DOI: https://dx.doi.org/10.4314/gjpas.v27i2.9

GLOBAL JOURNAL OF PURE AND APPLIED SCIENCES VOL. 27, 2021: 171-179
COPYRIGHT© BACHUDO SCIENCE CO. LTD PRINTED IN NIGERIA ISSN 1118-0579
www.globaljournalseries.com, **Email: info@globaljournalseries.com**

PERFORMANCE EVALUATION OF CONVENTIONAL EXPONENTIALLY WEIGHTED MOVING AVERAGE (EWMA) AND P-VALUE CUMULATIVE SUM (CUSUM) CONTROL CHART

IKPOTOKIN, O., BRAIMAH, J. O AND OBOH, H. E.

(Received 15 March 2021; Revision Accepted 15 April 2021)

ABSTRACT

This paper is aimed at comparing the performances of the conventional Exponentially Weighted Moving Average (EWMA) and p-value Cumulative Sum (CUSUM) control chart. These charts were applied in monitoring the outbreak of pulmonary tuberculosis in Delta State University Teaching Hospital (DELSUTH), Oghara for a period of eighty four (84) calendar months. Line chart and histogram were plotted to test for stationary and normality of the data. Autocorrelation plot was also used to study the randomness of the data. The results of the control charts show that conventional EWMA chart detects shifts faster in monitoring process mean than the p-value CUSUM control chart.

KEYWORDS AND PHRASES: Exponentially Weighted Moving Average (EWMA), p-value, Cumulative Sum (CUSUM), Autocorrelation, Randomness.

INTRODUCTION

Control charts are classified in view of whether to use or not to use the historical values of control statistic. Shewhart control charts are applied based on the information of the process contained in the present observations only and it disregards any information given by the entire series of points [1]. Hence, Shewhart control chart is categorized as control chart that has no memory. Therefore, Shewhart control chart has been established to be less susceptible in detecting smaller shifts, mainly smaller than 1.5 times the standard deviation [2]. When smaller shifts are of significance, an alternative to Shewhart control chart is Cumulative Sum (CUSUM) control chart and the Exponentially Weighted Moving Average (EWMA) control chart [3]. These charts are based on reminiscence and perform better than Shewhart chart in detecting smaller shifts. The information from the past observations is cumulated up to the recent sample and then the decision about the mean process is taken [4].

Also for CUSUM chart, the EWMA chart is better in detecting small shifts in the process mean. These charts are used for monitoring the mean of a process on the basis that samples are taken from the process at set times (i.e, shifts, hours, days, weeks, months, year). The measurements of these samples at a certain time comprise a subgroup [5]. The EWMA chart depends on the specification of target value and a reliable approximate of the standard deviation. For this reason, it is better to adopt the moving average chart after process control has been established. The exponentially weighted moving average (EWMA) control chart scheme was first introduced by Roberts in 1959, which is a good substitute to Shewhart chart when one is interested in a small process shift [6].

In a situation of hypothesis testing, early testing practice make decisions using the idea of rejection and acceptance region [7]. A null hypothesis will be rejected when the observed value of the associated test statistic falls within the rejection region. This usual way of hypothesis testing has therefore been replaced by the p-value since the p-value approach cannot just make a decision about the hypotheses, but as well tell us how strong the data in the observed values that is against the null hypothesis [8].

Motivated by this p-value method in hypothesis testing, the p-value approach is used to design a CUSUM chart also. By p-value approach, for a set control chart, the in-control (IC) distribution of control charting statistic is initially estimated [9]. Subsequently, at a given point, the p-value equivalent to the observed value of such

Ikpotokin, .O., Department of Mathematics and Statistics, P.M.B 14, Ambrose Alli University, Ekpoma, Edo State, Nigeria

Braimah, J. O., Department of Mathematics and Statistics, P.M.B 14, Ambrose Alli University, Ekpoma, Edo State, Nigeria

Oboh, H. E., Department of Mathematics and Statistics, P.M.B 14, Ambrose Alli University, Ekpoma, Edo State, Nigeria

charting statistic is then obtained. If the p-value is smaller than a specified significance level, the chart indicates a process distributional shift [10].

In contrast to conventional control charts using control limits, the p-value method has an advantage that at a known time point, even if a shift in the process is not detected, the p-value can give us a quantitative measure of the possibility of a potential shift, so that the subsequent sampling interval can be adjusted properly. All control charts using p-value approach have the same format, such that the vertical axis always ranges between 0 and1, denoting the p-values, and only one control limit corresponds to the significance level, therefore making the charts more suitable to use. The p-value computation of the charting statistic of the control chart was discussed in [11]. When the in-control process distribution is assumed to be normal with a known variance, [12] gave an approximate formula for the in control distribution of the statistic of the CUSUM chart. [8] studied the rate in multistage process mean monitoring, wherein p-value calculation of the charting statistics were applied in different stages of the process monitoring. Nevertheless, comparison of EWMA and p-vale CUSUM is still not much available in literature.

