Multivariate Statistical Quality Control Based on Ranked Set Sampling

Hakan Eygü[1] & M. Suphi Özçomak[2]

[1] Department of Econometrics, University of Atatürk, Erzurum, Turkey

[2] Department of Econometrics, University of Atatürk, Erzurum, Turkey

Correspondence: Hakan Eygü, Faculty of Economics and Administrative Sciences, Ataturk University, 25240, Erzurum, Turkey. E-mail: hakaneygu@atauni.edu.tr

Abstract

The sample of the study was formed using simple random sampling, ranked set sampling, extreme ranked set sampling and median ranked set sampling. At the end of this process, the researcher created Hotelling's T^2 control charts, a multivariate statistical process control method. The performances of SRS, RSS, ERSS and MRSS sampling methods were compared to one another using these control charts. A simulation was performed to see the average run-length values for Hotelling's T^2 control charts, and these findings were also used for the comparison of the sampling performances.

At the end of the study, the researcher formed a sample using median ranked set sampling and created the Hotelling's T^2 control chart. As a result of this operation, the researcher found that there was an out-of-control signal in the process, while there was no such signal in other sampling methods. When the average run-length values obtained from Hotelling's T^2 control charts were compared, it was seen that a shift in the process was detected by the ranked set sampling earlier, when compared to other sampling methods. This paper it can be said that the methods used are unique to the literature because they are applied to multivariate data.

Keywords: Statistical quality control; multivariate control charts; Hotelling T^2 control charts; ranked set sampling; average run length

1. Introduction

Statistical process control can be defined as a procedure that uses statistical methods to check whether or not a manufacturing and service process is working normally, and that detects an abnormal incident and eliminates it by determining the reasons for it (Burnak, 1997:61). The most important purpose of the statistical process control is to eliminate the specific reasons for a change in the process and to keep it in control (Işığıçok, 2012:151). Univariate or multivariate statistical methods can be used depending on the type of the process. Thus, the difference between univariate or multivariate process control methods should be addressed primarily. Statistical process control, consists of a number of powerful tools for problem solving and improvement of quality control through reducing variability in industrial manufacturing processes.

Throughout our study of multivariate statistical methods for quality and productivity improvement, we was concerned with essentially the same problem: using data obtained from the process to draw conclusions known as inferences about how it is or has been operating. We may be concerned with sampling the process. In each case we will find it useful to the improvement of quality and productivity to collect data from the process and model it using statistical concepts and methods.

Multivariate statistical process control is defined as a methodology based on control charts, and it is used to observe multivariate processes. The number of variables included is a significant difference between the methods, but actually there are much more important differences. One of these differences is the fact that the variables obtained from the multivariate processes are often related to one another. These variables should be examined together since this type of variable is inter-dependent (Mason, Young, 2002: 6). Multivariate statistical control methods have been of great importance in foreign literature over the last couple of years, and these methods have been used in industry in particular. Control charts are often used since they make it possible to visually monitor a change in the process and can be easily used and interpreted (Eygü, 2015:80). Multivariate process control is a methodology, based on control charts, that is used to monitor the stability of a multivariate process.

Schwertman et al. (2002) used Hotelling's T^2 method together with regression analysis and specified the changes

occurring in the control charts over time.

Aparisi et al. (2004) suggested that Hotelling's T^2 method can be used to determine whether or not a process is out-of-control with sample data randomly obtained from the process. They compared the average-run-length values of chi-square control charts using certain assumptions.

1.1 Hotelling's T^2 quality control chart

The multivariate quality control chart, developed by Harold Hotelling (1947), is based on T^2 values with the assumption that the distribution of random variables is normal. This chart is often called the multivariate Shewhart control chart due to its resemblance to Shewhart control chart in many sources. It is also called a chi-square chart when it has a chi-square distribution with *p*-value degrees of freedom (Montgomery, 2009).

Consider *p* correlated characteristics as being measured simultaneously and these characteristics as being measured simultaneously and these characteristics follow a multivariate normal distribution with mean vector μ'_0 = ($\mu_{0.1}$, $\mu_{0.2}$,... $\mu_{0.p}$), and covariance matrix Σ_0 when the process is in control, where μ_{0j} is the mean for the *j*th characteristic and Σ_0 is a *p x p* matrix consisting of the variances and covariances of the *p* characteristics. When an *i*th sample of size *n* is taken, we have *n* values of each characteristic and it is possible to calculate the $\bar{X}_i$ vector, which represents the *i*th sample mean vector for the p characteristics. The charting statistic

$$T_i^2 = n\left(\bar{X}_i - \mu_0\right)' \Sigma^{-1}\left(\bar{X}_i - \mu_0\right) \tag{1}$$

is called Hotelling's T^2 statistic that, when the process is in control, is distributed as a chi-square variate with *p*

degrees of freedom ($T_i^2 \approx \chi_p^2$). If $T_i^2 > \chi_{p,\alpha=0.03}^2$ is in question, the process is considered to be out-of-control.

(Aparisi, 2007).

