Journal of Material Science and Mechanical Engineering (JMSME)
p-ISSN: 2393-9095; e-ISSN: 2393-9109; Volume 3, Issue 2; January-March, 2016 pp. 95-98

http://www.krishisanskriti.org/Publication.html

Effect of Autocorrelation on Performance of Statistical Process Control Charts

Gaurav Dadheech[1] and Karan Ghule[2]

[1,2]BITS Pilani
E-mail: [1]h2015150@pilani.bits-pilani.ac.in, [2]h2015151@pilani.bits-pilani.ac.in

Abstract—*Traditionally Shewhart control charts have served as a fundamental tool for statistical quality control. These charts, however, are based on the assumption that the samples are independent and thus have no correlation. This assumption is challenged in many of the process industries where the production and the sampling rates are high. In such a scenario, Shewhart control charts fail as they predict excessive false alarms resulting in unnecessary expenditure in investigating the special causes. The effect of autocorrelation on control charts has been analyzed in literature and some the findings have been presented in this term paper.*

Keywords*: Autocorrelation, control chart, ARIMA, EWMA.*

1. INTRODUCTION

Autocorrelation is the measure of dependency between series of observations collected at various time intervals. To quantify the degree of interaction, the auto correlation coefficient defined by Russo *et al*. as:

$$r_k = \frac{\sum_{T-1}^{N-K}(X_t-\bar{X})(X_{t+k}-\bar{X})}{\sum_{t=1}^{N}(X_t-\bar{X})}$$

X_k = Auto covariance of lag k

N= Total number of observations in data set

X_t= Time series

$\bar{X}$ = Mean of observations

Lag refers to the time interval between successive observations. Autocorrelation in the data can be analyzed easily by constructing a scatter plot of the successive observations in software such as MINITAB. An example of a scatter plot for a positive autocorrelation has been presented below, where x_k and $x_{(k+1)}$ denote the observations at k and (k+1) time interval respectively.

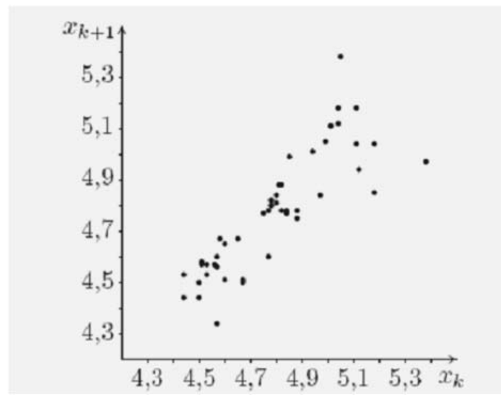

Fig. 1 Scatter plot for positive autocorrelation

When the sampling is carried out frequently then the samples are not independent and significant relation between them. This may also occur as the error in setting up the machine process parameters influences the entire lot. If this machine is running continuously and samples are taken, then the samples will not be independent. In such a scenario, traditional Shewhart control charts fail used as it violates the underlying assumption that the samples are independent. To take account the effect of auto correlation various time series models are employed which are explained below.

2. ARIMA MODEL

ARIMA or the Autoregressive Moving Integrated Average Model is a time series model which takes into account the dependency of the previous observations on the current observation. The ARIMA Model is a combination of the Autoregressive and Moving Average Model and also includes the differencing term to make the model stationary (i.e. to keep the variance as constant). The ARIMA model has been presented below.

$$X_{t=\xi+}\phi_1 X_{t-1+}\phi_2 X_{t-2}+\phi_1 X_{t-1}+\cdots\phi_p X_{t-p+}\varepsilon_t \quad (2)$$

X_t = Observation at time t

ε_t = Independent white noise at time period t ξ = Unknown constant

ϕ = Estimation coefficient

3. METHODOLOGY

The most effective method to remove the influence of auto correlation from control charts is the use of a time series ARIMA model explained above. After choosing ARIMA model with suitable parameters, residuals must be calculated for each data point. The residual is the difference between the original data and the data obtained from the ARIMA time series model. An important property of residuals is that the residuals are uncorrelated and follow normal distribution. This can be easily verified by testing for normality and plotting the Auto correlation function in MINITAB.

