



ELSEVIER

journal homepage: [www.elsevier.com/locate/jmatprotec](http://www.elsevier.com/locate/jmatprotec)

# Determination of sampling errors in natural stone plates through Single Linkage Cluster Method

Huseyin Ankara<sup>a</sup>, Suheyla Yerel<sup>b,\*</sup>

<sup>a</sup> Eskisehir Osmangazi University, Department of Mining Engineering, Eskisehir, Turkey

<sup>b</sup> Bilecik University, Bozuyuk Vocational School, Cerrahlar Sokak, 11300 Bozuyuk/Bilecik, Turkey

## ARTICLE INFO

### Article history:

Received 12 April 2007

Received in revised form

16 May 2008

Accepted 28 May 2008

### Keywords:

Natural stone

Range (R) control chart

Single Linkage Cluster Method

Sampling error

## ABSTRACT

In this study, the variations in the thicknesses of the plates cut from natural stone blocks with a circular saw block cutter machine have been investigated using Range (R) control charts. Erratic data caused by sampling errors in range values were encountered when these charts are examined. Erratic data among the range values were determined by Single Linkage Cluster Method and were eliminated from the processes. After the erratic data were sorted out, the parameters concerning the R control charts were recalculated and the charts were reconstructed. Using the reformed R control charts, it was shown that the processes are better than the previous R control charts. Finally, it was determined that the erratic data caused by sampling errors could reliably be estimated by the Single Linkage Cluster Method.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

Statistical process control (SPC) is widely used as a powerful method to measure, classify, analyze and interpret process data to improve the quality of products and service by detecting instabilities and justifying possible causes. SPC tools provide a graphical display of quality characteristics and data series versus the sample number (Fazel Zarandi et al., 2008). The main SPC tool is the control charts (De Vries and Conlin, 2005). Control charts, also known as Shewhart control charts, monitor whether or not a process is under control (Anagun, 1998).

Some studies conducted in earth sciences using control charts are as follows: Bhattacharjee and Samanta (2002) studied the use of Shewhart control charts in mining applications. Their work which considered the Shewhart control charts may still provide useful qualitative information. Ankara et al. (2006) suggested that Shewhart mean control charts could be used in determining the cutting losses in marble blocks. The study

indicates that plate losses can be minimized using control charts. Eleveli (2006) carried out a study in which the variations in the characteristics of coal delivered to a power plant have been investigated using control charts. The results of control charts show that the variations in the characteristics of coal are significantly high. Yerel et al. (2007) determined a preventive maintenance policy for a mineral processing plant using a mean quality control chart and the Kolmogorov–Smirnov test. Their work indicated that a 4-month preventive maintenance policy could be better than preventive maintenance performed once a year.

R control chart is one of the Shewhart control charts (Montgomery, 1996). The erratic data causing sampling errors are encountered while examining the R control chart. This data lead to acquiring incomplete and inaccurate information concerning the process. Therefore the erratic data are determined using the Single Linkage Cluster Method and are eliminated from the process. After, the R control chart is reconstructed. Using the reconstructed chart, the variations

\* Corresponding author.

E-mail address: [syerel@gmail.com](mailto:syerel@gmail.com) (S. Yerel).

0924-0136/\$ – see front matter © 2008 Elsevier B.V. All rights reserved.

doi:10.1016/j.jmatprotec.2008.05.048

in natural stone process are monitored and cutting settings of block cutter machine are evaluated according to this.

In the literature, several studies have reported on the Single Linkage Cluster Method. Yucel and Demir (2004) studied cluster method and reported it to be a useful approach for the characterization of marbles based on chemical properties. The investigation of 25 marbles samples originating from three regions of Turkey involving 10 elements revealed a great variety of distribution of these elements without any apparent connection to their geographic location. Cluster method could be applied to display how samples form groups according to their provenance properties. This result suggests that cluster method is a useful tool in marble industry for primary evaluation category similarity. Muri (2004) has investigated basic physical and chemical characteristics of the water in lakes using cluster method. Boyacioglu and Boyacioglu (2008) studied the cluster method which was applied to assess water quality. This analysis is believed to assist water managers to understand complex nature water quality issues and determine priorities to improve water quality.