Let's assume that μ_x and Σ parameters are known to be in normal multivariate distribution, and suppose that μ_x is a Y observation vector obtained from a normal multivariate distribution that has a different mean vector but the same covariance matrix. This observation vector is demonstrated by

$$T_y^2 = \left(Y - \mu_x\right)' \Sigma^{-1}\left(Y - \mu_x\right) \tag{2}$$

The statistic here cannot be defined with the χ^2, the center of which is μ_x and μ_y. Thus, multivariate normal distribution defined with *(Y − μ_x)* vector has a mean value that is different than 0. However, the center of the *(Y − μ_x)* normal vector can be determined considering the mean μ_x and μ_y vectors and in parallel to the $\bar{X}_1$ and $\bar{X}_2$ axes. If we keep

$$\left(Y - \mu_x\right) = \left(Y - \mu_x + \mu_y - \mu_y\right) = \left[\left(Y - \mu_y\right) + \delta\right] \tag{3}$$

in mind, $\delta = (\mu_y − \mu_x)$ indicates the mean deviation figure. With this result, the distribution of T_y^2 is given by

$$T_y^2 = \left[\left(Y - \mu_y\right) + \delta\right]' \Sigma^{-1}\left[\left(Y - \mu_y\right) + \delta\right] \sim \chi_{(p,\lambda)}'^2 \tag{4}$$

where $\chi'^2_{(p,\lambda)}$ is a non-central chi-square distribution with *p* degrees of freedom. A major difference between this distribution and the central chi-square is the additional parameter λ, labeled the non-centrality parameter (Mason and Young, age., 2002). This parameter is demonstrated as $\lambda = nd^2$. When the in-control mean vector μ_0 shifts to $\mu_1 = \mu_0 + \delta (\delta \neq 0)$, the magnitude of this shift is often expressed by

$$d = \sqrt{(\mu_1 - \mu_0)' \Sigma^{-1}(\mu_1 - \mu_0)} = \sqrt{\delta' \Sigma^{-1} \delta} \tag{5}$$

the Mahalanobis distance and μ is the p characteristics mean vector. The subgroup statistic T_i^2 then follows a non-central chi-square distribution with *p* degrees of freedom and non-centrality parameter $\lambda = nd^2$, that is $T_i^2 \sim \chi_p^2(\lambda)$. (Rakitzis, Antzoulakos, 2011). The average run length value of Hotelling's T^2 control chart is dependent upon the mean μ vector and covariance matrix, $\sum$, thanks to the non-central parameter demonstrated by *d*. In this equation, μ_0 indicates the target mean vector. It is possible to consider average run length as a function of *d*. We denote in-control state with *d* = 0, hence we have $\alpha = P(T^2 > UCL| d = 0)$. When the process is out of control with shift $d \neq 0$, hence we have $\beta = P(T^2 < UCL| d \neq 0)$ (Faraz & Moghadam, 2009:904). Faraz and Parsian (2006) suggested that schemes result in more rapid detection of lack of control and hence, reduce the

costs associated with nonconforming products. They showed that adding an additional warning line improves the performance of T^2 control chart using variable sample size and variable sampling interval scheme.

1.2 Average run length (ARL)

ARL of the chart is a common measure of how well a chart performs in detecting an out-of-control process. The run length is the number of the sampling stage at which the chart first signals. Woodall (1985) explained that the run length of a control procedure is the number of samples required before an out-of-control signal is given. An out-of-control signal indicates that a shift in the mean is likely to have occurred and that action should be taken to find and correct the assignable cause of this shift. We use the average run length (ARL) to measure the performance of a control procedure, although a percentage point of the run length distribution may be a more appropriate measure in some applications. Also he said that it is often not practical to detect quickly shifts from the target value which are too small to be of practical importance. The proper choice of a control procedure depends on the selection of in-control and out-of-control regions of parameter values. Woodall and Montgomery (1999) explained that ARL is a performance measure that is widely used to evaluate control charts. In the present study, in-control ARL was used to compare the performance of the control charts. *ARL* is defined as the average number of observations required for the control chart to detect a change under the in-control process.

Aparisi (1996) states that ARL is commonly used to compare quality control charts. ARL indicates the number of samples to be taken before a sample indicates an out-of-control event in any process. ARL is used to determine which sample in the control charts will be out of control first. ARL is expected to be a large numerical value in a control process. However, ARL is expected to be a small value to detect a shift faster when the process is out of control (Aparisi, 2007).

The average run length (ARL) for a control procedure is defined as

$$ARL = \frac{1}{p} \tag{6}$$

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 a Type I error. The ARL has a number of uses in both univariate and multivariate control procedures. (Mason & Young, 2002). Javaheri and Houshmand (2001) determined different covariance structures using p-dimensional multivariate data. Various amount of shift in the mean vector is induced and the resulting ARL is computed. They evaluated the effectiveness of ARL by considering five different methods including Hotelling's T^2, Shewhart Control Charts, Discriminant Analysis, Decomposition Method, and Multivariate Ridge Residual Chart.

