so that the variability of  $\mathbf{Z}$  represents some proportion of the  
22  
23 total variability of  $\mathbf{X}$ . Applying the inverse function of (4) and using (3), the  
24  
25 matrix  $\mathbf{X}_{(1)}$ , representing an approximation of  $\mathbf{X}$  free of noise, can be computed  
26  
27 as:  
28  
29

$$30 \quad \mathbf{X}_{(1)} = \mathbf{Z}\mathbf{U}'_{(J)} + \mathbf{1}\hat{\boldsymbol{\mu}}'. \quad (5)$$

31  
32 Note that the mean vector of  $\mathbf{X}_{(1)}$  is  $\hat{\boldsymbol{\mu}}$ . We can then interpret our original dataset  
33  
34 as:  
35  
36

$$37 \quad \mathbf{X} = \mathbf{X}_{(1)} + \mathbf{X}_{(2)}, \quad (6)$$

38  
39 where  $\mathbf{X}_{(2)}$  will be assumed as noise, with zero mean and independent of  $\mathbf{X}_{(1)}$ .  
40  
41 Then, in order to simulate a mean shift on  $\mathbf{X}$  we can use the new set of uncorre-  
42  
43 lated variables  $\mathbf{Z}$ .  
44  
45  
46

47  
48 STEP 2 Obtain the  $J$  independent components of  $\mathbf{Z}$ , denoted as  $\mathbf{Y}$ , by using the noise  
49  
50 free ICA model in (1), which is:

$$51 \quad \mathbf{Z} = \mathbf{Y}\mathbf{A}, \quad (7)$$

52  
53 where  $\mathbf{A}$  is the  $J \times J$  orthogonal mixing matrix, and  $\mathbf{Y}$  is the  $n \times J$  matrix of the  
54  
55 estimated independent components. The FastICA algorithm used to obtain  $\mathbf{Y}$   
56  
57  
58  
59  
60

and  $\mathbf{A}$  needs a whitened matrix  $\check{\mathbf{Z}}$ ; that is, a matrix with variables uncorrelated and with variances equal to unity. The whitening transformation of  $\mathbf{Z}$  can be obtained using the diagonal matrix with the first largest  $J$  eigenvalues of  $\mathbf{D}$ , denoted as matrix  $\mathbf{D}_{(J)}$ , since its diagonal components  $\lambda_1, \lambda_2, \dots, \lambda_J$  are the variances of  $Z_1, Z_2, \dots, Z_J$  respectively. Then, the whitened matrix  $\check{\mathbf{Z}}$  is:

$$\check{\mathbf{Z}} = \mathbf{Z}\mathbf{D}_{(J)}^{-1/2}, \quad (8)$$

and it is easy to check that  $E(\check{\mathbf{Z}}'\check{\mathbf{Z}}) = \mathbf{I}$ . Now, using this matrix  $\check{\mathbf{Z}}$ , the ICA model (7) can be written as:

$$\check{\mathbf{Z}} = \mathbf{Y}\mathbf{H}, \quad (9)$$

where  $\mathbf{H} = \mathbf{A}\mathbf{D}_{(J)}^{-1/2}$ . This ICA model can also be written as:

$$\mathbf{Y} = \check{\mathbf{Z}}\mathbf{B}. \quad (10)$$

where  $\mathbf{B} = \mathbf{H}^{-1}$  is the  $J \times J$  demixing matrix. The matrix  $\mathbf{B}$  is obtained by using the FastICA algorithm described below. A detailed description of this algorithm can be found in Hyvärinen and Oja (2000) and Hyvärinen et al.(2001). Here we will summarize the steps of the so-called symmetric orthogonalization algorithm, which is the one used in the empirical part of this article. This method is based on a fast fixed-point optimization method of the approximated negentropy. In this description, the symbol  $\leftarrow$  means that the result on the right-hand side is assigned to the variable on the left-hand side. The steps are the following:

- (a) Initialize the  $J \times 1$  vectors  $\mathbf{b}_i$ ,  $i = 1, 2, \dots, J$  using  $J$  random numbers and scale them to unit norm. Orthogonalize the matrix  $\mathbf{B}$  as in (c) below.
- (b) For every  $i = 1, 2, \dots, J$ , let  $\mathbf{b}_i \leftarrow E \left\{ g(\check{\mathbf{Z}}\mathbf{b}_i)\check{\mathbf{Z}} \right\} - E \left\{ g'(\check{\mathbf{Z}}\mathbf{b}_i) \right\} \mathbf{b}_i$ , where  $g$  is the first derivative of  $G$  and  $g'$  is the second derivative of  $G$ .
- (c) Do a symmetric orthogonalization of the matrix  $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_J)'$  by  $\mathbf{B} \leftarrow (\mathbf{B}\mathbf{B}')^{-1/2}\mathbf{B}$ .

(d) Repeat (b)-(c) until the convergence of  $\mathbf{B}$ .

Once we obtain  $\mathbf{B}$ , we can compute  $\mathbf{Y} = \check{\mathbf{Z}}\mathbf{B}$ .

STEP 3 Determine an ordering of the independent components with the aim of determining an ordering of the PA. Use the euclidean norm of the columns of the demixing matrix, as proposed in Lee et al.(2004). The demixing matrix  $\mathbf{B}$  in terms of the variables  $\mathbf{X}$ , denoted as  $\mathbf{L}$ , can be found using (10) and (14) as follows:

$$\mathbf{L} = \mathbf{U}_{(J)}\mathbf{B}\mathbf{D}_{(J)}^{-1/2}. \quad (11)$$

The euclidean norm of each column of  $\mathbf{L}$  gives the contribution of each independent component to the variances of  $\mathbf{X}$  (note that in Lee et al., 2004, the norm corresponds to the rows of the demixing matrix because the data matrix used is  $\mathbf{X}'$  instead of  $\mathbf{X}$ ). It will be assumed in the sequel that  $\mathbf{Y}$  has columns sorted by this euclidean norm. Then, the first PA will be obtained by modifying the first column of  $\mathbf{Y}$ .

