Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 69 (2014) 1449 – 1458

Procedia Engineering

www.elsevier.com/locate/procedia

24th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2013

Guidelines for Applying Statistical Quality Control Method to Monitor Autocorrelated Prcoesses

Karin Kandananond*

Rajabhat University Valaya-Alongkorn, Prathumthani, Thailand 13180

Abstract

The implementation of statistical control charts under autocorrelated situations is a critical issue since it has a significant impact on the monitoring capability of manufacturing processes. The objective of this study is to assess the performance of control charts under different scenarios and to optimize the design of control charts to best deal with autocorrelated processes. To achieve the proposed objective, two autoregressive integrated moving average models, ARIMA (1, 0, 1) and ARIMA (0, 1, 1), are utilized to characterize stationary and non-stationary processes. These process models were simulated to achieve the response, average run length (ARL), which is the performance measure of this study. The factorial design of experiment was conducted to quantify the effect of critical factors, i.e., ARIMA coefficients, types of charts (exponentially weighted moving average: EWMA and moving range: MR) and shift sizes on the ARL. The experimental results show that EWMA chart is the most appropriate control chart to monitor autocorrelated observations. Additionally, both AR and MA parameters along with shift sizes have a significant effect on the performance of control charts. Therefore, this study has pointed out a suitable tool for use under the different scenarios of autocorrelation. The validation of the above experimental results was conducted on another ARIMA model, ARIMA (1, 0, 0). If the performance of control charts under autocorrelated disturbances is correctly characterized, practitioners will have guidelines for achieving the highest possible performance potential when deploying SPC.

Keywords: Autocorrelation; Autoregressive integrated moving average (ARIMA); Average run length (ARL); Exponentially weighted moving average (EWMA) chart; Moving range (MR) chart

* Corresponding author. Tel.: +66-02-529-3829 fax: +66-02-909-3036.
E-mail address: kandananond@hotmail.com

doi:10.1016/j.proeng.2014.03.141

1. Introduction

Statistical process control (SPC) is a methodology used for monitoring and reducing the variation in manufacturing processes and the main tools of SPC are control charts. Normally, SPC works under the assumption that observed data is independent. However, because of the advanced measurement technology, shortened sampling interval and the nature of processes, especially in continuous processes, e.g., chemical processes, the independence of each observation has been violated in many scenarios. The lack of independence among each sample usually comes in the form of serial autocorrelation. The behaviour of process outputs significantly deteriorates the performance of control charts. Therefore, the consequence is that control charts signal fault alarms more often or does not signal when there is a shift. The selection of inappropriate type of control chart in this scenario might lead to the excessive number of false alarms or the losing ability to detect a special cause. As a result, the characterization of widely used control charts was performed in order to understand the characteristics of each control chart under the autocorrelated situations.

2. Literature review

According to the literature, there are many authors suggesting different approaches to solve the autocorrelation issues of SPC. These options include non-standard SPC charts and some sophisticated techniques which are difficult for practitioners to implement in real-life situations. As a result, the object of this study focuses on selecting the available quality tools that most practitioners are familiar with and they are simple for them to use. However, the characterization of these tools under autocorrelation situations should be fully understood. The first step leading to the performance characterization of standard charts is the capability to simulate different types of autocorrelation. Under normal and uncorrelated conditions, the process model has a fixed mean (μ), and the fluctuation around the mean is the result of white noise (a_t). However, when observations are correlated, the correlation structure and drift in the mean are characterized by disturbances. If process observations vary around a fixed mean and have a constant variance, this type of variability is called stationary behavior. Otherwise, the behaviour is non-stationary. MacGregor [1] indicates that there are two types of disturbances, deterministic and stochastic disturbances. Stochastic disturbances are random and might be stationary or non-stationary so it is the main source of autocorrelation in the data. Basically, deterministic disturbances are in the form of step shifts or ramp in the process mean. On the other hand, a stochastic difference equation is utilized to forecast one-step ahead disturbances [2]. These are represented by autoregressive integrated moving average models, ARIMA, as shown in (1).

$$\Delta_d Y_t = \mu + \phi_1 \Delta_d Y_{t-1} + \phi_2 \Delta_d Y_{t-2} + \ldots + \phi_p \Delta_d Y_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \tag{1}$$