The aim of this study was to examine whether or not the thicknesses of the plates cut from natural stone blocks were under control using R control charts. The erratic data caused by sampling errors among the range values were determined using Single Linkage Cluster Method and were eliminated from the process. Then, the R control charts were reconstructed.

## 2. Statistical analysis

### 2.1. R control chart

Grant and Leavenworth (1980) proposed that the following steps are applied for the construction of the R control chart:

- Statistical parameters related to R control chart are computed for each subgroup.
- After the computation of statistical parameters, the ranges of subgroups ( $R_j$ ) and the average of ranges ( $\bar{R}$ ) are calculated.
- The three basic parameters for the chart are calculated. These parameters are: the Central Line ( $CL_R$ ), the Upper Control Limit ( $UCL_R$ ) and the Lower Control Limit ( $LCL_R$ ).
- The R control chart is drawn and is interpreted.

#### 2.1.1. The calculation of $R_j$

$R_j$  is the difference between the largest ( $X_{\max}$ ) and the smallest ( $X_{\min}$ ) observations in a subgroup (Montgomery, 1996).  $R_j$  is presented by Eq. (1):

$$R_j = X_{\max} - X_{\min}, \quad j = 1, 2, 3, \dots, m \quad (1)$$

#### 2.1.2. The calculation of $\bar{R}$

$\bar{R}$  is the sum of ranges of subgroups divided by the number of subgroups in the group ( $m$ ) (Montgomery, 1996).  $\bar{R}$  is equal to Eq. (2):

$$\bar{R} = \frac{\sum_{j=1}^m R_j}{m} \quad (2)$$

#### 2.1.3. The calculation of R control chart parameters

The three basic parameters for the R control chart are computed as Eqs. (3)–(5):

$$UCL_R = D_4 \bar{R} \quad (3)$$

$$CL_R = \bar{R} \quad (4)$$

$$LCL_R = D_3 \bar{R} \quad (5)$$

where  $D_3$  and  $D_4$  are the constants taken from the table of the control charts.

### 2.2. Single Linkage Cluster Method

The steps in the Single Linkage Cluster Method are given as follows (Sharma, 1996):

- The number of clusters and the number of observations are determined.
- The Euclidean distances between the pairs of cluster are calculated.
- $N$ -dimensional symmetric Euclidean Distance Matrix ( $DM_E$ ) is formed from using Euclidean distances.
- The Similarity Matrix ( $SIM_E$ ) is computed using  $DM_E$ .  $SIM_E$  is represented by the similarity values between pairs of cluster.
- Maximum similarity value between all pairs of cluster is selected and this cluster pair is joined to form a new cluster.
- Above steps are repeated  $n - 1$  times, until all clusters are combined in a single cluster.
- After, clusters are represented graphically in what is called a dendrogram.
- The dendrogram is examined and the erratic data caused by sampling errors are determined.

#### 2.2.1. The calculation of the Euclidean distance

The Euclidean distance among the pairs of cluster is calculated as Eq. (6) (Johnson and Wichern, 2002). Then,  $DM_E$  is formed using the Euclidean distances (Everitt, 1993).

$$d_{it} = \sqrt{\sum_{k=1}^n (c_{ik} - c_{tk})^2}, \quad i = 1, 2, 3, \dots, m, \quad t = 1, 2, 3, \dots, m \quad (6)$$

where  $d_{it}$  is the Euclidean distance between  $i$ th and  $t$ th clusters;  $c_{ik}$  is the value of the  $k$ th variable for the  $i$ th cluster;  $c_{tk}$  is the value of the  $k$ th variable for the  $t$ th cluster;  $n$  is the number of variables.

