Journal of Statistics: Advances in Theory and Applications
Volume 18, Number 1, 2017, Pages 1-44
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/jsata_7100121868

MULTIVARIATE STATISTICAL PROCESS OF HOTELLING'S T^2 CONTROL CHARTS PROCEDURES WITH INDUSTRIAL APPLICATION

M. S. HAMED

Department of Administrative Sciences
Faculty of Business Administration
Taibah University
Kingdom of Saudi Arabia

Department of Applied Statistics
Faculty of Commerce
Benha University
Egypt
e-mail: moswilem@gmail.com

Abstract

A control chart is a tool that monitors quality characteristics of a process to ensure that process control is being maintained when monitoring multiple characteristics that are correlated, it is imperative to use multivariate control chart. Hotelling's T^2 quality control chart is used to determine whether or not the process mean vector for two or more variables is in-control. It is allowing us to simultaneously monitor whether two or more related variables are in control, and it is shown that multivariate quality control chart do not indicate which variables cause the out-of control signal so that the interpretation of the out-of-control signal. Also, in this paper develops the multivariate quality control charts of the out of-control signal, this will be maintained by making and industrial application on the fertilizers factory. It is important to know that this

2010 Mathematics Subject Classification: 62P30.
Keywords and phrases: quality control, multivariate analysis, average run lengths performance, monitoring process.
Received August 21, 2017

factory has a special department for quality control, in order to maintain International Standardization Organization (ISO).

1. Introduction

The statistical control chart is a well-known tool in today's industry, and it is one of the most powerful tools in quality control. First developed in the 1920's by Walter Shewhart, the control chart found widespread use during World War II and has been employed, with various modifications ever since. The drawbacks to multivariate charting schemes is their inability to identify which variable was the source of the signal.

With today's use of computers, it is common to monitor several correlated quality characteristics simultaneously. Various types of multivariate control charts have been proposed to take advantage of the relationships among the variables being monitored. Alt [1]; Jackson [7]; Lowry and Montgomery [11]; and Mason et al. [12] discuss much of the literature on this topic. The formatter will need to create these components, incorporating the applicable criteria that follow.

The rapid growth data acquisition technology and the uses of online computers for process monitoring led to an increased interest in the simultaneous control of several related quality characteristics. These techniques are often referred to multivariate statistical process control procedures. The use of separate univariate control chart for each quality characteristic has proved to be inappropriate. This is because, it neglects the correlation between the multiple quality characteristics; and this leads to incorrect results.

The modern statistical process control took place when Walter Shewhart [20] developed the concept of a control chart based on the monitoring of the process mean level through sample mean $(\overline{X}\text{ chart})$ and process dispersion through sample range (R chart) or sample standard deviation chart. In the multivariate setting, Harold Hotelling [4] published what can be called the first major works in multivariate quality control. Hotelling developed the T^2 statistic and the statistics based on the sample variance-covariance matrix S procedure, and its extensions to control charts to combine measurements taken on variables in several dimensions into a single measure of excellence.

After Hotelling there was no significant work done in this field until the early sixties, when with the advances in computers, interest in multivariate statistical quality control was revived. Since then, some authors have done some work in this area of multivariate quality control.

Houshmand and Javaheri [5] presented two procedures to control the covariance matrix in a multivariate setting. The advantages of these procedures are that they allow the investigators to identify the sources of the out-of-control signal. These procedures are based on constructing tolerance regions to control the parameters of the correlation matrix.

Linna et al. [10] presented a model for correlated quality variables with measurement error. The model determined the performance of the multivariate control charting methods. The usual comparison of control chart performance does not directly apply in the presence of measurement error.

The most familiar multivariate process monitoring and control procedure is the Hotelling's T^2 control chart for monitoring the mean vector of the process. It is a direct analog of the univariate Shewhart $\overline{X}$ chart. Shewhart, a pioneer in the development of the statistical control chart (Shewhart charts), first recognized the need to consider quality control problems as multivariate in character. Hotelling [4] did the original work in multivariate quality control. He applied his procedure, which assumed that P-quality characteristics are jointly distributed as P-variate normally and random samples of size n are collected across time from the process. T^2 is sensitive to shifts in the means, as well as to shifts in the variance, but it cannot distinguish between location shifts and scale shifts.

Multivariate charts are also useful for monitoring quality profiles as discussed by Woodall et al. [24]. Alt [1] defined two phases in constructing multivariate control charts, with phase I divided into two stages. In the retrospective Stage 1 of phase I, historical data (observations) are studied for determining whether the process was in control and to estimate the in-control parameters of the process. The Hotelling's T^2 control chart is utilized in this stage (Alt and Smith [2]; Tracy et al. [21]; and Wierda [23]). In phase II, control charts are used with future observations for detecting possible departure from the process parameters estimated in phase I. In phase II, one uses charts for detecting any departure from the parameter estimates, which are considered in the in-control process parameters (Vargas [22]).

An important aspect of the Hotelling's T^2 control chart is how to determine the sample variance-covariance matrix used in the calculation of the chart statistics, the upper control limit (UCL) and the lower control limit (LCL).

Onwuka and Hotelling [15] discussed the principal component analysis and Hotelling's T^2 tests were used, with the 3-characteristics measured showing negligible low correlation with nearly all the correlation coefficients small.

2. Multivariate Quality Control Chart

Multivariate quality control charts are a type of variables control that how correlated, or dependent, variables jointly affect a process or outcome. The multivariate quality control charts are powerful and simple visual tools for determining whether the multivariate process is in-control or out-of-control. In other words, control charts can help us to determine whether the process average (center) and process variability (spread) are operating at constant levels. Control charts help us focus problem – solving efforts by distinguishing between common and assignable cause variation. Multivariate control chart plot statistical from more than one related measurement variable. The multivariate control chart shows how several variables jointly influence a process or outcome.

It is demonstrated that if the data include correlated variables the use of separate control chart is misleading because the variables jointly affect the process. If we use separate univariate control chart in a multivariate situation, type I error and probability of a point correctly plotting in-control are not equal to their expected values the distortion of those values increases with the number of measurement variables.

It is shown that multivariate control chart has several advantages in comparison with multiply univariate charts:

■ The actual control region of the related variables is represented.

■ We can maintain specification type I error.

■ A signal control limit determines whether the process is in control.

■ Multivariate control chart simultaneously monitors two or more correlated variables. To monitor more than one variable using univariate charts, we need to create a univariate chart for each variable.

■ The scale on multivariate control charts unrelated to the scale of any of the variables.

■ Out-of-control signals in multivariate charts do not reveal which variable or combination of variables cause the signal.

A multivariate control chart consists of:

■ Plotted points, each for which represents a rational subgroup of data sampled from the process, such as a subgroup mean vector individual observation, or weighted statistic.

■ A center line, which represents the expected value of the quality characteristics for all subgroups.

■ Upper and lower control limits (UCL and LCL), which are set a distance above and below the center line. These control limits provide a visual display for the expected amount for variation. The control limits are based on the actual behaviour of the process, not the desired behaviour or specification limits. A process can be in control and yet not be capable of meeting requirements.

3. Construction of Hotelling's T^2 Control Chart

The Hotelling multivariate control chart signals that a statistically significant shift in the mean has occurred as soon as:

$$\chi^2 = (\overline{X}_i - \mu_o)' \sum\nolimits^{-1} (\overline{X}_i - \mu_o). \tag{1}$$

If the sample covariance matrix $\sum$ and the sample mean vector μ_0 are known, but if $\sum$ and μ_0 are known, then the T^2 statistic is the appropriate statistic for the Hotelling multivariate control chart. In this case, the sample covariance matrix, S and sample mean vector $\overline{X}$, are used to estimate $\sum$ and μ_0, respectively.

This statistic has the from:

$$T^2 = (\overline{X}_i - \overline{\overline{X}})' S^{-1} (\overline{X}_i - \overline{\overline{X}}). \tag{2}$$

Suppose that we have a random sample from a multivariate normal distribution – Say, $X_1, X_2, X_3, \dots, X_n$, where the i-th sample vector contains observations, $X_{i1}, X_{i2}, X_{i3}, \dots, X_{ip}$.