The ARIMA (p, d, q) model indicates p as the order of the autoregressive part, d as the amount of difference and q as the order of moving average part. As recommended by [2], ARIMA (1, 0, 0) and ARIMA (1, 0, 1), are likely to be the most suitable models to represent stationary processes while ARIMA (0, 1, 1) is the appropriated choice for non-stationary processes. Therefore, several authors ([3], [4], [5], [6], [7]) have raised an interesting remark that traditional chart (Shewhart), exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) chart, might have a better performance than non-standard charts when data is correlated. As a result, there are a number of works carried out to implement standard charts with correlated processes simulated by ARIMA models. ARIMA (1, 0, 1) is one of the models which is used to simulate autocorrelated data in order to assess the performance of traditional charts when the tested data is modeled by ARIMA (1, 0, 1) [8]. According to their work, average run length (ARL) was used to measure the robustness of designated chart. Moreover, the performance of each standard chart was also assessed by benchmarking with each other. One of these works is the performance comparison of $\overline{X}$ and EWMA chart [9]. The objective is to assess the capability of each chart to detect special causes when the data is autocorrelated. Another approach to improve the performance of standard charts is to filter correlated data with ARIMA models and the residual from filtering process is monitored by selected control charts. This technique is used to monitor the residual based data with EWMA chart [10].

In some cases, the actual data is utilized instead of the simulated data and these sets of data are acquired from different sources. For example, two different charts (CUSUM and EWMA) were used to monitor the autocorrelated data collected from a production process [11]. Another type data used to benchmark control charts is the biochemical quality data [12]. The application of SPC and autocorrelated data is not limited to only industrial data but information technology data as well. For example, EWMA chart is applied to detect analogous changes in the event intensity for intrusion detection while the data is correlated and simulated by deploying ARIMA (1, 0, 0) model [13].

In conclusion, SPC charts are widely utilized in different areas which are not limited to only the industrial area but also other fields. However, the critical problem is that the autocorrelated data usually downgrades the performance of SPC significantly. As a result, a number of studies are introduced to solve this problem but the downside of these approaches is their complication or not friendly use for people on the shop floor. For this reason, the utilization of these techniques might not be suitable to practitioners who are familiar with traditional standard charts. The objective of this study is the use of empirical study for characterizing the performance of standard charts so practitioners will have the guidelines for deploying available charts to monitor the autocorrelated data at its best.

3. Research Procedures

The basis of analysis in this paper is a mathematical model used to study the effects of process autocorrelation on the performance of SPC charts. Process disturbances are controlled by adjusting the level of autocorrelation in the form of ARIMA parameters. Moreover, the situation could be more complicated when there is a special cause. As a shift occurs in the process, moving range (MR) and exponentially weighted moving average (EWMA) charts are utilized to monitor the individual measurement of a process mean to detect a shift. The autocorrelated process in this study is a continuous process with only one quality characteristic, represented by Y. The evaluation of control chart performance is measured by considering the average run length (ARL) which is the average number of points plotted before a point indicated an out-of-control state. The schematic presentation of process model is shown in Fig. 1.

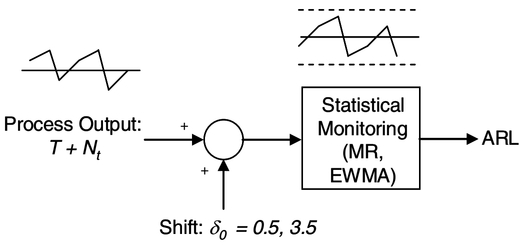

Fig. 1. process model.

The observation of a process is considered from period 1 to 550 (t = 1, 2, 3,…, 550) and the process output (Y_{t+1}) is equal to

$$Y_{t+1} = T + N_{t+1} + \delta(t) \tag{2}$$

The source of autocorrelation is process disturbances, characterized by the autoregressive moving average model, ARIMA (1, 0, 1) and ARIMA (0, 1, 1), as shown in (3) and (4):

$$N_{t+1} = \phi N_t + a_{t+1} - \theta a_t \tag{3}$$

$$N_{t+1} = N_t + a_{t+1} - \theta a_t \tag{4}$$

where N_{t+1}, N_t are the disturbances at time t+1 and t, a_{t+1}, a_t are the random errors at time t+1 and t, ϕ is the autoregressive (AR) parameter and θ is the moving average (MA) parameter. The values of ϕ and θ are between -1 and 1. Afterwards, MR and EWMA charts are utilized to monitor the autocorrelated observations.

