Thailand Statistician
October 2023; 21(4): 824-838
http://statassoc.or.th
Contributed paper

Profile Monitoring of Residual Control Charts under Gamma Additive Models

Engy S. Abozaid, Shereen H. Abdel Latif* and Salah M. Mohamed
Department of Applied Statistics and Econometrics, Faculty of Graduate studies for Statistical Research, Cairo University, Giza, Egypt.
*Corresponding author; e-mail: shereen_hamdy_@cu.edu.eg

Received: 27 April 2021
Revised: 29 October 2021
Accepted: 29 December 2021

Abstract

One of the most powerful tools in quality control is the statistical control chart. In some process control applications, the quality of a product or process can be characterized by a relationship between two or more variables, which is typically referred to as a profile. There are many techniques in literature for monitoring profiles. Generalized additive models (GMAs) have been used frequently in many different applications for modeling non-linear effects in regression models with non-Gaussian response Hastie and Tibshirani (1987). The scheme is based on the deviance residuals, and Pearson residuals for detecting any disturbance in the control variables. This paper investigates the performance of exponentially weighted moving average (EWMA) control charts to monitor the response variability for gamma additive regression models. The simulation study shows that using deviance residuals under log link function seems more suitable than Pearson residuals.

Keywords: Average run length, EWMA control charts, gamma additive residuals, generalized additive models.

1. Introduction

Statistical control chart is the most powerful tool in statistical process control (SPC), Shewhart 1920 presented the idea of graphical summary of information for monitoring of the production process. In quality control to describe the shape of the response data regardless of whether there is a set of explanatory variables or not the word “profile” is usually used. The term of “profile” refers to the relationship between the dependent and one or more independent variables and the main objective of profile monitoring is to check the stability of the functional relationship between the response and the explanatory variables Mahmoud (2011, 2012).

GAMs are semi-parametric extensions of generalized linear models (GLMs), the only underlying assumption made is that the functions are additive and that the components are smooth Hastie and Tibshirani (1987). GAMs like GLMs, use a link function to establish a relationship between the mean of the response variable and a ‘smoothed’ function of the explanatory variables. The strength of the GAMs is their ability to deal with highly non-linear and non-monotonic relationships between the response and the set of explanatory variables. In addition, the data determine the nature of the relationship between the response and the set of explanatory variables rather than assuming some form of parametric relationship as in GLMs Yee and Mitchell (1991).

The EWMA chart methodology was first introduced by Roberts (1959) who used simulation and evaluated its properties, purposed it for use by some organizations, particularly in the industrial process, financial and management control as the basis of new control performance charts system. Also, Roberts showed that the EWMA is an effective alternative to the traditional Shewhart control chart when small shifts in the process parameters are monitored. Hunter (1986) pointed out that the EWMA can be thought of as a compromise between the Shewhart *X*-bar and CUSUM charting procedures. For smoothing constant $(\lambda = 1)$, the EWMA places all its weight on the most recent observation, as does the *X*-bar chart. For λ close to zero, the most recent observation receives little weight, and the EWMA resembles the CUSUM. The choice of λ between 0 and 1 determines how much weight the most recent observation will receive Crowder (1987) (See Lucas and Saccucci (1990)) presented the EWMA control chart, as good choice to detect small change in process average. Various modification and supplemental criteria have also suggested that control charts based on moving average are also effective charts in detecting small process change Lucas and Saccucci (1990). Woodall et al. (2004) stated that the Shewhart charts are slow in detecting small shifts as they use only the information contained in the last sampling point. Knoth (2007) studied the monitoring of EWMA control charts to obtain accurate average run length (ARL) values of EWMA charts by using standard methods like the Markov chain approach, which provide quick and nearly accurate results, but do not attain high accuracy over time.