STEP 4 Introduce a mean shift in this matrix  $\mathbf{Y}$  following the methodology proposed in González and Sánchez (2006); that is, a mean shift is introduced in  $\mathbf{Y}$  by adding a constant term in some of its columns. In this way, the  $k$ -th PA will be obtained by calculating

$$\mathbf{Y}^{(k)} = \mathbf{Y} + \mathbf{1}\gamma', \quad (12)$$

where:

$$\gamma = (0, \dots, 0, c, 0, \dots, 0)', \quad (13)$$

and  $c$  is the  $k$ -th element of the vector  $\gamma$ , that determines the size of the shift.

STEP 5 Translate this mean shift in terms of the original variables  $\mathbf{X}$  by applying the suitable transformation. In order to compute this transformation, note that using

(4) expression (8) can be written in terms of  $\mathbf{X}$  as:

$$\check{\mathbf{Z}} = \tilde{\mathbf{X}}\mathbf{U}_{(J)}\mathbf{D}_{(J)}^{-1/2}. \quad (14)$$

Then, using (5), (8) and (9) we can obtain the new matrix  $\tilde{\mathbf{X}}_{(1)}^{(k)}$  that contains the data with the simulated mean shift. That is:

$$\tilde{\mathbf{X}}_{(1)}^{(k)} = \mathbf{Y}^{(k)}\mathbf{B}^{-1}\mathbf{D}_{(J)}^{1/2}\mathbf{U}'_{(J)}. \quad (15)$$

STEP 6 Since  $\tilde{\mathbf{X}}_{(1)}^{(k)} = \mathbf{X}_{(1)}^{(k)} - \mathbf{1}\hat{\boldsymbol{\mu}}'$  and the average vector of  $\mathbf{X}_{(1)}^{(k)}$  is  $\hat{\boldsymbol{\mu}}^{(k)}$ , the mean shift vector  $\boldsymbol{\delta}^{(k)} = \hat{\boldsymbol{\mu}}^{(k)} - \hat{\boldsymbol{\mu}}$  can be obtained by averaging  $\tilde{\mathbf{X}}_{(1)}^{(k)}$ . This mean shift vector  $\boldsymbol{\delta}^{(k)} = (\delta_1^{(k)}, \dots, \delta_j^{(k)})'$  is the estimation of the  $k$ -th principal alarm of  $\mathbb{X}$  given the available data. Note that  $\hat{\boldsymbol{\mu}}^{(k)}$  depends on the value of  $c$ , because a PA is not a fixed shift but a direction of change of the mean vector.

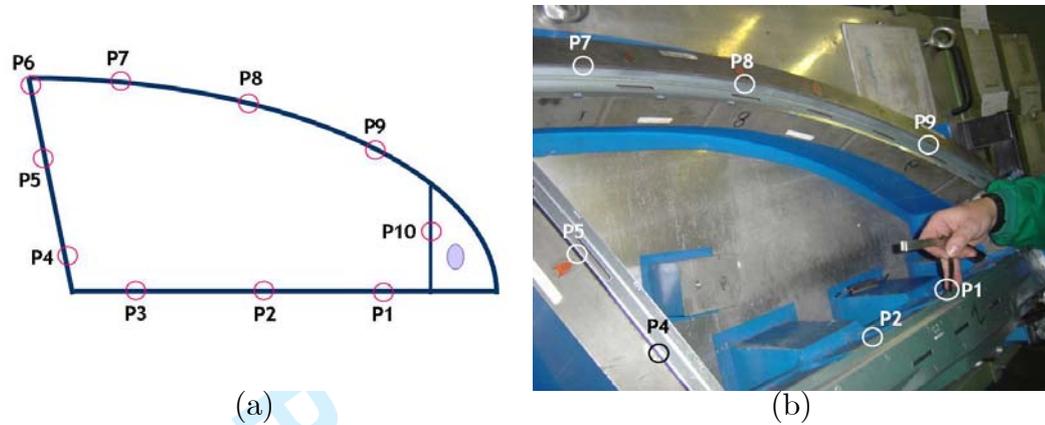
In the next section, we will apply the proposed methodology to analyse the PA using data from two real production processes.

## 4 Application in two real cases

### 4.1 Case 1: Manufacturing of a window frame

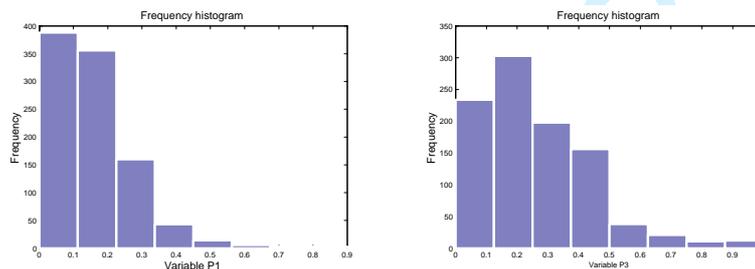
This process corresponds to the manufacturing of the window frame for the door of a vehicle. The shape of this frame has very restrictive dimensional tolerances since it must be fitted into the bodywork and it should allow the allocation of the window. The shape is obtained by a forming and stamping process (using a progressive die). The control of each frame implies the measuring of 10 points. The measurements are the gaps between the frame and a special gauge built for this purpose, as shown by the inspector that appears in Figure 1 (b). The nominal dimension of this gap is 0 and the tolerance is +1.5. Figure 1 (a) shows the profile of the frame with the 10 points.

Figure 1 (b) also shows the window frame fitted in this gauge, with some of the control points. **Figure 1 (b) also shows the inspector measuring point P1 using a low-precision measurement instrument.**



**Figure 1:** Window frame of the door of a vehicle. (a) Profile with the ten measured control points. (b) Picture of the window frame fitted in the gauge used to measure the control points.

