# INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

# SOME ALTERNATIVE METHODS FOR MONITORING MULTIPLE-STREAM PROCESSES

by

Jeff P. Colbeck

A Practicum
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Statistics
University of Manitoba
Winnipeg, Manitoba

(c) April, 1999

National Library of Canada

Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services

Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-41655-0

Canada

**THE UNIVERSITY OF MANITOBA**

**FACULTY OF GRADUATE STUDIES**
*****
**COPYRIGHT PERMISSION PAGE**

# SOME ALTERNATIVE METHODS FOR MONITORING MULTIPLE-STREAM PROCESSES

**BY**

**JEFF P. COLBECK**

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree**

**of**

**Master of Science**

**Jeff P. Colbeck©1999**

**Permission has been granted to the Library of The University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to Dissertations Abstracts International to publish an abstract of this thesis/practicum.**

**The author reserves other publication rights, and neither this thesis/practicum nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.**

## TABLE OF CONTENTS

# Abstract

An example of a "multiple-stream process" is a multiple-head machine which fills bottles to a common level. Instead of setting up a control chart for each head, it is convenient to monitor all of the heads simultaneously on a single control chart.

Various methods mentioned in the literature about this topic are discussed. The problem with the conventional method is discussed and two new methods which can be used to monitor multiple-stream processes are proposed.

Four methods are discussed in detail. The first method is based on the distribution of the maximum and minimum values in a sample selected from the process. The second method uses correction factors to widen the control limits. The third method is the conventional technique used to monitor these types of processes. The fourth method is an empirical method based upon the results of a Monte Carlo computer simulation.

The average run lengths of the four methods are compared and the methods are applied to real-life data sets to see how they perform in practice.

# Acknowledgments

I would like to thank my advisor Dr. Brian Macpherson for suggesting this topic and for his help. I appreciate all of the suggestions that he made. I would like to also thank Hansheng Xie for helping me with the computer simulations.

# 1. Introduction

Control charts are used to monitor a process over time and to quickly detect when a process goes out of statistical control. Usually, control charts monitor only a single stream of output. The thickness of a circuit board is one example where a control chart monitors a single variable. However, there are times when a process has several streams of output. Many industrial processes have multiple-head machines which, for example, fill bottles to a desired level. Each head could be thought of as a stream. Another example is a manufacturing process in which measurements are taken at various specific locations on a single part. In this context, the individual locations are the streams. For these types of situations, several possible control chart procedures may be followed. One method is to use a separate control chart for each stream. The problem with this approach is that it is often not practical because of the large number of control charts that may be involved. Another approach is to use a single control chart to monitor all of the streams simultaneously.

This practicum summarizes many of the methods mentioned in the literature used to monitor processes with multiple streams, as well as discussing two new methods. One of the new methods is based on the distribution of the maximum and minimum values in a sample selected from the process. The second is an empirical method based upon the results of a Monte Carlo computer simulation.

Performance results will be provided for several different methods used in monitoring multiple-stream processes. These methods will be applied to real-life data sets to observe their performances in practice.

# 2. Alternative Methods for Monitoring Multiple-Stream Processes

## 2.1 Group Control Chart

The most common method used to monitor processes with several streams of output is the group control chart. This technique is discussed by both Nelson (1986) and Montgomery (1996).

The procedure has two objectives. The first is to detect when the output of one stream has shifted away from the target value. The second is to detect when the outputs of all the streams have shifted away from the target value.

Sampling is done just as if separate control charts were set up on each stream. Usually, subgroups of size one to five are chosen from each stream. Suppose that we are dealing with a process with six streams and we chose a subgroup size of three. That means three items will be taken from each of the six streams at each time period. If we further suppose that there are 20 time periods, we would therefore have 20 samples of size three for each of the six streams (see Table 1).

The averages and ranges are calculated for all of the subgroups. For our example, there would be a total of 120 averages and 120 ranges. The grand average $\overline{\overline{X}}$ and the average range $\overline{R}$ are calculated as well. We then can construct the control limits for the group control chart.

The upper and lower control limits for the $\overline{X}$ chart are

$$UCL = \overline{\overline{X}} + A_2\overline{R} = \overline{\overline{X}} + \frac{3}{\sqrt{n}}\left(\frac{\overline{R}}{d_2}\right) \tag{1a}$$

$$LCL = \overline{\overline{X}} - A_2\overline{R} = \overline{\overline{X}} - \frac{3}{\sqrt{n}}\left(\frac{\overline{R}}{d_2}\right) \tag{1b}$$

Table 1

Group Control Chart Example

| Sample or Time Period | Stream 1 | | | ••••••••••• | Stream 6 | | |
|---|---|---|---|---|---|---|---|
| | Value | Subgroup Average | Subgroup Range | | Value | Subgroup Average | Subgroup Range |
| 1 | $X_{111}$ | | | | $X_{161}$ | | |
| | $X_{112}$ | $\overline{X}_{11}$ | $R_{11}$ | ••••••••••• | $X_{162}$ | $\overline{X}_{16}$ | $R_{16}$ |
| | $X_{113}$ | | | | $X_{163}$ | | |
| 2 | $X_{211}$ | | | | $X_{261}$ | | |
| | $X_{212}$ | $\overline{X}_{21}$ | $R_{21}$ | ••••••••••• | $X_{262}$ | $\overline{X}_{26}$ | $R_{26}$ |
| | $X_{213}$ | | | | $X_{263}$ | | |
| ⋮ | | ⋮ | | | | ⋮ | |
| 20 | $X_{2011}$ | | | | $X_{2061}$ | | |
| | $X_{2012}$ | $\overline{X}_{201}$ | $R_{201}$ | ••••••••••• | $X_{2062}$ | $\overline{X}_{206}$ | $R_{206}$ |
| | $X_{2013}$ | | | | $X_{2063}$ | | |

Grand Average: $\overline{\overline{X}} = \dfrac{\sum_{i=1}^{20}\sum_{j=1}^{6}\overline{X}_{ij}}{120}$   Average Range: $\overline{R} = \dfrac{\sum_{i=1}^{20}\sum_{j=1}^{6}R_{ij}}{120}$

First number of subscript denotes sample number.
Second number of subscript denotes stream number.
Third number of subscript (if applicable) denotes observation number.

where $\frac{\bar{R}}{d_2}$ is the unbiased estimate of the process standard deviation $\sigma$.

The upper and lower control limits for the R chart are

$$UCL = D_4\bar{R} \tag{2a}$$

$$LCL = D_3\bar{R} \tag{2b}$$

The centre lines are at $\bar{\bar{X}}$ for the $\bar{X}$ chart and $\bar{R}$ for the R chart. The control limits given are three standard deviations above and below the centre lines. The control chart constants $A_2$, $d_2$, $D_3$, and $D_4$ are determined from the subgroup size (n=3 for the example), not the number of streams (six for the example). Therefore, the appropriate values in this situation are $A_2$=1.023, $d_2$=1.693, $D_3$=0, and $D_4$=2.574 (see Appendix A).

It is only necessary to plot the highest and lowest mean of each sample on the $\bar{X}$ chart. This is because if these points lie within the control limits, so will all of the other points. Similarly, for our example, only the highest range for each sample is plotted on the range chart. The chart signals that the process may be out of control if a maximum (or minimum) point is above the upper (or below the lower) control limit on the $\bar{X}$ chart. The process may also be out of control if the largest range plots above the upper control limit of the R chart.

If a particular stream gives the largest (or smallest) value several times in a row, this is an additional characteristic that may provide sufficient evidence to conclude that the stream is performing differently from the others.

Let *s* be the number of streams and let *r* be the number of consecutive times that a particular stream gives the largest (or smallest) value. If *r* is sufficiently large such that it exceeds some critical value, we may conclude that there is evidence that the stream is performing differently from the rest.

One way to determine the critical value for *r* is to evaluate the average run length using various values of *r* as a control chart "out-of-control" signal. The average run length (ARL) is the number of samples that it would require, on average, to produce a signal on the control chart. There are two types of average run lengths. When the process is in control, the average number of points occurring before a false out-of-control signal is called the in-control ARL ($ARL_0$). Since we wish to have as few false alarms as possible, it is desirable to have the in-control ARL as high as possible. The out-of-control average run length ($ARL_1$) is the average number of points between the time an assignable cause occurs and the time an out-of-control signal occurs. Since we want to detect out-of-control conditions quickly, the out-of-control ARL should be as small as possible.

In this situation, the average number of samples to produce *r* consecutive maximum (or minimum) values from a particular stream when the process is in a state of statistical control is given by

$$ARL_0 = \frac{s^r - 1}{s - 1} \qquad (3)$$

assuming all streams are identical (Montgomery (1996)).

The ARL depends upon both the number of streams being simultaneously monitored and the choice of the value for *r*. Table 2 gives the $ARL_0$ values for a selection of pairings of *s* and *r*. These pairings are constructed so that the one-sided in-control ARL is roughly consistent with the in-control ARL of a conventional Shewhart control chart. The Shewhart control chart $ARL_0$ is approximately 370 using 3-sigma control limits. Since there are six streams in our example, Table 2 tells us that a stream producing the largest (or smallest) value five times in a row gives strong evidence that

Table 2

Recommended Pairings for s and r

| s | r | One-sided In-control ARL |
|---|---|---|
| 2 | 9 | 511 |
| 3 | 7 | 1093 |
| 4 | 6 | 1365 |
| 5 | 5 | 781 |
| 6 | 5 | 1555 |
| 7 | 4 | 400 |
| 8 | 4 | 585 |
| 9 | 4 | 820 |
| 10 | 4 | 1111 |

The number of streams is denoted by s. The number of consecutive times that a particular stream is the largest/smallest value is denoted by r.

the stream is different from the others. This is because a stream producing an extreme value five consecutive times only occurs by chance once every 1555 samples (on average).

## 2.2 Plotting the Raw Data/Analysis of Residuals/Analysis of Variance

Ott and Snee (1973) discuss three techniques which can be used to analyze multiple-stream data. The example given in their article involves a machine with 24 heads. The heads are divided into four groups of six heads each and only one group is examined at a particular time. The raw data consist of five samples from each of the groups taken at 15-minute intervals and are shown in Table 3.

### 2.2.1 Plotting the Raw Data

The first method discussed is to plot the 24 head averages (Figure 1a) and the 20 sample or time period averages (Figure 1b). We see from Figure 1a that the averages for Heads 23, 24, 1, 2, 3, 4, 5, and 6 all fall below the median and that the averages for Heads 7 to 13 inclusive are all above the median. A run in this case is defined as several points in a row all falling above or below the median. We see two long runs in Figure 1a, one is of length eight and the other is of length seven. The run of length eight occurs because sampling is done sequentially across the heads so that Head 1 is sampled after Head 24 in this sampling scheme. The two long runs indicate significant differences in the filling levels of the heads because the longest expected run in 24 observations is five with $\alpha = 0.01$. We can therefore conclude that the heads are not all filling to the same level.

We also see a long run in the sample (time) averages plot (Figure 1b). Seven points plot above the median (samples 10 to 16). This shows significant differences between time periods since a run of length seven is statistically significant at $\alpha = 0.01$.

### 2.2.2 Analysis of Residuals

Ott and Snee (1973) describe a method which is based upon the calculation of the row or time semi-residuals. Using the data in Table 3, the semi-residuals are found by subtracting the average value of each row from

Table 3

Raw Data of the Ott & Snee Example

Group I:

| Time | Sampling Time Order | Head #1 | Head #2 | Head #3 | Head #4 | Head #5 | Head #6 | $\bar{X}_{sample}$ |
|---|---|---|---|---|---|---|---|---|
| 9:00 a.m. | 1 | 1211 | 1222 | 1225 | 1207 | 1230 | 1229 | 1220.7 |
| 12:40 p.m. | 5 | 1229 | 1233 | 1222 | 1211 | 1242 | 1236 | 1228.8 |
| 1:40 p.m. | 9 | 1210 | 1227 | 1228 | 1235 | 1234 | 1228 | 1227.0 |
| 2:40 p.m. | 13 | 1241 | 1246 | 1247 | 1239 | 1254 | 1227 | 1242.3 |
| 3:40 p.m. | 17 | 1238 | 1239 | 1236 | 1238 | 1222 | 1245 | 1236.3 |
| | $\bar{X}_{head}$ = | 1225.8 | 1233.4 | 1231.6 | 1226.0 | 1236.4 | 1233.0 | $\bar{\bar{X}}$=1231.0 |

Group II:

| Time | Sampling Time Order | Head #7 | Head #8 | Head #9 | Head #10 | Head #11 | Head #12 | $\bar{X}_{sample}$ |
|---|---|---|---|---|---|---|---|---|
| 9:15 a.m. | 2 | 1268 | 1265 | 1275 | 1257 | 1232 | 1270 | 1261.2 |
| 12:55 p.m. | 6 | 1256 | 1252 | 1255 | 1248 | 1265 | 1247 | 1253.8 |
| 1:55 p.m. | 10 | 1240 | 1251 | 1252 | 1236 | 1249 | 1245 | 1245.5 |
| 2:55 p.m. | 14 | 1284 | 1287 | 1288 | 1273 | 1234 | 1270 | 1272.7 |
| 3:55 p.m. | 18 | 1250 | 1252 | 1252 | 1250 | 1245 | 1261 | 1251.7 |
| | $\bar{X}_{head}$ = | 1259.8 | 1261.4 | 1264.4 | 1252.8 | 1245.0 | 1258.6 | $\bar{\bar{X}}$=1257.0 |

Group III:

| Time | Sampling Time Order | Head #13 | Head #14 | Head #15 | Head #16 | Head #17 | Head #18 | $\bar{X}_{sample}$ |
|---|---|---|---|---|---|---|---|---|
| 9:30 a.m. | 3 | 1259 | 1232 | 1232 | 1239 | 1229 | 1229 | 1236.7 |
| 1:10 p.m. | 7 | 1252 | 1234 | 1235 | 1230 | 1233 | 1239 | 1237.2 |
| 2:10 p.m. | 11 | 1256 | 1240 | 1244 | 1246 | 1232 | 1235 | 1242.2 |
| 3:10 p.m. | 15 | 1261 | 1249 | 1255 | 1260 | 1260 | 1240 | 1254.2 |
| 4:10 p.m. | 19 | 1253 | 1224 | 1243 | 1246 | 1232 | 1232 | 1238.3 |
| | $\bar{X}_{head}$ = | 1256.2 | 1235.8 | 1241.8 | 1244.2 | 1237.2 | 1235.0 | $\bar{\bar{X}}$=1241.7 |

Group IV:

| Time | Sampling Time Order | Head #19 | Head #20 | Head #21 | Head #22 | Head #23 | Head #24 | $\bar{X}_{sample}$ |
|---|---|---|---|---|---|---|---|---|
| 9:45 a.m. | 4 | 1221 | 1226 | 1241 | 1220 | 1218 | 1229 | 1225.8 |
| 1:25 p.m. | 8 | 1235 | 1216 | 1239 | 1230 | 1219 | 1220 | 1226.5 |
| 2:25 p.m. | 12 | 1259 | 1250 | 1268 | 1255 | 1247 | 1240 | 1253.2 |
| 3:25 p.m. | 16 | 1253 | 1247 | 1263 | 1249 | 1240 | 1254 | 1251.0 |
| 4:25 p.m. | 20 | 1241 | 1228 | 1247 | 1238 | 1227 | 1226 | 1234.5 |
| | $\bar{X}_{head}$ = | 1241.8 | 1233.4 | 1251.6 | 1238.4 | 1230.2 | 1233.8 | $\bar{\bar{X}}$=1238.2 |

Figure 1a: Average Filling Weight by Head

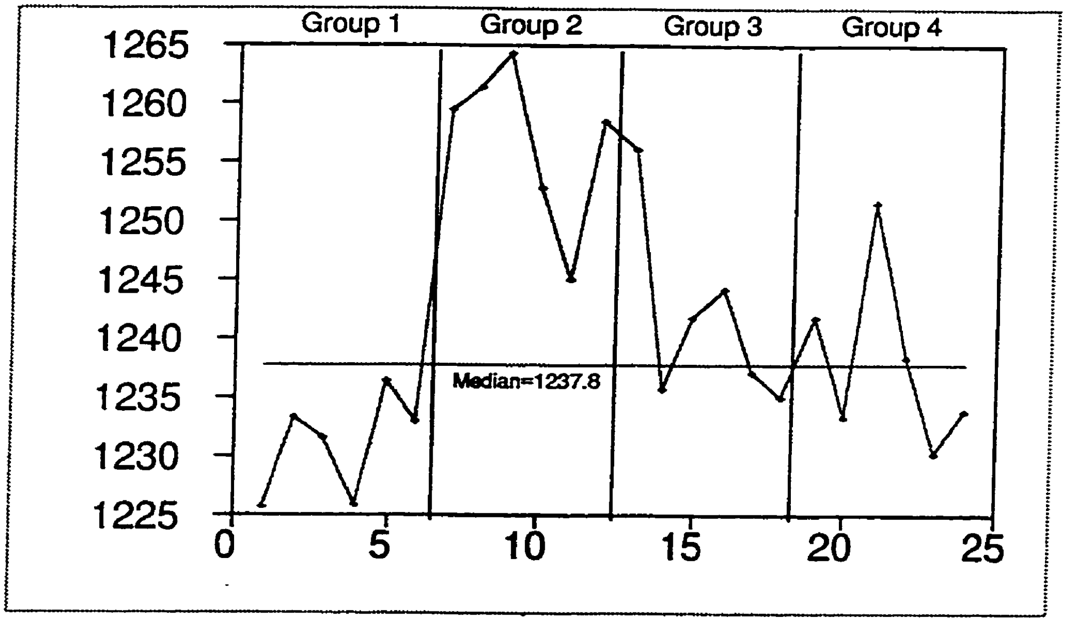
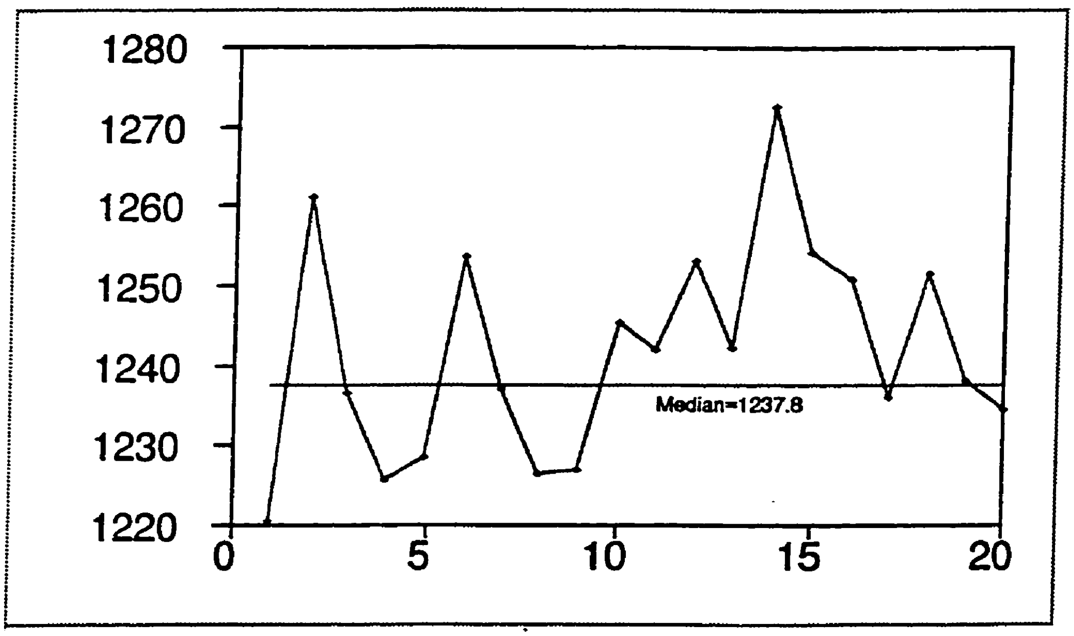



Figure 1b: Average Filling Weight by Sample



each observation in the row. These values are called semi-residuals because they remove the effect of the sampling time, while leaving the relationship between the heads unchanged.

Using the semi-residuals, the averages and ranges for each of the columns or heads are then calculated. The results are shown in Table 4. Plots of the averages and ranges are shown in Figures 2a and 2b.

Heads 11, 13, and 21 go beyond the control limits of the $\overline{X}$ chart (Figure 2a). This tells us that these heads perform differently from their neighbours. Head 11 also exceeds the upper control limit in the range chart (Figure 2b) giving further evidence that this head should be investigated.
A second set of semi-residuals is obtained from Table 3 by subtracting the average of each column (head) from each observation in the column. The averages and ranges for each time period are then calculated. This is shown in Table 5 and plotted in Figures 3a and 3b.

From the $\overline{X}$ chart (Figure 3a), we note that a big jump occurs at time period 12. Ott and Snee mention that a possible explanation for the jump is that the machine was adjusted at that time. We see from the range chart (Figure 3b) that no points exceed the upper control limit.

The analysis of residuals method assumes that the residuals are normally distributed and the machine variability is homogeneous for all heads and time periods.

2.2.3 Analysis of Variance

The third method discussed is an analysis of variance for each group and for all groups combined. The assumptions for this method are the same as for the analysis of residuals method (normality and homogeneous variance).

Table 4

Row (Time) Semi-residuals

Group I:

| | Head #1 | Head #2 | Head #3 | Head #4 | Head #5 | Head #6 |
|---|---|---|---|---|---|---|
| | -9.7 | 1.3 | 4.3 | -13.7 | 9.3 | 8.3 |
| | 0.2 | 4.2 | -6.8 | -17.8 | 13.2 | 7.2 |
| | -17.0 | 0.0 | 1.0 | 8.0 | 7.0 | 1.0 |
| | -1.3 | 3.7 | 4.7 | -3.3 | 11.7 | -15.3 |
| | 1.7 | 2.7 | -0.3 | 1.7 | -14.3 | 8.7 |
| $\bar{X}$ = | -5.2 | 2.4 | 0.6 | -5.0 | 5.4 | 2.0 |
| R = | 18.7 | 4.2 | 11.5 | 25.8 | 27.5 | 24.0 |

Group II:

| | Head #7 | Head #8 | Head #9 | Head #10 | Head #11 | Head #12 |
|---|---|---|---|---|---|---|
| | 6.8 | 3.8 | 13.8 | -4.2 | -29.2 | 8.8 |
| | 2.2 | -1.8 | 1.2 | -5.8 | 11.2 | -6.8 |
| | -5.5 | 5.5 | 6.5 | -9.5 | 3.5 | -0.5 |
| | 11.3 | 14.3 | 15.3 | 0.3 | -38.7 | -2.7 |
| | -1.7 | 0.3 | 0.3 | -1.7 | -6.7 | 9.3 |
| $\bar{X}$ = | 2.6 | 4.4 | 7.4 | -4.2 | -12.0 | 1.6 |
| R = | 16.8 | 16.1 | 15.0 | 9.8 | 49.9 | 16.1 |

Group III:

| | Head #13 | Head #14 | Head #15 | Head #16 | Head #17 | Head #18 |
|---|---|---|---|---|---|---|
| | 22.3 | -4.7 | -4.7 | 2.3 | -7.7 | -7.7 |
| | 14.8 | -3.2 | -2.2 | -7.2 | -4.2 | 1.8 |
| | 13.8 | -2.2 | 1.8 | 3.8 | -10.2 | -7.2 |
| | 6.8 | -5.2 | 0.8 | 5.8 | 5.8 | -14.2 |
| | 14.7 | -14.3 | 4.7 | 7.7 | -6.3 | -6.3 |
| $\bar{X}$ = | 14.5 | -5.9 | 0.1 | 2.5 | -4.5 | -6.7 |
| R = | 15.5 | 12.1 | 9.4 | 14.9 | 16.0 | 16.0 |

Group IV:

| | Head #19 | Head #20 | Head #21 | Head #22 | Head #23 | Head #24 |
|---|---|---|---|---|---|---|
| | -4.8 | 0.2 | 15.2 | -5.8 | -7.8 | 3.2 |
| | 8.5 | -10.5 | 12.5 | 3.5 | -7.5 | -6.5 |
| | 5.8 | -3.2 | 14.8 | 1.8 | -6.2 | -13.2 |
| | 2.0 | -4.0 | 12.0 | -2.0 | -11.0 | 3.0 |
| | 6.5 | -6.5 | 12.5 | 3.5 | -7.5 | -8.5 |
| $\bar{X}$ = | 3.6 | -4.8 | 13.4 | 0.2 | -8.0 | -4.4 |
| R = | 13.3 | 10.7 | 3.2 | 9.3 | 4.8 | 16.4 |

$\bar{R}$=15.71

Figure 2a: $\overline{X}$ Chart of the Column (Head) Semi-residuals

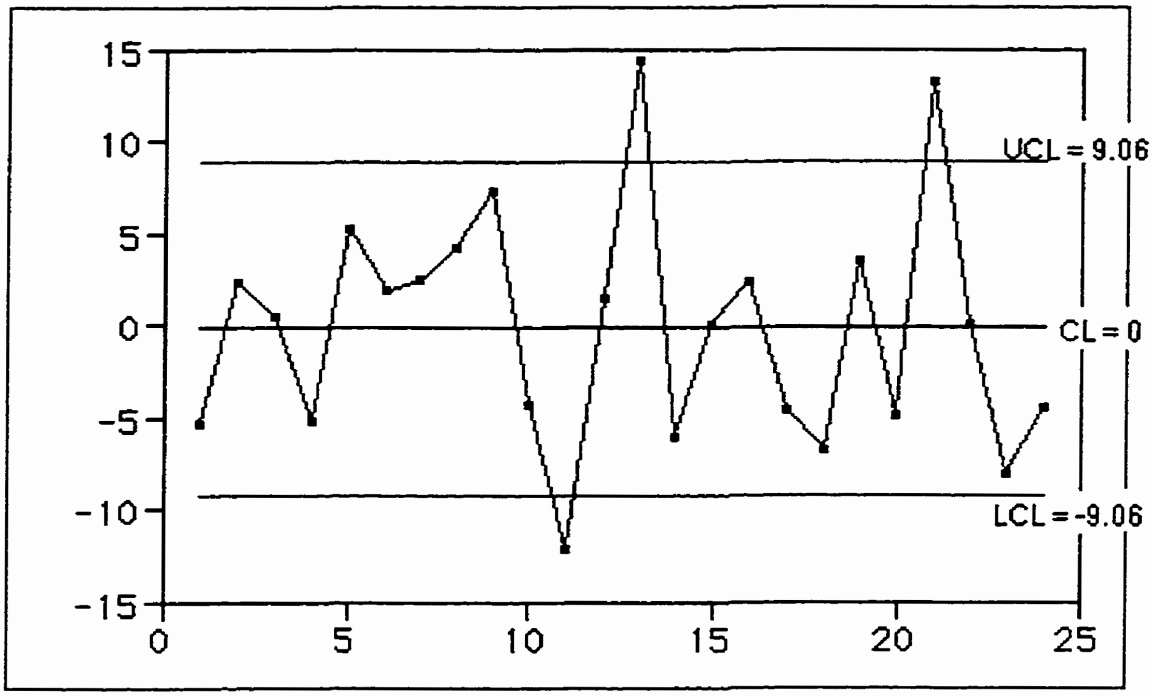



Figure 2b: Range Chart of the Column (Head) Semi-residuals

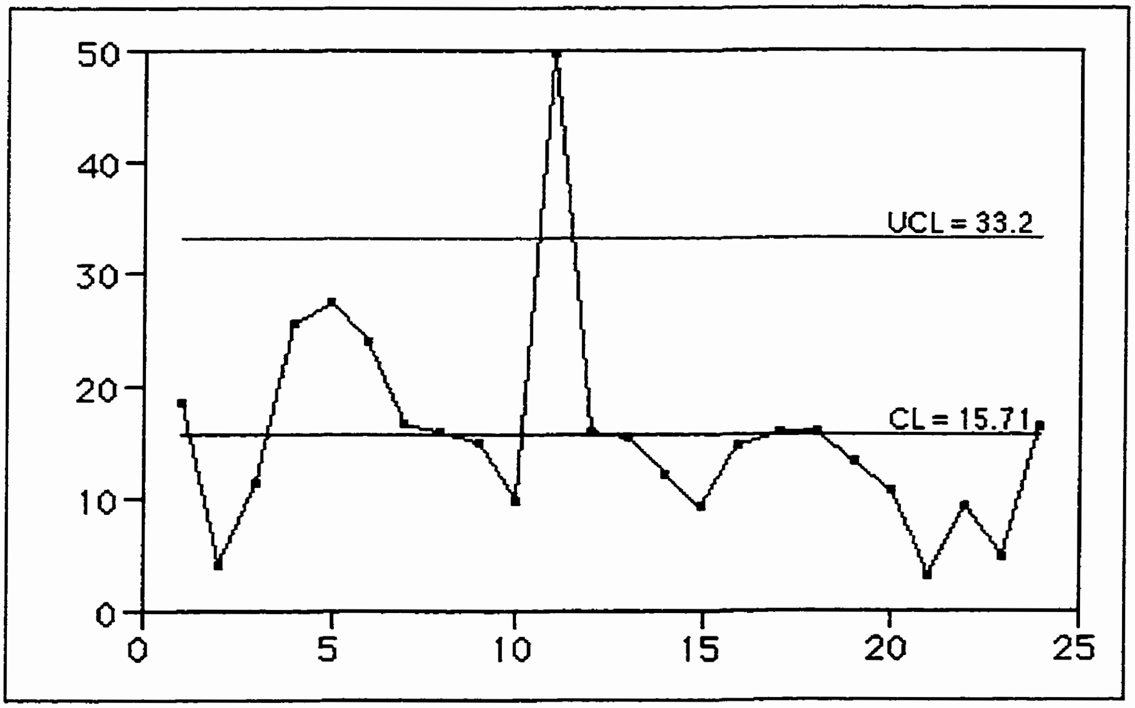


Table 5

Column (Head) Semi-residuals

Group I:

| Time Sequence | Head #1 | Head #2 | Head #3 | Head #4 | Head #5 | Head #6 | $\bar{X}$ | R |
|---|---|---|---|---|---|---|---|---|
| 1 | -14.8 | -11.4 | -6.6 | -19.0 | -6.4 | -4.0 | -10.4 | 15.0 |
| 5 | 3.2 | -0.4 | -9.6 | -15.0 | 5.6 | 3.0 | -2.2 | 20.6 |
| 9 | -15.8 | -6.4 | -3.6 | 9.0 | -2.4 | -5.0 | -4.0 | 24.8 |
| 13 | 15.2 | 12.6 | 15.4 | 13.0 | 17.6 | -6.0 | 11.3 | 23.6 |
| 17 | 12.2 | 5.6 | 4.4 | 12.0 | -14.4 | 12.0 | 5.3 | 26.6 |

Group II:

| Time Sequence | Head #7 | Head #8 | Head #9 | Head #10 | Head #11 | Head #12 | $\bar{X}$ | R |
|---|---|---|---|---|---|---|---|---|
| 2 | 8.4 | 4.6 | 10.6 | 4.2 | -13.0 | 11.4 | 4.4 | 23.6 |
| 6 | -3.6 | -9.4 | -9.4 | -4.8 | 20.0 | -11.6 | -3.1 | 31.6 |
| 10 | -19.6 | -10.4 | -12.4 | -16.8 | 4.0 | -13.6 | -11.5 | 23.6 |
| 14 | 24.4 | 25.6 | 23.6 | 20.2 | -11.0 | 11.4 | 15.7 | 36.6 |
| 18 | -9.6 | -9.4 | -12.4 | -2.8 | 0.0 | 2.4 | -5.3 | 14.8 |

Group III:

| Time Sequence | Head #13 | Head #14 | Head #15 | Head #16 | Head #17 | Head #18 | $\bar{X}$ | R |
|---|---|---|---|---|---|---|---|---|
| 3 | 2.8 | -3.8 | -9.8 | -5.2 | -8.2 | -6.0 | -5.0 | 12.6 |
| 7 | -4.2 | -1.8 | -6.8 | -14.2 | -4.2 | 4.0 | -4.5 | 18.2 |
| 11 | -0.2 | 4.2 | 2.2 | 1.8 | -5.2 | 0.0 | 0.5 | 9.4 |
| 15 | 4.8 | 13.2 | 13.2 | 15.8 | 22.8 | 5.0 | 12.5 | 18.0 |
| 19 | -3.2 | -11.8 | 1.2 | 1.8 | -5.2 | -3.0 | -3.4 | 13.6 |

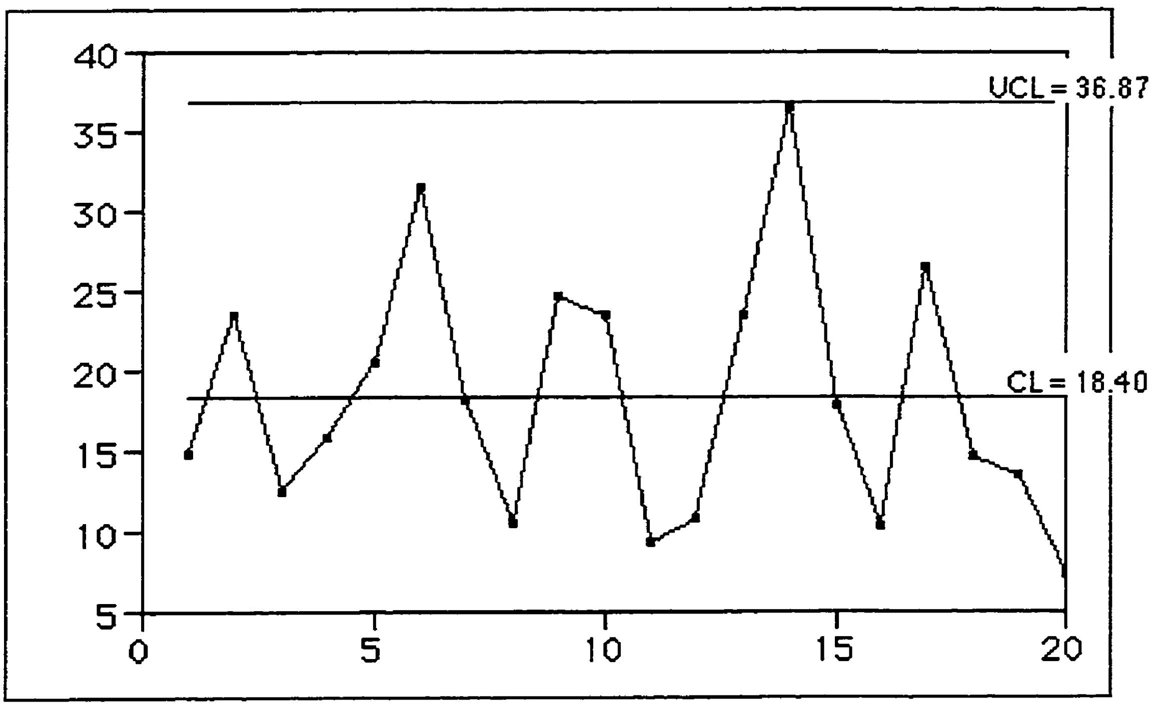
Group IV:

| Time Sequence | Head #19 | Head #20 | Head #21 | Head #22 | Head #23 | Head #24 | $\bar{X}$ | R |
|---|---|---|---|---|---|---|---|---|
| 4 | -20.8 | -7.4 | -10.6 | -18.4 | -12.2 | -4.8 | -12.4 | 16.0 |
| 8 | -6.8 | -17.4 | -12.6 | -8.4 | -11.2 | -13.8 | -11.7 | 10.6 |
| 12 | 17.2 | 16.6 | 16.4 | 16.6 | 16.8 | 6.2 | 15.0 | 11.0 |
| 16 | 11.2 | 13.6 | 11.4 | 10.6 | 9.8 | 20.2 | 12.8 | 10.4 |
| 20 | -0.8 | -5.4 | -4.6 | -0.4 | -3.2 | -7.8 | -3.7 | 7.4 |

$\bar{R} = 18.40$

Figure 3a: $\overline{X}$ Chart of the Row (Time Period) Semi-residuals


Figure 3b: Range Chart of the Row (Time Period) Semi-residuals



The ANOVAs are shown in Table 6. We see that there are significant differences among the heads in Groups III and IV. We also see that there are significant differences among the sampling times in all groups.

