Proceedings of the World Congress on Engineering and Computer Science 2013 Vol II
WCECS 2013, 23-25 October, 2013, San Francisco, USA

Performance of Conventional X-bar Chart for Autocorrelated Data Using Smaller Sample Sizes

D. R. Prajapati

***Abstract*—Control charts are used to determine whether or not a manufacturing or business process is in a state of statistical control. Various schemes of the standard (Shewhart) $\overline{X}$ chart for autocorrelated data at sample sizes of 2 and 3 are developed and compared with the schemes of the same chart for IID data at various process shifts in the process mean. It is clear from the comparison that the in-control ARL of 370 is achieved by employing the $\overline{X}$ chart for the IID and autocorrelated data but as the level of correlation increases, the performance of conventional (Shewhart) $\overline{X}$ chart deteriorates. It is concluded that for faster detection in the process mean, the larger sample sizes (n) may be used. Thus the $\overline{X}$ chart for autocorrelated data should be used with carefully and judiciously.**

***Index Terms*— Autocorrelated data, In-control ARLs, conventional $\overline{X}$ chart, Out-of control ARLs**

I. INTRODUCTION

Control charts; one of the important tools of quality control, are also known as Shewhart charts or process behavior charts. Control charts were developed by Dr. Walter F. Shewhart [1] in 1931. When control charts are used to monitor a process, it is assumed that the observations from the process output are independent and identically distributed (IID). However, for many processes; the observations are correlated and when this correlation builds-up automatically in the entire process, it is known as autocorrelation. In order to deal with autocorrelated data, modification in the design of control charts for monitoring the process mean is done. The performance of a control chart is measured in terms of average run length (ARL). Page [2] defined ARL as, *"as the average number of articles inspected between two successive occasions when rectifying action is taken"*

A. AUTOCORRELATION

One of the assumptions of implementing the chart is that the process outputs must be IID but usually there is some correlation among the data. When this correlation builds up automatically in the entire process, this phenomenon is called autocorrelation. The observations from the process output are usually positively correlated in most of the cases.

In this case, if the current observation is on one side of the mean, the next observation will most likely be found on the same side of the mean. Positively correlated data are characterized by runs above and below the mean. Positive correlation is more often encountered in practice than negative autocorrelation.

Autocorrelation is inherent to many processes like: chemical, manufacturing processes and service sectors. When $\overline{X}$ charts are applied to the autocorrelated data, the false alarm rate increases and performance of the chart is suspected. So improvement in the (Shewhart) $\overline{X}$ chart is needed to improve its performance for the correlated data. The following section deals with the literature review in this area.

B. CONVENTIONAL (SHEWHART) $\overline{X}$ CHART

In conventional $\overline{X}$ chart, means of small samples are taken at regular intervals, plotted on a chart, and compared against two limits. The limits are known as upper control limit (UCL) and lower control limit (LCL). These limits are defined as under:

LCL = $\overline{X}$ - 3σ'/√n
UCL = $\overline{X}$ + 3σ'/√n

The process is assumed to be in a state of out-of-control when the sample average falls beyond these limits. Figure 1 represents the plotting of sample averages on the $\overline{X}$ chart.

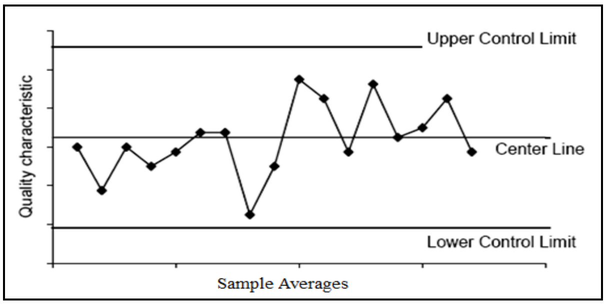

Fig. 1 Plotting of sample averages in the $\overline{X}$ chart

II. NOMENCLATURE

Following symbols have been used in this paper:

$\overline{X}$ = Target mean
σ' = Population standard deviation
n = Sample size
ARL = Average Run Length
UCL = Upper control limit
LCL = Lower control limit
Φ = Level of correlation
δ = Amount of shift in the process average from its target

Dr. D. R. Prajapati is working with the Department of Mechanical Engineering, PEC University of Technology (Formerly Punjab Engineering College), Chandigarh-160012 (India).
Ph. no. +919876166331 and E-mail: drprajapati@pec.ac.in,

III. LITRETURE REVIEW

Several researchers examined the performance of control charts in the presence of autocorrelation. Consequently, there has been considerable research in recent years on designing the control charts suitable for autocorrelated processes. Goldsmith and Whitfield [3] showed that positive autocorrelation can increase the false alarm rates for CUSUM control charts, while negative autocorrelation decreases the rates. Bagshaw and Johnson [4] investigated the effect of autocorrelation on the run length distribution when the process follows either an autoregressive of order one model, AR (1) model, or a moving-average of order one model, MA (1) model, where the run length is the number of observations plotted before an out-of-control signal. They also approximated the average run length (ARL) of CUSUM control charts for the autocorrelated data.

According to Alwan and Roberts [5]; more than 70% of the studied processes subject to change detection in quality control are autocorrelated. Because of more complexity, more effort is involved in designing change detection procedures for autocorrelated processes than for independent and identically distributed (IID) processes. If the presence of autocorrelation is ignored it may seriously reduce the effectiveness of the designed control charts. The run length properties of traditional control charts like Shewhart control chart, cumulative sum (CUSUM) control chart and exponentially weighted moving average (EWMA) control charts are strongly affected by the presence of autocorrelation in the data. Harris and Ross [6] discussed the impact of autocorrelation on CUSUM and EWMA control charts and claimed that the in-control ARL of these charts is sensitive to the presence of autocorrelation. Maragah and Woodall [7] provided results on the effect of autocorrelation on the performance of the Shewhart individuals control chart. Woodall and Faltin [8] then gave a brief summary on the effect of autocorrelation on the performance of control charts and explained methods to deal with autocorrelation. VanBrackle and Reynolds [9] evaluated the performance of EWMA and CUSUM charts for the process mean when the observations are from an AR(1) process with additional random error. They concluded that positive correlation may decrease the in-control ARL, shorten the time required to detect small to moderate shifts, and lengthen the detection time for large shifts. They also provided tables to aid in the design of the control charts. Noorossana et al. [10] presented an artificial neural network model for detecting and classifying three types of non-random disturbances referred to as level shift, additive outlier and innovational outlier, which were common in autocorrelated processes. An autoregressive of order one, AR(1) model, was considered to characterize the quality characteristic of interest in a continuous process where autocorrelated observations were generated over time.

Yang and Yang [11] found that autocorrelation had a significant effect on the performance of the control chart. They considered the problem of monitoring the mean of a quality characteristic X on the first process step and the mean of a quality characteristic Y on the second process step, in which the observations X could be modeled as, an autoregressive of order one, AR(1) model, and observations Y could be modeled as a transfer function of X. Thaga and Yadavalli [12] proposed EWMA chart, that was capable of detecting changes in both process mean and standard deviation for autocorrelated data (referred to as the Maximum Exponentially Weighted Moving Average Chart for Autocorrelated Process, or MEWMAP chart). This chart is based on fitting a time series model to the data, and then calculating the residuals. The observations are represented as a first order autoregressive process plus a random error term. The ARLs for fixed decision intervals and reference values (h, k) are calculated. Vermaat et al. [13] investigated that the serial correlation can seriously affect the performance of the traditional control charts. They derived explicit easy-to-use expressions of the variance of a EWMA statistic, when the process observations are autoregressive of order 1 or 2. These variances can be used to modify the control limits of the corresponding EWMA control charts. Cisar et al. [14] observed that intrusion detection was used to monitor and capture intrusions into computer and network systems which compel to compromise their security. Many intrusions manifested in changes in the intensity of events occurring in computer networks. Because of the ability of EWMA charts to monitor the rate of occurrences of events based on their intensity; this technique is appropriate for implementation in control limits based algorithms.