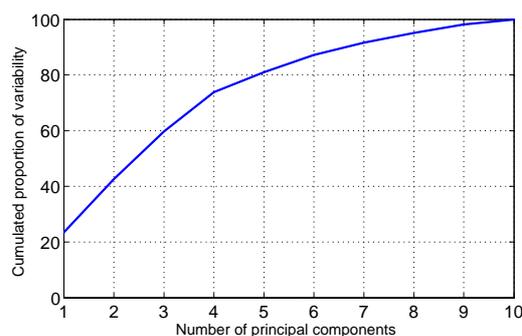
The dataset consists of 970 measured frames, obtained when the process is in control. The analysis of the data revealed that the variables were not normally distributed, showing a high positive asymmetry with long tails. Transforming data using optimal Box-Cox transformation did not produce normality either, rejecting normality of univariate and multivariate normality tests. Figure 2 shows the histogram of variables P1 and P3 and illustrates the non-normality.



**Figure 2:** Histograms of variables P1 and P3.

The first step to estimate the PA is the dimension reduction using PCA. Since the variables are of similar nature and of similar order of magnitude, the PCA was made

using the covariance matrix. Keeping in mind the goal of the PA analysis, the main latent independent sources of variability of  $X$  will correspond to the principal components with large eigenvalues. It is important to note that, as seen in Figure 1 (b), the measuring instrument used in this process is not very accurate. This lower accuracy reduces the correlation between the measured points. As a result, there is an added variability due to this effect with respect to the case of perfect measurements. From a statistical point of view, this lower correlation will induce a larger determinant of the correlation matrix, and a larger uniformity in the eigenvalues. Therefore, it will be difficult for a reduced number of principal components (say, 4 or 5) to explain a large portion of variability (for instance 95%).



**Figure 3:** Window frame of the door of a vehicle: cumulated proportion of variability as a function of the number of principal components.

Figure 3 shows the cumulated proportion of variability as a function of the number of principal components. This figure shows that the first four principal components are much larger than the rest, accounting for 74% of the variability. The fifth principal component adds only 8% of variability. Then, we will use  $J = 4$  independent components in **STEP 1** of the method presented in Section 3.2. Using FastICA, we can obtain the matrix  $B$ , of dimension  $10 \times 4$  and the matrix  $Y$ , of dimension  $970 \times 4$ , containing the four independent components (**STEP 2**). Afterwards, the four PA were

1  
2  
3  
4 generated by modifying each of these 4 independent components following  
5  
6 STEP 4. In each case the independent components were modified by adding  
7  
8 a shift  $c$  ranging in the interval  $[0, 2]$ . The ordering of the PA was done  
9  
10 according to the euclidean norm of the columns of the estimated demixing  
11  
12 matrix  $L$ , of dimension  $10 \times 4$ , shown in expression (11) of STEP 3. For  
13  
14 instance, the first column of  $L$  has the values  $l_1 = [0.1538; 0.1594; -0.5646;$   
15  
16  $-0.3256; 0.1116; -0.8431; -1.6502; -3.0604; -2.6638; 2.2850]$  and its euclidean  
17  
18 norm is 5.06. The euclidean norm of the remaining columns is, respectively,  
19  
20 5.17, 4.67 and 5.09. Then, the first PA is obtained modifying the second  
21  
22 column of  $Y$ , the second PA is obtained modifying the fourth column of  $Y$ ,  
23  
24 the third PA is related to the second column of  $Y$ , and the fourth PA is  
25  
26 related to the third column of  $Y$ .  
27

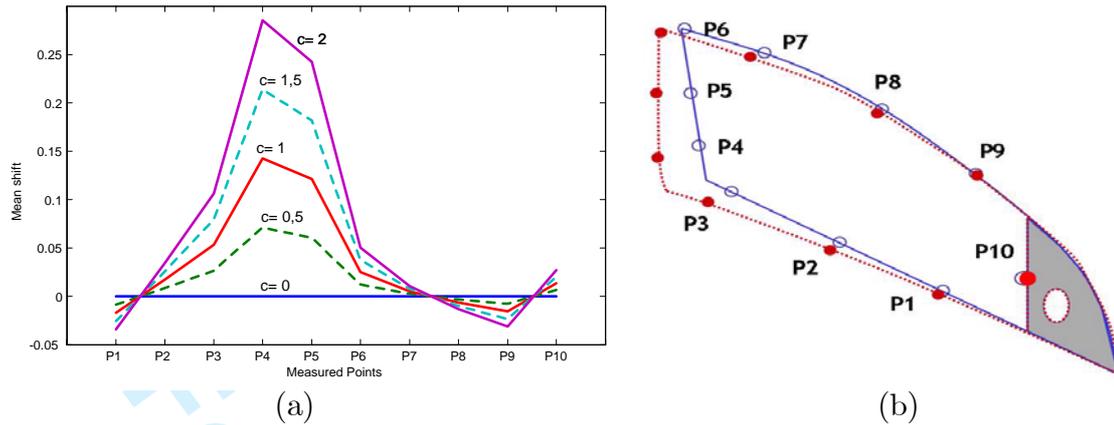
28  
29 The mean shifts  $\hat{\mu}^{(k)} - \hat{\mu}$ ,  $k = 1, \dots, 4$  are shown in Figures 4 to 7. The first PA is  
30  
31 shown in Figure 4. Figure 4 (a) shows the mean shift  $\hat{\mu}^{(k)} - \hat{\mu}$  obtained as a function  
32  
33 of  $c = 0, 0.5, 1.0, 1.5, 2.0$ . For  $c < 0$ , the mean shifts are the same but with opposite  
34  
35 sign. This figure suggests that the main shifts due to the first PA will mainly affect  
36  
37 points P4 and P5. Figure 4 (b) displays the original mean frame (solid line) with an  
38  
39 approximated representation of the deformation using some large value of  $c$ . This figure  
40  
41 can help the engineers understand and foresee potential deformations.  
42

43  
44 The second PA is shown in Figure 5. This figure shows that the main shifts due to  
45  
46 the second PA affect P9 and P10.

