Science World Journal Vol. 15(No 3) 2020
www.scienceworldjournal.org
ISSN 1597-6343
Published by Faculty of Science, Kaduna State University
https://doi.org/10.47514/swj/15.03.2020.013

Full Length Research Article

COMPARISON OF CHANGE-POINTS IN MULTIVARIATE STATISTICAL PROCESS CONTROL USING THE PERFORMANCE OF LAPAGE-TYPE (NONPARAMETRIC)

*S.S. Umar[1], M.K. Dauda[2], A.G. Sadiq[3] and A.S. Magaji[4]

[1,3]Department of Mathematics and Statistics, Kaduna Polytechnic, Nigeria
[2,4]Department of Mathematical Sciences, Kaduna State University, Nigeria

*Corresponding Author's Email Address: umarss83@yahoo.com

ABSTRACT
The inability of the Shewhart's, the EWMA, and the CUSUM, Hotelling's T^2 and many other control charts to indicate the time of shift poses great problems in production, Medicine, etc. To overcome the problems the need to identify the period of change (shift) in the process becomes inevitable. The study used Lapage-type Change-point (LCP) to detect the simultaneous shift in both mean and variance. In the study we compare the performance of generalized likelihood ratio change-point (GLRCP) a parametric-base with our proposed method (LCP) at different varying start-ups using real life data. We run the data on Normal, Laplace and Lognormal distributions and also Average Run Length (ARL_0) to assess the performance of the methods. Evaluating in-control ARLs (IC-ARLs) for each of the methods at change-point 250 and ARL_0 500 indicates the same performance irrespective of the start-up value; LCP and GLR methods have rather a similar performance IC-ARLs at change-point 50 and change-point 100 under the normality assumptions, but under non-normal distributions, LCP has substantially higher IC-ARLs compared to GLRCP at 20. The LCP outperformed the GLRCP when applied to children bronchial pneumonia status. We therefore recommend that new method be used in short-run situations and also when underlying distributions are usually unknown.

Keywords: Lapage-type Change-point, nonparametric statistic, Laplace and Lognormal distributions, Rank-Test Statistics

1.0 Introduction

Statistical Process Control (SPC) is about continuous monitoring or surveillance of a process to ensure that neither the mean nor the variability of the process distribution has changed Hawkins and Zamba, (2005); McCracken and Chakraborti, (2013). Some examples of process control may include: monitoring some quality characteristics of manufactured item to ensure compliance to certain standards; detection of an increased birth rate of infants with congenital malformations; surveillance of health data to detect an outbreak of a disease or increased rate of incidence of disease such as COVID-19, Salisu, Edokpa, Elakhe and Shaib (2018); the observance of a natural phenomenon such as water salinity levels, or adverse drug reaction etc. in public health practice, Flowers (2009). Note that process control is very useful in showing variation that exists in health outcomes or performance between groups or institution, and claimed that it is often a starting-point for needs assessment, targeting service and epidemiological understanding. In general, the target goal of process control is to detect the change in the process occurring at an unknown epoch of time as soon as possible after it has occurred, and at the same time controlling the rate of false alarm Dong, Hedayat and Sinha, (2008).

Change- Point Approach

Control chart, play an important role in SPC applications and distinguish between common causes of variations and special causes, they do not indicate when these special causes actually occurred Snoussi and Limam, (2007). The process of estimating the period of change as a result of shift detection is commonly referred to as change-point. In other words, the change-point approach to SPC, intuitively, is a technique targeted at knowing the period of process shift. Along with indication of a loss of control, they provide estimates of when the shift occurred and (if needed) of the pre and post-shift process parameters. By the change-point approach, primarily, the signal of process shift and time of shift can be obtained simultaneously Li, Qiu, Chatterjee, and Wang, (2013). Meanwhile, one major challenge in SPC is whether there has been a shift of the distribution from the target in a process quality; another is locating the point of change. Thus, it is necessary to proffer solution in terms of required and corrective methods to address the unavoidable question of the process shifts and of particular interest lies in the position of shift in an out of control process in order to cushion the consequences brought about by the change. In practice, detecting if the process is in control as well as the position of shift in an out of control processes are critical statistical tasks. When a control chart produces an out-of-control signal, a search must be initiated to find the assignable causes of the out-of-control state. Knowing the exact time of a change in a process restrict the range of searches for the assignable causes which in turn accelerate the assigned causes identification and appropriate corrective action implementation. In cases when a variable sampling scheme is adopted, even if a shift is not detected at a given time point, it would still be helpful to know the likelihood of a potential shift Li et al. (2013).

