Short-Run Multivariate Control Charts for Process Mean and Variability

L. Jaupi, D. E. Herwindiati, Ph. Durand, and D. Ghorbanzadeh

Abstract—**Statistical process control methods for monitoring short-run processes with multivariate measurements are considered and new multivariate short-run control charts to monitor process mean and variability are proposed. To monitor the process mean the influence function of mean is proposed and to investigate process variability control charts based on the influence function of eigenvalues are suggested. The proposed techniques are general, and the influence functions may be used to build up short-run multivariate control charts relative to either nominal values or estimates. The method is further illustrated with real datasets, from a flexible job shop manufacturing system producing spare parts for classical cars.**

Index Terms—**Eigenvalue, influence function, statistical process control, principal components**

I. Introduction

STANDARD multivariate control charts are powerful techniques that have been applied effectively on continuous and large batch size processes. Another trend in manufacturing characterized by a high degree of flexibility is job shop production. A typical example would be a multi-purpose CNC machining centre, which produces small batches of complex parts for local customers. In a short-run environment, it is difficult or perhaps impossible to gather data necessary to implement traditional multivariate SPC techniques, because production runs are usually short and change frequently from one part to another. For univariate short-run manufacturing processes, numerous control charts have been introduced to overcome the problems faced by conventional charts, [1], [2], [3], [4], [7], [8], [10], [11]. But more realistic scenarios involve measurements of several related variables. Another problem encountered in a short-run production of complex parts is that there are many different types of measurements and the number of monitored variables is not the same from one batch to another. So standardized control charts are necessary to present data from different runs in a single analysis.

L. Jaupi is with Conservatoire National des Arts et Métiers, 292 rue Saint Martin, 75003 Paris, France (corresponding author phone: +33 1 40 27 23 73; fax: +33 1 40 27 27 46; e-mail: jaupi@cnam.fr).

D. E. Herwindiati is with Tarumanagara University, Jakarta, Indonesia (e-mail: herwindiati@untar.ac.id).

Ph. Durand is with Conservatoire National des Arts et Métiers, Paris, France (e-mail: durand@cnam.fr).

D. Ghorbanzadeh is with Conservatoire National des Arts et Métiers, Paris, France (e-mail: ghorbanzadeh@cnam.fr).

For multivariate short-run environment, Quesenberry in [9] suggested the use of snapshot Q chart by plotting all Q statistics on the same chart. KHOO et al., in [5] proposed short-run multivariate control charts to monitor process mean and in [6] short-run multivariate control charts to monitor process variability based on individual measurements and sub-grouped normal data. When there is a lack of or limited knowledge about the underlying process distribution, nonparametric rank-based control charts are useful for short-run or start-up situations, [12].

In a multivariate process, when assignable causes are present, they may affect different process parameters: process mean, and/or process variability, and/or process orientation. Indeed, a multivariate quality characteristic vector posses both magnitude and direction. In this paper we propose new multivariate short-run control charts to monitor process mean and variability. To monitor the process mean the influence function of mean is proposed, to investigate process variability control charts based on the influence functions of eigenvalues are suggested and to describe process orientation control charts based on the influence functions of eigenvectors may be employed. The proposed techniques are general, and the influence functions may be used to build up short-run control charts for any process parameter.

To make the presentation clear the remainder of the paper is organized as follows: in Section II we introduce the influence function; control charts based on influence functions are presented in Section III; an application is given in Section IV; remarks on the use of the influence function for process monitoring and possible extensions complete the paper in Section V.

II. Influence function

A. Formulation

In what follows we suppose that $X = (X_1, X_2, \ldots, X_p)$ is the vector of the measured variables made on a given part. We assume that under a stable process the distribution of X is F with mean μ and covariance matrix Σ, ideally multivariate normal. When special causes are present in the process X has an arbitrary distribution noted G. A distribution function which describes the two sources of variation in a process is the contaminated model, [13], given by:

$$F_{\varepsilon H} = (1 - \varepsilon) F + \varepsilon G \tag{1}$$

with $0 \leq \varepsilon \leq 1$.

ISBN: 978-988-19251-0-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

If process is under control we have ε=0. When process is not stable, roughly a proportion ε of output subgroups will be contaminants.