47  
48 The third PA is shown in Figure 6. In this case, the main shifts are seen in the  
49  
50 zone around P7, P8 and P9.

51  
52 The last PA is shown in Figure 7. The main shifts due to this PA are in P2 and  
53  
54 P3.

55  
56 It is important to note that the deformations provoked by these four PA can be  
57  
58 assumed as independent. That is, they are provoked at different stages of the forming  
59  
60



**Figure 4:** Simulated mean shifts provoked by the first PA in each measured point for different values of  $c$ . (a) Mean shift at each point. (b) Simulated deformation (dotted line) from the nominal frame (solid line) provoked by a large  $c$ .

and stamping process. These results have important practical implications, in that a single multivariate control chart based on the 10 points could be an inefficient control tool, whereas four univariate control charts might be preferred. These four univariate control charts can be based on the points that are more affected by these independent deformations; P4, P10, P8, and P2, respectively.

## 4.2 Case 2: Manufacturing of a drug substance

This case study is based on the dataset provided by González and Rodríguez (2003). This process corresponds to the manufacturing of a crystalline drug substance. The production process is done in lots by a multistep procedure. The quality of the drug substance obtained in a given lot is evaluated by the so-called impurity profile. This impurity profile consists of 7 major organic impurities, labelled from A to G, measured using a high-performance liquid chromatography method and reported in ppm values. Figure 8 shows the histogram of impurities B and C, where the non-normality of this dataset can be seen .

The dataset provided in González and Rodríguez (2003) is split in three phases. In

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

phases 1 and 3 the process can be considered to be in control, whereas in phase 2 the production process is clearly out of control. We will take advantage of this circumstance to illustrate the validity and usefulness of the proposed PA analysis. Data from phases 1 and 3 will be used to estimate the PA of the impurity profile. Then, phase 2 will be used to check if the observed departures from the in-control situation are similar to those predicted by the estimated PA. Since the order of magnitude of the impurities is very different, all variables are standardized by the standard deviation computed by using only the observations in phases 1 and 3.

#### 4.2.1 Estimation period

Phases 1 and 3 have a total of 63 lots. We will perform the PA analysis of the process using these in control observations. The first step in the PA analysis is the reduction of the dimension using PCA. **From the eigenvalues and the cumulated proportion of variability of each dimension, we chose  $J = 5$ . These 5 principal components account for 88% of the variability.** The five PA were obtained by shifting each independent component by a value  $c$  ranging in the interval  $[0, 2]$ . The ordering of the PA was done according to the euclidean norm of the columns of the demixing matrix. These euclidean norms are: 1.08, 1.04, 0.99, 0.96, and 0.81, respectively, which can be considered rather similar.

Figure 9 shows the mean shift  $\hat{\boldsymbol{\mu}}^{(k)} - \hat{\boldsymbol{\mu}}$ ,  $k = 1, \dots, 5$ , provoked by each PA as a function of  $c$ . This figure reveals the following conclusions. The first PA, shown in Figure 9 (a) has mainly two effects on the impurity profile: an increase of impurities D and E together with a decrease of impurities B and G. The second PA, showed in Figure 9 (b), is mainly an increase of impurity C, with a possible decrease of impurities E and G, but of smaller order of magnitude than the increase of C. The third PA, shown in Figure 9 (c), is clearly a large increase of impurity A. The fourth PA, shown in Figure 9 (d), implies an increase of impurity D together with a decrease of C and F.

1  
2  
3 And the fifth PA, shown in Figure 9 (e), is an increase in all the impurities except A.  
4  
5 These five PA should be analysed by the chemical experts of the process in order to  
6  
7 assign them to real causes. For instance, a PA could be related to a deficient control  
8  
9 of the temperature, or the speed in the addition of reagents.  
10

11  
12 It is important to note that impurity A is only affected by the third PA. The other  
13  
14 PAs do not produce any change in this impurity. Also, by looking at the relative  
15  
16 changes of each impurity in the PAs, a simplified control procedure might be proposed  
17  
18 by monitoring only impurities A, C and D using univariate control charts for each  
19  
20 one. The interpretation of an alarm in the control chart of A is that the third PA has  
21  
22 taken place. If only the chart of C triggers an alarm, then it might be caused by the  
23  
24 second PA. If only the chart of D is out of control, then it might be caused by the  
25  
26 first PA. However, if both C and D show a drift of the same sign in their respective  
27  
28 control charts, it might be caused by the fifth PA. Finally, if the chart C and D show  
29  
30 drifts of opposite sign, the cause could be the fourth PA. These three univariate control  
31  
32 charts would be complementary to the multivariate control chart of the 7 impurities,  
33  
34 but with the advantage that the univariate control charts will detect the most frequent  
35  
36 out-of-control situations more efficiently than the multivariate chart.  
37  
38  
39