To simulate a special cause, a shift of size δ_0 which is in the form of a step function is applied into a process at time t =50 as:

$$\delta(t) = \begin{cases} 0; t < t_{50} \\ \delta_0; t \geq t_{50} \end{cases} \quad (5)$$

where δ_0 is the magnitude of a shift and t_0 is the time that a shift occurs.

4. Experimental Design

The empirical analysis is conducted to examine the effect of input factors on the responses. The screening design of experiment is the 2^k factorial. For every experiment, the fixed categorical variable is the type of control charts (MR and EWMA charts) used to monitor processes. However, other variables are numeric and they are set to low and high levels because of the factorial design condition. Therefore, small and large shift magnitudes are set to $0.5\sigma_a$ and $3.5\sigma_a$ respectively while the parameters of ARIMA equations (AR parameters: ϕ and MA parameter: θ) for each model are between -1 and 1 as shown in Table 1.

Table 1. Input factors and levels for ARIMA (1, 0, 1) and ARIMA (0, 1, 1).

Factor	ARIMA(1, 0, 1)		ARIMA (0, 1, 1)	
	Low	High	Low	High
A (AR parameter; ϕ)	-1	1	-	-
B (MA parameter; θ)	-1	1	-1	1
C (Types of charts)	MR	EWMA	MR	EWMA
D (Shift size)	$0.5\sigma_a$	$3.5\sigma_a$	$0.5\sigma_a$	$3.5\sigma_a$

5. Control Chart Characterization

To characterize the performance of control charts, the experiments are divided into two cases based on two process models: ARIMA (1, 0, 1) and ARIMA (0, 1, 1). Regarding the simulation, each run is composed of 10,000 iterations which are accomplished by using Palisade's @Risk® version 5.5. The random errors (at) from each period are simulated by following normal distribution with zero mean and a constant variance as: $a_t \sim N(0,1)$. A factor screening experiment is designed using a statistical package, Design Expert® version 8.0, to analyze the effect of autocorrelation and other factors on the response.

5.1. Stationary Processes

For stationary processes, ARIMA (1, 0, 1) is utilized to represent the processes. The experimental results for ARIMA (1, 0, 1) are shown in Table 2. For the analysis, the inversed square root transformation was applied to the response in order to satisfy all the residual conditions regarding model validation. The analysis of variance (ANOVA) and the half- normal plot were utilized to reveal the significant factors and their interactions. According to the half normal plot in Fig. 2, the types of charts (C) contributes the highest effect on the ARL, followed by AR parameter (A) and shift sizes (D). The MA parameter (B) is included in the model because of the hierarchical design. Moreover, due to the analysis of variance (ANOVA) in Table 3, the interaction effects exist and are based mostly on the above factors, with the highest-order terms being ABD.

Table 2. Design matrix and results for ARIMA (1, 0, 1) case.

Run	ϕ	θ	Chart	Shift	ARL
1	-1	-1	MR	0.5	42.6971
2	1	-1	MR	0.5	2.2831
3	-1	1	MR	0.5	46.611
4	1	1	MR	0.5	5.4945
5	-1	-1	EWMA	0.5	1.4106
6	1	-1	EWMA	0.5	1.0561
7	-1	1	EWMA	0.5	1.0281
8	1	1	EWMA	0.5	1.3537
9	-1	-1	MR	3.5	2.0983
10	1	-1	MR	3.5	1.38
11	-1	1	MR	3.5	45.1229
12	1	1	MR	3.5	1.32
13	-1	-1	EWMA	3.5	1.0495
14	1	-1	EWMA	3.5	1.1544
15	-1	1	EWMA	3.5	1.097
16	1	1	EWMA	3.5	1.0252

According to the interaction plot in Fig. 3, EWMA chart should be deployed since it can detect a shift rapidly for the whole range of θ: $-1 < \theta < 1$. It is interesting to note that MR chart might be used instead of EWMA only when θ is negatively low since its ARL at $\theta = -1$ is lower than the one at $\theta = +1$.

Table 3. ANOVA for ARIMA (1, 0, 1) case.