METHODS / CALCULATIONS

Cumulative Sum (CUSUM) Chart

The main feature of the CUSUM control chart scheme is that it cumulates the difference between observed value and a set target value, μ_a. The cumulative sum of deviations from the target value, μ_a given by

$$C_i = \sum_{j=1}^{i}(X_j - \mu_a) \quad (1)$$

which is plotted on a chart or tabulated in order to detect an upward or downward shift (change) from target. The tabular CUSUM statistic

$$C_i^{+} = \max[0, X_i - (\mu_a + K) + C_{i-1}^{+}] \quad (2)$$

$$C_i^{-} = \min[0, X_i - (\mu_a - K) + C_{i-1}^{-}] \quad (3)$$

is used to discover an increase and a decrease in the process mean shift, where μ_a is the acceptable process mean. The process is said to be out-of-control if $C_i \geq H$ where H is the decision interval [13].

DETERMINING THE VALUE OF K AND H

Parameter K is known as the reference value for the CUSUM control chart scheme and is chosen to be between acceptable process mean μ_a, which the CUSUM scheme is said to detect quickly. [14] concludes that for a control scheme designed to detect a specific mean shift of δ, a value of $\frac{\delta}{2}$ be used, gave the formula for determining K as

$$K = \frac{\delta}{2}\sigma = \frac{|\mu_d - \mu_a|}{2} \quad (4)$$

where δ is the variation from the acceptable process mean. This reference value is also known as the reference value for hypothesis testing.

$$H_0: \mu = \mu_a \quad against \quad H_1: \mu = \mu_d \quad (5)$$

Subsequent to the selection of K, the decision interval H is obtained from extracted table from CUSUM Average Run Length (ARL) without or with First Initial Response (FIR). The value of H should give appropriately large ARL when the process is at its desired Acceptable Quality Level (AQL) and appropriately small ARL when the process has shifted to undesired Reject-able Quality Level (RQL). Usually the FIR feature is used to ascertain a faster detection in case of start-up problems after a control action. An alarm is signaled whenever the CUSUM statistic exceeds H.

THE p-VALUE CUSUM APPROACH

A null hypothesis would be rejected when the observed value of the related test statistic falls within the rejection region [2]. This conventional way of hypothesis testing has been replaced by the p-value approach, because the p-value method cannot only make a decision about the hypotheses, but also tell us the evidence in the observed data is against the null hypothesis.

By the p-value approach in hypothesis testing, suggestion is made in designing a control charting scheme using the p-value approach. By the p-value approach, for a known control chart, the in-control (IC) distribution of the charting statistic is computed first. Subsequently, at a given time point, the p-value corresponding with the observed value of the charting statistic is then obtained. If the p-value is less than a pre-specified level, then the chart signals a process shift [15, 16].

ASSESSMENT OF THE PERFORMANCE OF p-VALUE CUSUM CHART

Several authors have studied CUSUM control scheme on the basis of Average Run Length (ARL) computation. Instead of using ARL at a specified quality level which is the average number of samples taken before an out of control signal is detected, the P-value is used. The chart gives a signal of mean shift when $C_t^{+} > h$, where h is a control limit (Decision interval) chosen to achieve a pre-specified IC ARL (denoted as ARL0) value. Instead of comparing the charting statistic C_t^{+} with the control limit (Decision interval) value h; here, computing the p-value corresponding to the value of C_t^{+} and thereafter comparing the p-value with a pre-specified significance level α for making a decision whether the process is out-of-control (OC) are used. Hence, there is a need to find the in-control distribution first. Some researchers provided an approximation formula for this IC distribution in cases when the IC process distribution is normal with a known variance [17, 18].

At a given time point and a given allowance constant k, generate the Phase II observations $X_1, X_2, \ldots, X_T$ and compute the reference value. This process is then repeated several times (e.g., a million times), and the experimental distribution of C_t^{+} can be determined by the computed C_t^{+} values. For a given observed value of C_t^{+}, denoted as C_t^{+}, the corresponding p-value is then computed by:

$$P_t^{+} = P(C_t^{+} > C_t^{+*}) \text{ and } d_n = T_{n+1} - T_n \quad (6)$$

where C_t^{+*} is the CUSUM values greater than zero.

THE EXPONENTIAL WEIGHTED MOVING AVERAGE (EWMA)

The Exponential Weighted Moving Average (EWMA) chart is used for monitoring process by dividing the data such that it gives less weight to older data as samples are taken and give a more weight to most recent data. It is also effective in detecting smaller shifts.

The EWMA charting procedure is used to monitor the rate of occurrence of uncommon events where the time between successive occurrences is exponentially distributed. This procedure can be used extensively in time series forecasting and modeling, [19]. The EWMA for individual value may be defined as:

$$Z_i = wX_i + (1-w)Z_{i-1} \quad (7)$$

where $0 < w \leq 1$ and $i = 1,2,\ldots,n$

Assuming the following notations: Z_t = Exponentially Weighted Moving Average, w = EWMA weight parameter (0 < w < 1), μ_0 = Process mean, X_{ij}= jth measure of ith subgroup, with σ = standard deviation of the process, n_i = Sample size of ith group, q = σlimit, $\bar{X}_i$= Mean of measurement in ith subgroup. If $n_i = 1$, then subgroup mean trim down to the observation in the group.

If a known value (μ_0) for μ is specified, $E_0 = \mu_0$, otherwise $E_0 = X$. The previous equation can be rewritten as:

$$Z_t = Z_{i-1} + w(X - Z_{i-1}) \quad (8)$$

which states the current EWMA plus the weighted error in the prediction of the current mean based on the previous mean. The EWMA for the ith subgroup can be written as:

$$Z_t = w(1-w)jX_i - j + (1-w)^i E_0 \quad (9)$$

expresses the EWMA as weighted average of historical subgroup means where the weight reduces exponentially as more weight is assigned to recent subgroup mean.