Let the sample mean vector is:

$$\underline{\overline{X}}_{1\times p} = (\overline{X}_1 \ \overline{X}_2 \dots \overline{X}_p),$$

where

$$\overline{X}_i = \sum_{L=1}^{n} X_{iL} \qquad (i = 1, 2, \dots, p),$$

and the sample covariance matrix is:

$$S = \begin{bmatrix} S_1^2 & S_{12} & \cdots & S_{1p} \\ S_{21} & S_2^2 & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_p^2 \end{bmatrix}, \tag{3}$$

where s_i^2 is the variance of the i-th variable and (ij)-th element of S-matrix is the estimated covariance between the variables i and j,

$$S_{ij} = \frac{1}{n-1}\sum_{L=1}^{n}(X_{iL} - \overline{X}_i)(X_{jL} - \overline{X}_j).$$

Note that we can show that the sample mean vector and the sample covariance matrix are unbiased estimators of the corresponding population quantities that is

$$E(\underline{\overline{X}}) = \underline{\mu} \quad \text{and} \quad E(S) = \sum. \tag{4}$$

Seber [18] gives the distribution properties of this estimate as follows:

(i) $\overline{X} \sim N_p\left(\mu, \frac{1}{n}\sum\right).$ (5)

(ii) If $\overline{X}$ distribution as in (i) then: $n(\overline{X} - \mu)'\sum^{-1}(\overline{X} - \mu) \sim \chi_p^2.$ (6)

(iii) $(n-1)S \sim W_p(n-1, \sum),$

where $W_p(n-1, \sum)$ stands for the Wishart distribution.

(iv) If Z and D are independent, random variables distributed, respectively as:

$$Z \sim N_p(0, \textstyle\sum_Z)$$

$$(n-1)D \sim W_p(n-1, \textstyle\sum_Z),$$

then the quadratic form:

$$T^2 = Z'D^{-1}Z, \tag{7}$$

is distributed as:

$$T^2 \sim \frac{(n-1)p}{(n-p)} F_{p,\, n-p}. \tag{8}$$

(v) If $Z \sim N_p(0, \sum_Z)$ and $(n-1)D \sim W_p(n-1, \sum_Z)$, $(n-1) > p$, where $(n-1)D$ can be decomposed as:

$$(n-1)D = (n-2)D_1 + ZZ',$$

where

$$(n-2)D_1 \sim W_p(n-2, \textstyle\sum_Z),$$

and Z is independent of D_1 then the quadratic form:

$$T^2 = Z'D^{-1}Z, \tag{9}$$

is distributed as:

$$T^2 \sim (n-1)\beta(p,\, n-p-1),$$

where

$\beta(p,\, n-p-1)$ is the central Beta distribution.

(vi) If the sample is composed of k subgroups of size n with subgroup means $\overline{X}_j$, $j = 1, 2, 3, \ldots, k$ and grand mean $\overline{\overline{X}},$ i.e.,

$$\overline{X} = \sum_{j=1}^{k} \overline{X}_j / k = \sum_{j=1}^{k} \sum_{i=1}^{n} X_{ij} / kn,$$

then

$$\sqrt{\frac{kn}{k-1}}\left(\overline{X}_j - \overline{\overline{X}}\right) \sim N_p(0, \textstyle\sum). \tag{10}$$

(vii) If the sample is composed of K subgroups of n identically distributed multivariate normal observations and if S_j is the sample covariance matrix from the j-th subgroup, $j = 1, 2, 3, \ldots, K$), then

$$\sum (n-1)S_j \sim W_p(K_{n-1}, \textstyle\sum), \tag{11}$$

these distributional properties of $\overline{X}, S$ and T^2 are used in the multivariate quality control procedures.

Now, we present two versions of Hotelling T^2 chart:

(a) Subgroup data

Suppose that P-related quality characteristic $X_1, X_2, X_3, \ldots, X_p$ are controlled jointly according to the P-multivariate normal distribution. The procedure requires computing the sample mean for each of the P-quality characteristics from a sample of size n.

Let the set of quality characteristic means is represented by the $(p \times 1)$ vector $\overline{X}$ as:

$$\overline{X} = \begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \\ \overline{X}_p \end{bmatrix}.$$

Then the test statistic plotted on the Chi-square control chart for each sample is:

$$\chi_0^2 = n\left(\underline{\overline{X}} - \underline{\mu}\right)' \textstyle\sum^{-1} \left(\underline{\overline{X}} - \underline{\mu}\right), \tag{12}$$

where

$\underline{\mu}' = (\mu_1, \mu_2, \ldots, \mu_p)$ is the $(p \times 1)$ vector of in-control means for each quality characteristic and $\sum$ is covariance matrix.

Now, suppose that m-subgroup are available. The sample means and variances are calculated from each subgroup as usual that is:

$$\overline{X}_{jk} = \frac{1}{n}\sum_{i=1}^{n} X_{ijk},$$

$$S_{jk}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_{ijk} - \overline{X}_{jk}\right)^2,$$

$$j = 1, 2, \ldots, p; \quad k = 1, 2, \ldots, m, \tag{13}$$

where X_{ijk} is the i-th observation on the j-th quality characteristic in the k-th subgroup.

$$S_{jhk} = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_{ijk} - \overline{X}_{jk}\right)\left(X_{ihk} - \overline{X}_{hk}\right).$$

$$k = 1, 2, \ldots, n, \quad j \neq h, \tag{14}$$

represents the covariance between quality characteristic j and quality characteristic h in the k-th subgroup.

The statistics $\overline{X}_{jk}, S_{jk}^2, S_{jhk}$ are the averaged over all m-subgroups to obtain

$$\overline{\overline{X}}_j = \frac{1}{m}\sum_{k=1}^{m}\overline{X}_{jk}, \quad j = 1, 2, \ldots, p,$$

$$\overline{S}_j^2 = \frac{1}{m}\sum_{k=1}^{m} S_{jk}^2, \quad j = 1, 2, \ldots, p,$$

and

$$\overline{S}_{jh} = \frac{1}{m}\sum_{k=1}^{m} S_{jhk}, \tag{15}$$

where $j \neq h$ and $\overline{\overline{X}}_j$ are the i-th elements of the $(p \times 1)$ sample mean vector $\underline{\overline{\overline{X}}}$ and $(p \times p)$ average of sample covariance matrices S is formed as:

$$S = \begin{bmatrix} \overline{S}_1^2 & \overline{S}_{12} & \cdots & \overline{S}_{1p} \\ & \overline{S}_2^2 & \cdots & \overline{S}_{2p} \\ & & \cdots & \\ & & & \overline{S}_p^2 \end{bmatrix}. \tag{16}$$

There are consider the unbiased estimate of μ and $\sum$ when the process is in control. If, we replace μ with $\overline{\overline{X}}$ and $\sum$ with S in (12), the test statistic now becomes

$$T^2 = n\left(\overline{X} - \overline{\overline{X}}\right)' S^{-1}(\overline{X} - \overline{\overline{X}}). \tag{17}$$

Alt [1] has pointed out that there are two distinct phases of control chart using. Phase I is the use of the chart for establishing control, that is, testing whether the process was in control when m-subgroups were drawn and the sample statistic $\overline{\overline{X}}$ and S computed. The objective in phase I is to obtain an in-control set of observations, so that control limits can be established for phase II which is the monitoring of future production.

In the phase I the control limits for the T^2-control chart is given by:

$$\left.\begin{aligned} \text{UCL} &= \frac{P(m-1)(n-1)}{mn - m - p + 1} F_{\alpha,\, P,\, mn-m-p+1} \\ \text{LCL} &= 0 \end{aligned}\right\}. \tag{18}$$

In the phase II when the chart is used for monitoring future production, the control limits are as follows:

$$\left.\begin{aligned} \text{UCL} &= \frac{P(m+1)(n-1)}{mn-m-p+1} F_{\alpha,\, P,\, mn-m-p+1} \\ \text{LCL} &= 0 \end{aligned}\right\}, \tag{19}$$

when the parameters μ and $\sum$ are estimated from a large number of subgroups, it is often to use $\text{UCL} = \chi^2_{\alpha,\, p}$ as the upper limit in both phases. Retrospective analysis of samples to test for statistical control and establish control limits also occurs in the univariate control chart setting. For the $\overline{X}$-chart, it is well-known that if use $m \geq 20$ or 25, samples, the distribution between phase I and phase II limits is usually unnecessary, because the phase I and phase II limits will nearly coincide. However, with multivariate control charts, we must be careful.

Lowry and Montgomery [11] showed that in many situations a large number of samples would be required before the exact phase II control limits are well approximate by the Chi-square.

(b) Individual observations

In some situation the subgroup size is naturally $n = 1$. Suppose that m samples each of size $n = 1$ are available and that p is the number of quality characteristics observed in each sample. The Hotelling T^2 statistic becomes:

$$T^2 = (X - \overline{X})' S^{-1} (X - \overline{X}). \tag{20}$$

Ryan [17] defined the phase II control limits for this statistic as:

$$\left.\begin{aligned} \text{UCL} &= \frac{P(m+1)(m-1)}{m(m-p)} F_{\alpha,\, P,\, m-p} \\ \text{LCL} &= 0 \end{aligned}\right\}. \tag{21}$$

Jackson [7] suggested that for large $m (m > 100)$ then we can use an approximate control limit, either

$$\text{UCL} = \frac{P(m-1)}{(m-p)} F_{\alpha,\, P,\, m-p}, \tag{22}$$

or

$$\text{UCL} = \chi^2_{\alpha,\, p}. \tag{23}$$

Equation (23) is only appropriate if the covariance matrix is known.

