**The effects of large round-off errors on the performance of control charts for the mean when the quality characteristic is normally distributed with a known variance**

Benson-Karhi Diamanta[[1]](#footnote-1)\*1, Ben-David Michal1,2, Levi Ofer1, Schechtman Edna2

1Department of Mathematics and Computer Science, the Open University of Israel,

P.O.B. 808, Raanana 43107, Israel

2Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, P.O.B. 653, Beer-Sheva 84105, Israel

August 2023

**ABSTRACT**

This research discusses the effects of large round-off errors on the performance of control charts for means when a process is normally distributed with a known variance and a fixed sample size. Quality control in practice uses control charts for means as a process monitoring tool, even when the data is significantly rounded. The objective of this research is to demonstrate how ignoring the round-off errors and using a standard Shewhart chart affects the quality control of a measured process.

The first part of the research includes theoretical calculations for estimating the values of alpha and beta, relating to the unrounded data and the large-rounded data. For the rounded data, normality can no longer be assumed because the data is discrete, therefore the multinomial distribution is used. The results show that given an in-control process, alpha indicates that false alarms are more frequent, whereas given an out-of-control process, the influence on beta is minor and inconsistent. For some rounding levels, there is a decline in the control chart performances, and in others, there is an improvement. In the second part, a simulation study is used to evaluate the performances of the control chart based on a single sample, checking whether the conclusion (reject or fail to reject) for a sample is consistent for rounded and unrounded data. The results of the simulation match the theoretical calculations.

***Key Words:***

 ***Average Run Length (ARL), Control Chart, Control Limits, Large Round-Off, Measurement Error, Round-Off Error***

**Introduction**

Shewhart control charts are an important tool for monitoring processes in an ongoing manner. There is a natural variability in process variables in every production process due to random causes. A process where the variability is exclusively due to random causes is one that is statistically controlled. At times however, the variability not only stems from random causes but can also be attributed to problems with the machinery, the worker, the raw material, or other influencing factors. The manufacturer chooses the variables to monitor, which are known as the quality characteristics. The primary role of control charts is to monitor parameters such as the mean and variance of the quality characteristics, in an ongoing manner, with the primary goals being to quickly identify deviations in the process, estimate the process parameters, and determine its capacity, while minimizing the incidences of false alarms.

There is an increasing number of papers on statistical process monitoring (SPM) methods that contain misleading justifications for the methods they propose. These flawed methods and the incorrect theory associated with them threaten the integrity of the SPM research field. For example, methods have been proposed that involve using one chart statistic as the input into another chart statistic calculation instead of the original data. For example, Abbas et al. (2013) use the EWMA statistic as the input into a cumulative sum (CUSUM) chart in their approach. Similarly, the approach of Taboran et al. (2021) uses a moving average as the input into a double EWMA chart. Knoth et al. (2023) have shown that the justification of some of these ad hoc methods is inadequate. Studying the mixed MA-EWMA and EWMA-MA approaches of Sukparungsee et al. (2020) and others, Haq and Woodall (2023b) showed that the weightings assigned to data values using these charts increase with the age of the data, with some older data values given a considerably greater weighting than current values. This again leads to poor conditional expected delay and steady-state performance. In this paper we aim to demonstrate firstly how large round-off errors – such as those produced by the approaches mentioned above – affect the performance of control charts for means, and secondly how ignoring the round-off errors and using a standard Shewhart chart affects the quality control of a measured process. This study was the basis for a follow-up study in which we suggested a new SPM approach for large rounded data.

 Deviations caused by random and systematic measurement errors are generally disregarded when using control charts. However, significant measurement errors are liable to affect the control chart, and in some cases even lead to incorrect conclusions regarding the process monitoring. This could be due to an increased false alarm rate (indicating that the process has gone out of statistical control when there have been no deviations) or an increased probability of signal (the possibility of not recognizing a deviation when one in fact exists).

 The literature dealing with Shewhart’s classic control charts contains the assumption that the measured data are the true values of the monitored variable. Bennett (1954) addressed the matter in his research, proposing that when the measurement error variance was smaller than the process variance it could be disregarded, as its effect on the performance of the control chart was negligible. Abraham (1977) proposed calculating the control charts without taking measurement errors into account and then adding a fixed value to represent the measurement error. Kanazuka (1986) proposed increasing the sample size in order to solve the problem of the diminishing performance of control charts due to measurement errors. Walden (1990) also suggested solving the problem by increasing the sample size. As another possible solution, he proposed a process of repeated sampling, or a combination of increasing the sample size and performing repeated sampling. The results of their research along with others are mentioned in Linna and Woodall’s review article (2001).

