# Problem of non-normality in statistical quality control: a case study in a surface mine

by B. Samanta* and A. Bhattacherjee†


## Synopsis

**A quality characteristic analysed through construction of the Shewhart control chart is required to be normally distributed. In mining applications, this assumption is often violated as many quality characteristics follow skewed or lognormal distribution. In this paper, two methods were investigated using a heuristic approach for setting up the control limits for a skewed population. Based on these approaches a Mode chart, and a Weighted Variance (WV) control charts were constructed. The performance of the proposed methods was compared with the normal based Shewhart method through a case study conducted in a surface mine supplying iron ore to a steel plant. A simulation study was also conducted to evaluate the effectiveness of the proposed methods over the Shewhart method. The study revealed that the Shewhart control charts generated more out of control alarms compared with the Mode and Weighted Variance control charts for a skewed population of Fe%. The simulation results also revealed that for extremely high skewed population, the WV control chart and Mode control chart provided higher probability of coverage than the Shewhart method; although the detection rate (probability of detection) of out of control condition for these methods was slightly reduced compared with Shewhart method.**


## Introduction

Quality of ore varies in the complete spectrum of mining operations. Depending upon the quality control practices at mines, grade variation may follow a systematic pattern or an erratic fluctuation. The different statistical quality control techniques, especially control charts, are used for continuous monitoring of grade variation. The main purpose of the application of control charts is to identify the root causes of quality variation based on which corrective actions are taken to remove irregular grade fluctuation.

Some studies on the application of control charts were conducted in the mineral industry for improving the quality of ore (Khuntia *et al.*, 1991; Samanta, *et al.*, 1998). These studies explored the application of the Shewhart $\overline{X}$ and R charts. The constructions of these charts were based on the normality assumption of sample observations.

However, many quality characteristics in mining operations do not follow the normality assumption. If a characteristic is not normally distributed, but normal based techniques like the Shewhart charts are used, serious errors can be made. Yourstone and Zimmer (1992) showed that significant departures from normality can have serious effects on the error probability and interpretation associated with such charts. Particularly in the cases of ore deposits which are formed under complex geological phenomena, the process of mineralization largely affects the grade distribution. This process leads to formation of deposits, which are widely varying in nature. Specifically in erratic deposits, a zone with a low average value usually contains a small proportion of relatively high values and vice versa.

Distributions of widely varying ore grades usually do not follow the normality assumption but they appeared to be shaped as lognormal or skewed. For skewed populations, the Type I risk probabilities of the Shewhart control charts grow larger as the skewness increases.

One approach, which is widely practised to deal with this problem, is to increase the number of observations in a sample, as the statistical theory of construction of the Shewhart control chart is based upon the distribution of the sample means, which generally tends to follow the normal curve as sample size increases.

It was noted by Shewhart that the distribution of many individual measurements is non-normal, although the distribution of sample means of size four will, in many cases, follow the normal curve as predicted by the central limit theorem. Shewhart arrived at this conclusion on the basis of a set of experiments involving random drawings from populations


* *School of Mineral Engineering, UAF, USA.*

† *Indian Institute of Technology, Kharagpur, India.*


 *Paper received Jul. 2003; revised paper received Apr. 2004.*

with rectangular and right-triangular distributions (Grant and Leavenworth, 1980). However, according to Spedding and Rawlings (1994), these distributions did not represent significant departure from normality and more extreme population distributions may require larger sample sizes to achieve a normal distribution of the sample means. Chatfield (1976) suggested that the centre limit theorem should not be relied on for sample sizes of less than 30 for non-normal population; however, it may not be practically possible to go for large sample sizes due to increased sampling costs. Therefore, a method is needed to alleviate this problem by suitably modifying control charts with alternative approaches.

## Background

The problem of non-normality in control charting procedure was dealt with by several researchers. Chou *et al.* (1998) proposed an approach to solve this problem by transforming non-normal data to normality using Johnson's system of distribution. They used the sample quantile ratio, in conjunction with the Shapiro-Wilk test of normality, to find a suitable transformation for non-normal data. Burr (1967), Schilling and Nelson (1976), and Chan *et al.* (1988) examined the effect of non-normality on the Shewhart $\overline{X}$ and R charts for skewed population. Burr (1967) developed tables of quality control constants for various combinations of skewness and kurtosis factors while constructing the $\overline{X}$ and R charts. He recommended using his table for a skewed population. Ferrel (1958) proposed the geometric mid-range and geometric range charts for a lognormal population instead of the Shewhart $\overline{X}$ and R charts. Nelson (1979) derived control limits of median, range, scale and location charts for the Weibull distribution. Some researchers proposed heuristic approaches in dealing with the skewness problems (Cowden, 1957; Choobineh and Ballard, 1987; Bai and Choi, 1995) and suggested asymmetric control limits while constructing control charts. Cowden (1957) presented a split distribution method in which he divided the asymmetric distribution into two parts at its mode and treated each part as a half of a distinct normal distribution, each having the same mean but different standard deviations. Choobineh and Ballard (1987) proposed an interesting approach namely, weighted variance (WV) method, which was based on the semi-variance approximation of Choobineh and Branting (1986). This method provided asymmetric control limits of the Shewhart $\overline{X}$ and R charts for the skewed distributions using the standard deviations of sample means and ranges, which were estimated based on between sample variation. However, Abel (1987) suggested that the standard deviation should be estimated from within sample variation and not from between sample variation. Bai and Choi (1995) constructed the $\overline{X}$ and R charts constants for the skewed population based on the WV method. Further, a comparative evaluation of the Shewhart and WV methods was performed through a simulation study which revealed that the modified charts of the WV method for the skewed population performed better than the Shewhart method in detecting an out of control condition. In this paper, two heuristic control charts: (a) WV Control Chart, and (b) Mode Chart were investigated for the skewed distribution of a quality characteristic encountered in grade control.