## 2.3 Mortell and Runger's $Y_t$ and $R_t$

Mortell and Runger (1995) have also developed a control scheme that monitors a change in all streams and a scheme that monitors a change in one stream relative to the others.

In order to detect a change affecting all streams, Mortell and Runger suggest using the average of the subgroup means across all of the streams for each time period. The plotted statistic for this method is called $Y_t$. The standard Shewhart control chart is used with the mean of the $Y_t$'s as the centre line. This technique improves the ability of the control chart to detect an assignable cause shifting the mean of all streams simultaneously.

The method suggested by Mortell and Runger to detect a change in one stream relative to the others is the use of the range of the streams at each time period. This control variable, called $R_t$, is calculated by subtracting the minimum subgroup mean from the maximum subgroup mean across all of the streams for each time period. The reasoning behind this method is that only changes in one stream with respect to the others are monitored, not changes in the process as a whole.

## 2.4 Fractional Sample Adaptive Monitoring Techniques

Most of the monitoring methods for multiple stream processes generally require that all streams be sampled at a given point in time. However, it is not always possible to collect samples from each stream at every time period in

Table 6

Analysis of Variance for Each Group of Heads and All Heads Combined

| | | Group I | | Group II | | Group III | | Group IV | |
|---|---|---|---|---|---|---|---|---|---|
| Source | df | MS | F | MS | F | MS | F | MS | F |
| Sampling Times (ST) | 4 | 426.5 | 4.95* | 650.3 | 5.00* | 319.3 | 9.36* | 1037.0 | 43.94* |
| Heads (H) | 5 | 91.3 | 1.06 | 245.1 | 1.88 | 316.5 | 9.28* | 290.0 | 12.29* |
| ST x H | 20 | 86.2 | | 130.1 | | 34.1 | | 23.6 | |
| Total | 29 | | | | | | | | |

| | Combined | | |
|---|---|---|---|
| Source | df | MS | F |
| Sampling Times (ST) | 4 | 1425.8 | 13.79* |
| Heads (H) | 23 | 674.9 | 6.53* |
| ST x H | 92 | 103.4 | |
| Total | 119 | | |

*Significant at $\alpha = 0.01$

processes that have a very large number of streams.

Three adaptive monitoring techniques are discussed by Lanning, Montgomery, and Runger (1998) in which only a fraction of the total streams are required to be sampled. They are called the variable sample interval (VSI), variable sample size (VSS), and variable sample size and interval (VSSI) monitoring techniques.

The VSI approach to adaptive monitoring is designed to make improvements over conventional charts by varying the frequency with which samples are taken. Using information contained in the previous sample, the technique determines when the next sample should be taken. Samples are taken infrequently while the process is relatively close to target. The frequency is increased as the plotted points approach the chart control limits.

The VSS approach to adaptive monitoring gives improved performance relative to conventional charts by varying the size of samples taken. The previous sample provides information by which we determine how large the next sample taken should be. Samples are relatively small in size when the process is close to the target and the size is increased when the chart limits are approached.

The VSSI approach is a combination of the VSI and VSS methods. It allows both the frequency and sample size to vary. Again information contained in the previous sample is used to determine when not only the next sample should be taken but also how large the sample should be. Samples are small and taken infrequently when the process is close to target with both the size and frequency being increased as the plotted values approach the chart limits.

The sample size for the three techniques is the fraction of streams sampled, denoted by *s*. If a process has 50 streams and 10% of the streams are sampled, the sample size is s = 5.

Threshold limits are added to the standard Shewhart chart to help in the decision of when to change the frequency and/or size of samples. Lanning, Montgomery, and Runger suggest that small samples be taken with long intervals between them when the samples plot within the threshold limits, but large frequent samples are to be taken when a sample plots beyond the threshold limits (but still within the chart control limits).

# 3. Modifications to the Traditional Group Control Chart

In this section we will consider the most commonly used method to monitor multiple-stream processes, the group control chart. Modifications to the group control chart are proposed.

## 3.1 Distribution of Maximum/Minimum

The group control chart uses the traditional 3-sigma control limits appropriate for a single stream but plots on it the maximum and minimum of the set of values from the s streams. There are problems with this approach.

We consider a multiple-stream process where each stream provides values which are generated from a normally distributed process with mean zero and standard deviation one. The upper and lower control limits on the group control chart are determined assuming the sampling is done from the standard normal process. At each time or sampling period, there is one stream that has the maximum value, and one that has the minimum value among the s streams. We plot these maximum and minimum values at each sample period. The problem with basing the control limits on the standard normal is that the plotted statistics do not follow the standard normal.

If we calculate the mean and standard deviation of the maximum values for all of the time periods, we find that the mean of the maximum values is higher than zero and the standard deviation is less than one. Also, if we calculate the mean and standard deviation of the minimum values for all of the time periods, we find that the mean of the minimum values is less than zero and the standard deviation is less than one.

Godwin (1949) gives the means and standard deviations of the maxima/minima in a sample from a process following the standard normal distribution. These values are shown in Table 7. Note that the mean gets farther and farther from zero as the sample size increases and the standard deviations of the extreme values decrease as the sample size increases. In our situation, the number of streams in the process is the sample size.

Table 7

Means and Standard Deviations of the Maxima/Minima
For a Standard Normal Process

| Number of Streams (s) | Mean of the Maxima/Minima $\mu_{max\,(s)}$ | $\mu_{min\,(s)}$ | Standard Dev. of the Maxima/Minima ($\sigma_{max/min\,(s)}$) | Corresponding Figures |
|---|---|---|---|---|
| 2 | 0.564 | -0.564 | 0.826 | Figures 4a & 4b |
| 3 | 0.846 | -0.846 | 0.748 | --------- |
| 4 | 1.029 | -1.029 | 0.701 | Figures 4c & 4d |
| 5 | 1.163 | -1.163 | 0.669 | ---------- |
| 6 | 1.267 | -1.267 | 0.645 | Figures 4e & 4f |
| 7 | 1.352 | -1.352 | 0.626 | ---------- |
| 8 | 1.423 | -1.423 | 0.611 | Figures 4g & 4h |
| 9 | 1.485 | -1.485 | 0.598 | ---------- |
| 10 | 1.539 | -1.539 | 0.587 | Figures 4i & 4j |

To calculate the mean of the maxima/minima for any other normal process, multiply the value from this table by the standard deviation and add the mean of the process.

To calculate the standard deviation, multiply the value from this table by the standard deviation of the process.

Figure 4a
Distribution of Maxima (2 Streams)

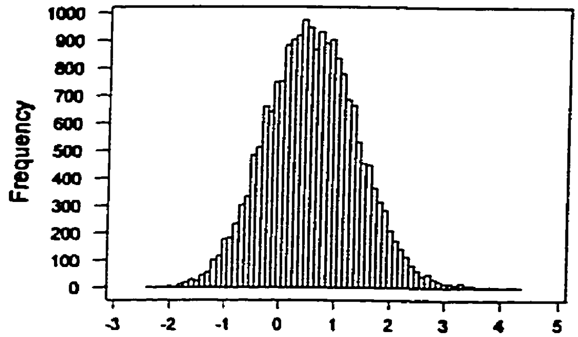



Figure 4b
Distribution of Minima (2 Streams)

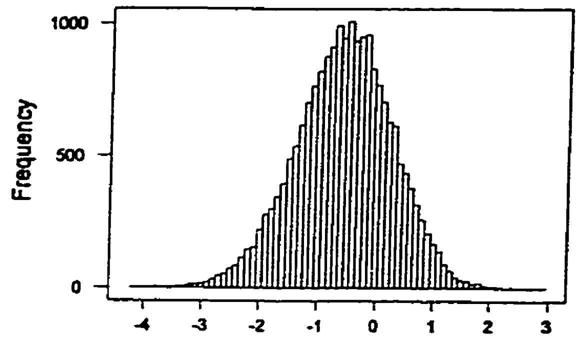



Figure 4c
Distribution of Maxima (4 Streams)

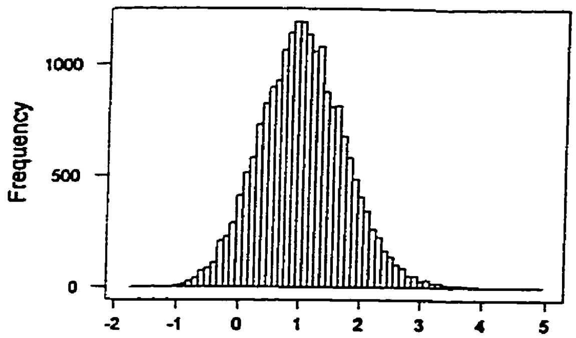



Figure 4d
Distribution of Minima (4 Streams)

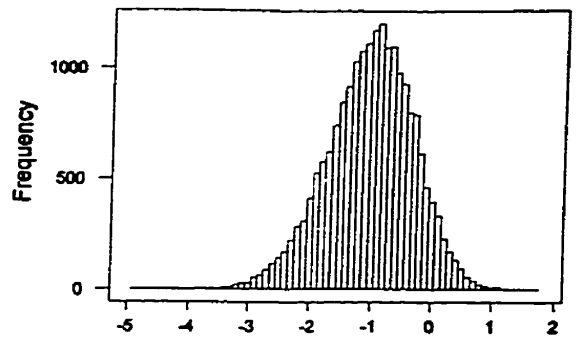



Figure 4e
Distribution of Maxima (6 Streams)

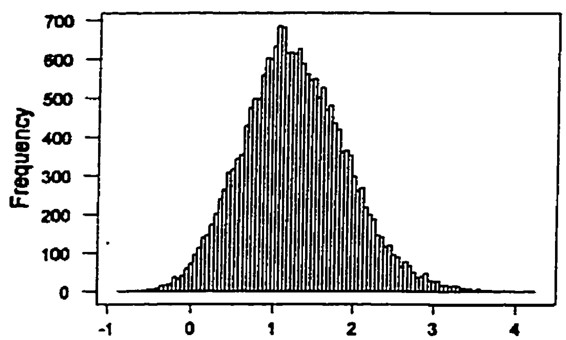



Figure 4f
Distribution of Minima (6 Streams)

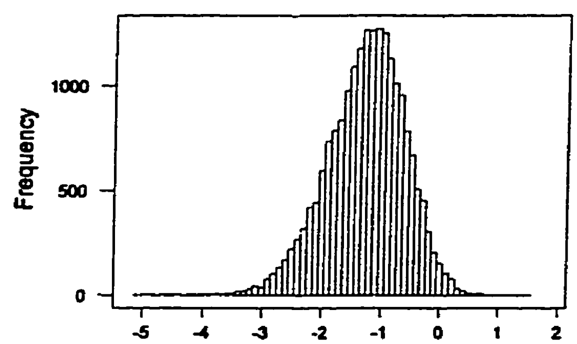


Figure 4g
Distribution of Maxima (8 Streams)

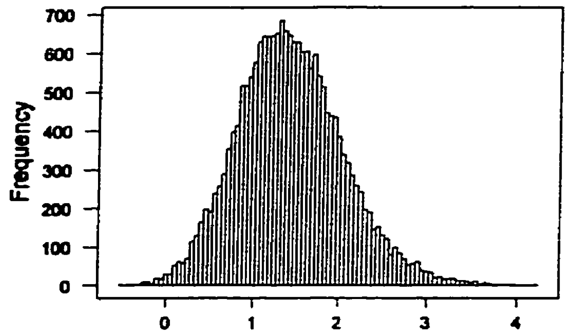



Figure 4h
Distribution of Minima (8 Streams)

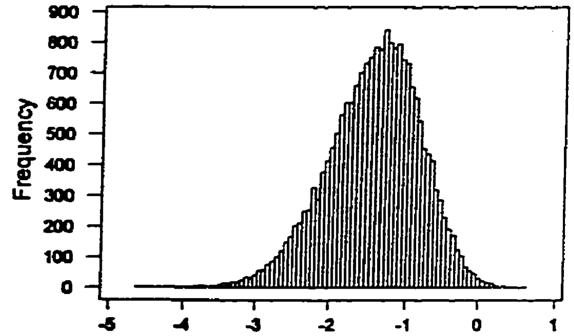



Figure 4i
Distribution of Maxima (10 Streams)

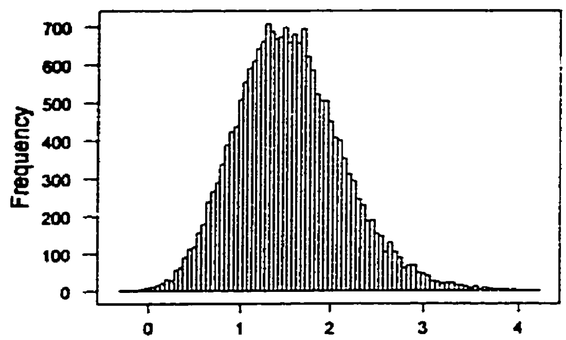
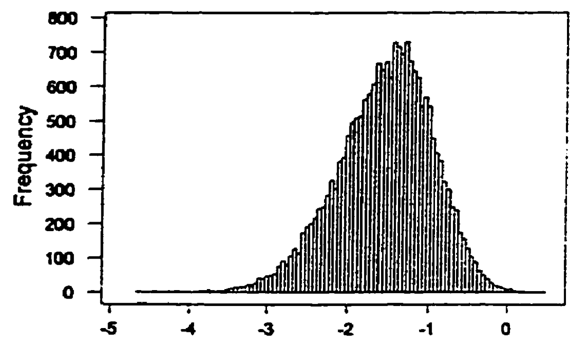



Figure 4j
Distribution of Minima (10 Streams)




Computer simulations were conducted for processes with two to ten streams. Each simulation consisted of 20,000 randomly generated standard normal values. Histograms of the distributions of the maximum and minimum values for the simulated standard normal processes with two, four, six, eight, and ten streams are shown in Figures 4a to 4j.

The problem with the traditional group control chart is that it does not take into account the distribution of the maximum and minimum values. Although the shapes of the distributions are not unlike the normal distribution, we can see that the means and standard deviations are different from 0 and 1 respectively. This results in control limits which are too narrow, especially when there are a large number of streams.

This is the reasoning behind the introduction of one of the new methods to be discussed. This new method has its control limits based on the distribution of the maximum and minimum values, rather than the overall mean and standard deviation of the process.

## 3.2 Correction Factors

It is well known that the traditional 3-sigma Shewhart control chart has a false alarm rate of 0.0027 when the process is in control.

As was noted earlier, one approach to monitor a multiple-stream process would be to use a Shewhart chart for each stream. In using the group control chart method, the plotting of the maximum and minimum of the set of stream values at a sample period is like plotting all streams on separate control charts. This is because, if the maximum and minimum values are inside the 3-sigma limits then the values from all streams must be inside these limits as well.

In our example with six streams, the probability that any one of the six control charts gives a false alarm would be 0.0027. As a result, assuming independence of the six streams and hence the six control charts, the false alarm rate for the set of six charts at any sample period is $1 - (1 - 0.0027)^6 = 0.016$. One could argue therefore that the real false alarm rate associated with the group control chart using traditional 3-sigma limits is not 0.0027 but much higher.

We now consider a modification to the group control chart on which the maximum and minimum values are plotted as suggested by Bingham (1999). The modification consists of replacing the traditional 3-sigma control limits by L-sigma control limits with L to be determined. The value of L is found such that the true false alarm rate for the group control chart with s streams is 0.0027.

Let *s* be the number of streams, and let *p* be the false alarm rate of the Shewhart chart which monitors one stream. The probability that the maximum or minimum points plot above the L-sigma UCL (or below the L-sigma LCL) when the process is in control is given as

$$P\,[(\text{Max} > \text{UCL}) \text{ or } (\text{Min} < \text{LCL})] = 2\,[1 - (1 - p)^{s}] \tag{4}$$

Setting this equation equal to 0.0027, and solving for *p* we get

$$p = 1 - (0.99865)^{1/s} \tag{5}$$

We now have an equation that gives us the false alarm rate of the Shewhart chart which monitors a single stream when the false alarm rate for the group control chart is set equal to 0.0027. This results in the following control limits for the group control chart using correction factors:

$$\text{UCL} = \bar{\bar{X}} + Z_{1-\frac{p}{2}} \left( \frac{\bar{R}}{(d_2)(\sqrt{n})} \right) \tag{6a}$$

$$\text{CL} = \bar{\bar{X}} \tag{6b}$$

$$\text{LCL} = \bar{\bar{X}} - Z_{1-\frac{p}{2}} \left( \frac{\bar{R}}{(d_2)(\sqrt{n})} \right) \tag{6c}$$

where $\bar{\bar{X}}$ = grand average (same value as traditional group control chart)

$\bar{R}$ = average range (same value as traditional group control chart)

$d_2$ = control chart constant (same value as traditional group control chart) (see Appendix A)

n = size of the subgroup from each stream (same value as traditional group control chart)

p = false alarm rate of the Shewhart control chart which monitors a single stream (see Equation (5))

$Z_{1-\frac{p}{2}}$ = value of z from the standard normal table such that

$P[Z < z] = 1 - p/2$

The standard normal value $Z_{1-\frac{p}{2}}$ is used to determine the L-sigma control limits. The correction formulas are made so that the false alarm rate of the group control chart is consistent with the ordinary Shewhart control chart.

## 3.3 <u>The Four Methods to be Evaluated</u>

Four different methods used to monitor multiple-stream processes were evaluated. Methods 1 and 4 are new techniques being proposed for the first time in this practicum. Method 2 is the technique based upon the correction factors proposed by Bingham (1999) for the group control chart. Method 3 is the conventional group control chart.

### 3.3.1 <u>Method 1: Distribution of Maximum/Minimum</u>

The control limits for Method 1 are based on the distribution of the maximum and minimum values in a sample selected from the process.

Table 8 gives the upper control limits for standard normal processes with the subgroup size n = 1 (meaning one observation is taken from each stream at each time period) and with 2 to 10 streams (see Appendix B for processes

Table 8

Method 1 — Upper Control Limits ($UCL^1_{(s)}$)
for a Standard Normal Process (n=1)

| No. of Streams | $UCL^1_{(s)}$ |
|---|---|
| 2 | 3.041 |
| 3 | 3.090 |
| 4 | 3.133 |
| 5 | 3.170 |
| 6 | 3.202 |
| 7 | 3.230 |
| 8 | 3.256 |
| 9 | 3.278 |
| 10 | 3.299 |

up to 30 streams for Methods 1, 2, and 4). The lower control limits are simply the negative of the values found in Table 8.

The values in Table 8 can be calculated from Table 7. Suppose we have a standard normal process with six streams and a subgroup size of one. The upper control limit is three standard deviations above the mean of the distribution of the maximum. Substituting the appropriate values from Table 7, we get:

$$\begin{aligned} UCL = \mu_{max\,(6)} + 3\ \sigma_{max/min\,(6)} &= 1.267 + 3\,(0.645) \\ &= 3.202 \\ &= UCL^1_{(6)} \end{aligned} \tag{7a}$$

The lower control limit is three standard deviations below the mean of the distribution of the minimum:

$$
\begin{aligned}
LCL = \mu_{\min(6)} - 3\,\sigma_{\max/\min(6)} &= -1.267 - 3\,(0.645) \\
&= -3.202 \\
&= -\,UCL^1{}_{(6)}
\end{aligned}
\tag{7b}
$$

The centre line is equal to the mean of the process. Since the process for our example is the standard normal distribution, the centre line is equal to zero.

$$CL = \mu = 0 \tag{7c}$$

To calculate the control limits for a normal process which has a mean different than zero and/or a standard deviation other than one, the following formulae can be used:

$$UCL = \mu + UCL^1{}_{(s)}\left(\frac{\sigma}{\sqrt{n}}\right) \tag{8a}$$

$$CL = \mu \tag{8b}$$

$$LCL = \mu - UCL^1{}_{(s)}\left(\frac{\sigma}{\sqrt{n}}\right) \tag{8c}$$

where $\mu$ = mean of the process

$\sigma$ = standard deviation of the process

$UCL^1{}_{(s)}$ = upper control limit for standard normal process (see Table 8)

n = size of the subgroup from each stream

For example, the control chart for a process with six streams, a mean of five, a standard deviation of ten, and a subgroup of size four (meaning four observations are taken from each stream at each time period) would have the following control limits:

$$\text{UCL} = 5 + (3.202)\left(\frac{10}{\sqrt{4}}\right) = 21.01 \tag{9a}$$

$$\text{CL} = 5 \tag{9b}$$

$$\text{LCL} = 5 - (3.202)\left(\frac{10}{\sqrt{4}}\right) = -11.01 \tag{9c}$$

In practice, however, we never know the true mean and standard deviation of the process. Therefore, we need to estimate the values of $\mu$ and $\sigma$. The following formulae give estimates for $\mu$ and $\sigma$:

$$\hat{\mu} = \overline{\overline{X}} \tag{10a}$$

$$\hat{\sigma} = \frac{\overline{R}}{d_2} \text{ if } n > 1 \quad \left(\frac{\overline{\overline{MR}}}{d_2} \text{ if } n = 1 \text{ (using } d_2 \text{ for } n=2)\right) \tag{10b}$$

Note that if the subgroup size is one, then the overall average moving range is used instead of the average range. The moving range (MR) is the absolute value of the difference between two consecutive observations. The average moving range ($\overline{MR}$) is simply the average of all of the moving ranges for each stream (there should be as many $\overline{MR}$'s as there are streams). The overall average moving range ($\overline{\overline{MR}}$) is the average of the average moving ranges for each stream. By substituting the mean and standard deviation estimates into Equations (8a) to (8c), we get the following control limits for Method 1:

**Control Limits For Method 1**

$$UCL = \bar{\bar{X}} + UCL^1_{(s)} \left( \frac{\bar{R}}{(d_2)(\sqrt{n})} \right) \qquad (11a)$$

$$CL = \bar{\bar{X}} \qquad (11b)$$

$$LCL = \bar{\bar{X}} - UCL^1_{(s)} \left( \frac{\bar{R}}{(d_2)(\sqrt{n})} \right) \qquad (11c)$$

where $\bar{\bar{X}}$ = grand average of the data (substitute $\bar{\bar{X}}$ with $\bar{X}$ if n=1)

$\bar{R}$ = average range of the data (substitute $\bar{R}$ with $\overline{\overline{MR}}$ if n=1)

$UCL^1_{(s)}$ = upper control limit for standard normal process (see Table 8)

$d_2$ = control chart constant (see Appendix A)

n = subgroup size

3.3.2 Method 2: Correction Factors

This is the method that makes use of correction factors. Suppose the process we are studying follows the standard normal distribution and has a subgroup size of one. The upper control limit using Method 2 is $Z_{1-\frac{p}{2}}$, the value of z from the standard normal table satisfying P [Z < z] = 1 - $\frac{p}{2}$, where p is the false alarm rate taking into account all of the streams in the process. The lower control limit is $-Z_{1-\frac{p}{2}}$. The value of $Z_{1-\frac{p}{2}}$ could be called $UCL^2_{(s)}$ because it is Method 2's upper control limit for the standard normal process. Table 9 gives the values of $UCL^2_{(s)}$ for 2 to 10 streams (see Appendix B for processes up to 30 streams).

Table 9

Method 2 — Upper Control Limits ($UCL^2_{(s)}$)
for a Standard Normal Process (n=1)

| Streams | $UCL^2_{(s)}$ |
|---|---|
| 2 | 3.399 |
| 3 | 3.509 |
| 4 | 3.585 |
| 5 | 3.642 |
| 6 | 3.689 |
| 7 | 3.728 |
| 8 | 3.762 |
| 9 | 3.791 |
| 10 | 3.817 |

If we further suppose that the process has six streams, we see from Table 9 that the upper control limit is 3.6890 and the lower control limit is -3.6890. The centre line is at zero since the process follows the standard normal distribution.

Control limits for normal processes with a mean different than zero and/or a standard deviation other than one can be calculated from the following formulae:

$$UCL = \mu + UCL^2_{(s)}\left(\frac{\sigma}{\sqrt{n}}\right) \tag{12a}$$

$$CL = \mu \tag{12b}$$

$$LCL = \mu - UCL2_{(s)}\left(\frac{\sigma}{\sqrt{n}}\right) \tag{12c}$$

For example, a process with six streams, a mean of five, a standard deviation of ten, and a subgroup of size four would have the following control limits:

$$UCL = 5 + (3.6890)\left(\frac{10}{\sqrt{4}}\right) = 23.45 \tag{13a}$$

$$CL = 5 \tag{13b}$$

$$LCL = 5 - (3.6890)\left(\frac{10}{\sqrt{4}}\right) = -13.45 \tag{13c}$$

Note that these control limits are wider than the limits of Method 1 (Equations (9a) to (9c)).

Again, we do not know the true mean and standard deviation of the process in practice. Substituting the estimates for $\mu$ and $\sigma$ (Equations (10a) and (10b)) into Equations (12a) to (12c), we arrive at the following control limits for Method 2:

**Control Limits For Method 2**

---

$$UCL = \bar{\bar{X}} + UCL2_{(s)}\left(\frac{\bar{R}}{(d_2)(\sqrt{n})}\right) \tag{14a}$$

$$CL = \bar{\bar{X}} \tag{14b}$$

$$LCL = \bar{\bar{X}} - UCL2_{(s)}\left(\frac{\bar{R}}{(d_2)(\sqrt{n})}\right) \tag{14c}$$

where $\bar{\bar{X}}$ = grand average of the data (substitute $\bar{\bar{X}}$ with $\bar{X}$ if n=1)

$\bar{R}$ = average range of the data (substitute $\bar{R}$ with $\overline{\overline{MR}}$ if n=1)

$UCL^2_{(s)}$ = upper control limit for standard normal process (see Table 9) (equivalent to $Z_{1-\frac{p}{2}}$, the value of z satisfying P [Z < z] = 1 - $\frac{p}{2}$)

$d_2$ = control chart constant (see Appendix A)

n = subgroup size

---

3.3.3 Method 3: Traditional Method

Method 3 is the traditional group control chart. Regardless of the number of streams, a process following the standard normal distribution with a subgroup size of one has an upper control limit of 3.0, a lower control limit of -3.0, and a centre line at zero. This means that $UCL^3_{(s)}$ = 3.0 for all values of s.

Control limits for processes which do not have a mean of zero and/or a standard deviation of one can be calculated from the following formulae:

$$UCL = \mu + 3\left(\frac{\sigma}{\sqrt{n}}\right) \tag{15a}$$

$$CL = \mu \tag{15b}$$

$$LCL = \mu - 3\left(\frac{\sigma}{\sqrt{n}}\right) \tag{15c}$$

A process with six streams, a mean of five, a standard deviation of ten, and a subgroup size of four has the following control limits:

$$UCL = 5 + 3\left(\frac{10}{\sqrt{4}}\right) = 20.00 \tag{16a}$$

$$CL = 5 \tag{16b}$$

$$LCL = 5 - 3\left(\frac{10}{\sqrt{4}}\right) = -10.00 \tag{16c}$$

A process with ten streams, a mean of five, a standard deviation of ten, and a subgroup of size four has the same control limits as the six stream case because the number of streams has no effect on the limits. However, it is more likely that points will plot outside the limits if there are ten streams than when there are only six. That is why the number of streams should be considered in the construction of the control limits.

Since we do not know the true mean and standard deviation of the process in practice, we can use the estimates of $\mu$ and $\sigma$ (Equations (10a) and (10b)) to get the following control limits for Method 3:

**Control Limits For Method 3**

$$UCL = \bar{\bar{X}} + 3\left(\frac{\bar{R}}{(d_2)(\sqrt{n})}\right) = \bar{\bar{X}} + A_2\,\bar{R} \tag{17a}$$

$$CL = \bar{\bar{X}} \tag{17b}$$

$$LCL = \bar{\bar{X}} - 3\left(\frac{\bar{R}}{(d_2)(\sqrt{n})}\right) = \bar{\bar{X}} - A_2\,\bar{R} \tag{17c}$$

where $\bar{\bar{X}}$ = grand average of the data (substitute $\bar{\bar{X}}$ with $\bar{X}$ if n=1)

$\bar{R}$ = average range of the data (substitute $\bar{R}$ with $\overline{\overline{MR}}$ if n=1)

$d_2$ = control chart constant (see Appendix A)

n = subgroup size

$A_2$ = control chart constant (substitute $A_2$ with $E_2$ if n=1) (see Appendix A)

### 3.3.4 Method 4: Simulated Control Limits

Method 4 was constructed by computer simulations. The simulations were done so that an out-of-control signal would occur 0.27% of the time on average when the process is in control.

An out-of-control signal is defined to be when one of the following situations occurs:

(1) the stream with the largest subgroup mean is higher than the upper control limit

(2) the stream with the smallest subgroup mean is lower than the lower control limit

(3) the stream with the largest subgroup mean is higher than the UCL and the stream with the smallest subgroup mean is lower than the LCL in the same sample or time period

Note that the occurrence of Situation 3 counts as **one** out-of-control signal. Regardless of the number of streams, the two-sided in-control average run length for Method 4 is 370.37.

Table 10 gives the values for $UCL^4_{(s)}$, the estimated upper control limits for standard normal processes with a subgroup size of one and with 2 to 10 streams (see Appendix B for processes up to 30 streams). The lower control limits are simply the negative of the values found in Table 10. For example, a standard normal process with six streams would have a UCL of 3.51 and a LCL of -3.51.

Control limits for processes which are not standard normal can be calculated from the following formulae:

Table 10

Method 4 — Estimated Upper Control Limits ($UCL^4_{(s)}$)
for a Standard Normal Process (n=1)

| No. of Streams | $UCL^4_{(s)}$ |
|---|---|
| 2 | 3.23 |
| 3 | 3.35 |
| 4 | 3.41 |
| 5 | 3.47 |
| 6 | 3.51 |
| 7 | 3.55 |
| 8 | 3.59 |
| 9 | 3.62 |
| 10 | 3.65 |

$$UCL = \mu + UCL^4_{(s)} \left(\frac{\sigma}{\sqrt{n}}\right) \tag{18a}$$

$$CL = \mu \tag{18b}$$

$$LCL = \mu - UCL^4_{(s)} \left(\frac{\sigma}{\sqrt{n}}\right) \tag{18c}$$

For example, a process with six streams, a mean of five, a standard deviation of ten, and a subgroup size of four would have the following control limits:

$$UCL = 5 + (3.51)\left(\frac{10}{\sqrt{4}}\right) = 22.55 \tag{19a}$$

$$CL = 5 \tag{19b}$$

$$LCL = 5 - (3.51)\left(\frac{10}{\sqrt{4}}\right) = -12.55 \tag{19c}$$

Since we do not know the true mean and standard deviation in practice, we can use the estimates of $\mu$ and $\sigma$ (Equations (10a) and (10b)) to get the following control limits for Method 4:

**Control Limits For Method 4**

---

$$UCL = \bar{\bar{X}} + UCL^4_{(s)}\left(\frac{\bar{R}}{(d_2)(\sqrt{n})}\right) \tag{20a}$$

$$CL = \bar{\bar{X}} \tag{20b}$$

$$LCL = \bar{\bar{X}} - UCL^4_{(s)}\left(\frac{\bar{R}}{(d_2)(\sqrt{n})}\right) \tag{20c}$$

where $\bar{\bar{X}}$ = grand average of the data (substitute $\bar{\bar{X}}$ with $\bar{X}$ if n=1)

$\bar{R}$ = average range of the data (substitute $\bar{R}$ with $\overline{\overline{MR}}$ if n=1)

$UCL^4_{(s)}$ = estimated upper control limit for a standard normal process constructed by computer simulations (see Table 10)

$d_2$ = control chart constant (see Appendix A)

n = subgroup size

---

# 4. Average Run Lengths

Computer simulations were conducted to determine estimates of the in-control and out-of-control average run lengths for the four methods discussed in the last section. The average run length (ARL) for the group control chart is the average number of samples or time periods occurring before at least one of the two plotted points indicates an out-of-control condition. The ARL is the reciprocal of the probability that at least one of the two plotted points goes beyond the upper and/or lower control limits. If a process is in control, we do not want to see frequent out-of-control signals because these are false alarms. It is therefore desirable to have the in-control average run length ($ARL_0$) as high as possible. A Shewhart 3-sigma control chart has an in-control ARL of approximately 370. This value forms a useful baseline against which to compare other charts. On the other hand, if the process is out-of-control, we would like to detect this situation as soon as possible. Therefore, the out-of-control average run length ($ARL_1$) should be as small as possible.

The computer programs used for the simulations were written in the FORTRAN language. Four simulation runs were used to estimate each of the average run lengths, with 20,000 normally distributed values generated in each run.

Table 11 gives the upper control limits for the computer simulated standard normal processes for Methods 1 to 4. The values given in Table 11 are the same as the values for $UCL^i_{(s)}$ (i = 1, 2, 3, 4) given in the previous section. The centre line for all of the control charts is zero and the lower control limit is simply the negative of the UCL.

Table 11

Upper Control Limits of Standard Normal Processes

| No. of Streams | Method 1 | Method 2 | Method 3 | Method 4 |
|---|---|---|---|---|
| 2 | 3.04 | 3.40 | 3.00 | 3.23 |
| 3 | 3.09 | 3.51 | 3.00 | 3.35 |
| 4 | 3.13 | 3.58 | 3.00 | 3.41 |
| 5 | 3.17 | 3.64 | 3.00 | 3.47 |
| 6 | 3.20 | 3.69 | 3.00 | 3.51 |
| 7 | 3.23 | 3.73 | 3.00 | 3.55 |
| 8 | 3.26 | 3.76 | 3.00 | 3.59 |
| 9 | 3.28 | 3.79 | 3.00 | 3.62 |
| 10 | 3.30 | 3.82 | 3.00 | 3.66 |

### 4.1 In-control ARLs

Table 12 gives the estimated in-control average run lengths ($ARL_0$) for the four methods. The values in Table 12 are averages of four computer runs. The four runs were each started with a different initial seed in order to generate a different set of standard normal observations.

We see from Table 12 that the in-control ARLs for Methods 1 and 3 decrease as the number of streams increase. The ARLs for Method 2 seem to increase slightly as the number of streams increase. Method 4 was designed so that its in-control ARL is 370.37 regardless of the number of streams.

If we look at the ARL for Method 3 and ten streams (in bold type in Table

Table 12

In-control Two-sided Average Run Lengths

| No. of Streams | Method 1 | Method 2 | Method 3 | Method 4 |
|---|---|---|---|---|
| 2 | 196.15 | 698.57 | 169.19 | 370.37 |
| 3 | 156.18 | 677.03 | 114.28 | 370.37 |
| 4 | 141.75 | 755.98 | 89.56 | 370.37 |
| 5 | 128.34 | 729.94 | 73.54 | 370.37 |
| 6 | 120.19 | 718.01 | 61.56 | 370.37 |
| 7 | 114.77 | 769.75 | 53.30 | 370.37 |
| 8 | 113.00 | 758.78 | 47.04 | 370.37 |
| 9 | 112.94 | 788.82 | 42.16 | 370.37 |
| 10 | 106.21 | 780.19 | **37.88** | 370.37 |

12), we see that the value is 37.88. Getting an out-of-control signal when the process is really in control every 38 samples is much too frequent. This again explains why there is a problem with the traditional group control chart and why alternative methods are being discussed.

### 4.2 Out-of-control ARLs

Two situations were constructed in order to get estimates of the out-of-control average run lengths ($ARL_1$). The first was having the mean value of one stream shift by a specified amount and the second was having the mean values of all streams shift by the same specified amount.