#### 2.2.2. The calculation of the similarity value

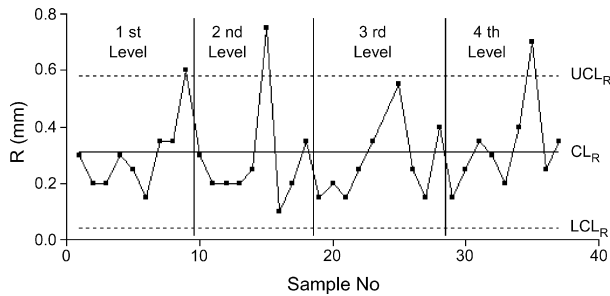
The similarity value ( $Sim(i, t)$ ) is calculated using Eq. (7). Similarity matrix ( $SIM_E$ ) is formed for similarity values (Ozdamar, 2004).

$$Sim(i, t) = 100 \left( 1 - \frac{d_{it}}{\max(d_{it})} \right) \quad (7)$$

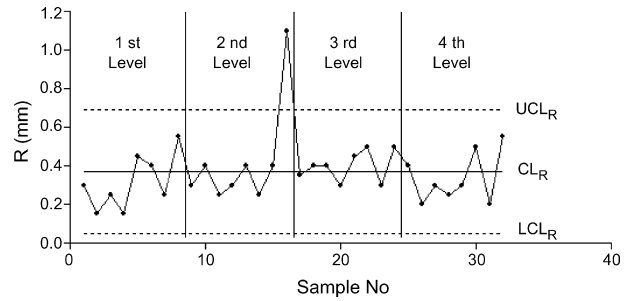
where  $\max(d_{it})$  defines maximum Euclidean distance on  $DM_E$ .

**Table 1 – The R control chart parameters**

Parameters	Mugla White Marble	Denizli Travertine
$\bar{R}$ (mm)	0.32	0.37
$UCL_R$ (mm)	0.59	0.69
$CL_R$ (mm)	0.32	0.37
$LCL_R$ (mm)	0.04	0.05
$D_3$	0.14	0.14
$D_4$	1.86	1.86



**Fig. 1 – The R control chart for Mugla White Marble.**



**Fig. 2 – The R control chart for Denizli Travertine.**

### 3. Result and discussion

#### 3.1. Materials and methods

Plates from natural blocks, cut at four different levels by a circular saw block cutter machine, were used in this study. A total of 36 plates from Mugla White Marble natural stone blocks and 32 plates from Denizli Travertine natural stone blocks were selected for analysis. Afterwards, thickness measurements were taken from eight different points at these plates.

#### 3.2. Construction of R control charts

Using the thickness values obtained by making measurements at eight different points from the plates, the R control chart parameters were calculated and were presented in Table 1.

Then, using the parameters from Table 1, the R control charts were presented in Figs. 1 and 2, for Mugla White Marble and for Denizli Travertine, respectively. It can be seen that, in Fig. 1, the 6th, 7th and 9th points at every level and, in Fig. 2,

the 8th point at every level are not compatible with the other points.

#### 3.3. Determination of sampling errors

The Single Linkage Cluster Method was utilized to evaluate the incompatible points as seen in Figs. 1 and 2. This method accepts the range values in each level that are in the same row as a cluster and evaluates the similarity of a cluster with the other clusters. The number of clusters for Mugla White Marble and Denizli Travertine was selected as 9 and 8, respectively, and the number of observations for these clusters was determined as 4 when using the Single Linkage Cluster Method. Then, for each cluster pair belonging to Mugla White Marble and Denizli Travertine, Euclidean distances were calculated and distance matrices were formed. Using the distance matrices, similarity matrices were constructed. Cluster pairs that have the highest similarity among the similarity matrices were chosen and were joined to form as a new cluster. The steps were repeated eight times for Mugla White Marble and seven times for Denizli Travertine, and all clusters were joined in one cluster (Table 2).

Dendrograms were constructed using the clusters joined versus the similarity values (Figs. 3 and 4). Examining the dendrogram for Mugla White Marble, it was seen that the 6th cluster joins 27.26% with the other clusters, the 7th cluster joins 49.32%, the 9th joins 53.74% and the 5th joins 80.27% (Fig. 3). While examining the dendrogram for Denizli Travertine, it was seen that the 8th cluster joins 22.06% with the other clusters and the 5th joins 76.24% (Fig. 4).