It is to be recognized that various statistical control charts, applied in this study, have been used by several researchers in other industrial applications. The present study examines their suitability in mining applications. However, it is believed that these charts, except the Shewhart control chart, have not been presented before for quality control in the mineral industry.

## Weighted Variance (WV) Chart

The WV method is based on the idea that a skewed distribution can be split into two segments at its mean, and each segment is used for creating a new symmetric distribution (Choobineh and Ballard, 1987). The two new distributions created from the original skewed distribution have different standard deviations. The WV method uses these two distributions for setting up the limits of a control chart, one for computing standard deviation for the upper control limit (UCL) and the other for estimating standard deviation for the lower control limit (LCL). If the population is skewed to the right, then the distance of the UCL from the mean is larger than that of the LCL. Similarly, if the population is skewed to the left, then the distance of the LCL from the mean is larger than that of the UCL. For a symmetric population, however, the distances of the UCL and the LCL from the mean are the same, and the chart based on the W.V. method reduces to the Shewhart Chart.

The WV method, like the Shewhart method, uses the standard deviation of sample means to set up the limits of the control chart. However, it differs from the Shewhart method in that it considers two factors that account for the skewness of a distribution. One factor is used for the UCL, while the other factor is used for the LCL. Let $P_X$ be the weight to be assigned to the right side of mean $\mu_X$ so that the variance is split asymmetrically. Then, the UCL factor is $\sqrt{2P_X}$, and the LCL factor is $\sqrt{2(1-P_X)}$.

The control limits of the $\overline{X}$ chart based on the WV method are:

$$UCL_{\overline{X}} = \mu_x + 3\frac{\sigma_x}{\sqrt{n}}\sqrt{2P_x} \quad [1]$$

$$\text{Or, } UCL_{\overline{X}} = \mu_x + A_2\overline{R}\sqrt{2P_x} \quad [2]$$

$$CL_{\overline{X}} = \mu_x \quad [3]$$

$$LCL_{\overline{X}} = \mu_x - 3\frac{\sigma_x}{\sqrt{n}}\sqrt{2(1-P_x)} \quad [4]$$

$$\text{Or, } LCL_{\overline{X}} = \mu_x - A_2\overline{R}\sqrt{2(1-P_x)} \quad [5]$$

where, $\mu_X$ = overall mean of the sample means
$\sigma_X$ = standard deviation of the population
$A_2$ = control chart constant
$\overline{R}$ = mean range of the samples
$n$ = sample size.

Similarly, if the weight $P_R$ is assigned to the right side of the $\overline{R}$, then the UCL factor is $\sqrt{2P_R}$ and the LCL factor is $\sqrt{2(1-P_R)}$. Then, the control limits of the R chart are:

$$UCL_R = \overline{R} + 3\sigma_R\sqrt{2P_R} \quad [6]$$

$$\text{Or, } UCL_R = \overline{R} + d_3\overline{R}\sqrt{2P_R} \quad [7]$$

$$CL_R = \overline{R} \quad [8]$$

$$LCL_R = \overline{R} - 3\sigma_R\sqrt{2(1-P_R)} \quad [9]$$

$$\text{Or, } LCL_R = \overline{R} - d_3\overline{R}\sqrt{2(1-P_R)} \quad [10]$$

where, $\overline{R}$ = mean range of the samples
$\sigma_R$ = standard deviation of sample range
$d_3$ = control chart constant.

Construction of control charts using the WV method requires the estimation of the factors $P_X$ and $P_R$. For estimation of the factors $P_X$ and $P_R$, the following equations are proposed by the authors.

$$P_X = \frac{n_{RX} \cdot S_{RX}^2}{n_{LX} \cdot S_{LX}^2 + n_{RX} \cdot S_{RX}^2} \quad [11]$$

where, $n_{LX}$ = number of samples below the mean
$n_{RX}$ = number of samples above the mean
$S_{LX}^2$ = average sum squared deviation of sample values from the mean to the left side
$S_{RX}^2$ = average sum squared deviation of sample values from the mean to the right side.

$$P_R = \frac{n_{RR} \cdot S_{RR}^2}{n_{LR} \cdot S_{LR}^2 + n_{RR} \cdot S_{RR}^2} \quad [12]$$

where, $n_{LR}$ = number of range values below the mean range
$n_{RR}$ = number of range values above the mean range
$S_{LR}^2$ = average sum squared deviation of sample ranges from the mean range to the left side
$S_{RR}^2$ = average sum squared deviation of sample ranges from the mean range to the right side.

## Mode chart

The concept behind the construction of the Mode chart is based on the earlier theory of Watermeyer (1919) and Cowden(1957) on skewed population. According to them, a skewed distribution can be imagined as compounded halves of two symmetric normal distributions on either side of mode, each parts having different standard deviations. Hence, a skewed distribution can be spilt into two halves at mode, as if mode is now considered as mean, and the construction of upper and lower control limits can be based on the observed standard deviations of the values with respect to mode on the either side of the mode. According to Watermeyer (1919), the mode is determined to satisfy the condition that the standard deviations with respect to mode on either side thereof, divided by the corresponding sum of frequencies, are in equilibrium. Therefore, the following condition is satisfied at mode.

$$\frac{sdl}{ml} = \frac{sdr}{mr} \quad [13]$$

$$\text{Where, } sdl = \sqrt{\frac{1}{n_1}\sum\left(X_{il} - mode\right)^2}$$

$$sdr = \sqrt{\frac{1}{n_r}\sum\left(X_{ir} - mode\right)^2}$$

$ml$ = sum of frequencies of the values below mode
$mr$ = sum of frequencies of the values above mode
$x_{il}$ = values of the observations below mode
$x_{ir}$ = values of the observations above mode
$n_l$ = number of observations below mode
$n_r$ = number of observations above mode.