Table 13 gives the ARL values for the shift of one stream for processes with two, four, six, eight, and ten streams. Table 14 gives the ARL values for shifts affecting all of the two, four, six, eight, and ten streams. Shifts in the mean of 0.5 to 3.0 standard deviations are shown, as well as the in-control ARLs (no shift) which were shown previously in Table 12. The ARL values for an individual Shewhart control chart using 3-sigma limits are provided for comparison purposes.

Starting with Table 13 and Method 1, we see that the out-of-control ARLs increase as the number of streams increase for shifts of one standard deviation or more. The same situation occurs for Method 2. Again, the ARLs increase as the streams increase. However, the ARLs for Method 2 are much higher than for Method 1. For example, a process with ten streams and a 2-sigma shift in one of its streams will take Method 1 on average 9.46 samples to detect the shift, as compared to 27.56 samples for Method 2.

Turning our attention to Method 3, we see that the ARLs decrease as the number of streams increase. We also see that the ARLs are much smaller for all of the different numbers of streams, as compared with Methods 1 and 2.

For Method 4, the out-of-control ARLs increase as the number of streams increase. The ARLs are larger than those for Methods 1 and 3, but lower than those for Method 2.

Looking at Table 14, we see that the out-of-control ARLs generally decrease as the number of streams increase for all four methods. The highest ARLs belong to Method 2, with Method 4 having the second highest and Method 1 having the third highest ARLs. Method 3 has the lowest average run lengths. This ordering is the same as the ordering for the shift of one stream situation.

Table 13

Average Run Lengths
(Shift of One Stream)

| No. of Streams | No Shift | ±0.5 $\sigma$ | ±1.0 $\sigma$ | ±1.5 $\sigma$ | ±2.0 $\sigma$ | ±2.5 $\sigma$ | ±3.0 $\sigma$ |
|---|---|---|---|---|---|---|---|
| Method 1: | | | | | | | |
| 2 | 196.15 | 121.14 | 44.48 | 15.70 | 6.62 | 3.39 | 2.06 |
| 4 | 141.75 | 100.62 | 46.33 | 18.12 | 7.51 | 3.73 | 2.22 |
| 6 | 120.19 | 95.23 | 48.36 | 19.45 | 8.37 | 4.05 | 2.36 |
| 8 | 113.00 | 92.52 | 49.87 | 21.21 | 9.08 | 4.34 | 2.47 |
| 10 | 106.21 | 88.11 | 51.52 | 22.19 | 9.46 | 4.53 | 2.57 |
| Method 2: | | | | | | | |
| 2 | 698.57 | 367.43 | 111.02 | 34.50 | 12.36 | 5.41 | 2.90 |
| 4 | 755.98 | 469.57 | 172.54 | 51.67 | 17.65 | 7.14 | 3.58 |
| 6 | 718.01 | 541.98 | 211.44 | 64.51 | 21.34 | 8.54 | 4.06 |
| 8 | 758.78 | 580.31 | 256.16 | 76.45 | 24.99 | 9.72 | 4.47 |
| 10 | 780.19 | 664.35 | 291.10 | 88.32 | 27.56 | 10.43 | 4.77 |
| Method 3: | | | | | | | |
| 2 | 169.19 | 108.98 | 40.59 | 14.51 | 6.21 | 3.22 | 1.99 |
| 4 | 89.56 | 68.20 | 32.90 | 13.40 | 6.03 | 3.19 | 1.98 |
| 6 | 61.56 | 50.01 | 27.66 | 12.68 | 5.87 | 3.14 | 1.97 |
| 8 | 47.04 | 40.92 | 24.66 | 12.04 | 5.76 | 3.11 | 1.96 |
| 10 | 37.88 | 33.39 | 21.69 | 11.17 | 5.52 | 3.07 | 1.95 |
| Method 4: | | | | | | | |
| 2 | 370.37 | 211.04 | 70.40 | 23.21 | 9.10 | 4.30 | 2.44 |
| 4 | 370.37 | 256.81 | 101.70 | 33.70 | 12.32 | 5.44 | 2.93 |
| 6 | 370.37 | 276.14 | 119.65 | 40.52 | 14.89 | 6.31 | 3.25 |
| 8 | 370.37 | 298.10 | 143.98 | 48.72 | 17.41 | 7.23 | 3.58 |
| 10 | 370.37 | 329.50 | 162.21 | 55.96 | 19.32 | 7.92 | 3.88 |
| Shewhart: | 370.37 | 156.25 | 43.86 | 14.97 | 6.30 | 3.24 | 2.00 |

Table 14

Average Run Lengths
(Shift of All Streams)

| No. of Streams | No Shift | $\pm 0.5\sigma$ | $\pm 1.0\sigma$ | $\pm 1.5\sigma$ | $\pm 2.0\sigma$ | $\pm 2.5\sigma$ | $\pm 3.0\sigma$ |
|---|---|---|---|---|---|---|---|
| Method 1: | | | | | | | |
| 2 | 196.15 | 87.76 | 25.04 | 8.37 | 3.63 | 1.99 | 1.37 |
| 4 | 141.75 | 57.32 | 15.62 | 5.26 | 2.37 | 1.41 | 1.10 |
| 6 | 120.19 | 47.56 | 12.32 | 4.16 | 1.93 | 1.23 | 1.04 |
| 8 | 113.00 | 42.26 | 10.67 | 3.59 | 1.70 | 1.15 | 1.02 |
| 10 | 106.21 | 38.57 | 9.74 | 3.26 | 1.56 | 1.10 | 1.01 |
| Method 2: | | | | | | | |
| 2 | 698.57 | 252.53 | 61.38 | 17.83 | 6.46 | 2.99 | 1.76 |
| 4 | 755.98 | 233.04 | 52.33 | 13.90 | 4.79 | 2.22 | 1.37 |
| 6 | 718.01 | 225.66 | 47.35 | 11.92 | 4.05 | 1.90 | 1.22 |
| 8 | 758.78 | 221.17 | 44.19 | 10.83 | 3.63 | 1.71 | 1.15 |
| 10 | 780.19 | 213.85 | 41.93 | 10.17 | 3.37 | 1.59 | 1.11 |
| Method 3: | | | | | | | |
| 2 | 169.19 | 77.62 | 22.80 | 7.76 | 3.43 | 1.92 | 1.34 |
| 4 | 89.56 | 38.54 | 11.52 | 4.12 | 2.01 | 1.30 | 1.07 |
| 6 | 61.56 | 25.86 | 7.72 | 2.93 | 1.55 | 1.12 | 1.02 |
| 8 | 47.04 | 19.72 | 5.90 | 2.34 | 1.33 | 1.06 | 1.004 |
| 10 | 37.88 | 16.05 | 4.86 | 2.00 | 1.22 | 1.03 | 1.001 |
| Method 4: | | | | | | | |
| 2 | 370.37 | 153.77 | 38.96 | 12.26 | 4.84 | 2.43 | 1.54 |
| 4 | 370.37 | 134.11 | 31.63 | 9.31 | 3.55 | 1.82 | 1.23 |
| 6 | 370.37 | 125.51 | 27.78 | 7.90 | 2.98 | 1.56 | 1.13 |
| 8 | 370.37 | 124.95 | 26.59 | 7.19 | 2.69 | 1.44 | 1.08 |
| 10 | 370.37 | 126.85 | 26.62 | 6.93 | 2.55 | 1.37 | 1.06 |
| Shewhart: | | | | | | | |
| | 370.37 | 156.25 | 43.86 | 14.97 | 6.30 | 3.24 | 2.00 |

## 4.3 Comparison of the Four Methods

To help determine what method may work best, it is useful to summarize the analysis. This summary is given in Table 15. Note that the "magnitude ordering" can also be thought of as the method's ranking because the ordering goes from highest to lowest for the in-control ARLs and from lowest to highest for the out-of-control ARLs.

From examination of Table 15, it appears that both Methods 2 and 3 are poor. Looking at Method 2, we see that the in-control ARLs are high which is desirable. These ARLs are much higher than the standard Shewhart in-control ARL of 370.37. The consequence of this is that the out-of-control ARLs are too high. On the other hand, Method 3 has low out-of-control ARLs but the in-control ARLs are much too low.

We now consider Methods 1 and 4. We know that the in-control ARLs for Method 4 are exactly the same as the in-control ARL for the standard Shewhart chart. We also know that the in-control ARLs for Method 1 are lower than the traditional in-control ARL. But how do the out-of-control ARLs for Methods 1 and 4 compare with the traditional out-of-control ARLs? Looking into this may help decide which of these two remaining methods compares favourably with the traditional Shewhart chart. Table 16 gives a comparison of Method 1, Method 4, and the Shewhart X chart.

Looking at Table 16, it is difficult to decide which of the two methods is better. The main reason for this is the large difference in out-of-control ARLs when only one stream shifts compared to when all streams shift. For both methods, the out-of-control ARLs are better than the standard when all streams shift, but are not as good when only one stream shifts. The only exception is when one stream shifts $\pm 0.5\sigma$ for Method 1. In that case, Method 1 is better than the standard.

Table 15

Summary of the Average Run Lengths

| Method | In-control ARL: Behaviour as Streams Increase | In-control ARL: Magnitude | In-control ARL: Magnitude Ordering (Highest to Lowest) |
|---|---|---|---|
| 1 | decreasing | fairly low | 3rd |
| 2 | increasing | high | 1st |
| 3 | decreasing | much too low for medium/large no. of streams | 4th |
| 4 | stays constant | same as the Shewhart chart | 2nd |

| Method | Out-of-control ARL: One Stream Shifted: Behaviour as Streams Increase | One Stream Shifted: Magnitude | All Streams Shifted: Behaviour as Streams Increase | All Streams Shifted: Magnitude | Magnitude Ordering (Lowest to Highest) |
|---|---|---|---|---|---|
| 1 | increasing for shifts > 1.0 $\sigma$ | fairly low | decreasing | fairly low | 2nd |
| 2 | increasing | high | decreasing | high | 4th |
| 3 | decreasing | low | decreasing | low | 1st |
| 4 | increasing | fairly high | decreasing | fairly high | 3rd |

Table 16

Average Run Lengths for Method 1,
Method 4, and the Standard Shewhart X Chart
for Two and Ten Streams

| Method | $ARL_0$ | | | $ARL_1$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2 Streams | 10 Streams | | $\pm 0.5\sigma$ | | $\pm 1.0\sigma$ | | $\pm 1.5\sigma$ | |
| | | | | 2 | 10 | 2 | 10 | 2 | 10 |
| 1 | 196.15 | 106.21 | One Stream | 121.14 | 88.11 | 44.48 | 51.52 | 15.70 | 22.19 |
| | | | All Streams | 87.76 | 38.57 | 25.04 | 9.74 | 8.37 | 3.26 |
| 4 | 370.37 | 370.37 | One Stream | 211.04 | 329.50 | 70.40 | 162.21 | 23.21 | 55.96 |
| | | | All Streams | 153.77 | 126.85 | 38.96 | 26.62 | 12.26 | 6.93 |
| Standard Shewhart X Chart | 370.37 | | | 156.25 | | 43.86 | | 14.97 | |

| Method | | $\pm 2.0\sigma$ | | $\pm 2.5\sigma$ | | $\pm 3.0\sigma$ | |
|---|---|---|---|---|---|---|---|
| | | 2 | 10 | 2 | 10 | 2 | 10 |
| 1 | One Stream | 6.62 | 9.46 | 3.39 | 4.53 | 2.06 | 2.57 |
| | All Streams | 3.63 | 1.56 | 1.99 | 1.10 | 1.37 | 1.01 |
| 4 | One Stream | 9.10 | 19.32 | 4.30 | 7.92 | 2.44 | 3.88 |
| | All Streams | 4.84 | 2.55 | 2.43 | 1.37 | 1.54 | 1.06 |
| Standard Shewhart X Chart | | 6.30 | | 3.24 | | 2.00 | |

Therefore, if we want to hold to the rule of having an in-control ARL of 370.37, then we can use Method 4 because even though the out-of-control ARLs are larger than those for Method 1, they are still reasonably close to the Shewhart chart. But, if we want to be able to detect out-of-control situations sooner and we can allow lower in-control ARLs, then Method 1 would be the appropriate method.

# 5. Data Set Analysis

## 5.1 <u>Data Set 1</u>

The data from "Data Set 1" are shown in Table 17. The data come from measurements taken from six locations around the circumference of a manufactured part and reported as deviations from the nominal value. With locations as streams, there are thus six streams in this process and they are labeled E0 to E5. Figure 5 displays a histogram and boxplot of the entire data set and Figures 6a to 6f display the data by individual stream. The histograms and boxplots show that the data are not normally distributed. There are a number of outliers present for Streams E1 to E5. The most extreme outlier for each of Streams E1 to E5 are shown in bold type in Table 17.

### 5.1.1 <u>Control Chart Using the Streams as a Subgroup</u>

The control chart method used by the company from which the data were obtained was the common Shewhart 3-sigma $\overline{X}$ and R chart treating the measurements from the six locations on each part as a subgroup. Figure 7 shows the control charts and it appears that the process is in statistical control.

### 5.1.2 <u>Testing the Equality of the Streams</u>

Means diamonds and quantile boxes for the six streams are shown in Figure 8 to help determine the equality of the streams. Figure 8 indicates that the streams do not appear to be identical with respect to their location. Stream E1 appears to be the highest and Stream E3 appears to be the lowest of the six streams. Since the data are nonnormal, the Friedman two-way analysis of variance by ranks procedure was used to test the equality of the streams. First, the data were ranked for each time period. Then the ranks

Table 17

Data Set 1

| Sample No. | E0 | E1 | E2 | E3 | E4 | E5 |
|---|---|---|---|---|---|---|
| 1 | 0.006 | 0.021 | 0.014 | **0.034** | **0.032** | -0.012 |
| 2 | 0.003 | 0.018 | 0.005 | -0.022 | -0.014 | -0.015 |
| 3 | 0.009 | 0.019 | 0.007 | -0.019 | -0.013 | -0.013 |
| 4 | 0.010 | 0.022 | 0.010 | -0.018 | -0.012 | -0.010 |
| 5 | 0.002 | **0.002** | **-0.018** | -0.015 | 0.001 | **0.017** |
| 6 | 0.013 | 0.029 | 0.012 | -0.020 | -0.013 | -0.010 |
| 7 | 0.006 | 0.021 | -0.002 | -0.021 | -0.010 | -0.005 |
| 8 | 0.009 | 0.024 | 0.005 | -0.021 | -0.013 | -0.011 |
| 9 | 0.009 | 0.020 | 0.002 | -0.022 | -0.015 | -0.010 |
| 10 | 0.008 | 0.017 | 0.001 | -0.029 | -0.021 | -0.013 |
| 11 | 0.009 | 0.022 | 0.001 | -0.026 | -0.019 | -0.011 |
| 12 | 0.005 | 0.022 | -0.001 | -0.022 | -0.014 | -0.019 |
| 13 | 0.009 | 0.020 | 0.002 | -0.025 | -0.020 | -0.013 |
| 14 | 0.006 | 0.021 | 0.001 | -0.028 | -0.022 | -0.020 |
| 15 | 0.011 | 0.024 | 0.000 | -0.030 | -0.024 | -0.016 |
| 16 | 0.010 | 0.021 | 0.000 | -0.027 | -0.020 | -0.010 |
| 17 | 0.005 | 0.022 | 0.005 | -0.019 | -0.007 | -0.010 |
| 18 | 0.008 | 0.017 | -0.004 | -0.021 | -0.009 | 0.006 |
| 19 | 0.013 | 0.028 | 0.004 | -0.027 | -0.017 | -0.008 |
| 20 | 0.004 | 0.015 | -0.002 | -0.022 | -0.012 | -0.006 |

Figure 5: Histogram and Boxplot of Data Set 1

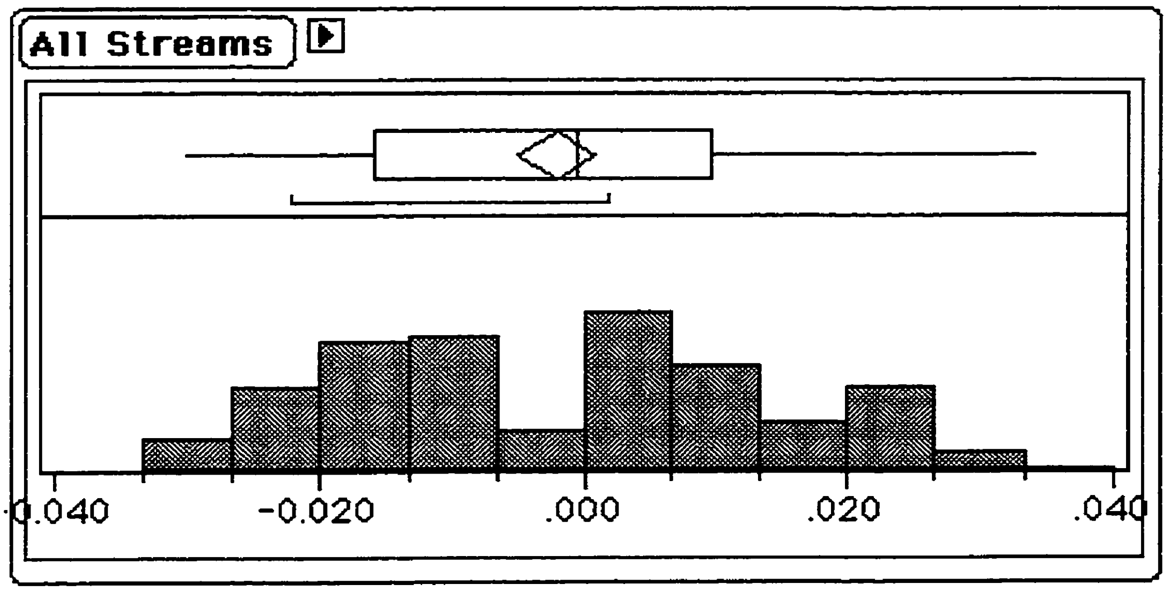


Figure 6a

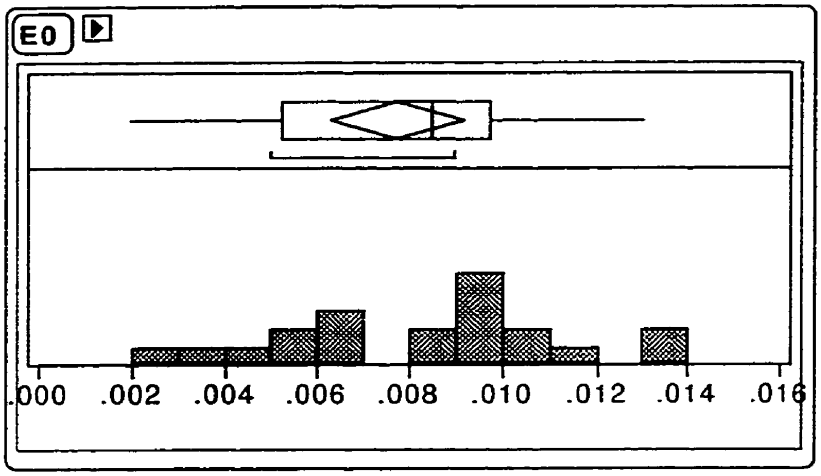



Figure 6b

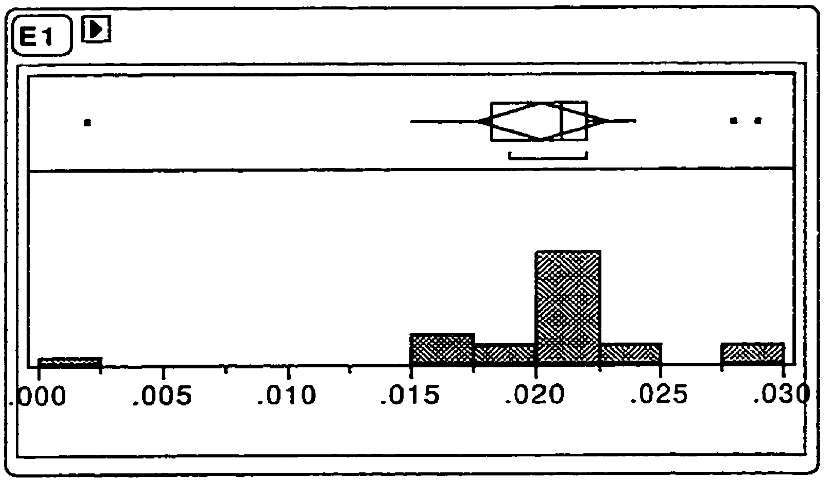



Figure 6c

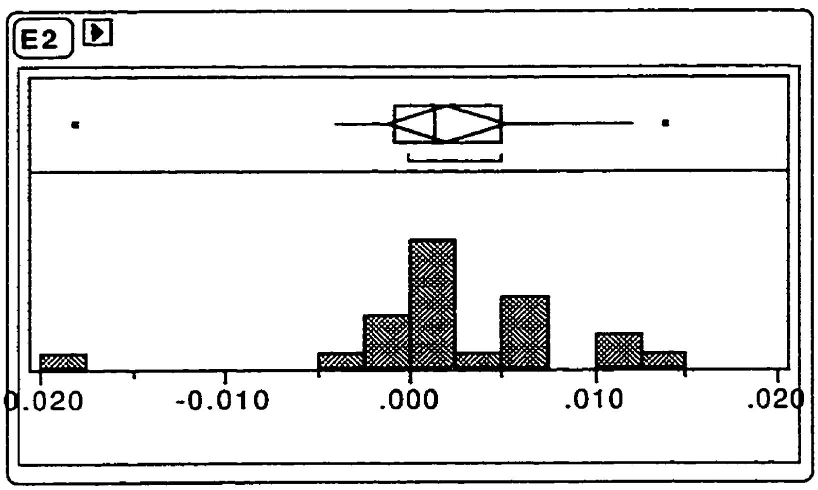


Figure 6d

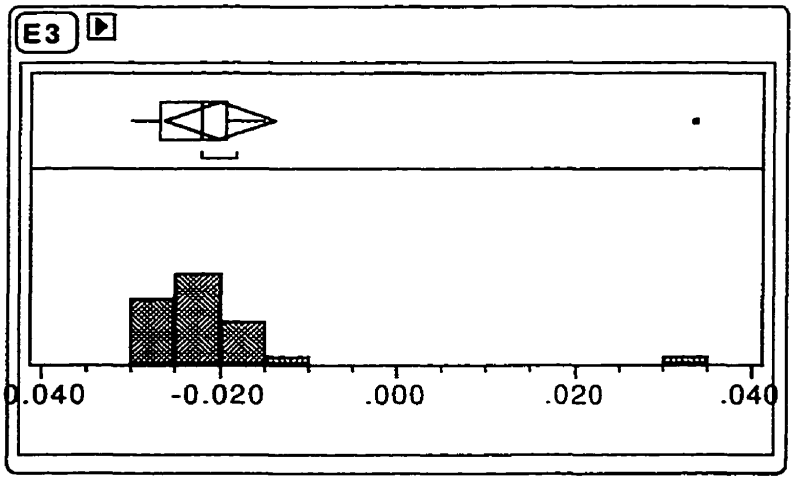



Figure 6e

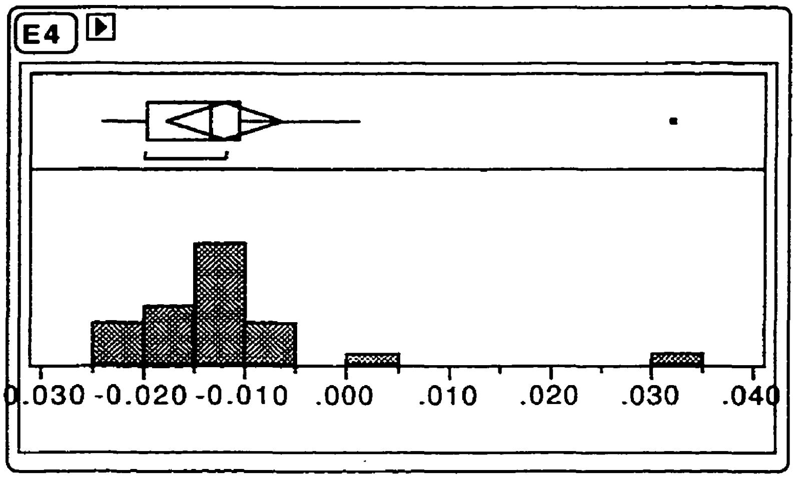



Figure 6f

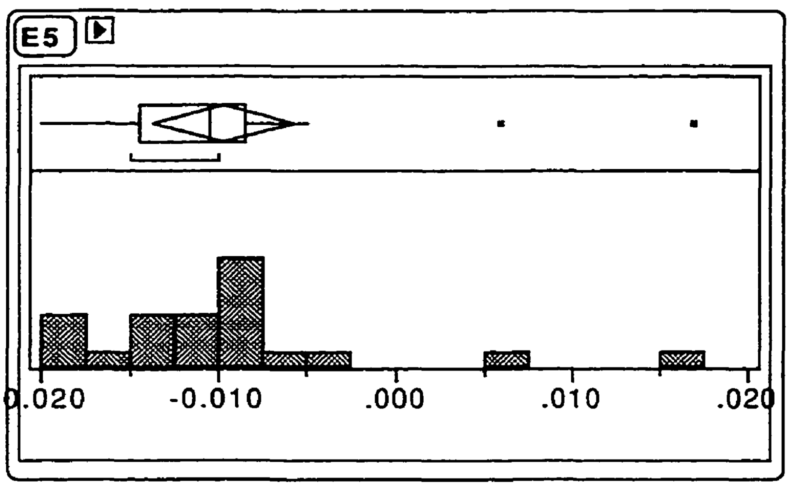


Figure 7: Control Charts Treating the Streams as a Subgroup

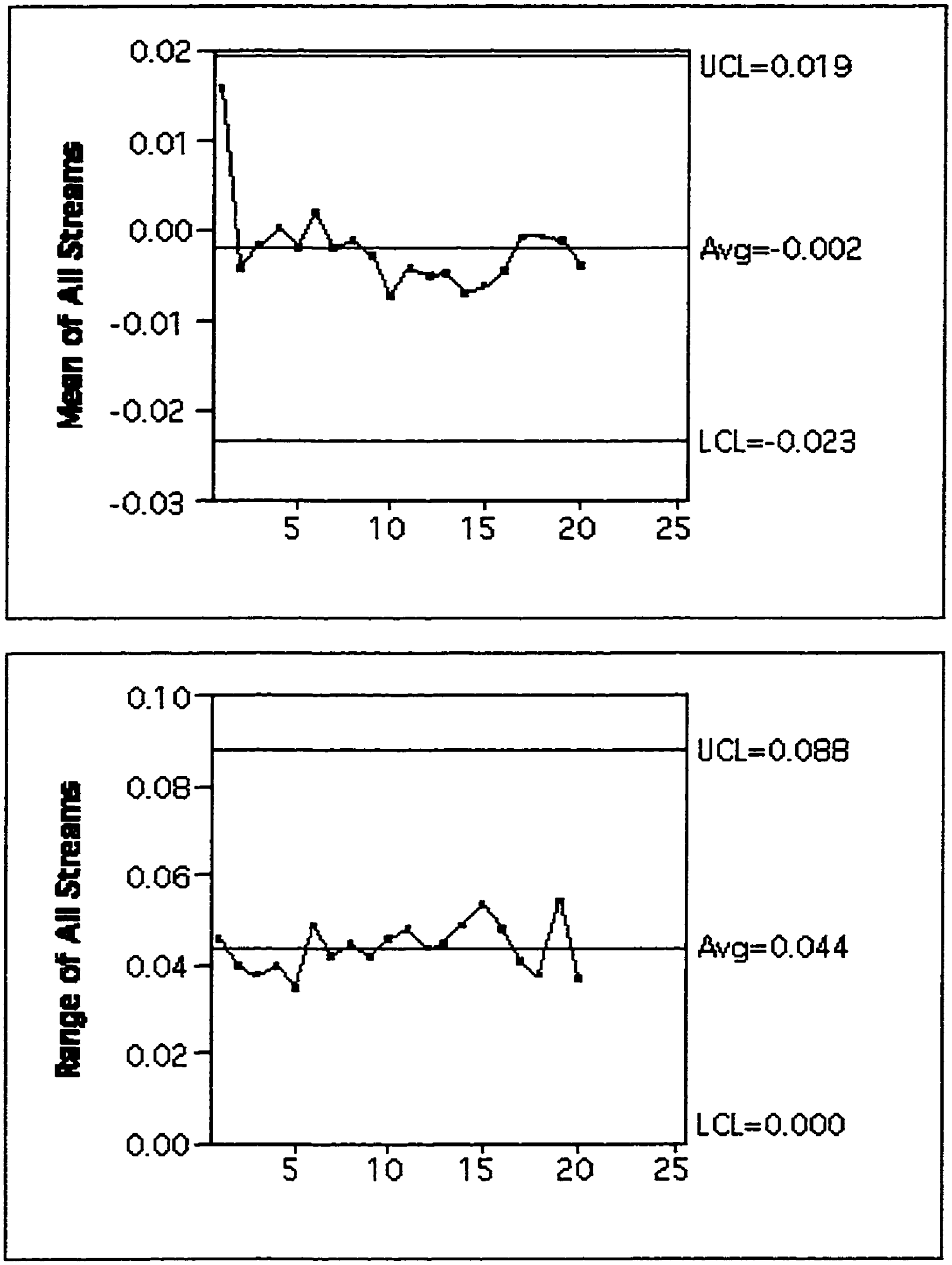
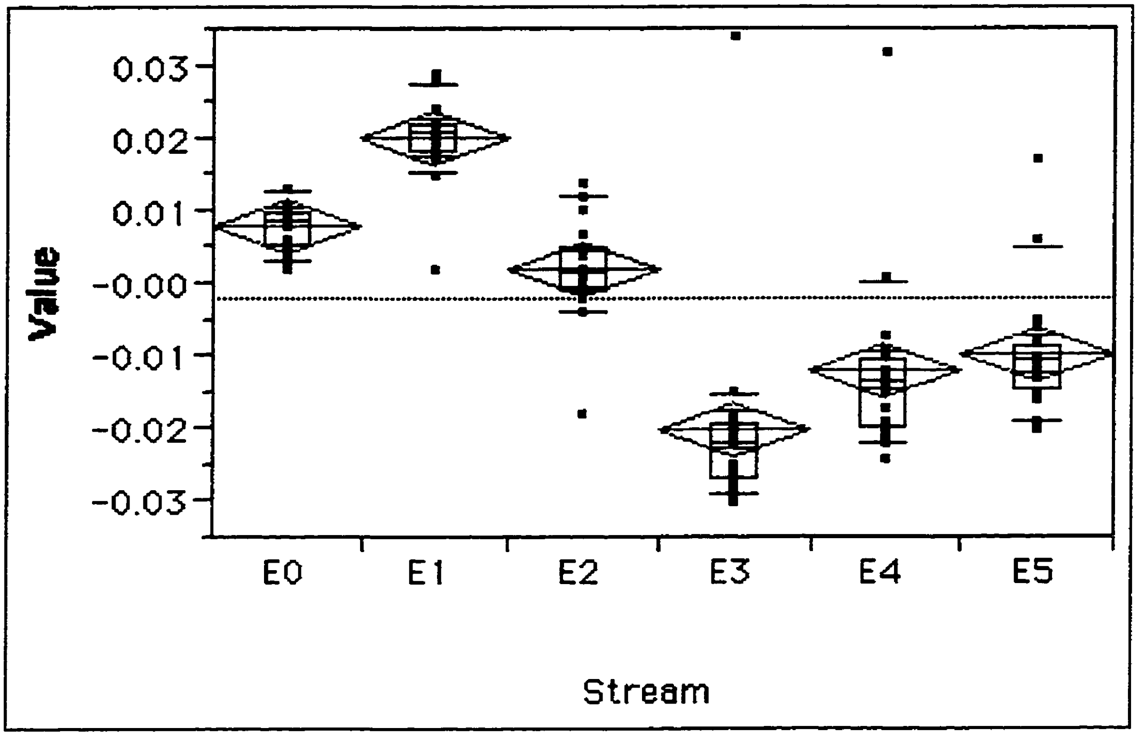


Figure 8: Means Diamonds and Quantile Boxes for Data Set 1




were summed for each stream. This is displayed in Table 18.

The distribution of the Friedman test statistic is usually approximated by the chi-square distribution. The following equation given by Sheskin (1997) is used to compute the chi-square approximation:

$$\chi^2_{s-1} = \frac{12}{ns(s+1)}\left[\sum_{j=1}^{s} R_j^2\right] - 3n(s+1) \tag{21}$$

where n = number of time periods

s = number of streams

j = stream number

$R_j$ = sum of ranks of stream j

Table 18

Rankings for Data Set 1

| Sample No. | E0 | | E1 | | E2 | | E3 | | E4 | | E5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Value | Rank | Value | Rank | Value | Rank | Value | Rank | Value | Rank | Value | Rank |
| 1 | 0.006 | 2 | 0.021 | 4 | 0.014 | 3 | 0.034 | 6 | 0.032 | 5 | -0.012 | 1 |
| 2 | 0.003 | 4 | 0.018 | 6 | 0.005 | 5 | -0.022 | 1 | -0.014 | 3 | -0.015 | 2 |
| 3 | 0.009 | 5 | 0.019 | 6 | 0.007 | 4 | -0.019 | 1 | -0.013 | 2.5 | -0.013 | 2.5 |
| 4 | 0.010 | 4.5 | 0.022 | 6 | 0.010 | 4.5 | -0.018 | 1 | -0.012 | 2 | -0.010 | 3 |
| 5 | 0.002 | 4.5 | 0.002 | 4.5 | -0.018 | 1 | -0.015 | 2 | 0.001 | 3 | 0.017 | 6 |
| 6 | 0.013 | 5 | 0.029 | 6 | 0.012 | 4 | -0.020 | 1 | -0.013 | 2 | -0.010 | 3 |
| 7 | 0.006 | 5 | 0.021 | 6 | -0.002 | 4 | -0.021 | 1 | -0.010 | 2 | -0.005 | 3 |
| 8 | 0.009 | 5 | 0.024 | 6 | 0.005 | 4 | -0.021 | 1 | -0.013 | 2 | -0.011 | 3 |
| 9 | 0.009 | 5 | 0.020 | 6 | 0.002 | 4 | -0.022 | 1 | -0.015 | 2 | -0.010 | 3 |
| 10 | 0.008 | 5 | 0.017 | 6 | 0.001 | 4 | -0.029 | 1 | -0.021 | 2 | -0.013 | 3 |
| 11 | 0.009 | 5 | 0.022 | 6 | 0.001 | 4 | -0.026 | 1 | -0.019 | 2 | -0.011 | 3 |
| 12 | 0.005 | 5 | 0.022 | 6 | -0.001 | 4 | -0.022 | 1 | -0.014 | 3 | -0.019 | 2 |
| 13 | 0.009 | 5 | 0.020 | 6 | 0.002 | 4 | -0.025 | 1 | -0.020 | 2 | -0.013 | 3 |
| 14 | 0.006 | 5 | 0.021 | 6 | 0.001 | 4 | -0.028 | 1 | -0.022 | 2 | -0.020 | 3 |
| 15 | 0.011 | 5 | 0.024 | 6 | 0.000 | 4 | -0.030 | 1 | -0.024 | 2 | -0.016 | 3 |
| 16 | 0.010 | 5 | 0.021 | 6 | 0.000 | 4 | -0.027 | 1 | -0.020 | 2 | -0.010 | 3 |
| 17 | 0.005 | 4.5 | 0.022 | 6 | 0.005 | 4.5 | -0.019 | 1 | -0.007 | 3 | -0.010 | 2 |
| 18 | 0.008 | 5 | 0.017 | 6 | -0.004 | 3 | -0.021 | 1 | -0.009 | 2 | 0.006 | 4 |
| 19 | 0.013 | 5 | 0.028 | 6 | 0.004 | 4 | -0.027 | 1 | -0.017 | 2 | -0.008 | 3 |
| 20 | 0.004 | 5 | 0.015 | 6 | -0.002 | 4 | -0.022 | 1 | -0.012 | 2 | -0.006 | 3 |
| | $R_0 = 94.5$ | | $R_1 = 116.5$ | | $R_2 = 77$ | | $R_3 = 26$ | | $R_4 = 47.5$ | | $R_5 = 58.5$ | |

By substituting the appropriate values into the equation, the value $\chi^2_{s-1} = 76.94$ was computed.

$$\chi^2_{s-1} = \frac{12}{(20)(6)(6+1)}\left[94.5^2 + 116.5^2 + 77^2 + 26^2 + 47.5^2 + 58.5^2\right] - (3)(20)(6+1)$$
$$= 76.94 \qquad (22)$$

For five degrees of freedom, the tabled critical 0.01 chi-square value is 15.08. Since 76.94 is greater than 15.08, it can be concluded that there is a significant difference between at least two of the six streams at the 0.01 level of significance.