B. Influence Function

Let T=T(F) be a statistical functional. The influence function $IF(x,T,F)$ of the statistical functional T at F is defined as the limit as $\varepsilon \to 0$ of

$$\left\{T\left[(1-\varepsilon)F+\varepsilon\delta_x\right]-T(F)\right\}/\,\varepsilon \quad (2)$$

where δ_x denotes the distribution giving unit mass to the point $x \in R^p$. The perturbation of F by δ_x is denoted as

$$F_{\varepsilon x}=(1-\varepsilon)F+\varepsilon\delta_x \quad (0\le\varepsilon\le 1) \quad (3)$$

As such the influence function measures the rate of change of T as F is shifted infinitesimally in the direction of δ_x, [13]. The importance about the influence function lies in its heuristic interpretation: it describes the effect of an infinitesimal contamination at point x on the estimate. Our idea is that output segments that have a large influence on monitored parameters show up the time when special causes are present in a manufacturing process. The influence functions may be calculated for almost all process parameters. Therefore, based on influential measures derived from them, multivariate control charts for different process parameters and with different sensitivities are be set up.

C. Influence Measures for Classical Estimators

Let (μ,Σ) denote the location scale parameter defined by

$$\int(X-\mu)dF=0$$
$$\int[(X-\mu)(X-\mu)^t-\Sigma]dF=0 \quad (4)$$

In order to calculate the influence function of the location scale parameter (μ,Σ), we substitute F by $F_{\varepsilon x}$ in (4) and take the derivative with respect to ε at $\varepsilon=0$. The differentiation of the mean equation gives

$$IF(x,\mu,F)=x-\mu \quad (5)$$

and the differentiation of the covariance matrix equation gives

$$IF(x,\Sigma,F)=(x-\mu)(x-\mu)^t-\Sigma \quad (6)$$

We assume that Σ has distinct eigenvalues $\lambda_1>\lambda_2>...>\lambda_p$ and we denote by $\alpha_1,\alpha_2,...,\alpha_p$ the associated eigenvectors. Under regularity conditions, (cf. [14], [15], [16]), we find the following expressions for the influence function of the j[th] eigenvalue and the associated eigenvector of the covariance matrix respectively

$$IF(x,\lambda_j,F)=\alpha_j^t IF(x,\Sigma,F)\alpha_j \quad (j=1,\dots,p)$$
$$IF(x,\alpha_j,F)=\sum_{k=1,k\ne j}^{p}(\lambda_j-\lambda_k)^{-1}\alpha_k\alpha_k^t IF(x,\Sigma,F)\alpha_j \quad (7)$$

Generally in the applications of the influence function the unknown distribution function F has to be estimated by $\hat{F}$ the empirical distribution function based on a random sample $x_1,x_2,...,x_n$ from F. Replacing F by $\hat{F}$ and taking $x=x_i$ in (5) and (7) we have for the empirical influence function of the eigenvalues and the eigenvectors the following expressions

$$IF(x_i,\mu,\hat{F})=x_i-\hat{\mu}$$

$$IF(x_i,\lambda_j,\hat{F})=\hat{\alpha}_j^t\left[(x_i-\hat{\mu})(x_i-\hat{\mu})^t-\hat{\Sigma}\right]\hat{\alpha}_j \quad (8)$$

$$IF(x_i,\alpha_j,\hat{F})=\sum_{k=1,k\ne j}^{p}(\hat{\lambda}_j-\hat{\lambda}_k)^{-1}\hat{\alpha}_k\hat{\alpha}_k^t\left[(x_i-\hat{\mu})(x_i-\hat{\mu})^t-\hat{\Sigma}\right]\hat{\alpha}_j$$

III. Control charts based on Influence function

A. Control Charts for Process Mean

When one is monitoring the process mean, what is calculated and plotted on a control chart is the value of a quadratic form. In general, a quadratic form can be written as Y^tMY, where M is a definite positive matrix and $Y \in R^P$. We use as vector Y the influence function of process mean. To build up control charts for the phase I we take as matrix M the inverse of sample covariance matrix: $M=\hat{\Sigma}^{-1}$. Then the quadratic form Y^tMY gives a sample version of Hotelling's T^2 statistic. For the i[th] subgroup, (i=1,...,n), one would calculate and plot in the control chart the quantity

$$T_i^2=(x_i-\hat{\mu})^t\hat{\Sigma}^{-1}(x_i-\hat{\mu})=(x_i-\overline{x})^t\hat{\Sigma}^{-1}(x_i-\overline{x}) \quad (9)$$

Each of T^2 values, (i=1,...,n), of (9) would be compared with

$$UCL=\frac{(n-1)p}{(n-p)}F_{\alpha,p,n-p} \quad (10)$$

The upper control limit is an approximation however, that should be used only for a large number of observations. When the standards are known, they are used to build up the control charts. In this case the quadratic form Y^tMY has a χ^2 distribution with p degrees of freedom. The control charts that are based on T^2 statistic or χ^2 distribution are efficient as long as the number of quality characteristics is not greater than four. If the number of variables is greater

ISBN: 978-988-19251-0-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

than four only huge shifts can be detected by using these charts.