Objectives of the Study

The study of the statistical performance of charts is very important became it provides insight into how charts work in practice and it provides the only way to effectively compare competing methods in a fair and objective manner. In order to implement the idea of simultaneously monitoring process quality in practice, performances of the proposed and the competing methods are evaluated and compared. This is practically meant to demonstrate the inappropriateness of Generalized Likelihood Ratio (GLR)

Science World Journal Vol. 15(No 3) 2020
www.scienceworldjournal.org
ISSN 1597-6343
Published by Faculty of Science, Kaduna State University
https://doi.org/10.47514/swj/15.03.2020.013

parametric-based method of Hawkin and Zamba (2005b), as well as the inadequacy of Deng (2009) nonparametric based Mann-Whitney method (purposefully designed for a process shift in mean). The overall aim of this study is to apply Change-Point to multivariate in Statistical Process control.

Specific Objectives

The specific objectives are to:

(a) Develop multivariate control chart for detecting a change/shifts in mean and variance.

(b) Determine the required number of observations before the actual process monitoring commences;

(c) Compare the performance of multivariate Lapage Change-point (MLCP) with those of the nonparametric Wilcoxon-Mann-Whitney (WMW) and parametric Generalized Likelihood Ratio (GLR) methods.

(d) Practically demonstrate the application of the proposed method using real-life and simulated data.

2.0 Materials and Method

Lepage-Type Rank Test

The nonparametric two-sample Lepage test was developed by Lepage 1971. The test is designed to carrying out equality of the location and scale parameters test simultaneously against the alternative that at least for one of the parameters, the equality does not hold. Basically, it is a combination of the Wilcoxon-Mann-Whitney and the Ansari-Bradley statistics Hutchinson, (2002); Rublik, (2005). That is, it converts both the Wilcoxon-Mann-Whitney and the Ansari-Bradley statistics to square-standardized deviations from their respective expectations and adds the results. Perhaps the most widely used two-sample rank test of equality of location parameters is the Wilcoxon-Mann-Whitney test. The Ansari-Bradley test is also used in two-sample rank test for equality of the scale parameters, though the ranking procedure is not that straightforward. However, the possibility of a two-sample test statistic which combines the Wilcoxon-Mann-Whitney and the mood statistics (with ranking procedure is more direct and straightforward) has been suggested in the literature Rublik, (2009). The nonparametric two-sample Lepage test's variant would be called Lepage-type test (for testing equality of the location and scale parameters against the alternative that at least for one of the parameters the equality does not hold) in the course of this study. Lepage-type test's proposition is straightforward in concept and simple to carry out.

Suppose that $\{y_{11}, y_{21}, \cdots, y_{n1}\}$ and $\{y_{21}, y_{22}, \cdots, y_{n2}\}$ are independent random samples from random variables Y_1 and Y_2 respectively. Assume their distributions are given as follows:

$$F(y_1; \mu_1, \sigma^2); \quad F(y_2; \mu_2, \sigma^2) \qquad (1)$$

Where μ_1, μ_2 denote the location parameters, $\sigma_1^2 > 0, \sigma_2^2 > 0$ are scale parameters and $F(.)$ is a continuous distribution function. Let $\{R_1, R_1, \cdots, R_n\}$, $n = n_1 + n_2$, denote the rank of the pooled sample $\{y_{11}, y_{21}, \cdots, y_{n1}, y_{12}, \cdots, y_{n2}\}$ random variables. Under the assumption of no change, this model (1) can be summarized in terms of a joint hypothesis as:

$$H_o: \mu_1 = \mu_2 \;\; ; \sigma_1 = \sigma_2^2 \qquad (2)$$

Vs

a) $H_1: \mu_1 \neq \mu_2 \;\; ; \sigma_1 = \sigma_2$ (representing shift in process location only)

b) $H_1: \mu_1 = \mu_2 \;\; ; \sigma_1 \neq \sigma_2$ (representing shift in process variability only)

c) $H_1: \mu_1 \neq \mu_2 \;\; ; \sigma_1 \neq \sigma_2$ (representing shift in both process location and variability)

Let T_W^2 denote the square of the standard Wilcoxon-Mann-Whitney two-sample test statistic; T_M^2 the square of the standardized mood two-sample test statistic; and L the combination of the Wilcoxon-Mann-Whitney and the Mood statistics, then Lapage-type test denoted by L is of the form:

$$L = T_W^2 + T_M^2 \qquad (3)$$

Where

$$T_W^2 = \frac{(S_W - E(S_W/H_0))^2}{Var(S_M/H_0)}$$

$$T_M^2 = \frac{(S_M - E(S_M/H_0))^2}{Var(S_M/H_0)}$$

$$S_W = \sum_{i=1}^{n_1} R_i \;\; and \; S_M = \sum_{i=1}^{n_1} \left(R_i - \frac{n+1}{2}\right)^2$$