 Wheeler (2001a) addressed control charts for characteristics when products are incorrectly classified as intact or defective due to measurement errors. He proposed improving the quality of the production process rather than investing resources in a perfect measurement system, arguing that improving the measurement systems when sorting products so they match the specifications would increase the overhead, while improving the production process would lower costs.

 The measurement process is influenced by many factors, such as environmental conditions, differences between measurers, rounding errors, calibration errors, errors in reading the results, and measurement times. Therefore, when a certain variable is measured several times, there is likely to be variability between the measurements. The different factors influencing this variability can be divided into two main categories: those stemming from the measurand itself, and those stemming from the measurement process and instruments, i.e. measurement errors (Gertsbakh, 2003).

Measurand Y can be described as:

 (1)

– the mean of X

– the difference between the mean of X and its true value

– the noise of the true value, the random error

 – the measurement error

The premise is that the two error components are independent (Gertsbakh, 2003).

In this paper we address measurement errors caused by rounding. Every measuring instrument displays measurements that are rounded-off, depending on the accuracy of the device. Round-off errors stem from two main causes. The first is the measurement device itself. Sometimes, due to technical, financial or other constraints, measurements are performed using cheap and fast measuring instruments with large measurement units. For example, if the measurement units of a scale are displayed in intervals of whole units (e.g. kilograms), the only results obtained will be whole numbers with no decimals, even though the value of the weighed variable could be a decimal (Gertsbakh, 2003). The second cause is the accuracy of the systems receiving the measurement results. For example, the measurement device may send a 25-digit value but the system reading the data is only able to preserve 18 digits (Zhidong et al. 2009).

One of the implications of rounding is that the final values obtained depend on the measurement units of the equipment and not the “true” value of the measured variable. This leads to rounded data: Y = with Y representing the rounded observations of X (Gertsbakh, 2003). Rounded data affects statistical analysis because the rounding-off process itself leads to a data discretization process. In other words, the data goes from being treated as a continuous random variable to being treated as a discrete random variable, and this has a significant impact on statistical inference.

In many cases it is safe to disregard rounding; the parameters can be estimated traditionally using rounded data and the estimation is considered accurate enough for the purpose of statistical inference. However, in some cases disregarding the nature of the round-off process can lead to significant inaccuracies in the evaluation of the parameters, and as a result, the use of theoretical statistical tools could lead to statistical inference errors (Gertsbakh, 2003). Consider for example a statistical process control (SPC) of the process variance where the random variable being monitored is the normally distributed capacity of a pierced irrigation pipe. If the capacity measurement instrument was originally designed for large pipes, using it on newer, much smaller pipes may cause significant round-off measurement errors, leading to an incorrect SPC inference.

In SPC, both the mean and the variance are generally unknown and need to be estimated, but there are situations where the mean can be assumed to be known. Although the standard deviation is a feature of the instrument, the mean can be controlled by the user by adjusting the instrument to a desirable value. On the other hand, there are cases where the standard deviation, which is an inherent characteristic of the machine, is given and is well controlled, while the mean cannot be easily calibrated and hence needs to be estimated and monitored (Montgomery, 2019).

The degree of rounding is determined by the ratio between the standard deviation of the measured variable (σ) and the measuring unit of the measurement device (h). The ratio is denoted by δ ( ), with the measuring unit of the measurement device defined as the difference between two consecutive values on it. It is generally presumed to be known.

The effect of a rounding error on the performance of the control chart depends on the size of δ. The smaller δ is, the cruder the rounding and the greater its influence. Gertsbakh (2003) presented some rules of thumb for classification: when δ > 2, the measurement process is considered regular and allows for the use of traditional theoretical statistical tools. When δ < 0.5, the rounding is considered crude and the measurement process is considered special; in such cases the use of theoretical statistical tools may lead to incorrect statistical inference and therefore adaptations must be made. When 0.5 < δ < 2, the round-off level has no particular definition and this is considered an intermediate state (Gertsbakh, 2003).

The true value of the measurand, X, and the rounded results, Y, can be expressed thusly:

(2)

 – the measurement error caused by crude rounding

We assume that the measurement errors () and the random error () are negligible compared to the crude rounding error. Therefore in Equation 2 replaces in Equation 1. The observations obtained are with .

Theoretically, Y can assume any value that is a multiple of h. However, in practice, when Y is measured with a measurement error stemming from crude rounding (δ < 0.5), it can assume no more than five different values with a significant probability (greater than 0.00001). The other values obtained will have a very small probability and can therefore be disregarded.