COMPUTATION OF CONTROL LIMIT

The central line on EWMA chart indicates an estimate for μ, which is computed as:

$$\hat{\mu} = \bar{\bar{X}} = \frac{n_1\bar{X}_1 + \cdots + n_N\bar{X}_N}{n_1 + \cdots + n_N} \quad (10)$$

Therefore, when the subgroup sample sizes remain constant, the control limit width increases monotonically with i. For probability limits, replace k with $\emptyset - 1(1 - \frac{\alpha}{2})$ in the earlier equations [7], where $\emptyset$ =shift size and α = level of significance.

Note that the EWMA z_i is a weighted average of all observations that precede it. For example:

For $i = 1$, $z_1 = \lambda x_1 + (1-\lambda)z_0$

For $i = 2$, $z_2 = \lambda x_2 + (1-\lambda)z_1$

$= \lambda(1-\lambda)^0 x_2 + (1-\lambda)^1 \lambda x_1 + (1-\lambda)^2 z_0$

For $i = 3$, $z_3 = \lambda x_3 + (1-\lambda)z_2$

$= \lambda x_3 + (1-\lambda)$

$= (1-\lambda)^0 \lambda x_3 + (1-\lambda)^1 \lambda x_2 + (1-\lambda)^2 \lambda x_1 + (1-\lambda)^3 z_0$

We recursively can write each z_i (if 0 < λ < 1) as:

$$z_i = \lambda \sum_{j=0}^{i-1} (1-\lambda)^j x_{i-j} + (1-\lambda)^i z_0 \quad (11)$$

Recall $\sum_{j=0}^{i=1} p^j = \frac{1-p^i}{p}$ for $|p| < 1$. If $p = 1 - \lambda$, then the sum of the weights in (11) is

$$z_i = \lambda \sum_{j=0}^{i-1} (1-\lambda)^j + (1-\lambda)^i \quad (12)$$

The fact that the weights decrease exponentially is the reason it is called an exponentially weighted moving average control chart.

As i increases, the control limits approach constant values.

If the observations x_i are said to be independent with common variance σ^2, then the variance z_i is:

$$\sigma_{z_i}^2 = Var\left(\lambda \sum_{j=0}^{i-1} (1-\lambda)^j x_{i-j} + (1-\lambda)^i z_0\right)$$

$$= \lambda^2 \sum_{j=0}^{i-1} (1-\lambda)^{2j} \sigma^2 + 0$$

$$= \lambda^2 \frac{1-(1-)^{2i}}{1-(1-)^2} \sigma^2$$

$$\lambda^2 \frac{1-(1-)^{2i}}{1-(1-)^2} \sigma^2 = \frac{\lambda}{2-\lambda}(1-(1-)^{2i})\sigma^2 \quad (13)$$

On replacing λ with w, when μ_0 and σ^2 are known, the EWMA chart is build up by plotting z_i versus the sampled number *i* with control limits at:

$$\mathrm{UCL} = \text{Uppercontrol limit} = \bar{\bar{\mathrm{X}}} + \mathrm{q}\sigma_{\mathrm{w}} \sqrt{\frac{\sum_{\mathrm{j}=0}^{\mathrm{i}-1}(1-\mathrm{w})^2}{\mathrm{n_i}-\mathrm{j}}} \quad (14)$$

Centerline = μ_0

$$\mathrm{LCL} = \text{Lowercontrollimit} = \bar{\bar{\mathrm{X}}} - \mathrm{q}\sigma_{\mathrm{w}} \sqrt{\frac{\sum_{\mathrm{j}=0}^{\mathrm{i}-1}(1-\mathrm{w})^2}{\mathrm{n_i}-\mathrm{j}}} \quad (15)$$

Equation (14) and (15) can be transformed to become:

$$\mathrm{UCL} = \bar{\bar{\mathrm{X}}} + \mathrm{q}\sigma \sqrt{\frac{\mathrm{w}}{\mathrm{n}(2-\mathrm{w})}} \quad (16)$$

$$\mathrm{LCL} = \bar{\bar{\mathrm{X}}} - \mathrm{q}\sigma \sqrt{\frac{\mathrm{w}}{\mathrm{n}(2-\mathrm{w})}} \quad (17)$$

where 'n' is the subgroup size, the control limits becomes:

The resulting control limits are:

$$\mathrm{UCL} = \bar{\bar{\mathrm{X}}} + \mathrm{q}\sigma \sqrt{\frac{\mathrm{w}}{\mathrm{n}(2-\mathrm{w})}} \quad (18)$$

$$\mathrm{LCL} = \bar{\bar{\mathrm{X}}} - \mathrm{q}\sigma \sqrt{\frac{\mathrm{w}}{\mathrm{n}(2-\mathrm{w})}} \quad (19)$$

DISCUSSION OF RESULTS AND APPLICATION

The set of data used in this study are secondary data on the pulmonary tuberculosis cases in DELSUTH, which consist of the monthly record from 2012 to 2018 as shown in the appendix. Statistical software, MINITAB 17 was implored to carry out the statistical analysis and plotting of control charts while *anygeth* software was used in obtaining the target values (k) and reference values (h). The findings from the study and application are presented below.

TEST FOR STATIONARY AND NORMALITY OF THE DATASET.