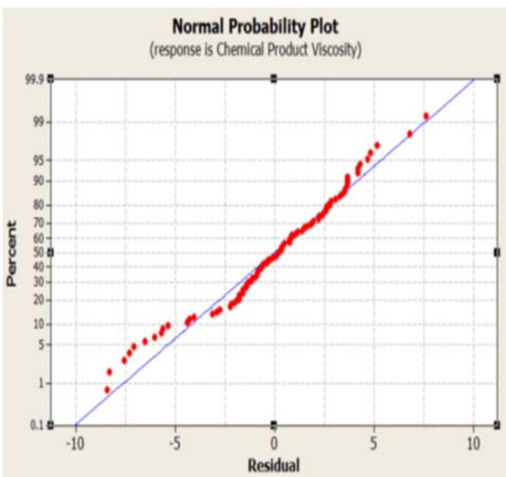

Fig. 2 Normality Plot for Residuals

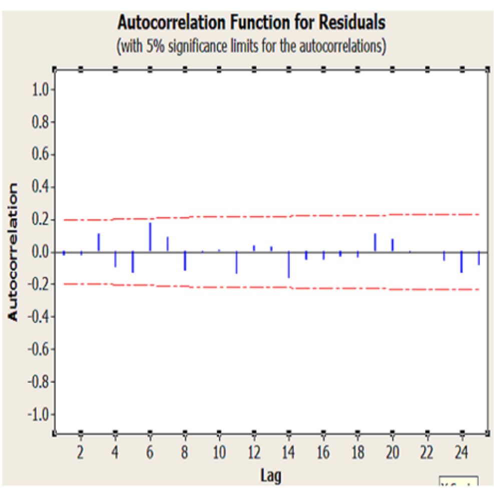

Fig. 3 ACF for Residual

Some of the typical plots for residuals have been shown in Fig 2 &3. Since the residuals are independent and follow normal distribution, then the traditional Shewhart control charts can be applied without any hesitation. This ARIMA time series model can be employed effectively to the desired control charts. The literature review has been presented below to examine the effects of auto correlation on the control charts and the methods to deal with it.

4. LITERATURE REVIEW

Darja Noskievičová [7] proposed an iterative algorithm to set the control limits when the data is auto correlated and a time series ARIMA model is used. The algorithm begins with selection of ARIMA model and its parameters. The next step is to calculate the residuals. He pointed out that the residuals of such model should be normally distributed and must have constant variance. In cases, when the residuals do not have a constant variance, the model parameters must be altered accordingly. Then the ARIMA model must be applied to a suitable control chart. To validate his algorithm, the author has compared his algorithm with traditional methods by constructing control charts for the output of a blast furnace. The control limits proposed by the algorithm agreed with that proposed by traditional methods.

Wheeler [2] argued that the usual control limits are affected only when the autocorrelation becomes excessive (say 0.80 or larger) and the traditional methods for control charts can be safely applied when the autocorrelation coefficient is below 0.8. However, his proposal has not been well accepted by many quality control engineers who feel that the effect of autocorrelation must not be neglected even when the coefficient as low as 0.4.

Adem Tasdemir [12] analyzed the effect of autocorrelation on control charts used to monitor the quality characteristics of fine coal using the Stats graphics software. He used a scatter plot to find the autocorrelation and found out that there was a high correlation of about 0.6 for the ash content. The X/MR charts were not effective as they indicated many false alarms and he therefore fitted the ARIMA time series model to the data by varying the parameters of the ARIMA model and then constructed the control charts for residuals. Akaike Information Criterion (AIC) which was used to evaluate the quality of the statistical model has been used in this study. Hence, the model with the lowest value of the AIC was chosen as the best model. The ARIMA (0, 1, 2) & ARIMA (0, 1, 0) model has been used to recommended to be used for generating time series of ash and moisture content respectively. Reduced number of out-of control points in the residual control charts highlighted the importance of taking autocorrelation into account while constructing the control charts for the coal industry.

Kovářík Martin [11] et. al. used the ARIMA, EWMA and CUSUM control charts to analyze the stability in cash flows, the data considered was the GDP of Britain from 1960 to 1997. These control charts are sensitive to mean shifting while calculating the autocorrelation coefficient. Hence, the objective of the case study was to illustrate the sensitivity of the time series models in detecting the small shifts when the data is auto correlated. The nature of auto correlation in the data they considered was positive first degree (linear) and this was found out by constructing of plot of k^{th} observation and $(k+1)^{Th}$ observation. The value of auto correlation was found out to be 0.86. The results presented showed that time series charts are found to be sensitive in detecting small shifts and we utilize the fact that these control charts can be used in certain situations where the data is auto correlated.