Many studies have been done by researchers for monitoring different types of profiles such as: gamma profiles, Poisson profile and so on. Braimah et al. (2014) studied the performance of the EWMA control charts that monitoring road crashes and compared the charts performance by using different values of smoothing parameter (λ) to detecting small shifts in the process mean. Maravelakis et al. (2019) stated that the control chart is the main technique to monitor a process. Shewhart control charts are used to detect large shifts in a process, whereas EWMA charts are more efficient to detect small to moderate shifts in a process. Khan et al. (2018) defined an EWMA chart, it is time weighted, and it is a powerful tool for detecting small shifts in a parameter of a process more rapidly than a Shewhart chart with an equal sample size.

According to the residual EWMA control chart that is used to monitor the mean of the residuals to detect small and moderate shifts that occurs in the mean, Lu and Reynolds (1999) suggested EWMA control chart based on the residuals from the forecast values of the model using an integral equation method. Messaoud et al. (2008) mentioned the idea behind residual control charts is that if the time series model fits the data well, the residuals will be approximately independent. Then, traditional control charts designed to monitor independent observations can be applied to the residuals. Areepong (2013) compared the effectiveness of the Shewhart *X*-bar, EWMA, and geometric moving average control chart (GMA) residual control charts for auto-correlation observations with upward and downward linear trend, the comparison of the control charts was based on the ARL. They showed that the performance of EWMA charts is superior to the Shewhart X-bar and GMA residual charts for small shifts; however, the performance of Shewhart X-bar residual chart is superior to EWMA and GMA residual charts for large shifts.

The purpose of this paper is to fit gamma additive model by using the spline smoothing (SS) as a nonparametric regression estimation method in case of two link functions; identity and log link functions respectively, extracting two types of residuals; Pearson residuals, and deviance residuals, then monitoring these residuals, that is, (the differences between the reference line and the sample profiles) by using X-bar chart/S-chart in phase I analysis, and using EWMA chart in phase II analysis. Finally, make a comparison between the performances of residuals at EWMA charts by using ARL measure, where the run length is the number of samples taken before a sample fall outside the control limits.

2. Generalized Additive Models

An important statistical development of the last 30 years has been the advance in regression analysis provided by GAMs. The introduction of models that automatically identify appropriate transformations was a second important step forward in regression analyses. This led to a wider generalization of GLMs known as GAMs, Hastie and Tibshirani (1990). One can envision the different regression models as being nested within each other, with simple and multiple linear regression (SLR and MLR) being the two most limiting cases, and GAMs the most general: SLR $\subset$ MLR $\subset$ GLMs $\subset$ GAM $\subset$ GAMs are parameterized just like GLMs, except that some predictors can be modeled non-parametrically in addition to linear and polynomial terms for other predictors. The probability distribution of the response variable must still be specified, and in this respect, they are more aptly named semi-parametric models. A crucial step in applying GAMs is to select the appropriate level of the 'smoother' for a predictor. This is best achieved by determining the level of smoothing using the concept of effective degrees of freedom. A reasonable balance must be maintained between the total number of observations and the total number of degrees of freedom used when fitting the model. One of the considered methods was a generalized additive model GAM. A GAM is a non-parametric, regression technique not restricted by linear relationships, and it is flexible regarding the statistical distribution of the data (Swartzman et al. 1995).

3. Gamma Additive Regression Models

A class of models for gamma distributed random variables is presented. These models are shown to be more flexible than the classical linear models with respect to the structure that can be imposed on the expected value. Both additive, multiplicative, and combined additive-multiplicative models can be formulated. The gamma regression model is called gamma additive regression model under the conditions.

1. The response y is distributed as gamma distribution.
2. The relation between the response and the predictors is nonlinear.
3. The link function is one of the following forms: identity link $(\eta = \mu)$, log link $(\eta = \log(\mu))$, and inverse link $(\eta = 1/\mu^2)$.

As an example of the additive gamma model, Hastie and Tibshirani (1990) reanalyzed data on the clotting times of blood. The response was clotting time in seconds for normal plasma diluted to nine different percentage concentrations with prothrombin-free plasma for different lots of thromboplastins. They fit an additive gamma model with inverse (canonical) link and used smoothing spline with degrees of freedom equal two for the smooth term.