Under the assumption that the measurand prior to rounding, X, is normally distributed , the probability function of Y is:

(3) 

(Gertsbakh, 2003)

 – the mode, i.e., the most frequent rounded value (the one with the highest probability).

Benson et al. (2013) studied the estimation of the variance when the mean is known and the data is distributed normally with a measurement error stemming from crude rounding. Benson et al. (2015) then went on to study the estimation of the variance using a confidence interval as well, for data normally distributed with a known mean that was collected with a measurement error stemming from crude rounding.

 This paper aims to shed light on the effect of crude rounding errors on the performance of Shewhart control chart for the mean. The paper is structured as follows: in Section 2, we present the theoretical model for calculating the performances of Shewhart control charts for unrounded versus crudely rounded data. In Section 3 we compare, theoretically, the performances of control charts for rounded versus unrounded data. In Section 4, we present an analysis at the level of a single sample, using a simulation. Finally, in Section 5 we present our summary and conclusions.

**The theoretical model for calculating the performances of control chart for unrounded and crudely rounded data**

In order to compare the performances of a control chart for unrounded data with its performances for rounded data, we performed several theoretical calculations for the following indices:

Alpha (α) – The probability of type 1 error – the probability that the sample mean will fall outside the control limits when in fact the process mean has not shifted (i.e., the process is in-control).

Beta (β) – The probability of type II error – the probability that the sample mean will fall within the control limits when in fact the process mean has shifted.

We performed the calculations under the assumption that the observations were independent and the process was normally distributed with a known standard deviation and a fixed sample size.

In the first stage, the control limits were calculated for rounded and unrounded data, under the assumption that the process was in-control. When the data was unrounded, the indices were calculated using formulas for calculating probabilities of a normal distribution, as is customarily done for Shewhart’s control charts (Montgomery, 2019).

(4)

The beta values were calculated under the assumption that the process mean had deviated by k standard deviations. Using the control limits that were calculated under the assumption that the process was in-control, we calculated the probability of the shift in the mean not being identified, that is, the probability of the sample mean falling within the control limits despite a shift in the process mean having occurred.

 Formally, let k be the deviation of the mean from the original mean in standard deviation units, then:

 (5)

When the data is rounded it **is no longer** normally distributed, but discretely distributed according to the degree of rounding. Therefore, the indices were calculated using a multinomial distribution. First, we calculated the distribution of averages () that form the basis for all the calculations of rounded data. Then we calculated the grand average () and the control limits according to which we calculated the values of alpha (α) and beta (β).

The distribution of the averages () was calculated in several stages: **In the first stage** the distribution of rounded data Y was built by finding the upper and lower limits of the *X* values [the original values (] for every possible Y (a half h from each direction). In addition, the probability of obtaining y was calculated by

(6)

We obtained a table with five values for Y in intervals of h in each direction from the most frequent value of Y, h0, as well as the probability of each of these values (*P(Y*)). The other values had very small probability and could be disregarded. We represent the possible values for Y using a vector with five values (L):

with denoting the mode.

 **In the second stage,** sample size n was selected with replacement, and the mean was calculated for every combination of n possible values under the assumption that the order in which the data was sampled was of no importance.

 **In the third stage** we calculated the probability distribution of the mean (as several combinations could have the same mean) - a table showing the possible values and the probability of each of them occurring. This table served us later on in building the control chart.

 The number of combinations (b) was determined by calculating the number of partitions, the number of ways c out of d objects could be chosen, with repetitions and with no importance placed on the order. There are c+d-1 possible patterns, out of which c need to be chosen while taking into account the patterns that repeat themselves in a different order. In general:

b

Specifically, let us use LL to denote the number of possible values out of which values are chosen for a combination. In cases of crude rounding, there are no more than five values with significant probability, meaning d = LL = 5. The sample size is the number of objects selected (c = n), and we arrive at:

with

LL – as the length of vector L, LL = 5

n – as the sample size

b – as the number of combinations

For example, when the sample size is 7, the number of combinations is:

Let us denote:

i – an index of a random variable in the combination, i=1,…, n

j – an index of a combination out of total combinations, j=1,…, b

Yi,j –the *i-*th observationin combination j

We calculated the average ( for every combination (j), and calculated the probability (P*j*) using the multinomial distribution formula:

with

*l* – as the index of values on vector L (1…5)

 – as the number of times the *l*-th value of vector L appears in combination j.

Note that for every combination j, *.*

For example, if for n = 3 h = 2.5

The probability of each value of the vector L is calculated by

and is given by

[0.001,0.1056,0.7887,0.1056,0.0001].