Therefore, the control limits of chart for the mode chart are:

$$UCL_{\overline{X}} = mode_x + 3\frac{sdr_x}{\sqrt{n}} \quad [14]$$

$$CL_{\overline{X}} = mode_x \quad [15]$$

$$LCL_{\overline{X}} = mode_x - 3\frac{sdl_x}{\sqrt{n}} \quad [16]$$

Similarly, control limits of R chart can be constructed from the distribution of R values, and the control limits of R chart will be:

$$UCL_R = mode_R - 3sdr_R \quad [17]$$

$$CL_R = mode_R \quad [18]$$

$$LCL_R = mode_R - 3sdl_R \quad [19]$$

## Case study

An investigation was carried out for a surface mine supplying iron ore to a steel plant. The quality norm set for the supply of Fe% from the mine to the steel plant is as follows: 63.5 ±0.5. In order to obtain improved grade control practices at the mine, it was required to continuously monitor the supplied ore grade of the despatched material. The objective of this study was to observe the occurrences of any unsystematic behaviour of the supplied ore grade in addition to finding out the root causes of quality variation using suitable control charts. For evaluation of this study, samples collected from the despatched point were used. Information about the grade of supplied ore material for Fe% was available for each of the rake (lots) separately, and it was used as an input data in this analysis.

### *Problem of non-normality*

The Shewhart $\overline{X}$ and R charts are generally prepared based on the assumption of normality. However, as previously discussed, if the quality characteristic under study does not follow the normal distribution, a gross error can be made on the interpretation of an out of control condition. Therefore, the normality assumption should be verified and if it is not satisfied, proper control limits should be set for determining an out of control condition. For verifying the distributional validity of the quality characteristic Fe%, a statistical analysis of the data set was conducted. One of the simplest diagnostic tests for normality is a visual check of the histogram that compares the observed data values with a normal distribution. Figure 1 shows a histogram plot of the variable Fe%. A visual inspection of this plot reveals that the distribution does not follow a normal curve as it is skewed towards the left side of the mean. Although the visual portrayal of a histogram is a good indication of non-normality, a more

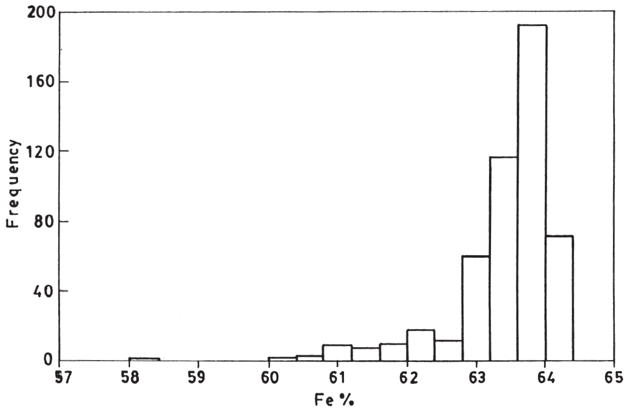


**Figure 1—Histogram of Fe% for individual observations**

reliable approach for the normality test is plotting the data on a normal probability paper. Figure 2 shows the normal probability plot of the variable, which reveals that the cumulative distribution of the data is not fitted by a straight line. In addition to examining the histogram and normal probability plot of the data set, the skewness and kurtosis values were also examined to observe the presence of non-normalitiy. The estimated values of skewness and kurtosis for the variable Fe% were -1.97 and 10.53, respectively. The estimated values clearly revealed that the distribution is far from normality. One can complement this analysis with various statistical tests which can be performed to assess the normality. One of the widely used tests, is the Kolmogorov-Smirnov test. The Kolmogorov-Smirnov test statistic showed that the estimated probability level for normality of Fe% was zero. This test strongly advocated that the normality assumption failed for the variable Fe% at $\alpha$=0.05 significance level. It also verified whether the data assumed lognormal distribution or not, since many quality characteristics in mining application follow lognormal distribution. Consequently, a 2-parameter lognormal distribution was tried to fit with the data. The test statistic also showed that the data were not lognormally distributed. In complement with 2-parameter lognormal distribution, it was also verified if the data could be fitted with the 3-parameter lognormal distribution, since some of the quality characteristics, especially gold values, are some times better represented by 3-parameter lognormal distribution (Krige,1981). In this regard, maximum likelihood estimator (Cohen,1951; Giesbrecht and Kempthorne,1976) was used to fit the data with the 3-parameter lognormal distribution. The study results also indicated that 3-parameter lognormal distribution was not fitted with the data.