Lee and Apley [15] investigated that the Residual-based control charts for autocorrelated processes are sensitive to time series modeling errors, which can seriously inflate the false alarm rate. They proposed a design approach for a residual-based EWMA chart that mitigates this problem by modifying the control limits based on the level of model uncertainty. Using a Bayesian analysis, they derived the approximate expected variance of the EWMA statistic, where the expectation is with respect to the posterior distribution of the unknown model parameters. They compared their approach to two other approaches for designing robust residual-based EWMA charts and claimed that their approach generally results in a more appropriate widening of the control limits. Hachicha and Ghorbel [16] concluded that the Control Chart Pattern Recognition

(CCPR) is a critical task in Statistical Process Control (SPC). Abnormal patterns exhibited in control charts can be associated with certain assignable causes, adversely affecting the process stability. Abundant literature treats the detection of different Control Chart Patterns (CCPs). They surveyed and proposed a new conceptual classification scheme, based on content analysis method, to classify past and current developments in CCPR research. More than 120 papers published on CCPR studies within years 1991 to 2010, were classified and analyzed by them.

The next section deals with various schemes of conventional chart.

IV. FORMILATION OF AUTOCORRELATED DATA

In the autocorrelated series of observation, each individual observation is dependent upon the previous observation. A

ISBN: 978-988-19253-1-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

series of positively autocorrelated numbers with a mean of zero and standard deviation of one is generated, using the MATLAB at various levels of correlation (Φ). Assuming N pairs of observations on two variables, x and y. The correlation coefficient between x and y is given by equation (1). Some authors use coefficient of correlation (Φ) instead of "r".

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum (x_i - \bar{x})^2\right]^{1/2} \left[\sum (y_i - \bar{y})^2\right]^{1/2}} \qquad (1)$$

Where, the summations are over the N observations.

The series generated are positively correlated in nature. For each level of correlation (Φ), various schemes of chart are developed, using MATLAB. The detailed procedure to implement the chart is explained in the following section.

V. PPOCEDURE TO IMPLEMENT THE $\overline{X}$ CHART

Step 1 Take the observations from industry at random basis.
Step 2 Observations are generated randomly at a given mean and standard deviation.
Step 3 For simulation, 10,000 observations with a sample sizes (n) of 5 are generated.
Step 4 The observations are generated in such a way that there should be positive correlation with their previous data.
Step 5 Those sets of parameters of the $\overline{X}$ chart, which give the in-control ARLs of approximately 370 are considered for comparison.
Step 6 For the selected combinations, the ARLs are calculated at various shifts in process mean at different width of the control limits (L) and at the each levels of correlation (Φ).
Step 7 Optimal schemes of modified chart are obtained for levels of correlation (Φ) of 0.00, 0.50 and 1.00.

Next section deals with the computation of ARLs of the chart, at different sample sizes and levels of correlation (Φ).

VI. VARIOUS SCHEMES OF THE CONVENTIONAL $\overline{X}$ CHART

In the conventional (Shewhart) $\overline{X}$ chart, two control limits are used to decide the state of the process. The control limits in the conventional (Shewhart) chart are also assumed at 'L' times sample standard deviation on both sides from center line.

A. *Average run length of conventional $\overline{X}$ chart for sample size of two*

If Shewhart $\overline{X}$ chart is applied to an IID data, with no change in the process mean, the average occurrence of false alarm is 371. Which means the chart will give a false signal after every 371 samples, while the process mean is under control. This ARL is also known as the in control ARL as there is no drift in the process. When the process average shifts from its target, the corresponding ARL is called out-of-control ARL. The out-of-control ARL depends upon:

(i) The sample size (n)
(ii) Amount of shift (δ) of the process average from its target.
(iii) Level of correlation (Φ) present in the measured data.