S_W and S_M are the Wilcoxon-Mann-Whitney and the Mood rank test statistic respectively; and $E(.)$ and $Var(.)$ denote the corresponding expected value and variance of S_W and S_M under H_0. Thus;

$$E(S_W) = \frac{n_1(n+1)}{2} \quad \text{and} \quad Var(S_W) = \frac{n_1 n_2 (n+1)}{12}$$

$$E(S_M) = \frac{n_1(n^2+1)}{12} \quad \text{and} \quad Var(S_M) = \frac{n_1 n_2 (n+1)(n^2-4)}{180}$$

Hence, the Lepage-type statistic L of equation (3) could be expressed as:

$$L = \frac{12}{n_1 n_2 (n+1)} \left(S_W - \frac{n_1(n+1)}{2}\right)^2 + \frac{180}{n_1 n_2 (n+1)(n^2-4)} \left(S_M - \frac{n_1(n^2-1)}{12}\right) \qquad (4)$$

Since the distribution function F in equation (1) is continuous, the statistic L is distribution-free whenever the null hypothesis in equation (2) holds Rublik, (2009), and testing the null hypothesis (2) at α $(0 < \alpha < 1)$, level of significance, H_0 is rejected whenever $LP \geq C_\alpha$ (critical value is chosen so that the type I error rate is α).

The Proposed LCP Method

We suggest a nonparametric Lepage-type Change Point (LCP) approach to Statistical Process Control (SPC) based on Lapage-type test (which combines the Wilcoxon-Mann -Whitney and the Mood statistics for shift in process location and variability, respectively).

LCP Method Formulation

Suppose that the independent process observations $\{y_1, y_2, \cdots, y_n\}$ came from a continuous cumulative distribution function $F(y, \mu_i, \sigma_i)$, where μ_i and σ_i are the location and variability parameters respectively. Also, consider the existence of time τ (change-point) when there is a shift in mean and or in standard deviation of the process. The process reading, parallel Hawkins and Zamba (2005b), can be modeled by:

$$Y_i \sim \begin{cases} F(y; \mu_1, \sigma_1) & if \; i \leq \tau \\ N(y; \mu_2, \sigma_2) & if \; i > \tau \end{cases} \qquad (5)$$

Under the assumption of no shift, this model (2) can be summarized in terms of the hypothesis in equation (2) in which location shift occurs if $\mu_1 \neq \mu_2$ and variability shift occurs if

Science World Journal Vol. 15(No 3) 2020
www.scienceworldjournal.org
ISSN 1597-6343
Published by Faculty of Science, Kaduna State University
https://doi.org/10.47514/swj/15.03.2020.013

$\sigma_1 \neq \sigma_2$. If there is no enough evidence to reject the null hypothesis in equation (2), we will claim that the process is in state of ''statistical control'' (in-control, IC) and stable with "random causes" which cannot be removed easily from the process without fundamental changes in the process itself. On the other hand, if it shows enough evidence to reject the null hypothesis in equation (2), we will conclude that the control chart issues a signal and the process is out of "statistical control" (out-of-control, OOC) and undergoes an unusual variation due to "special causes". In principle, either or both of these shifts could occur. In addition to the τ (change-point) being an unknown parameter, μ_i (location parameter) and σ_i (variability parameter) are also unknown. let the change-point $\tau = k$ and R_i be the rank of y_i observations. And, for the fact that the sample size keeps increasing in Phase II analysis as new observation comes in, we express the Lepage-type test statistic as:

$$L_{max,n} = max_k \left| L_{k,n} \right| \quad (6)$$

where

$$L_{k,n} = \frac{12}{k(n-k)(n+1)} \left[S_W - \frac{k(n+1)}{2} \right]^2 + \frac{180}{k(n-k)(n+1)(n^2-4)} \left[S_M - \frac{k(n^2-1)}{12} \right]^2$$

$S_W = \sum_{i=1}^{k} R_i$ and $S_M = \sum_{i=1}^{k} \left(R_i - \frac{n+1}{2} \right)^2$

LCP Implementation Procedure

When the sample size is not fixed but increase, the procedure for adapting the formulation in equation (3) for use in the Statistical Process Control (SPC), setting is described similar to Hawkins and Zamba, (2005b) and Ross *et al*., (2011) as follows:

(a) Find $L_{max,n}$, after observation n has been added to the total record of the process, by

i. Obtain the standardized

$T_W = \frac{(S_W - k(n+1)/2)}{\sqrt{k(n-k(n+1)/12)}}$ and $T_M = \frac{(S_M - k(n^2-1)/12)}{\sqrt{\frac{k(n-k)(n+1)(n^2-4)}{180}}}$

ii. Calculate the sum of squares of the standardized statistics, $L_{k,n}$

iii. Determine the maximum of $L_{k,n}$ over all the possible k, $L_{max,n} = max_k \left| L_{k,n} \right|$

(b) If $L_{max,n} \leq h_n$, where h_n is some suitable control limit, then conclude that there is no evidence of a shift in either mean or variance, and leave the process running uninterrupted.