For the skewed population, one of the approaches is to increase the sample size as the sample means follow the normal distribution with increased sample size, as predicted by the central limit theory. An earlier study on the central limit effect for a variety of populations conducted by Bradley (1973) showed that to achieve normality, small sample size sufficed when the sampled population was not appreciably skewed. For a highly skewed population, sample size running into hundreds or thousands may be required to fit the normality. He also inferred that the rapidity of the central limit effect in the general case has been greatly exaggerated. Similarly, Spedding and Rawlings (1994) showed in a study that a sample size of 30 was required for a reasonably well behaved population where skewness and kurtosis ranged from 0.9 to 1.0 and 2.5 to 3.0 respectively, otherwise sample sizes between 30 to 50 were required for a highly skewed population. In this study, to determine the validity of the central limit effect, the normality of the sample mean was investigated by increasing the sample size step by step. The normal distribution has a skewness of zero and a kurtosis of three, hence if the properties of the central limit theorem are to be used to obtain a normal distribution of the sample means, the shape statistics (skewness and kurtosis) should quickly assume their normal values with increasing sample sizes. In addition to that, if the Kolmogorov-Smirnov test is applied, then the 2-tailed probability will also increase as the distribution approaches normality. Table I shows the statistics of skewness, kurtosis, and 2-tailed probability of the K-S test for the variable Fe%. It reveals that there was no clear discernible trend found for the statistics of skewness and kurtosis against increased sample size. The 2-tailed

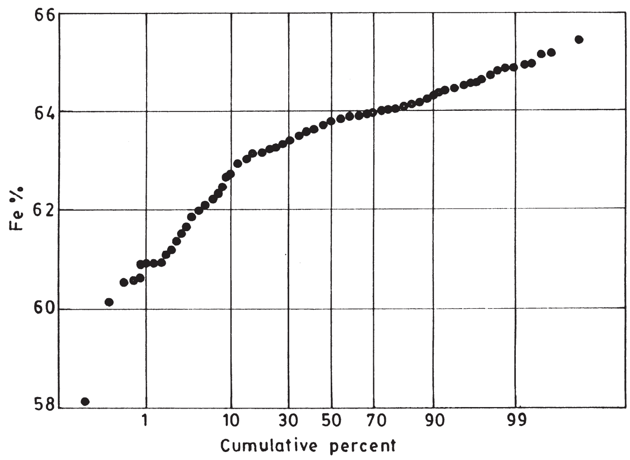


**Figure 2—Normal probability plot of Fe% for individual observations**

*Table I*

**Normality test of Fe% with increased sample size**

| Sample size | Skewness | Kurtosis | K-S Test | | |
|---|---|---|---|---|---|
| | | | Max difference (Absolute) | KS-Z | 2-tailed probability (*p*) |
| 01 | -1.974 | 10.530 | 0.1560 | 3.610 | 0.000 |
| 02 | -1.770 | 7.304 | 0.1537 | 2.174 | 0.002 |
| 03 | -2.030 | 10.857 | 0.1489 | 1.817 | 0.003 |
| 04 | 1.530 | 6.559 | 0.1298 | 1.475 | 0.026 |
| 05 | -2.195 | 11.406 | 0.2145 | 1.214 | 0.093 |
| 06 | -1.418 | 6.365 | 0.1293 | 1.206 | 0.108 |
| 07 | -0.946 | 3.684 | 0.1247 | 1.083 | 0.193 |
| 08 | -1.137 | 3.669 | 0.1716 | 1.384 | 0.043 |
| 09 | -1.357 | 5.579 | 0.1180 | 0.904 | 0.386 |
| 10 | -1.497 | 6.443 | 0.1470 | 1.062 | 0.209 |
| 15 | -1.130 | 4.708 | 0.1350 | 0.801 | 0.541 |
| 20 | -0.606 | 2.983 | 0.1170 | 0.569 | 0.902 |
| 25 | -1.160 | 5.383 | 0.1490 | 0.686 | 0.733 |
| 30 | -0.707 | 2.655 | 0.1120 | 0.461 | 0.983 |

probability of the K-S test statistics shows a somewhat better picture between the sample size and the probability $p$ of the K-S test.

As the normality assumption-as well as 2-and 3-parameter lognormal distribution, was not fitted with the data, two heuristic based non-parametric approaches using Mode and WV control charts were investigated.

### *Construction of control charts*

For constructing the control charts, five consecutive rakes' values were considered as observations of a sample as they represented average grade on a day-to-day basis. The control charts were prepared using the data for a period of 9 months. There are 525 observations available from the mine. Taking a subgroup size of five, the control charts were prepared for 105 samples. The distribution of the sample mean of size five also showed that data were highly skewed to the left side (Figure 3). As the shape of the distribution was deviating from the normal curve, the WV and Mode charts were used for construction of control limits. For construction of the WV control chart, the underlying distribution was split into two distributions at the mean value of Fe%, which was computed as 63.59. Using Equation [11], the value of $P_x$ was calculated as 0.29. The distribution of the range revealed that it was skewed to the right side. For construction of the R chart, the PR value was computed from the distribution of R, which was estimated as 0.68. Using Equations [1] through [10], the $\overline{X}$ and R chart limits were calculated for the WV chart.

For construction of the Mode chart, the distribution was split at mode, and the centre line, upper and lower control limits were calculated using Equations [13] through [19] for the $\overline{X}$ and R charts.

To compare the performance of the WV and Mode charts over the conventional Shewhart method, the Shewhart $\overline{X}$ and R charts were also constructed. Figure 4 presents the $\overline{X}$ chart of Fe% for the WV, Mode and Shewhart methods. Figure 4 shows that the upper and lower control limits for the WV and Mode charts are asymmetric control limits due to the skewness, and they are placed at 64.20 and 62.64 respectively for the WV chart, and 64.52 and 62.66 for the mode chart. For the Shewhart chart, the upper and lower control limits are placed at 64.39 and 62.78 respectively. Figure 4 also reveals that 4 samples are out of control below the lower control limit and one sample is above the upper control limit in the WV method; 4 samples are below the lower control limit of the Mode chart. On the other hand, in the Shewhart method, six samples are below the lower control limit. Similarly, the R chart of Fe% in Figure 5 shows that the upper control limit of the WV Chart, Mode chart and Shewhart chart is placed at 3.15, 4.47, and 2.92 respectively, and the lower control limit is placed at 0.17, .19 and 0 respectively. Five samples are out of control in the WV method, and only 1 sample is out of control in the Mode chart, whereas, six samples are out of control in the Shewhart method. Thus, it is observed that for the R chart there is a high discrepancy in the number of out of control conditions observed in the mode chart compared with other two charts.