B. Control Charts for Process Variability

The total variance of X is equal to the sum of the eigenvalues of the covariance matrix Σ

$$Var(X) = \lambda_1 + \lambda_2 + ... + \lambda_p \quad (11)$$

Assignable causes that affect the variability of the output do not increase significantly each component of total variace of X. Instead, they may have a large influence in the variability of some components and small effect in the remaining directions. Therefore an approach to design control charts for variability consists to detect any significant departure from the stable level of the variability of each component. To build up such control charts one may use either the principal components or the influence functions of the eigenvalues of the dispersion matrix. That is, if one wants to monitor the process variability according to the j^{th} direction, (j=1,...,p), what would be calculated and plotted on a control chart are the values of the influence function of the j^{th} eigenvalue of the covariance matrix Σ. Thus, for the i^{th} subgroup, (i=1,...,n), one would calculate and plot

$$IF(x_i, \lambda_j, \hat{F}) = C_{ij}^2 - \hat{\lambda}_j \quad (12)$$

where C_{ij} is the j^{th} principal component, (j=1,...,p), of the i^{th} observation. The control limits are three sigma control limits as in any Shewhart control chart.

C. Control Charts for Short-Runs

For job-shop manufacturing systems control charts based on influence functions are straightforward extensions of conventional multivariate control charts. To combine data from several runs into a single analysis it will be necessary to standardize measurements and standardize also the control chart limits according to the properties of each run. Different strategies may be employed. Here we present one of them:

Step 1: Calculate the influence functions of a multivariate location-scale parameter for the class of M-estimators (classical or pseudo-covariance estimators). If the process parameters and the number of quality characteristics used to describe each item quality are the same for each run, then use the influence functions without any further standardization. Else,

Step2: Standardize the influence functions as follows:

$$Q_{\lambda j}(x_i) = IF(x_i, \lambda_j) / \sqrt{\lambda_j} \quad (13)$$

$$Q_T(x_i) = C_p^{-1} \|Y_i\|^2 \quad (14)$$

where

$$C_p = \chi^2_{1-\alpha,p} \quad (15)$$

$$Y = A \cdot IF(x, \mu) \quad (16)$$

$$Var(IF(x, \mu) = A^{-t} A^{-1} \quad (17)$$

Step3: Set up control charts for process variability and mean.

IV. Application

A. Case Study

The case study considers a flexible multi-purpose CNC machining centre, which produces small batches of spare parts for classical cars. To illustrate the procedure we consider the data from two types of parts manufactured at the center: pump bodys and crankcase caps. We will refer them as part A and part B respectively. The first order from an auto-club of classical cars consists on manufacturing a batch of 25 pump bodies and 14 crankcase caps. For part A there are 21 quality characteristics, but only 3 are important and checked at all items that are produced. These quality characteristics are called *intrinsic quality characteristics*. For the part B there are 16 quality characteristics but only three are considered as intrinsic quality characteristics. The influential charts based on empirical influence functions for the first, second and third eigenvalue of the covariance matrix are displayed in Fig. 1, Fig. 2 and Fig. 3 respectively. Average influence is zero. The standardized three sigma control limits of influential charts of eigenvalues are: 0 ± 3√2. The statistic that is plotted in the chart is given in (13). The control, chart based on the influence function of mean is displayed in Fig. 4. The statistic that is plotted in the chart is given in (14). It's standardized upper control limit is equal to 1.