The gamma distribution will be proposed as a general model for the analysis of response times. It will be shown that this probability density function (PDF) has a larger scope than the classical linear models. In particular, the parameters of additive, multiplicative, and combined additive-multiplicative models can be estimated without analytical difficulties. Moreover, this PDF exhibits certain established characteristics of response time distributions.

$$f\left(y|\theta,\psi\right) = \frac{\psi^{\theta} y^{\theta-1} e^{-\psi y}}{\Gamma(\theta)} I_{(0,\infty)}\left(y\right), \tag{1}$$

where $\theta,\psi > 0, \Gamma(\cdot)$ denotes the gamma function, and $I(\cdot)$ is the indicator function. Under this parameterization, the mean and variance of y are given by $E(y) = \theta / \psi$ and $var(y) = \theta / \psi^2 = \mu^2 / \theta,$ given that $\psi = \theta / \mu.$

4. Exponentially Weighted Moving Average Control Charts

One of the most famous quality control charts is EWMA control chart which introduced by Roberts (1959). EWMA is an effective alternative to the traditional Shewhart control chart when small shifts in the process parameters are monitored. The EWMA control chart has received much importance during the last few years, in phase II analysis, the application of the EWMA technique was made simple as the Shewhart control chart.

4.1. EWMA control chart formulas

Following Crowder (1987) (See Lucas and Saccucci (1990)) the EWMA statistic is defined as

$$z_i = \lambda x_i + (1-\lambda) z_{i-1}. \quad (2)$$

The value of z_0 is set to the target mean μ_0. Sometimes the average $\overline{x}$ is used as the starting value of the EWMA. The value of the constant λ is specified by the user where $0 \le \lambda \le 1$. But for commonly used the value of λ is between $0.1 \le \lambda \le 0.25,$ because these values make the chart more sensitive in detecting small and median changes.

4.2. Estimating the EWMA chart center line

$$\mu_0 = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^{k} n_i}. \quad (3)$$

If the subgroups are of equal size, the above equation for the CL reduces to

$$\mu_0 = \frac{\sum_{i=1}^{k} \overline{x}_i}{k} = \frac{\overline{x}_1 + \overline{x}_2 + \ldots + \overline{x}_k}{k},$$

the lower control limit (LCL) and upper control limit (UCL) for the EWMA chart are calculated using these formulas

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]},$$

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]}. \quad (4)$$

After we plot on the chart several observations or, expressing it in a different way if some time passes, the term $\left[1-(1-\lambda)^{2i}\right]$ approaches unity as i becomes larger. Thus, the asymptotic limits of EWMA chart are as follow

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda}},$$

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda}}, \quad (5)$$

where L is the width of the control limits (often set to 2.7 or 3), and λ is the exponential smoothing constant. It is possible to choose these parameters $(L,\ \lambda)$ to give ARL performance for the EWMA control chart that closely approximates CUSUM ARL performance for detecting small shifts or the ARL of Shewhart performance. And as mentioned before, the criteria of performance which is used in this paper is the ARL, and it is the expected number of samples until a control chart signals, given the process is out of control. It is desirable for ARL to be as small as possible.

5. Gamma Regression Residuals

Most residuals are based on the differences between the observed responses and fitted conditional mean. Gipe (1976) mentioned that the analysis of residuals plays a major role in predictive model formation. When numerical-valued functions are fitted to sample data, all the information about lack of fit is contained in the residuals which can be used to provide necessary feedback on the modeling process. These residuals from a fitted model can be plotted to help detect unequal variances and relationships over time.