It is reasonable to point out that the abnormal fluctuation portion of the variability was included in setting up the control limits of various charts since the out of control data points were also used in calculating the control limits. This resulted in a wider range of control limits than, expected which incorporates only the natural variability.

The real importance and significance of a control chart would be judged based on its ability to detect a real out of control against a false alarm. In real life practical applications, once a control chart triggers an out of control condition, one should find whether the alarm of out of control is a true indicator of abnormal fluctuation or falsely represents the normal fluctuation as an out of control condition. It may be reasonable to conclude that sample numbers 5, 39, 41 and 45 in the $\overline{X}$ chart indicated the

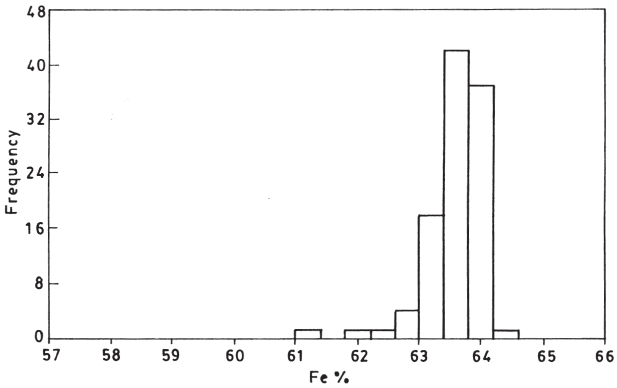


**Figure 3—Histogram of sample—means of subgroup size five**

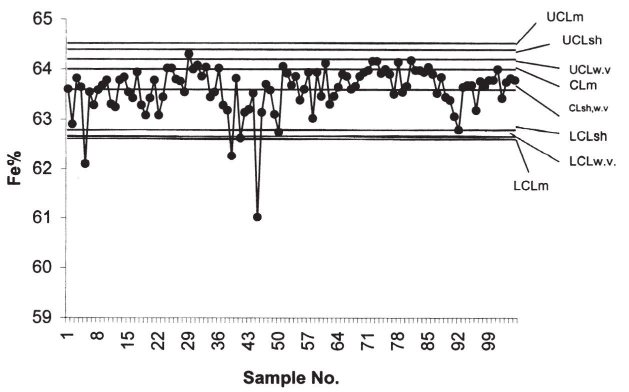


**Figure 4—$\overline{X}$ chart of Fe% for 105 samples**

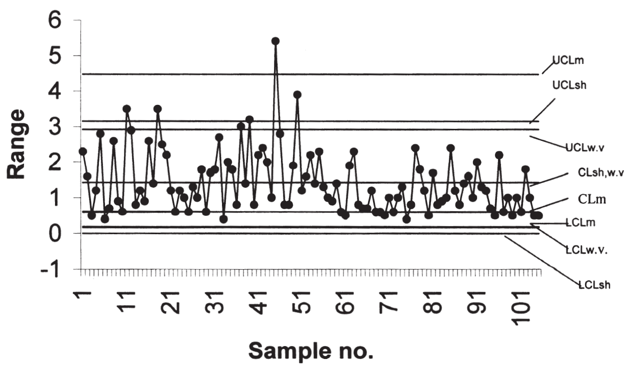


**Figure 5—R chart of Fe% for 105 samples**

abnormal fluctuations since all the control charts triggered out of control alarms. However, a problem occurred on deciding the true identity of sample numbers 29, 50 and 92 and whether they attributed real or merely false out of control conditions, since different control charts represented conflicting identities. Therefore, suitability of a particular control chart would lie in the true identification of these samples. A detailed discussion was held with the mine management about this issue. However, they could not provide any evidence/information to justify the true indications in the conflicting cases. Hence it was not possible to evaluate any legitimate comparison of various control charts from the case study application. This is also true for R control chart. Therefore, it was decided to conduct a simulation study to assess a comparative performance evaluation of various control charts considering a wide range of skewed populations. Although the case study application could not confirm that alternative control charts work better than the Shewhart chart for a skewed population, however, the simulation study showed an indication of improved performance of the WV and Mode control charts for highly skewed distribution (as will be presented later). It may be noted that performance evaluation of a control chart based on simulation study is a reasonable approach since no information was available from the case study.

It was previously mentioned that the objective of the application of a control chart was to find out the possible causes of the abnormal fluctuations of supplied grade from the mine, which are attributed to assignable causes by continuously monitoring the grade. A detailed discussion was held with the mine management to identify the possible assignable causes of the abnormal fluctuations of grade supplied from the mine. The authors were in full agreement with the mine management about the causes. The discussion revealed that the various assignable causes are the following: (i) improper estimation of *in situ* grade of ore, (ii) ore dilution, (iii) improper blending, and (iv) improper washing.

The grade control practices at the mine primarily depend on the estimated grades of the mining blocks. However, estimated grade differs considerably from the actual grade as the deposit is very erratic in nature. Further, ore dilution is another factor which virtually degraded the quality of ore. The deposit is so erratic that, in the same mining block, waste material is entwined with the ore material. Sometimes it was possible to segregate the waste from the ore by sorting, depending upon the visual inspection. Another reason for abnormal fluctuation of the grade is the improper blending. The mine is not in a position to blend the whole material regularly because of non-functioning of blending equipment at times. Even though washing is practiced at the mine to further improve the grade, it was revealed that the washing capacity is inadequate to wash all the material on a day-to-day basis. Moreover, the washing machinery frequently encounters breakdowns. Ultimately, all these causes led to abnormal fluctuations of the supplied ore grade. However, proper management action should be taken to remove these assignable causes.