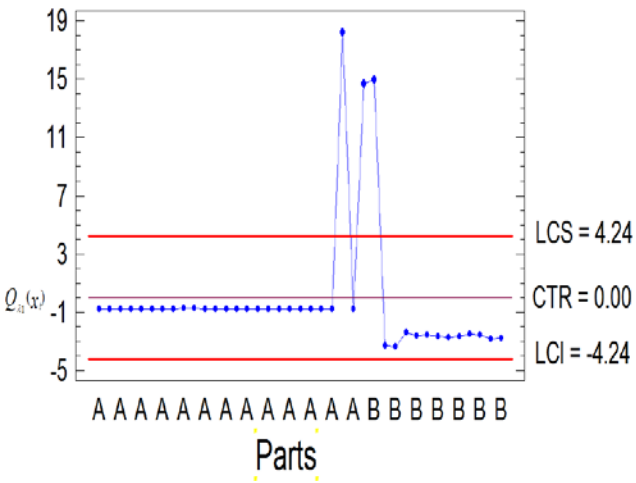

Fig. 1. Control chart for short-runs based on the standardized influence functions of the first eigenvalue.

ISBN: 978-988-19251-0-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

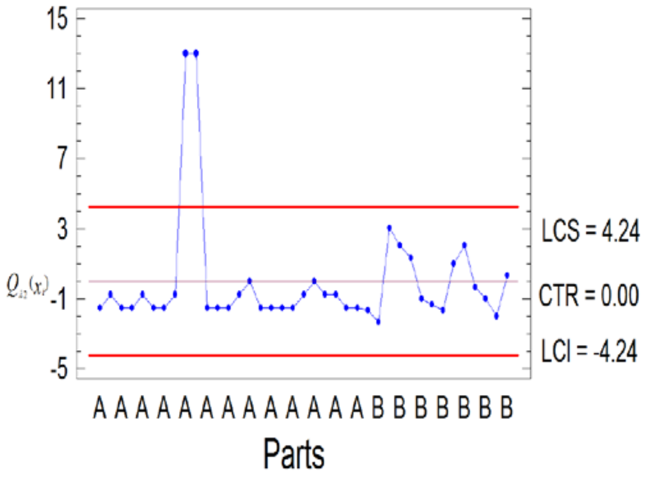

Fig. 2. Control chart for short-runs based on the standardized influence functions of the second eigenvalue.

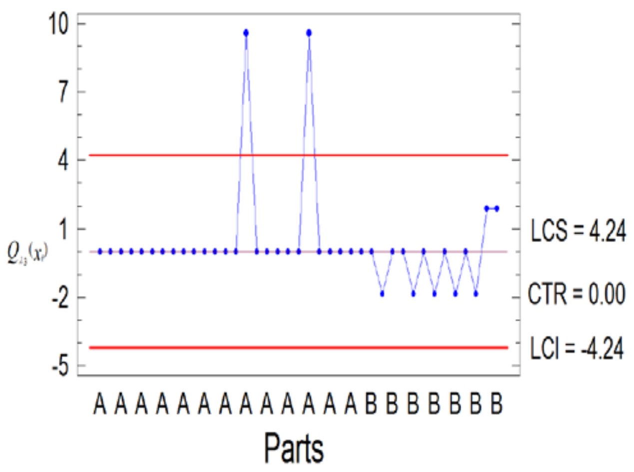

Fig. 3. Control chart for short-runs based on the standardized influence functions of the third eigenvalue.

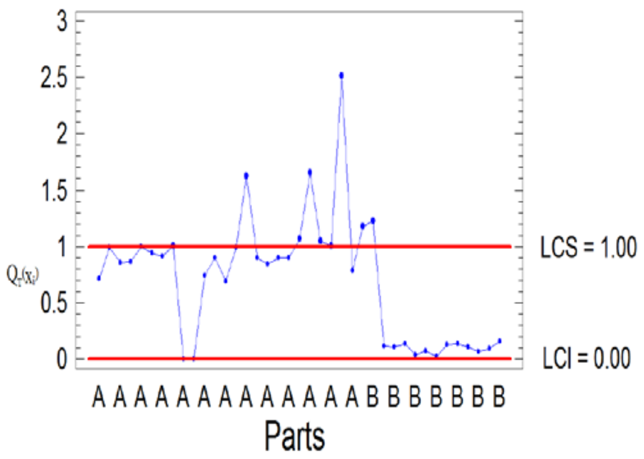

Fig. 4. Multivariate control chart to monitor the process mean.

B. Interpretation of results

An inspection of these charts shows that: there are three very highly influential subgroups for the first component of the total variance; there are two influential points for the second eigenvalue and two influential points for the third component of the total variance; the control chart for the process mean shows that the operator has difficulties during the manufacturing of part A and in the early beginning of part B, but then the process seems under control. In process logbook there are clear explanations for all these assignable causes.