In a sense of modeling structure, once it became feasible to compute different kinds of residual in a straightforward way, various methods have focused on their underlying properties and their effectiveness. Residuals may be employed as a measure for model selection, play a significant role in model diagnostics and assessment of its fitting and used for testing the validity of the assumptions of statistical models. For example, residuals are used to verify homoscedasticity, linearity of effects, normality, and independence of the error. In classical linear models, residuals are usually standardized so that they become scale free and have the same precision, and this makes it more convenient to compare residuals at various locations in the region of experimentation, Shao and Lin (2013). Some types of residuals such as classical residuals, Pearson residuals, and deviance residuals will be presented in the following subsections.

5.1. Pearson residuals

Residuals in GLMs were first discussed by Pregibon (1981), though ostensibly concerned with logistic regression models, then McCullagh and Nelder (1989) provided a survey of GLMs with substantial attention to definition of residuals. Pearson residuals are the most used measures of overall fit for GLMs, it can be used to check the model for each observation, Cordeiro and Simas (2009) defined Pearson residuals in GRM as follows

$$r_i^p = \frac{y_i - \hat{\mu}_i}{\sqrt{Var(\hat{\mu}_i)}} = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i(n_i - \hat{\mu}_i)/n_i}},\ i = 1,2,\ldots,n, \tag{6}$$

where $\hat{\mu}_i$ is the fitted mean value and the denominator follows from the fact that $y_i = n_i\pi_i(1-\pi_i)$ are the fitted variance function of y_i.

5.2. Deviance residuals

Cuervo et al. (2016) defined alternative residual was based on the deviance or likelihood ratio, which for GRM is given by

$$r_i^d = -2\sum_{i=1}^{n} \log\left(\frac{y_i}{\hat{\mu}_i}\right) - \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}, \tag{7}$$

where $\hat{\mu}_i = g^{-1}\left(x_i'\hat{\beta}\right)$.

Two statistics used to assess the goodness of fit of the GLMs are the deviance and the Pearson chi squared statistic χ^2. Linear models use raw residuals for testing the model diagnostics, whereas the GLMs provides several structures for residuals such as the Pearson, deviance, likelihood, and working. The two common types of residuals are then the components of these statistics. Other definitions for residuals in GLMs have been proposed in Pierce and Schafer (1986).

6. Simulation Study

This simulation was implemented to evaluate the performance of two residuals control charts: deviance residuals and Pearson residuals for two gamma additive regression models using spline smoothing nonparametric estimation method.

6.1. Generating the data

The observations X_i, $i = 1,2,3$ were generated such that values of X_1 were generated from a uniform distribution on the interval (0, 30); values of X_2 were generated from a uniform distribution on the interval (0, 15); and values of X_3 were generated from a uniform distribution on the interval (10, 20). Gamma regression model was built. This model has two different link functions; identity link function and log link function, Cuervo et al. (2016). Different sample sizes (n) were chosen such that; 5 and 10 and different number of samples (m) were chosen such that; 25 and 30. The following are the steps with more details.

i. The gamma regression model with mean and shape structures given by (8) and (9), respectively, is considered

$$\mu_i = x_i'\theta, \tag{8}$$

$$\log(\theta_i) = z_i'y, \tag{9}$$

where (8) describes the variable of interest Y and the regressors X of the mean structure and (9) describes the regressors Z for the shape structure, with log-link function.

ii. The values Y_i were generated from a gamma distribution with mean and shape parameters given by

$$\mu_i = 15 + 2x_{2i} + 3x_{3i} \text{ and } \theta_i = \exp\left(0.2 + 0.1x_{2i} + 0.3x_{3i}\right)$$

Second Step: Mean structure function with log link function.

iii. The gamma regression model with mean and shape structures given by (10) and (9), respectively, is considered:

$$\log(\mu_i) = x_i'\theta. \tag{10}$$

iv. The values Y_i were generated from a gamma distribution with mean and shape parameters given by

$$\mu_i = \exp\left(-5 + 0.2x_{2i} - 0.03x_{3i}\right) \text{ and } \theta_i = \exp\left(0.2 + 0.1x_{2i} + 0.3x_{3i}\right).$$