This study uses Augmented Dickey Fuller test to test for stationary of the dataset in order to ascertain the validity and the result is shown in Table 1 and Histogram chart is to ascertain Non-normality if it is not bell shape as shown in Figure 1.

Table 1: Augmented Dickey Fuller test with p-value

Dickey Fuller	p-value
-2.8872	0.212

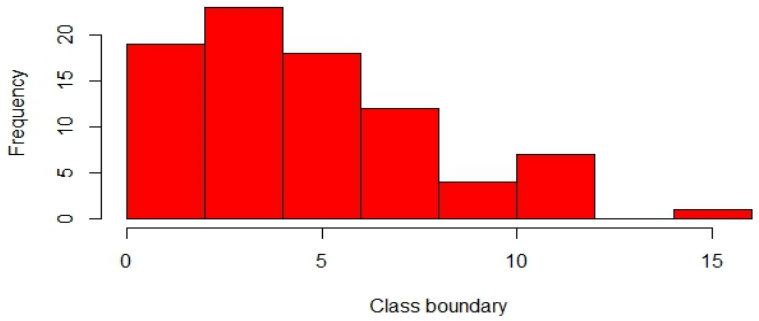

Figure 1: Histogram of the pulmonary tuberculosis
The histogram above is not normally distributed since it has long tail to the right (i.e., positively skewed).

AUTO-CORRELATION

The study considers the Autocorrelation chart to test for randomness and the significance of the lags. Any lag that is above the line dot is random and significant as shown in Figure 2 and Figure 3.

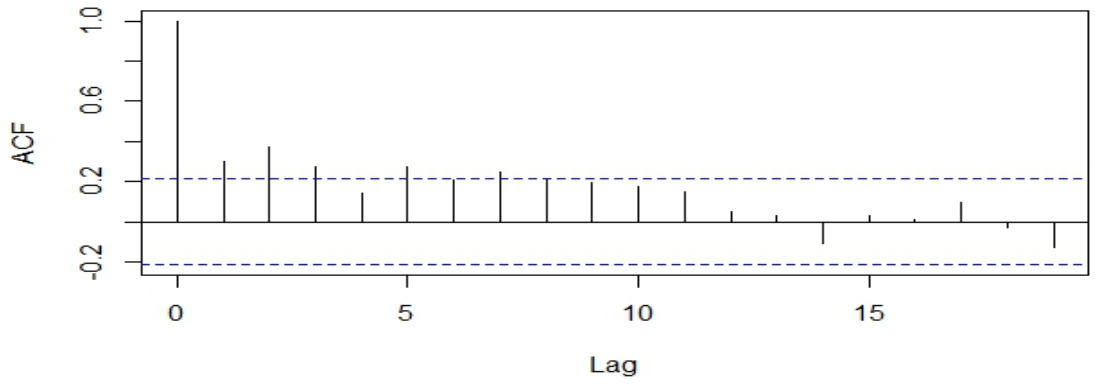

Figure 2: Correlogram for the Pulmonary Tuberculosis
Autocorrelation Function (ACF) chart in Figure 2 shows that lag 0, 1, 2,3, 5and 7are random and significant

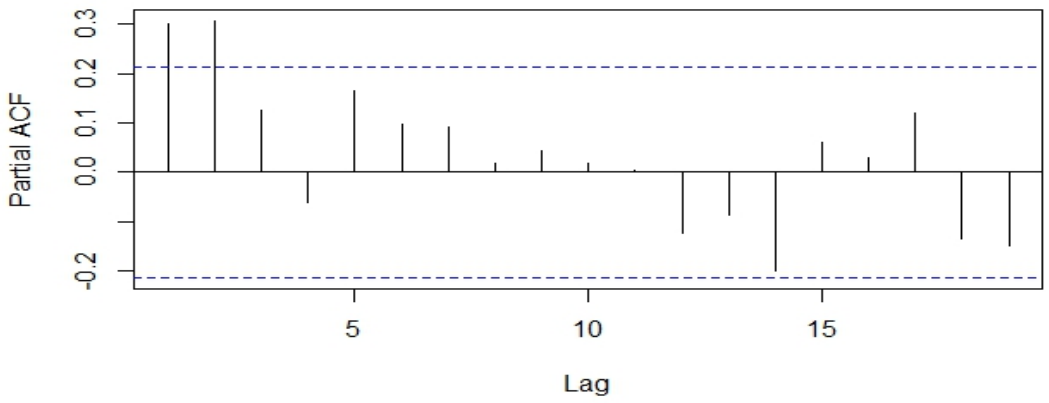

Figure 3: The chart showing the partial correlogram for the pulmonary tuberculosis

The Partial Autocorrelation Function (PACF) chart in Figure 3 shows that lag 1 and 2 are random and significant.

The idea of CUSUM chart design for the data on pulmonary tuberculosis cases in DELSUTH is detecting the first shift when the process is actually in control given that $\sigma = 1.5$ and 2.0 as shown in Figures 4 and 5.