For combination (7.5, 10, 10) we get

In the next stage we calculated the probability of each possible (the mean can be identical for several different combinations) by summing the P*j*values of identical values. This stage resulted in a probability table for all possible values. The values in the table appear in intervals of . In total the table has 4n + 1 values.

Next, we calculated the theoretical mean for rounded numbers according to the formula put forth by Benson et al. (2013):

The lower and upper control limits for rounded data (*LCLr, UCLr*)were calculated according to Shewhart’s formulas for control charts for the mean ( is assumed to be known):

A = 3/sqrt(n)

The value of alpha was calculated by summing the probabilities of the values that were outside the control limits when the process was in-control.

For every level of rounding, the beta values were calculated as dependent on the size of the shift of the mean in standard deviation units (k).

The method for calculating the distribution of means () given is identical to the method presented above. The value of beta was calculated by summing the probabilities of the values within the existing control limits.

**Comparing the performances of control charts for rounded versus unrounded data**

In this section, we present several examples of theoretical calculations of performance indices of control charts for unrounded and rounded data. The results presented are for unrounded values from a normal distribution with a mean of 10 and a standard deviation of 1, and for rounded data with crude rounding levels within the range of 0.3 < < 0.5, in intervals of 0.005, with a total of 41 rounding levels. For the purpose of illustration, we chose three sample sizes, n = 7, 15, and 25. A small n (e.g. n = 4), as often used for Shewhart charts, might not cover enough possible values of the discreet random variable Y, in which case the control chart of will not reflect the distribution of Y.

The sensitivity of the charts was tested with 12 deviation values (k), k = ± 0.1, 0.3. 0.5, 0.7, 1, 1.25.

 We note that only one combination of values of the mean and standard deviation is presented, as the charts’ performances are influenced by the mean only through the value of their central line (CL), and the value of the standard deviation also does not affect the indices. The size of the standard deviation bears no significance on alpha calculations, which are only influenced by the number of standard deviations according to which the control limits are determined. For beta calculations, the shift of the mean in relation to the process standard deviation is important. Therefore, we checked small, medium, and large shifts in relation to the standard deviation we chose.

***Analysis of the results of the alpha values***

Whereas for unrounded data the value of alpha is determined by the user and is known in advance, for rounded data the distribution is not normal and the value of alpha is not set and known in advance. In Figure 1 we present a comparison of alpha values for rounded and unrounded data.

[Figure 1 near here]

The chart above clearly demonstrate that the performances of the control chart are diminished when the data is rounded. The value of alpha is higher for every rounding level within the studied range for each of the sample sizes tested. This results in significantly higher false alarm rates that impede the continuity of work and the stability of the production line.

 Two salient phenomena can be seen in this chart. Firstly, there are spikes—sharp ups and downs appearing on the chart. This phenomenon appears for every rounding level in all three sample sizes. Secondly, there is a difference between the alpha values when the rounding level is between 0.32–0.38 and when it is in the ranges 0.3–0.32 and 0.38–0.5, with the damage to the performance being more prominent when the round-off level is in the range 0.32–0.38.

 In an attempt to understand the spiking in alpha values we repeated the experiment while **neutralizing the relative position** of the mean within the interval created as a result of the level of rounding. We set the mean precisely at the mode of the rounded value (in other words, the mean of X was equal to the mode of Y). The mode h0 was set as and the other possible values were set around it.

L= [,, , , ]

As can be seen in Figure 2, when the mean was set exactly on the mode, the chart obtained presents sawtooth behavior for each of the three sample sizes.

[Figure 2 near here]

Alpha calculations depend on the rounding level, the sample size, and the control limits of the control chart for the mean. Each rounding level δ has a fixed number of values (the average of a combination of size n out of all possible values), for a specific scale size (h) and sample size (n). The size of the interval between values (the difference between two adjacent values)

 is .

For example, for n = 7 and a rounding level of δ = 0.4 (i.e., h = 2.5), the interval size between values is . If we use D to denote the number of times the size appears within the control limits, we arrive at:

When the value of D is an integer, this indicates a spike, or in other words, the start of a new round of values in relation to the control limits. For example, for a sample size of 7, D receives a value of 5 when δ = 0.315. In the next round, the size (δ = 0.4410) is entered once again on each side of the control chart so that D = 7. In this situation, the probability of being within the control limits significantly rises, and therefore the alpha values drop.

