Time Series Approaches to Statistical Process Control

Ajit Goswami, Prof. (Dr.) H.Nr. Dutta

Abstract- In traditional Statistical Process Control (SPC) procedure, a standard assumption is that observation from the process at different time points are independent random variable. However, this independent assumption is not always true.. In fact, in the last decade, the time-series approach to Statistical Process Control has been a topic of interest of many quality scientists. In this paper, an attempt has been made to highlight some of the works in this area and a few models will be discussed to analyze the effects of autocorrelation on some standard control charts techniques.

Index Terms—**Autocorrelation, Dependent observation, EWMA control chart, SPC**

I. INTRODUCTION

Most of the industrial processes often have complex behaviors, when successive units are related to previous one. When there are carry over effect from the earlier observations, the standard control charts may exhibit an increased frequency of false alarms. There is an increased likelihood that the data will exhibit autocorrelation in systems where the process time is longer than the time between sample collections [1]. Autocorrelation results from many factors- such as work shift, operator rotations, mechanic or technician changes. Sometimes some processes, inherently produces autocorrelated data. Traditional Shewhart control charts are sensitive to autocorrelated data and even at low levels of correlation, a significant changes may occur in chart properties including short Average Run Length (ARL). Hence in recent times, studies on autocorrelation data is an important area for SPC users and more attention is being paid by many quality scientists to study the behavior of control chart performance in presence of autocorrelation.

II. EFFECT OF AUTO CORRELATION IN PROCESS DATA

When there is significant autocorrelation in the process data, it is not advisable to use traditional control chart technique without modification. Two general approaches have been considered by the scholars of quality control to deal with the auto correlation in recent times, they are:

(a) Traditional control charts are used, but methods used to estimate the process parameters and finally the control limits are adjusted in order to account for the auto correlation. This method is recommended when the level of auto correlation is not extremely high.

(b) A time series model is fitted to the process observations and the residuals from this model are used in traditional control charts

III. REVIEW OF PAST WORKS ON AUTOCORRELATED PROCESS DATA

Dutta & Phukan [2] reviewed the effect of autocorrelation on traditional variable control charts and other modified variable control charts like Cumulative Sum Chart (CUSUM), Exponentially Weighted Moving Average (EWMA) control chart and Multivariate (T^2) control chart covering the period 1978-2008. They, however, did not considered the past works done by the quality scientists in the area of autocorrelated attribute control charts. In our present study, we shall try to include (as far as possible) most of the current research works in these area, both for variable (section A) as well as for attribute control charts (section B). However, considering the fast growing nature of the topic, studies on autocorrelation effects on variable sampling intervals (VSI) control charts and non-parametric control charts could not be discussed in this section and it will be reported in a future study.

A. Past Works on Variable Control Charts

Since 2008, many papers have been published by the scholars in studying the effect of autocorrelation in variable control charts and more are offing.

Sheu & Lu [3] presents a useful discussion of a method that enables the detecting ability of the EWMA control chart to be enhanced and shows that when the observations are drawn from an AR(1) process with random error, the EWMA control chart is far more useful than the Shewhart control chart in detecting small shifts. They found that The Generalized Weighted Moving Average (GWMA) control chart of observations is shown to be superior to the EWMA control chart in detecting small shifts in the process mean and variance. The GWMA control chart of observations

Manuscript received November 22, 2013.

Ajit Goswami, Research Scholar, Dibrugarh University, Dibrugarh, Assam, India (e-mail: agoswami09@gmail.com).
Dr. H. Nr. Dutta, Professor, Department of Statistics, Dibrugarh University, Dibrugarh, Assam, India (e-mail: hnd_stat@rediffmail.com).

requires less time to detect small process mean and/or variance shifts as the level of autocorrelation declines.

Keoagile [4] considers the problem of monitoring a process in which the observations can be represented as a first-order autoregressive model following a heavy tailed distribution. He propose a chart based on computing the control limits using the process mean and the standard error of the least absolute deviation for the case when the process quality characteristics follows a heavy tailed *t*-distribution.

Chang & Wu [5] developed a general and unified approach based on the use of discretization and the finite Markov chain imbedding technique to investigate the run length properties for various control charts when the process observations are autocorrelated. Also numerical results are presented for illustrative purposes.