**Table 2 – Similarity values, clusters joined and new cluster for Mugla White Marble and Denizli Travertine**

Mugla White Marble					Denizli Travertine				
Step	Number of clusters	Similarity values (%)	Clusters joined	New cluster	Step	Number of clusters	Similarity values (%)	Clusters joined	New cluster
1	8	86.05	2–3	2	1	7	86.28	1–3	1
2	7	84.72	2–4	2	2	6	85.34	2–7	2
3	6	84.72	2–8	2	3	5	84.45	2–4	2
4	5	81.29	1–2	1	4	4	83.60	1–2	1
5	4	80.27	1–5	1	5	3	78.00	1–6	1
6	3	53.74	1–9	1	6	2	76.24	1–5	1
7	2	49.32	1–7	1	7	1	22.06	1–8	1
8	1	27.26	1–6	1	–	–	–	–	–

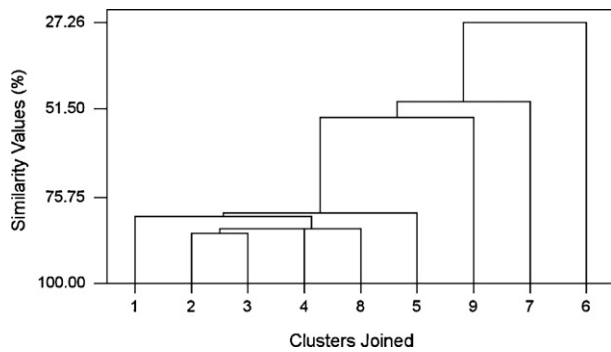


Fig. 3 – Dendrogram for Mugla White Marble.

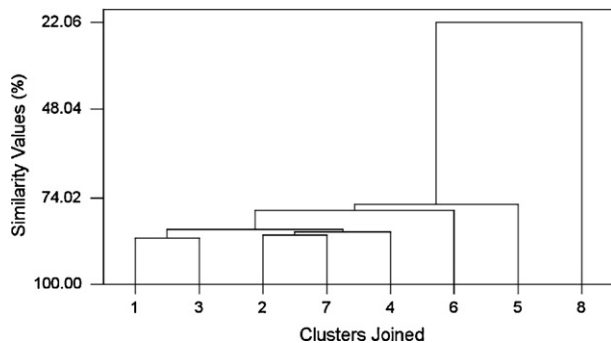


Fig. 4 – Dendrogram for Denizli Travertine.

Comparing Figs. 1 and 3 for Mugla White Marble, it was determined that the 6th, the 7th and the 9th points were erratic data caused by sampling errors. While examining Figs. 2 and 4 for Denizli Travertine, it was seen that 8th point is erratic data caused by sampling errors. These points that were determined using R control charts and dendrograms should be removed from the processes.

### 3.4. Reconstruction of R control charts

After the removal of the erratic data belonging to Mugla White Marble and Denizli Travertine from the processes, the parameters for the R control charts were recalculated and were presented in Table 3.

Using these parameters, the R control charts were reconstructed and could be seen in Figs. 5 and 6. It was determined that using the reconstructed R control charts, compared to using the previous R control charts, the processes belonging to natural stones can be tracked more accurately.

Table 3 – The recalculated R control chart parameters

Parameters	Mugla White Marble	Denizli Travertine
$\bar{R}$ (mm)	0.24	0.33
$UCL_R$ (mm)	0.44	0.61
$CL_R$ (mm)	0.24	0.33
$LCL_R$ (mm)	0.03	0.04
$D_3$	0.14	0.14
$D_4$	1.86	1.86

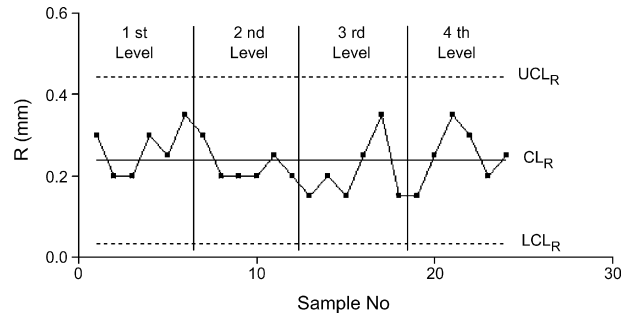


Fig. 5 – The reconstructed R control chart for Mugla White Marble.