6.2. Design of the EWMA control chart (Phase II analysis)

To construct an EWMA chart suppose that l and λ, the multiple of σ_{zi} used in the control limits and the weighting or smoothing constant are the design parameters respectively. Steps to conduct the EWMA chart:

a. Compute the values of parameters form historical data through *X*-bar chart / *S*-chart such as the target mean or process mean (μ_0) and the estimated sigma $(\hat{\sigma})$ by using the S -chart method.

b. Calculate the EWMA statistic (z_i) and calculate the control limits (UCL, LCL) with the combination of $(l = 2.962,\ \lambda = 0.2)$ for each sample (i) (Lucas and Saccucci 1990).

c. Plot the statistic (z_i) for each sample, with the control limits.

d. Declare the process to be in-control (IC) if $LCL_{(i)} \le z_i \le UCL_{(i)}$; otherwise, declare the process to be out-of-control (OC).

e. If the process is declared OC, then count the number of subgroups as the run length, i.e., the process remains IC before it is declared to be OC.

f. Compute the ARL for EWMA control charts.

6.3. EWMA control charts and its performance results

This section is divided into two subsections; for two different sample sizes; sample size equals five, and sample size equals ten. Each subsection is also divided into two subsections according to the number of samples. More details of these two different numbers of samples will be presented below. But before this the Tables 1 and 2, which contain the results of ARL, and the relative ARL (RARL) respectively as measures for the performance of the chart will be presented.

Table 1 Average run length results

Sample size n	Number of Samples m	Identity Link Function		Log link Function	
		Deviance Residuals	Pearson Residuals	Deviance Residuals	Pearson Residuals
$n = 5$	25	495.7451	495.7933	495.7655	495.7935
	30	495.7704	495.7917	495.7650	495.7935
$n = 10$	25	495.7705	495.7933	495.7693	495.7933
	30	495.7745	495.7931	495.7572	495.7933

To facilitate the comparison between different link functions of the ARL, we use the formula.

$$RARL = \frac{\text{Proposed estimate value}}{\text{Deviance residuals with identity L. F.}},$$

where L. F. indicates link function. (i.e., RARL value for Pearson residuals using identity link function with $n = 5$ and $m = 25$ equals 495.7933 divided by 495.7451 which finally equals 1.000097).

The following Table 2 shows the relative performance (RARL) for all ID, LD, IP, and LP values relative to ID value, as a measure for the performance of the chart.

Table 2 Relative average run length results

Sample size n	Number of Samples m	Identity Link Function		Log link Function	
		Deviance Residuals	Pearson Residuals	Deviance Residuals	Pearson Residuals
$n = 5$	25	1	1.000097	1.000041	1.000098
	30	1	1.000043	0.999989	1.000047
$n = 10$	25	1	1.000046	0.999998	1.000046
	30	1	1.000038	0.999965	1.000038

6.3.1. Sample size equals five

This section is also divided into two subsections according to the number of samples which equals twenty-five or thirty. More details about the descriptive for EWMA control charts are shown below.

i. Number of samples equals twenty-five

The following charts from Figures 1 to 4 show the EWMA values for the 25 samples for deviance residuals with identity link function (ID) and with log link function (ID), and for Pearson residuals with identity link function (IP) and with log link function (LP) respectively.