(c) If however, $L_{max,n} > h_n$, then conclude that there is evidence of a shift in the mean, the variance or both.

One of the main objectives of a control chart is to detect unusual variation as soon as possible, and at the same time keeping the probability of erroneous signal below a reasonable level, using the initial framework of Hawkins, Qiu, and Kang, (2003), while the process is in-control, the sequence of control limits (h_n) is chosen so that the conditional probability of a false alarm at each observation n given that there was no false alarm prior to n, is fixed at desirably selected constant level α. According to Dong, Hedayat, and Sinha, (2008), the type I error is usually characterized by the in-control average run length (ARL_o) to a false alarm. That is, $ARL_o = \frac{1}{\alpha}$ Likening Hawkins and Zamba (2005b), this can be written in symbols as

$$P\left[L_{max,n} > h_n | L_{max,j} < h_j; j < n \right] = \alpha \quad (7)$$

Theoretically, similar to the literature (such as Hawkins and Zamba, (2005a, 2005b); Zhou, Zou, Zhang, and Wang, (2009), submission, it does not seen possible to solve for these h_n values. So, a simulation through the use of R "cpm" package Ross, (2013) is used to estimate them.

After-Signal Diagnosis

A primary consideration when choosing a control chart to detect and estimate special causes should be the ability to signal quickly after a special cause occurs. It is, therefore, necessary to be able to point out which parameter or parameters have shifted after a signal occurs.

Change-point detection implies that there has been a process shift. The signaling of a process shift in location (mean) and/or in variability, however, normally poses the challenges of identifying which of the process parameters have shifted. Addressing this challenge would be approached by splitting the process history at the estimate change-point and carrying out a two-sample comparison test between the two resulting segments (pre-shift and post-shift), using nonparametric tests. This procedure is called after-signal diagnosis. Our proposed after-signal diagnosis would not be a problem in the face of advanced technological statistical software already developed for this task, in today's environment Zhang, Zou, and Wang, (2011). It is known that after a signal is raised, statistical software can easily be used when needed to help in diagnosing which parameters have changed.

Suppose that a process shift is detected at time t, corresponding to the value of k which maximized $L_{max,t}$. The accrued observations can then be partitioned into the subsets

$$\{Y_1, \cdots, Y_k\} \text{ and } \{Y_{k+1}, \cdots, Y_t\} \quad (8)$$

A two-sample Wilcoxon-Mann-Whitney and Mood tests can then be carried out on these two subsets and the p-value evaluated compared. The shift is most likely to constitute a location shift or variability shift as in Wilcoxon-Mann-Whitney or Mood test respectively gives the lower p-value Ross, Tasoulis, and Adams, (2011).

Performance Evaluation Method

The study of the statistical performance of charts provides insight into how charts work in practice and is the only way to effectively compare competing methods in a fair and objective manner. The choice of any control charts is primarily based on how effective they are. Thus, the main goal would be to know how to measure the effectiveness of a chart.

Under the mathematical framework used to determine the statistical performance of the charts, performance measures are chosen to give the chart a good balance between failing to signal a real shift in process level (a type II error) and signaling a shift when none has occurred (a type I error). In phase II, the probability of a signal on any one sample is sometimes used if the successive statistics plotted are independent, as may be the case with a basic Shewhart-type chart. The distribution of the false alarms is usually expressed in terms of the Average Run Length (ARL). In this study, evaluating the likely performance of a process monitoring method is therefore achieved by the ARL.

Science World Journal Vol. 15(No 3) 2020
www.scienceworldjournal.org
ISSN 1597-6343
Published by Faculty of Science, Kaduna State University
https://doi.org/10.47514/swj/15.03.2020.013

The Average Run Length (ARL)

Suppose r (which is referred to as the run length for the method) is the period at which a process monitoring method first signals. The run length is the number of samples required for a signal to occur. The probability distribution of r is, therefore, known as the run length distribution, and the average value of this distribution is called the Average Run Length (ARL) for the process monitoring method. That is, $ARL = \mu_r$. Generally, the power of control chart methods is characterized by their run lengths. Some parameter of the run length distribution is used. A measure that is often used in quality control is based on the average run length properties Ramzy and Peiris, (2013).

Quantifying what a given monitoring method might be expected to do is necessary when considering alternative methods for indicating out-of-control signals based on process monitoring data. Therefore, evaluation of the ARL plays a crucial role in the design of control charts and performance comparison. The challenges of designing control chart procedures thus involves finding control procedure parameters such that the ARL is sufficiently large when the controlled parameter is IC and sufficiently small when the controlled parameter is OOC. Essentially, understanding the ARL characteristics of the control charts can lead to timely detection of changes in the process quality levels, and promptly subsequent appropriate corrective measures taken to return the process into a state of control, if the change is unfavorable.