Source	SS	df	MS	F	p-value
A-θ	0.172882	1	0.172882	19.73447	0.0113
B-ϕ	0.030441	1	0.030441	3.474872	0.1358
C-Chart	0.792649	1	0.792649	90.48066	0.0007
D-Shift	0.115838	1	0.115838	13.22284	0.0220
AB	0.001451	1	0.001451	0.165678	0.7048
AC	0.176303	1	0.176303	20.12497	0.0109
AD	0.000367	1	0.000367	0.041885	0.8478
BC	0.04309	1	0.04309	4.918747	0.0908
BD	0.004751	1	0.004751	0.542334	0.5023
CD	0.060569	1	0.060569	6.913986	0.0582
ABD	0.079155	1	0.079155	9.035473	0.0397
Residual	0.035042	4	0.00876		
Total	1.512539	15			

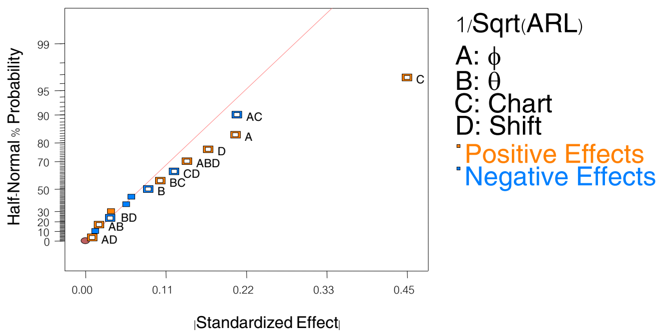

Fig. 2. Half-normal plot for ARIMA (1, 0, 1) case.

On the other hand, due to the cube plot in Fig. 4, when ϕ is in consideration, the ARLs at ϕ = -1 is higher than the ones at ϕ = + 1. However, there are exceptions under these conditions (chart = EWMA, θ = +1, shift = 0.5 and chart = EWMA, θ = -1, shift = 3.5).

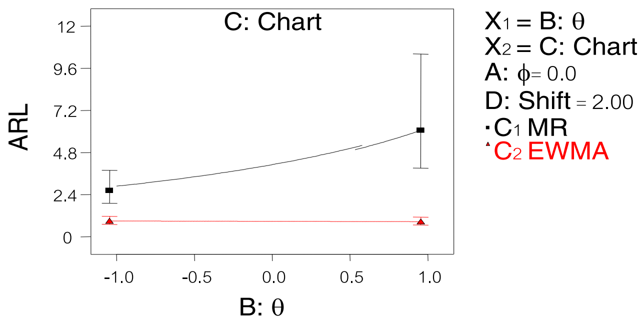

Fig. 3. Interaction plot of BC.

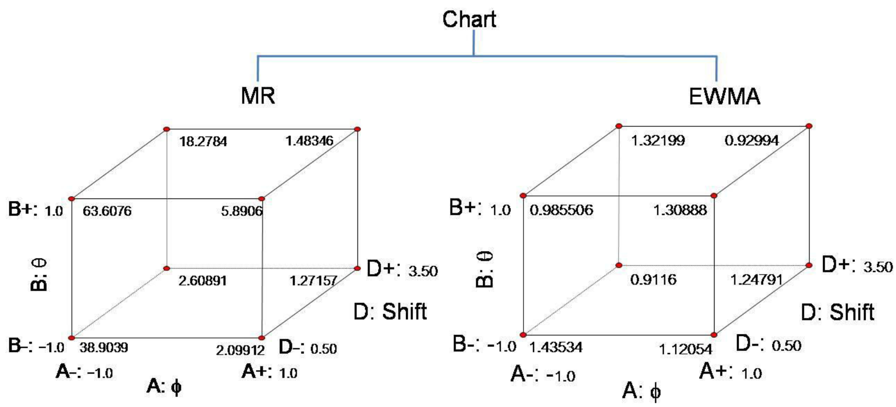

Fig. 4. Cube plot of ABD interaction.

In conclusion, when compared with MR, EWMA charts should be the most appropriate control chart to monitor correlated and stationary process. These results are still valid whether shift sizes are small or large. However, besides the EWMA chart, the MR chart might be utilized under some situations.

5.2. Non-Stationary Processes

The ARIMA (0, 1, 1) model is deployed to represent non-stationary processes and the design matrix for non-stationary case is shown in Table 4.

Table 4. Design matrix and results for ARIMA (0, 1, 1) case.