Lowry and Montgomery [11] suggested that if p is large-say $p \geq 10$ then at least 250 samples must be taken $(m \geq 250)$ before Chi-square upper control limit is a reasonable approximation to the correct value.

Tracy et al. [21] point out that if $n = 1$, the phase I limits should be based on a beta distribution that is, the phase I limits defined as:

$$\left.\begin{array}{l} \text{UCL} = \dfrac{(m-1)^2}{m} \beta_{\alpha,\, \frac{P}{2},\, \frac{m-p-1}{2}} \\ \text{UCL} = 0 \end{array}\right\}. \tag{24}$$

The average run length performance for a Hotelling's T^2

Mason et al. [12] suggested that the average run length (ARL) for a control procedure is defined as: $\text{ARL} = \frac{1}{p}$, where p represents the probability of being outside the control region. For a process that is in-control, this probability is equal to α, the probability of type I error. The ARL has a number of uses in both univariate and multivariate control procedures. They suggested that it can be used to calculate the number of observations that one would expect to observe, on average, before a false alarm occurs. This given by: $\text{ARL} = \frac{1}{p}$.

Another use of the ARL is to compute the number of observations one would expect to observe before detecting a given shift in the process. The probability of a type II error. The ARL for detecting the shift is given by:

$$\text{ARL} = \frac{1}{1-\beta}.$$

Multivariate control charts using Hotelling's T^2 statistic is popular and easy to use. A major advantage of Hotelling's T^2 statistic is that it can be shown to be the optimal test statistic for detecting a general shift in the process mean vector for an individual multivariate observation. However, the technique has several practical drawbacks. A major drawback is that when the T^2 statistic indicates that a process is out of control, it does not provide information in which variable or set of variables is out of control. Further, it is difficult to distinguish location shifts from scale shifts since the T^2 statistic is sensitive to both types of process changes.

The MYT decomposition

Mason, Young and Tracy (MYT) extended that the interpretation of signals from a T^2 chart to the setting where there is more than process variables MYT decomposition. The MYT decomposition is the primary tool used in this effort, and they examined many interesting properties associated with it. They showed that the decomposition terms contained information on the residuals generated by all possible linear regressions of one variable on any subset of the other variables. And they add that to being an excellent aid in locating the source of a signal in terms of individual variables or subsets of variables, this property has another major function. It can be used to increase the sensitivity of the T^2 statistic in the area of small process shifts.

Mason et al. [14] presented decomposition procedure. They considered that, the T^2 statistic for a p-dimensional observation vector $\overline{X} = (X_1, X_2, \ldots, X_p)$ can be represented as

$$T^2 = \left(X - \overline{X}\right)^{\backslash} S^{-1}\left(X - \overline{X}\right), \tag{25}$$

where $\overline{X}$ and S are the common estimators of the mean vector and covariance matrix obtained from historical data set (HDS), they partitioned the vector $(X - \overline{X})$ as:

$$\left(X - \overline{X}\right)' = [(X^{(p-1)} - \overline{X}^{(p-1)}),\ \left(X_p - \overline{X}_p\right)]',$$

where $X^{(p-1)'} = (X_1,\ X_2,\ \ldots,\ X_{p-1})$ represented the $(p-1)$-dimensional variable vector excluding the p-th variable X_p and $\overline{X}^{(p-1)}$ represented the corresponding $p-1$ elements of the mean vector. They also partition the matrix S so that

$$S = \begin{bmatrix} S_{XX} & s_{xX} \\ s'_{xX} & s_{xx} \end{bmatrix}, \tag{26}$$

where S_{XX} is the $(p-1) \times (p-1)$ covariance matrix for the first $(p-1)$ variables, S_p^2 is the variance of X_p, and s_{xX} is a $(p-1)$-dimensional vector the containing the covariances between X_p and the remaining $(p-1)$ variables.

The T^2 statistic in (25) can be partitioned into two independent parts (see Rencher [16]). These components are given by

$$T^2 = T_{p-1}^2 + T_{p,\ 1,\ 2,\ \ldots,\ p-1}^2. \tag{27}$$

The first term in (27),

$$T_{p-1}^2 = (X^{(p-1)} - \overline{X}^{(p-1)})' S_{XX}^{-1} (X^{(p-1)} - \overline{X}^{(p-1)}), \tag{28}$$

uses the first $(p-1)$ variables and is itself a T^2 statistic.

Mason et al. [14] proved that the last term in (27) was the p-th component of the vector X_i adjusted by the estimates of the mean and standard deviation of the conditional distribution X_p given $(X_1, X_2, \ldots, X_{p-1})$. It is given by

$$T^2_{p, 1, 2, \ldots, p-1} = \frac{(X_p - \overline{X}_{p, 1, 2, \ldots., p-1})^2}{S^2_{p, 1, 2, \ldots., p-1}}, \tag{29}$$

where

$$\overline{X}_{p, 1, 2, \ldots., p-1} = \overline{X}_p + B_p^{\grave{}}(X^{(p-1)} - \overline{X}^{(p-1)}),$$

and $B_p^{\grave{}} = S_{XX}^{-1} s_{xX}$, is the $(p-1)$-dimensional vector estimate of the coefficients from the regression of X_p on the $(p-1)$ variables $X_1, X_2, \ldots, X_{p-1}$. It can be shown that the estimate of the conditional variance is given as $S^2_{p, 1, 2, \ldots, p-1} = S_p^2 - s_{xX}^{\grave{}} S_{XX}^{-1} s_{xX}$, since the first term of (27) is a T^2 statistic, it too can be separated into two orthogonal parts:

$$T^2_{p-1} = T^2_{p-2} + T^2_{p-1, 1, 2, \ldots, p-2}.$$

The first term, T^2_{p-2}, is a T^2 statistic, on the first $(p-2)$ components of the X vector, and the second term $T^2_{p-1, 1, 2, \ldots, p-2}$, is the square of X_{p-1} adjusted by the estimates of the standard deviation of the conditional distribution of X_{p-1} given $(X_1, X_2, \ldots, X_{p-2})$. They proposed one from of MYT decompositions of a T_2 statistic. It is given by.

$$\begin{aligned} T^2 &= T_1^2 + T^2_{2, 1} + T^2_{3, 1, 2} + \ldots + T^2_{p, 1, 2, \ldots, p-1} \\ &= T_1^2 + \sum_{j=1}^{p-1} T^2_{j+1, 1, \ldots, j}. \end{aligned} \tag{30}$$

The T_1^2 term in (30) is the square of the univariate for the first variable of the vector X and is given as

$$T_1^2 = \frac{(X_1 - \overline{X}_1)^2}{S_1^2}. \tag{31}$$

This term is not a conditional term, as its value does not depend on a conditional distribution. In contrast, all other terms of the expansion in (30) are conditional terms, since they represent the value of a variable by the mean and standard deviation from the appropriate conditional distribution.

Computing the decomposition terms

They considered that the first $(p-1)$ terms of (30) correspond to the T^2 value of the sub vector $X_{p-1}^{\backprime} = (X_1, X_2, \ldots, X_{p-1})$; i.e.,

$$T^2_{(X_1, X_2, \ldots, X_{p-1})} = T_1^2 + T_{2,1}^2 + T_{3,1,2}^2 + \ldots + T^2_{p-1,1,2,\ldots,p-2},$$

similarly, the first $(p-2)$ terms of this expansion correspond to the sub vector $X_{p-2}^{\backprime} = (X_1, X_2, \ldots, X_{p-2})$; i.e.,

$$T^2_{(X_1, X_2, \ldots, X_{p-2})} = T_1^2 + T_{2,1}^2 + T_{3,1,2}^2 + \ldots + T^2_{p-2,1,2,\ldots,p-3},$$

continuing in this fashion, they compute the T^2 values for all sub vectors of the original vector X. The last sub vector, consisting of the first component $X_{(1)} = (X_1)$ is used to compute the unconditional T^2 term given in (31); i.e., $T_{X_1}^2 = T_1^2$.