The ARL is the most frequently used parameter, although the run length distribution is often skewed to the right. Thus, the performance of the chart is measured in term of it ARL (the number of samples or subgroups that need to be collected before the first out-of-control signal is given by a chart) following a step change in process distribution Ghute and Shirke, (2012) and Hawkins and Zamba, (2005b). It is probably used more often in the literature because it is conceptually easy to understand.

Both the in-control and out-control cases will be considered. The in-control robustness information is important in the implementation of the chart, particularly for small to moderate sample sizes. This ARL is also known as the in-control ARL when there is no shift in the process. When there is process shift from its target, the corresponding ARL is called out-of-control ARL. It is desirable that a process monitoring method produces a large ARL when the process is stable at standard values for process parameters and small ARL under OOC condition.

Finding the in-control ARL or out-of-control ARL actually determines the starting point at which practitioners begin to count the number of plotted observations Maravelakis, Panaretos, and Psarakis, (2005). The pre-specified ARL until an alarm, when there is no change in process distribution is usually denoted as ARL_0. In the course of this study, ARL_0 370 and 500 corresponding to a false alarm rate $\alpha = 0.0027$ and $\alpha = 0.002$ respectively would be applied to ease comparison and interpretability.

In the tables 1and 2 the LCP and GLR summary of the IC-ARL for both ARL_0 370 and 500 is presented. Except for a small bit of error, setting start-up observations above 20, the observed IC-ARLs are more or less the same accordingly to the specified ARL (ARL_0). The effect of setting start-up values (20, 25, 30, 40 and 50) at different change-point values is insignificant at the different pre-specified ARL_0 of 370 and 500. See tables below:

Table 1: LCP In-Control ARLs by Varying Start-ups

ARL_0	Start-up	Change-point			
		50	100	250	500
370	20	363.5	360.4	367.0	351.8
	25	363.1	360.7	366.7	352.3
	30	362.5	360.4	365.9	351.9
	40	362.0	360.8	365.8	352.1
	50	361.4	360.3	365.5	352.2
500	20	496.8	498.7	500.3	489.5
	25	496.2	498.9	500.6	490.6
	30	495.7	498.7	500.1	489.8
	40	497.2	499.6	499.9	490.2
	50	496.4	498.5	498.8	490.0

Table1: showing the LCP's In-Control ARLs for start-up values: 20, 25, 30, 40 and 50 at specified $ARL_0 = 370$ and $ARL_0 = 500$, respectively.

Table 2: GLR In-Control ARLs by Varying Start-ups

ARL_0	Start-up	Change-point			
		50	100	250	500
370	20	370.1	355.3	366.4	366.7
	25	370.0	355.3	367.3	366.1
	30	369.8	355.2	367.3	366.1
	40	369.7	355.6	367.4	367.1
	50	369.2	356.1	366.9	367.2
500	20	495.6	485.4	489.8	488.2
	25	495.8	485.3	490.5	487.8
	30	496.4	485.8	490.6	490.2
	40	496.3	486.2	490.8	489.9
	50	496.3	486.5	491.5	489.0

Table 2: showing the GLR's In-Control ARLs for start-up values: 20, 25, 30, 40 and 50 at specified $ARL_0 = 370$ and $ARL_0 = 500$, respectively.

In Table 3, the resulting IC-ARLs for the five distributions considered in this study are presented. The two methods had a rather similar performance with respect to the estimated IC-ARLs, under the normal distribution. Their IC-ARL was 497 (496) and 489 (488) for LCP (GLR) at change-points 50 and 500 respectively. Essentially, this finding will enable comparison of OOC-ARLs of the two methods since both methods have comparable IC-ARLs (under the normal setting). However, the LCP IC-ARL (497) was slightly higher compared to GLRCP IC-ARL (493) at the initial change point 20 while at change-point

Science World Journal Vol. 15(No 3) 2020
www.scienceworldjournal.org
ISSN 1597-6343
Published by Faculty of Science, Kaduna State University
https://doi.org/10.47514/swj/15.03.2020.013

250, the LCP (GLR) IC-ARL was 500 (490).

On the other hand, the LCP had substantially higher IC-ARL compared to the GLRCP when considering the other distributions, as shown in Table 3: The LCP IC-ARLs proportion (with respect to the specified ARL of 500) remains about 100% to 97% and constant for the distributions considered. However, for all the non-normal distributions considered, the GLRCP IC-ARLs are relatively lower and could not in any way attain the specified ARL as expected. The result shows that Laplace and Logistic distributions which are heavy-tailed symmetric distributions respectively are about one-third and two-third of the specified ARL. Also, in the case of Gamma and Lognormal distributions which are skewed distributions, the GLR IC-ARLs are about one-fourth and one-twenty-fifth of the specified ARL respectively. Also, both methods (LCP & GLR) maintain their IC-ARL distributions at the different change-points considered. For example, LCP (GLR) IC-ARLs are 499 (154) and 496 (150) at change-points 20 and 100 respectively, under Laplace distribution. The same pattern of IC-ARLs is exhibited for the other change-points (50, 250 and 500) considered, and for the other distributions (logistic, gamma and lognormal) examined.