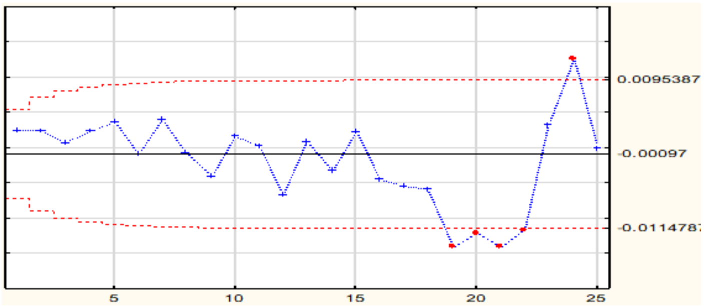

Figure 1 EWMA control chart of ID with $n = 5$ and $m = 25$

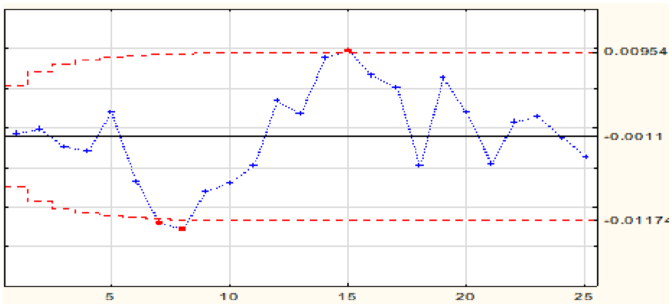

Figure 2 EWMA control chart of LD with $n = 5$ and $m = 25$

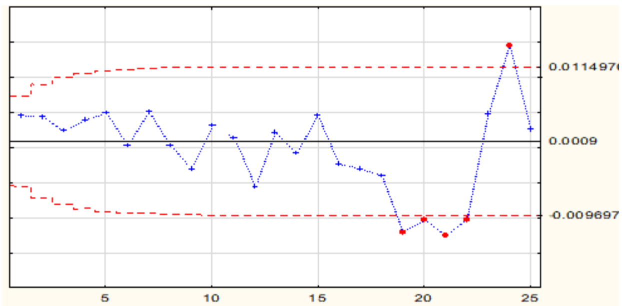

Figure 3 EWMA control chart of IP with $n = 5$ and $m = 25$

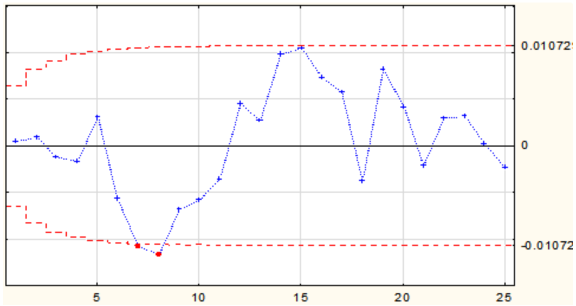

Figure 4 EWMA control chart of LP with $n = 5$ and $m = 25$

Figures 1 and 3, using identity link function, show that 16 % of all samples OC (lower) and 4% OC (upper). But with log link function, Figure 2 shows that 8% of all samples OC (lower) and 4% OC (upper) and Figure 4 shows that 8% OC (lower) of all samples.

Therefore, and from Table 2, the values of RARL indicates that the performance of identity link function with deviance residuals (i.e., equals one) is less than RARL for Pearson residuals (i.e., equals 1.000097) which is as same as the RARL values under log link function.

ii. Number of samples equals thirty

The following charts from Figures 5 to 8 show the EWMA values for the 30 samples for ID, LD, IP, and LP, respectively.