When $\chi^2_{s-1}$ is significant, it does not indicate whether only two or if more than two streams differ significantly from another. In order to answer this question, comparisons contrasting specific streams with one another are necessary.

To do this, we must calculate $CD_F$, the minimum required difference between the sums of the ranks for any two streams in order for them to be considered different at the prespecified level of significance. The following equation is used to determine $CD_F$:

$$CD_F = z_{adj}\sqrt{\frac{ns(s+1)}{6}} \qquad (23)$$

In order to determine the value for $z_{adj}$, we must first establish the familywise Type I error rate. This was chosen to be equal to 0.05. Since there are $\frac{s(s-1)}{2} = \frac{6(5)}{2} = 15$ comparisons in total, the per comparison Type I error rate is $.05/15 = 0.003$. The z value satisfying $P\,[Z > z] = 0.003$ is 2.72 meaning that $z_{adj}$ is 2.72. Substituting the appropriate values into Equation (23), we get

$$CD_F = (2.72)\sqrt{\frac{(20)(6)(6+1)}{6}} = 32.18 \tag{24}$$

In order to use the $CD_F$, we must calculate the absolute value of the difference between the sums of the ranks of each pair of streams. Table 19 shows these "difference scores". A comparison is declared significant if the value from Table 19 is equal to or greater than $CD_F$. Nine of the fifteen comparisons are therefore significant.

We can check for the homogeneity of the variances by looking at the moving ranges of the streams. Figure 9 shows side-by-side boxplots of the moving ranges for all six streams. The shaded area represents 95%

Figure 9: Boxplots of the Moving Ranges

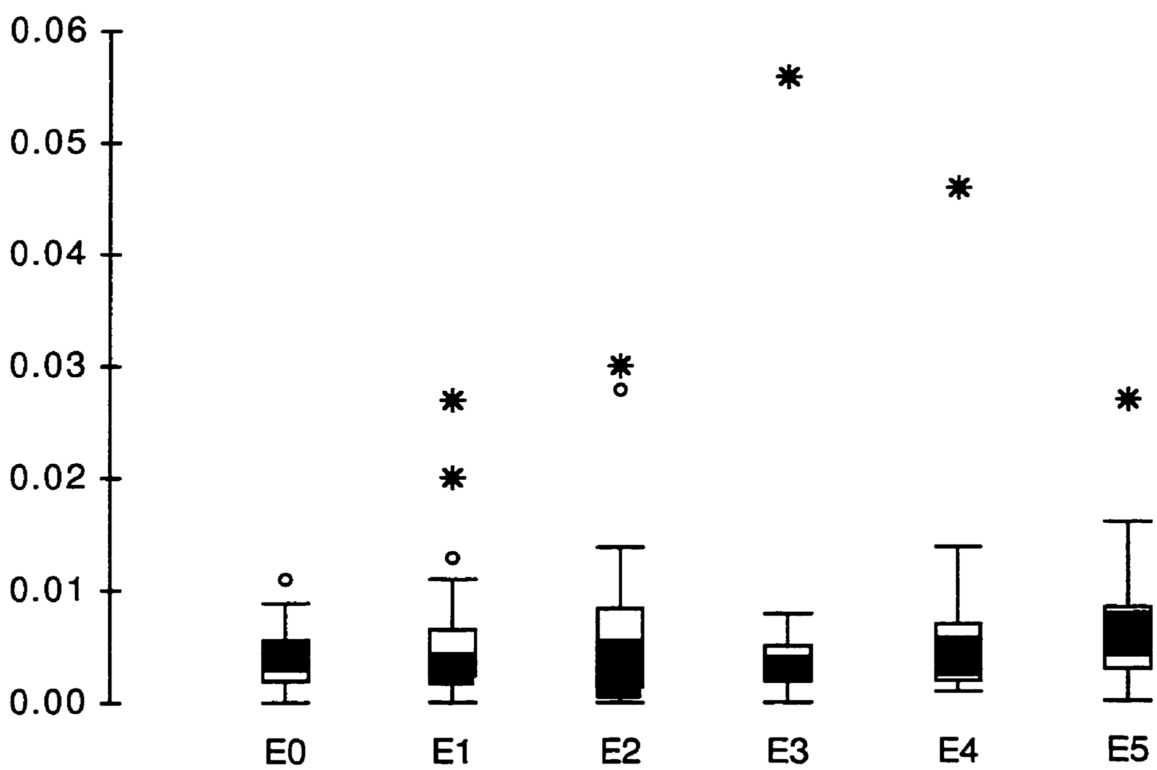


Table 19

Difference Scores Between Pairs of Sums of Ranks for Data Set 1

---

$| R_0 - R_1 | = | 94.5 - 116.5 | = 22$

$| R_0 - R_2 | = | 94.5 - 77 | = 17.5$

$| R_0 - R_3 | = | 94.5 - 26 | = 68.5^*$

$| R_0 - R_4 | = | 94.5 - 47.5 | = 47^*$

$| R_0 - R_5 | = | 94.5 - 58.5 | = 36^*$

$| R_1 - R_2 | = | 116.5 - 77 | = 39.5^*$

$| R_1 - R_3 | = | 116.5 - 26 | = 90.5^*$

$| R_1 - R_4 | = | 116.5 - 47.5 | = 69^*$

$| R_1 - R_5 | = | 116.5 - 58.5 | = 58^*$

$| R_2 - R_3 | = | 77 - 26 | = 51^*$

$| R_2 - R_4 | = | 77 - 47.5 | = 29.5$

$| R_2 - R_5 | = | 77 - 58.5 | = 18.5$

$| R_3 - R_4 | = | 26 - 47.5 | = 21.5$

$| R_3 - R_5 | = | 26 - 58.5 | = 32.5^*$

$| R_4 - R_5 | = | 47.5 - 58.5 | = 11$

*Denotes a significant difference between streams at 0.05 familywise Type I error rate

---

confidence intervals for the medians. From the general appearance (except for the outlier values), we can conclude that the streams appear to have similar variability.

### 5.1.3 Overlay Plots

Another way we can see the differences in the streams is to use overlay plots, shown in Figures 10a and 10b. In Figure 10a, the points are connected for all of the streams showing how the values change for a particular stream in comparison to the other streams. In Figure 10b, the points are connected for only streams E1 and E3. Note that even though both E1 and E3 have the same target value, they clearly behave quite differently from each other.

### 5.1.4 Control Charts of the Four Methods

Figures 11a to 11d show the control charts of the four different multiple-stream monitoring techniques. We see that there are no sample periods in which both the maximum and minimum points plot inside the control limits for Methods 1 and 3. Both the maximum and minimum points plot within the control limits only five times for Method 2 and four times for Method 4.

We also examine the charts for the number of consecutive times a particular stream plots as the maximum, or the minimum. It should be noted that this characteristic is independent of the four methods being examined as it is purely a property of the sequence of plotted values and not the control limits.

For a process with six streams, Table 20 (Table 2 reproduced for convenience) says that it is unusual for a stream to be the largest or smallest value five consecutive times. The control charts show that E1 is the maximum stream 15 consecutive times (starting in time period 6), and E3 is the minimum stream 14 times in a row (beginning in time period 6). This shows us once again that the six streams are not identical.

Figure 10a: Overlay Plot (All Streams Connected)

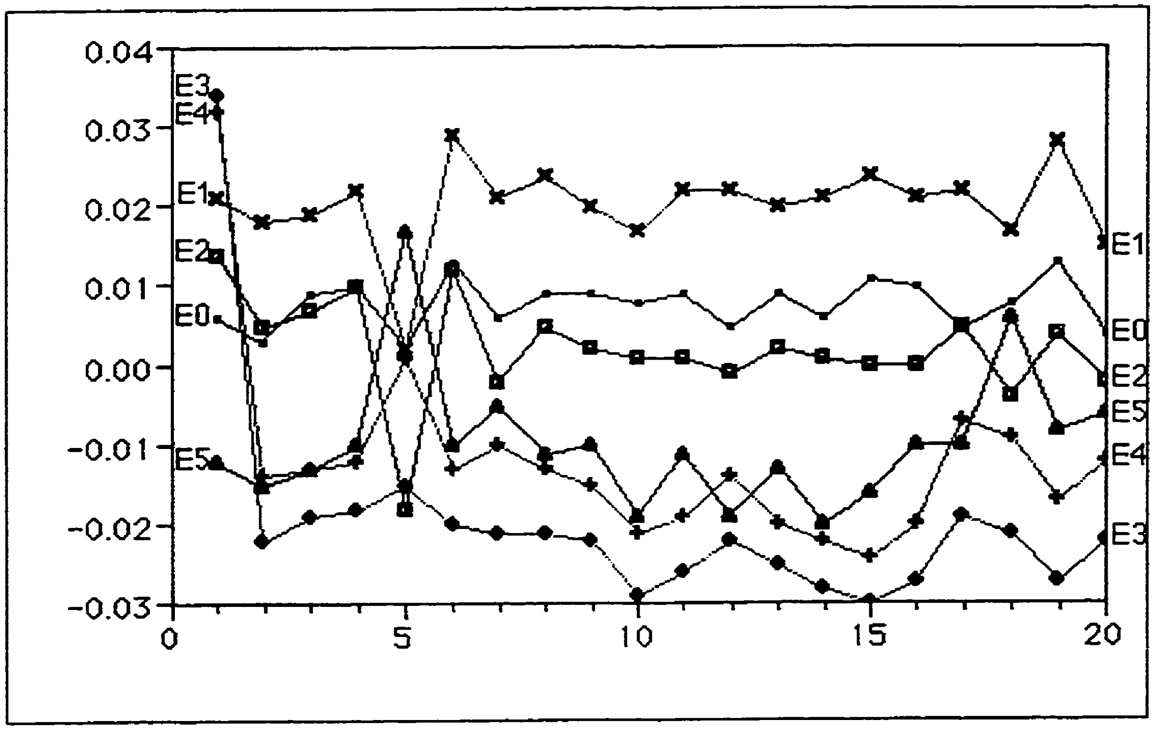



Figure 10b: Overlay Plot (Only E1 and E3 Connected)

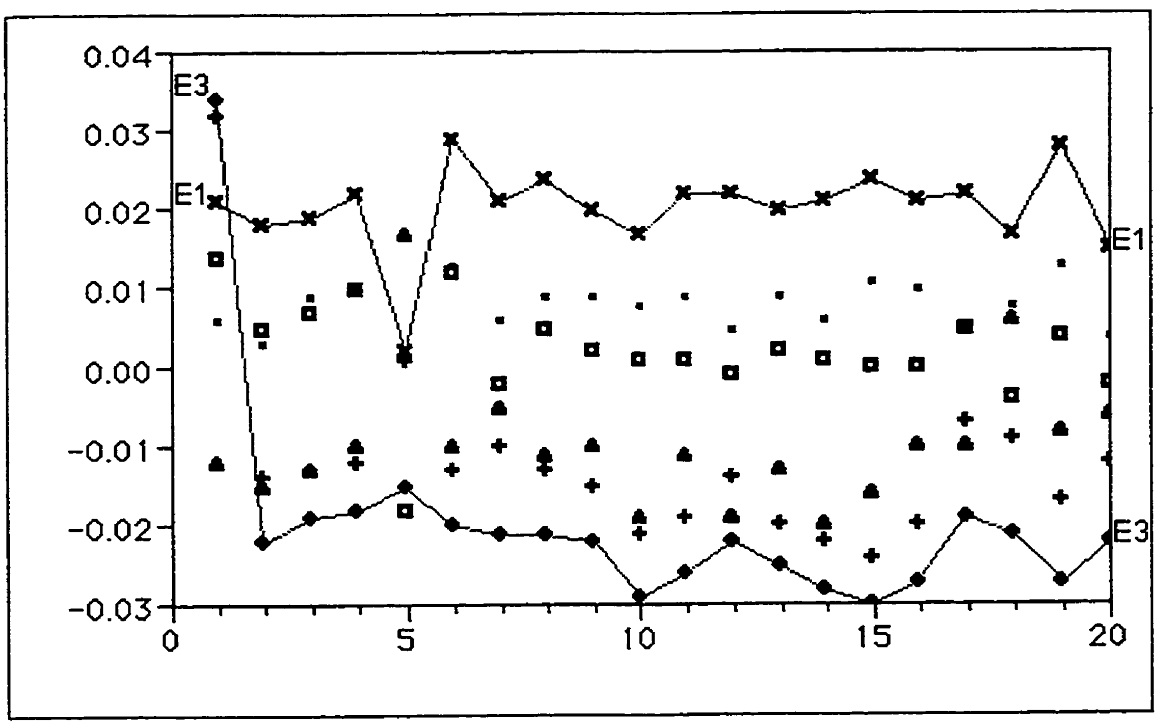


Figure 11a: Group Control Chart (Method 1)

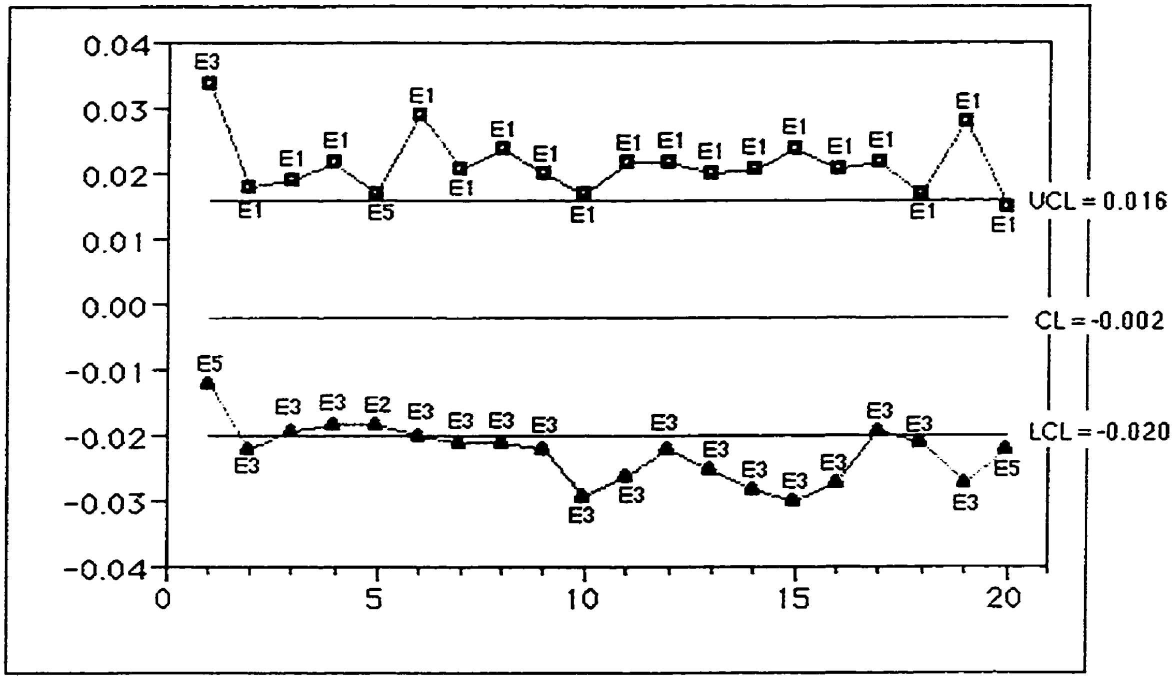



Figure 11b: Group Control Chart (Method 2)

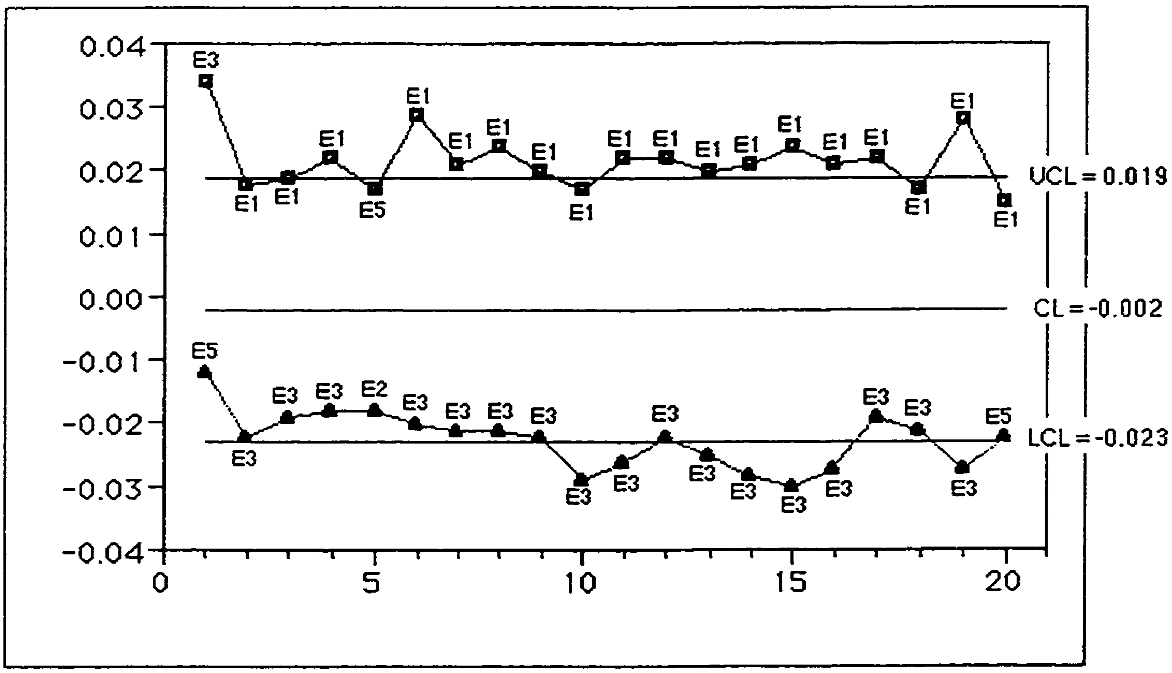


Figure 11c: Group Control Chart (Method 3)

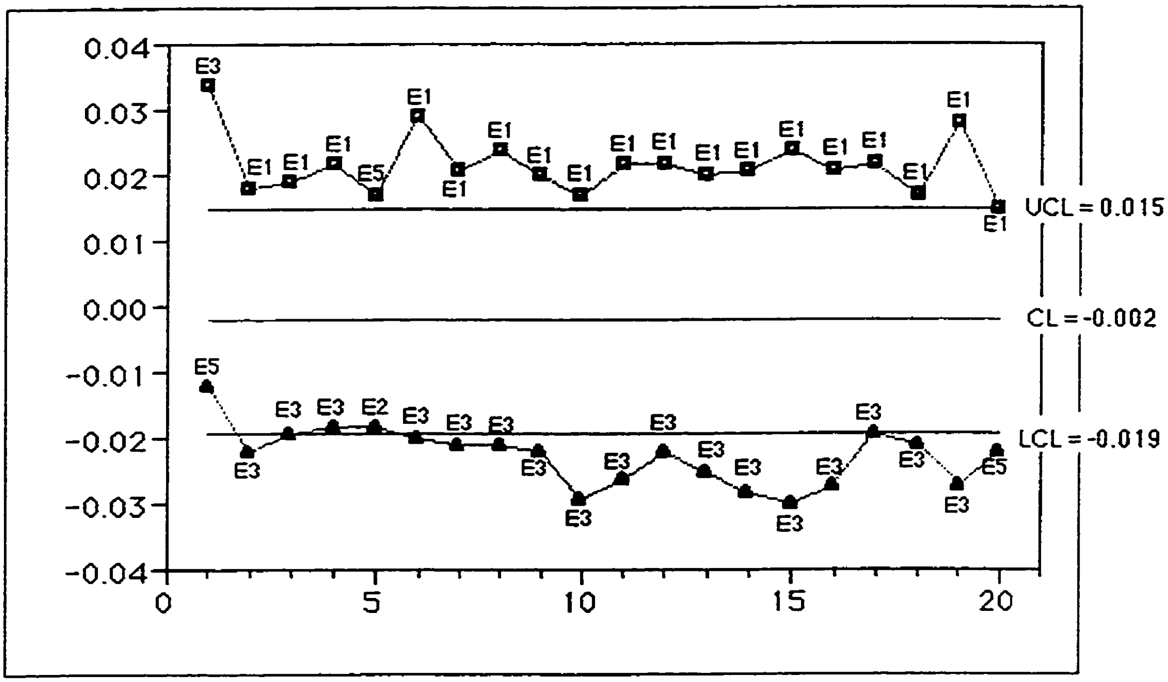



Figure 11d: Group Control Chart (Method 4)

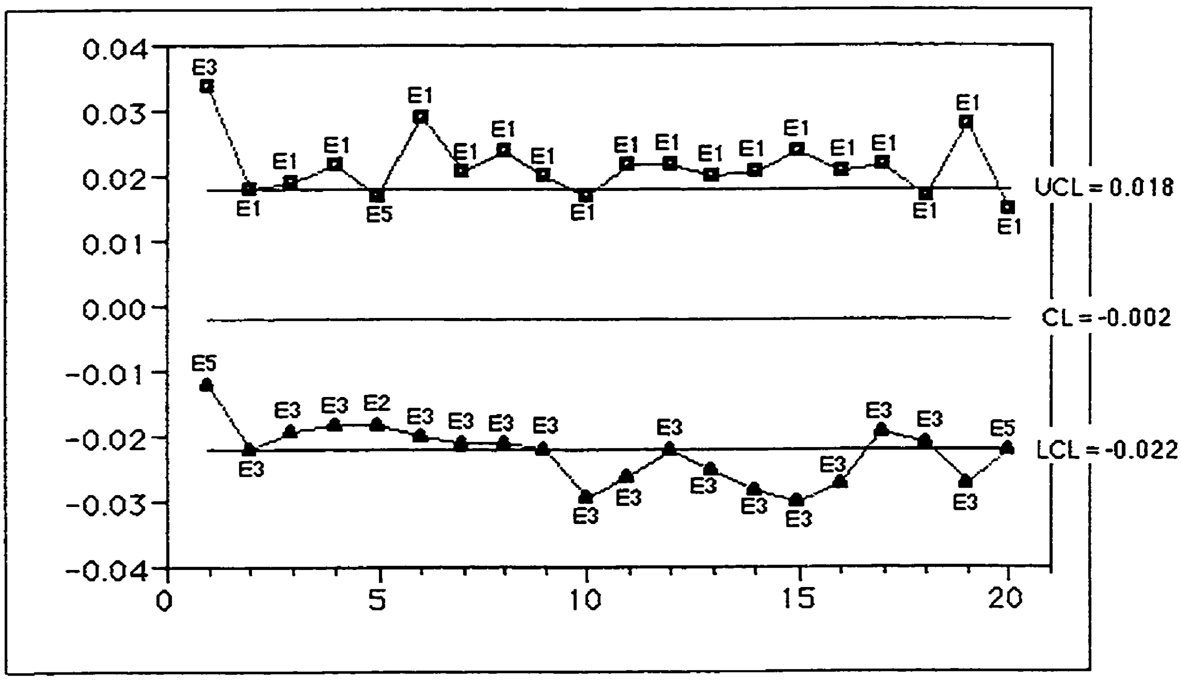


Table 20

Recommended Pairings for s and r

| s | r | One-sided In-control ARL |
|---|---|---|
| 2 | 9 | 511 |
| 3 | 7 | 1093 |
| 4 | 6 | 1365 |
| 5 | 5 | 781 |
| 6 | 5 | 1555 |
| 7 | 4 | 400 |
| 8 | 4 | 585 |
| 9 | 4 | 820 |
| 10 | 4 | 1111 |

The number of streams is denoted by s. The number of consecutive times that a particular stream is the largest/smallest value is denoted by r.

### 5.1.5 Control Charts for Each Stream

Control charts for each individual stream are shown in Figures 12a to 12f. Stream E0 appears to be in-control, but all the other streams have a point plotting outside the control limits on the X chart. The out-of-control point occurs in time period 5 for Streams E1, E2, and E5. The out-of-control point occurs in time period 1 for Streams E3 and E4. Note that even though three streams have an out-of-control point at time period 5, there is nothing unusual about time period 5 in Figures 11a to 11d. This is because the group control chart only plots the maximum and minimum streams for each time period.

Figure 12a: X and MR Control Charts for Stream E0

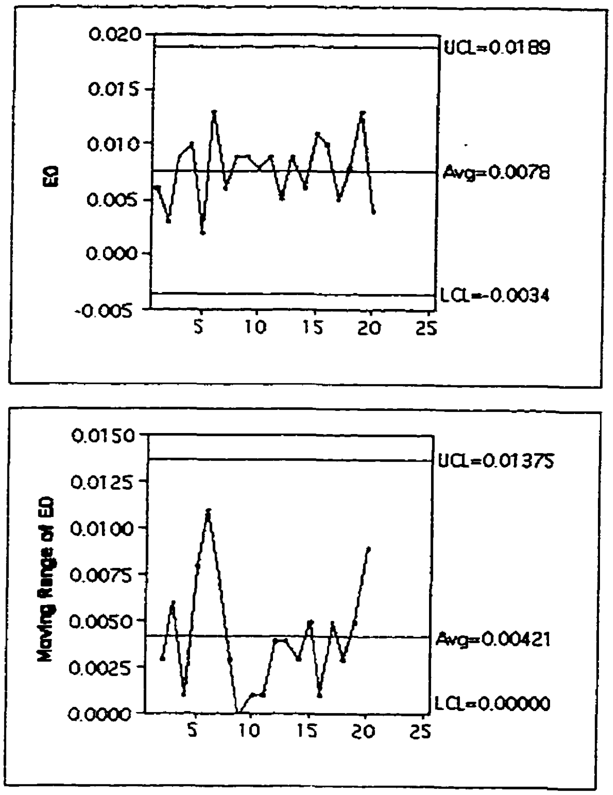



Figure 12b: X and MR Control Charts for Stream E1

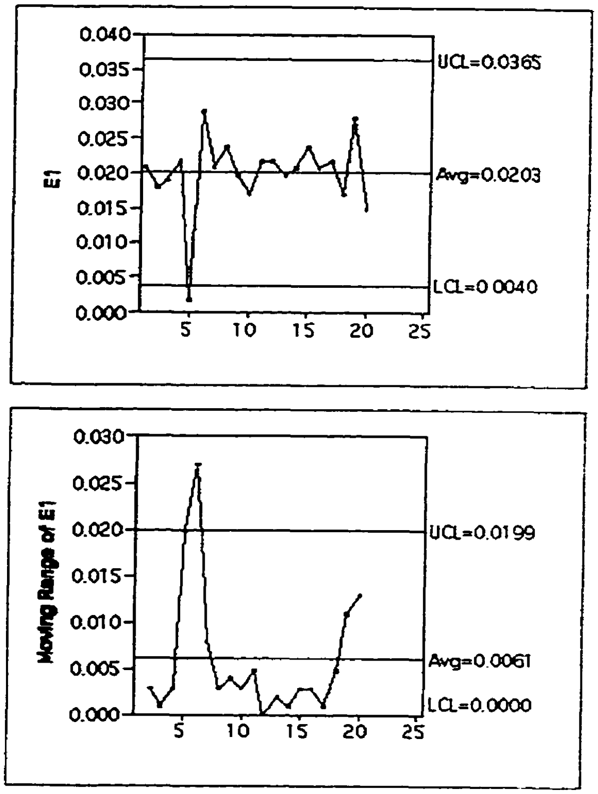



Figure 12c: X and MR Control Charts for Stream E2

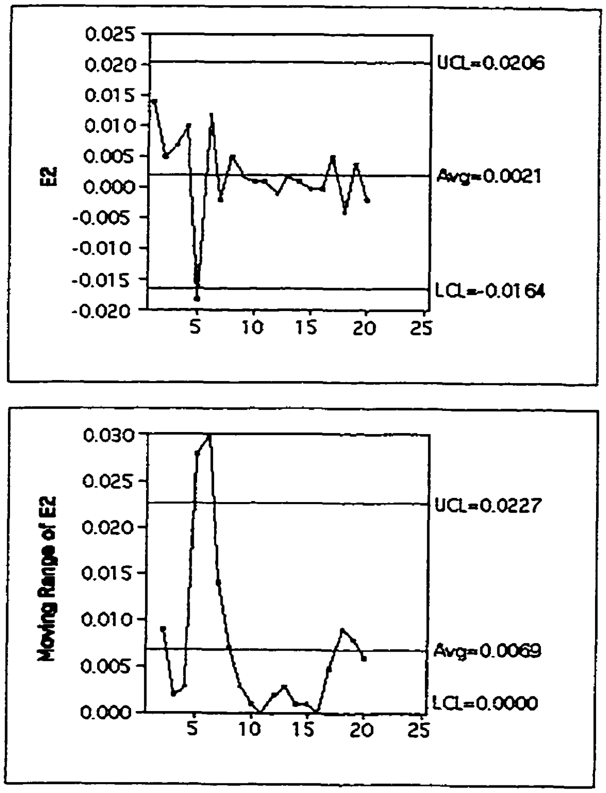



Figure 12d: X and MR Control Charts for Stream E3

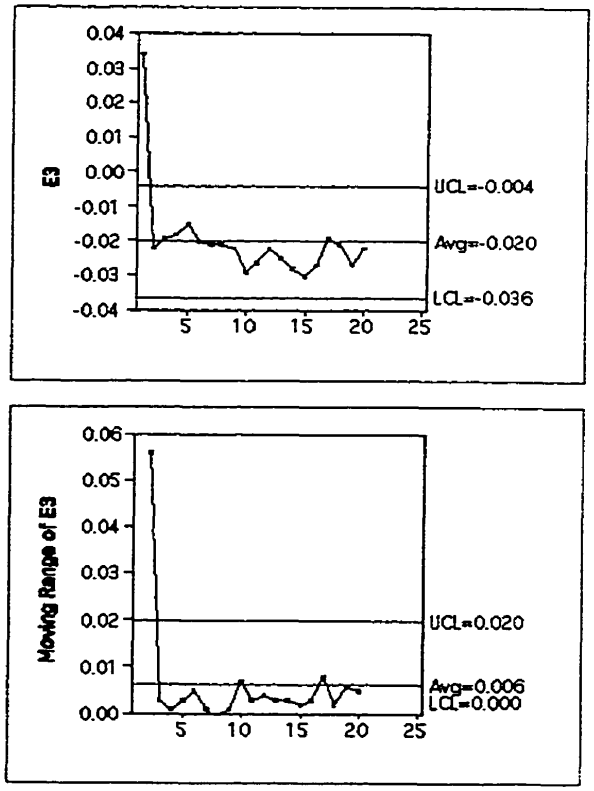


Figure 12e: X and MR Control Charts for Stream E4

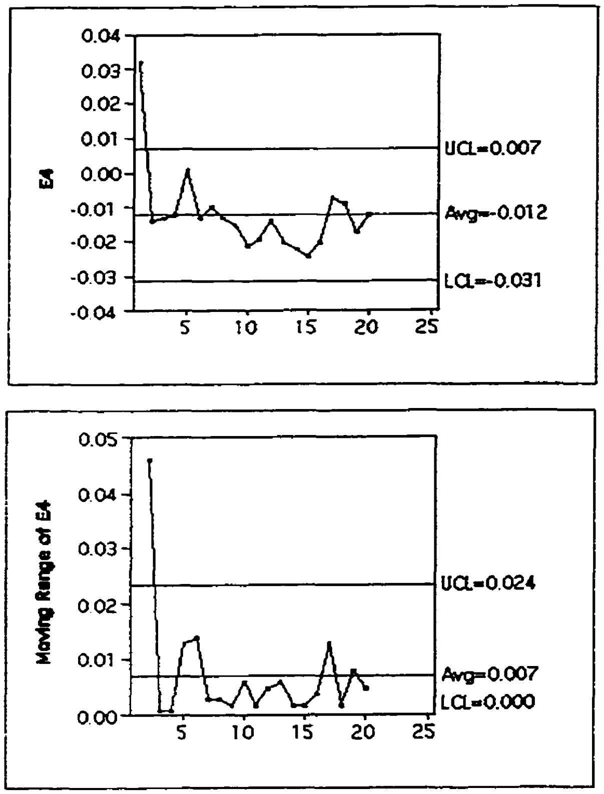
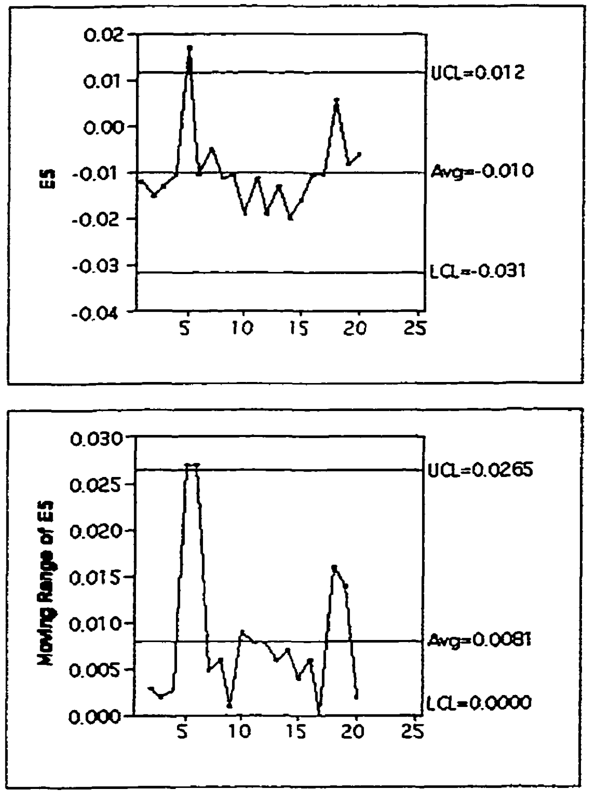



Figure 12f: X and MR Control Charts for Stream E5




5.1.6 <u>The Disadvantage of Using the Streams as a Subgroup</u>

We get very different results depending upon the method used to monitor the process. It was shown that if we treat the streams as a subgroup, the process will appear to be in control. However, the group control charts had most of their points outside the control limits. This shows why it is misleading to use streams as the subgroup. We have shown that the process is really not in control because the six streams are behaving differently. The $\bar{X}$ chart averages out the six streams thereby hiding the differences in the streams.

## 5.2 Data Set 2

Table 21 shows the data from "Data Set 2". The data are reported as measurements taken from five locations from a single part. The process therefore has five streams and the streams are labeled X1 to X5. Figure 13 displays a histogram and boxplot of the entire data set and Figures 14a to 14e show the data by individual stream. These histograms and boxplots show that the data are not normally distributed. The only extreme outlier in the data is the value of 10.7 for Stream X5 (in bold type).