Two cases are presented for ARL. The first is that the process mean is at the targeted values which is called controlled ARL and indicated as ARL_0. The second case is when the process mean deviates – called out-of-control ARL – and indicated as ARL_1. If the process mean is at the targeted values, the signal indicated by the control chart is false. Thus, the expected ARL value should be large in this case. When the process mean deviates, the signal indicated by the control chart is correct, thus the expected ARL value should be small (Cox, 2001).

ARL can also be used to calculate the number of mean expected observations before an out-of-control signal is present in the process, in other words when the process is in control. Thus, the ARL value can be obtained using the following Eq. (7),

$$ARL_0 = \frac{1}{\alpha} \tag{7}$$

When a shift is present in the process, another use of ARL is to calculate the number of observations of the shift before the shift itself is detected. The probability of detecting any shift, and possible deviation for the β value is equal to $(1-\beta)$. The β value represents Type II error probability. If a shift is present, this probability deviations can be determined using standard statistical equals. The ARL for detecting the shift is given by

$$ARL_1 = \frac{1}{1 - \beta} \tag{8}$$

The probability $(1-\beta)$ represents the power of the test of a statistical hypothesis that the mean has shifted. This result produces another major use of the ARL, which consists of comparing one control procedure to another. This is done by comparing the ARLs of the two procedures for a given process shift (Mason & Young, age., 2002). As will be seen, the probability β is a function of the distributional parameters μ, μ_0, and Σ_0 of the

distribution of the vector of quality measurements, X, only through the value *d*, where $d = \sqrt{(\mu - \mu_0)'\Sigma^{-1}(\mu - \mu_0)}$ with *μ* the out-of-control value of the mean vector provided $\Sigma = \Sigma_0$ (Champ and Aparisi, 2008:155).

1.3 Ranked set sampling (RSS)

Patil et al. (1999) suggested that RSS is a sampling method in which the selected units are ordered first and measurements are performed on these units later, without the necessity to measure all units; thus, the population parameters are estimated. Estimations made with this method use different characteristics such as objectivity and consistency.

In the RSS method recommended by George A. Mclntyre (1952), sampling is selected by following the steps below:

- First, a random sample of size n^2 units is selected from the target population, and divided randomly into the sample which consists of *n* sets, each of size *n*. The units within each set are then ranked with respect to a variable of interest.

- Units in the sets are ordered from the smallest to the largest using an inexpensive and easy measurement – so that a precise measurement is not required. This process can be visually conducted or a variable can be used for this purpose that can be measured with little expense.

- The first unit is selected from the first set – in which the units have been ordered. The second unit is selected from the second set, and the nth unit is selected from the nth set.

- The procedure continues in this manner until the largest unit has been selected from n[th] set. The cycle may be repeated *r* times until *nr* units have been measured.

Samples with a size of *n* are taken at low size to facilitate visual ordering. For RSS, literature recommends $n = 2$, 3, 4, 5 or 6. When a sample with greater size is needed, the sampling selection is performed *r* times to obtain an RSS with a magnitude of *nr*.

Balakrishnan and Cohen (1991) explained unbiased estimator of the population mean by RSS as follows:

$$\bar{X}_{SK} = \frac{1}{rn} \sum_{j=1}^{r} \sum_{i=1}^{n} X_{i[i:n]j} \tag{9}$$

To this end, we let $X_{i[i:\{n\}]j}$ indicates the ith row statistics of the ith set with *n* magnitude at the jth repetition with the assumption that no error has been made ($j = 1,2,\dots r$, $i = 1,2,\dots n$). Ordered statistics will be different from one another in this sampling.

Variance of $\bar{X}_{SK\ddot{o}}$ is defined as:

$$Var(\bar{X}_{SK}) = \frac{1}{n^2 r} \sum_{i=1}^{n} \sigma^2_{(i:n)} \quad \text{and} \quad \sigma^2_{(i:n)} = E\left[X_{(i:n)} - E(X_{(i:n)}) \right]^2 \tag{10}$$

$\sigma^2_{(i:n)}$ is the variance of the ith row of statistics in a random sample with size *n*.

Champ and Aparisi (2008) explained sampling is a simple addition to a control chart that significantly increases the ability of the chart in detecting various changes in the process. This is the case with the sampling control chart based on Hotelling's T^2 statistics introduced with paper. They differ only in how they use the second sample to make a decision about the process.

1.4 Extreme ranked set sampling (ERSS)

In cases where the size of the sample is great, it will be easier to determine the largest and smallest units with visual ordering. Thus, ERSS design has been recommended in place of RSS in these situations.

Samawi et al. (1996) stated that ERSS was performed using the following steps:

- *n* samples at *n* magnitude are randomly selected.

- Units in each set are ordered considering the variable processed with visual or inexpensive methods. It is accepted that this ordering is as good as precise measurement. Sampling selection in ERSS varies depending on *n* (whether it is odd or even).

- If the set size *n* is even, select from n/2 sets the smallest unit and from the other n/2 sets, the largest unit for actual measurement.