CUSUM CHART

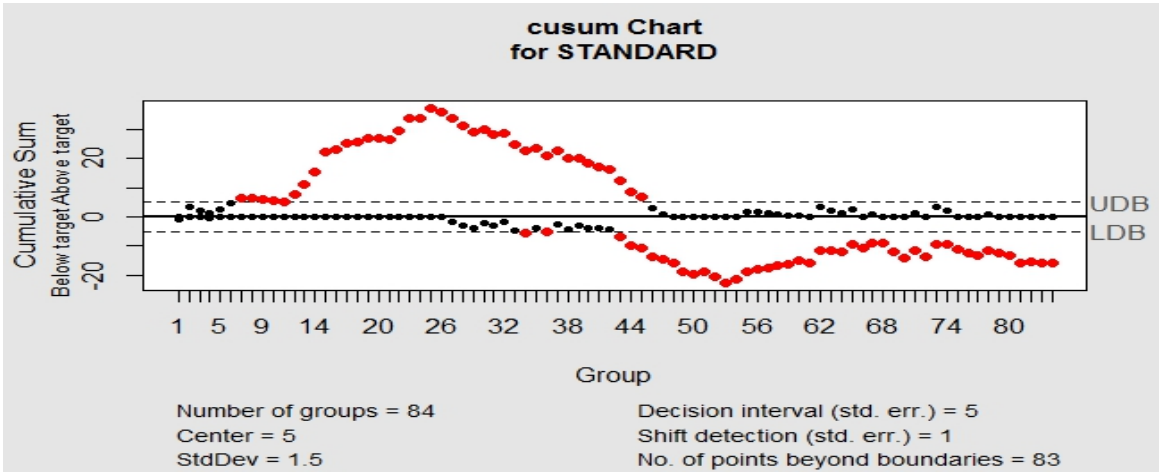

Figure 4: The CUSUM chart for σ = 1.5

The first shift is detected at month 6 in the chart of Figure 4 with ARL = 6. This indicates that at the sixth month, the tuberculosis control system was out of control in DELSUTH. Therefore, the management system needs to intensify more efforts to bring the process in control so that the outbreak can be under control.

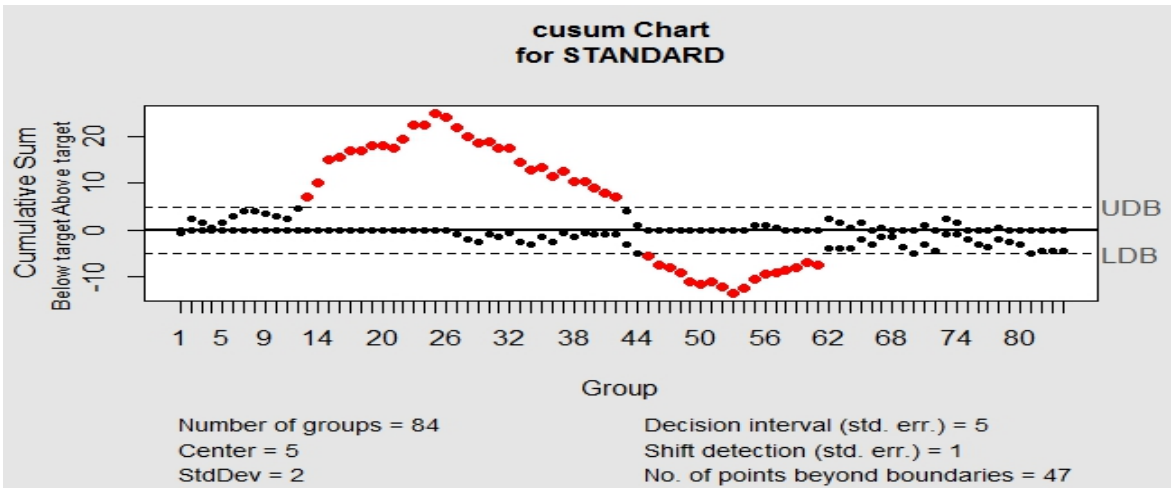

Figure 5: The CUSUM chart for σ = 2.0

The first shift is detected at month 12 in the chart of Figure 5 with ARL = 12. This indicates that at the twelfth month, the tuberculosis control system was out of control in DELSUTH. Therefore, the management system needs to intensify more efforts to bring the system in control so that the outbreak can be under control.

Generally, it is observed that there were shift in the process mean for the various values of σ, and the smaller the value of σ, the faster the shift is detected.

p-VALUE FOR THE STATISTICAL PROCESS CONTROL

The p-Value CUSUM method is also used to monitor the outbreak of tuberculosis data used in this study in order to compare the performances of the two schemes. In designing p-value CUSUM chart, the dataset is conditioned to be stationary at p-value < 0.05 and also be in statistical control as shown in Figures 6 and 7.