Table 21

Data Set 2

| Sample No. | X1 | X2 | X3 | X4 | X5 |
|---|---|---|---|---|---|
| 1 | 11.1 | 11.3 | 10.5 | 9.6 | 9.4 |
| 2 | 11.0 | 11.1 | 9.5 | 9.1 | 9.3 |
| 3 | 11.7 | 10.8 | 10.1 | 10.2 | 9.8 |
| 4 | 12.1 | 11.5 | 9.3 | 9.6 | 9.4 |
| 5 | 11.2 | 11.0 | 10.6 | 10.8 | 10.1 |
| 6 | 11.1 | 10.4 | 11.6 | 10.9 | **10.7** |
| 7 | 11.5 | 10.1 | 10.4 | 9.5 | 9.2 |
| 8 | 12.1 | 10.5 | 10.6 | 9.4 | 9.1 |
| 9 | 11.8 | 11.1 | 10.1 | 10.3 | 9.5 |
| 10 | 11.4 | 11.1 | 10.1 | 10.3 | 9.1 |
| 11 | 10.3 | 10.5 | 9.3 | 9.1 | 9.2 |
| 12 | 11.3 | 11.1 | 10.1 | 10.2 | 9.5 |
| 13 | 11.8 | 11.6 | 11.1 | 9.9 | 9.8 |
| 14 | 11.8 | 11.1 | 9.7 | 9.1 | 9.3 |
| 15 | 11.3 | 11.3 | 9.4 | 9.1 | 9.3 |
| 16 | 11.1 | 10.3 | 9.8 | 10.2 | 9.6 |
| 17 | 11.6 | 11.1 | 10.1 | 10.3 | 9.7 |
| 18 | 10.4 | 10.4 | 10.6 | 9.9 | 9.7 |
| 19 | 11.3 | 10.9 | 9.9 | 9.7 | 9.8 |
| 20 | 11.4 | 11.3 | 10.1 | 9.4 | 9.7 |

Figure 13: Histogram and Boxplot of Data Set 2

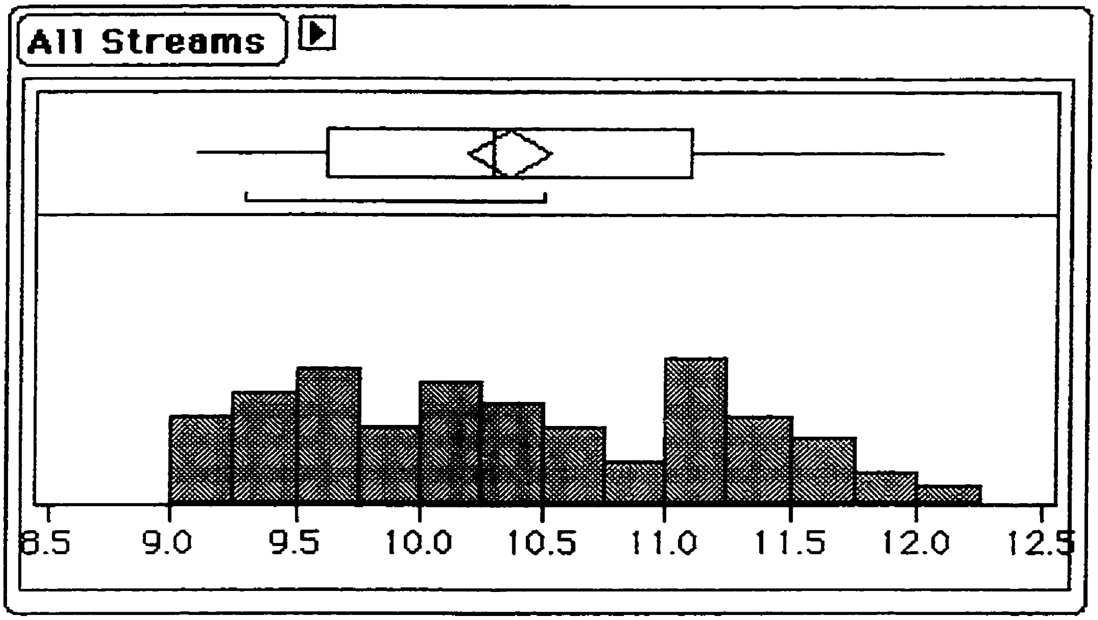



Figure 14a

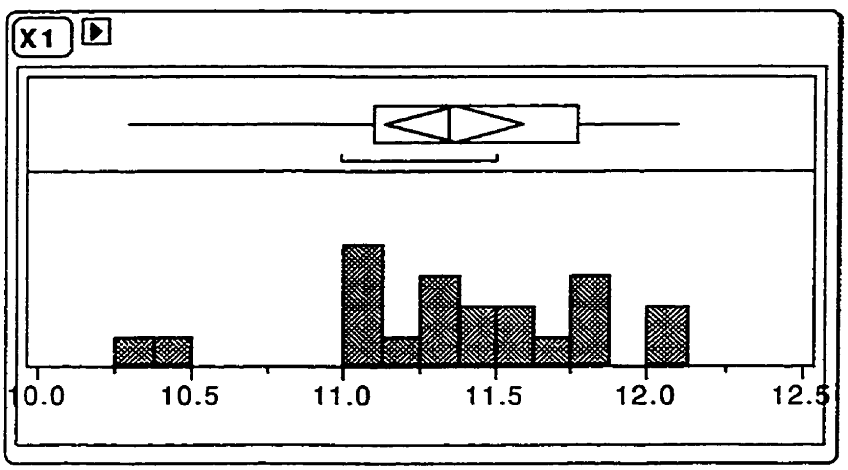


Figure 14b

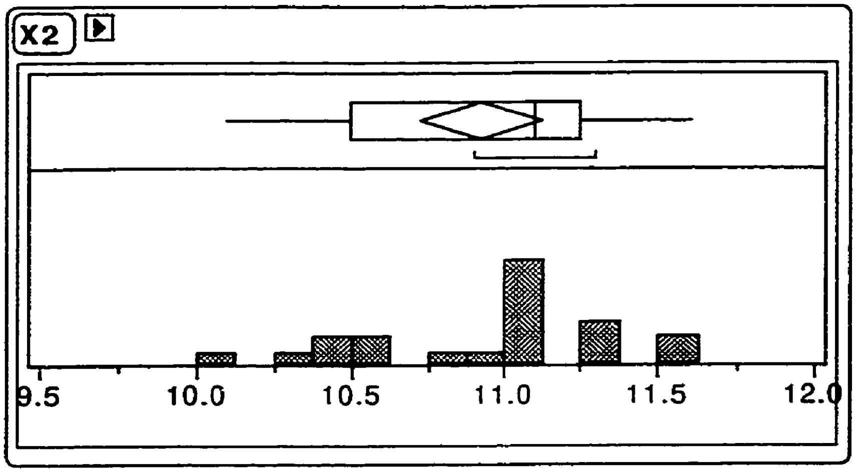



Figure 14c

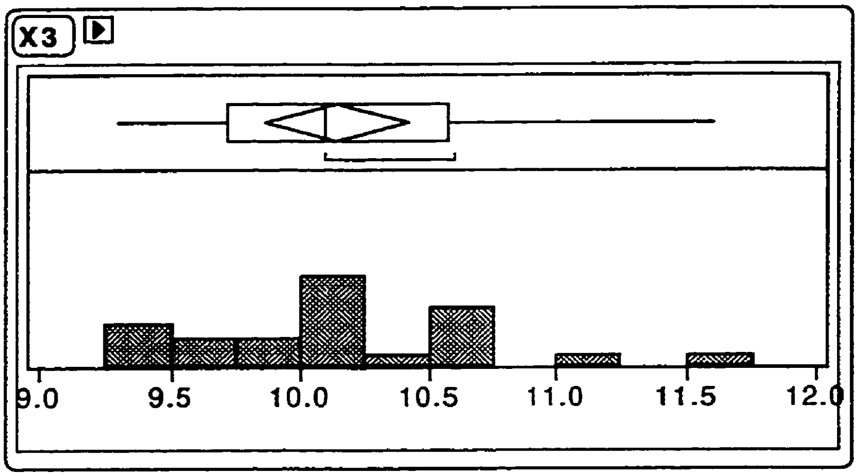



Figure 14d

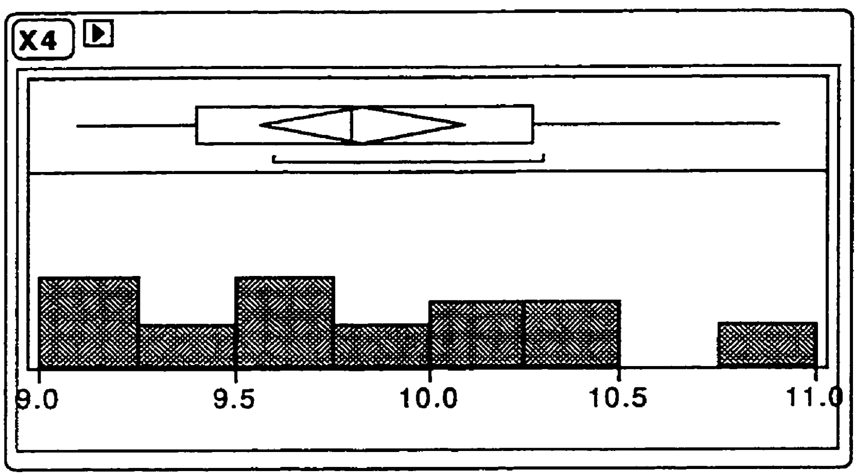
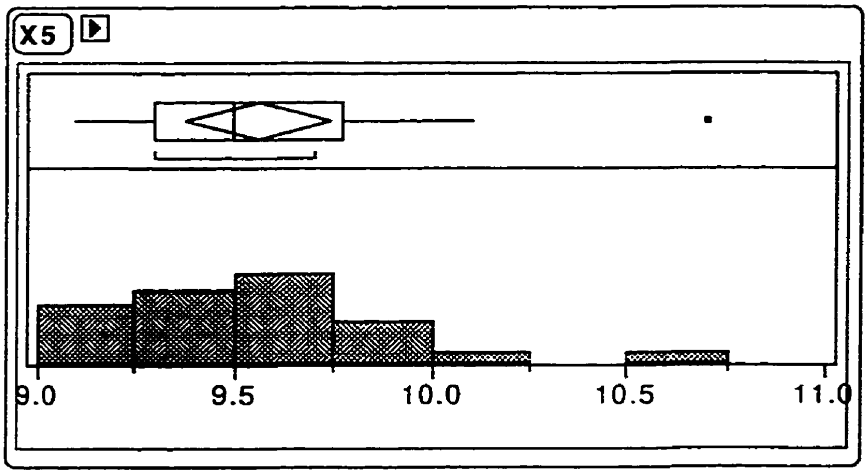


Figure 14e




### 5.2.1 Testing the Equality of the Streams

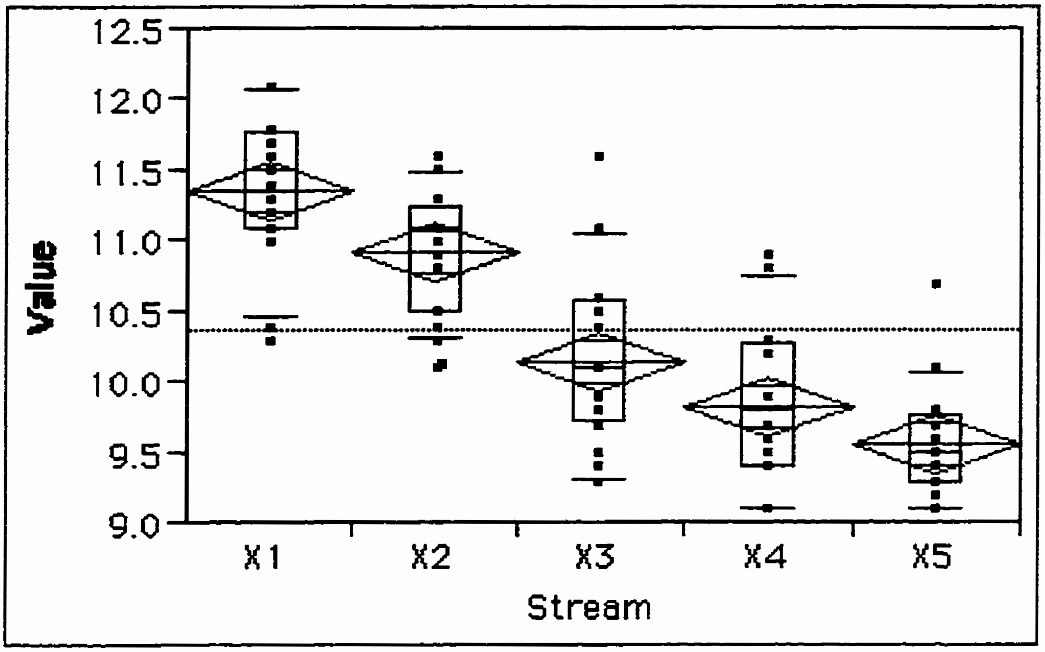
Means diamonds and quantile boxes for the five streams are shown in Figure 15 to help determine the equality of the streams. Figure 15 shows that the streams are not identical with respect to their location. Stream X1 appears to be the highest and Stream X5 appears to be the lowest of the five streams. The Friedman two-way analysis of variance by ranks was again used to test the equality of the streams. The data were ranked for each time period and the ranks were summed for each stream (see Table 22). Substituting the appropriate values into Equation (21), the value $\chi^2_{s-1} = 56.04$ was computed.

$$\chi^2_{s-1} = \frac{12}{(20)(5)(5+1)}\left[94^2 + 78^2 + 57^2 + 43^2 + 28^2\right] - (3)(20)(5+1)$$
$$= 56.04 \qquad (25)$$

The tabled critical 0.01 chi-square value is 13.28 (four degrees of freedom). Since 56.04 is greater than 13.28, there is a significant difference between at least two of the five streams at the 0.01 level of significance.

Choosing a familywise Type I error rate of 0.05, the per comparison

Figure 15: Means Diamonds and Quantile Boxes for Data Set 2




Type I error rate is equal to $\frac{0.05}{10} = 0.005$ (there are $\frac{(5)(4)}{2} = 10$ total comparisons). The z value satisfying $P[Z > z] = 0.005$ is 2.5758. Substituting the appropriate values into Equation (23), we get

$$CD_F = (2.5758)\sqrt{\frac{(20)(5)(5+1)}{6}} = 25.758 \qquad (26)$$

Table 23 shows the absolute value of the difference between the sums of the ranks of each pair of streams. Six out of ten comparisons are significant.

We can again check for the homogeneity of the variances by looking at the moving ranges of the streams. Figure 16 shows side-by-side boxplots of the moving ranges for all five streams. Since the boxplots are all at about the same location, we can conclude that the streams appear to have similar variability.

Table 22

Rankings for Data Set 2

| Sample No. | X1 | | X2 | | X3 | | X4 | | X5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Value | Rank | Value | Rank | Value | Rank | Value | Rank | Value | Rank |
| 1 | 11.1 | 4 | 11.3 | 5 | 10.5 | 3 | 9.6 | 2 | 9.4 | 1 |
| 2 | 11.0 | 4 | 11.1 | 5 | 9.5 | 3 | 9.1 | 1 | 9.3 | 2 |
| 3 | 11.7 | 5 | 10.8 | 4 | 10.1 | 2 | 10.2 | 3 | 9.8 | 1 |
| 4 | 12.1 | 5 | 11.5 | 4 | 9.3 | 1 | 9.6 | 3 | 9.4 | 2 |
| 5 | 11.2 | 5 | 11.0 | 4 | 10.6 | 2 | 10.8 | 3 | 10.1 | 1 |
| 6 | 11.1 | 4 | 10.4 | 1 | 11.6 | 5 | 10.9 | 3 | 10.7 | 2 |
| 7 | 11.5 | 5 | 10.1 | 3 | 10.4 | 4 | 9.5 | 2 | 9.2 | 1 |
| 8 | 12.1 | 5 | 10.5 | 3 | 10.6 | 4 | 9.4 | 2 | 9.1 | 1 |
| 9 | 11.8 | 5 | 11.1 | 4 | 10.1 | 2 | 10.3 | 3 | 9.5 | 1 |
| 10 | 11.4 | 5 | 11.1 | 4 | 10.1 | 2 | 10.3 | 3 | 9.1 | 1 |
| 11 | 10.3 | 4 | 10.5 | 5 | 9.3 | 3 | 9.1 | 1 | 9.2 | 2 |
| 12 | 11.3 | 5 | 11.1 | 4 | 10.1 | 2 | 10.2 | 3 | 9.5 | 1 |
| 13 | 11.8 | 5 | 11.6 | 4 | 11.1 | 3 | 9.9 | 2 | 9.8 | 1 |
| 14 | 11.8 | 5 | 11.1 | 4 | 9.7 | 3 | 9.1 | 1 | 9.3 | 2 |
| 15 | 11.3 | 4.5 | 11.3 | 4.5 | 9.4 | 3 | 9.1 | 1 | 9.3 | 2 |
| 16 | 11.1 | 5 | 10.3 | 4 | 9.8 | 2 | 10.2 | 3 | 9.6 | 1 |
| 17 | 11.6 | 5 | 11.1 | 4 | 10.1 | 2 | 10.3 | 3 | 9.7 | 1 |
| 18 | 10.4 | 3.5 | 10.4 | 3.5 | 10.6 | 5 | 9.9 | 2 | 9.7 | 1 |
| 19 | 11.3 | 5 | 10.9 | 4 | 9.9 | 3 | 9.7 | 1 | 9.8 | 2 |
| 20 | 11.4 | 5 | 11.3 | 4 | 10.1 | 3 | 9.4 | 1 | 9.7 | 2 |
| | $R_1 = 94$ | | $R_2 = 78$ | | $R_3 = 57$ | | $R_4 = 43$ | | $R_5 = 28$ | |

Table 23

Difference Scores Between Pairs of Sums of Ranks for Data Set 2

| | | |
|---|---|---|
| $\lvert R_1 - R_2 \rvert$ | $= \lvert 94 - 78 \rvert$ | $= 16$ |
| $\lvert R_1 - R_3 \rvert$ | $= \lvert 94 - 57 \rvert$ | $= 37^*$ |
| $\lvert R_1 - R_4 \rvert$ | $= \lvert 94 - 43 \rvert$ | $= 51^*$ |
| $\lvert R_1 - R_5 \rvert$ | $= \lvert 94 - 28 \rvert$ | $= 66^*$ |
| $\lvert R_2 - R_3 \rvert$ | $= \lvert 78 - 57 \rvert$ | $= 21$ |
| $\lvert R_2 - R_4 \rvert$ | $= \lvert 78 - 43 \rvert$ | $= 35^*$ |
| $\lvert R_2 - R_5 \rvert$ | $= \lvert 78 - 28 \rvert$ | $= 50^*$ |
| $\lvert R_3 - R_4 \rvert$ | $= \lvert 57 - 43 \rvert$ | $= 14$ |
| $\lvert R_3 - R_5 \rvert$ | $= \lvert 57 - 28 \rvert$ | $= 29^*$ |
| $\lvert R_4 - R_5 \rvert$ | $= \lvert 43 - 28 \rvert$ | $= 15$ |

*Denotes a significant difference between streams at 0.05 familywise Type I error rate

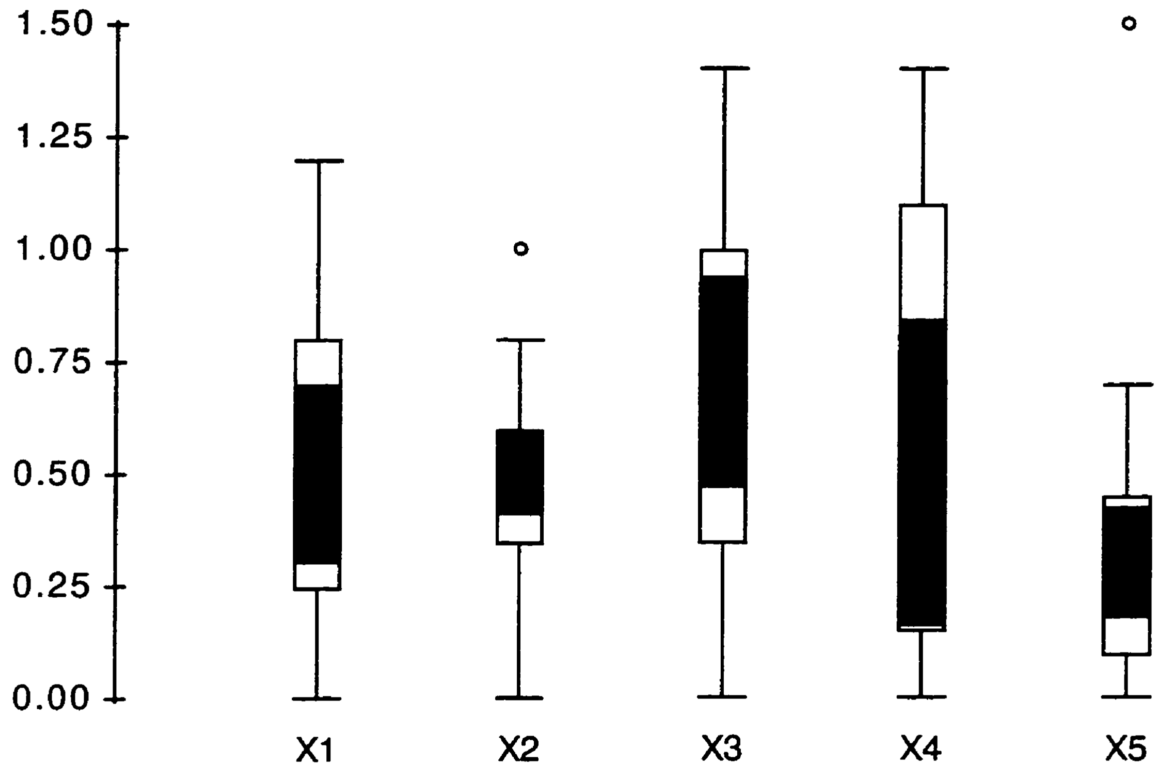
### 5.2.2 Overlay Plots

Figure 17a is an overlay plot in which the points are connected for all five streams. Figure 17b connects the points only for Streams X1 and X5. This clearly shows that there is a difference between those two streams. This is unfortunate because X1 and X5 have the same target value.

### 5.2.3 Control Charts of the Four Methods

Figures 18a to 18d show the control charts of the four different multiple-stream monitoring methods. If we focus initially on points plotting outside the control limits, we see that the control charts using Methods 1, 2, and 4 signal

Figure 16: Boxplots of the Moving Ranges




the potential presence of assignable causes at periods 4 and 8 where the maximum value from all streams plots above the UCL. Method 3 gives an out-of-control signal at periods 4 and 8 as well, but it has three additional signals at periods 9, 13, and 14 where again the maximum point plots above the UCL. There are no out-of-control signals for any method in which the minimum point plots below the LCL.

We also examine the chart for the number of consecutive times a particular stream plots as the maximum, or the minimum. For a process with five streams, Table 20 tells us that it is unusual for a stream to be the largest or smallest value five consecutive times. The control charts show that X1 is the maximum stream six consecutive times beginning in period 12 (note that X1

Figure 17a: Overlay Plot (All Streams Connected)

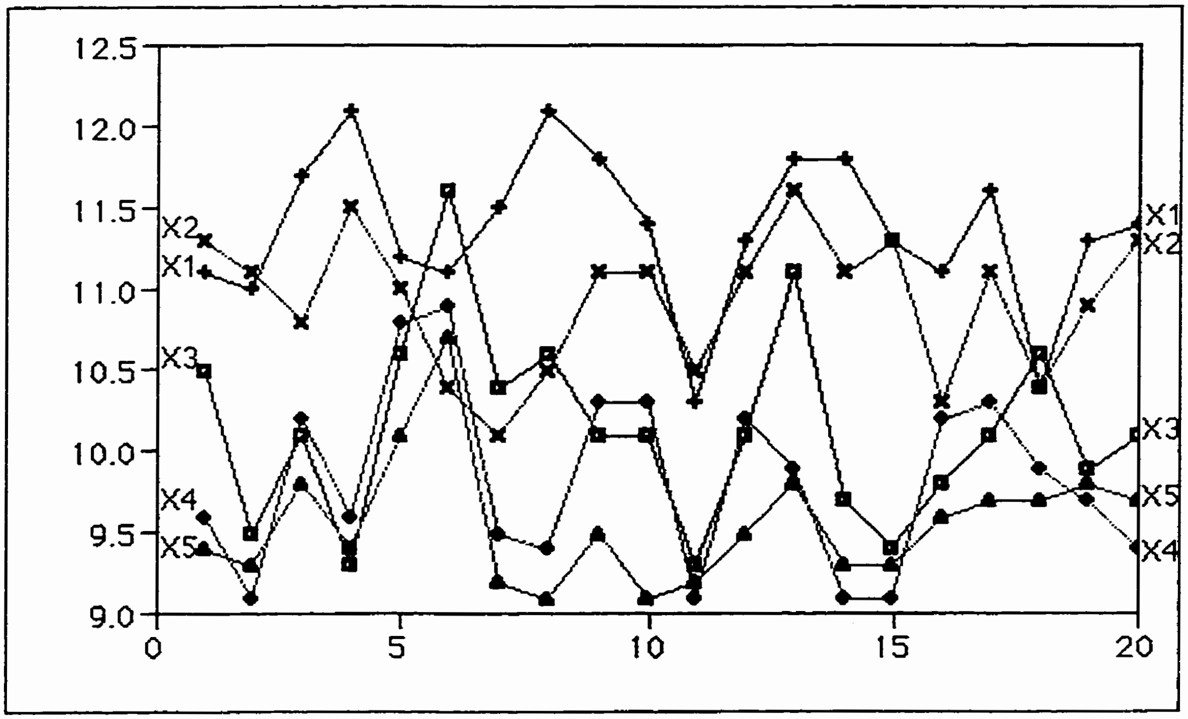



Figure 17b: Overlay Plot (Only X1 and X5 Connected)

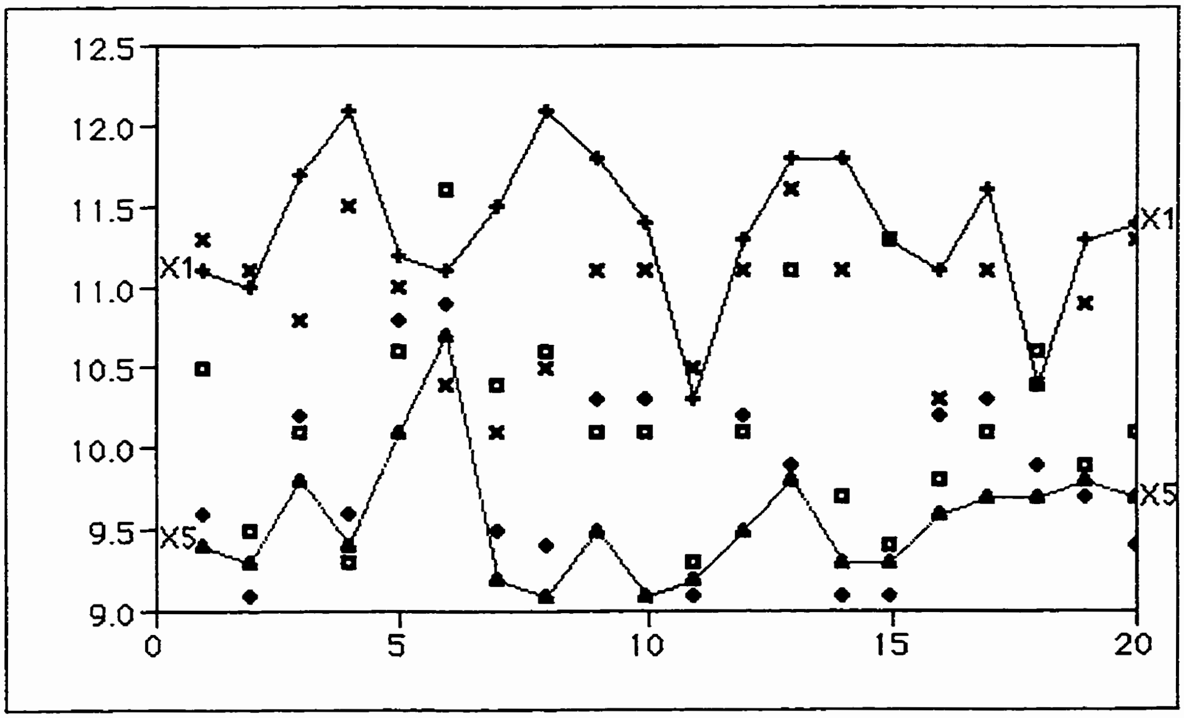


Figure 18a: Group Control Chart (Method 1)

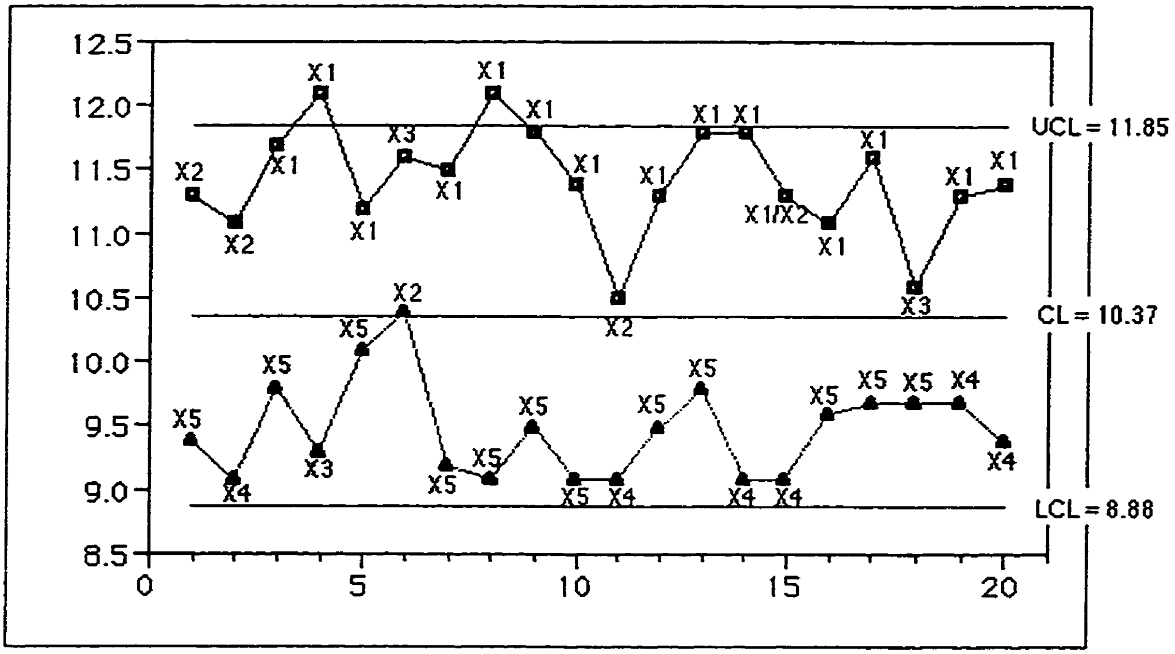



Figure 18b: Group Control Chart (Method 2)

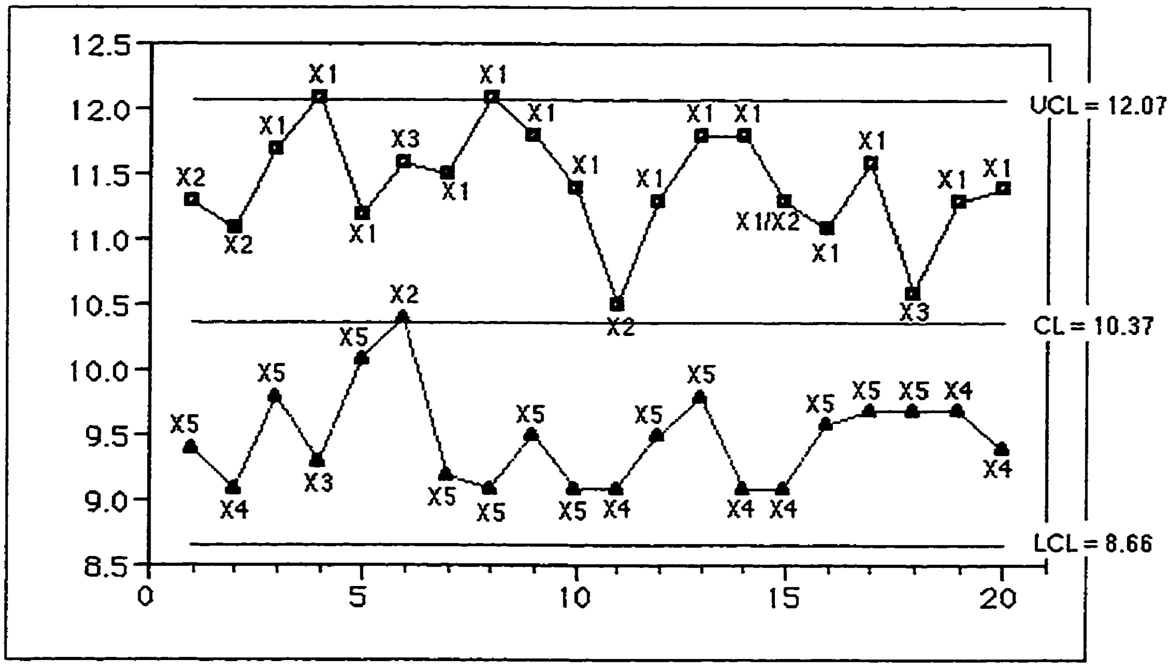


Figure 18c: Group Control Chart (Method 3)

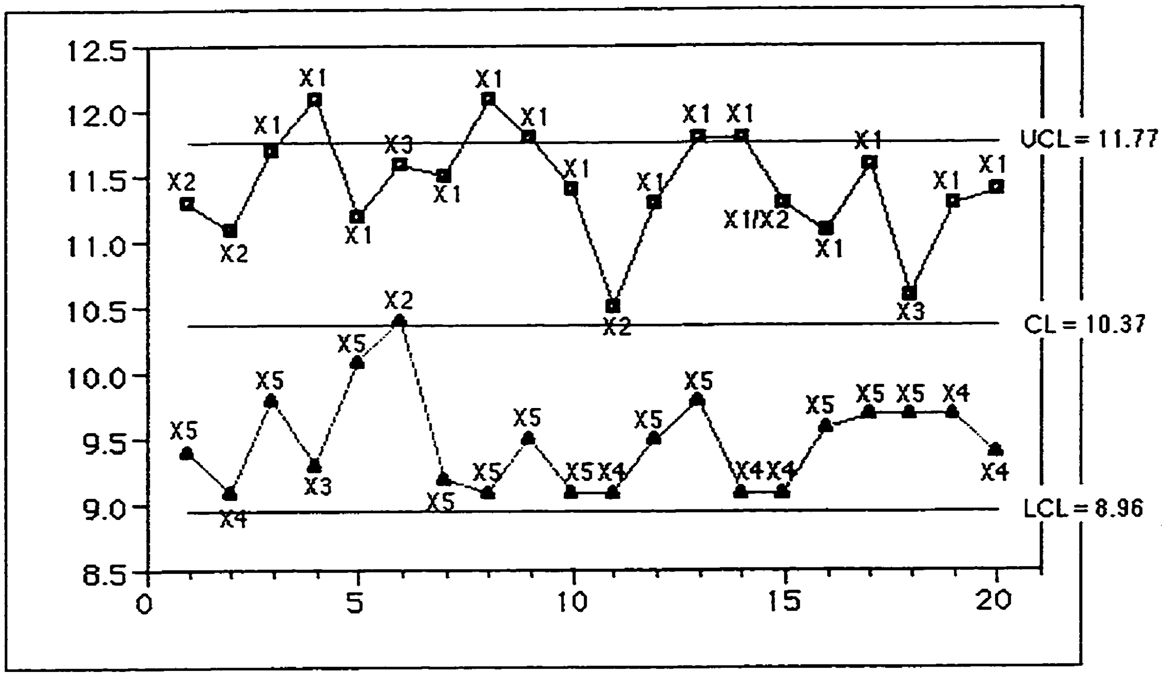
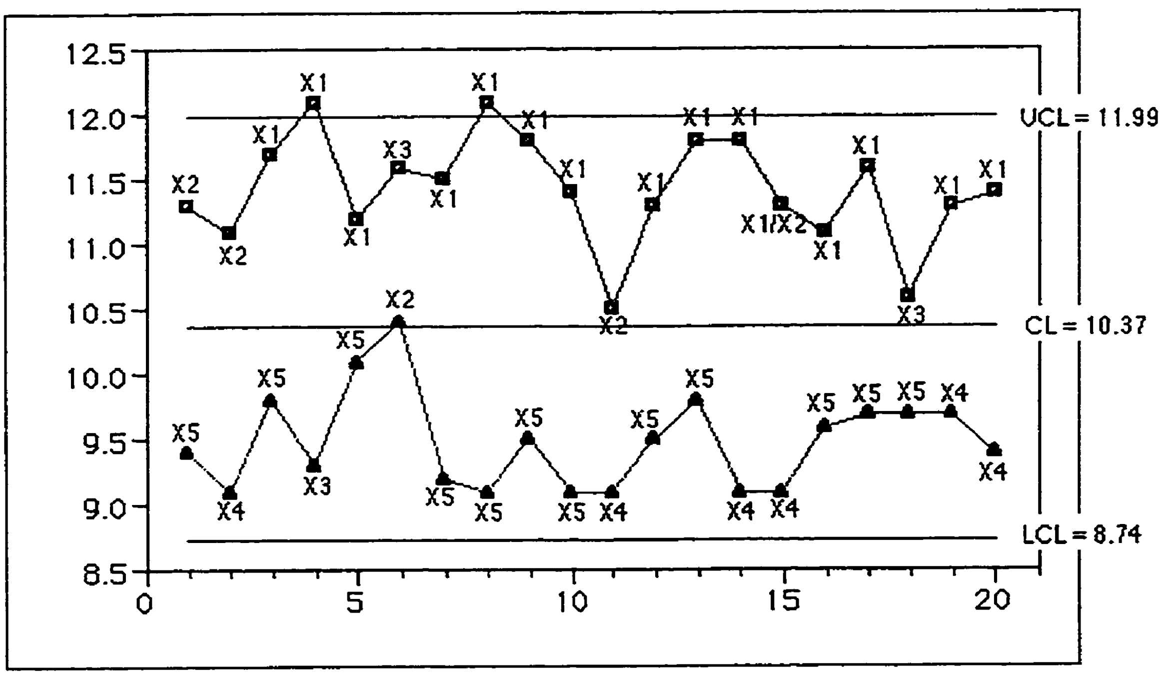



Figure 18d: Group Control Chart (Method 4)



is tied with X2 for the maximum stream in period 15). This gives further evidence that the streams are not equal. Looking at the minimum streams, the longest run is for Stream X5 with four consecutive points starting in period 7. According to Table 20, four consecutive points for a process with five streams is not that unusual. However, notice that Streams X4 and X5 are the minimum streams for 14 time periods in a row beginning in period 7. It seems that this is telling us that Streams X4 and X5 may be different from the rest of the streams.