[Table 1 near here]

The table above shows the values of rounding levels that contain complete rounds of values within the range of the control limits and the central axis for the various sample sizes. It is possible to see that within the range of rounding levels studied, 0.3 < δ < 0.5, the number of times in which this situation occurs increases along with the sample size (when n = 7, it happens twice, as opposed to when n = 15, where it happens three times, and when n = 25, where it happens four times). In Figure 2 the two intervals of alpha values within the studied range for sample size 7 are at rounding levels 0.3150 and 0.4410, as noted in Table 1.

***Analysis of the results of the beta values***

The following charts present beta values for unrounded and rounded data at various round-off levels, according to sample and shift sizes. Twelve sensitivity tests were conducted for six shift sizes, k = ±0.1, 0.3, 0.5, 0.7, 1, 1.25.

[Figure 3 near here]

[Figure 4 near here]

[Figure 5 near here]

These charts present beta values for eight different round-off levels, together with beta values for unrounded data (marked in red), for three sample sizes. The three charts demonstrate that in some cases the beta values of rounded data are smaller than the beta values for the unrounded data (values that are under the beta chart for unrounded data, represented by the dashed red line). This means that the control chart’s performances improved in terms of beta when the data were rounded. On the other hand, in other cases the beta values for the rounded data were higher than the beta values for the unrounded data (values that are above the beta chart for unrounded data, represented by the dashed red line), meaning that the chart’s performances were diminished when the data were rounded. It is important to note, and we will expound on this below, that the loss in performance in terms of alpha values (i.e., increase) when rounding is much more profound than the gain (i.e., decrease) in terms of beta values.

 When the data is not rounded there is complete symmetry in beta values for deviations from the mean at identical sizes and with opposite signs. When the data is rounded, the symmetry is maintained for some of the rounding levels while for others the symmetry is marred.

**Analysis at the level of a single sample using a simulation**

In the previous section we presented general comparisons. In the current section we present comparisons at the level of a single sample. The goal of the simulation is to compare the percent of samples in which the outcome of the monitoring process is identical for unrounded and rounded data, versus the percent of samples for which the results of the monitoring process for unrounded data would be the opposite of those for rounded data. In other words, we aim to address cases where the sample mean is within the control limits when the data is unrounded, but falls outside the control limits when the data is rounded and vice versa.

[Table 2 near here]

States oo and ii represent cases where the round-off error has no impact (i.e. an identical result). State io represents cases where when the data is unrounded the sample mean is inside the control limits (i) and when the data is rounded the sample mean is outside the control limits (o). In an in-control process, this situation indicates an increase in alpha (diminished performances of the control chart) when the data is rounded. In an out-of-control process, this situation improves the performances of the control chart when the data is rounded, meaning it reduces beta.

 The oi state represents cases where the sample mean is outside the control limits when the data is unrounded and inside the control limits when the data is rounded. In an in-control process, this situation indicates a decrease in alpha (i.e., improved performances of the control chart) when the data is rounded. In an out-of-control process, this situation indicates diminished performances of the control chart when the data is rounded, meaning it increases beta.

 The value of alpha when the data is **unrounded** is the sum of both states in which the sample mean is outside the control limits when the data is unrounded, despite the process being in statistical control (the sum of the values in the left column in Table 2).

α = oo + oi

The value of alpha when the data is **rounded** is the sum of the two states in which the sample mean is outside the control limits when the data is rounded, despite the process being in statistical control (the sum of the values in the first row of Table 2).

αr = oo + io

The value of beta when the data are **unrounded** is the sum of the two states in which the sample mean is inside the control limits when the data is unrounded, despite the fact that the process is no longer in statistical control (the sum of the values in the right column in Table 2).

β = ii + io

The value of beta when the data are **rounded** is the sum of the two states in which the sample mean is inside the control limits when the data is rounded, despite the fact that the process is no longer in statistical control (the sum of the values in the second row of Table 2).

βr = ii + oi

[Table 3 near here]

*Simulation stages:*

*In an in-control process*

1. Drawing m sets (m = 100k) of various sizes of numbers from a normal distribution with known mean and standard deviation:

.

Seven sample sizes were tested to determine the influence of sample size on chart performances:

n = [7, 10, 15, 20, 25, 30, 40]

1. Calculating the lower and upper control limits (LCL, UCL) on chart

For each of the seven databases (the different n sizes), we built control limits calculated according to three standard deviations:

1. Testing for every sample j out of the m samples whether or not was within the control limits and summarizing the binary results in a table with m rows (each row representing the sample id number). If the value 1 is registered, otherwise the value 0 is registered.
2. Rounding the data set from stage a according to various δ levels.
3. Estimating the mean for the rounded numbers based on a naïve estimate—averaging the n rounded numbers.
* Calculating new control limits (LCLr, UCLr) for the mean (calculated as the distance of three standard deviations from the estimated mean).