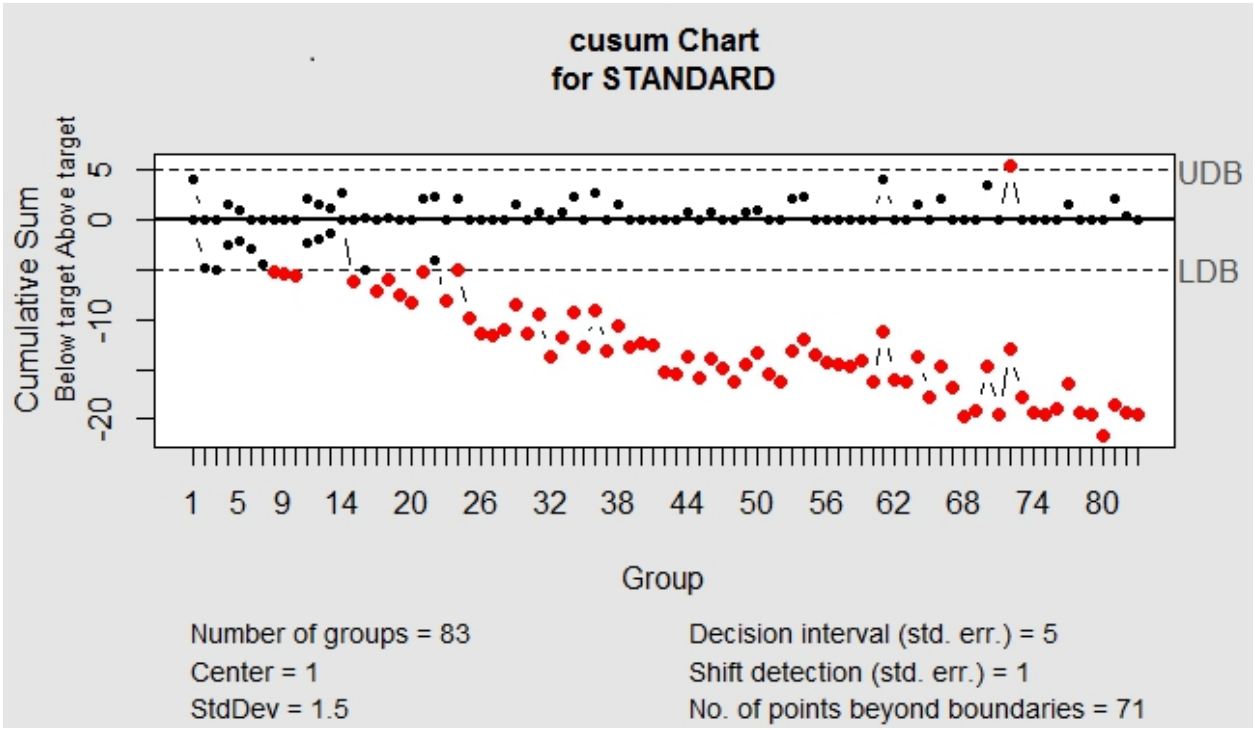

Figure 6: The p-value CUSUM chart for σ = 1.5
The first shift is detected at month 7 in the chart of Figure 6 with the ARL = 7

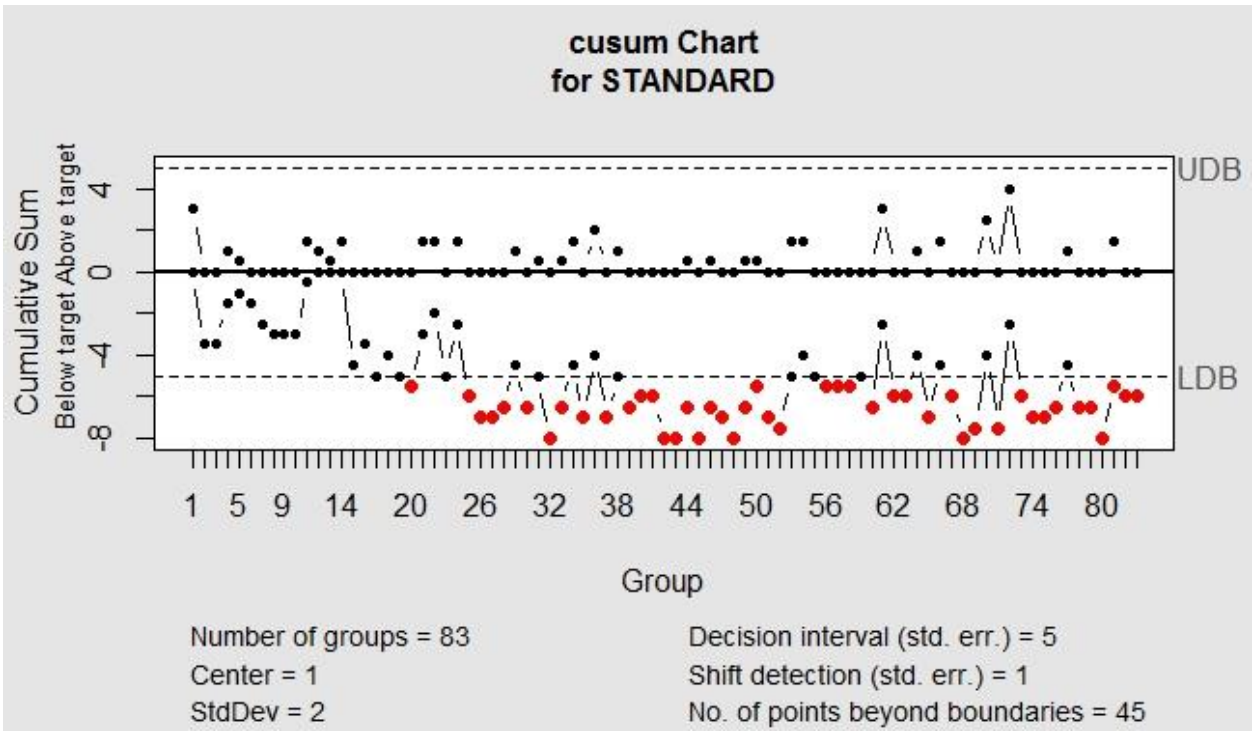

Figure 7: The p-value CUSUM chart for σ = 2.0
The first shift was detected at month 19 in the chart of Figure 7 with the ARL = 19**.**

DESIGN OF EWMA CONTROL CHART SCHEME

The EWMA control chart scheme was also applied on the data set. The overall mean per month is approximately$\bar{X} = 0.366$, the standard deviation from the historical data is $\sigma = 3.358$ and the subgroup size (n) is one. The EWMA control chart is then plotted taken w = 0.15 and 0.20 as shown in Figures 8 and 9 respectively.