#### 40 41 **4.2.2 Validation period**

42  
43 As reported in González and Rodríguez (2003), during phase 2 the process was out of  
44  
45 control due to lack of experience of the production personnel and deficient manufac-  
46  
47 turing instructions. The 134 lots of phase 2 will be used to analyse if the observed  
48  
49 shifts are similar to those predicted by the five estimated PA. This analysis consists of  
50  
51 comparing the observed impurity profile of each lot with the impurity profile of each  
52  
53 PA. As a result of this comparison, each lot is classified as an in-control observation,  
54  
55 as provoked by a specific PA, or as provoked by 'other alarms'. In order to perform  
56  
57 this classification, a distance of each observation to each alternative pattern (PA or  
58  
59  
60

in-control situation), as well as a decision rule to evaluate such distance have to be established. We have used the euclidean distance to measure the distance of each lot profile to each alternative pattern. The distance of the  $i$ -th lot,  $i = 1, \dots, 134$ , to the in-control pattern is then

$$D_i^c = \sum_{j=1}^7 (x_{ij} - \mu_j)^2. \quad (16)$$

Measuring the distance to a given PA is not that straightforward, since a PA is not a fixed shift but only a direction of change. The size of the shift is not determined. In order to obtain a distance to the  $k$ -th PA, we have used the mean  $\hat{\boldsymbol{\mu}}^{(k)} = (\hat{\mu}_1^{(k)}, \dots, \hat{\mu}_7^{(k)})$ ,  $k = 1, \dots, 5$ , obtained with  $c = 2$  in (12) using the data of phases 1 and 3. Then, the distance of observation  $X_i$  to a shift provoked by the  $k$ -th PA is computed as

$$D_i^{(k)} = \sum_{j=1}^7 \left( x_{ij} - \hat{\alpha}_{ik} - \hat{\beta}_{ik} \hat{\mu}_j^{(k)} \right)^2. \quad (17)$$

where

$$\hat{\beta}_{ik} = \frac{\sum \tilde{x}_{ij} \tilde{\mu}_j^{(k)}}{\sum \left( \tilde{\mu}_j^{(k)} \right)^2}, \quad (18a)$$

$$\hat{\alpha}_{ik} = \tilde{x}_{ij} - \hat{\beta}_{ik} \tilde{\mu}_j^{(k)}, \quad (18b)$$

$$\tilde{x}_{ij} = x_{ij} - \frac{\sum_{j=1}^7 x_{ij}}{7}, \quad (18c)$$

$$\tilde{\mu}_j^{(k)} = \hat{\mu}_j^{(k)} - \frac{\sum_{j=1}^7 \hat{\mu}_j^{(k)}}{7}. \quad (18d)$$

The evaluation of the distances  $D_i^c$  and  $D_i^{(k)}$ ,  $k = 1, \dots, 5$  for each observation in phase 2 is done using their empirical distributions computed with the 63 observations of phases 1 and 3. These empirical distributions are computed as follows. For the in-control case, the 63 values of  $D_i^c$  are computed as in (16). Let us denote as  $\hat{F}_c$  the empirical cumulative distribution function of  $D_i^c$ . For each PA, the 63 values of  $D_i^{(k)}$  are computed by provoking a shift in the  $k$ -th PA with  $c = 2$ . Then, 18 is computed using  $x_{ij} + \delta_j^{(k)}$  instead of  $x_{ij}$ . The empirical cumulative distribution function of  $D_i^{(k)}$  is denoted as  $\hat{F}_k$ . For each observation in phase 2, we compare the distances  $D_i^c$  and

1  
2  
3  $D_i^{(k)}$ ,  $k = 1, \dots, 5$ , with their empirical cumulative distributions  $\hat{F}_c$  and  $\hat{F}_k$ , respectively.  
4  
5  
6 If all the distances are larger than their respective 95th percentile, the observation is  
7  
8 classified as produced by 'other alarms'. Otherwise, the observation is classified as  
9  
10 produced by the pattern that has minimum empirical cumulative distribution. For  
11  
12 instance, an observation is classified as in-control if  $\hat{F}_c(X_i) < \hat{F}_k(X_i)$ ,  $k = 1, \dots, 5$ .  
13

14 According to this classification rule, the 134 observations of phase 2 have been  
15  
16 classified as follows: 31 observations (23%) have been classified as being in control; 69  
17  
18 observations (51%) have been classified as provoked by the third PA; 15 observations  
19  
20 (11%) have been classified as provoked by the fourth PA; and 19 observations (14%)  
21  
22 have been classified as provoked by 'others' alarms. The solid lines in Figure 10 show  
23  
24 the average mean shift of the 69 observations classified as belonging to the third PA,  
25  
26 and the average mean shift of the 15 observations classified as belonging to the fourth  
27  
28 PA. This figure also displays (dotted line) the mean shift of the third and fifth PA,  
29  
30 respectively, as also seen in Figure (9) for  $c = 2$ . Both lines show the similarity between  
31  
32 the estimated PAs and the observed shifts.  
33

34  
35 Figure 10 (a) shows a large shift in impurity A and very small shifts in the remaining  
36  
37 impurities, probably due to randomness. This pattern is similar to the one foreseen by  
38  
39 the third PA. Figure 10 (b) shows a large shift in impurity D and shifts of opposite  
40  
41 sign in impurities C and F, just as predicted by the fourth PA.  
42  
43  
44  
45

## 46 **5 Concluding remarks**

47  
48

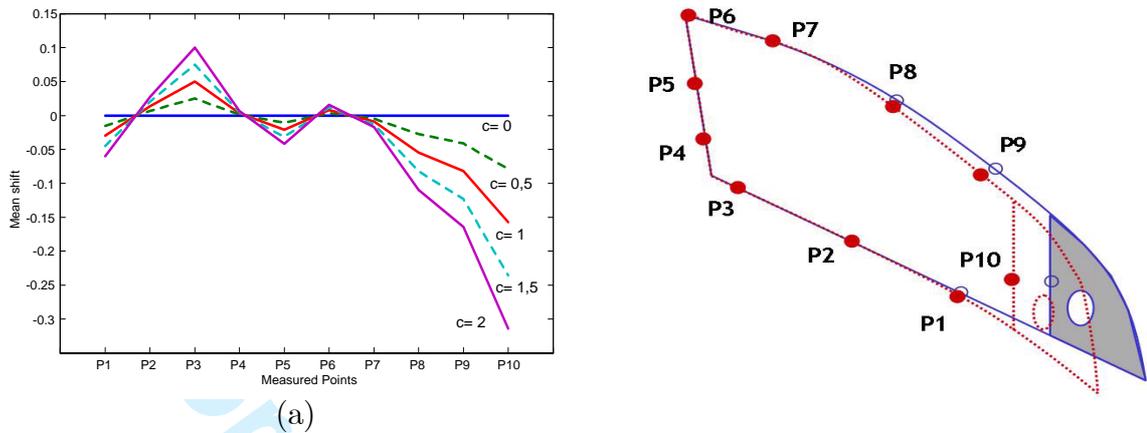
49 The PA analysis proposed by González and Sánchez (2006) for normal processes is  
50  
51 extended in this article to non-normal processes by using ICA. Non-normal processes  
52  
53 can be very frequent in real applications. For instance, processes where the measures  
54  
55 of quality are related to non-negative deviations from some ideal levels can produce  
56  
57 highly asymmetric distributions. This is the case, for instance, of the manufacturing of  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