V. Remarks

The influence function may be calculated for all real situations on the background of robust statistics either for classical estimators or for robust estimators. Therefore based on the influence functions, or on influential measures that may be derived from them, control charts for different process parameters and with different sensitivities may be set up such as: control charts for process mean-variability-orientation or the structure of relationships between the variables; control charts for attributes and CUSUM charts for measurable data as well as for attributes.

The control charts that are based on T^2 statistic or χ^2 distribution are efficient as long as the number of quality characteristics is not greater than four. If the number of intrinsic quality characteristics is greater than four only huge shifts can be detected by using these charts. Therefore for processes where huge amounts of multidimensional data are available, the reduction of their dimensionality in terms of space and in terms of original variables should be considered.

The influence functions are not scale invariant, therefore someimes it is better to apply the above method to standardized data rather than to raw data. In such cases the covariance matrix is replaced by the corresponding correlation matrix.

References

[1] G. Capizzi and G. Masarotto, "An Enhanced Control Chart for Start-Up Processes and Short Runs", *Quality Technology & Quantitative Management*, 2012, Vol. 9, No. 2, pp. 189-202.
[2] G. Celano, Ph. Castagliola, E. Trovato and S. Fichera, "Shewhart and EWMA t Control Charts for Short Production Runs", *Quality and Reliability Engineering International*, 2011, Vol. 27, n° 3, pp. 313–326.
[3] E. Del Castillo, J.M. Grayson, D.C. Montgomery, G.C. Runger, "A review of statistical process control techniques for short run manufacturing systems". *Communications in Statistics—Theory and Methods*, 1996, Vol. 25, pp. 2723–2737.
[4] L. L. Ho & A. L. G. Trindade, "Economic design of an X chart for short-run production", *International Journal of Production Economics*, 2009, Vol. 120, n° 2, pp. 613–624.
[5] B.C. Khoo and S.H. Quah, "Proposed Short Runs Multivariate Control Charts for the Process Mean", *Quality Engineering*, 2002, Vol. 14, n° 4, pp. 603-621.

ISBN: 978-988-19251-0-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

[6] B.C. Khoo, S.H. Quah, H.C. Low, C.K. Chung' "Short runs multivariate control chart for process dispersion", *International Journal of Reliability, Quality and Safety Engineering*, 2005; Vol. 12, No. 2, pp.127–147.

[7] G. Nenes and G. Tagaras, "The economically designed CUSUM chart for monitoring short production runs", *International Journal of Production Research*, 2006, Vol. 44, pp.1569–1587.

[8] G. Nenes and G. Tagaras, "The CUSUM chart for monitoring short production runs", *Proceedings of the 5th International Conference on Analysis of Manufacturing Systems Production Management*, Zakynthos, Greece, 2005, pp. 43–50.

[9] C. P. Quesenberry, "The Multivariate Short-Run Snapshot Q Chart", *Quality Engineering*, 2001, Vol. 13 No. 4, pp. 679-683.

[10] C.P. Quesenberry, "On properties of Q charts for variables (with discussion)", *Journal of Quality Technology*, 1995, Vol. 27, pp. 204–213.

[11] C.P. Quesenberry, "SPC Q charts for start-up processes and short or long runs", *Journal of Quality Technology*, 1991; Vol. 23, pp. 213–224.

[12] C. Zou, Z. Wang, and F. Tsung, "A spatial rank-based multivariate EWMA control chart", *Naval Research Logistics*, Vol 59 Issue 2, 2012, pp. 91-110.

[13] F.R. Hampel, E.M. Ronchetti, P.J. Rousseew, W.A. Stahel, "Robust Statistics - The Approach Based on Influence Functions", Wiley, 1986, New-York.

[14] L. Jaupi and G. Saporta, "Using the Influence Function in Robust Principal Components Analysis". In S. Morgenthaler, E. Ronchetti and W.A. Stahel, eds., *New Directions in Statistical Data Analysis and Robustness*, Birkhäuser Verlag, 1993, Basel, pp. 147-156.

[15] L. Jaupi and G. Saporta, "Control Charts for Multivariate Process Based on Influence Functions". *Proceedings of the Conference on Statistical Science.* Monte Verità, Zwitzerland, Birkhäuser Verlag Basel, 1997, pp. 193-201.

[16] L. Jaupi, "Multivariate Control Charts for Complex Processes". In C. Lauro, J. Antoch, V. Esposito, G. Saporta eds., *Multivariate Total Quality Control*, Springer, 2001, pp. 125-136.

ISBN: 978-988-19251-0-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)