- If the set size is odd, select from (n-1)/2 sets the smallest unit, from the other (n-1)/2 the largest unit and from one set the median of the set for actual measurement. The cycle may be repeated *r* times to get *nr* units. These *nr* units from the ERSS data.

If *n* is even, estimator of the mean population is calculated using the sample having a size of *n* obtained with the ERSS equals to:

$$\bar{X}_{USK} = \frac{1}{n}\left[\sum_{i=1}^{n/2} X_{i[1,n]} + \sum_{i=(n/2)+1}^{n} X_{i[n,n]}\right] \tag{11}$$

with variance

$$Var(\bar{X}_{USK}) = \frac{1}{n^2}\left[\sum_{i=1}^{n/2} Var(X_{i[1,n]}) + \sum_{i=(n/2)+1}^{n} Var(X_{i[n,n]})\right] \tag{12}$$

If *n* is odd, the ERSS estimator is

$$\bar{X}_{USK} = \frac{1}{n^2}\left[\sum_{i=1}^{\frac{n-1}{2}} Var(X_{i[1,n]}) + X_{\frac{n+1}{2}\left[\frac{n+1}{2},n\right]} + \sum_{i=\frac{n+1}{2}+1}^{n} X_{i[n,n]}\right] \tag{13}$$

with variance

$$Var(\bar{X}_{USK}) = \frac{1}{n^2}\left[\sum_{i=1}^{\frac{n-1}{2}} Var(X_{i[1,n]}) + Var\left(X_{\frac{n+1}{2}\left[\frac{n+1}{2},n\right]}\right) + \sum_{i=\frac{n+1}{2}+1}^{n} Var(X_{i[n,n]})\right] \tag{14}$$

1.5 Median ranked set sampling (MRSS)

The median ranked set sampling (MRSS) method aims to select *n* random sets each of size *n* from the population, and rank the units within each set with respect to a variable of interest (Jabeen, 2011). Muttlak and Al-Sabah (2003) developed various quality control charts in their study to detect the deviations from the mean value using RSS, and ERSS and MRSS – which are modifications of RSS. Control charts, RSS and the modifications to RSS were used within that study, and out-of-control points were detected with smaller samples by calculating ARL values.

Muttlak (1997) suggested that MRSS yielded better results than RSS in estimating the mean population for unimodal symmetrical distributions such as normal distribution. This design is created by selecting the median values in each set. The stages of selecting the samples with MRSS are as follows:

- *n* samples at *n* magnitude are randomly selected.

- Units in each set are ordered by considering the variables processed by visual or inexpensive methods. It is accepted that this order is as good as precise measurement. The sampling selection in MRSS varies depending on *n* (whether it is odd or even).

- If the set size *n* is odd, from each set select for measurement the ((n+1)/2)th smallest rank. If the set size is even, select for measurement from the first n/2 sets the (n/2)th smallest rank and from the second n/2 sets the ((n+2)/2)th smallest rank. As we can see in both cases we will get a sample of size *n* measured units. The cycle may be repeated *r* times to get *nr* units. These *nr* units form the MRSS data.

If *n* is even, estimator of the mean population is calculated using the sample with *n* size obtained with MRSS using the following equal:

$$\bar{X}_{MSK} = \frac{1}{n}\left[\sum_{i=1}^{n/2} Var(X_{i[n/2,n]}) + \sum_{i=n/2+1}^{n} Var(X_{i[(n/2)+1,n]})\right] \tag{15}$$

with variance

$$Var(\bar{X}_{MSK}) = \frac{1}{n^2}\left[\sum_{i=1}^{n/2} Var(X_{i[n/2,n]}) + \sum_{i=n/2+1}^{n} Var(X_{i[(n/2)+1,n]})\right] \tag{16}$$

If *n* is odd, the MRSS estimator is

$$\overline{X}_{MSK} = \frac{1}{n}\left[\sum_{i=1}^{n} X_{i\,\frac{1}{n+1}}, n\right] \tag{17}$$

with variance

$$Var(\overline{X}_{MSK}) = \frac{1}{n^2}\left[\sum_{i=1}^{n} Var\left(X_{i\left[\frac{n+1}{2},n\right]}\right)\right] \tag{18}$$

2. Application

Statistical process control was performed in a cement plant using Hotelling's T^2 quality control chart, one of the multivariate statistical process control methods. During the application, silicon oxide (SiO_2), aluminum oxide (Al_2O_3), calcium oxide (CaO) and Sulphur oxide (SO_3), chemicals that affect the quality of the cement produced in the plant, and endurance (N/mm^2) data were collected over 28 days.

Although there are plenty of variables that affect the quality of the cement, five variables with high correlation rates were used within the study. Thus, a dataset with 100 samples was obtained from the above-mentioned variables. Multivariate statistical process control was performed using these variables.