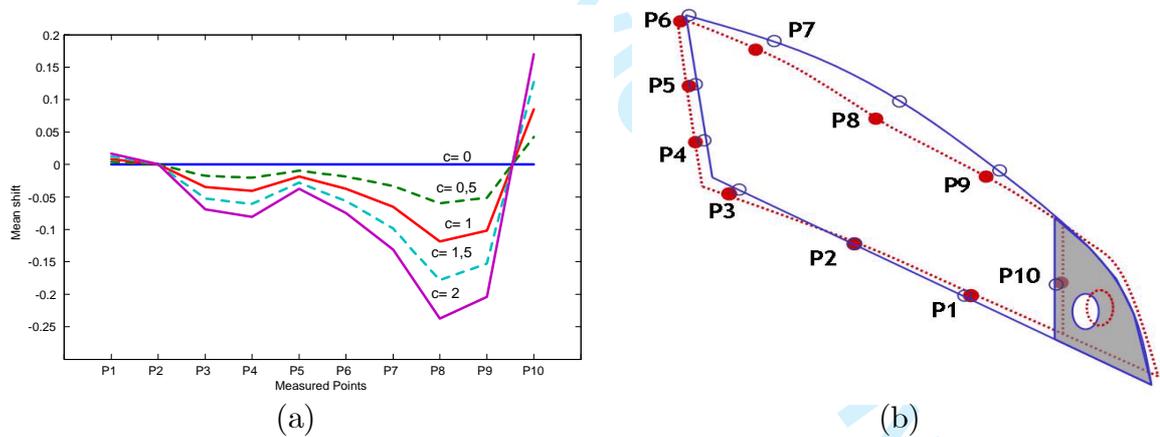
mechanical parts, like the one in Section 4.1, where one-side tolerances are established. Other examples can be found in chemical processes such as the one shown in Section 4.2, where the impurity profile can only have positive deviations from a zero level. The multivariate analysis of such processes has the added complexity of the non-normality. In those cases, the analysis of PA using ICA helps to foresee the potential deviations of the process using only in-control data. That is, we do not need to be out of control to have some insights about what is going to happen. The analysis of PA is related to the principles of the so-called Pareto analysis, since it helps to identify a small number of alarms that will account for most of the out-of-control situations. Such analysis can be used for many purposes: to identify weak points in the design of the process, to design more specific control charts, and to facilitate the interpretation of out-of-control signals.

## 6 Acknowledgements

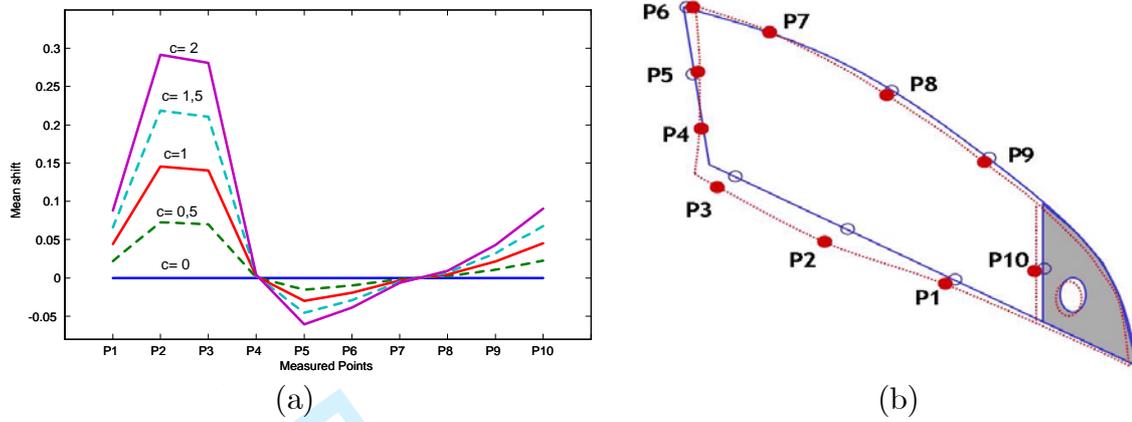
Authors are grateful to Alcalá Industrial SA for providing the case study, the data, and the permission for using Figure 1 (b). We are especially grateful to his Quality Engineer, Jose A. Delgado-Echague, for his comments and support. González's research is partly supported by the CICYT grant DPI2005-08018. Sánchez's research is partly supported by the grant CAM 06/HSE/0174/2004 and the CICYT grant SEJ2004-03303. Authors are also grateful to the anonymous referee for insightful comments that significantly improved the manuscript.