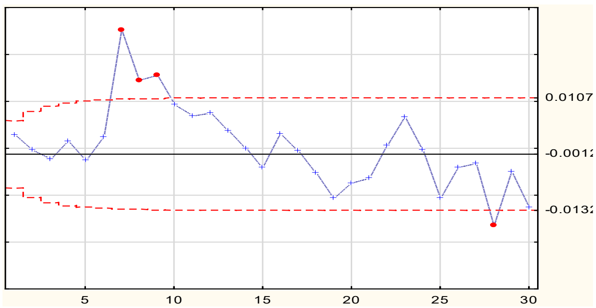

Figure 5 EWMA control chart of ID with $n = 5$ and $m = 30$

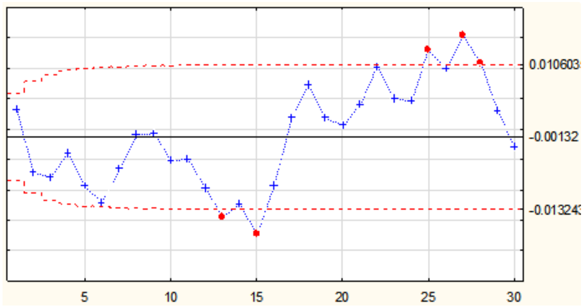

Figure 6 EWMA control chart of LD with $n = 5$ and $m = 30$

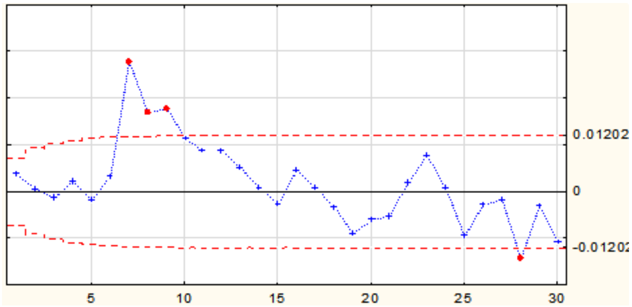

Figure 7 EWMA control chart of IP with $n = 5$ and $m = 30$

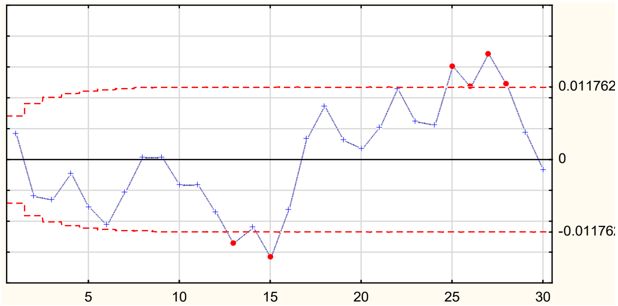

Figure 8 EWMA control chart of LP with $n = 5$ and $m = 30$

Using identity link function, Figures 5 and 7 show that 3% OC (lower) and 10% OC (upper) of all samples. But using log link function, Figure 6 shows that 7% OC (lower) and 10% OC (upper) of all samples and Figure 8 shows that 7% OC (lower) and 13% OC (upper) of all samples.

And this result was consistent with the values of RARL, Table 2, so, the values of RARL indicates that the performance of deviance residuals with identity link function [i.e., equals one] is less than RARL for Pearson residuals (i.e., equals 1.000043) and the same result was achieved with the RARL values using log link function.

6.3.2. Sample size equals ten

Here sample size equal ten and as it was mentioned before there are two different numbers of samples: twenty-five, and thirty. More details with these two different numbers of samples will be presented below.

i. Number of samples equals twenty-five

The following charts from Figures 9 to 12 show the EWMA values for the 25 samples for ID, LD, IP, and LP, respectively.