Suriyakat et.al [6] derived an explicit formula for the characteristic of EWMA control chart for trend stationary exponential AR (1) processes. They compare the results for Average Run Length (ARL) obtained from the explicit formula with values obtained from the integral equation and found that the new results are simple, easy to programming, which make it attractive to be used in practice by performers.

Karaoglan & Bayhan [7] computed ARL performances of control charts for peroxide data from two batches, for which trend stationary first order autoregressive (trend AR(1) for short) model is a representative model.

B. Autocorrelated Attribute Control Charts

To our knowledge little attention has been given to the development of control charts in the case of correlated attribute data. A few work in this area are Deligonul and Mergen [8], Bhat and Lal [9]. They assumed a two-state Markov chain model for auto correlated attribute data. Harvey and Fernandes [10] and Wisnowski and Keats [11] shows that correlated count data can be modeled with a EWMA approach.. Stimson and Mastrangelo [12] studied the monitoring of serially dependent processes with attributes data obtained from multistations of production. Lai et al. [13] examined control procedures based on the conforming unit run length applied to near-zero-defect processes in the presence of serial correlation. Lai et al. [14] also studied the problem of process monitoring when the process is of high quality and measurement values possess a certain serial dependence. Nembhard et.al [15] studied a demerits control charts (U-chart) for autocorrelated data. Their study is related to injection-modeling production lines produced by various models of leak proof plastic containers. Tang and Cheong [16] proposed a control scheme that is effective in detecting changes in fraction nonconforming for high yield processes with correlation within each inspection group. Shepherd et al. [17] proposed two control chart schemes. These control charts are based on a sequence of random variables that are used to classify an item as conforming or nonconforming under a stationary Markov chain model and 100% sequential sampling.

IV. TIME SERIES MODEL

To apply control chart for residual, we can modeled quality characteristics X_t as follows-

$$X_t = \xi + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \dots\dots + \phi_p X_{t-p} + \varepsilon_t \tag{1}$$

Here, X_t is a p^{th} *order autoregressive or AR (p)* Process where, ξ and ϕ (-1 < ϕ <1) are unknown constant and it is normally and independently distributed with mean 0 and standard deviation σ .

If we modeled $X_t = \xi + \phi_1 X_{t-1} + \varepsilon_t$ (2)

then it is called *first order autoregressive AR (1) model;* the observations X_t from such a model have mean $\xi/(1-\phi)$, standard deviation $\sigma/(1-\phi^2)^{1/2}$ and the observations that are *k* periods apart $(X_t - X_{t-k})$ have correlation coefficient ϕ^k . Suppose that $\hat{\phi}$ is an estimate of ϕ , obtained from analysis of sample data from the process, and $\hat{X}_t$ is the fitted value of X_t . Then the residuals

$$e_t = X_t - \hat{X}_t$$

are approximately normally and independently distributed with mean zero and constant variance. Conventional control charts could now be applied to the sequence of residuals. Similarly, the second order autoregressive model AR (2) will be

$$X_t = \xi + \phi_1 X_{t-1} + \phi_2 X_{t-2} + .\varepsilon_t \tag{3}$$

Another possibility is to model the dependency through the random component ε_t . A simple way to do this is

$$X_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1} \tag{4}$$

This is called a *first-order moving average model*. In this model, the correlation between X_t and X_{t-1} is $p_1 = -\theta/(1+\theta^2)$ and is zero at all other lags. Thus, the correlative structure in X_t only extends backwards one time period. Sometimes combinations of autoregressive and moving average terms are useful. A *first order mixed model* is

$$X_t = \xi + \phi X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \tag{5}$$

We also encounter the *first-order integrated moving average* model

$$X_t = X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \tag{6}$$

in some applications. Whereas the previous models are used to describe stationary behavior (that is X_t wanders around a "fixed" mean), the model in equation (6) describes non-stationary behavior (the variable X_t "drifts" as if there is no fixed value of the process mean).

V. Calculation of Autocorrelation

The autocorrelation coefficient for data that are k time period apart r_k is defined as

$$r_k = \frac{\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}, k = 0,1,2,3,.. \tag{7}$$

where n is the total number of observations in the data set.