The width of control limits (L) is taken as 3. The average run lengths (ARLs) of various schemes of $\overline{X}$ chart for sample size of two at the levels of correlation (Φ) of 0.00 and 0.25 are shown in Table I.

Table I
ARLs of $\overline{X}$ chart for sample size of two at Φ = 0.00 and 0.25

ARLs of conventional $\overline{X}$ chart for n = 2						
Shift (in mean)	Φ = 0.00			Φ = 0.25		
	L = 2.95	L= 3.0	L= 3.05	L = 2.95	L= 3.0	L= 3.05
0.00	370.4	370.4	370.4	370.4	370.4	370.4
0.25	263.2	277.8	312.5	271.5	285.5	303.0
0.50	158.7	169.5	169.5	161.2	175.4	204.1
0.75	80.0	84.6	86.7	85.5	95.2	104.2
1.00	48.6	55.5	58.2	52.9	56.3	61.5
1.25	24.6	33.2	35.4	26.8	34.7	39.6
1.50	15.2	15.6	18.2	16.3	16.8	20.5
1.75	9.5	13.9	14.4	10.3	14.5	15.7
2.00	6.2	6.3	6.5	6.8	6.8	8.1
2.50	3.3	3.3	3.6	3.3	3.6	4.4
3.00	2.0	2.5	2.6	2.1	2.6	3.1
3.50	1.4	1.5	1.5	1.5	1.5	1.9
4.00	1.2	1.2	1.2	1.2	1.2	1.5

The ARLs of various schemes of $\overline{X}$ chart for sample size of two at the levels of correlation (Φ) of 0.50 and 0.75 are shown in Table II.

ISBN: 978-988-19253-1-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table II
ARLs of $\overline{X}$ chart for sample size of two at Φ = 0.50 and 0.75

Schemes of conventional $\overline{X}$ chart for n = 2					
Shift (in mean)	Φ = 0.50		Φ = 0.75		
	L= 3.0	L= 3.05	L = 2.95	L= 3.0	L= 3.05
0.00	370.4	370.4	370.4	370.4	370.4
0.25	285.5	333.3	287.5	312.5	333.3
0.50	185.2	204.1	200.0	262.2	277.7
0.75	99.3	113.6	106.4	128.2	142.8
1.00	58.8	70.4	75.8	90.4	81.0
1.25	37.4	41.5	44.4	53.8	54.4
1.50	23.4	28.6	28.8	39.5	41.2
1.75	14.6	16.0	17.6	28.1	31.0
2.00	10.4	10.5	12.4	17.0	21.4
2.50	4.9	5.2	6.2	8.2	9.6
3.00	2.9	3.0	3.5	4.7	5.8
3.50	2.8	1.9	2.4	3.4	4.1
4.00	1.4	1.5	1.7	2.7	3.0

The ARLs of various schemes of $\overline{X}$ chart for sample size of two at the levels of correlation (Φ) of 1.00 are shown in Table III.

Table III
ARLs of $\overline{X}$ chart for sample size of two at Φ = 1.00

Schemes of conventional $\overline{X}$ chart for n =2		
Shift (in mean)	Φ = 1.0	
	L=3.0	L= 3.05
0.00	370.4	370.4
0.25	333.3	312.5
0.50	285.7	244.0
0.75	156.3	161.3
1.00	98.2	93.5
1.25	64.5	65.0
1.50	46.6	41.3
1.75	36.3	27.0
2.00	23.4	17.5
2.50	12.3	9.5
3.00	7.2	5.4
3.50	4.4	3.4
4.00	3.4	2.4

Out of above suggested schemes, those schemes which have the in-control ARL of approximately 370 and having the out-of-control ARLs consistently lower than other schemes are selected as the optimal schemes. Table IV shows optimal schemes of the conventional X chart for n = 2, for various levels of correlation (Φ)