Young [4] et. al analyzed the effect autocorrelation on the X/MR and EWMA process control charts for the continuous production of forest products in a real time environment. The authors stated that when the data is positively correlated, there is a roller coaster type pattern on the graph and when the data is negatively correlated, a saw tooth pattern is exhibited on the graph. To deal with the auto correlated data(r=0.57), the authors used a first order stationary autoregressive AR (1) model. In this method the quality characteristic at time t+1 is expressed as a linear combination of the quality characteristic at time t and the white noise (error). Then the model was employed to develop X/MR and EWMA chart. The authors also pointed out that adjusting the weight factor for EWMA as a function of auto correlation coefficient is necessary, lest the control chart will predict the excess of false alarms

Leoni [14] et.al investigated the effect of autocorrelation on the performance of Hoteling T^2 Chart. The Hoteling T^2 control chart is the most referenced control scheme for detecting mean shifts in multivariate processes. The T^2 chart was constructed for a bivariate process and the effect of autocorrelation and cross correlation was analyzed by assuming rational subgroups of size n=4. The performance of the chart was measured by the ARL. It was pointed out that the ARL is not affected by the autocorrelation if only one of the two independent variables is auto correlated and the chart is robust to detect the occurrence of special causes. In the absence of autocorrelation or in the presence of two variables with the same level of autocorrelation, it was observed that a higher dependence between the two variables improves the performance of the T^2 chart when indicating shift of only one variable. The authors reported that the ARL increases when the assignable causes shifts both the variables... Also the T^2 chart has a better performance if two data quantities of highly correlated.

Khediri et. al [8] analyzed control charts for nonlinear auto correlated process. The authors have used the method of Support Vector Regression (SVR) to analyze the control charts for nonlinear processes. The disadvantage with time series model is the requirement of process structure beforehand. Thus, application of the time series model requires a lot of experience and knowledge about the process, which is quite often not available. Thus, the authors have implemented residual SVR control chart which allow analyzing the process control for complex systems without the need of a predefining process structure. The control chart limits were defined such that, for in-control process, the probability of the Run Length being lower to 100 was approximately equal to 5%. Results showed that the used control charts can effectively monitor the process behavior and is robust mechanism to prevent false alarms.

Aminu Ahmad Magaji [15] et al used the data used by Montgomery and Johnson to study the effect of control chart performance when data is auto correlated. The work concluded that ARIMA (2.0.0) fits most suitably among all identified models. This is due to the ability of the model to detect the autocorrelation in the observed data. The result was concluded after comparing the value of Akaike Information Criterion (AIC) selected from the model to obtain a lowest value. The performance of methods was compared before and after application of data. The analysis revealed that if the data shows the autocorrelation then EWMA control charts can be used to detect false alarms.

In his work R. Noorossana [5] found that the performance of multivariate cumulative sum (MCUSUM) control chart decreases with autocorrelation. The paper also presented a solution based on time series methods which improves the Average run length (ARL). Healy's multivariate control procedure was the base for proposed method. The study suggests that if residuals instead of original data are used, then the ARL of MCUSUM charts improves considerably. The method also reduces the unnecessary large values of ARL.

The problem of detecting special cause due to which changes in the variance and mean are produced was dealt by Chao-Wen Lu [3] et al. Their work considered a process model with AR (l) along with a random error. Four parameters were considered that is overall mean, two variance parameters and one autoregressive parameter. Change in the process can affect any one or set of these parameters such as, in batch production special cause increase the within batch variability or in-between batch variability. The paper mainly compares the various schemes of monitoring of variance. For the case of low to moderate autocorrelation Shewhart control chart and a EWMA chart of observation are recommended.

Timothy M Young [4] et al in his work compared two adjusted X- control charts with unadjusted X-control charts. The analysis was done for forest manufacturing industry, sample size of 10 was taken at 10 minutes interval. A total of 100 such samples were taken of fiber moisture from continuous medium density fiber board manufacturing process. The false signals of special cause were increased as the autocorrelation was increased. The conventional control chart reported 10 false alarms for the data set, but when the model was applied no such alarm was observed. The authors also indicated that the autocorrelation increases if the process is continuous. The study also found that the auto correlated control chart of Wheeler detected more false alarms then the adjusted control chart of Gilbert. The detection of the special cause variation also depends upon the degree of the positive autocorrelation.