## Comparative performance of charts using simulation study

The performance evaluation of the three types of control charts was carried out using a Monte Carlo simulation study. For conducting the simulation study, the Weibull distribution was selected for generation of data since it can take on a wide variety of shapes, from highly skewed to nearly symmetrical, as the shape parameter changes. For convenience of the study, the scale parameter of the Weibull distribution was chosen as 1.0. The effectiveness of the control charts was measured by the probability of coverage and detection rate. The probability of coverage (PC) computes the probability of the samples falling between the chart limits. High probability of coverage of a control chart indicates a reduced false alarm rate. On the other hand, detection rate implies how quickly a control chart detects an out of control condition when an actual out of control condition occurs. Therefore, a control chart based on a particular method will show an improved performance if it provides a higher detection rate along with a higher probability of coverage over the other methods.

For the simulation study, 5000 i.i.d (independent and identically distributed) observations were drawn from the Weibull distribution for each of the shape parameters studied. Then, taking a sample size of 5, control charts were prepared for 1000 samples. The probability of coverage and detection rate for each of the control chart methods was then computed. For the computation of probability of coverage and detection rate, the simulation for a particular shape parameter ran for 100 iterations. The probability of coverage and detection rate was evaluated by taking the average of the 100 simulation runs. For assessing the detection rate, three situations of out of control conditions were considered: (i) small shift in mean (1 sigma shift), (ii) moderate shift in mean (2 sigma shift), and (iii) large shift in mean (3 sigma shift).

Table II presents the values of probability of coverage using the three charting methods for the $\overline{X}$ and R charts. Tables III and IV show the detection rate of the three control charting procedures for the $\overline{X}$ and R charts respectively. Table II reveals that for both the $\overline{X}$ and R charts, the WV and Mode control charts consistently attain higher probability of coverage than the Shewhart chart. It is also observed that

*Table II*

**Probability of coverage for different combination of β values of the Weibull distribution for the Shewhart, WV and Mode methods of the $\overline{X}$ and R charts**

| β | Shewhart Chart | | WV Chart | | Mode Chart | |
|---|---|---|---|---|---|---|
| | PC | | PC | | PC | |
| | $\overline{X}$ | R | $\overline{X}$ | R | $\overline{X}$ | R |
| 0.25 | 0.9407 | 0.9002 | 0.9915 | 0.9913 | 0.9837 | 0.9852 |
| 0.50 | 0.9510 | 0.8872 | 0.9913 | 0.9885 | 0.9652 | 0.9810 |
| 0.75 | 0.9731 | 0.9216 | 0.9949 | 0.9902 | 0.9607 | 0.9882 |
| 1.0 | 0.9849 | 0.9525 | 0.9966 | 0.9921 | 0.9606 | 0.9924 |
| 1.5 | 0.9934 | 0.9845 | 0.9975 | 0.9949 | 0.9775 | 0.9960 |
| 2.0 | 0.9963 | 0.9946 | 0.9978 | 0.9969 | 0.9919 | 0.9977 |
| 2.5 | 0.9974 | 0.9967 | 0.9978 | 0.9972 | 0.9960 | 0.9979 |
| 3.0 | 0.9975 | 0.9980 | 0.9976 | 0.9979 | 0.9973 | 0.9985 |
| 3.5 | 0.9976 | 0.9982 | 0.9976 | 0.9981 | 0.9975 | 0.9987 |
| 4.0 | 0.9975 | 0.9979 | 0.9975 | 0.9979 | 0.9974 | 0.9986 |
| * | 0.9962 | 0.9908 | 0.9974 | 0.9964 | 0.9940 | 0.9973 |

* indicates γ (location ) = 58.0, α (scale) = 5.87, β (shape) = 9.67

*Table III*

**Probability of detection (detection rate) for the $\overline{X}$ Chart using Shewhart, WV and Mode methods**

| β | Shewhart Chart Detection rate | | | WV Chart Detection rate | | | Mode Chart Detection rate | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 sigma shift | 2 sigma shift | 3 sigma shift | 1 sigma shift | 2 sigma shift | 3 sigma shift | 1 sigma shift | 2 sigma shift | 3 sigma shift |
| 0.25 | 1.0000 | 1.0000 | 1.0000 | 0.9491 | 0.9994 | 1.0000 | 0.9874 | 1.000 | 1.000 |
| 0.50 | 0.9991 | 1.0000 | 1.0000 | 0.9812 | 0.9993 | 1.0000 | 0.9881 | 0.9900 | 0.9900 |
| 0.75 | 0.9980 | 1.0000 | 1.0000 | 0.9877 | 0.9996 | 1.0000 | 0.9988 | 1.0000 | 1.0000 |
| 1.0 | 0.9975 | 1.0000 | 1.0000 | 0.9914 | 0.9998 | 1.0000 | 0.9993 | 1.0000 | 1.0000 |
| 1.5 | 0.9970 | 1.0000 | 1.0000 | 0.9931 | 0.9999 | 1.0000 | 0.9989 | 1.0000 | 1.0000 |
| 2.0 | 0.9966 | 1.0000 | 1.0000 | 0.9948 | 0.9999 | 1.0000 | 0.9982 | 1.0000 | 1.0000 |
| 2.5 | 0.9961 | 0.9999 | 1.0000 | 0.9950 | 0.9999 | 1.0000 | 0.9971 | 0.9999 | 1.0000 |
| 3.0 | 0.9964 | 0.9999 | 1.0000 | 0.9964 | 0.9999 | 1.0000 | 0.9969 | 0.9999 | 1.0000 |
| 3.5 | 0.9969 | 0.9999 | 1.0000 | 0.9970 | 0.9999 | 1.0000 | 0.9969 | 0.9999 | 1.0000 |
| 4.0 | 0.9967 | 0.9999 | 1.0000 | 0.9970 | 0.9999 | 1.0000 | 0.9965 | 0.9999 | 1.0000 |
| * | 0.9966 | 1.0000 | 1.0000 | 0.9952 | 0.9999 | 1.0000 | 0.9979 | 1.0000 | 1.0000 |