Meanwhile, unlike the GLR, there appeared to be a slight reduction in LCP IC-ARLs when the change-point gets larger (i.e. >250) for all the distributions considered. In particular, at change-point 250 (500), the LCP IC-ARLs are 500 (489) and 497 (484) for normal and logistic distributions respectively. This could be attributed to the fact that the method is designed to starts process monitoring as fast as possible, especially for short run processes.

Application of the Methods to Real-Life Data

The performance comparison and assessment of charts is very important, not only because it provides insight into how charts work in practice but also, it's the only way to effectively compare competing methods in a fair and objective manner. As noted by Gan (1995), the comparisons will, in no small measure, allow practitioners to have an understanding of the different competing methods. Besides the fact that the control charts are useful tools for the practitioners, illustrations are highly essential. Hence, both the proposed method and the considered existing methods presented in this study are applied to real-life data obtained from Specialist Hospital Irua, Edo state, Nigeria

Application of the Methods to Children Suffering from Bronchial Pneumonia

There is no doubt that the challenges pose by bronchial pneumonia affect the entire population, but children are most vulnerable because of their unique physiologic, anatomic, morphologic and socio-economic characteristics (W.H.O, 2015). In the same vain, the Millennium Development Goals (MDGs) advanced increased on international attention focusing on child bronchial pneumonia in the developing world, with major aim to reduce under-five child mortality Annim, Awusabo-Asare and Amo-Adjei, (2013). This assertion is characterized by the Nigeria situation particularly in the Savannah and Rain forest region of the country. The report by the World Health Organisation, 2016 resolved that National Development for Health Services (NDHS) in Hyper, Hypo and Meso-endemic area in Africa are charged with the responsibility of collecting data on the bronchial pneumonia of children. NDHS measure the weight of all children under age 5 (60 months) in selected households in Nigeria. The scope of the data used in the computation was based on the reported cases of bronchial pneumonia as obtained from the records section of Irua Specialist Hospital.

3.0 RESULTS

Data Application and Methods

The study considered the existing medical records on the report cases of children who suffered from the bronchial pneumonia disease in Irrua Specialist Hospital. Children with $z_\alpha < -2\sigma$ from the reference population median are considered thin (bronchial pneumonia) or acutely pneumonia.

The performances of conventional change-point charts rely on the normality assumption of process distribution. However, the distribution of process is not only skewed, but also heavy-tailed as indicated using the most commonly Shapiro-Wilk Normality Test. The data set distribution does not follow a normal distribution. Collaborating this, Shapiro-Wilk Normality test statistic: $W = 0.9268$ $(p = 0.001463)$ confirms that there is sufficient evidence to conclude that the data set has not been drawn from a normal population.

Application of GLR-based Chart

Figure 1: below shows a GLR parametric-based change-point control chart on acute bronchial pneumonia of under 5 children; while its estimated change-point is shown in figure 1a.

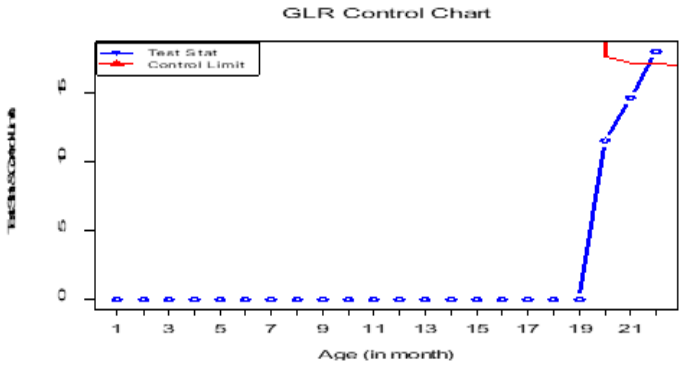

Figure 1: GLR parametric-based control chart on acute bronchial pneumonia under 5 children

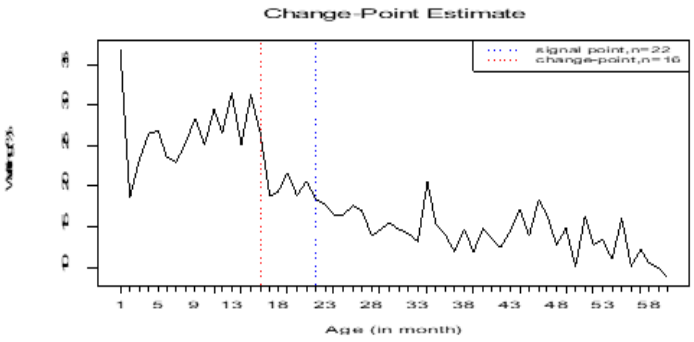

Figure 1a: GLR estimated change-point, along with the detection time (signal point) of acute bronchial pneumonia of less than 5 children