Adjusting the weight factor as a function of the autocorrelation is also necessary for the detection of the special-cause variation.

5. RESULTS AND DISCUSSION

The reviewed literature suggests that independence of the observed data is the most critical assumption while plotting the control charts and must be taken care of. The dependency of data is compensated by models such as ARIMA model. The model is applied after considering the Akaike information criteria (AIC), and if the value of AIC is lowest among all the models then that particular model is applied.

It was pointed out that when two variables have the same level of autocorrelation, the bivariate T^2 chart for individual observations has a superior performance than the bivariate T^2 chart for rational subgroups of size more than one. In case of MCUSUM charts when residuals are used on place of raw data the performance of the chart increases. It was also concluded that X/MR chart is better for processes that exhibit stationary time series model, while EWMA is better when the data has a non-stationary time series model. The type of data influences the choice of model and the control chart to be used and it is the skill of the quality control engineer to choose the appropriate model while dealing with auto correlated data.

6. CONCLUSION

The basic assumption of the Shewhart control chart is that the observation data is independent in nature, but in various cases the assumption is not so true. The observations are related with each other such data is called as auto correlated. When standard control charts are applied with auto correlation, there is an increase in the number of false alarms. To deal with the problem of autocorrelation, one approach is to nullify the effect of autocorrelation by a time series model and to use residuals for plotting the control chart. As the residuals are statistically uncorrelated and follow normal distribution, the theory of traditional control charts can be used effectively. Some case studies involving the use of X/MR, Exponential weighted moving average (EWMA) and ARIMA time series models has been explained in this literature review.

REFERENCES

[1] W. A. Shewhart, "*Economic control of quality of manufacturing product*", Van No strand, New York1, 1931.

[2] Wheeler, *Advances in Statistical Quality Control*, 1995.

[3] Chao-Wen, Lu Reynolds, Marion R, *Cu sum charts for monitoring an auto correlated process*, Journal of Quality Technology*;* Jul 2001; 33, 3;

[4] Timothi M. Young and Paul M., *Winistorfer Forest products journal* Vol 51 No- 11/12 November/ December 2001.

[5] R. Noorossana1 and S. J. M. Vaghefi QUALITY AND RELIABILITY ENGINERING INTERNATIONAL *Qual. R eliab. E ngng. Int.* 2006; **22**:191–197 Published Online 27 June 2005.

[6] Soren Bisgaard and Murat Kulahci Quality Engineering, 17:481–489, 2005 copyright# Taylor & Francis Inc. ISSN: 0898-2112 print =1532-4222 online DOI: 10.1081/QEN-200068575.

[7] Darja Noskievičová, *Control Chart Limits Setting when Data are Auto correlated*, Lund University Publication (2007)

Sermin Elevili, *Control Charts for auto correlated colemite data*, Journal of Scientific & Industrial Research, vol. 68, January 2009.

[8] Issam Ben Khediri, Claus Weihs, Mohamed Limam, *Support Vector Regression control charts for multivariate nonlinear autocorrelated processes,* Chemometrics and Intelligent Laboratory Systems 103 (2010) 76–81

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

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

[11] Kovářík Martin, Klímek Petr, *The Usage of Time Series Control Charts for Financial Process Analysis*, Journal of Competitiveness, Vol. 4, Issue 3, pp. 29-45, September 2012

[12] Adem Tasdemir, *Effect of Autocorrelation on the Process Control Charts in Monitoring of a Coal Washing Plant*, Physicochem. Probl. Miner. Process. 48(2), 2012, 495−512.

[13] D.R. Prajapati Proceedings of the World Congress on Engineering and Computer Science 2013 Vol II WCECS 2013, 23-25 October, 2013, San Francisco, USA

[14] Roberto Campos Leoni, Antonio Fernando Branco Costa, Marcela Aparecida Guerreiro Machado, *The effect of the autocorrelation on the performance of the T^2 Chart*, European Journal of Operational Research 247 (2015) 155–165

[15] Aminu Ahmad Magaji, Abubakar Yahaya, Osebekwin Ebenezer Asiribo, International Journal of Mathematics and Statistics Invention (IJMSI) E-ISSN: 2321 – 4767 P-ISSN: 2321 – 4759.