Run	*θ*	Shift	Chart	ARL
1	-1	0.5	MR	2.3325
2	1	0.5	MR	5.4765
3	-1	3.5	MR	1.3765
4	1	3.5	MR	1.3185
5	-1	0.5	EWMA	1.3589
6	1	0.5	EWMA	2.8043
7	-1	3.5	EWMA	1.1572
8	1	3.5	EWMA	1.0252

For non-stationary case, before the regression equation is constructed, the transformation is required to ensure that residuals satisfy the i.i.d. conditions. After applying the inverse transformation to ARL, the half-normal plot (Fig. 5) and analysis of variance (ANOVA) in Table 5 show that the type of chart (C), shift size (B), AR model coefficients (A) and their interactions (AB and BC) contribute the significant effects on ARL. The experimental results also point out that EWMA chart is still robust to both outliers and the correlation structure of observations as implied in its ARLs. According to the interaction plots in Fig. 6 and 7, both EWMA and MR charts are not sensitive to shift size when θ is highly negative (the ARLs are significantly low at θ = -1). Moreover, the ARL of EWMA is still lower than that of MR chart.

Table 5. ANOVA for ARIMA (0, 1, 1).

Source	SS	df	MS	F	p-value
Model	0.509985	5	0.101997	106.5878	0.0093
A-θ	0.012174	1	0.012174	12.72167	0.0704
B-Shift	0.131001	1	0.131001	136.8978	0.0072
C-Chart	0.256607	1	0.256607	268.1564	0.0037
AB	0.044778	1	0.044778	46.79333	0.0207
BC	0.065425	1	0.065425	68.37001	0.0143
Residual	0.001914	2	0.000957		
Total	0.511899	7			

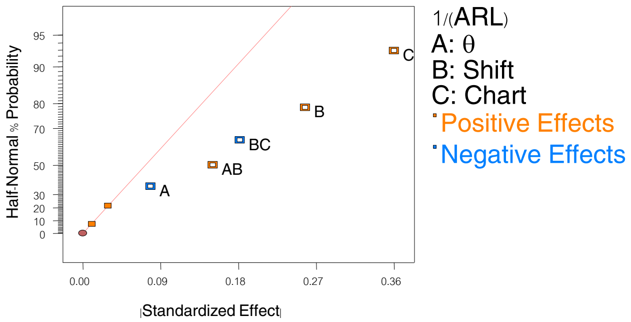

Fig. 5. Half-normal plot for ARIMA (0, 1, 1) case.

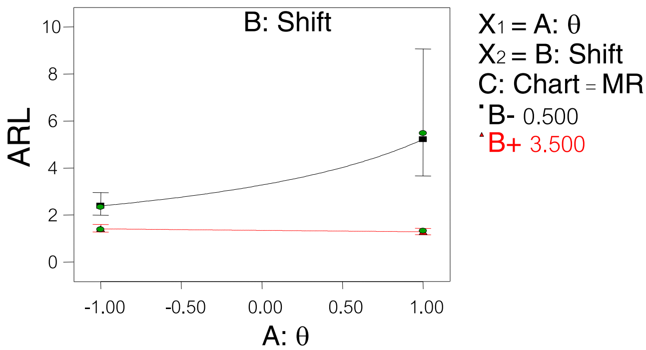

Fig. 6. Interaction plot AB (MR chart).

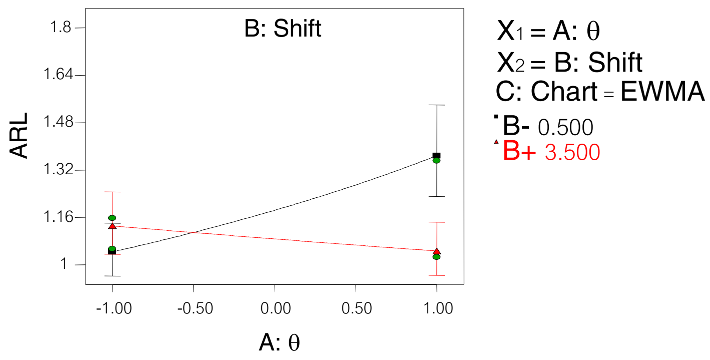

Fig. 7. Interaction plot AB (EWMA chart).

The interaction plots (Fig. 8 and 9) point out that both charts have the same performance when the shift size is large (3.5). However, the capability of EWMA is superior to MR charts. Similar to the results from a stationary case, EWMA charts should be selected to monitor non-stationary processes since its ARLs are lower than those of MR chart in every scenarios.