All the T^2 values, $T^2_{(X_1, X_2, \ldots, X_p)}, T^2_{(X_1, X_2, \ldots, X_{p-1})}, T^2_{(X_1)}$, are computed using the general formula

$$T^2_{(X_1, X_2, \ldots, X_j)} = \left(X^{(j)} - \overline{X}^{(j)}\right)^{\backprime} S_{jj}^{-1}(X^{(j)} - \overline{X}^{(j)}), \tag{32}$$

where $X^{(j)}$ represents the appropriate sub vector, $\overline{X}^{(j)}$ is the corresponding sub vector mean, and S_{jj} denotes the corresponding covariance sub matrix obtained from the overall S matrix given in (26) by deleting all unused rows and columns. The terms of the MYT decomposition can be computed as follows:

$$T^2_{p,\,1,\,2,\,\ldots,\,p-1} = T^2_{(X_1,\,X_2,\,\ldots,\,X_p)} - T^2_{(X_1,\,X_2,\,\ldots,\,X_{p-1})},$$

$$T^2_{p-1,\,1,\,2,\,\ldots,\,p-2} = T^2_{(X_1,\,X_2,\,\ldots,\,X_{p-1})} - T^2_{(X_1,\,X_2,\,\ldots,\,X_{p-2})},$$

$$\ldots\ldots$$

$$\ldots\ldots$$

$$\ldots\ldots$$

$$T^2_{2,\,1} = T^2_{(X_1,\,X_2)} - T^2_1,$$

$$T^2_1 = \frac{(X_1 - \overline{X}_1)^2}{s_1^2}. \tag{33}$$

Properties of the MYT decomposition

Many properties are associated with the MYT decomposition. Consider they dimensional vector defined as $X^{\setminus} = (X_1,\, X_2,\, \ldots,\, X_p)$ they interchange the first two components to form another vector $(X_2,\, X_1,\, \ldots,\, X_p)$ so that the only difference between the two vectors is the two vectors such that the first two components have been permuted the T^2 value of the two vectors is the same; i.e.,

$$T^2_{(X_1,\,X_2,\,\ldots,\,X_p)} = T^2_{(X_2,\,X_1,\,\ldots,\,X_p)}.$$

This occurs because T^2 values cannot be changed by permuting the components of the observation vector. This invariance property of permuting the T^2 components that each ordering of an observation vector will produce the same overall T^2 value. Since there are $p! = (p)(p-1)(p-2)\ldots(2)(1)$ permutations of the components of the vector $(X_1, X_2, \ldots, X_p)$ this implies that we can partition a T^2 value in $p!$ different ways. To illustrate his result, suppose $p = 3$. There are $3! = (3)(2)(1) = 6$ decompositions of the T^2 value for an individual observation vector. These are listed below:

$$\begin{aligned} T^2 &= T_1^2 + T_{2,1}^2 + T_{3,1,2}^2 \\ &= T_1^2 + T_{3,1}^2 + T_{2,1,3}^2 \\ &= T_2^2 + T_{3,2}^2 + T_{1,2,3}^2 \\ &= T_2^2 + T_{1,2}^2 + T_{3,1,2}^2 \\ &= T_3^2 + T_{1,3}^2 + T_{2,1,3}^2 \\ &= T_3^2 + T_{2,3}^2 + T_{1,2,3}^2. \end{aligned} \tag{34}$$

Each row of (34) corresponds to a different permutation of the components of the observation vector. For example, the first row corresponds to the vector written in its original form as (X_1, X_2, X_3), whereas the last row represents (X_3, X_2, X_1). Note that all six possible permutations of the original vector components are included.

The importance of this result is that it allows one to examine the T^2 statistic from many different perspectives. The p terms in any particular decomposition are independent of one another, although the terms across the decompositions are not necessarily independent. With p partition and $p!$ partitions, there are $p \times p!$ possible terms to evaluate in a total MYT decomposition of a particular partition, as certain terms occur more than once. In general, there are $p \times 2^{(p-1)}$ distinct terms among the possible decompositions. These unique terms are the ones that need to be examined for possible contribution to a T^2 signal, when p is large. Computing all these terms can be cumbersome.

They considered the MYT decomposition given in (30) and suppose T_1^2 dominates the overall value of the T^2 statistic. This indicates that the observation on the variable X_1 is contributing to the signal. However, to determine if the remaining variables in this observation contribute to the signal, we must examine the T^2 value associated with the sub vector $(X_2, X_3, \ldots, X_p)$ which excludes the X_1 component. Small values of the T^2 statistic for this sub vector imply that no signal is present. They also indicate that one need not examine any term of the total decomposition involving these $(p-2)$ variables.

They considered another important property of T^2 statistic is the fact that the $p(2^{(p-1)}-1)$ unique conditional terms of a MYT decomposition contain the residuals from all possible linear regressions of each variable on all subsets of the other variables. This property of the T^2 statistic provides a procedure for increasing the sensitivity of the T^2 sensitivity of the T^2 statistic to process shifts.

Locating signaling variables

They seek to relate a T^2 signal and its interpretation to the components of the MYT decomposition. They consider signaling observation vector:

$$X^{\setminus} = (X_1, X_2, \ldots, X_p) \text{ such that } T^2_{(X_1, X_2, \ldots, X_p)} > \text{UCL}.$$

They proposed two methods one for locating the variables contributing to the signal is to develop a forward iterative scheme. This was accomplished by finding the subset of variables that do not contribute to signal from (27) and (29) such that a T^2 statistic can be constructed on any subset of the variables $X_1, X_2, \ldots, X_p$. Construct the T^2 statistic for each individual variable X_j, $j = 1, 2, \ldots, p$, so that

$$T_j^2 = \frac{(X_j - \overline{X}_j)^2}{S_j^2},$$

where $\overline{X}_j$ and S_j^2 are the corresponding mean and variance estimates as determined from the HDS. Compare these individual T_j^2 values to their UCL, where

$$\begin{aligned} \text{UCL}_{(X_j)} &= \left\{\frac{p(n+1)(n-1)}{n(n-p)}\right\} F_{(\alpha,\ p,\ n-p)} \\ &= \left\{\frac{n+1}{n}\right\} F_{(\alpha, 1,\ n-1)}, \end{aligned} \tag{35}$$

is computed for an appropriate α level and for a value of $p = 1$. Exclude from the original set of variables all X_j for which

$$T_j^2 > \text{UCL}_{(X_j)},$$

since observations on this subset of variables are definitely contributing to the signal.

From the set of variables not contributing to the signal. Compute the T^2 statistic for all possible pairs of variables. For example, for all (X_i, X_j) with $i \neq j$ compute $T^2_{(X_i, X_j)}$, and compare these values to the upper control limit,

$$\text{UCL}_{(X_i, X_j)} = \left\{\frac{2(n+1)(n-1)}{n(n-2)}\right\} F_{(\alpha,\ 2,\ n-2)}.$$

Exclude from this group all pairs of variable for which

$$T^2_{(X_i, X_j)} > \text{UCL}_{(X_i, X_j)}.$$

The excluded pairs of variables in addition to exceeded single variable comprise the group of variables contributing to the overall signal continue to iterate in this fashion so as to exclude from the remaining group all variables of signaling groups of three variables four variables etc. The procedure produces a set of variables that contribute to the signal. And another method of locating the vector components contributing to a sing al is to examine the individual terms of the MYT decomposition of a signaling observation vector and to determine which are large in value. This method was accomplished by comparing each term to its corresponding critical value.

Mason et al. [14] proposed that the distribution governing the components of the MYT decomposition for the situation where there are no signals is F distribution. For the case of p variables, these are given by

$$T_j^2 \sim \left\{\frac{n+1}{n}\right\} F_{(1,\, n-1)}, \tag{36}$$

for unconditional terms, and by

$$T_{j,\, 1,\, 2,\, \ldots,\, j-1}^2 \sim \left\{\frac{(n+1)(n-1)}{n(n-k-1)}\right\} F_{(1,\, n-k-1)}, \tag{37}$$

for conditional terms. Where k equals the number of conditioned variables. For $k = 0,$ the distribution in (37) reduces to the distribution in (36). Using these distributions, critical values for a specified α level and HDS sample of size n for both conditional and unconditional terms are obtained as follows:

$$\text{unconditional terms :} \qquad CV = \left\{\frac{n+1}{n}\right\} F_{(\alpha,\, 1,\, n-1)},$$

$$\text{conditional terms :} \qquad CV = \left\{\frac{(n+1)(n-1)}{n(n-k-1)}\right\} F_{(\alpha,\, 1,\, n-k)}. \tag{38}$$

They add that we can compare each individual term of the decomposition to its critical value and make the appropriate decision.