The standard error at lag k is

$$se_k = \sqrt{1/n} \quad ;k=1 \tag{8}$$

$$= \sqrt{1/n}(1 + 2\sum_{i=1}^{k-1} r_i^2) \quad ; k>1 \tag{9}$$

VI. EWMA Control Chart

Roberts introduced exponentially weighted moving average chart in 1959 [18] . This chart is popular for the control of industrial processes where the individual observations arrive one by one. The EWMA, $\bar{y}_t$ is computed sequentially as a linear interpolation between the present observation z_t and $\bar{y}_{t-1}$, the previous EWMA

$$\bar{y}_t = \lambda z_t + (1-\lambda)\bar{y}_{t-1} \tag{10}$$

Where λ is a constant $0 \le \lambda \le 1$. Hunter [19] has shown that for independent and normally distributed data, the control limits for the EWMA $\bar{y}_t$ are given by

$$\begin{aligned} UCL_t &= \bar{z} + 3\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2t}]} \\ CL &= \bar{z} \\ LCL_t &= \bar{z} - 3\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2t}]} \end{aligned} \tag{11}$$

where the estimate of the process variability, $\hat{\sigma}$, typically is estimated using the same method as for the individual control chart.

VII. Analysis

To analyze the control chart model in presence of autocorrelation, we have studied the finished product of Formalin from a chemical factory in Assam. It may be mentioned here that formalin product of the chemical factory is set as 37 ± 0.5 % weight of formaldehyde gas. If the finished Product is below 36.5%, the customers don't accept it. If the finished product is above 37.5%, it is not affordable to the management so far its cost benefit margin is concerned. To analyze the data, we have collected 268 set of raw data of formalin and deal with using statistical process control tools. First, we calculate the autocorrelation function (ACF) of the formalin (chemical) product data which will indicate the presence of the autocorrelation in the data. Graph of ACF and PACF are shown in figure 1 and 2 respectively.

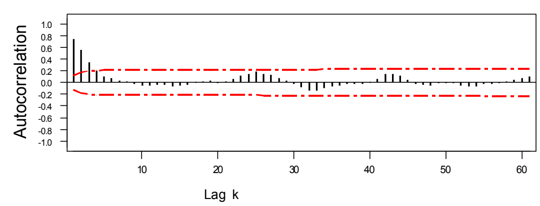

Fig.1. Autocorrelation Function (ACF) of Purity of Chemical Product (Formalin)

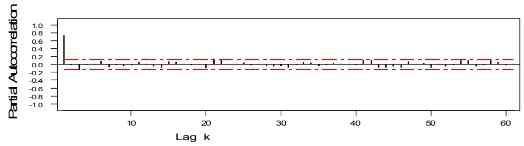

Fig. 2. Partial Autocorrelation Function (PACF) of Purity of Chemical Product (Formalin)

From the visual inspection of the figure 1, we can easily conclude that there is autocorrelation in the original set of data. Also, from the ACF plot fig 1., it is clear that the lag(s) is significantly different from zero and the series is not white noise i.e the data has auto correlation.

A. Removing Autocorrelation from the Observed Data

To achieve an independent, normally distributed data set, Montgomery [1] recommends modeling the correlative structure and control charting the residuals directly. For the formalin data, the predicted purity of formalin (chemical) product at period time't' is (from the fig 1)

$$\hat{X}_t = \xi + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \phi_4 X_{t-4} \qquad (12)$$

Only four points from the previous data were used because of the high autocorrelation coefficient for lags 1-4 (fig-6.1). To determine the parameters of this model multiple linear regression can be performed.

Using Minitab Software, the regression parameters ϕ_1, ϕ_2, ϕ_3 and ϕ_4 are calculated which is given below.

ξ =11.424

ϕ_1 = 0.756

ϕ_2 =0.0971

ϕ_3 =-0.138

ϕ_4 =-0.0116

To check the model, we show in figure 3, a normal plot of the residuals, and in figure 4, a plot of the residuals in time order. Both plots indicate that the model fits the data well. The ACF and the PACF of the residual provide a further check. Ideally, if the model fits well, all autocorrelation would have been removed from the data and the residual behave like white noise. Figure 5 and 6 show the ACF and the PACF for the residual after fitting the AR (4) model to the formalin data. Both the ACF and the PACF are essentially zero for all lags.