Table IV
ARLs of optimal schemes $\overline{X}$ chart for sample size of two

Shift (in mean)	Optimal schemes of the conventional				
	Φ = 0.00	Φ = 0.25	Φ = 0.50	Φ = 0.75	Φ = 1.00
	L=2.95	L=2.95	L=3.0	L=2.9 5	L=3.0
0.00	370.4	370.4	370.4	370.4	370.4
0.25	263.2	271.5	285.5	287.5	333.3
0.50	158.7	161.2	185.2	200.0	285.7
0.75	80.0	85.5	99.3	106.4	156.3
1.00	48.6	52.9	58.8	75.8	98.2
1.25	24.6	27.1	37.4	44.4	64.5
1.50	15.2	16.3	23.4	28.8	46.6
1.75	9.5	10.3	14.6	17.6	36.3
2.00	6.2	6.8	10.4	12.4	23.4
2.50	3.3	3.3	4.9	6.2	12.3
3.00	2.0	2.1	2.9	3.5	7.2
3.50	1.4	1.5	2.8	2.4	4.4
4.00	1.2	1.2	1.4	1.7	3.4

The out-of-control average run lengths (ARLs) of conventional $\overline{X}$ chart for sample size of three have also been computed by keeping the in-control ARLs of approximately 370. Various schemes at the levels of correlation (Φ) of 0.00, 0.25, 0.50, 0.75 and 1.00 are obtained but not included in this paper due to limitations of space. Only optimal schemes of the conventional $\overline{X}$ chart for n = 3, for various levels of correlation (Φ) are shown in Table V.

Table V
ARLs of optimal schemes $\overline{X}$ chart for sample size of three

Shift (in mean)	Optimal schemes of the conventional				
	Φ = 0.00	Φ = 0.25	Φ = 0.50	Φ = 0.75	Φ = 1.00
	L=2.95	L=2.95	L=3.0	L=2.95	L=3.0
0.00	370.4	370.4	370.4	370.4	370.4
0.25	256.5	268.8	277.8	281.1	304.1
0.50	140.8	151.5	181.2	195.4	225.4
0.75	78.3	82.5	91.7	103.6	130.0
1.00	46.4	49.7	56.0	72.2	81.0
1.25	23.0	27	33.7	41.4	49.2
1.50	14.2	15.6	18.1	27.1	36.1
1.75	9.1	10.1	12.3	16.5	23.6
2.00	6.1	6.4	8.6	11.6	16.3
2.50	3.2	3.2	4.3	6.0	8.9
3.00	1.9	2.0	2.5	3.4	4.9
3.50	1.4	1.5	1.8	2.3	3.3
4.00	1.2	1.2	1.4	1.7	2.6

Following facts are summarized from the Tables 1 to 5:

1) The false alarm rate (in-control ARL) of approximately 370 is maintained, for all the optimal schemes of traditional X chart.

2) For a particular sample size; when the level of correlation (Φ) increases, the sensitivity of the conventional $\overline{X}$ chart to detect shift in the process mean decreases. For sample size of two and level of correlation (Φ) of zero, the conventional $\overline{X}$ chart detects 1σ shift in the process mean after about 48 samples whereas at level of correlation (Φ) of one, it detects same shift in the process mean after 98 samples.

3) The performance of optimal schemes of conventional $\overline{X}$ chart improves on increasing the sample size.

4) The in-control and out-of-control ARLs of the optimal schemes of the conventional $\overline{X}$ chart also depends on the width of control limit (L).

VI. CONCLUSIONS

Taguchi (1989) and others researchers have recommended that inspection cannot be relied upon to judge the quality of product. So, instead of 100% inspection, it is better to design a suitable statistical tools e.g. control charts, which can be used to monitor the manufacturing process**.** The false alarm rate (in-control ARL) of approximately 370 is maintained, for all the optimal schemes of conventional X chart. Control charts are used to detect the presence of assignable causes of variation by checking the desired stable state of the process. Reduction of variation is thus achieved via rapid detection and elimination of such special causes. A process is said to be in a state of statistical control, if it operates under common causes of variation and the probability distribution representing the quality characteristic is constant over time. If there are some changes over time in this distribution, the process is said to be out-of-control.