Interpretation of a signal on a T^2 component

They (Mason et al. [12]) considered one of the p possible unconditional, terms resulting from the decomposition of the T^2 statistic associated with a signaling observation. As stated earlier, the term

$T_j^2 = \dfrac{(X_j - \overline{X}_j)^2}{s_j^2}$, $j = 1, 2, \ldots, p$ is square of a univariate t statistic for the observed value of the j-th variable of an observation vector X. For control to be maintained, this component must be less than its critical value, i.e., $T_j^2 < \left\{\dfrac{n+1}{n}\right\} F_{(\alpha,\, 1,\, n-1)}$, since $t_{(\frac{\alpha}{2},\, n-1)} = \sqrt{F_{(\alpha,\, 1,\, n-1)}}$ they re-expressed this condition as T_j being in the following interval:

$$-\sqrt{\left(\frac{n+1}{n}\right)} t_{(\frac{\alpha}{2},\, n-1)} < T_j < \sqrt{\left(\frac{n+1}{n}\right)} t_{(\frac{\alpha}{2},\, n-1)}, \tag{39}$$

or as

$$\overline{X} - \sqrt{\left(\frac{n+1}{n}\right)} t_{(\frac{\alpha}{2},\, n-1)} < X_j < \overline{X} + \sqrt{\left(\frac{n+1}{n}\right)} t_{(\frac{\alpha}{2},\, n-1)}, \tag{40}$$

where $t_{(\frac{\alpha}{2},\, n-1)}$ is, the appropriate value from a t-distribution with $n-1$ degrees of freedom. This is equivalent to using a univariate Shewhart control chart for the j-th variable. And they considered the form of a general conditional term given as

$$T_{j,\,1,\,2,\,\ldots,\,j-1}^2 = \frac{(X_j - \overline{X}_{j,\,1,\,2,\,\ldots,\,j-1})^2}{s_{j,\,1,\,2,\,\ldots,\,j-1}^2}, \tag{41}$$

if the value in (41) is to be less than its control limit,

$$T_{j,\,1,\,2,\,\ldots,\,j-1}^2 < \left\{\frac{(n+1)(n-1)}{n(n-k-1)}\right\} F_{(\alpha,\, 1,\, n-k-1)}.$$

Its numerator must be small, as the denominator of these terms is fixed by the historical data this implies that component X_j from the observation vector $X^{\backslash} = (X_1, X_2, \ldots, X_j, \ldots, X_p)$ is contained in the conditional distribution of X_j given $X_1, X_2, \ldots, X_{j-1}$ and falls in the elliptical control region.

A signal occurs on the term in (41) when X_j is not contained in conditional distribution of X_j given $X_1, X_2, \ldots, X_j$, i.e., when

$$T^2_{j, 1, 2, \ldots, j-1} > \left\{\frac{(n+1)(n-1)}{n(n-k-1)}\right\} F_{(\alpha, 1, n-k-1)}.$$

Regression perspective

Mason et al. [13, 14] proposed that, in general, $T^2_{j, 1, 2, \ldots, j-1}$ is a standardized observation on the j-th variable adjusted by the estimated of the mean and variance form the conditional distribution associated with $(X_j \mid X_1, X_2, \ldots, X_{j-1})$. The general from of this term was given in (41). They considered the estimated mean of X_j adjusted for $X_1, X_2, \ldots, X_{j-1}$, and estimated this mean by using the prediction equation.

$$\overline{X}_{j, 1, 2, \ldots, j-1} = \overline{X}_j + B^{\setminus}{}_j \left(X^{(j-1)} - \overline{X}^{(j-1)}\right), \tag{42}$$

where $\overline{X}_j$ is the sample mean of X_j obtained from the historical data. The sub vector $X^{(j-1)}$ is composed of the observation on $(X_1, X_2, \ldots, X_{j-1})$ and $\overline{X}^{(j-1)}$ is the corresponding estimated mean vector, S_{jj}, is the covariance matrix of the first j components of the vector X. To obtain S_{jj} partition S as follows:

$$S = \begin{bmatrix} S_{jj} & S_{j(p-j)} \\ S^{\setminus}_{j(p-j)} & S_{(p-j)(p-j)} \end{bmatrix}.$$

Further, partition the matrix S_{jj} as

$$S_{jj} = \begin{bmatrix} S_{(j-1)(j-1)} & S_{j(j-1)} \\ S^{\setminus}_{j(j-1)} & S^2_j \end{bmatrix}.$$

Then

$$B_j = S^{-1}_{(j-1)(j-1)} S_{j(j-1)},$$

since the left-hand side of (42) contains $\overline{X}_{j, 1, 2, \ldots, j-1}$, which is the predicted value of X_j, the numerator of (41) is a regression residual represented by

$$r_{j, 1, 2, \ldots, j-1} = (X_j - \overline{X}_{j, 1, 2, \ldots, j-1}),$$

rewriting the conditional variance as

$$S^2_{j, 1, 2, \ldots, j-1} = S^2_j (1 - R^2_{j, 1, 2, \ldots, j-1}).$$

(see, e.g., Rencher [16] and substituting $r_{j, 1, 2, \ldots, j-1}$ for $(X_j - \overline{X}_{j, 1, 2, \ldots, j-1})$, they expressed

$$T^2_{j, 1, 2, \ldots, j-1} = \frac{(r_{j, 1, 2, \ldots, j-1})^2}{s^2_j (1 - R^2_{j, 1, 2, \ldots, j-1})}. \tag{43}$$

The above results indicate a T^2 signal may occur if something goes astray with the relationships between subsets of various variables. This situation can be determined by examination of the conditional T^2 terms. A signaling value indicates that a contradiction with historical relationship between the variables has occurred either (1) due to a standardized component value that is significantly larger or smaller than that predicted by a subset of the remaining variables or (2) due to a standardized component value that is marginally smaller or larger than that predicted by a subset of the remaining variables when there is a very severe collinearity (i.e., a large R^2 value) among the variables. Thus, a signal results when an observation on a particular variable or set of variables, is out of control and/or when observations on a set of variables are counter to the relationship established by the historical data.

Improving the sensitivity of the T^2 statistic

Mason et al. [12, 13] used the decomposition for improving the sensitivity of the T^2 in signal detection. They showed that the T^2 statistic to be a function of all possible regressions existing among a set of process variables. Furthermore, they showed that the residuals of the estimated regression models are contained in the conditional terms of the MYT decomposition. Large residuals produce large T^2 components for the conditional terms and are interpreted as indicators of counter relationships among the variables. However, a large residual also could imply an incorrectly specified model. This result suggests that it may be possible to improve the performance of the T^2 statistic by more carefully describing the functional relationships existing among the process variables. Minimizing the effects of model misspecification on the signaling ability of the T^2 should improve its performance in detecting abrupt process shifts. They showed when compared to other multivariate control procedures, the T^2 lacks the sensitivity of detecting small process shifts. They showed that this problem can be overcome by monitoring the error residuals of the regressions contained in the conditional terms of the MYT decomposition of T^2 statistic. Furthermore, they showed that such monitoring can be helpful in certain types of on-line experimentation within a processing unit.

They proposed an alternative form of condition terms, they considered the conditional term of the MYT decomposition in (51). This is the squares of the j-th variable of the observation vector which adjusted by the estimates of the mean and variance of the conditional distribution of X_j given $X_1, X_2, \ldots, X_{j-1}$. They showed that (51) could be written as

$$T^2_{j, 1, 2, \ldots, j-1} = \left(\frac{r_{j, 1, 2, \ldots, j-1}}{s_{j, 1, 2, \ldots, j-1}} \right)^2, \tag{44}$$

this was achieved by noting that $\overline{X}_{j, 1, 2, \ldots, j-1}$ can be obtained from the regression of X_j on $X_1, X_2, \ldots, X_{j-1}$: i.e.,

$$\overline{X}_{j, 1, 2, \ldots, j-1} = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_{j-1} X_{j-1}, \tag{45}$$

where b_j are the estimated regression coefficients. Since $\overline{X}_{j, 1, 2, \ldots, j-1}$ is the predicted value of X_j. The numerator of (41) is the raw regression residual,

$$r_{j, 1, 2, \ldots, j-1} = (X_j - \overline{X}_{j, 1, 2, \ldots, j-1}) \tag{46}$$

given in (44).

Another form of the conditional term in (41) is obtained by substituting the following quantity for the conditional variance contained in (44), i.e., by substituting

$$S^2_{j, 1, 2, \ldots, j-1} = S^2_j (1 - R^2_{j, 1, 2, \ldots, j-1}),$$

where $R^2_{j, 1, 2, \ldots, j-1}$ is the squared multiple correlation between X_j and $X_1, X_2, \ldots, X_{j-1}$ this yields:

$$T^2_{j, 1, 2, \ldots, j-1} = \frac{r^2_{j, 1, 2, \ldots, j-1}}{s^2_j (1 - R^2_{j, 1, 2, \ldots, j-1})}. \tag{47}$$

Much information is contained in the conditional terms of the MYT decomposition. Since these terms are, in fact, squared residuals from regression equations they can be helpful in improving the sensitivity of the T^2 statistic in detecting both abrupt process changes and gradual shifts in the process.

They considered the better the fit of a model, the more sensitive the T^2 control procedure will be to departures from the model. This suggests that more effort should be taken in phase I operations, during construction of the historical database, to insure proper functional forms are chosen for the process variables and useful models are created.