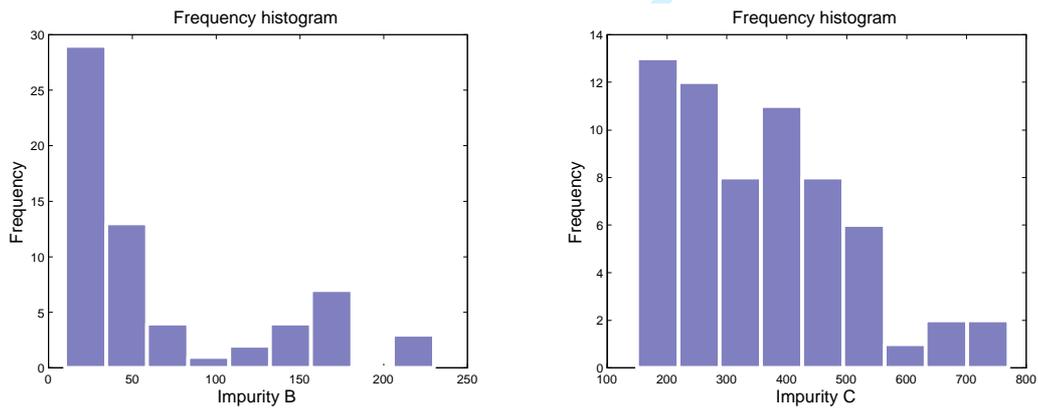
**Figure 5:** Simulated mean shifts provoked by the second PA in each measured point for different values of  $c$ . (a) Mean shift at each point. (b) Simulated deformation (dotted line) from the nominal frame (solid line) provoked by a large  $c$ .



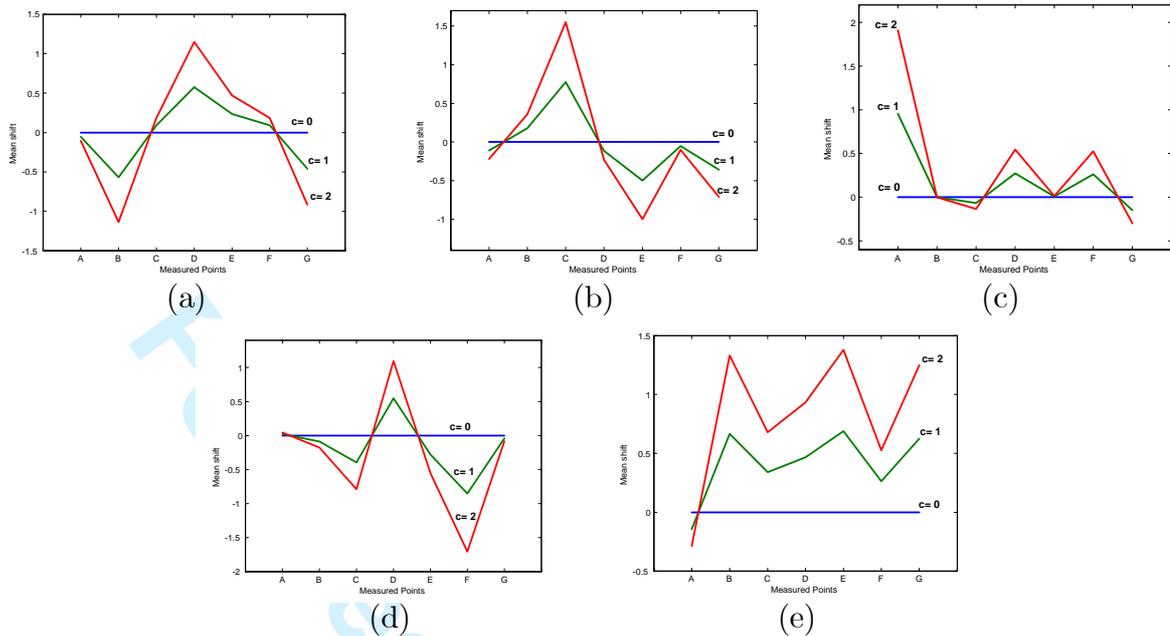
**Figure 6:** Simulated mean shifts provoked by the third PA in each measured point for different values of  $c$ . (a) Mean shift at each point. (b) Simulated deformation (dotted line) from the nominal frame (solid line) provoked by a large  $c$ .



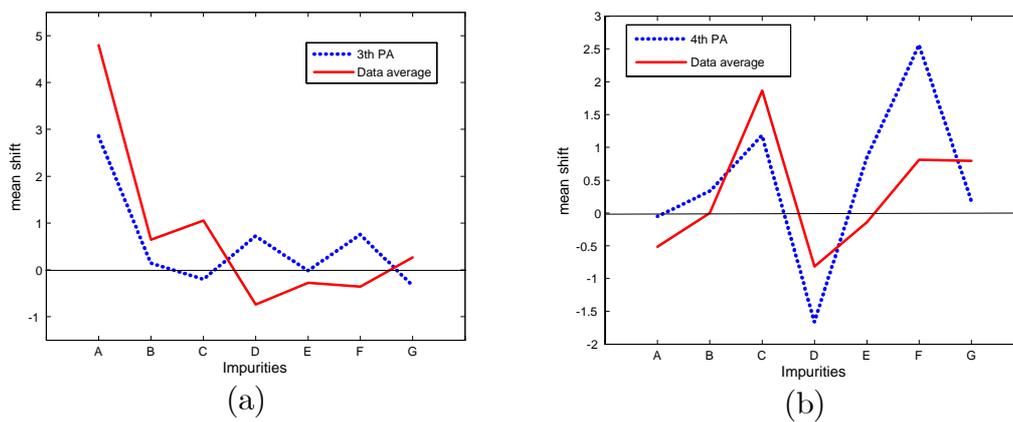
**Figure 7:** Simulated mean shifts provoked by the fourth PA in each measured point for different values of  $c$ . (a) Mean shift at each point. (b) Simulated deformation (dotted line) from the nominal frame (solid line) provoked by a large  $c > 0$ .



**Figure 8:** Histograms of impurities B and C.



**Figure 9:** Simulated mean shifts provoked by the five PA in each impurity for different values of  $c > 0$ . (a) First PA. (b) Second PA. (c) Third PA. (d) Fourth PA. (e) Fifth PA.



**Figure 10:** Solid lines: average mean shift of the observations classified as belonging to: (a) the third PA (a) and (b) to the fourth PA. Dotted lines: average mean shift of: (a) the third, and (b) the fifth PA.