Samples were selected from the manufacturing process using SRS, RSS, ERSS and MRSS methods, and recursion was performed 10,000 times using the simulation method. The mean of these recursions was calculated using Hotelling's T^2 statistics. ARL values were calculated with different sampling methods and compared to one another to see how successfully they detected deviations in the process mean.

2.1 Comparing SRS, RSS, ERSS and MRSS Methods with ARL

The ARL value was determined for Hotelling's T^2 quality control chart. The real data were collected using the SRS, RSS, ERSS and ERSS in cases of perfect and imperfect ranking to make this determination. SRS, RSS, ERSS and MRSS methods were used in each recursion, and samples were taken at a size of $n=3$ and $n=4$. The ARL value was calculated to be $1/0.0027=370$ the process in control conditions. Therefore, we realized that the control chart gives an out-of-control signal approximately once every 370 samples when the process is in control. Each method was revised in a way to make the ARL value approximately 370 when the process was in control. Each deviation valued was determined to be 0, 0.50, 1.00, 1.50, 2.00, 2.50, 3.00, 3.50 and 4.00, respectively. Hotelling's T^2 control chart values for SRS, RSS, ERSS and MRSS methods are presented in Table 1 and 2 for $p=3$, $n=3$ and $p=4$, $n=4$. In addition, a certain shift magnitude and different upper control line for the p variables were presented too, to ensure that ARL is equal to 370 when the process is in control.

The ARL values of Hotelling's T^2 control chart are presented in Table 1 for $p=3$ and $n=3$ sampling size and for different of d.

Table 1. Comparison of the ARL for $p=3$

	T^2			
	$CL=X^2_{3;\,0.0027}$			
d	SRS	RSS	ERSS	MRSS
0.00	370.37	368.18	368.18	368.59
0.50	156.25	151.16	165.77	134.45
1.00	43.86	40.80	40.80	33.22
1.50	14.97	15.07	15.07	14.98
2.00	6.31	6.25	6.24	6.05
2.50	3.24	3.46	3.31	3.16
3.00	1.98	1.99	1.98	1.07
3.50	1.37	1.36	1.36	1.00
4.00	1.02	1.01	1.00	1.00

CL=central line; SRS=simple random sampling; RSS=ranked set sampling; ERSS=extreme ranked set sampling; MRSS=median ranked set sampling

As presented in Table 1, the ARL was found to be 134 for MRSS for a small shift magnitude of d=0.5. The MRSS method determined whether or not the process was an out-of- control with the smallest samples. Thus, a

small ARL value indicates that MRSS is more effective than other methods in the early detection of shifts with smaller samples.

Table 2 presents ARL values of Hotelling's T^2 control charts for different d for four different quality variables: Al_2O_3, CaO, SO_3 and endurance (N/mm^2) data over 28 days, ordered in relation to type II variable (SiO_2) and an $n = 4$ sampling size.

Table 2. Comparison of the ARL for $p=4$

	T^2			
	$CL = X^2_{4;\ 0.0027}$			
d	SRS	RSS	ERSS	MRSS
0.00	369.23	365.06	370.23	370.64
0.50	158.68	144.94	150.08	138.29
1.00	44.01	40.26	43.03	38.08
1.50	14.49	13.84	13.82	12.89
2.00	6.35	6.28	6.66	5.93
2.50	3.19	3.21	3.21	2.95
3.00	1.95	1.97	1.97	1.04
3.50	1.34	1.25	1.26	1.00
4.00	1.22	1.00	1.00	1.00

CL=central line; SRS=simple random sampling; RSS=ranked set sampling; ERSS=extreme ranked set sampling; MRSS=median ranked set sampling

The researcher used the sampling methods in Table 2 to calculate the average number of samples that should be taken to determine whether or not the process is out-of-control at a sampling size of $n = 4$. The ARL values were calculated as well. The researchers took the small shift values into consideration to see the magnitude of the samples that would be created for determining any out-of-control incidence with the assistance of Hotelling's T^2 control chart. Thus, as the table shows, the smallest sampling value to ensure a small shift magnitude of $d = 0.5$ is achieved by MRSS. It is clear that whether or not the process is out-of-control will be determined by taking approximately 138 samples. It is fair to say that the MRSS method is more effective than other methods in the early detection of whether or not the process is an out-of-control when using these magnitudes.

In the event that the process is in control – in other words if the magnitude value d is equal to zero – it is clear that the number of out-of-control signals does not increase when RSS, SRS, ERSS and MRSS methods are used. However, small declines are detected in ARL values.

Table 1 demonstrates that ARL values for $n = 3$ and $p = 3$ using SRS, RSS, ERSS and MRSS methods are 370, 368, 368 and 368, respectively. It also indicates that when the magnitude of the deviation in the mean process is $d = 0.5$ for $n=3$ and $p=3$, the ARL values are equal to 156, 151, 165 and 134, respectively. These numbers are the approximate numbers of samples to be taken for ARL to detect for a small shift magnitude of 0.5. It is clear that the RSS method will detect the out-of-control point faster with fewer samples when compared with SRS. However, it is fair to say that MRSS is the most effective method, using the least number of samples compared to the other methods.