Principal components procedure

Jakson [6, 8] recommended to use the principal components procedure to aid in the interpretation of an out-of-control signal. The principal component analysis (PCA) technique decomposes the T^2 statistic into a sum of p independent principal components, which are linear combinations of the original variables, and using these components to help for solving this identification problem. The PCA used to reduce to dimensionality of a data set which consists of a large number of interrelated variables. While retaining as much as possible of the variation present in the data set. This goal is achieved by transforming the original variables to a new set of uncorrelated variables, which are called the principal components. This transformation is a principal axis rotation of the variance and covariance matrix of the data set, and the elements of the characteristic vectors or the eigenvectors of the covariance matrix are direction cosines of the new axes related to the old.

The transformed new uncorrected variable or the principal components are normally numbered in descending order according to the amount at the variation. The use of the method of PCA in the field of multivariate quality control was first introduced by Jackson and Morris [9]. They identified a large number (p) of correlated variables that account for the quality of the process.

They notice that the use of Hotelling's T^2 may involve computational problems since the determinant of the variance and covariance matrix is near zero. The solution is to transform the original p variables to lesser k principal components.

Jackson [6] proposed using PCA to interpretation an out-of-control single by decomposing T^2 into independent component. Jackson proposed that the starting point of the statistical application of the method of principal components is the sample covariance matrix S, for a p-variate problem in (3).

If the covariances are not equal to zero, it indicates a relationship existing between those two variables, the strength of that relationship (if it is linear) being represented by the correlation $r_{ij} = \frac{S_{ij}}{(S_i S_j)}$.

A principal axis transformation will transform p correlated variables $X_1, X_2, \ldots, X_p$ into p new uncorrelated variables $Z_1, Z_2, \ldots, Z_p$ the coordinate axes of these new variables being described by the vectors u_i which make up the matrix U of direction cosines used in the following transformation:

$$Z = U^{\backslash}(X - \overline{X}), \tag{48}$$

the transformed variables are called the principal components of X. The j-th principal component would be

$$z_i = u_i^{\backslash}(X - \overline{X}). \tag{49}$$

If one wishes to transform a set of variables X by a linear transformation $Z = U^{\backslash}(X - \overline{X})$, whether U is orthogonal or not, the covariance matrix of the new variables S_z can be determined directly from the covariance matrix of the original variables S by the relationship:

$$S_z = U^{\backslash} SU. \tag{50}$$

However, the fact that U is orthonormal is not a sufficient condition for the new variables to be independent. Only a transformation such as the principal axis transformation will produce the diagonal element S_z of the matrix L, where

$$L = \begin{bmatrix} S_1 & 0 & \cdots & 0 \\ 0 & S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_p \end{bmatrix}.$$

The fact that S_z is diagonal elements of L means that principal components are uncorrected. He added that we can also determine the correlation of each principal component with each of the original variables; this is useful for diagnostic purposes. The correlation of the i-th principal component S_z and j-th original variable X_j can be determined as

$$r_{z_i X_j} = \frac{u_{ij}\sqrt{L_i}}{s_j}. \tag{51}$$

Another interesting property of principal components is the fact that the Equation (48) can be inverted to

$$X = \overline{X} + U_z, \tag{52}$$

by virtue of the fact that U is orthonormal so that $U^{-1} = U^{\backslash}$. This means that if we know the values of the principal components, we can determine what the original data were.

Principal component form of T^2

Jackson [8] proposed that the T^2 can be expressed as a function of the principal components of the estimated covariance matrix. He gives an alternative form of the T^2 statistic as:

$$T^2 = \left(X - \overline{X}\right)^{\backslash} S^{-1}\left(X - \overline{X}\right) = \sum_{i=1}^{p} \frac{z_i^2}{\lambda_i}, \tag{53}$$

where $\lambda_1 > \lambda_2 > \cdots > \lambda_p$ are the eigenvalues of the estimated covariance matrix S and the z_i, $i = 1, 2, \ldots, p$, are the corresponding principal components.

Each of these component is obtained by multiplying the vector quantity $\left(X - \overline{X}\right)$ by the transpose of the normalized eigenvector U_i of S corresponding to λ_i: i.e., $z_i = U_i^{\prime}\left(X - \overline{X}\right)$.

Each z_i is a scalar quantity and the T^2 statistic is expressed in terms of these values.

The representation in (53) is derived from the fact that the estimated covariance matrix S is a positive definite symmetric matrix. Thus, its singular value decomposition is given as $S = UAU^{\prime}$, where U is a $p \times p$ orthogonal matrix whose columns are the normalized eigenvectors U_i of S, and A is a diagonal matrix whose elements are the corresponding eigenvalues,

$$U = (U_1, U_2, \ldots, U_p) \qquad \text{and} \qquad A = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}.$$

Note that

$$S^{-1} = UA^{-1}U^{\prime}. \tag{54}$$

Substituting this quantity into the T^2 statistic of (53), we have

$$\begin{aligned} T^2 &= (X_i - \overline{X})^{\prime} UA^{-1}U^{\prime}(X_i - \overline{X}) \\ &= Z^{\prime}A^{-1}Z \end{aligned}$$

$$= \sum_{i=1}^{p} \frac{Z_i^2}{\lambda_i}, \tag{55}$$

where $z_i = U_i^{\backslash}(X - \overline{X})$ and $Z^{\backslash} = z_1, z_2, \ldots, z_p$.

A Hotelling's T^2 statistic for a single observation also can be written as

$$T^2 = (X - \overline{X})^{\backslash} S^{-1}(X - \overline{X}) = Y^{\backslash} R^{-1} Y,$$

where R is the estimated correlation matrix and Y is the standardized observation vector of x, i.e.,

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1p} \\ r_{21} & 1 & r_{23} & \cdots & r_{2p} \\ \cdots & r_{32} & 1 & \cdots & r_{3p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{p_1} & \cdots & \cdots & r_{p(p-1)} & 1 \end{bmatrix},$$

where $r_{ij} = \mathrm{corr}(X_1, Xj)$ and

$$\left[\frac{(X_1 - \overline{X})}{S_1}, \frac{(X_2 - \overline{X})}{S_2}, \ldots, \frac{(X_p - \overline{X})}{S_p}\right] = [y_1, y_2, \ldots, y_p].$$

The matrix R (obtained from S) is a positive definite symmetric matrix and can be represented in terms of its eigenvalues and eigenvectors. Using a transformation similar to (64), the above T^2 can be written as

$$T^2 = \sum_{i=1}^{p} \frac{W_i^2}{\gamma_i}, \tag{56}$$

where $\gamma_1 > \gamma_2 > \ldots > \gamma_p$ are the eigenvalues of the correlation matrix R, and $w_1, w_2, \ldots, w_p$ are the corresponding principal components of the matrix R, w_i can be determined by the following transformation: $w_i = v_i^{\backslash}(X - \overline{X})$, $i = 1, 2, \ldots, p$, where the $v_1, v_2, \ldots, v_p$, are the corresponding normalized eigenvectors of R Equation (56) is not to be confused with (53). The first equation is written in terms of the eigenvalues and eigenvectors of the covariance matrix, and the second is in terms of the eigenvalues and eigenvectors of the estimated correlation matrix. These are two different forms of the same Hotelling's T^2 as the mathematical transformations are not equivalent.

The principal component representation of the T^2 plays a number of roles in multivariate statistical process control (SPC). The control region can be defined by the UCL. The observations contained in the T^2 values less than the UCL, i.e., for each X_i,

$$T_i^2 < \text{UCL}.$$

Thus, by (56)

$$T^2 = \frac{w_1^2}{\gamma_1} + \frac{w_2^2}{\gamma_2} + \ldots + \frac{w_p^2}{\gamma_p} < \text{UCL}.$$

The control region is defined by the equality

$$\frac{w_1^2}{\gamma_1} + \frac{w_2^2}{\gamma_2} + \ldots + \frac{w_p^2}{\gamma_p} = \text{UCL},$$

which is the equation of a hyper ellipsoid in a p-dimensional space provided the eigenvalues $\gamma_1 > \gamma_2 > \ldots > \gamma_p$ are all positive. The fact that the estimated correlation matrix R is a positive definite matrix guarantees that all the γ_i's are positive.

Note that in special case of the principal component space of the estimated correlation matrix, T^2 can be reduced to

$$T^2 = \frac{w_1^2}{\gamma_1} + \frac{w_2^2}{\gamma_2} = \text{UCL}, \quad (57)$$

which gives the equation of the control ellipse. The length of the major axis of the ellipse in (57) is given by γ_1 and the length of the minor axis is given by γ_2. The axes of this space are the principal components, w_1 and w_2. The absence of a product term in this representation indicates the independency between w_1 and w_2. This is a characteristic of principal components, since they are transformed to be independent.

Assuming that the estimated correlation r is positive it can be shown that $\gamma_1 = (1 + r)$ and $\gamma_2 = (1 - r)$. For negative correlations, the γ_i values are reversed. One can also show that the principal components can be expressed as $w_1 = (y_1 + y_2)\sqrt{2}$, $w_2 = (y_2 - y_1)\sqrt{2}$. From these equations, one can obtain the principal components as functions of the original variables.