A = 3/sqrt(n)

Testing for every sample j out of the m samples whether or not was within the control limits and adding the binary results to the table from stage c (adding columns according to the number of rounding levels). If the value 1 is registered, otherwise, the value 0 is registered.

1. Comparing between the results with and without rounding for each data set. We examined the rate of agreement between the two states (oo and ii), with and without rounding, and the rates of incidences of disagreement (states io and oi).

*In an out-of-control process*

For every sample size eight new sets of m samples were drawn, each from a normal distribution with a mean deviated by k orders of magnitude of a standard deviation.

Nine k values were tested, placed as positive and negative.

k = [0, 0.25, 0.3, 0.4, 0.5, 0.75, 0.9, 1, 1.25, 1.5]

We then repeated stages b–f for every data set.

Below is an example of the analysis, for rounding level δ= 0.355 and sample size n =25:

[Table 4 near here]

The table above shows that the value of alpha for the rounded data is significantly greater than its value for the unrounded data (11 times greater), whereas the value of beta for the rounded data increases less significantly compared to its value when the data is unrounded (for small shifts k=±0.5, it decreases).

[Figure 6 near here]

For a relatively large sample n=25 and a rounding level of 0.355, the rate of agreement in relation to the level of deviation from the mean (in absolute values) increases. In an in-control process, the rate of agreement is higher than 95%.

[Figure 7 near here]

The chart above demonstrates that for a relatively large sample n=25 and a rounding level of 0.355, in an in-control process,the performances of the charts are diminished. In almost 100% of the disagreements regarding the average’s position in relation to the limits, the sample averages were found to be outside the control limits when the data were rounded. In other words, the value of alpha increased when the data was rounded (state io).

When there was a change in the mean (in an out-of-control process), we found that in most cases where the results with and without rounding were not in agreement, the sample average was outside the control limits when the data was rounded, meaning the value of beta decreased (the io state) and the chart performances improved.

It appears that when the deviation from the mean is positive there are fewer states in which the averages are found outside the control limits when the data is rounded. However, here too, in most cases it is possible to see that the value of beta is smaller.

In conclusion, for a large sample size and a rounding level of 0.355 the performances of the control chart are diminished for alpha but improved for beta.

**Summary of the findings based on the simulation at the level of a single sample**

* The agreement rate regarding the average’s position in relation to the control limits when comparing unrounded and rounded data is very high, especially when the process is in-control. However, there is still a percentage of disagreements that affect the performances of the control chart.
* For the cases that were tested, we found that when the size of the sample was small, the rate of agreement decreased the more the deviation from the mean increased, or that there was a decrease followed by an increase (in most cases a relatively small increase) in the rate of agreements when the deviation increased. When the sample size was large, we found that the rate of agreement increased the more the deviation from the mean increased.
* In cases of disagreement regarding the average’s position in relation to the control limits, and no change in the mean (in-control process), in all the incidents we tested the vast majority of samples were found to be outside the control limits when the data was rounded. In other words, the value of alpha had increased significantly and the performances of the control chart had been diminished.
* In cases of disagreement regarding the average’s position in relation to the control limits, it was not possible to decisively determine the state of beta when the data was rounded. We saw cases where most of the averages were within the control limits (the value of beta increased, the performances of the control chart decreased) and cases where most of the averages were outside the control limits (the value of beta decreased, the performances of the control chart improved). We also saw cases of near equilibrium. In all cases the change for the better or for the worse in the value of beta was relatively small compared to the change in the value of alpha.
* We found that for small deviations there was generally an improvement in the performances of the chart. In other words, the deviation is discovered sooner when the data is rounded, whereas for medium and large deviations the performances of the charts were diminished in most cases. The quality controller expects to receive a signal that the process has gone out of statistical control when there is a deviation from the process. When the deviations are large, the signal is expected sooner compared to when they are small. When the data is rounded, the opposite process takes place: it takes longer to identify that the system has gone out-of-control when the deviations are large compared to when they are small.

The findings of the simulation at the level of the sample are in alignment with the general findings from the analysis of the control charts and the alpha and beta indices.

**Summary and conclusions**

For alpha, there is a significant decrease in the performances of the control chart when the data is rounded and the control limits are calculated using Shewhart’s tools. The direct implication is that false alarms occur more frequently, disrupting the work routine on the production line. Additional damage that may result from the large amount of false alarms is that quality controllers on the production line could ignore signals received during the control process.