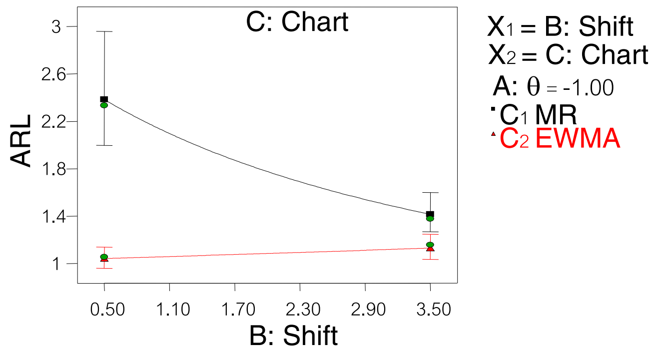

Fig. 8. Interaction plot BC (θ = -1).

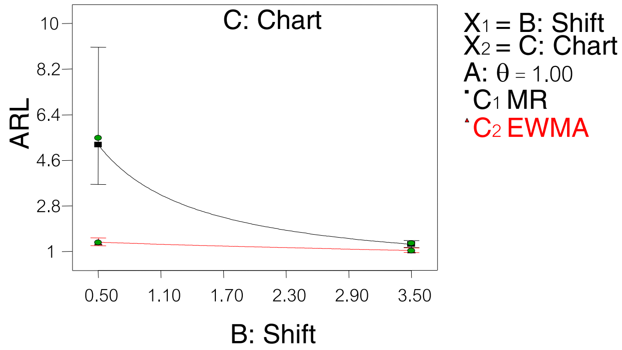

Fig. 9. Interaction plot BC (θ = +1).

6. EWMA Robustness

In reality, there is no single model which can accurately explain the observations. As a result, an extra study was conducted to ensure that EWMA charts are robust to different types of autocorrelated data. As a result, besides the previously used models, another ARIMA model, ARIMA (1, 0, 0), was used to assess the robustness of EWMA charts. The selected model in this study is ARIMA (1, 0, 0) which can be shown as: $N_{t+1} = \phi N_t + a_{t+1} - \theta a_t$. According to table 6, the AR parameter gradually increased by 0.2 from -1 to 1 while the shift size was varied from 0 to 4 (0, 1, 2, 3 and 4) and MR and EWMA charts were utilized to monitor the data

Table 6. ARL results for MR charts.

ϕ	Shift	ARL(MR)	ARL(EWMA)	ϕ	Shift	ARL(MR)	ARL(EWMA)
-1.0	0	84.44	104.36	1.0	0	62.55	5.13
	1	82.24	16.67		1	29.13	2.21
	2	81.97	7.07		2	8.00	1.85
	3	81.71	4.58		3	4.76	1.44
	4	81.10	3.43		4	4.03	1.03
-0.8	0	92.51	474.29	0.8	0	81.40	10.16
	1	94.72	24.54		1	74.61	4.03
	2	91.06	8.98		2	56.72	2.74
	3	85.47	5.58		3	27.89	2.28
	4	79.18	4.03		4	8.02	2.01
-0.6	0	65.33	456.34	0.6	0	48.87	25.03
	1	64.32	21.12		1	45.41	5.80
	2	63.70	8.09		2	33.78	3.41
	3	56.21	5.07		3	16.82	2.63
	4	47.46	3.71		4	4.97	2.22
-0.4	0	54.5	411	0.4	0	43.45	54.18
	1	52.91	17.53		1	40.52	7.47
	2	48.73	7.09		2	32.13	3.89
	3	41.9	4.51		3	19.24	2.85
	4	31.93	3.35		4	7.55	2.34
-0.2	0	46.59	334.89	0.2	0	42.44	115.8
	1	45.95	14.4		1	40.68	9.44
	2	41.11	6.12		2	33.37	4.49
	3	33.51	3.97		3	22.72	3.12
	4	21.59	2.99		4	10.35	2.48
0	0	43.87	219.69				
	1	41.97	11.67				
	2	36.97	5.24				
	3	26.41	3.50				
	4	12.07	2.69				

According to Table 6, when assignable causes exist in the process, EWMA charts are able to detect shifts much faster than MR charts (the ARL_1 of EWMA is lower than those of MR charts). Moreover, the empirical study also reveals that EWMA charts are more robust to the autocorrelation structure of data (the ARL_0 of MR is lower than those of EWMA charts) than MR charts when there is no shift in the process. However, this above conclusion might not be holistic since the ARL_0 of MR charts is much higher than those of EWMA charts when ϕ equals 0.6, 0.8 and 1.0. The results from the above study signify that EWMA charts still outperform MR charts when the observations follow ARIMA (1, 0, 0) model.