5.2.4 Revising the Control Limits

For "Data Set 1", the control charts for Methods 1 and 3 had no samples where both the maximum and minimum points plotted within the control limits. It happened only five times for Method 2 and four times for Method 4. With so few samples in control, it is impossible to come up with revised control limits by eliminating the out-of-control samples since that would leave hardly any samples for the new limits.

Many more plotted points are within the control limits in "Data Set 2", however. This allows us the opportunity to eliminate the out-of-control samples, assuming that assignable causes were detected and eliminated, and to recalculate the control limits. Two samples (4 and 8) were eliminated for Methods 1, 2, and 4, and five samples (4, 8, 9, 13, and 14) were eliminated for Method 3 to get the four revised control charts shown in Figures 19a to 19d. All of the plotted points are now within the control limits.

Looking at the revised control charts, Stream X1 is the maximum stream six times in a row for Methods 1, 2, and 4 (starting in period 10). Nothing seems too unusual for the maximum streams for Method 3.

Streams X4 and X5 are the minimum streams for thirteen time periods in

Figure 19a: Revised Group Control Chart (Method 1)

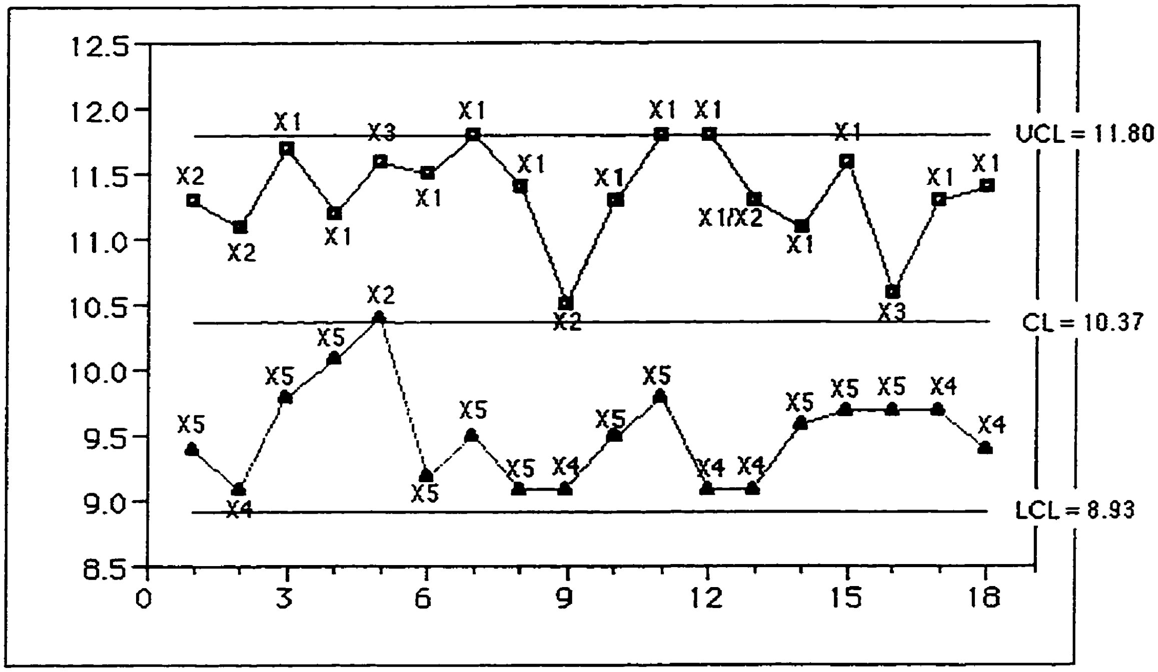



Figure 19b: Revised Group Control Chart (Method 2)

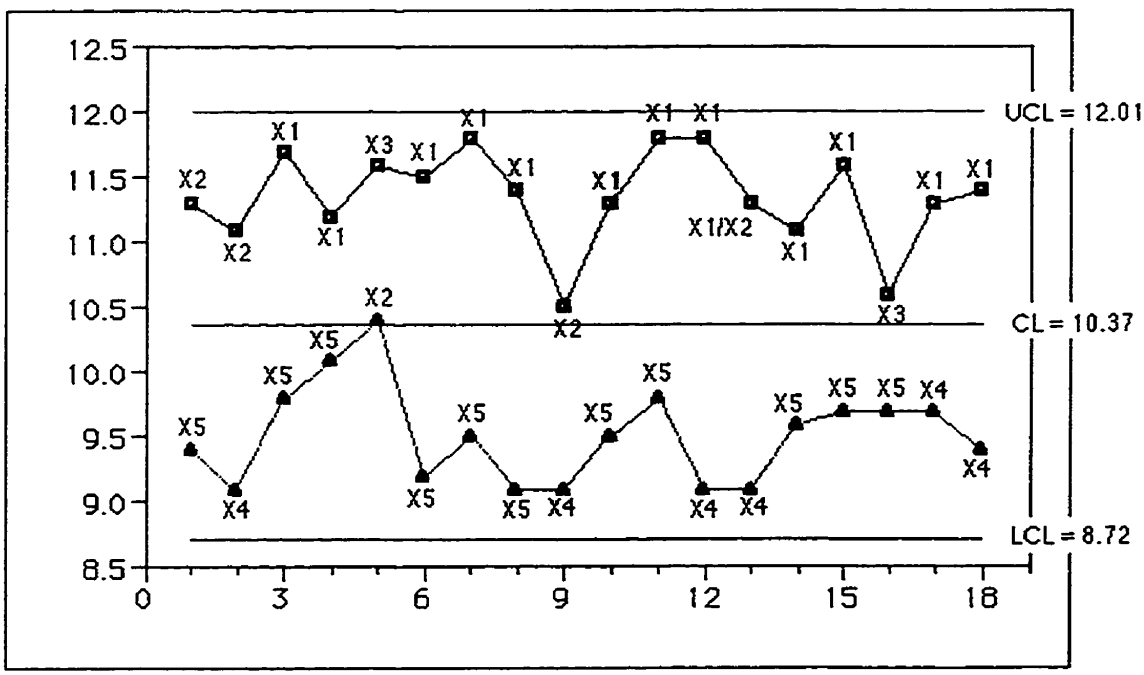


Figure 19c: Revised Group Control Chart (Method 3)

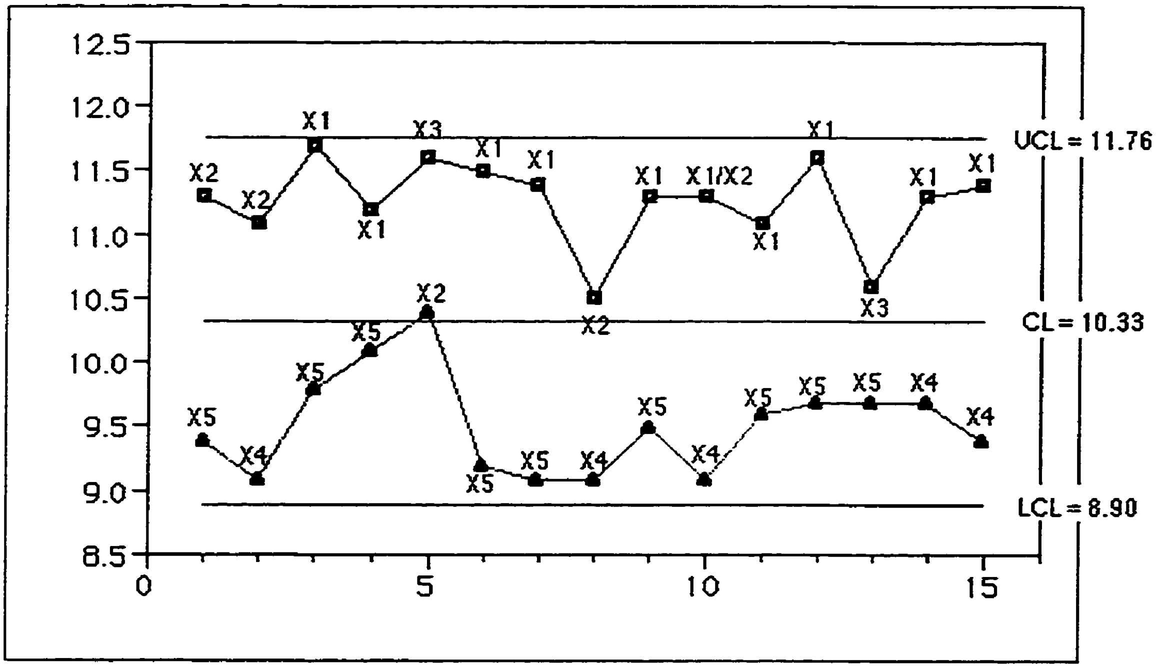
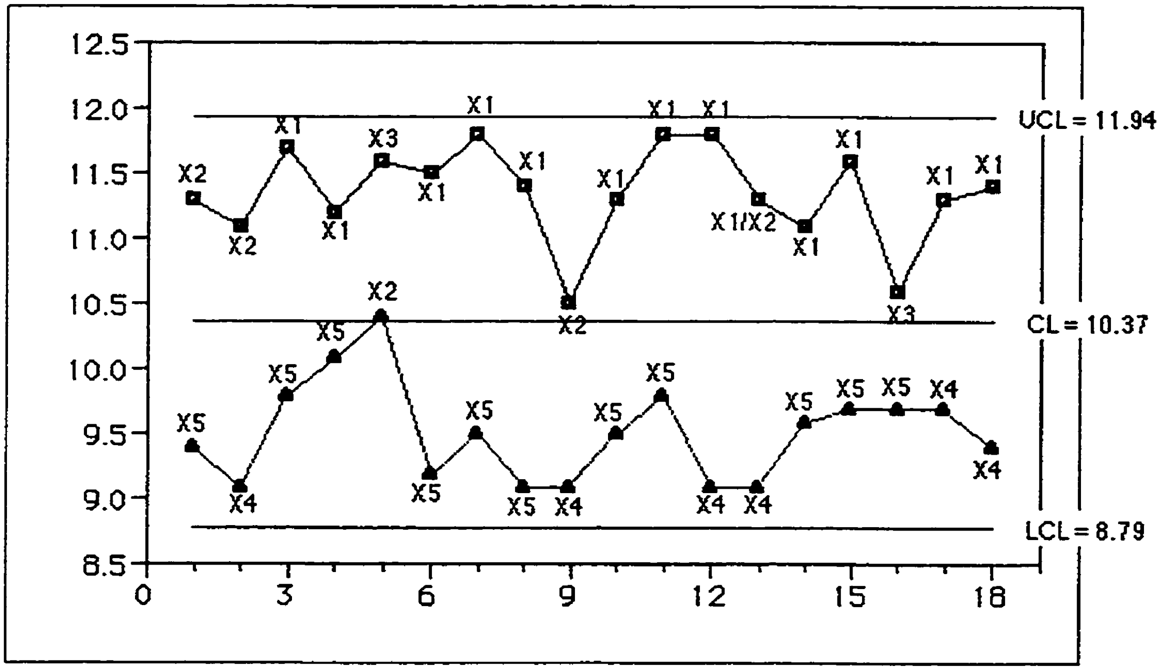



Figure 19d: Revised Group Control Chart (Method 4)



a row for Methods 1, 2, and 4 (starting in period 6), and for ten time periods in a row for Method 3 (also starting in period 6).

### 5.2.5 Control Charts for Each Stream

Control charts for each individual stream are shown in Figures 20a to 20e. All of the streams are in control except for X5. Stream X5 has an out-of-control point at time period 6. This does not show up in the four trial control charts (Figures 18a to 18d) or the four revised control charts (Figures 19a to 19d). Even though the value is high for stream X5, it is not high enough to become the maximum value for the time period and therefore is not shown in the four trial control charts and the four revised control charts.

### 5.2.6 Control Chart Using the Streams as a Subgroup

Figure 21 shows a $\overline{X}$ and R chart where the five streams are treated as a subgroup. The process appears to be in control, but this control scheme averages the streams and therefore hides their differences.

Figure 20a: X and MR Control Charts for Stream X1

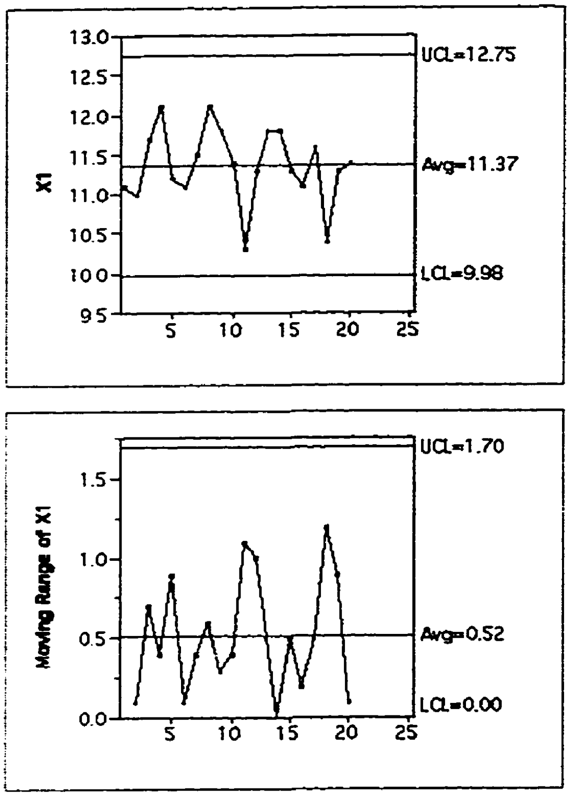



Figure 20b: X and MR Control Charts for Stream X2

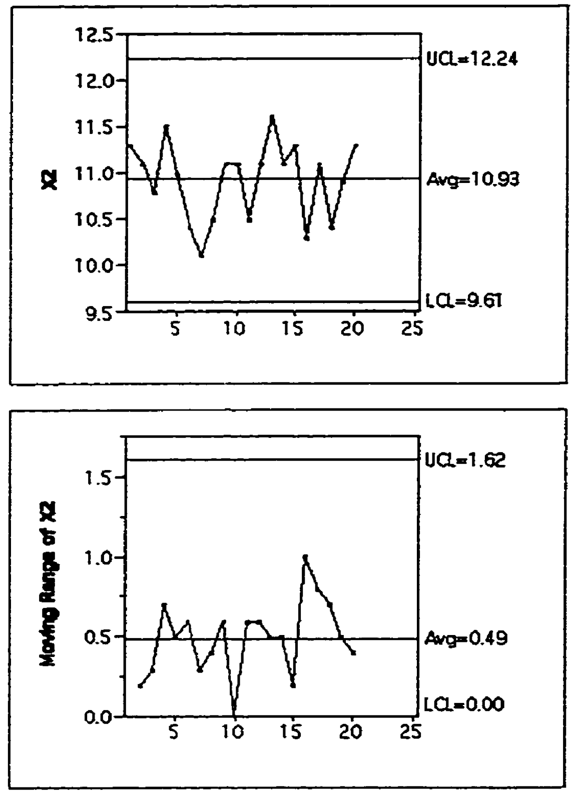



Figure 20c: X and MR Control Charts for Stream X3

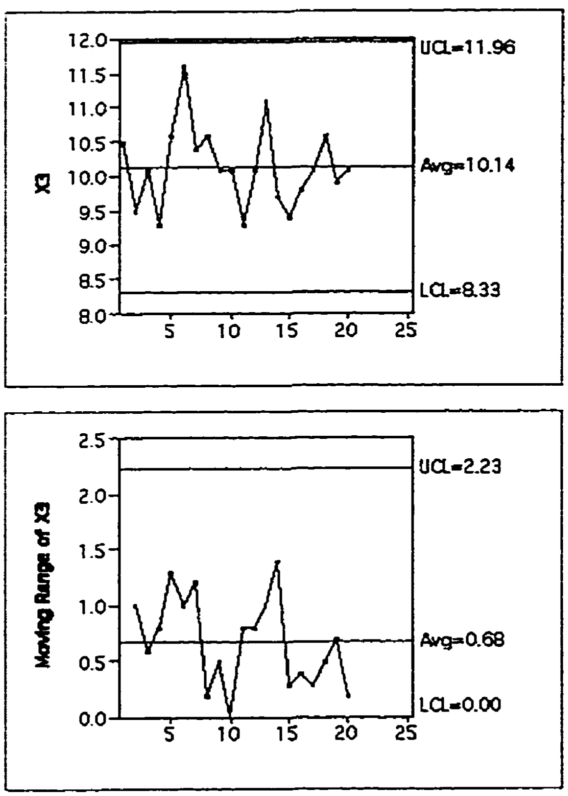



Figure 20d: X and MR Control Charts for Stream X4

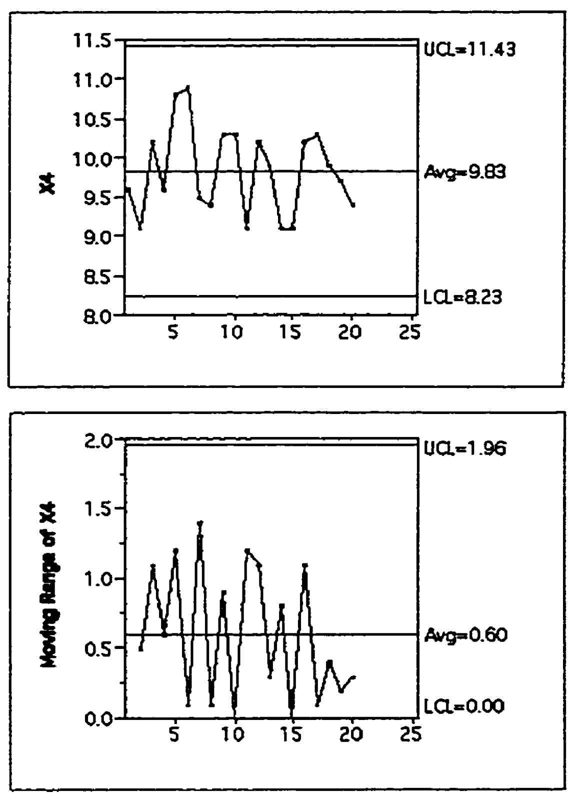
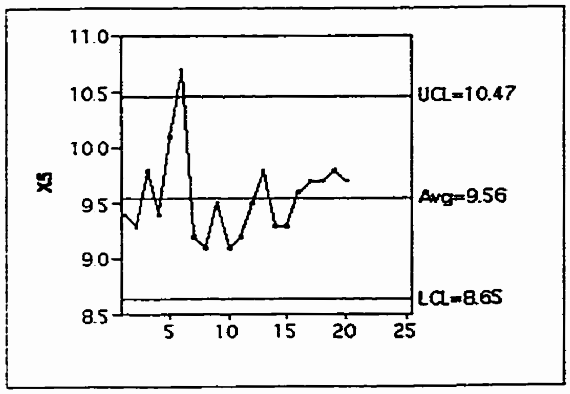


Figure 20e: X and MR Control Charts for Stream X5



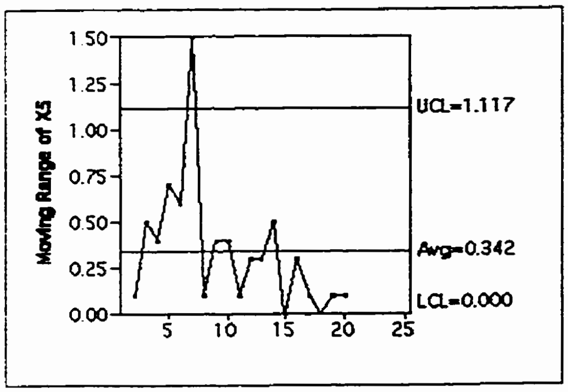
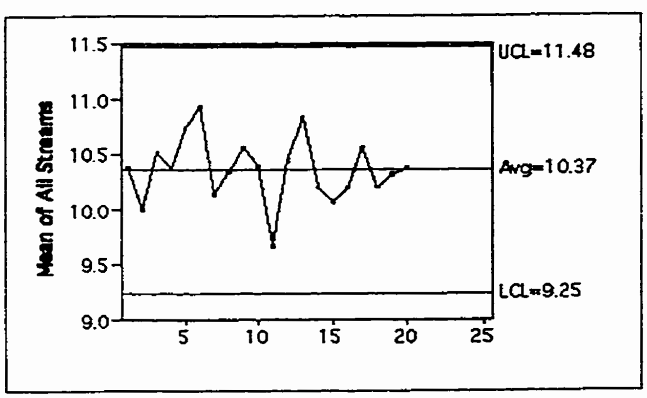
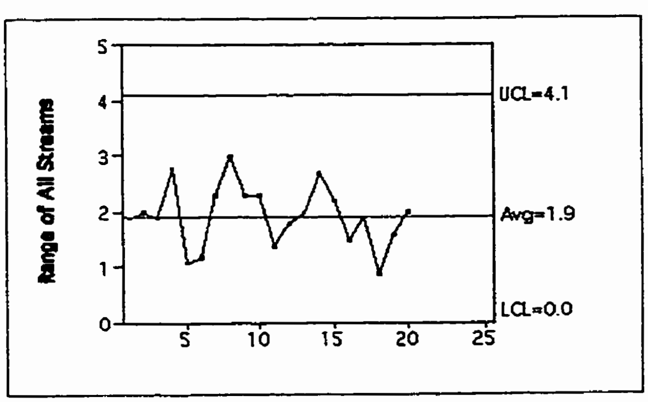




Figure 21: Control Charts Treating the Streams as Subgroups



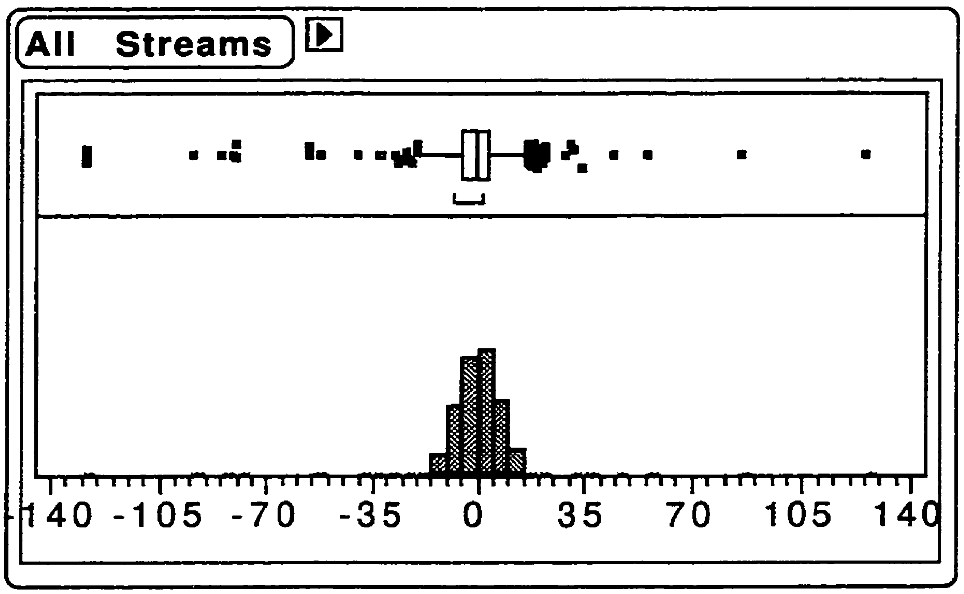
### 5.3 <u>Data Set 3</u>

This data set represents a process from a printing operation with 15 streams. The data are recorded as the difference between two images expected to line up together. If the two images line up perfectly, a value of zero is recorded. Figure 22 displays a histogram and boxplot of the data set for 450 consecutive time periods. We can see from Figure 22 that there are many outliers in the data.

In consultation with an expert for the process, it was decided that the extreme outliers could be eliminated because these values usually occur when the process is just starting up for the day. Therefore, those values should not be used in the construction of our control limits because they are transient and not expected to re-occur for the rest of the day.

For the purpose of illustration for this practicum, 50 consecutive time periods were selected from the original data set for the discussion of the four monitoring methods. The "revised" data set is shown in Table 24.

Figure 22: Histogram and Boxplot of Data Set 3 Before Revision



## Table 24

### Data Set 3 (Revised)

| Sample No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5 | 11 | 7 | 3 | -5 | -5 | 1 | -1 | -5 | 2 | 4 | -5 | -7 | 2 | -11 |
| 2 | -2 | -6 | 12 | -1 | 3 | 2 | -3 | 11 | 0 | -11 | 4 | 2 | -3 | 3 | -5 |
| 3 | -4 | 6 | 2 | 3 | 0 | -2 | -4 | 3 | 5 | 7 | 1 | -7 | 3 | -9 | -8 |
| 4 | -4 | 18 | -1 | -5 | -11 | 2 | -1 | 1 | -6 | -4 | 2 | 0 | 7 | 12 | -8 |
| 5 | 3 | 9 | 13 | 5 | -1 | -6 | 1 | 7 | 2 | -12 | -3 | -6 | 3 | 0 | -10 |
| 6 | 5 | 21 | 7 | 11 | 10 | 1 | 6 | -5 | -2 | -9 | -9 | -4 | -1 | 5 | 0 |
| 7 | -2 | -11 | 8 | -1 | -15 | -8 | 1 | 10 | 3 | 6 | -4 | 0 | -6 | -2 | -4 |
| 8 | -3 | 13 | 5 | -4 | -6 | -7 | -6 | 5 | -10 | -11 | -1 | -5 | 13 | 0 | 6 |
| 9 | 9 | 2 | 3 | -6 | -11 | -3 | -13 | -14 | -10 | -10 | -3 | -2 | 4 | -4 | 4 |
| 10 | -2 | 0 | 11 | 4 | 0 | 0 | 8 | -1 | -3 | 0 | -2 | 2 | -5 | -8 | -5 |
| 11 | 5 | 10 | 8 | 14 | 8 | 8 | 0 | 6 | 1 | 7 | -3 | 2 | -9 | 5 | -7 |
| 12 | 9 | 9 | 9 | 1 | 1 | 5 | -6 | -5 | 10 | 2 | -18 | 3 | 2 | 0 | -4 |
| 13 | 7 | 13 | 3 | -1 | -7 | -1 | -9 | -12 | -3 | -1 | 0 | -7 | -5 | 1 | -1 |
| 14 | 15 | -7 | 11 | 1 | -3 | -4 | 6 | 8 | -2 | 0 | -3 | -3 | -2 | 8 | 4 |
| 15 | 3 | 5 | -1 | 9 | 8 | 3 | -6 | -6 | 1 | -3 | -3 | 12 | 6 | 0 | 12 |
| 16 | 21 | 14 | 0 | 0 | -4 | -5 | -11 | -6 | -7 | -3 | 5 | -12 | -7 | 8 | -4 |
| 17 | 0 | 8 | 16 | 7 | 4 | -4 | 1 | -2 | -2 | -11 | -13 | -5 | 0 | -9 | -3 |
| 18 | 3 | 10 | 0 | -7 | -3 | -3 | -10 | 0 | 0 | 2 | -10 | 3 | 8 | 4 | -3 |
| 19 | 7 | 5 | 11 | 5 | -5 | 2 | -11 | -3 | 0 | -5 | 1 | 0 | 1 | 5 | -13 |
| 20 | 3 | 1 | 9 | -12 | 1 | -1 | 1 | 2 | 4 | -3 | 2 | -11 | -4 | 1 | 1 |
| 21 | -2 | 11 | 8 | 3 | 2 | -4 | -5 | -9 | 4 | -4 | 5 | 0 | -5 | -1 | -4 |
| 22 | -3 | -2 | 13 | 4 | 2 | -10 | -5 | -10 | -6 | 0 | 7 | 2 | -1 | 9 | 1 |
| 23 | 6 | -8 | 4 | 0 | -1 | -5 | 0 | -9 | -3 | 2 | -6 | 2 | 6 | 3 | 7 |
| 24 | 7 | 6 | 6 | -5 | -6 | -4 | 1 | -2 | -3 | -8 | -10 | -7 | -2 | 11 | -11 |
| 25 | 11 | -1 | 7 | 8 | 8 | -2 | -7 | -5 | 0 | 7 | 1 | 4 | 1 | 5 | 6 |
| 26 | -3 | 2 | -7 | -4 | -2 | 0 | -12 | 10 | -4 | -6 | -4 | -5 | -1 | 2 | 7 |
| 27 | 9 | -3 | 7 | -6 | 2 | 3 | -2 | -4 | -10 | -6 | 16 | 7 | 7 | 3 | -6 |
| 28 | 3 | -2 | -2 | -1 | -8 | -2 | -3 | 1 | -2 | 1 | 0 | -4 | 3 | 4 | 3 |
| 29 | 3 | -9 | 20 | 2 | -3 | -3 | 3 | 1 | -1 | -14 | -11 | 3 | 6 | 10 | 6 |
| 30 | 12 | 5 | 6 | -13 | 5 | -4 | -8 | -4 | -16 | 1 | -1 | -4 | -10 | 3 | -3 |
| 31 | 9 | -2 | 6 | 2 | -8 | -12 | -1 | 9 | 7 | 4 | -7 | -10 | -5 | -1 | 9 |
| 32 | 3 | 2 | 5 | -7 | -6 | 0 | -12 | 5 | -1 | 8 | -2 | -6 | -7 | 3 | 4 |
| 33 | 5 | -11 | 9 | 5 | -1 | 2 | -11 | -1 | -2 | -1 | 0 | 1 | 10 | 5 | -7 |
| 34 | 11 | 6 | 7 | 0 | 1 | 8 | -5 | 6 | -1 | -5 | -4 | 6 | 0 | 7 | 15 |
| 35 | 8 | 2 | 8 | -1 | -6 | -11 | -7 | -9 | -9 | -2 | 0 | 8 | 5 | 4 | 1 |
| 36 | 2 | 17 | 16 | 2 | 3 | -3 | -2 | -3 | -10 | -10 | -1 | -9 | 3 | 2 | 1 |
| 37 | 15 | 4 | 6 | 5 | 0 | -2 | 0 | -3 | 9 | 7 | 1 | -2 | 3 | 3 | -6 |
| 38 | -6 | -22 | -6 | -5 | -13 | -11 | -16 | -11 | -7 | -12 | 4 | -6 | -8 | 1 | 1 |
| 39 | 4 | 14 | 10 | 3 | -2 | 2 | -14 | -10 | 2 | 2 | 8 | 7 | 4 | 6 | -2 |
| 40 | 12 | -1 | 14 | -13 | 5 | 1 | -7 | -5 | 1 | 0 | -2 | -11 | -5 | -6 | 4 |
| 41 | 5 | -4 | 9 | -12 | -2 | -4 | -12 | -4 | -9 | -4 | 3 | 5 | 2 | -6 | -2 |
| 42 | -12 | 9 | 2 | 6 | 7 | -1 | -7 | 1 | 3 | -9 | 9 | -5 | -1 | -5 | -6 |
| 43 | 6 | 4 | 7 | 6 | 7 | 11 | -4 | 12 | 2 | 5 | 6 | -3 | 0 | 2 | -7 |
| 44 | 9 | 5 | -2 | -8 | -2 | 2 | -7 | -1 | -5 | 7 | 1 | 9 | 0 | 1 | 0 |
| 45 | 9 | 8 | 7 | -3 | -4 | -2 | -7 | -10 | -4 | -5 | -7 | -3 | -5 | -2 | 0 |
| 46 | 14 | 3 | 8 | -4 | -6 | -3 | -8 | -2 | -12 | -7 | -4 | -7 | -4 | -2 | 0 |
| 47 | 9 | 11 | 4 | 9 | 4 | 1 | -11 | -3 | -8 | -6 | 4 | 3 | -5 | 1 | -1 |
| 48 | -3 | -2 | 15 | -4 | 7 | 6 | 2 | 5 | -3 | -3 | 3 | -6 | -4 | 6 | 0 |
| 49 | 23 | 8 | 9 | 10 | 12 | -3 | -9 | -5 | 4 | -5 | 1 | 0 | -1 | -4 | 4 |
| 50 | -4 | 5 | 4 | -13 | 2 | 7 | -10 | -2 | -4 | -10 | -1 | -7 | 4 | -1 | 7 |

A histogram and boxplot of the revised data is shown in Figure 23. Figures 24a to 24o show the data by individual stream. These histograms and boxplots show that it is not unreasonable to assume that the data are approximately normally distributed.