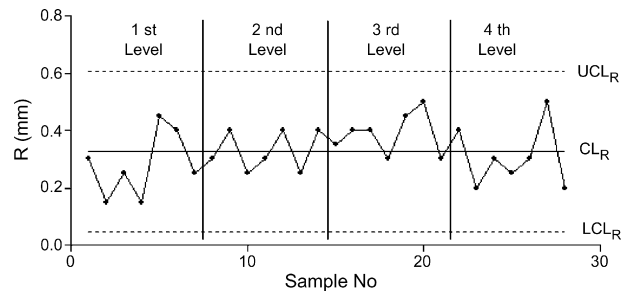


Fig. 6 – The reconstructed R control chart for Denizli Travertine.

## 4. Conclusions

In this study, the R control charts were used for examining the thicknesses of the plates cut with a circular saw block cutter machine.

- Examining the Mugla White Marble and Denizli Travertine for R control charts, 6th, 7th, 9th points at each level and 8th points at each level, respectively, were located in unsuitable places compared to the other values.
- These points were erratic data caused by sampling errors as it was determined by the Single Linkage Cluster Method and so were eliminated from the process.
- After, the R control charts were reconstructed. These charts showed that the processes are better than the previous R control charts.
- It is strongly suggested that the erratic data caused by sampling errors are determined and eliminated in the process by using a single linkage cluster method before the R control charts are constructed.

## REFERENCES

- Anagun, A.S., 1998. A neural network applied to pattern recognition in statistical process control. *Computers & Industrial Engineering* 35 (1–2), 185–188.
- Ankara, H., Yerel, S., Ersoy, A.H., Konuk, A., Yavuz M., 2006. Determination of plate losses in block cutter machine with Shewhart control charts. In: Ersoy M. (Ed.), *Proceedings of MERSEM 2006*, Afyon, Turkey, pp. 421–430.
- Bhattacharjee, A., Samanta, B., 2002. Practical issues in the construction of control charts in mining applications. *The*

- Journal of the South African Institute of Mining and Metallurgy 3, 173–180.
- Boyacioglu, H., Boyacioglu, H., 2008. Water pollution source assessment by multivariate statistical methods in the Tahtali Basin, Turkey. *Environmental Geology* 54, 275–282.
- De Vries, A., Conlin, B.J., 2005. A comparison of the performance of statistical quality control charts in a dairy production system through stochastic simulation. *Agricultural Systems* 85, 317–341.
- Elevli, S., 2006. Coal quality control with control charts. *International Journal of Coal Preparation and Utilization* 26, 181–199.
- Everitt, B.S., 1993. *Cluster Analysis*, third ed. John Wiley & Sons, Inc., New York.
- Fazel Zarandi, M.H., Alaeddini, A., Turksen, I.B., 2008. A hybrid fuzzy adaptive sampling—run rules for Shewhart control charts. *Information Sciences* 178, 1152–1170.
- Grant, E.L., Leavenworth, R.S., 1980. *Statistical Quality Control*, fifth ed. McGraw-Hill International Book Co.
- Johnson, R.A., Wichern, D.W., 2002. *Applied Multivariate Statistical Analysis*, fifth ed. Pearson Education International, pp. 668–745.
- Montgomery, D.C., 1996. *Introduction to Statistical Quality Control*, third ed. John Wiley & Sons, Inc., New York, pp. 179–230.
- Muri, G., 2004. Physico-chemical characteristics of lake water in 14 Slovenian Mountain Lakes. *Acta Chimica Slovenica* 51, 257–272.
- Ozdamar, K., 2004. *Paket Programlar ile Istatistiksel Veri Analizi 2 (Cok Degiskenli Analizler)*, fifth ed. Kaan Press, Turkey, pp. 279–323.
- Sharma, S., 1996. *Applied Multivariate Techniques*. John Wiley & Sons, Inc., New York, pp. 188–193.
- Yerel, S., Ankara, H., Konuk, A., Ozdag, H., 2007. Preventive maintenance policy supported by a quality control chart and Kolmogorov–Smirnov tests: Emet Colemanite Mineral Processing Plant, Turkey. *Minerals & Metallurgical Processing* 24 (3), 152–156.
- Yucel, Y., Demir, C., 2004. Principal component analysis and cluster analysis for the characterization of marbles by capillary electrophoresis. *Talanta* 63, 451–459.