7. Conclusion

This study signifies that EWMA charts outperform the traditional Shewhart charts under autocorrelation scenarios. Therefore, the utilization of EWMA charts will lead to the better performance of a control chart to detect a shift resulted from a special cause in the autocorrelated processes. The different categories of stationarity need a different chart design and it will facilitate the application of practitioners when the process is autocorrelated. The performance analysis of a statistical in this phase is also based on stationary and non-stationary processes based on two different models, ARIMA (1, 0, 1) and ARIMA (0, 1, 1). According to the analysis, the effects of AR parameter (ϕ), MA parameters (θ), the appropriate types of control charts and shift sizes on the ARL are determined. In summary, the resultant analysis is concluded as follows:

1. When the observations are stationary and follow ARIMA (1, 0, 1) pattern, both ϕ and θ have a siginificant effect on the ARL. The empirical analysis reveals that EWMA is the most suitable control chart to monitor stationary processes because of its robustness to shift size and autocorrelation structure. However, MR chart can also be utilized in a specific scenario that ϕ is highly positive.
2. When ARIMA (0, 1, 1) is utilized to characterize the non-stationary processes and θ was highly negative, both EWMA and MR charts are sensitive to small shift size.
3. For both stationary and non-stationary cases, the performance of the SPC to minimize ARL will be significantly improved if the EWMA chart is utilized to monitor the observations.
4. The robustness of EWMA charts is assessed by the performance comparison of MR and EWMA charts under ARIMA (1, 0, 0) scenario.

According to the results, the selection of appropriate control charts will assist practitioners to monitor the autocorrelated processes effectively. It is interesting to note that this empirical research was conducted on only two types of charts (MR and EWMA).

References

[1] J.F. MacGregor, Online statistical process control, Chem. Eng. Prog., 84 (1998) 21-31.
[2] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, Time Series Analysis: Forecasting and Control, John Wiley & Sons, New York, 1976.
[3] B.L. MacCarthy, T. Wasusri, Statistical process control for monitoring scheduling performance-addressing the problem of correlated data, J. Oper. Res. Soc, 52 (2001) 810-820.
[4] W. Jiang, K.L. Tsui, W.H. Woodall, A new SPC monitoring method: the ARMA chart, Technometrics, 42 (2000) 399-410.
[5] R. Noorossana, M. Farrokhi, A. Saghaei, Using neural network to detect and classify out-of-control signals in autocorrelated processes. Qual. Reliab. Eng. Int., 19 (2003) 493-504.
[6] S. Hwang, H. Chen, C. Chang, An exponentially weighted moving average method for identification and monitoring of stochastic systems, Ind. Eng. Chem. Res., 47 (2008) 8239-8249.
[7] D.W. Apley, Design of exponentially weighted moving average control charts for autocorrelated processes with model uncertainty, Tehcnometrics, 45 (2012), 187-198.
[8] D.G. Wardell, H. Moskowitz, R.D. Plante, Control charts in the presence of data correlation, Manag. Sci., 38 (1992) 1084-1105.
[9] J.R. English, S. Lee, T.W. Martin, C. Tilmon, Detecting changes in autoregressive processes with $\overline{X}$ and EWMA Charts. IIE Trans., 32 (2000), 1103-1113.
[10] C.W. Lu, M.R. Reynolds, Control charts for monitoring the mean and variance of autocorrelated processes, J. Qual. Technol., 31 (1999) 259-274.
[11] V.C.C. Vargas, L.F.D. Dias, A.M. Souza, Comparative study of the performance of the CUSUM and EWMA control charts, Comput. Ind. Eng., 46 (2004) 707-724.
[12] P. Winkel, N.F. Zhang, Serial correlation of quality control data–on the use of proper control charts, Scand. J. Clin. Lab. Inv., 64 (2004) 195-204.
[13] N. Ye, S. Vilbert, Q. Chen, Computer intrusion detection through EWMA for autocorrelated and uncorrelated Data, IEEE T. Reliab., 52 (2003), 75-82.