It is clear from all the tables that autocorrelation distort the performance of conventional (Shewhart) chart. It is also observed that the in-control and out-of-control ARLs of the optimal schemes of the conventional $\overline{X}$ chart also depends on the width of control limit (L) and sample size. The performance of conventional X chart also improves on increasing the sample sizes.

ACKNOWLEDGEMENT

I am very grateful to the reviewers, whose comments were really helpful to make this paper suitable for publication.

REFERENCES

[1] W. A. Shewhart, “Economic control of quality of manufacturing product”, Van Nostrand, New York1, 1931.

[2] E. S. Page, “Continuous Inspection Scheme", *Biometrika,* Volume 41, Issue 1-2, pp. 100-115, 1954.

[3] P. H. Goldsmith, and H. Whitfield, “Average run lengths in the cumulative chart quality control schemes”, *Technimetrics*, Volume 3, pp. 11-20, 1961.

[4] M. Bagshaw and R. A Johnson, “The effect of serial correlation on the performance of CUSUM tests II”, *Technimetrics,* Volume 17, pp. 73-80, 1975.

[5] L. C. Alwan and H. V. Roberts, “Time series modeling for statistical process control”, *Journal of Business & Economic Statistics*, Volume 6, Number 1, pp. 87-95, 1988.

[6] T. J. Harris and W. H. Ross, “Statistical process control procedures for correlated observations, *The Canadian Journal of Chemical Engineering,* Volume 69, Issue 1, pp. 48-57, 1991.

[7] H. Maragah and W. H. Woodall, “The effect of autocorrelation on the restropective X chart”, *Communications in Statistics-Simulation and Computation*, Volume 40, Issue 1, pp. 29-42, 1992.

[8] W. H. Woodall, and F.W. Faltin, “Autocorrelated data and SPC”, *ASQC Statistics Division, Newsletter*, Vol. 13, No. 4, pp.18–21, 1993.

[9] L. N. VanBrackle and M. R. Reynolds Jr., “EWMA and CUSUM control charts in the presence of correlation", *Communications in Statistics: Simulation and Computation*, Volume 26, Issue 3, pp. 979-1008, 1997.

[10] R. Noorossana, M. Farrokhi and A. Saghaei, “Using neural networks to detect and classify out-of-control signals in autocorrelated processes”, *Quality and Reliability Engineering International*, Volume 19, Issue 6, pp. 493 – 504, 2003.

[11] S. F. Yang, and C. M. Yang, “Effects of imprecise measurement on the two dependent processes control for the autocorrelated observations”, *The International Journal of Advanced Manufacturing Technology*, Volume 26, Number 5-6, pp. 623-630, 2005.

[12] K. Thaga, and V.S.S. Yadavalli, "Max- EWMA chart for autocorrelated processes (MEWMAP chart)", *South African Journal of Industrial Engineering*, Vol. 18, No. 2, pp.131–152, 2007.

[13] M. B. Vermaat, R. J. M. M. Does and S. Bisgaard, "EWMA control chart limits for first-order and second-order autoregressive processes", *Quality and Reliability Engineering International,* Volume 24, Issue 5, pp. 573 – 584, 2008.

[14] P. Cisar, S. Bosnjak and S.M. Cisar, "EWMA algorithm in network practice", *Int. J. Computers, Communications & Control*, Vol. 5, No. 2, pp.160–170, 2010.

[15] H. C. Lee, and D. W. Apley, "Improved design of robust exponentially weighted moving average control charts for autocorrelated processes", *Quality and Reliability Engineering International*, Volume 27, Issue 3, pp. 337-352, 2011.

[16] W. Hachicha and A. Ghorbel, "A survey of control-chart pattern-recognition literature (1991–2010) based on a new conceptual classification scheme", *Computers & Industrial Engineering,* Volume 63, Issue 1, pp. 204-222, 2012.

ISBN: 978-988-19253-1-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)