4. Application of Multivariate T^2 in Industrial

Delta Fertilizers and Chemical Industries is considered one of the leading companies in the field of fertilizers production in Egypt. About 4500 employees are working for it, on the various managerial levels. Urea production is one of the major products of the company. The production of urea occurs through three stages, summarized as follows:

A. High pressure stage

In this stage, urea is produced through two reactions; the first reaction occurs by condensation of Ammonia gas and Carbon dioxide under high pressure and temperature for the sake of the production of intermediate material, known as Carbamate. The second reaction happens by separating the water from the Carbamate in order to a chive urea. In this, stage the condensation of urea approximately 56%.

It contains 16 variables, these are:

1	X1	E-201Outlet Temperature
2	X2	Outlet cold NH3 from E- 201
3	X3	CO_2 to Train
4	X4	CO_2 pressure to synthesis
5	X5	CO_2 after E-22
6	X6	R-201
7	X7	Temperature in reactor R-201
8	X8	Temperature in reactor R-201
9	X9	Temperature in reactor R-201
10	X10	Temperature in reactor R-201
11	X11	Stripper level
12	X12	Liquid leaving the Stripper
13	X13	Stream from E-204 to j-201
14	X14	Conditioned water to scrubber E-204
15	X15	Conditioned water from scrubber E-204
16	X16	Stream from j-203

Table analysis of laboratory in this stage:

1	t1.1	NH_3	Reactor outlet
2	t1.2	CO_2	Reactor outlet
3	t1.3	UR	Reactor outlet
4	t1.4	B_1	Reactor outlet
5	t1.5	H_2O	Reactor outlet
6	t2.1	NH_3	Stripper outlet
7	t2.2	CO_2	Stripper outlet
8	t2.3	UR	Stripper outlet
9	t2.4	B_1	Stripper outlet
10	t2.5	H_2O	Stripper outlet

B. Low pressure stage

In this stage, the condensation of urea liquid rises from 56% to 71%. This happens through the decomposition of the remaining Carbamate and the elimination of water under low pressure.

It contains seven variables, these are:

1	y1	Urea solution from stripper E-202
2	y2	Steam to E-205
3	y3	Urea carbonate solution from stripper T-201 to E-205
4	y4	Gas leaving T-201
5	y5	Level in TK-201
6	y6	P-203
7	y7	Urea solution in TK-201

Table analysis of laboratory in this stage:

1	t3.1	NH_3	D 202 Outlet
2	t3.2	CO_2	D 202 Outlet
3	t3.3	UR	D 202 Outlet
4	t3.4	B_1	D 202 Outlet
5	t3.5	H_2O	D 202 Outlet
6	t4.1	NH_3	In TK 201
7	t4.2	CO_2	In TK 201
8	t4.3	UR	In TK 201
9	t4.4	B_1	In TK 201
10	t4.5	H_2O	In TK 201
11	t5.1	NH_3	In PI 302
12	t5.2	CO_2	In PI 302
13	t5.3	UR	In PI 302

C. Evaporation and prilling stage

This stage occurs by two stages:

(i) Evaporation stage

In this stage, the condensation of urea rises from 71% to 98.7% approximately and the urea liquid trams forms to urea melt. This happens under high pressure and temperature.

(ii) Prilling stage

In this stage, the urea melt is through formed into prilling in the prilling tower.

It contains four variables, these are:

1	Z1	Urea solution from D-204 to E-209
2	Z2	D-205 Vacuum
3	Z3	Urea to prilling tower X-202
4	Z4	E-211 Vacuum

Table analysis of laboratory in this stage:

1	t6.1	B_1
2	t6.2	H_2O
3	t6.3	Pills > 3.35
4	t6.4	Pills 3.35: 2.4
5	t6.5	Pills 2.4: 1.4
6	t6.6	Pills 1.4: 1.0
7	t6.7	Pills < 1.0
8	t6.8	UR

Data description:

For the application of multivariate quality control, chart data originate from urea production process, which consists of the three stages and the analysis of laboratory, which discussed above.

The number of the sample is 732 observations taken per hour.

The advantages of this sample that, it has several variables and several stages of the production. This advantage of the production is the basic reason for choosing this production to allow us to study the multivariate quality control charts.

In this application, we shall introduce the most common using technique of multivariate quality control charts; MEWMA control chart & Hotelling's T^2 control chart.

A Hotelling T^2 chart consists of:

■ Plotted points, each of which represents T^2 statistic for each observation.

■ A center line (green), which is the median of the theoretical distribution of T^2 statistic.

■ Control limits (red), which provide a visual means for assessing whether the process is in-control. The control limits represent the expected variation.

MINITAB marks points outside of the control limits with a red symbol.

MINITAB indicates which points is out-of-control by using decomposition of T^2 statistic, along with the P-value for each significant variable.

4.1. Hotelling T^2 chart of $\mathrm{X1}, \ldots, \mathrm{X16}$ and $\mathrm{t1.1}, \ldots, \mathrm{t2.5}$

Test results for T squared chart of $\mathrm{X1}, \ldots, \mathrm{X16}$ and $\mathrm{t1.1}, \ldots, \mathrm{t2.5}$.

TEST. One point beyond control limits.

Test Failed at points: (Less than LCL)

114	118	123	128	133	134	138	140	143	147
151	156	161	166	167	171	173	349	352	353
354	361	373	374	378	384	385	386	576	578
582	586	590	594	598	600	604	608	612	616
620	624	626	630	634	638	642	646	650	652
656	660								

Test Failed at points: (Greater than UCL)

13	20	30	40	43	45	48	50	52	56
60	66	80	100	245	247	250	252	254	259
261	265	276	448	458	483	636	647	657	658
663	664	703	714	715	718	721	724	727	730

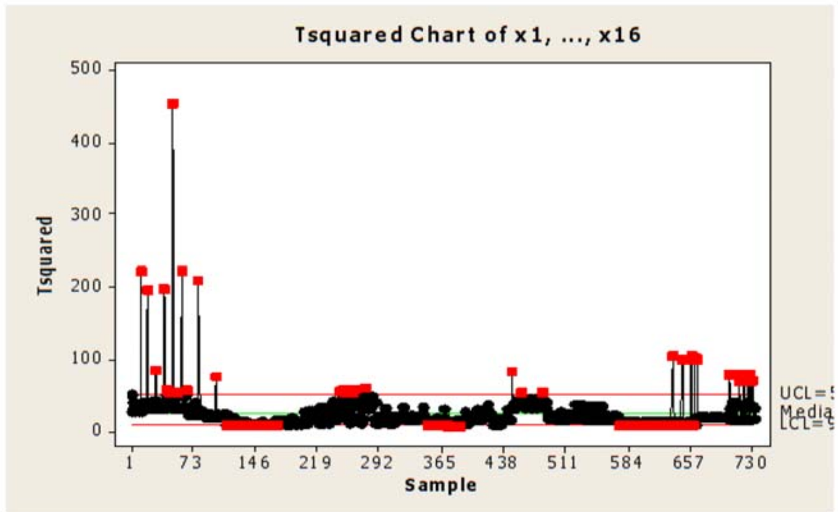

Figure 1. T^2 chart of X1, ... , X16 and t1.1, ... , t2.5.

The Hotelling T^2 chart of X1, ... , X16 and t1.1, ... , t2.5 can be summarized as follows:

- The lower and upper control limits are 9.7 and 52, respectively. Therefore, we expect the T^2 Statistics to fall between 9.7 and 52. The center line or median, is 25.3.

- Test results indicate that 52 points less than LCL, for example, point 114 exceeds the lower control limit.

- Test results indicate that 40 points greater than UCL, for example, the test results indicate that point 13 exceeds the upper control limit.

● Test results indicate 92 point through beyond the control limits. Then the out-of-control rate 12.6% and the in-control rate 87.4%.

4.2. Hotelling T^2 chart of y1, …, y7 and t3.1, …, t4.5

Test results for T squared chart of y1, …, y7 and t3.1, …, t4.5.

TEST. One point beyond control limits.