The researcher used the sampling methods in Table 2 to calculate the average number of samples that should be taken to determine whether or not the process is out-of-control at a sampling size of $n = 4$. The ARL values were calculated as well. The researchers took the small shift values into consideration to see the magnitude of the samples that would be created for determining any out-of-control incidence with the assistance of Hotelling's T^2 control chart. Thus, as the table shows, the smallest sampling value to ensure a small shift magnitude of $d = 0.5$ is achieved by MRSS. It is clear that whether or not the process is out-of-control will be determined by taking approximately 138 samples. It is fair to say that the MRSS method is more effective than other methods in the early detection of whether or not the process is an out-of-control when using these magnitudes.

In the event that the process is in control – in other words if the magnitude value d is equal to zero – it is clear that the number of out-of-control signals does not increase when RSS, SRS, ERSS and MRSS methods are used. However, small declines are detected in ARL values.

Table 1 demonstrates that ARL values for $n = 3$ and $p = 3$ using SRS, RSS, ERSS and MRSS methods are 370, 368, 368 and 368, respectively. It also indicates that when the magnitude of the deviation in the mean process is

$d = 0.5$ for $n=3$ and $p=3$, the ARL values are equal to 156, 151, 165 and 134, respectively. These numbers are the approximate numbers of samples to be taken for ARL to detect for a small shift magnitude of 0.5. It is clear that the RSS method will detect the out-of-control point faster with fewer samples when compared with SRS. However, it is fair to say that MRSS is the most effective method, using the least number of samples compared to the other methods.

2.2 Comparing SRS, RSS, ERSS and MRSS Methods Using Hotelling's T^2 Control Charts

Sampling is performed with each method to obtain the values of the variables displaying normal distribution with the simulation application to find the upper control line (UCT) value. Hotelling's T^2 control chart will generate an out-of-control signal if T^2 is greater than UCL. Data may be generated from standard multivariate normal distribution by simulation. Data are ordered after being chosen for the population. Estimated values for the ordered data are recurred 10,000 times, and only the mean value is calculated.

Hotelling's T^2 values are calculated separately for SRS, RSS, ERSS and MRSS methods at the end of this recursion. Following this calculation, the method(s) having the most stable orientation is determined. Hotelling's T^2 control chart construction with $n = 3$ and $n = 4$ sampling size is presented in Charts 1 and 2.
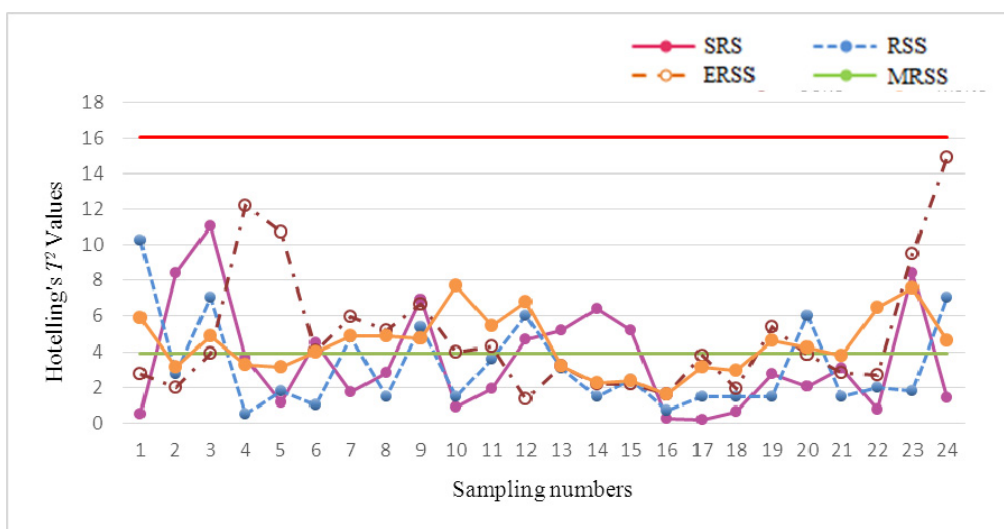