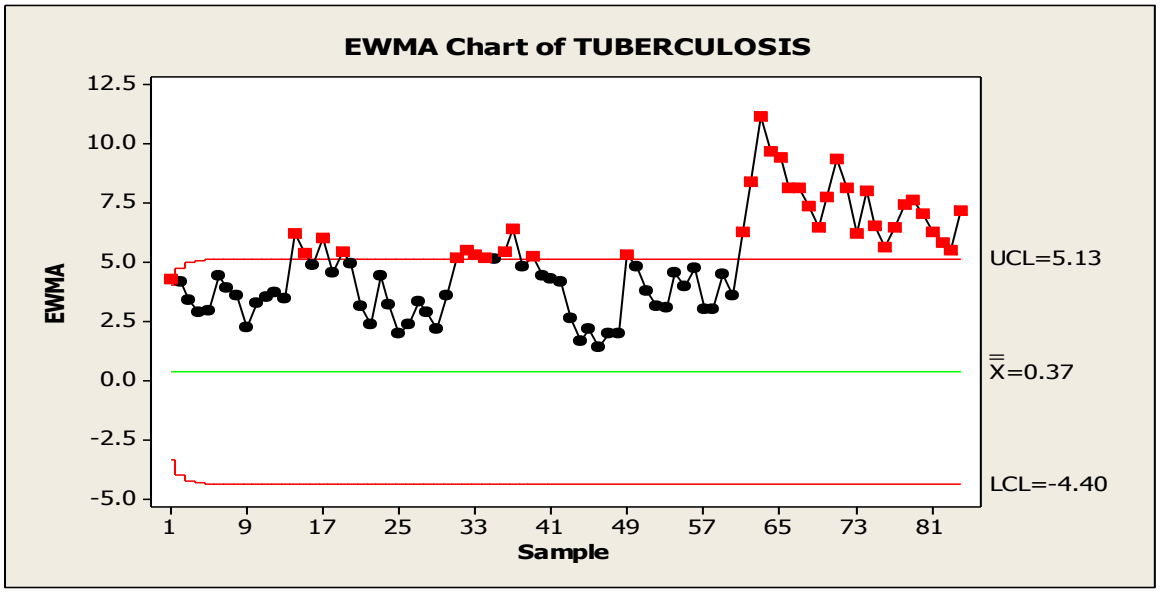

Figure 8: EWMA control chart for w = 0.15

From Figure 8, out of control signals or shift were observed at points: 1, 14, 15, 17, 19, 31, 32, 33, 34, 36, 37, 39, 49, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83 and 84.

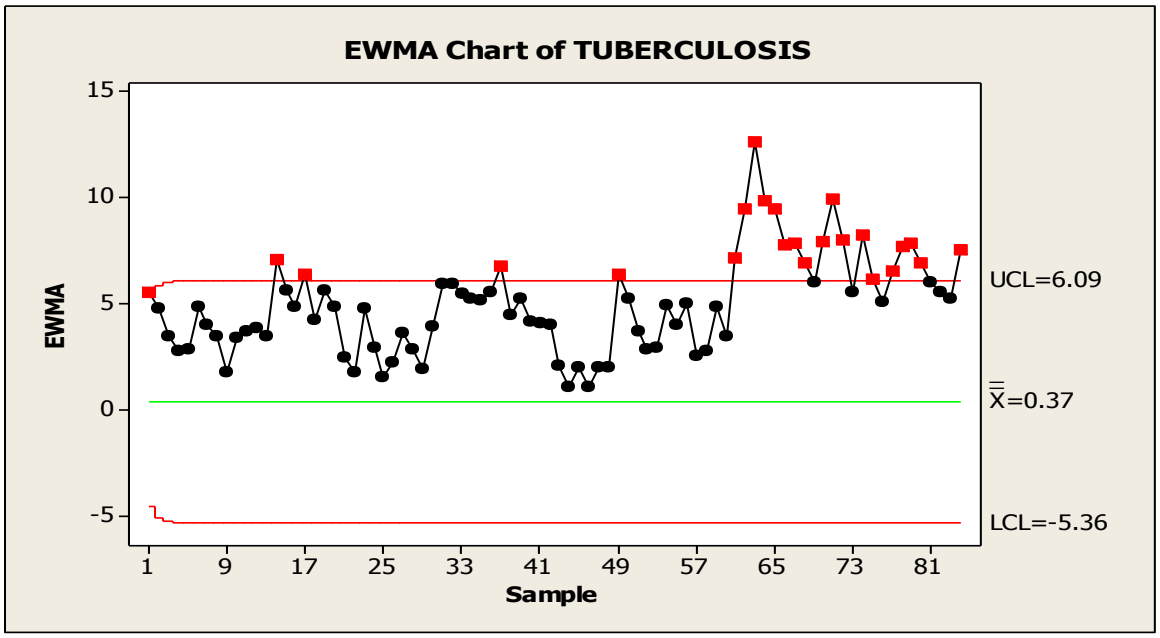

Figure 9: EWMA control chart for $\boldsymbol{w}$ = 0.20

From Figure 9, out of control signals or shift were also observed at points: 1, 14, 17, 37, 49, 61, 62, 63, 64, 65, 66, 67, 68, 70, 71, 72, 74, 75, 77, 78, 79, 80 and 84.
Generally, from the EWMA charts of Figures 8 and 9, the process is said to be out of control for the two weighted parameter (w) and this indicate an out of control signals at the same points. For w =0.15 and 0.20, the two charts detected the first out of control points (signals) at point 1 which correspond to January of the base year.