Science World Journal Vol. 15(No 3) 2020
www.scienceworldjournal.org
ISSN 1597-6343
Published by Faculty of Science, Kaduna State University
https://doi.org/10.47514/swj/15.03.2020.013

The inappropriateness of GLR control chart application to non-normal data is shown in figures 1 and 1a. Precisely, a shift in the process data is signaled at about 2 years old (22nd month) and estimated period of change is said to be at 16th month. This estimated point of shift (at the 16th observed month) is unrealistic as actual process monitoring begins at the 20th month.

Application of proposed LCP-based Chart

The figures 2 and 2a below respectively shows LCP nonparametric-based control chart applied to acute bronchial pneumonia of under 5 children

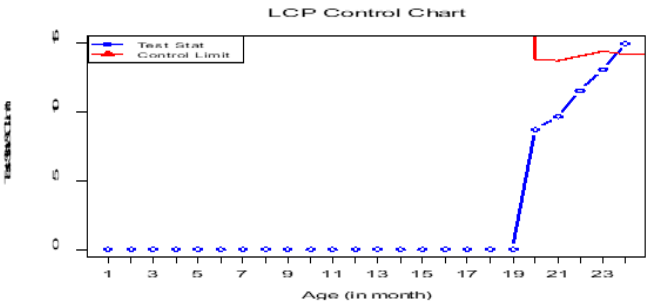

Figure 2: LCP nonparametric-based chart on acute bronchial pneumonia of under-5 children

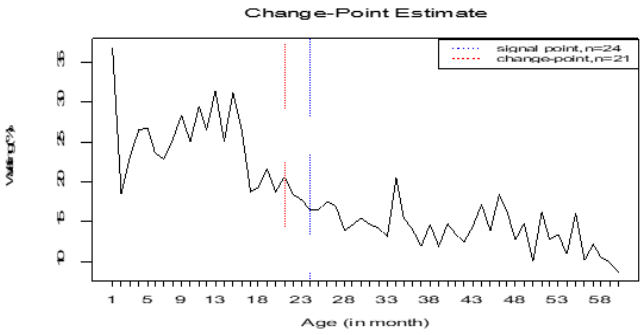

Figure 2a: LCP estimated change-point, along with its detection time of acute of under-5 children bronchial pneumonia.

Figures 2 and (2a) show the results of nonparametric-based LCP control chart to the data. The chart does not only detect the shift in location but also the shift in variability. Figure 2 displays the change-point statistic along with its control limit corresponding to in-control ARL of 500. The estimated change-point, along with the period at which the maximised test statistic (L_{max}) exceeds the control limit, is shown in figure 2a. The chart signals a shift in children's bronchial pneumonia, as measured by acute pneumonia, at about two years old (24th month), while it suggests that the actual change had started at the 21st month (observation). This is an indication of the LCP promptness in raising alarm of a process shift if indeed it exists.

After-Signal Diagnosis

A shift signal by LCP is an indication of either shift in mean, in variance or in both. Hence, it is important for after-signal diagnosis to be carried out on the pre-shift and post-shift data segments.

Table 3: Summary Statistics of Pre-shift and Post-shift Segments of bronchial pneumonia

	Pre-Shift (Segment 1)	Post-Shift (Segment 2)	2-sample test statistic (p-value)
Mean	25.07201	17.33677	W = 76 (0.002949)
Standard deviation	4.747507	0.910598	Z = -2.056 (0.03978)

The summary statistics in Table 3 confirm that the signal may have resulted in both a mean shift (p=0.002949) and a variability shift (p=0.03978) of children's bronchial pneumonia as measured.

4.0 Discussion

The inappropriateness of GLR control chart application to non-normal data is demonstrated with the use of Irrua Specialist Hospital data as applied to under-5 children bronchial pneumonia status, as measured by acute pneumonia diagnosis. The result is said to be unrealistic as the actual commencement of process monitoring was at 20th month. And indeed, this corroborated the stance of Hawkins and Zamba (2005b) claim that their proposed chart may not be suitable for non-normal data. This could be attributed to the fact that there exist joint location and variability shifts in the process data. In agreement with literature of Hawkins and Deng, (2009) and Zhang et al., (2010), however, results from the LCP nonparametric-based chart substantiated NPC and ICT (2014) findings which reported that majority (83%) of children less than 6 months old were not exclusively monitored, and that over 90% of children age 6-23 months were monitored inappropriately (based on recommended infant and young child monitoring practices). Bearing this in mind, application of LCP gave a clear pointer to the policy maker the need to urgently address the failure of children less than 2 years old to receive adequate monitoring. And, if the necessary corrective measures were taken to address this challenge, the tendency of wasting would be reduced to a large extent.