### 5.3.1 Testing the Equality of the Streams

Figure 25 shows means diamonds and quantile boxes for the 15 streams. We can conclude that the streams are not all identical with respect to their location because the confidence intervals shown by the means diamonds do not all overlap. It would appear that Streams 1, 2, and 3 are higher, and Stream 7 lower than the other streams. We can also see the stream differences from the analysis of variance table (Table 25). The p-value from the table is less than 0.0001. This shows strong evidence that at least one of the streams is significantly different from the rest.

Figure 26 shows side-by-side boxplots of moving ranges for all 15

Figure 23: Histogram and Boxplot of Data Set 3 (Revised)

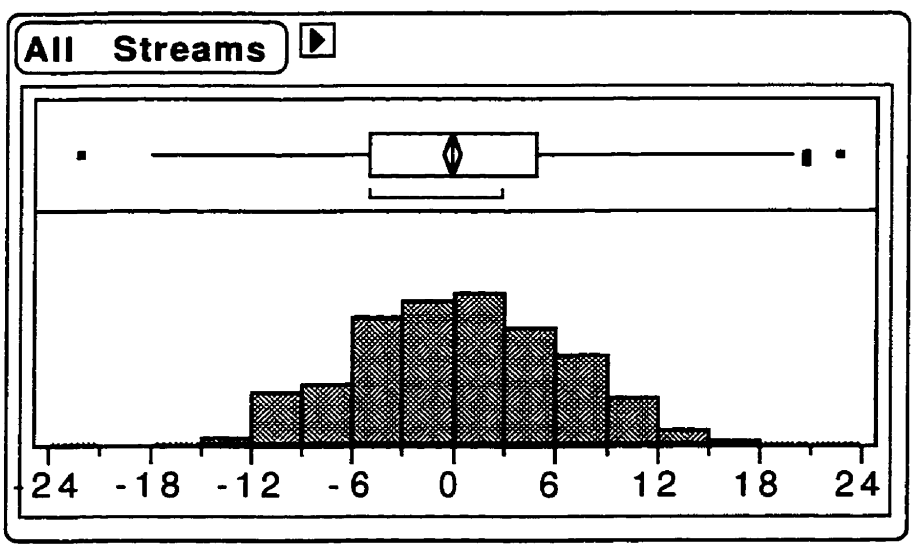


Figure 24a

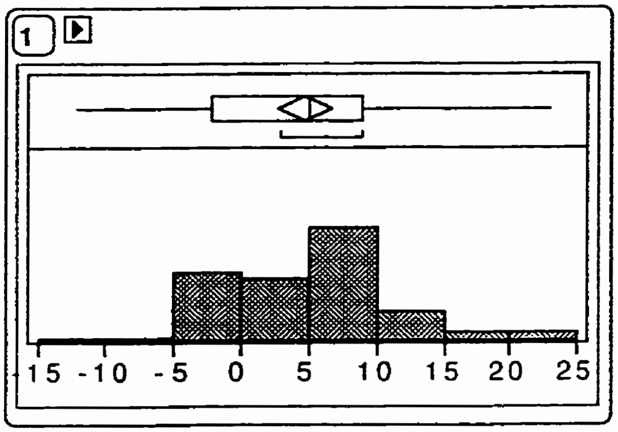



Figure 24b

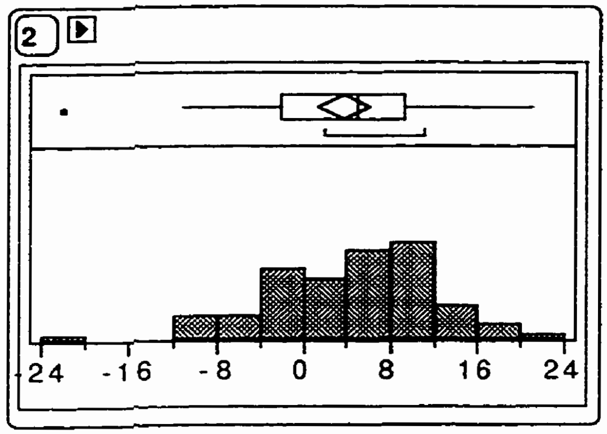



Figure 24c

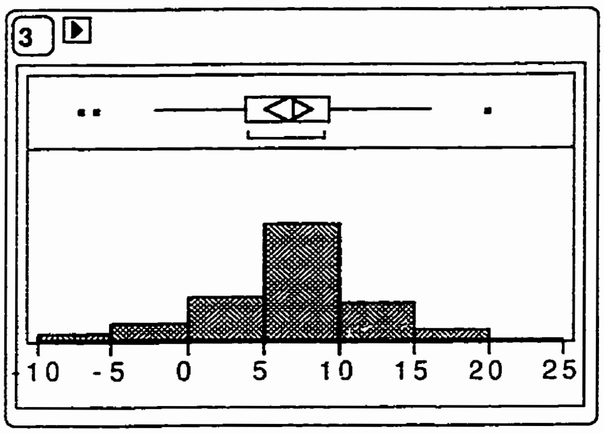



Figure 24d

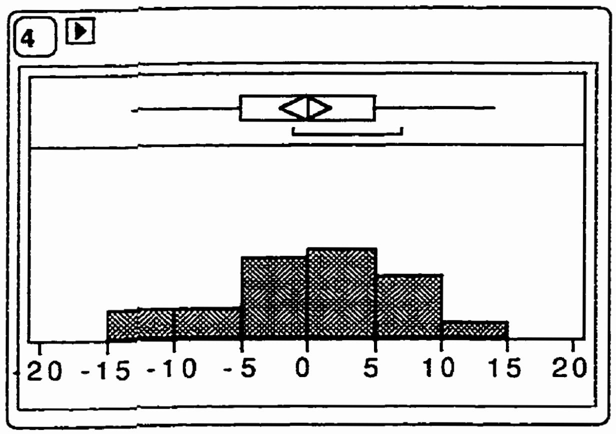



Figure 24e

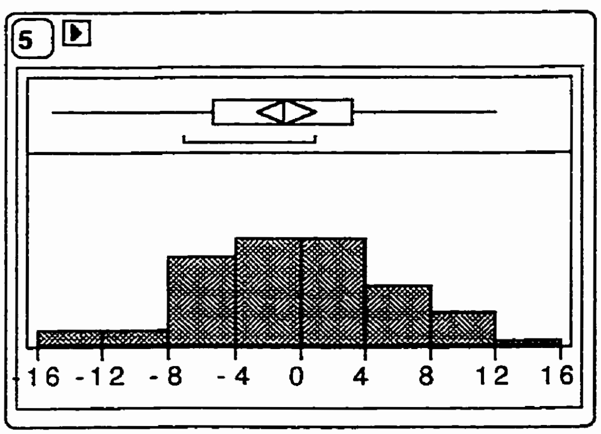



Figure 24f

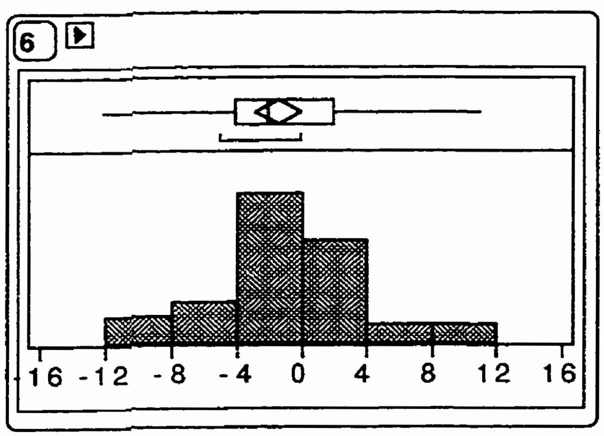


Figure 24g

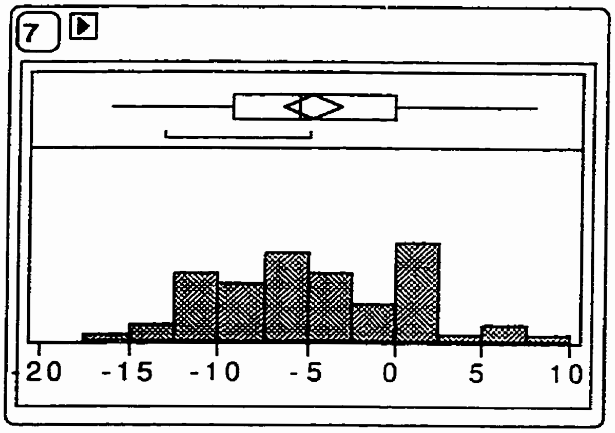



Figure 24h

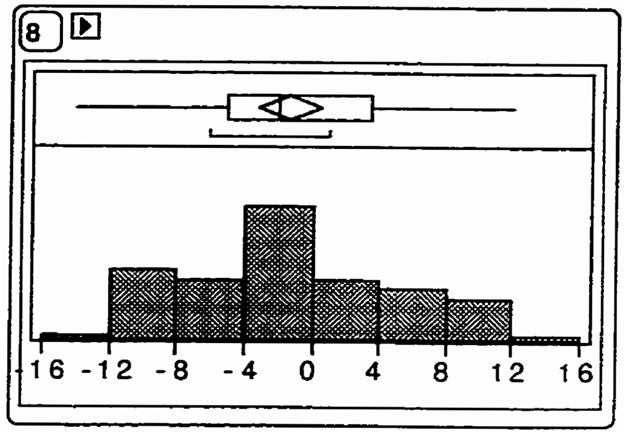



Figure 24i

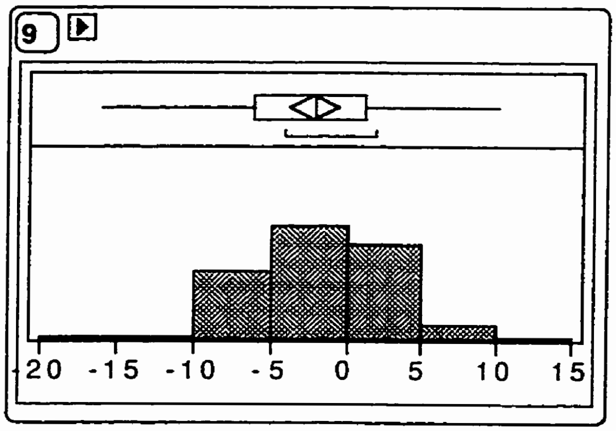



Figure 24j

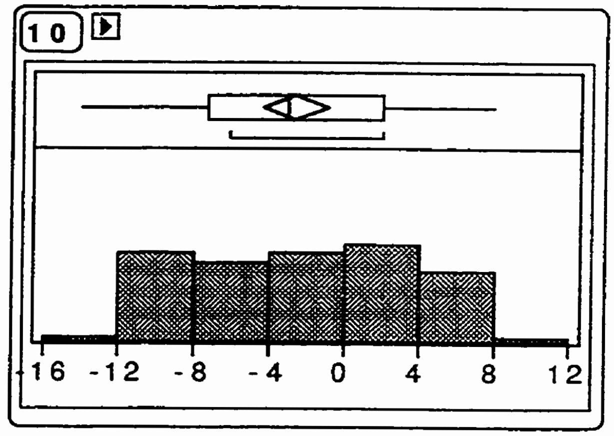



Figure 24k

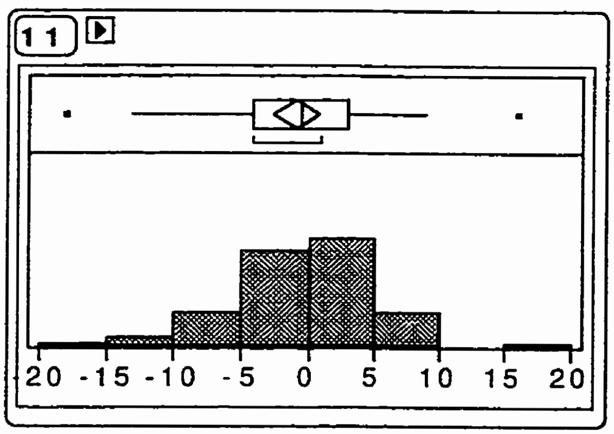



Figure 24l

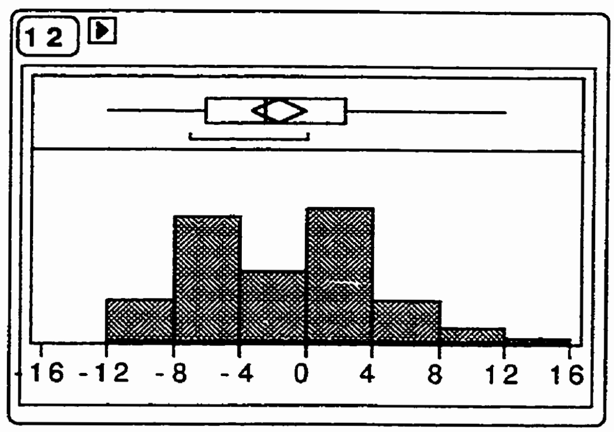


Figure 24m

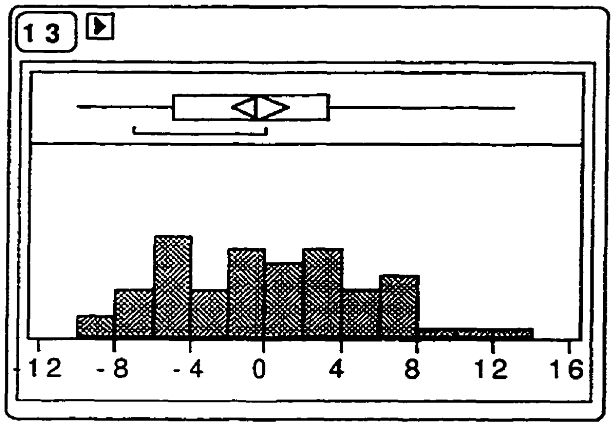



Figure 24n

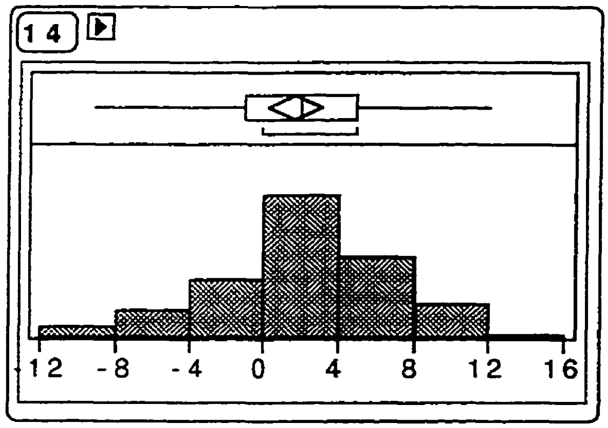



Figure 24o

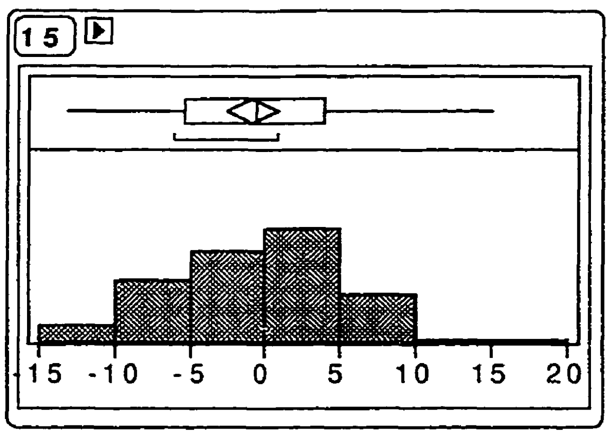


Figure 25: Means Diamonds and Quantile Boxes for Data Set 3

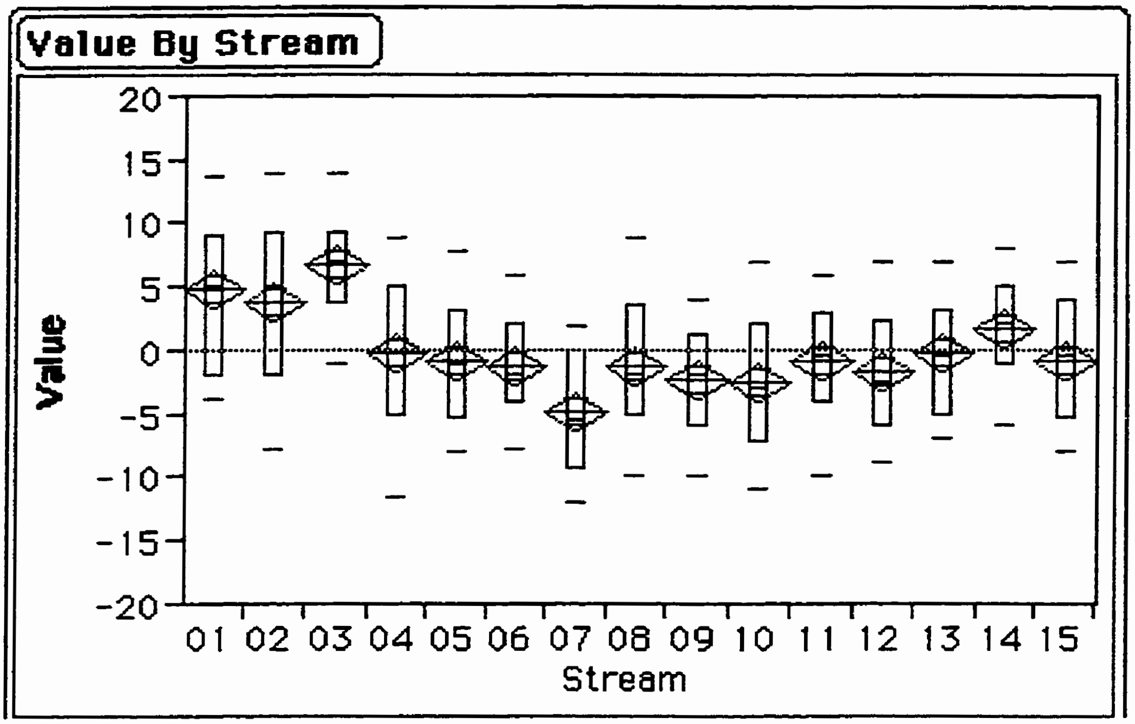
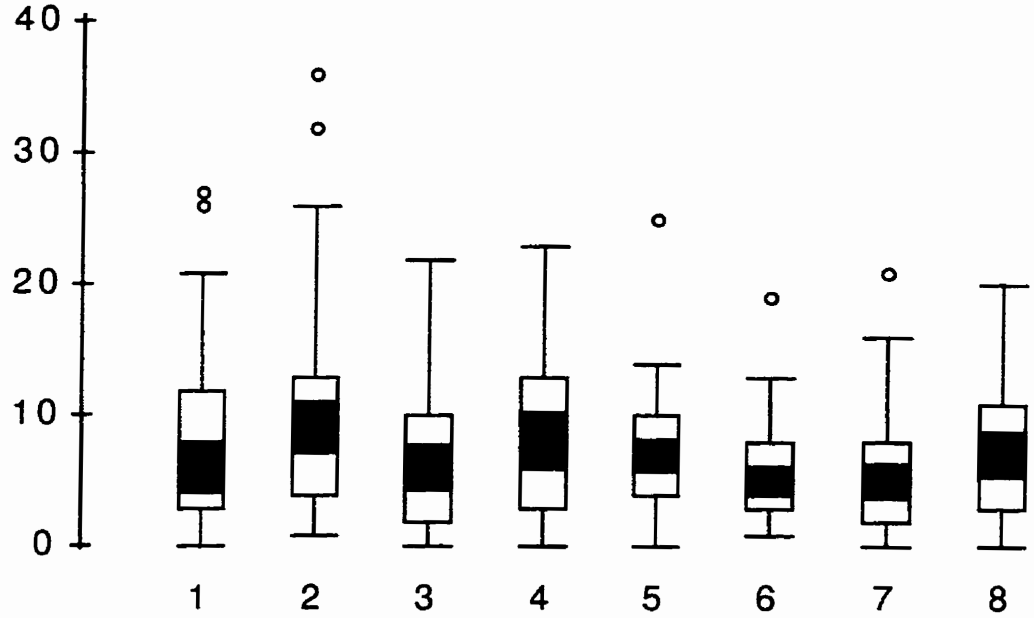


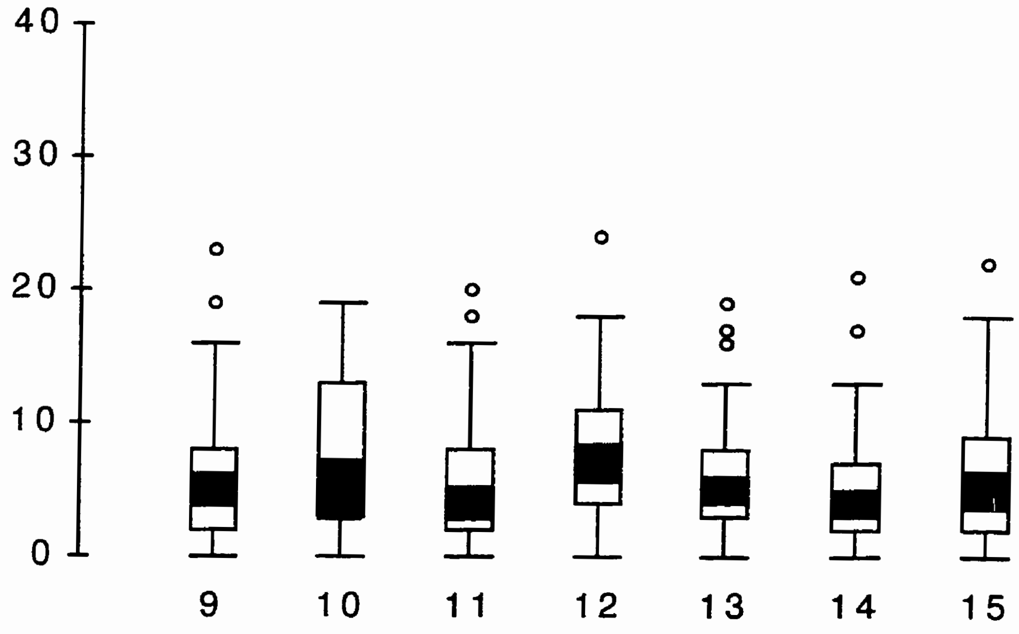

Table 25

Analysis of Variance for Data Set 3

| Source | DF | Sum of Squares | Mean Square | F Ratio | P-value |
|---|---|---|---|---|---|
| Streams | 14 | 6367.192 | 454.799 | 12.4839 | **<0.0001** |
| Error | 735 | 26776.760 | 36.431 | | |
| Total | 749 | 33143.952 | | | |

Figure 26: Boxplots of the Moving Ranges



streams. Since the boxplots are all at about the same location, we can conclude that the streams appear to have similar variability.

### 5.3.2 Control Charts of the Four Methods

Figures 27a to 27d show the control charts of the four different multiple-stream monitoring techniques. If we focus initially on points plotting outside the control limits we see that Methods 2 and 4 do not give any out-of-control signals. The control chart using Method 1 signals the potential presence of assignable causes at periods 6, 16, and 49 where the maximum value from all streams plots above the UCL, and at period 38 with the minimum of all streams falling below the LCL. Like Method 1, Method 3 gives a signal at periods 6, 16, 38, and 49 but it also signals at period 29 with a maximum value above the UCL.

We also examine the chart for the number of consecutive times a particular stream plots as the maximum, or the minimum. As mentioned earlier, this characteristic is independent of the four methods being examined as it is purely a property of the sequence of plotted values and not the control limits.

Examination of the plotted maxima and minima reveals three occurrences of a sequence or run of points from a particular stream greater than two. Stream 1 is the maximum stream on three consecutive periods beginning with period 44 (Stream 1 is tied with Stream 12 for time period 44), Stream 7 is the minimum stream on three consecutive periods beginning with period 32 (Stream 7 is tied with Stream 2 for period 33 and with Stream 10 for period 34), and Stream 8 is the minimum stream on three consecutive periods beginning with period 21 (Stream 8 is tied with Stream 6 for period 22).

From Equation (3) with 15 streams, we calculate that the one-sided

Figure 27a: Group Control Chart (Method 1)

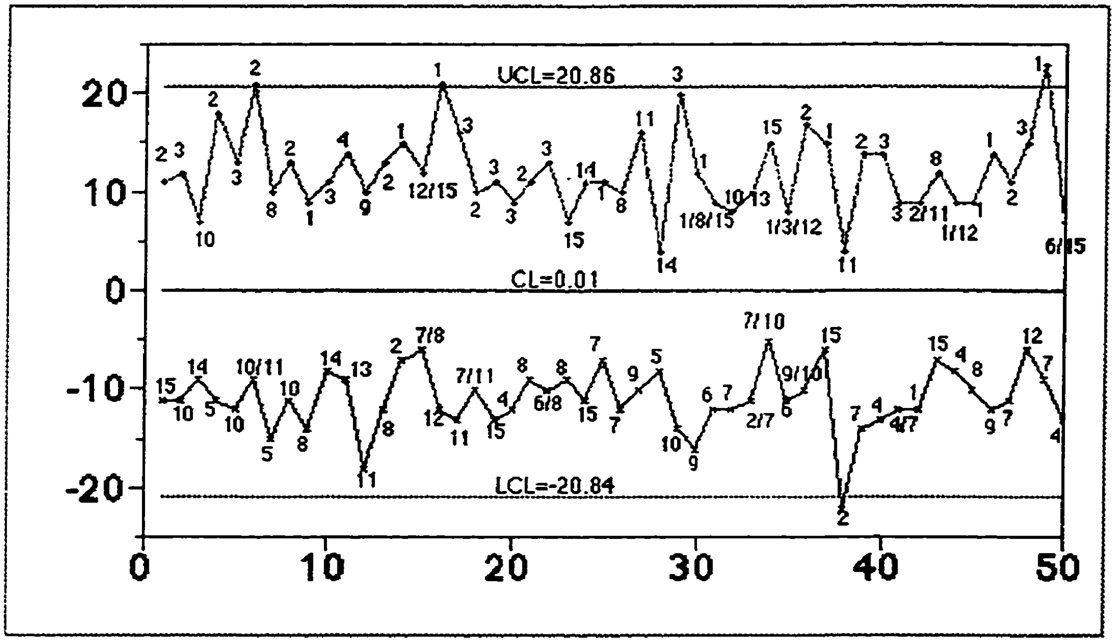



Figure 27b: Group Control Chart (Method 2)

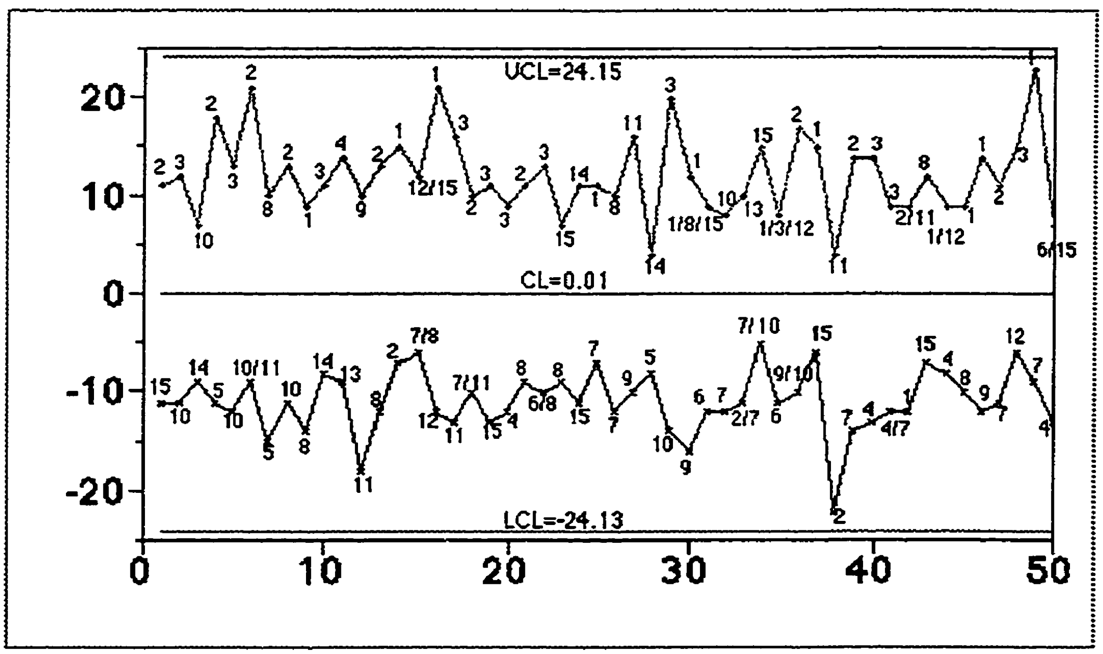


Figure 27c: Group Control Chart (Method 3)

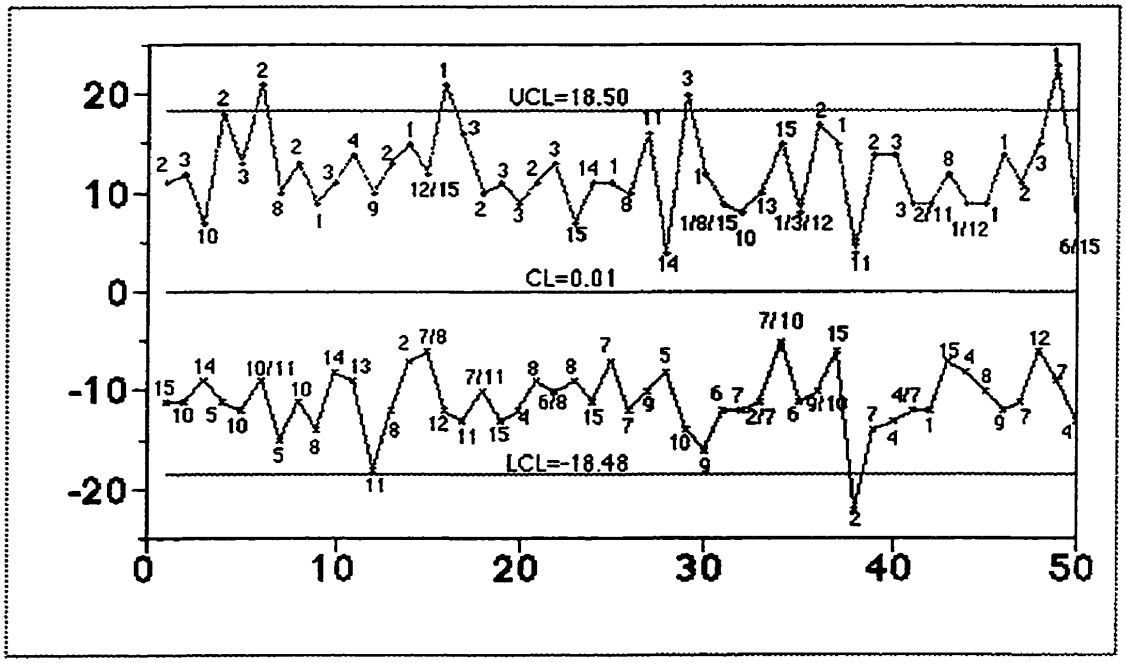
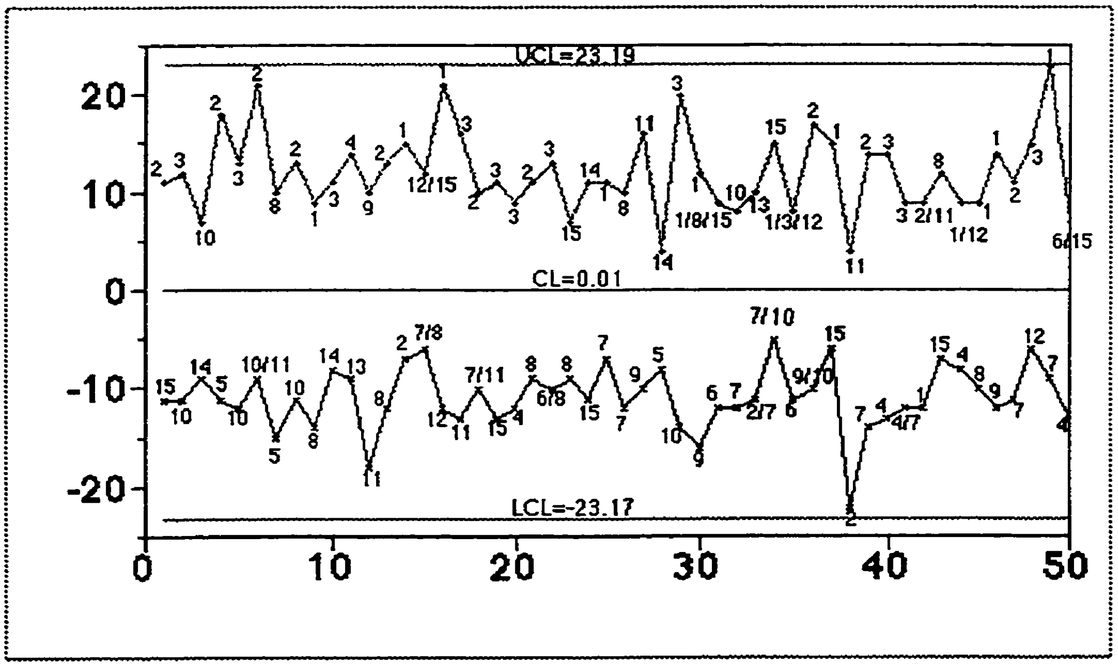



Figure 27d: Group Control Chart (Method 4)



average run length for a run of three consecutive values from the same stream is

$$ARL_0 = \frac{s^r - 1}{s - 1} = \frac{15^3 - 1}{15 - 1} = \frac{3374}{14} = 241 \tag{27}$$

If all of the streams are identical, then a stream plots as the maximum or minimum value three times in a row by chance an average of once every 241 time periods. Because this is a fairly rare occurrence, with probability of approximately 0.004, we have evidence suggesting that Stream 8 at this point behaves differently from the other streams.

We can also see that Streams 1, 2, and 3 plot as the maximum stream much more frequently than any of the other streams. Streams 1 and 3 are the maximum stream (or tied as the maximum stream) 12 times, and Stream 2 is the maximum stream (or tied as the maximum stream) 11 times. This agrees with what we saw in Figure 25. That figure showed that most of the values for Streams 1, 2, and 3 were higher than the values for the other streams. This gives further evidence that some of the streams are behaving differently.

### 5.3.3 Control Charts for Each Stream

Control charts for each individual stream are shown in Figures 28a to 28o. Streams 1, 2, 5, 6, 7, 9, 11, 14, and 15 have one point plotting above the upper control limit on the moving range chart. Stream 11 also has two points outside the control limits on the X chart (samples 12 and 27). Note that Stream 11 is plotted as the minimum stream for sample 12 and the maximum stream for sample 27 in the four group control charts (Figures 27a to 27d), but neither point goes outside the control limits for all four group control methods. This is because the control limits for the group control charts are wider than the control limits for Stream 11.