Test Failed at points: (Greater than UCL)

28	91	114	130	150	200	250	256	264	268
551	703	714	715	718	721	724	727	730	

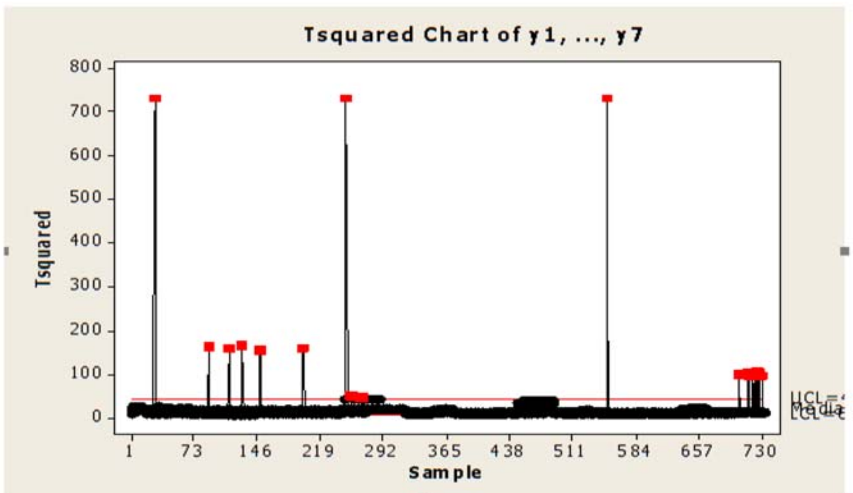

Figure 2. T^2 chart of y1, …, y7 and t3.1, …, t4.5.

The Hotelling T^2 chart of y1, …, y7 and t3.1, …, t4.5 can be summarized as follows:

■ The lower and upper control limits are 0.2 and 43.6, respectively. Therefore, we expect the T^2 statistics to fall between 0.2 and 43.6. The center line, or median, is 19.3.

■ Test results indicate that 19 point greater than UCL, for example test results indicate that point 91 exceeds the upper control limit.

■ Test results indicate 19 point that are beyond the control limit. Then the out of control rate 2.59% and the in-control rate 97.41%.

4.3. Hotelling T^2 chart of Z1, ..., Z4 and t6.1, ..., t6.8

Test results for T squared chart of Z1, ..., Z4 and t6.1, ..., t6.8

TEST. One point beyond control limits.

Test Failed at points: (Greater than UCL)

245	248	250	252	254	259	261	265	269	272	276
280	284	489	491	493	495	497	515	517	519	

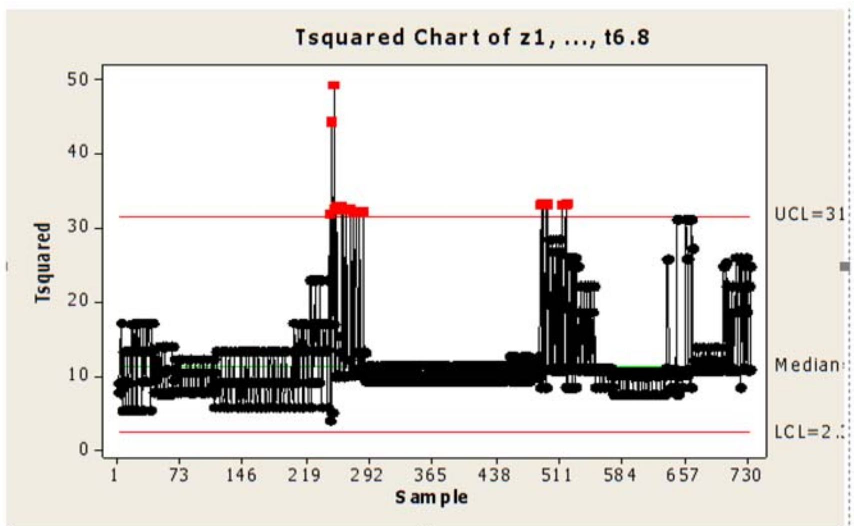

Figure 3. T^2 chart of Z1, ..., Z4 and t6.1, ..., t6.8

The Hotelling T^2 chart of Z1, ..., Z4 and t6.1, ..., t6.8 can be summarized as follows:

■ The lower and upper control limits are 2.37 and 31.63, respectively. Therefore, we expect the T^2 statistics to fall between 2.37 and 31.63. The center line, or median, is 11.35.

■ Test results indicate that 21 point greater than UCL, for example test results indicate that point 245 exceeds the upper control limit.

■ Test results indicate 21 point that are beyond the control limit. Then the out of control rate 2.87% and the in-control rate 97.13%.

5. Test Results of Application

The application is shown that in high process stage, test results of Hotelling T^2 chart indicate that the out-of-control percentage 87.4% and the in-control percentage 12.6%, and it shown that in low process stage, test results of Hotelling T^2 chart indicate that the out-of-control percentage 2.59% and the in-control percentage 97.41%.

It is shown that in the evaporation and prilling stage, test results of Hotelling T^2 chart indicates that the out-of-control percentage 2.87% and the in-control percentage 97.13%.

6. Conclusions

The results allow us to determine whether the joint process variability is in control or out-of-control. Hotelling's T^2 charts used to determine whether or not the process mean vector (a vector of the process means that accounts for the mean of each charted variable) for two or more variables is in-control. An in-control process exhibits only random variation with the control limits. An out-of-control process exhibits unusually variation, which may be due to the process of assignable causes (unusual occurrences that are not normally part of the process). T^2 charts allow us to simultaneously monitor whether two or more related variables are in control.

Finally

■ The company should use multivariate Hotelling's T^2 quality control chart to monitor the quality of the urea production.

■ Too, the company should use the Hotelling's T^2 chart to determine variables which causes the out-of-control signals.

Acknowledgement

The author is grateful to the associate editor and referees for careful reading and the Editor for very useful comments and suggestions.

References

[1] F. B. Alt, Multivariate Quality Control, Encyclopedia of the Statistical Sciences, John Wiley & Sons, New York, 1985.

[2] F. B. Alt and N. D. Smith, Multivariate Process Control, Handbook of Statistics, P. R. Krishnaiah and C. R. Rao, Eds., North-Holland, Amsterdam, Elsevier Science, B. 7 (1988), 333-351.

[3] M. S. Hamed, MEWMA control chart procedure: Average run length performance with application, Journal of Statistics: Advances in Theory and Applications 15(2) (2016), 163-203.

[4] H. Hotelling, Multivariate Quality Control in Techniques of Statistical Analysis, McGraw Hill, New York, 1947, 111-184.

[5] A. Houshmand and A. Javaheri, Multivariate Ridge Residual Charts, Ph.D. Dissertation, University of Cincinnati, 1998.

[6] J. E. Jackson, Principal components and factor analysis: Part 1: Principal components, Journal of Quality Technology 12 (1980), 201-213.

[7] J. E. Jackson, Multivariate quality control, Communications in Statistics-Theory and Methods 11 (1988), 2657-2688.

[8] J. E. Jackson, A User's Guide to Principle Component, John Wiley & Sons, New York, 1991.

[9] J. E. Jackson and R. H. Morris, An application of multivariate quality control to photographic processing, Journal of the American Statistical Association 52 (1957), 186-199.

[10] W. Linna, H. Woodal and L. BusBy, The performance of multivariate control charts in the presence of measurement error, journal of Quality Technology 33 (2001), 133-139.

[11] C. A. Lowry and D. C. Montgomery, A review of multivariate control charts, IIE Transactions 27 (1995), 800-810.

[12] R. L. Mason, N. D. Tracy and J. C. Young, Multivariate statistical process control with industrial application, The American Statistical Association and Society for Industrial and Applied Mathematics (2002).

[13] R. L. Mason, N. D. Tracy, S. T. Wierda and J. C. Young, Improving the sensitivity of T^2 statisitc in multivariate process control, Journal of Quality Technology 31 (1999), 155-165.

[14] R. L. Mason, N. D. Tracy and J. C. Young, Decomposition for multivariate control chart interpretation, Journal of Quality Technology 27 (1995), 99-108.

[15] Gerald I. Onwuka and H. Hotelling, T-Square & principal component analysis approaches to quality control sustainability, International Journal of Computational Engineering Research (IJCER) 2 (2012), 211-217.

[16] A. C. Rencher, The contribution of individual variables to Hotelling's T^2 Wilks A, R^2, Biometrics 49 (1993), 479-489.

[17] T. P. Ryan, Statistical Methods for Quality Improvement, John Wiley, New York, 1989.

[18] G. A. F. Seber, Multivariate Observations, John Wiley, New York, 1984.

[19] H. M. Shehata, Statistical Study of Multivariate Quality Control Procedures and its Applications, M. Sc. Thesis, Benha University, 2007.

[20] W. A. Shewart, Economic Control of Quality of Manufactured Product, D. Van Nostrand Company INC, New York, 1931.

[21] N. D. Tracy, J. C. Young and R. L. Mason, Multivariate control charts for individual observations, Journal of Quality Technology 24 (1992), 88-95.

[22] J. A. Vargas, Robust estimation in multivariate control chart for individual observations, Journal of Quality Technology 35(4) (2003), 367-376.

[23] S. J. Wierda, Multivariate statistical process control-Recent results and directions for future research, Statistica Neerlandica 48 (1994), 147-168.

[24] W. H. Woodall, D. J. Spitzner, D. C. Montgomery and S. Gupta, Using control charts to monitor Process and product quality profiles, Journal of Quality Technology 36 (2004), 309-320.

■