5.0 Conclusion and Recommendation

In conclusion, there is a clear indication that LCP performs quite better compared with the competing chart considered. It should therefore be the preferred method of choice in the face of unknown distribution of process data.

REFERENCES

Annim, S. K., Awusabo-Asare, K. and Amo-Adjei, J. (2013). Household nucleation, dependency and child health outcomes in Ghana. *DHS Working Papers No. 98,* retrieved from http://www.dhsprogram.com/publications/publication-wp98-working-papers.cfm.

Dong, Y., Hedayat, A. S. and Sinha, B. K. (2008). Surveillance strategies for detecting changepoint in incidence rate based on Exponentially Weighted Moving Average Methods. *Journal of the American Statistical Association*, Vol. 103, No. 482, pp.843- 853.

Deng, Q. (2009). A nonparametric change-point model for phase II analysis. Unpublished PhD thesis. University of Minnesota, pp. 185

Flowers, J. (2009). Statistical process control methods in public health intelligence. Technical Briefing 2, Association of Public Health Observatories [APHO], retrieved from http://www.apho.org.uk.

Gan, F. F. (1995). Joint monitoring of process mean and variance using exponentially weighted moving average control charts. *Technometrics*, Vol. 37, No. 4, pp.446 - 453.

Ghute, V. B. and Shirke, D. T. (2012). A nonparametric signed-rank control chart for bivariate process location. *Quality Technology & Quantitative Management*, Vol. 9, No. 4, pp.317-328.

Hawkins, D. M. and Zamba, K. D. (2005a). A change-point model for a shift in variance. *Journal of Quality Technology*, Vol. 37, No. 1, pp.21-31.

Hawkins, D. M. and Zamba, K. D. (2005b). Statistical process control for shifts in mean or variance using a change-point formulation. *Technometrics*, Vol. 47, No. 2, pp.164-173.

Hawkins, D. M., Qiu, P. and Kang, C. W. (2003). The changepoint model for statistical process control. *Journal of Quality Technology*, Vol. 35, No. 4, pp.355-366.

Hutchinson P. T. (2002). Technical Note: Should we routinely test for simultaneous location and scale changes? *Ergonomics,* Vol. 45, No. 3, pp.248-251.

Lapage, Y. (1971). A combination of Wilcoxon's and Ansari-Bradley's statistics. *Biometrica*, Vol. 58, No. 1, pp.213–217.

Li, Z., Qiu, P., Chatterjee, S., and Wang, Z. (2013). Using p-values to design statistical process control charts. *Statistical Papers*, Vol.54, No. 2, pp.523–539.

Maravelakis, P.E., Panaretos, J. and Psarakis, S. (2005). An examination of the robustness to nonnormality of the EWMA control charts of the dispersion. *Communications in Statistics –Simulation and Computation*, Vol. 34, No. 4, pp.1069–1079.

McCracken, A. K. and Chakraborti, S. (2013). Control chart for joint monitoring of mean and variance: an overview. *Quality Technology and Quantitative Management,* Vol. 10, No. 1, pp.17-36.

Razmy, A.M. and Peiri, T.S.G. (2013). Performance Comparison of Shewart joint Monitoring Scheme for Mean and Variance. *National Engineering Conference 2013, 19thEru Symposium. Faculty of Engineering. University of Maratuwa, Sri-linka. Pp. 99 – 104.*

Ross, G.J., Tasoulis, D.K. and Adams, N.M. (2011). Nonparametric monitoring of data streams for changes in location and scale. *Technometrics,* Vol. 53, No.4, pp.379-389.

Rublik, F. (2005). The multisample version of the Lepage test. *Kybernetika*, Vol. 41, No. 6, pp.713–733.

Salisu, S.U., Edokpa, O.S., and Shaib, I.O (2018) Evaluation of Parametric and Nonparametric Based Change-Point Control Chart. *Journal of the Nigerian Association of Mathematical Physics. Vol. 46 No1*

Snoussi, A. and Limam, M. (2007). The change point model: SPC method for short run autocorrelated data. *Quality Technology and Quantitative Management,* Vol. 4, No. 3, pp. 313-329.

Zamba, K. D. and Hawkins, D. M. (2006). A multivariate change-point model for statistical process control. *Technometrics,* Vol. 48, No. 4, pp.539–549.

Zhang, J., Zou, C. and Wang, Z. (2011). A new chart for detecting the process mean and variability. *Communications in Statistics-Simulation and Computation*, Vol. 40, No. 5, pp.728–743. doi: 10.1080/03610918.2011.552823.

Zhou, C., Zou, C., Zhang, Y., and Wang, Z. (2009). Nonparametric control chart based on change-point model. *Statistical papers*, Vol. 50, No. 1, pp.13-28.