Figure 28a: X and MR Control Charts for Stream 1

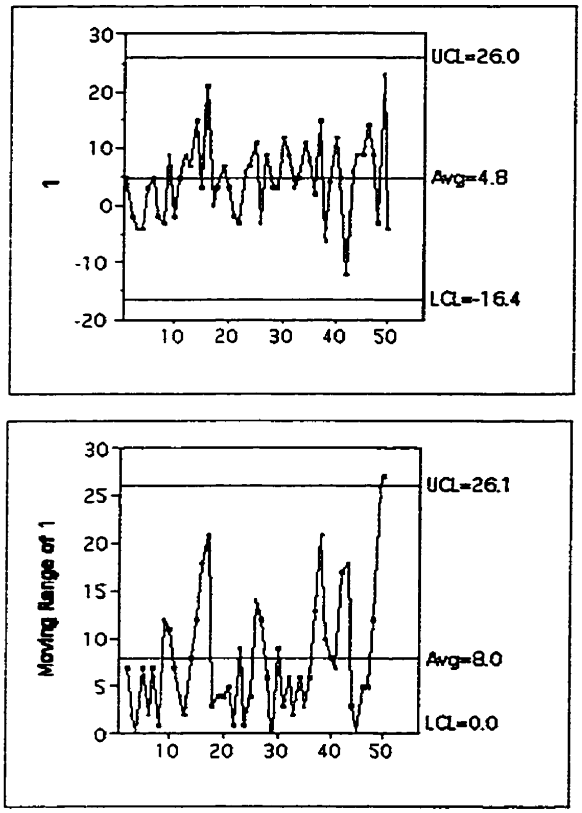



Figure 28b: X and MR Control Charts for Stream 2

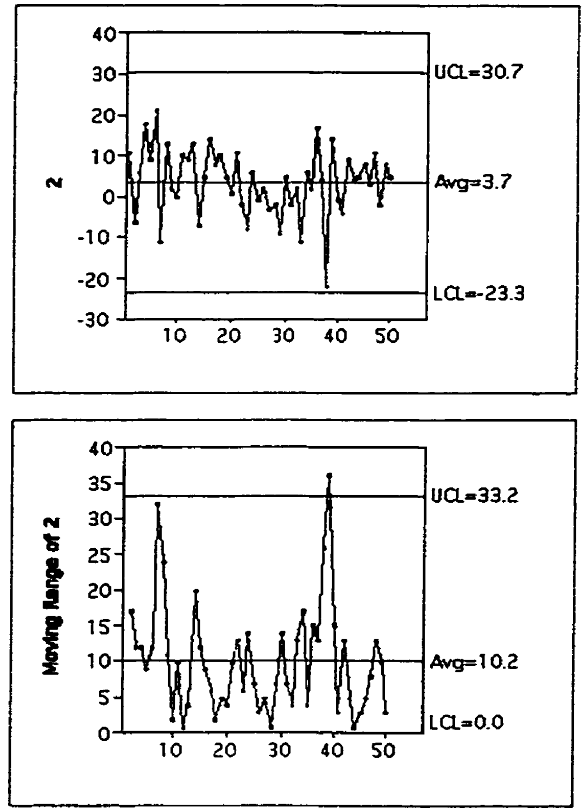



Figure 28c: X and MR Control Charts for Stream 3

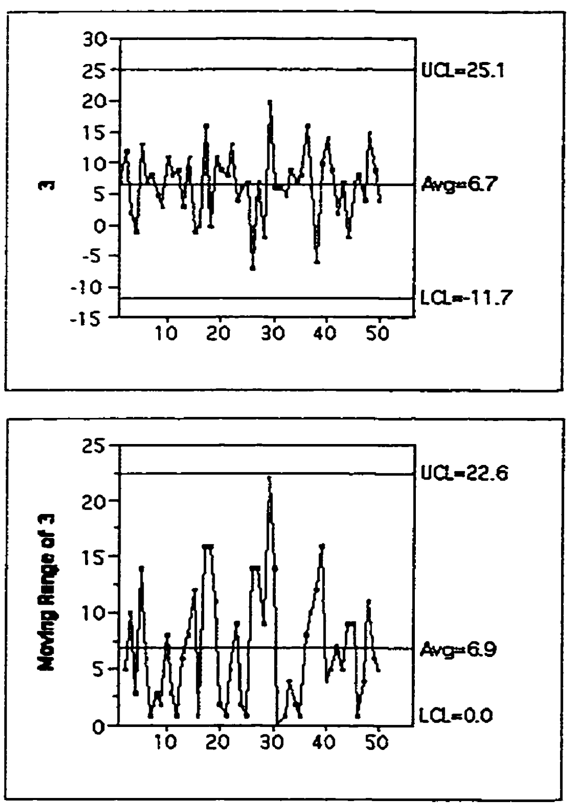



Figure 28d: X and MR Control Charts for Stream 4

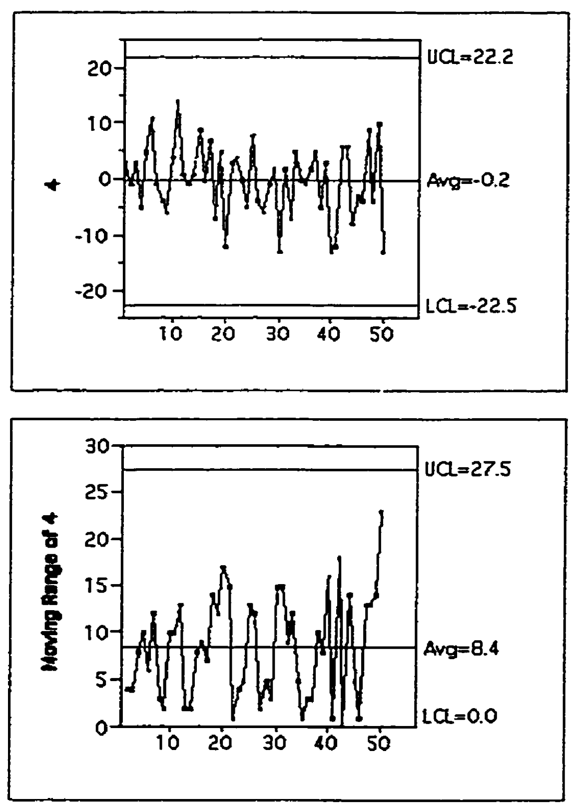


Figure 28e: X and MR Control Charts for Stream 5

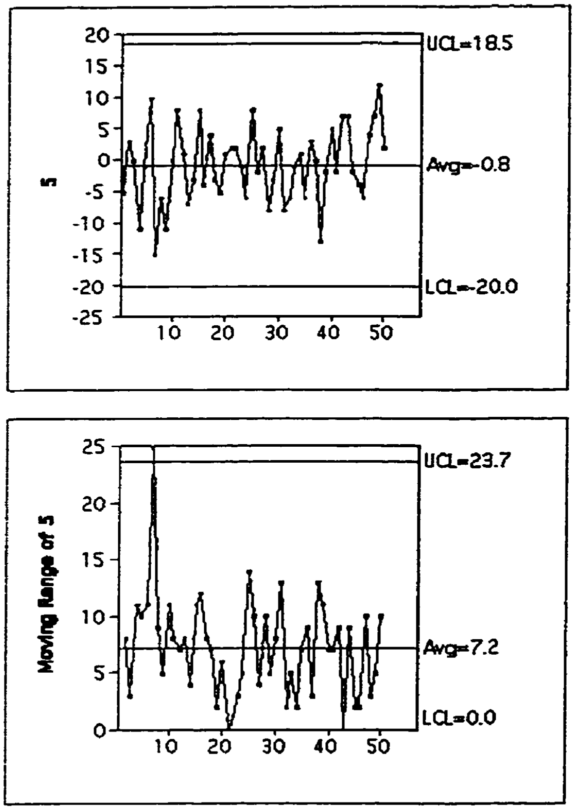



Figure 28f: X and MR Control Charts for Stream 6

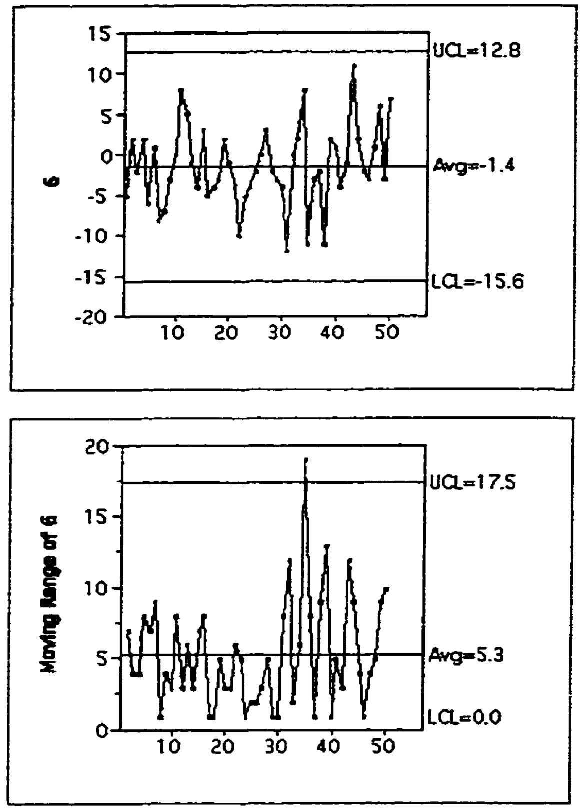



Figure 28g: X and MR Control Charts for Stream 7

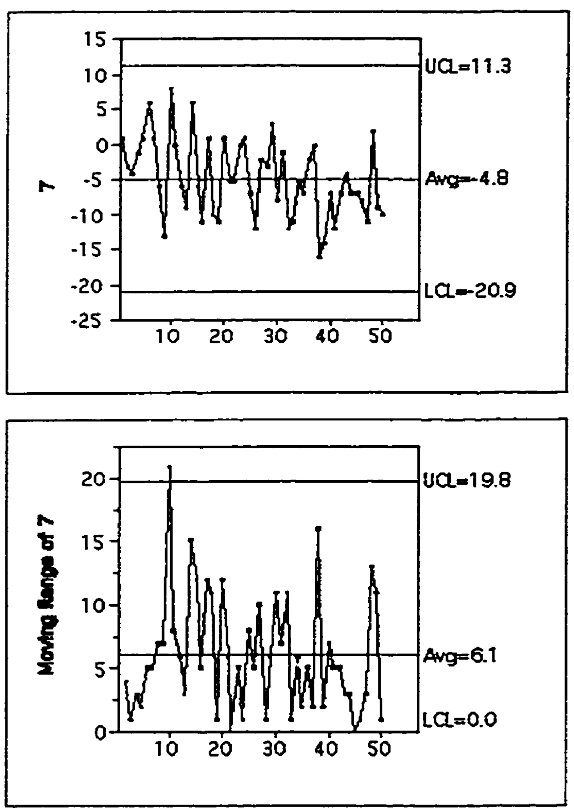



Figure 28h: X and MR Control Charts for Stream 8

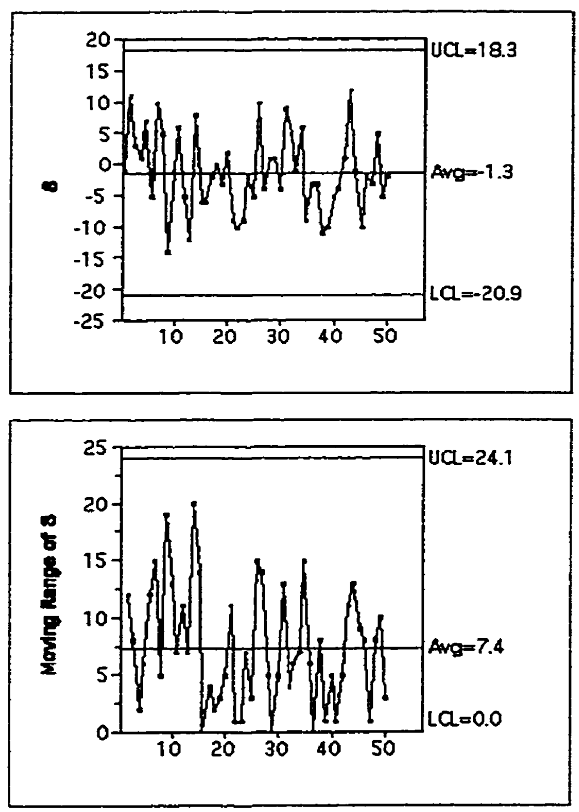


Figure 28i: X and MR Control Charts for Stream 9

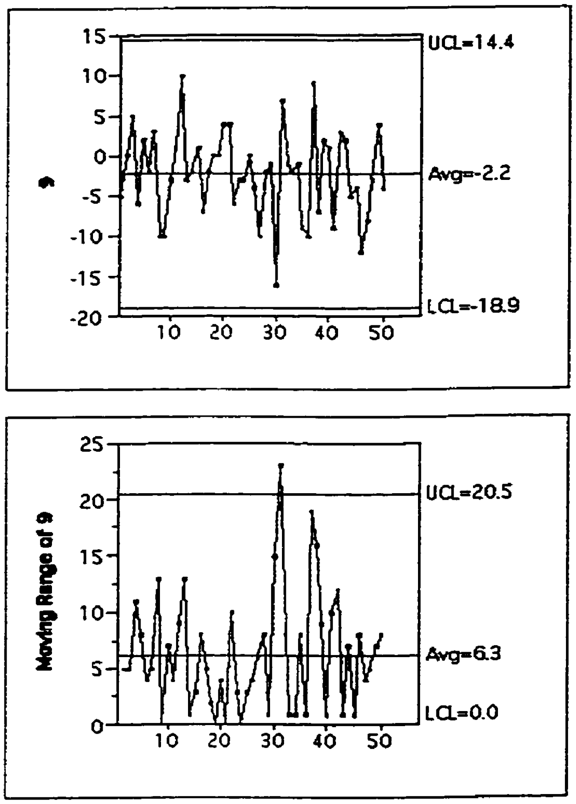



Figure 28j: X and MR Control Charts for Stream 10

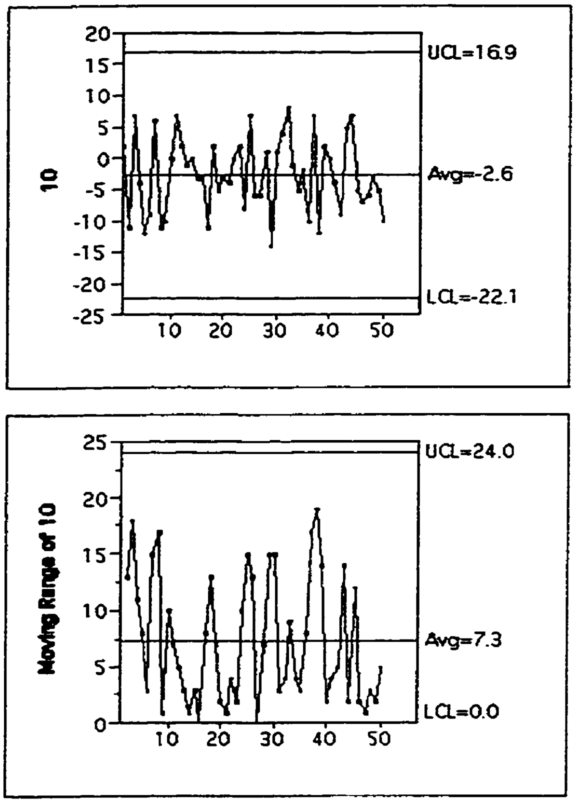



Figure 28k: X and MR Control Charts for Stream 11

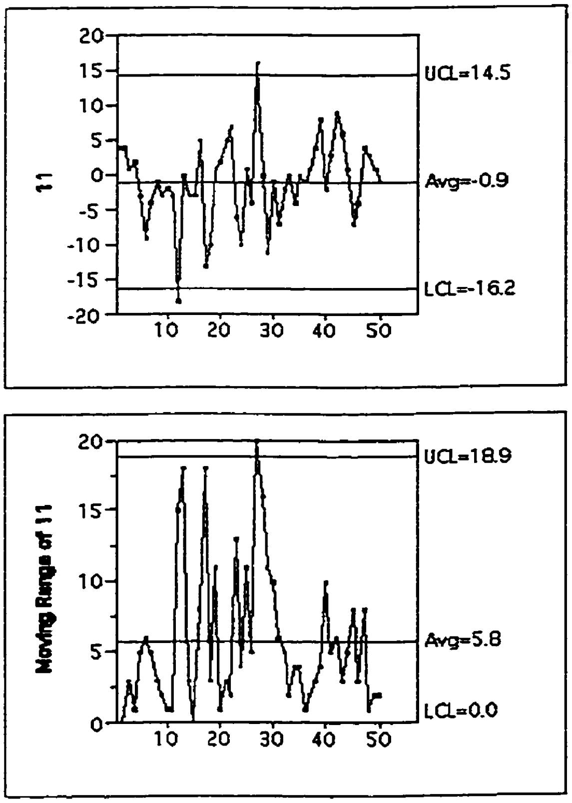



Figure 28l: X and MR Control Charts for Stream 12

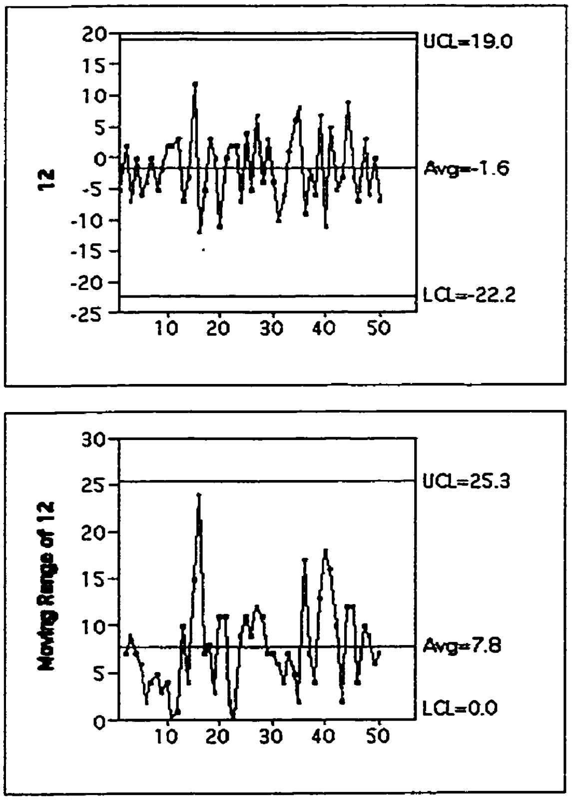


Figure 28m: X and MR Control Charts for Stream 13

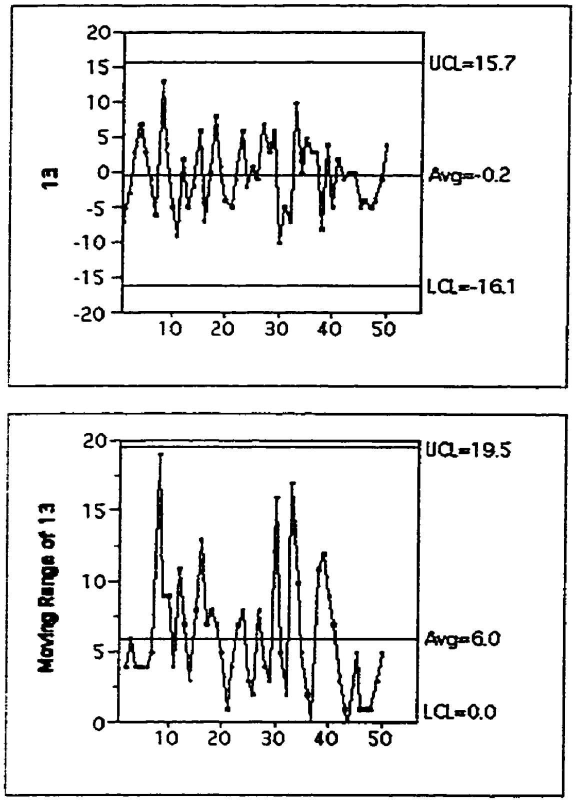



Figure 28n: X and MR Control Charts for Stream 14

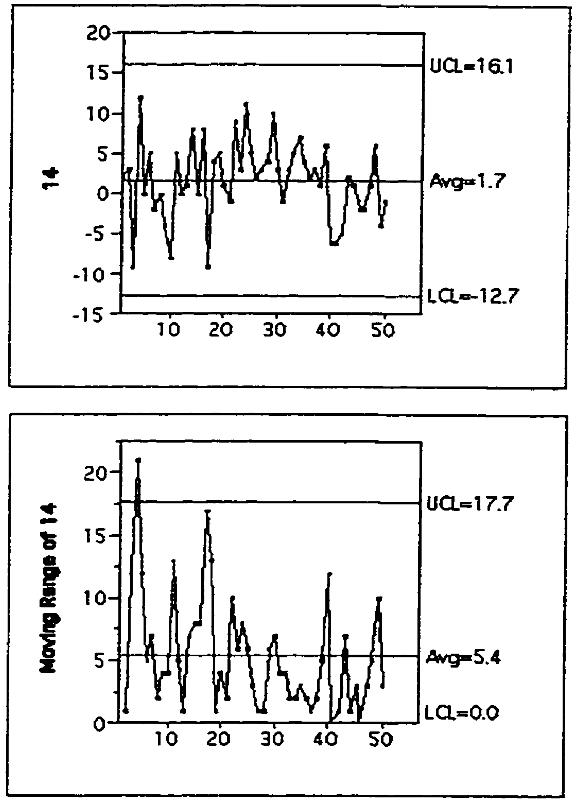
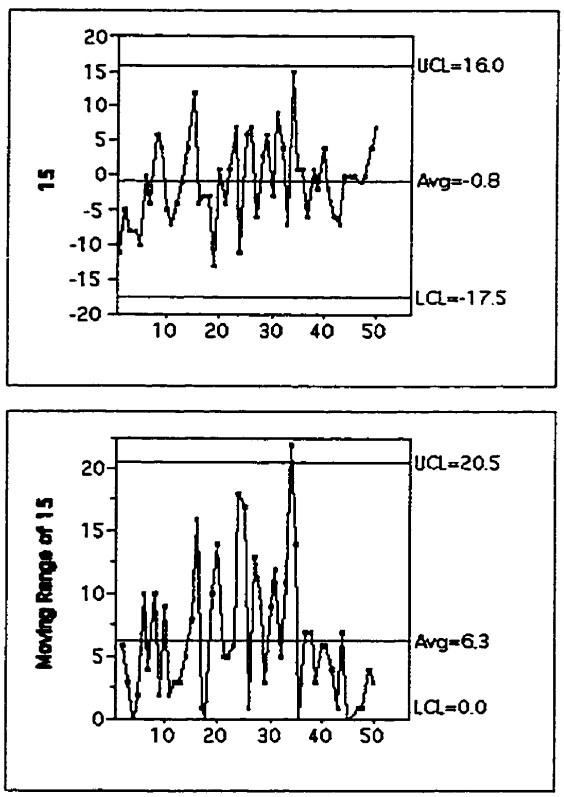



Figure 28o: X and MR Control Charts for Stream 15

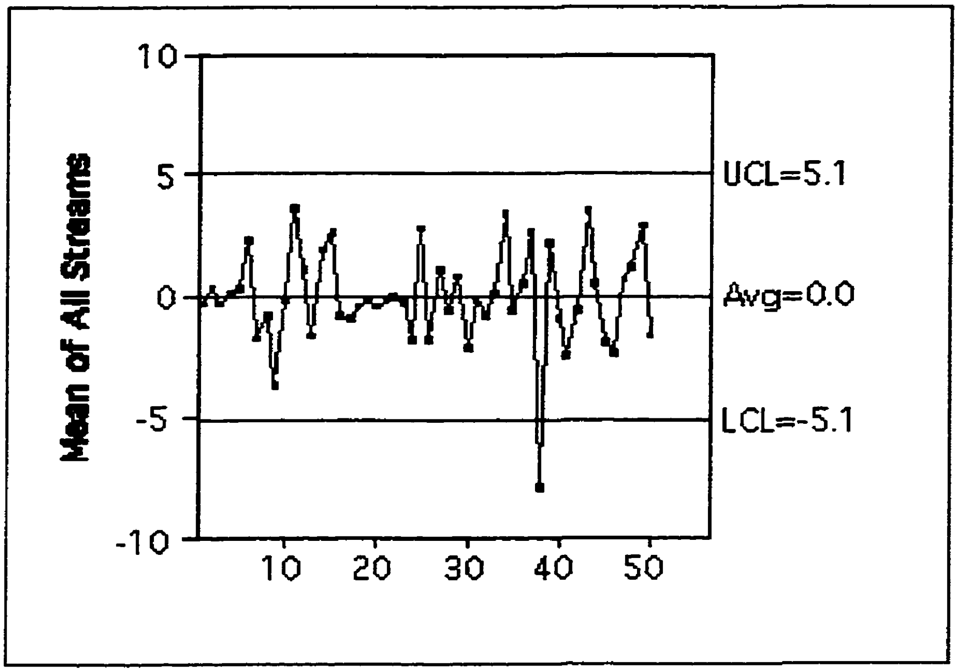


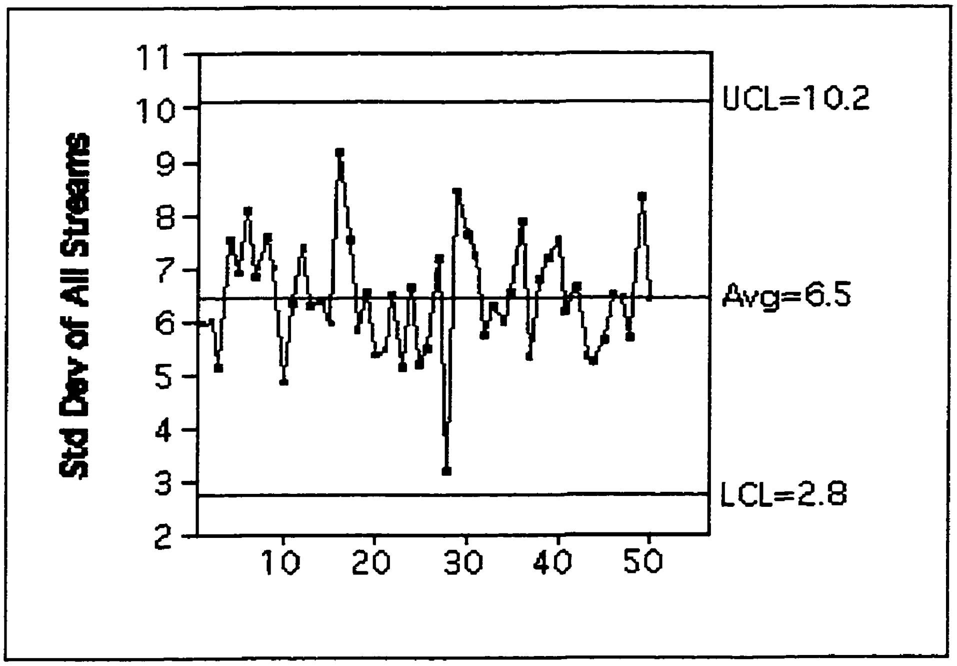
### 5.3.4 Control Chart Using the Streams as a Subgroup

If we were to incorrectly treat the 15 values from the streams at any time period as a subgroup and plot the subgroup averages and standard deviations on a $\overline{X}$ and S chart, only one signal would be given. Figure 29 shows that an out-of-control signal occurs at period 38, the same time period that provided a signal on the control charts using Methods 1 and 3.

This chart fails to detect other time periods when problems may have arisen or to identify which streams are operating differently.

Figure 29: Control Charts Treating the Streams as a Subgroup



# 6. Conclusion

In this practicum, we have primarily considered the group control chart and modifications to it for monitoring multiple-stream processes. To assess the performance characteristics of the four charts, simulation studies were done in order to determine in-control and out-of-control average run lengths. While it isn't possible to categorically state which of the four methods performs best, the average run lengths suggest that two of the methods perform quite well.

The traditional group control chart (Method 3) has excellent out-of-control detection characteristics but its in-control performance is unacceptable. Method 3 gives a false signal when in-control too frequently. A modification involving altering the control limits (Method 2) has very large in-control ARL values, but its out-of-control detection is very poor.

The group control chart modification using control limits based upon the mean and standard deviation of the maximum (Method 1) showed good out-of-control ARL values and acceptable in-control ARL values. The modification using control limits determined from the empirical distribution of the maximum obtained by simulation (Method 4) had excellent in-control ARL performance, equivalent to the usual Shewhart $\overline{X}$ control chart, and acceptable out-of-control ARL values.

The four methods were applied to three data sets to illustrate their application. Consistent with the average run length analysis, the traditional group control chart found the greatest number of out-of-control signals. Some of these signals may well be false alarms.

# Appendix A

Control Chart Constants[1]

| n | $d_2$ | $A_2$ | $E_2$ | $D_3$ | $D_4$ |
|---|---|---|---|---|---|
| 2 | 1.128 | 1.880 | 2.660 | 0 | 3.267 |
| 3 | 1.693 | 1.023 | 1.772 | 0 | 2.574 |
| 4 | 2.059 | 0.729 | 1.457 | 0 | 2.282 |
| 5 | 2.326 | 0.577 | 1.290 | 0 | 2.114 |
| 6 | 2.534 | 0.483 | 1.184 | 0 | 2.004 |
| 7 | 2.704 | 0.419 | 1.109 | 0.076 | 1.924 |
| 8 | 2.847 | 0.373 | 1.054 | 0.136 | 1.864 |
| 9 | 2.970 | 0.337 | 1.010 | 0.184 | 1.816 |
| 10 | 3.078 | 0.308 | 0.975 | 0.223 | 1.777 |

[1] Source: Cheng and Fu (1994)

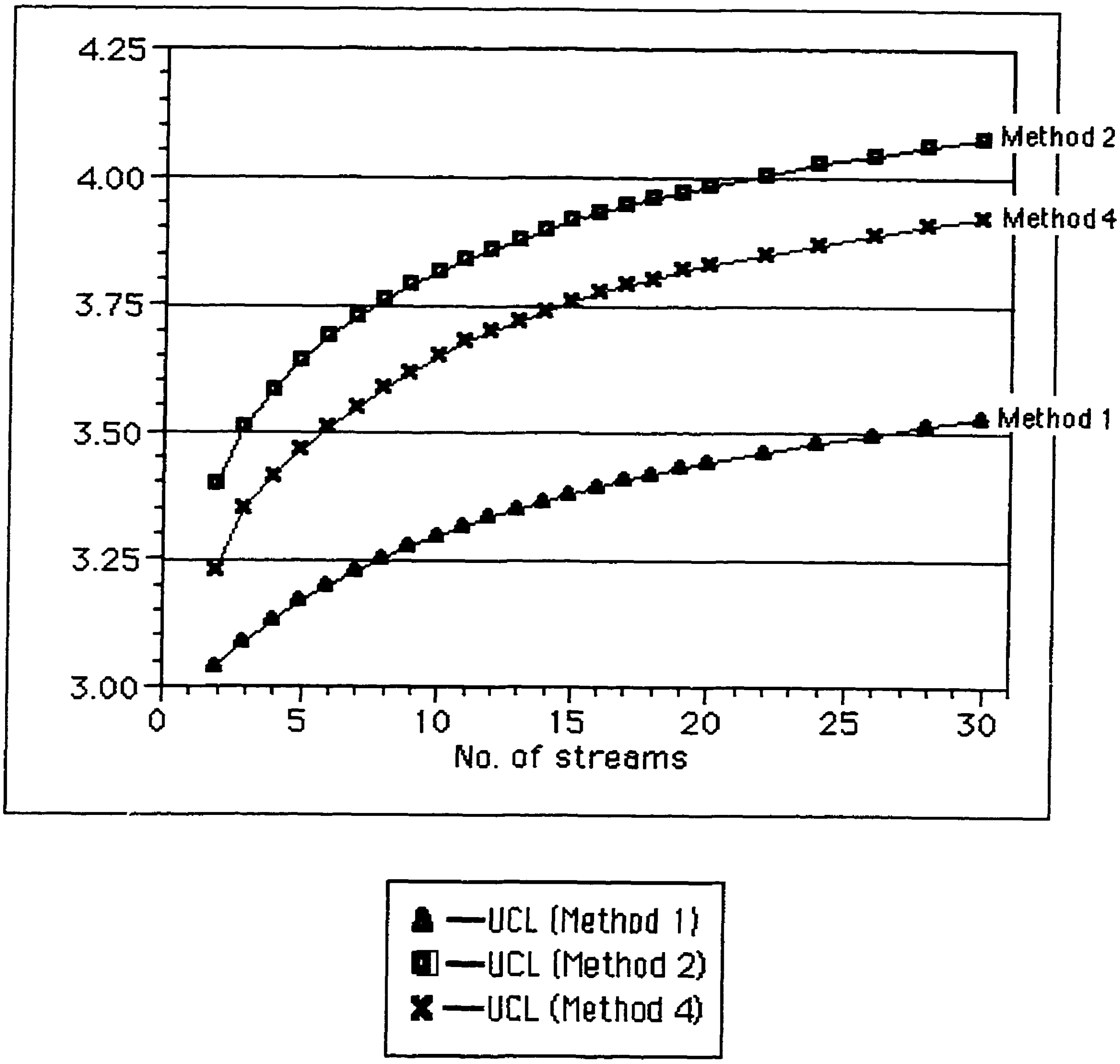
# Appendix B

Upper Control Limits for Standard Normal Processes (n=1)

| s | $UCL^1(s)$ | $UCL^2(s)$ | $UCL^4(s)$ |
|---|---|---|---|
| 2 | 3.041 | 3.399 | 3.23 |
| 3 | 3.090 | 3.509 | 3.35 |
| 4 | 3.133 | 3.585 | 3.41 |
| 5 | 3.170 | 3.642 | 3.47 |
| 6 | 3.202 | 3.689 | 3.51 |
| 7 | 3.230 | 3.728 | 3.55 |
| 8 | 3.256 | 3.762 | 3.59 |
| 9 | 3.278 | 3.791 | 3.62 |
| 10 | 3.299 | 3.817 | 3.65 |
| 11 | 3.318 | 3.841 | 3.68 |
| 12 | 3.336 | 3.862 | 3.70 |
| 13 | 3.352 | 3.881 | 3.72 |
| 14 | 3.368 | 3.899 | 3.74 |
| 15 | 3.382 | 3.916 | 3.76 |
| 16 | 3.395 | 3.932 | 3.78 |
| 17 | 3.408 | 3.946 | 3.79 |
| 18 | 3.420 | 3.960 | 3.80 |
| 19 | 3.432 | 3.973 | 3.82 |
| 20 | 3.443 | 3.985 | 3.83 |
| 22 | 3.463 | 4.007 | 3.85 |
| 24 | 3.482 | 4.028 | 3.87 |
| 26 | 3.499 | 4.047 | 3.89 |
| 28 | 3.515 | 4.064 | 3.91 |
| 30 | 3.530 | 4.080 | 3.92 |

Values for $UCL^i{}_{(s)}$ (i=1,2,4) are given for 2 to 30 streams.
Note that $UCL^3{}_{(s)} = 3.0$ for all values of s.

Figure 30: Overlay Plot – Number of Streams vs. UCL
for Methods 1, 2, and 4



## References


Bingham, T. C. (1999). Private communication.

Cheng, S. W. and Fu, J. C. (1994). *Statistical Tables for Classroom and Exam Room,* 2nd ed. University of Manitoba, Department of Statistics.

Godwin, H. J. (1949). "Some Low Moments of Order Statistics". *Annals of Mathematical Statistics* 20, pp. 279-285.

Lanning, J. W.; Montgomery, D. C.; and Runger, G. C. (1998). "Monitoring a Multiple Stream Filling Operation Using Fractional Samples". Technical Report, Arizona State University, Department of Industrial Engineering.

Montgomery, D. C. (1996). *Introduction to Statistical Quality Control,* 3rd ed. John Wiley & Sons, New York, NY.

Mortell, R. R. and Runger, G. C. (1995). "Statistical Process Control of Multiple Stream Processes". *Journal of Quality Technology* 27, pp. 1-12.

Nelson, L. S. (1986). "Control Chart for Multiple Stream Processes". *Journal of Quality Technology* 18, pp. 255-256.

Ott, E. R. and Snee, R. D. (1973). "Identifying Useful Differences in a Multiple-Head Machine". *Journal of Quality Technology* 5, pp. 47-57.

Sheskin, D. J. (1997). *Handbook of Parametric and Nonparametric Statistical Procedures,* CRC Press, Boca Raton, FL.