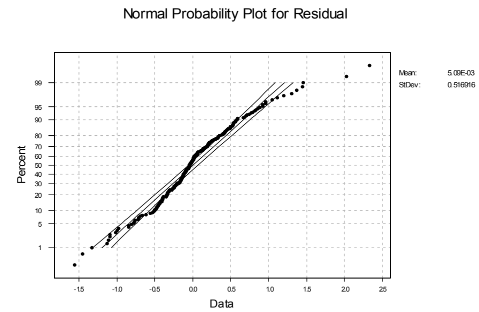

Fig. 3. Normal Plot of Residuals after Fitting an AR (4) Model to the Formalin Data

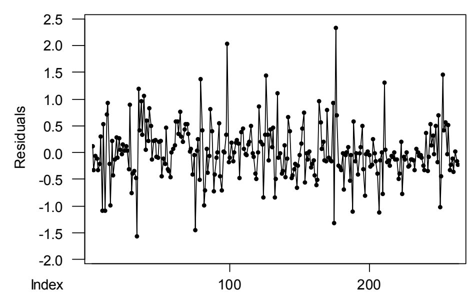

Fig. 4. Time Series Plot of the Residuals

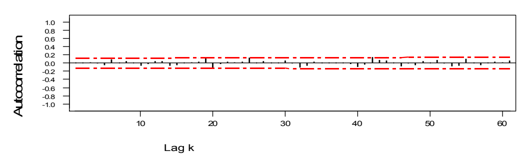

Fig. 5. The ACF of the Residuals after Fitting an AR (4) Model to the Formalin Data

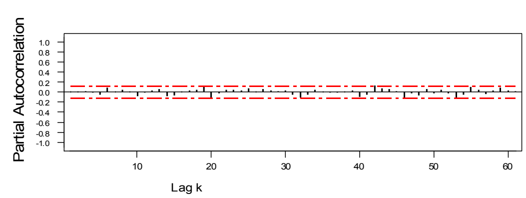

Fig. 6. The PACF of the Residuals after Fitting an AR (4) Model To the Formalin Data

B. EWMA Control Chart for Residual

In using the inflated limits for the individuals control chart, we emphasized the importance of reducing the false alarm rate, and making the chart easy to interpret. However, this approach desensitizes the chart and will likely increase the average run length (ARL) to signal an alarm in case of a real change.

For the current process, we could use an individual's control chart, a cumulative sum (CUSUM) chart or an EWMA chart. The residuals are not on a meaningful scale. Hence the practical interpretation argument for using the individuals control chart no longer applies. We therefore suggest using an EWMA chart.

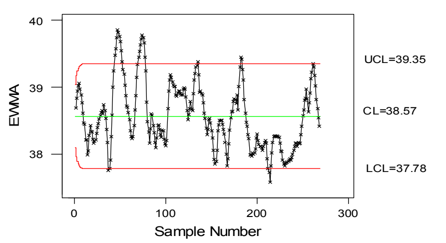

Fig. 7. An EWMA of Observations from the Original Formalin Data

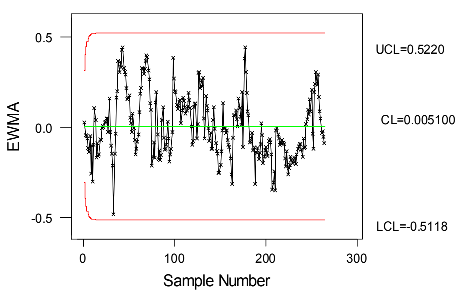

Fig. 8. An EWMA of Observations from the Residual Formalin Data

Using equation (11), the original formalin data is plotted for EWMA chart using Minitab 11 version with the typical default value $\lambda = 0.2$. Fig 7 shows an EWMA for original formalin data set. We see that the process with this chart appears to be out of control. But after removing the effect of autocorrelation when we use the EWMA control chart for the residual formalin data, the process is found in statistical control. (fig.8)

VIII. Conclusion

In modern approach of application of statistical process control, the effect of autocorrelation is increasingly becoming a fact of life and must not be ignored. In our study, we have tried to explain with a chemical data how to detect autocorrelation; illustrated it's consequences for standard control chart EWMA chart. Other control chart can be used. As demonstrated, modern software packages such as MINITAB, make it relatively easy to perform the computations needed when dealing with autocorrelated processes and using AR time series models.