* indicates γ (location ) = 58.0, α (scale) = 5.87, β (shape) = 9.67

*Table IV*

**Probability of detection (detection rate) for the R Chart using Shewhart, WV and Mode methods**

| β | Shewhart Chart Detection rate | | | WV Chart Detection rate | | | Mode Chart Detection rate | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 sigma shift | 2 sigma shift | 3 sigma shift | 1 sigma shift | 2 sigma shift | 3 sigma shift | 1 sigma shift | 2 sigma shift | 3 sigma shift |
| .25 | 1.0000 | 1.0000 | 1.0000 | 0.9064 | 0.9389 | 0.9698 | 0.9477 | 0.9794 | 0.9983 |
| .50 | 0.9981 | 1.0000 | 1.0000 | 0.9578 | 0.9798 | 0.9924 | 0.9738 | 0.9898 | 0.9975 |
| .75 | 0.9964 | 0.9994 | 1.0000 | 0.9625 | 0.9864 | 0.9954 | 0.9679 | 0.9887 | 0.9969 |
| 1.0 | 0.9957 | 0.9989 | 0.9999 | 0.9655 | 0.9905 | 0.9973 | 0.9625 | 0.9900 | 0.9973 |
| 1.5 | 0.9874 | 0.9968 | 0.9994 | 0.9639 | 0.9920 | 0.9982 | 0.9592 | 0.9900 | 0.9976 |
| 2.0 | 0.9760 | 0.9947 | 0.9990 | 0.9615 | 0.9923 | 0.9984 | 0.9491 | 0.9910 | 0.9980 |
| 2.5 | 0.9600 | 0.9939 | 0.9989 | 0.9568 | 0.9935 | 0.9989 | 0.9482 | 0.9928 | 0.9986 |
| 3.0 | 0.9572 | 0.9926 | 0.9987 | 0.9590 | 0.9922 | 0.9987 | 0.9539 | 0.9910 | 0.9984 |
| 3.5 | 0.9517 | 0.9931 | 0.9987 | 0.9582 | 0.9939 | 0.9989 | 0.9422 | 0.9926 | 0.9984 |
| 4.0 | 0.9593 | 0.9943 | 0.9983 | 0.9625 | 0.9946 | 0.9984 | 0.9498 | 0.9932 | 0.9980 |
| * | 0.9785 | 0.9950 | 0.9989 | 0.9622 | 0.9903 | 0.9977 | 0.932 | 0.9886 | 0.9975 |

* indicates γ (location ) = 58.0, α (scale) = 5.87, β (shape) = 9.67

probability of coverage is almost the same for the three methods when the population is nearly symmetrical. For a highly skewed population like β = .25, the probability of coverage of the WV chart and Mode chart is decisively higher than the Shewhart chart. However, it is also observed from Tables III and IV that the detection rate of WV and Mode charts is also degraded considerably, especially for the detection of a small shift (1 sigma shift). Considering the probability of coverage and the detection rate, the mode chart demonstrates the best performance for the β = 0.25. For the β values of 0.5, 0.75, the WV Method performs better than the Shewhart chart and the Mode chart for the $\overline{X}$ control chart, considering the probability of coverage and detection rate. In general it is seen that WV method performs better than the other two methods for the $\overline{X}$ chart for the skewed distribution, except for the β = 0.25. It can also be seen from the simulation study that for a highly skewed population, none of the control charts reached the probability of coverage of 0.9974 of ± 3 std control limits, although the WV and Mode charts are better than the Shewhart chart. For example, at the β value of 0.25, the probability of coverage of the Shewhart, WV and Mode charts are 0.9407, 0.9915 and 0.9837 respectively for the $\overline{X}$ control chart. It is also noteworthy to point out that in the case study application, the probability of coverage for the $\overline{X}$ chart of the Shewhart, WV and Mode methods are 0.9429 (6 out of 105), 0.9524 (5 out of 105), and 0.9619 (4 out of 105) respectively. These figures indicate that the WV and mode methods are somewhat closer to the ideal value of .9974 than the Shewhart method.

It is also obvious from the simulation study that the R chart is more affected than the $\overline{X}$ chart due to skewness. For the skewed population, the R chart of all three methods performs poorly compared with $\overline{X}$ chart.

The simulation study was also carried out using the estimated Weibull distribution parameters of the original data sets of Fe%. The result of this study is presented at the bottommost row of Tables II, III and IV. It is revealed that the probability of coverage, and the detection rate for the three methods are marginally different from each other for the $\overline{X}$bar and R charts. Therefore, it is also very difficult from a simulation study to attach any superior performance of a particular chart for the type of skewed population observed in the case study mine. However, application of the control charts on the real data set of Fe% showed high discrepancy in the three methods, especially for the R chart of Mode method. The reason might be the following: (a) inaccurate estimation of the parameters using small number of sample observations (525 observations), (b) influence of some of the extreme values on the estimation of the parameters of the control charts, and (c) lack of fitting of Weibull distribution to the observed Fe%, resulting in inaccurate estimation of shape parameter. However, the simulation study clearly demonstrated superior performance of WV and Mode charts for the highly skewed population over the Shewhart method. It is also revealed from the simulation study that the three methods worked equally well when the distribution is symmetric or less skewed. However, WV and Mode charts consistently performed better than the Shewhart chart when the distribution was highly skewed. It is really important to note that the WV and Mode chart never performed worse than the Shewhart chart, whatever the shape of the distribution. Hence, the WV and Mode methods are proved to be better for highly skewed situations, even though the case study application detracts from proving their superior performance Therefore, two proposed methods could be used as alternatives to the Shewhart method in constructing the control chart in mining applications for a highly skewed distribution such as gold.