 With regards to beta, there is no conclusive effect. Some of the rounding levels diminished the performance of the control charts, meaning it would have taken fewer samples to identify loss of statistical control had the data been unrounded. However other rounding levels actually improved the performances of the control chart, requiring fewer samples to identify the deviation compared to when the data was unrounded. Based on the analyses we conducted, we found that when the deviation from the mean was small there was an improvement, meaning the deviation was discovered sooner. Surprisingly, when the deviation was medium-sized or large, precisely when we would have expected to be informed of losing statistical control relatively quickly, the performances of the control charts were diminished, taking more samples to recognize the deviation.

 Our primary finding is that when the data is rounded and the control limits are calculated according to Shewhart’s classic theory, performances of the control charts in an in-control process are significantly diminished as shown by alpha, in relation to their values when the data is unrounded. In an out-of-control process, the performances of the charts are diminished or improved based on the size of the deviation from the mean. However, in all cases, the change in beta is relatively small compared to the change when the process is in-control.

Another important finding is that there was no conclusive direction of the change in beta values when the data was rounded. The analysis at the level of the single sample showed inconclusive results for all the samples. Where beta increased, we found substantial rates of samples behaving in the opposite manner. This finding indicates that standard control charts are inappropriate for crudely rounded data.

Acknowledgements:

Declaration of Interest: The authors report there are no competing interests to declare.

Biographical Notes:

**Dr. Diamanta Benson-Karhi** is a senior lecturer at the Department of Mathematics and Computer Science at the Open University of Israel. Dr. Benson-Karhi is the head of the Industrial Engineering and Management program. Her main background is in industrial engineering with expertise in innovative applied statistics tools, specifically statistical process control (SPC) methodologies. Her research includes original theoretical developments and unique applications in a wide range of disciplines, such as chemical engineering and medical/health fields.

**Michal Ben David** received a B.Sc in Industrial Engineering and Management, The Open University (2009); MBA, The Open University (2013); Thesis, "Development of control chart of the mean for variables measured with large round off errors", Ben Gurion University (2018). Michal Ben David is a Member of the IEM unit, the Open University of Israel and Shenkar College. She teaches: Quality Control, Work Study, Simulation, Projects Management, Operations Management and more. Michal is now a PH.D. student at Ariel University.

**Prof. Ofer Levi** is an Associate Professor in the department of Mathematics and Computer Science at the Open University of Israel. Ofer received his M.Sc. degree in Industrial Engineering and Management from Ben-Gurion University in Israel and his Ph.D. in Scientific Computing and Computational Mathematics from Stanford University. Ofer’s main expertise is developing computational and statistical tools and algorithms for effective solution of problems in various application fields such as imaging, signal processing and economics.

**Prof. Edna Schechtman** received a B.Sc. in Mathematics and Statistics, Hebrew University of Jerusalem (1971); M.A. in Statistics, Hebrew university (1976); Ph.D. in Statistics, Ohio State University (1980). Edna Schechtman is a professor emerita of statistics at Ben Gurion University, Israel. Her current research interests are in the field of measures based on the Gini index, and applied statistics in various areas such as quality control, medicine, traffic safety, economics and more.

References

Abbas, N., Riaz, M. and Does, R. J. M. M. 2013. Mixed Exponentially Weighted Moving Average – Cumulative Sum Charts for Process Monitoring. *Quality and Reliability Engineering International* 29 (3), 345–356.

Abraham, B. 1977. Control Charts and Measurement Error. *Annual Technical Conference of the American Society for Quality Control* 31, 370–374.

Bennett, C.A. 1954. Effect of Measurement Error on Chemical Process Control. *Industrial Quality Control* 10, 17–20.

Benson, D., E. Dvir-Harcabi, I. Regev, and E. Schechtman. 2013. Estimation of a Normal Process Variance from Measurements with Large Round-Off Errors. *IET Science, Measurement and Technology* 7. no. 3, 180–189.

Benson, D. and E. Schechtman. 2015. Using Measurements with Large Round-Off Errors Interval Estimation of Normal Process Variance. *IET Science, Measurement and Technology*, 1–7.

Gertsbakh, I. 2003. *Measurement Theory for Engineers*. Berlin: Springer Verlag.

Kanazuka, T. 1986. The Effects of *Measurement* Error on the Power of X̄ - R Charts. *Journal of Quality Technology* 18, 91–95.

Haq, A. and Woodall, W. H. 2023b. A Critique of the Use of Modified and Moving Averagebased EWMA Control Charts. *Quality and Reliability Engineering International* 39, 1269–1276.