Figure 1. Hotelling's T^2 control chart for SRS, RSS, ERSS and MRSS at $n = 3$

The chart demonstrates that all observation data selected with the four different methods are in the upper control line; in other words, the process is in control.
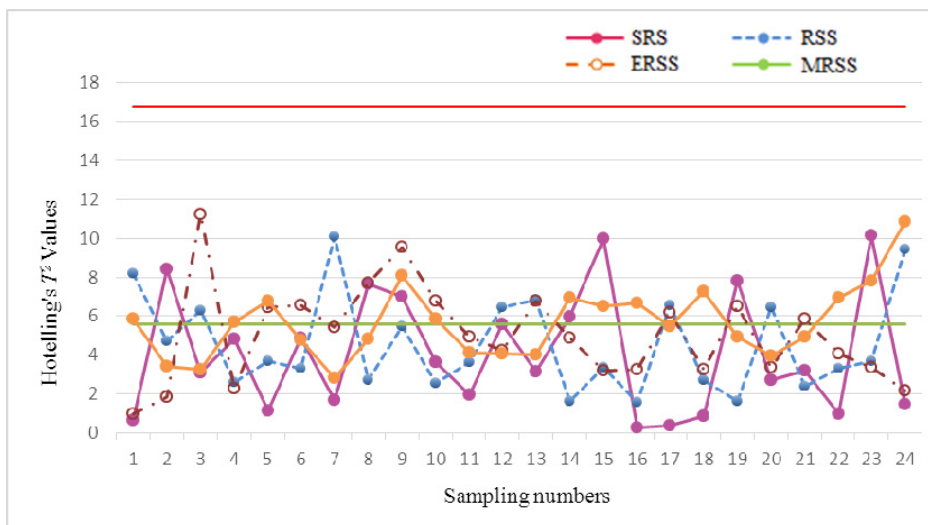

Figure 2. Hotelling's T^2 control chart for SRS, RSS, ERSS and MRSS at $n = 4$

After the upper control line is calculated for the SRS, RSS, ERSS and MRSS methods for these control charts, it becomes clear that the data observed when Hotelling's T^2 control chart is drawn are all in the line for each of the four methods. Although the process appears to be in control, it is seen that observation values obtained with the

SRS method are inclined to send an out-of-control signal. This is due to the fact that the observation values obtained with the SRS method give signals that are close to zero, the lower limit. The reason why the lower line is placed at zero is that an increase is seen in the T^2 value when a shift is present in the process mean, and this causes the lower line to lose its importance. However, the T^2 value is sensitive not only to shifts in the mean vector but also to switching in the covariance matrix. However, if the covariance matrix changes, abnormally small T^2 values may be seen. Thus, the lower line point is set at zero to detect these small switches. But, it is fair to say that MRSS has a more systematic approach compared to the other methods since a couple of consecutive points always display an increase or a decline.

3. Conclusions

Today, multivariate statistical process control methods have become highly important in addition to the sampling methods that would be used to create sample units that would enable determining whether these processes are under control. Sampling was performed using simple random sampling, ranked set sampling, extreme ranked set sampling and median ranked set sampling methods, and Hotelling's T^2 control charts – multivariate statistical process control methods – were created. The performances of the sampling methods were compared to one another using these control charts. Extensive simulation study has shown that multivariate control charts method is superior to Hotelling T^2 method, sampling methods, and multivariate methods in terms of ARL performance, especially when the magnitude of shift is small. The main problem with the multivariate control charts method is, when there is a shift in more than one variable, this method not only does not catch the shift and shows that the process is in control. Also when the magnitude of shift is more than three sigma, this method gives false alarm. And in this study they differ only in how they use of the sampling methods to make a decision about the process. The ranked set sampling has been demonstrated to be an efficient sampling method. The RSS method proved to be more efficient when units are difficult and costly to measure, but are easy and cheap to rank with respect to a variable of interest without actual measurement. In this study, we used Hotelling's T^2 control chart for SRS, RSS, ERSS and MRSS. Data may be generated from standard multivariate normal distribution by simulation. What's more, its modifications to develop several multivariate quality control charts for the variables of interest using the sample mean. It is clear that all the newly developed for multivariate control charts are more efficient than classical control chart, but some of them are better than others.

When a correlation was present between the variables of multivariate processes, out-of-control signals were detected when sampling was performed using RSS and MRSS methods and Hotelling's T^2 control chart was created, but this signal was not detected with other sampling methods. In addition, any shift in the process was detected earlier with the use of MRSS when methods were compared to one another for their ARL values. Therefore, the researcher suggests that the MRSS method should be used to determine the shifts in the process in a timely manner since this method is more effective than the others. It enables calculating the ARL value in advance to see the number of the samples to be formed for the determination of the possible shifts in the process, and making an approximate estimation of the size of these shifts.

The results show that the sampling plan to be applied is a function of the magnitude of the process shift. For small process shifts, we should employ large sample sizes infrequently, and for large process shifts, a small sample size should be taken very frequently. This sampling method will help detect the defective products in time, and minimize the cost and loss for establishments.

References

Aparisi, F. (1996). Hotelling's T^2 Control Chart with Adaptive Sample Sizes. *International Journal of Production Research, 34*(10), 2853-2862. https://doi.org/10.1080/00207549608905062

Aparisi, F., Champ, W. C. ve Diaz, G. C. J. (2004). A Performance Analysis of Hotelling's T^2 Control Chart with Supplementary Runs Rules. *Quality Engineering, 16*(3), 359-368. https://doi.org/10.1081/QEN-120027938

Aparisi, F. (2007). Sampling Plans for the Multivariate T^2 Control Chart. *Quality Engineering, 10*(1), 141-147. https://doi.org/10.1080/08982119708919117

Balakrishnan, N., & Cohen C. (1991). *Order statistics and inference: estimation methods*. San Diego: Academic Press.