CONCLUSIONS

The paper concluded that the early shift was detected when the σ is small in p-value CUSUM. Also, the conventional EWMA control chart is better in detecting early shift in the mean process than p-value CUSUM and; from the observation, it is noticed that there was always a high outbreak of Tuberculosis in month 7 which corresponds to the month of July for p-value CUSUM and January for EWMA control chart.

REFERENCES

Han D. and Tsung F. 2014. A Generalized EWMA Control Chart and its Comparison with the Optimal EWMA, CUSUM and GLR Schemes. The Annals of Statistics Vol. 32 issue 1, pp 316-339.

Montgomery D. C. 2007. SPC Research-Current Trends. Quality and Reliability Engineering International Vol. 23, pp 515-516.
Alpaben K. P and Jyoti D. 2016. Modified exponentially weighted moving average (EWMA) control chart for an analytical process data. Journal of Chemical Engineering and Materials Science Vol. 2 issue 1, pp. 12-20.

Zhang J., Li Z. and Wang Z. 2012. A New Adaptive Control Chart for Monitoring Process Mean and Variability. The International Journal of Advanced Manufacturing Technology Vol. 60 issue 9- 12, pp 1031-1038.

Wu S. 2011. Optimal Inspection Policy for Three-State Systems Monitored by Variable Sample Size Control Charts. The International Journal of Advanced Manufacturing Technology Vol. 55 issue 58, pp 689-697.

Montgomery D. C. 2009. Introduction to Statistical Quality Control, Seventh Edition, John Wiley and Sons, Inc.

Chatterjee S. and Qiu P. 2009. Distribution-free Cumulative Sum Control Charts using Bootstrap-based Control Limits. The Annals of Applied Statistics Vol. 3 issue 1, pp 349- 369.

Li Y. and Tsung F. 2009. False Discovery Rate-Adjusted Charting Schemes for Multistage Process Monitoring and Fault Identification. Technometrics Vol.51, pp 186-205.

Qiu P. and Li Z. 2011. On Nonparametric Statistical Process Control of Univariate Processes. Technometrics. Vol. 53, pp 390-405.

Jensen W. A., Jones-Farmer L. A., Champ C. W. and Woodall W. H. 2006. Effects of Parameter Estimation on Control Chart Properties: A Literature Review. Journal of Quality Technology Vol. 38 issue 4, pp 349-364.

Jones L. A., Champ C. W. and Rigdon S. E. 2004. The Run Length Distribution of the CUSUM with Estimated Parameters. Journal of Quality Technology Vol. 36 issue1, pp 95-108.

Grigg O. A. and Spiegelhalter D. J. 2008. An Empirical Approximation to the Null Unbounded Steady-State Distribution of the Cumulative Sum Statistic. Technometrics Vol. 50, pp 501-511.

Hawkins D. M., Qiu P. and Kang C. W. 2003. The Change Point Model for Statistical Process Control. Journal of Quality Technology Vol. 35 issue 4, pp 355-366.

Hawkins D. M. and Olwell D. H. 2007. Cumulative Sum Charts and Charting for Quality Improvement. Springer, Berlin.

Montgomery D.C. 2004. Introduction to Statistical Quality Control, Fifth Ed. John Wiley and Sons, New York.

Lucas J.M. 2012. Combined Shewart CUSUM Quality Control Schemes, Journal of Quality Technology Vol. 8, pp 1-89

Saniga E. M., McWilliams T. P., Davis D. J. and Lucas J. M. 2006. Economic Advantages of CUSUM Control Charts for Variables. In Frontiers in Statistical Quality Control Vol. 8, pp 185-198

Wu Z., Zhang S. and Wang P. 2007. A CUSUM Scheme with Variable Sample Sizes and Sampling Intervals for Monitoring the Process Mean and Variance. Quality and Reliability Engineering International Vol. 23, pp 157-170.

Hawkins D. M. 2015. A Fast Accurate Approximation for Average Run Lengths of CUSUM control charts. Journal of Quality Technology Vol. 24 issue 1, pp 37-43.

Appendix

Laboratory Confirmed Pulmonary Tuberculosis Cases for a Period of Seven Years

Year / Month	2012	2013	2014	2015	2016	2017	2018
JANUARY	3	11	11	8	0	3	11
FEBRUARY	11	12	4	2	3	11	4
MARCH	4	16	2	6	5	4	2
APRIL	4	7	2	3	2	4	2
MAY	8	9	3	4	1	8	3
JUNE	9	6	7	4	6	2	7
JULY	8	8	3	0	8	7	3
AUGUST	6	6	6	0	6	4	3
SEPTEMBER	5	5	0	3	5	0	0
OCTOBER	5	10	3	0	5	1	5
NOVEMBER	5	12	7	3	5	8	4
DECEMBER	10	6	2	2	4	1	4

Source: Delta State University Teaching Hospital (DELSUTH), Oghara.