## Conclusion

An application of the Shewhart control charts without verifying the normality assumption will produce a gross error as many quality characteristics in mining do not follow the normal distribution; rather, they are skewed. It is inferred that for a skewed population, the application of the WV and Mode charts will be useful for constructing the control charts as they provide asymmetric control limits in accordance with the actual variability pattern of the population. For the case study mine, it was revealed that the normality assumption of the variable Fe% was violated. The behaviour of control charts of the WV Mode and Shewhart charts is quite different. However, the simulation study on the three methods revealed that, although for highly skewed populations WV and Mode charts provided superior performance, the type of skewed distribution observed in the case study mine provided little clue to deciding the superior performance of either of the methods. However, it is suggested that the modified control charts using the WV method and Mode might be a viable alternative for a highly skewed population.

## Acknowledgement

The authors thank the management of the case study mine for their extended cooperation. The authors also sincerely acknowledge the two referees for their valuable comments and suggestions for improving the quality of the paper.

## References

1. Abel, V. Comment on: Control-Limits of QC Charts for skewed Distribution Using Weighted Variance, *IEEE Transactions on Reliability*, vol. 28 , no. 2, 1989. p. 265.
2. Bai, D.S. and Choi, I.S. $\overline{X}$ and R Control Charts for Skewed Populations, *Journal of Quality Technology*, vol. 27, no. 2 , 1995. pp. 120–131.
3. Bradley, J.V. The central limit effect for a variety of populations and the Influence of population moments, *Journal of Quality Technology*, vol. 5, no. 4, 1973. pp. 171–177.
4. Burr, I.W. The Effect of non-normality on constants for X and R charts, *Industrial Quality Control*, vol. 24, 1967. pp. 563–569.
5. Chan, L.K., Hapuarachchi, K.P., and Macpherson, B.D. Robustness of $\overline{X}$ and R charts', *IEEE Transactions on Reliability*, vol. 37, 1988. pp. 117–123.
6. Chatfield, C. *Statistics for Technology*, Chapman & Hall, London. 1976.
7. Choobineh, F. and Ballard, J.L. Control -Limits of QC charts for skewed distribution using Weighted Variance, *IEEE Transactions on Reliability*, vol. R-36, no. 4, 1987. pp. 473–477.
8. Choobineh, F. and Branting, D. A simple approximation for semivariance, *European Journal of Operational Research*, vol. 27,1986. pp. 364–370.
9. Chou, Y., Polansky, A.M., and Mason, R.L. Transforming nonnormal data to normality in statistical process control, *Journal of Quality Technology*, vol. 30, no. 2, 1998. pp. 133–141.
10. Cohen, A.C. Jr. Estimating parameters of logarithmic-normal distributions by maximum likelihood, *Journal of American Statistical Association*, vol. 46, 1951. pp. 206–212.
11. Cowden, D.J. *Statistical Methods in Quality Control*, Prentice Hall, Englewood Cliffs, NJ. 1957.
12. Ferrel, E.B. *Control Charts for log-normal universe*, *Industrial Quality Control*, vol. 15, 1958. pp. 4–6.
13. Giesbrecht, F. and Kempthorne. Maximum Likelihood estimation in the three-parameter Lognormal distribution, *Journal of the Royal Statistical Society. Series B* (Methodological), vol. 38, no. 3. 1976. pp. 257–264.
14. Grant, E.L. and Leavenworth, R.S. *Statistical quality control*, Fifth edition, McGraw-Hill, 1980. 684 pp.
15. Khuntia, G.S., Mishra, P.K., and Khare, A.K. Control charts—Statistical tool for quality control, *Transactions of the Mining, Geological and Metallurgical Institute of India*, vol. 88, no. 2, 1991. pp. 152–161.
16. Krige, D. Lognormal—De Wijsian geostatistics for ore evaluation, *South African Institute of Mining and Metallurgy*, Johannesburg, 1981. 51 pp.
17. Nelson, P.R. Control Charts for Weibull Process with Standards given, *IEEE Transaction on Reliability*, vol. R-28, 1979. pp. 114–116.
18. Samanta, B., Bhattacherjee, A., and Ray, S.C. A statistical quality control model for surface mine operations, *Proceedings of the International Conference on Surface Mining*, 9–11th July, 1998, Calcutta, 1998. pp. 306–320.
19. Schilling, E.G. and Nelson, P.R. The Effect of nonnormality on the Control Limits of Charts, *Journal of Quality Technology*, vol. 8, no. 4, 1976. pp.179–193.
20. Spedding, T.A. and Rawlings, P.L. Nonnormality in Statistical Process Control measurements, *International Journal of Quality & Reliability Management*, vol. 11, no. 6, 1994. pp. 27–37.
21. Watermeyer, G.A. Application of the theory of probability in the determination of ore reserves, *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, vol. XIX, no. 7, 1919. pp. 97–108.
22. Yourstone, S.A. and Zimmer, W.J. Nonnormality and the design of control charts for averages, *Decision Sciences*, vol. 23, 1992. pp. 1099–1113. ◆