## References

- [1] Alt, F.B., Multivariate quality control. In *Encyclopedia of Statistical Science*, edited by S.Kotz and N. Johnson, pp. 110-122.,1985 (New York: John Wiley: New York).
- [2] Amari, S., Cichocki, A., and Yang, H., A new learning algorithm for blind signal separation. *Advances of Neural Information Processing System*, 1996, **8**, 757-763.
- [3] Bell, A. and Sejnowski, T., A Information-maximization approach to blind separation and blind deconvolution. *Neural Computing*, 1995, **7**, 1129-1159.
- [4] Chen, G., Liang, J., and Qian, J., Process monitoring and fault detection based on multivariate statistical projection analysis, *IEEE International Conference on Systems, Man, and Cybernetics*, 2004, 2719-2723.
- [5] Cheung, Y. and Xu, L., Independent component ordering in ICA time series analysis. *Neurocomputing*, 2001, **41**, 145-152.
- [6] Crosier, R.B., Multivariate generalizations of cumulative sums quality-control schemes. *Technometrics*, 1988, **30**, 291-303.
- [7] Doganaksoy, N., Faltin, F.W., and Taker, W.T., Identification of out-of-control quality characteristics in a multivariate manufacturing environment. *Communications in Statistics-Theory and Methods*, 1991, **20**, 2775-2790.
- [8] González, M, and Rodríguez, P., Application of the multivariate  $T^2$  control chart and the Mazon-Tracy-Young decomposition procedure to the study of the consistency of impurity profiles of drug substances. *Quality Engineering*, 2003, **16**, 127-142.

- 1  
2  
3  
4 [9] González, I. and Sánchez, I., Principal alarms In multivariate statistical process  
5 control. Working paper, Universidad Carlos III de Madrid, Spain, 2006. Available  
6 on line at docubib.uc3m.es/WORKINGPAPERS/WS/ws061504.pdf.  
7  
8  
9  
10  
11 [10] Hawkins, D.M. Multivariate quality control based on regression-adjusted variables.  
12 *Technometrics*, 1991, **33**, 61–75.  
13  
14  
15 [11] Hawkins, D.M. Regression adjustment for variables in multivariate quality control.  
16 *Journal of Quality Technology*, 1993, **25** 170–182.  
17  
18  
19 [12] Hyvärinen, A., New approximations of differential entropy for independent compo-  
20 nent analysis and projection pursuit. *Advances of Neural Information Processing*  
21 *System*, 1998, **10**, 273-279.  
22  
23  
24 [13] Hyvärinen, A., Survey of independent component analysis. *Neural Computing*,  
25 1999, **2**, 94-128.  
26  
27  
28 [14] Hyvärinen, A. and Oja, E., Independent component analysis: algoritms amd  
29 applications. *Neural Networks*, 2000, **13**, 411-430.  
30  
31  
32 [15] Hyvärinen, A., Karhunen, J., and Oja, E., *Independent component analysis*, 2001,  
33 (John Wiley and Sons: New York).  
34  
35  
36 [16] Hotelling, H., Mutivariate quality control. Illustrated by the air testing of sample  
37 bombsights. In *Techniques of Statistical Analysis*, edited by C. Eisenhart, M.W.  
38 Hastay, and W.A. Wallis, pp. 111-184, 1947 (McGraw-Hill: New York).  
39  
40  
41 [17] Jackson, J.E. , Principal components and factor analysis: part1-principal compo-  
42 nents, *Journal of Quality Technology*, 1980, **12**, 201-213.  
43  
44  
45 [18] Jackson, J.E., Multivariate quality control, *Communications in Statistics, Theory*  
46 *and Methods*, 1985, **14**, 2657-2688.  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3  
4 [19] Jackson, J.E., and Muldholkar, G.S., Control procedures for residuals associated  
5 with principal component analysis. *Technometrics*, 21, 341-349.  
6  
7  
8  
9 [20] Jones, M.C. and Sibson R., What is projection pursuit?. *Journal of the Royal*  
10 *Statistical Society A*, 1987, **150**, 1-36.  
11  
12  
13 [21] Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H., Evolution of multivariate  
14 statistical process control: application of independent component analysis and  
15 external analysis. *Computers & Chemical Engineering*, **28**, 1157-1166.  
16  
17  
18 [22] Lee, T., *Independent Component Analysis: Theory and Application*, 1998 (Kluwer  
19 Academic Publishers: Boston).  
20  
21  
22 [23] Lee, J.M., Yoo, C, and Lee, I., Statistical process monitoring with independent  
23 component analysis. *Journal of the Process Control*, 2004, **14**, 467-485.  
24  
25  
26 [24] Lowry, C.A., Woodall, W.H., Champ, C.W., and Rigdon, S.E., A multivariate  
27 exponentially weighted moving average control chart. *Technometrics*, 1992, **34**,  
28 46–53.  
29  
30  
31 [25] Mason, R.L., Tracy, N.D., and Young, J.C., The composition of  $T^2$  for multivariate  
32 control chart interpretation. *Journal of Quality Technology*, 1995, **27**, 99-108.  
33  
34  
35 [26] Mason, R.L., Tracy, N.D., and Young, J.C., A practical approach for interpreting  
36 multivariate  $T^2$  control chart signals. *Journal of Quality Technology*, 1997, **29**,  
37 396–406  
38  
39  
40 [27] Murphy, B.J., Selecting out-of-control variables with the  $T^2$  multivariate quality  
41 control procedures. *The Statistician*, 1987, **36**, 571–583.  
42  
43  
44 [28] Wade, M. R. and Woodall, W. H., A review and analysis of cause selecting charts.  
45 *Journal of Quality Technology*, 1993, **25**, 3, 161-169.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

[29] Woodall, W.H., and Ncube, M.M., Multivariate CUSUM quality-control procedures. *Technometrics*, 1985, **27**, 285–292.

For Peer Review Only