References

[1] Montgomery, D.C., "Introduction to Statistical Quality Control", 6th Edition. John Wiley and Sons-New York.,2009

[2] Dutta, H.Nr. & Phukan, A, "Performance of Some Control Chart in Presence of Autocorrelation: A Review and Survey of Literature," Assam Statistical Review, 2008, 22, 1-2, 88-107.

[3] Sheu, S, S. Lu ., " *The effect of autocorrelated observations on a GWMA control chart performance,"* International Journal of Quality & Reliability Management., 2009, Vol. 26 Iss: 2 pp. 112 – 128

[4] Keoagile, T ., "Control Chart for Autocorrelated Processes with Heavy Tailed Distributions." *Economic Quality Control., 2010,* Volume 23, Issue 2, Pages 197–206, ISSN (Online) 1869-6147.

[5]] Chang, Y.M., & T. Wu, **"**On Average Run Lengths of Control Charts for Autocorrelated Processes," *Methodol Comput Appl Probab,* 2011,13:419–431

[6] Suriyakat, W, Y. Areepong, S. Sukparungsee and G. Mititelu, "*An Analytical Approach to EWMA Control Chart for Trend Stationary Exponential AR (1)Processes,"* Proceedings of the World Congress on Engineering, 2012, Vol. IJuly 4 – 6

[7] Karaoglan, A.D and G.M. Bayhan ., "ARL performance of residual control charts for trend AR (1) process: A case study on peroxide values of stored vegetable oil." *Scientific Research and Essays, 2012,* Vol. 7(13), pp. 1405-1414,

[8] Deligonul and Mergen, "Dependence bias in conventional p-charts and its correction with an appropriate lot quality distribution*," Jrnl. Of Applied Statistics,1987* 14(1),75-81.

[9] Bhat, U.N. and Lal, R., "Attribute Control Charts for Markov Dependent Production Processes*," IIE Transactions*, 22(2), 1990,181-188.

[10] Harvey, A.C. and Fernandes, C., "Time Series Modeling for count or Correlated Observations," *Jrnl. Of Business and economic Statistics*, 1989,7(4), 407-422.

[11] Wisnowski, J. W. and Keats, J. B., "Monitoring the availability of assets with binomial and correlated observations," *Quality Engineering*, 1999,11(3), 387-393.

[12] Stimson, W.A. and Mastrangelo, C.M., " *Monitoring Serially-dependent Processes with Attribute Data"* JQT,1996, 28(3), 279-288.

[13] Lai, C.D., Govindaraju, K. and Xie, M. , " Effects of Correlation on Fraction non conforming Statistical Process Control Procedures*," Jrnl. Of Applied Statistics*, 1998, 25(4),535-543.

[14] Lai, C.D., Govindaraju, K. and Xie, M., "Study of Markov Model for a high quality dependent process*," Jrnl. Of Applied Statistics, 2000,* 27(4),461-473.

[15] Nembhard D.A. , "*A Demerits Control Chart for Autocorrelated data"Quality Engineering, 2000,* 13(2), 179-190.

[16] Tang, L.C. and Cheong, W.-T. , "A control scheme for high-yield correlated Production under group inspection," *Journal of Quality Technology*, 2006, 38(1), 45-56.

[17] Shepherd, D.K., Champ, C.W., Rigdon, S.E. and Fuller, H.T. (2006):*Attribute Charts for monitoring a dependent process,* Quality and Reliability Engineering International

[18] Roberts, S.W., "Control Chart Test based on Geometric Moving Averages." Technometrices, 1959, 1.

[19] Hunter, J.S., "The Exponentially Weighted Moving Average," Journal of Quality Technology,1998, 18