Başar, A., & Oktay, E. (2001). *Applied Statistics I,* Erzurum: Aktif Publisher, 10th Edition.

Burnak, N. (1997). *Total Quality Management-Statistical Process Control*. Osmangazi University Ed., Eskişehir.

Champ, W. C., & Aparisi, F. (2008). Double Sampling Hotelling's T^2 Charts. *Quality and Reliability Engineering International, 24,* 153-166. https://doi.org/10.1002/qre.872

Costa, Antonio F. B. (1994). $\overline{X}$ Charts with Wariable Sample Size. *Journal of Quality Technology, 26*(3), 155-163.

Cox, M. A. A. (2001). Towards The Implementation of a Universal Control Chart and Estimation of Its Average Run Length Using a Spreadsheet: An Artificial Neural Network is employed to Model the Parameters in a Special Case. *Journal of Applied Statisitcs, 28*(3), 353-364. https://doi.org/10.1080/02664760120034090

Eygü, H. (2015). Application of Multivariate Statistical Process Control in Cement Industry. *Journal of Business Economics and Political Science, 4*(8), 67-82.

Faraz, A., & Parsian A. (2006). Hotelling's T^2 Control Chart With Double Warning Lines. *Statistical Papers, 47*, 569-593. https://doi.org/10.1007/s00362-006-0307-x

Faraz, A., & Moghadam, M. B. (2009). Hotellin's T^2 Control Chart with Two Adaptive Sample Sizes. *Quality & Quantity 43*(6), 903-913. https://doi.org/10.1007/s11135-008-9167-x

Jabeen, S. (2011). Ranked Set Sampling And its Application in Quality Control, Master of Philosoph, Srinagar: Post-graduate Department of Statistics University of Kashmir, 54-61.

Javaheri, A. ve Houshmand A. A. (2001). Average Run Length Comparison of Multivariate Control Charts. *Journal of Statistical Computation & Simulation, 69,* 125-140. https://doi.org/10.1080/00949650108812086

Işığıçık, E. (2012). *Statistical Quality Control in Terms of Total Quality Management* (2rd ed.). Bursa: Ezgi Bookstore Publications.

Kumpas, S. (2006). *Standard $\overline{X}$ -R Graphics with $\overline{X}$ -R Sampled in Two Stages Evaluation of Graphics Performance with ARL (Average Run Length) Help.* İstanbul Univ. Graduate School of Social Sciences.

Patil, G. P., Sinha, A. K., & Taillie, C. (1999). Ranked set sampling. A Bibliography. *Environmental and Ecological Statistics, 6,* 91-98. https://doi.org/10.1023/A:1009647718555

Pongpullponsak, A., & Sontisamran, P. (2013). Statistical Quality Control Based on Ranked Set Sampling for Multiple Characteristics. *Journal of the Science Faculty of Chiang Mai University, 40*(3), 485-498.

Rakitzis, C. A., & Antzoulakos, L. D. (2011). Chi-square Control Charts with Runs Rules. *Methodol Computing in Applied Probability, 13*(4), 657-669. https://doi.org/10.1007/s11009-010-9178-7

Mason, R. L., Tracy, N. D., & Young, J. C. (1995). Decomposition of for Multivariate Control Chart Interpretation. *Journal of Quality Technology, 27*(2), 99-108.

Mason, R. L., & Young, J. C. (2002). *Multivariate Statistical Process Control with Industrial Applications*, ASA-SIAM Series on Statistics and Applied Probability. https://doi.org/10.1137/1.9780898718461

McIntyre, G. A. (1952). A Method of Unbiased Selective Sampling Using Ranked Sets. *Australian Journal of Agricaltural Research 3*, 384-390. https://doi.org/10.1071/AR9520385

Montgomery, D. C. (2009). *Introduction to Statistical Quality Control* (6th ed.), New York: John Wiley.

Muttlak, H. A. (1997). Median ranked set sampling. *Applied Statistical Science*, 6(4), 245-255.

Muttlak, H. A., & Al-Sabah, W. S. (2003). Statistical Quality Control Based on Ranked Set Sampling. *Journal of Applied Statistics, 30*(9), 1055-1078. https://doi.org/10.1080/0266476032000076173

Samawi, H. M., Ahmed, M. S., & Abu-Dayyeh. (1996). Estimation the Population Mean Using Extreme Ranked Set Sampling. *Biometrical Journal, 38*(5), 577-586. https://doi.org/10.1002/bimj.4710380506

Schwertman, C. N., Ryan p.T. ve Adnan, R. (2002). A Multivariate Control Chart for Expected Changes Over Time. *Quality Engineering, 15*(2), 275-280. https://doi.org/10.1081/QEN-120015859

Woodall, W. H. (1985). The Statistical Design of Quality Control Charts. *Journal of the Royal Statistical Society. Series D (The Statistician), 34*(2), 155-160. https://doi.org/10.2307/2988154

Woodall, W. H., & Montgomery, D. C. (1999). Research İssues and İdeas in Statistical Process Control. *Journal of Quality Technology, 31*(4), 376-386.