Knoth, S., Saleh, N. A., Mahmoud, M. A., Woodall, W. H., and Tercero-Gómez, V. G. 2023. A Critique of a Variety of ‘Memory-Based’ Process Monitoring Methods. *Journal of Quality*

*Technology* 55(1), 18-42.

Linna, K., and W. Woodall. 2001. Effect of Measurement Error on Shewhart Control Charts. *Journal of Quality Technology*, 213­–222.

Montgomery, D.C. 2019. *Introduction to Statistical Quality Control.* 8th ed. NewYork: Wiley. Sukparungsee, S., Areepong, Y. and Taboran, R. 2020. Exponentially Weighted Moving Average—Moving Average Charts for Monitoring the Process Mean. *Plos One*, 15(2), p.e0228208.

Taboran, R., Sukparungsee, S. and Areepong, Y. 2021. Design of a New Tukey MA-EWMA Control Chart to Monitor Process and its Applications. *IEEE Access* 9, 102746-102757.

Vardeman, S.B. 2005. Sheppard’s Correction for Variances and the Quantization Noise Model. *IEEE Transactions on Instrumentation and Measurement* 54, 2117–2119.

Walden, C.T. 1990. An Analysis of Variables Control Charts in the Presence of Measurement Error. Unpublished master's thesis, Department of Industrial Engineering, Mississippi State University, UMI no. 1343291.

Wheeler, D. 2011. 100% Inspection and Measurement Error. *Quality Digest*, May 3.

<https://www.qualitydigest.com/inside/quality-insider-column/100-inspection-and-measurement-error.html>.

Zhidong B., Z. Shurong, Z. Baoxue, and H. Guorong. 2009. Statistical Analysis for Rounded Data. *Journal of Statistical Planning and Inference* 139, 2526–2542.

**Tables**

|  |  |  |  |
| --- | --- | --- | --- |
|   | **n=7** | **n=15** | **n=25** |
| **D** | **δ** | **δ** | **δ** |
| **3** | 0.1890 | 0.1291 | 0.1000 |
| **5** | 0.3150 | 0.2152 | 0.1667 |
| **7** | 0.4410 | 0.3012 | 0.2333 |
| **9** | 0.5669 | 0.3873 | 0.3000 |
| **11** | 0.6929 | 0.4734 | 0.3667 |
| **13** | 0.8189 | 0.5594 | 0.4333 |
| **15** | 0.9449 | 0.6455 | 0.5000 |
| **17** | 1.0709 | 0.7316 | 0.5667 |
| **19** | 1.1969 | 0.8176 | 0.6333 |
| **21** | 1.3229 | 0.9037 | 0.7000 |

Table 1. Rounding levels by sample sizes where there are complete rounds of values with the mean in the middle of the interval.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |  |
| --- | --- | --- |
|  | Unrounded – outside limits | Unrounded – inside limits |
| Rounded – outside limits | oo | io |
| Rounded – inside limits | oi | ii |

Table 2. Simulation states for evaluating control chart performances at the level of a single sample.

|  |  |  |
| --- | --- | --- |
|  | Large io | Large oi |
|  |  |  |
| H0 | Large alphaDiminished control chart performances | Small alphaImproved control chart performances |
| H1 | Small betaImproved control chart performances | Large betaDiminished control chart performances |

Table 3. Meaning of states io and oi in an in-control process (H0) and in an out-of-control process (H1).



Table 4. Distribution of the states in an in-control process (H0) and in an out-of-control process (H1) and calculation of alpha and beta values; rounding level 0.355, sample size 25.

**Figures**

Figure 1:



Figure 2:



Figure 3:



Figure 4:



Figure 5:



Figure 6:



Figure 7:



**Figure Captions**

Figure 1. Alpha values for rounded and unrounded data by the rounding level (delta) and sample size.

Figure 2. ARL0 values for rounded and unrounded data by the rounding level and sample size.

Figure 3. Alpha values by the level of round-off and sample size when the mean is at the mid-interval.

Figure 4. Beta values for unrounded and rounded data by the size of shift and round-off level, n = 7.

Figure 5. Beta values for unrounded and rounded data by the size of shift and round-off level, n = 15.

Figure 6. Beta values for unrounded and rounded data by the size of shift and round-off level, n = 25.

Figure 7. The rates of agreement between unrounded and rounded data; rounding level 0.355, sample size 25.

Figure 8. The rates of disagreement (by state types) between unrounded and rounded data; rounding level 0.355, sample size 25.

1. \* Corres.: Diamanta Benson-Karhi, Department of Mathematics and Computer Science, the Open University of Israel. diamanta@openu.ac.il [↑](#footnote-ref-1)