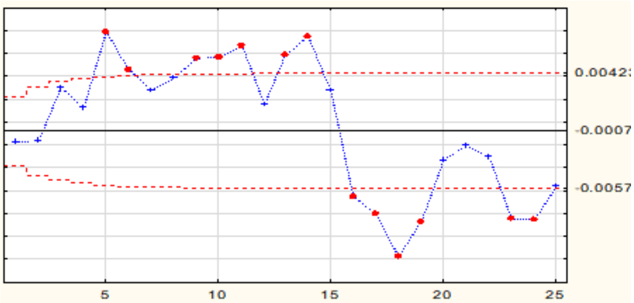

Figure 9 EWMA control chart of ID with $n = 10$ and $m = 25$

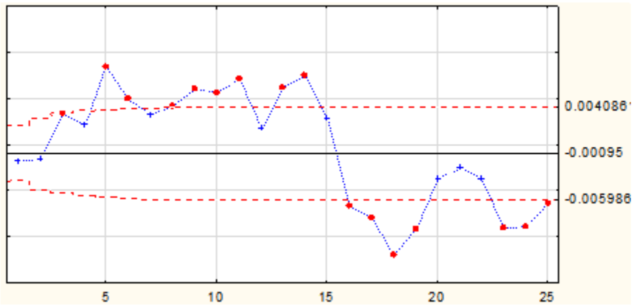

Figure 10 EWMA control chart of LD with $n = 10$ and $m = 25$

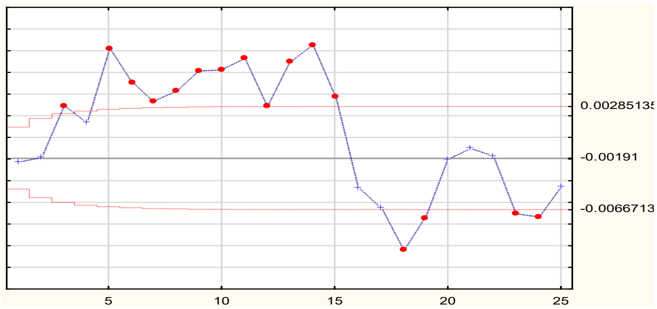

Figure 11 EWMA control chart of IP with $n = 10$ and $m = 25$

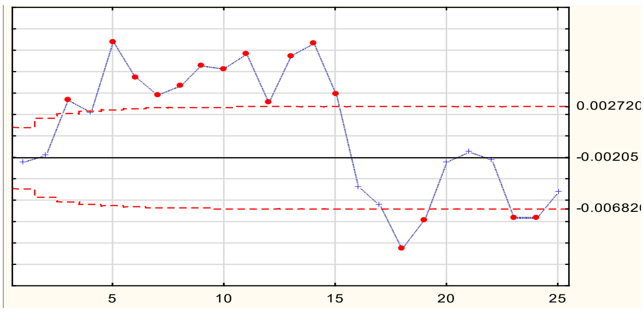

Figure 12 EWMA control chart of LP with $n = 10$ and $m = 25$

Using identity link function, Figure 9 shows that 24% OC (lower) and 28% OC (upper) of all samples and Figure 11 shows that 16% OC (lower) and 48% OC (upper) of all samples. But, using log link function, Figure 10 shows that 28% OC (lower) and 36% OC (upper) of all samples and Figure 12 shows that 16% OC (lower) and 48% OC (upper) of all samples.

On depending on the values of RARL, Table 2 show that the values of RARL indicates that the performance of deviance residuals with identity link function (i.e. equals one) is less than RARL of Pearson residuals (i.e., equals 1.000046) and the same result was achieved with the RARL values using log link function, the RARL values of deviance and Pearson residuals are 0.999998 and 1.000046, respectively.

ii. Number of samples equals thirty

The following charts from Figures 13 to 16 show the EWMA values for the 30 samples for ID, LD, IP, and LP, respectively.

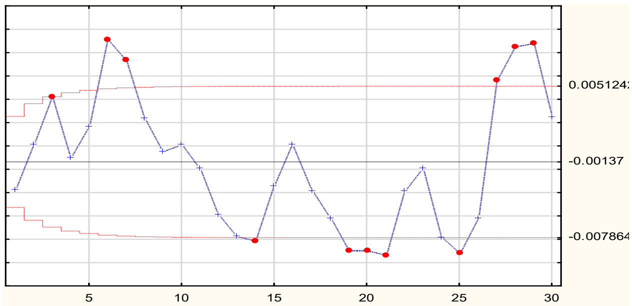

Figure 13 EWMA control chart of ID with $n = 10$ and $m = 30$

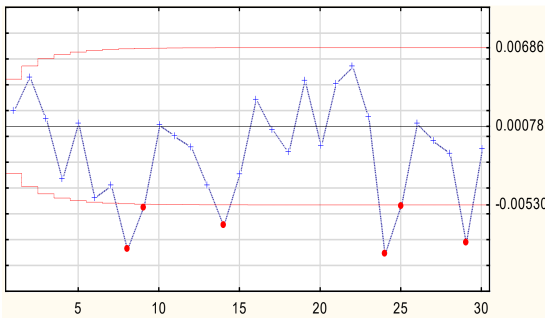

Figure 14 EWMA control chart of LD with $n = 10$ and $m = 30$

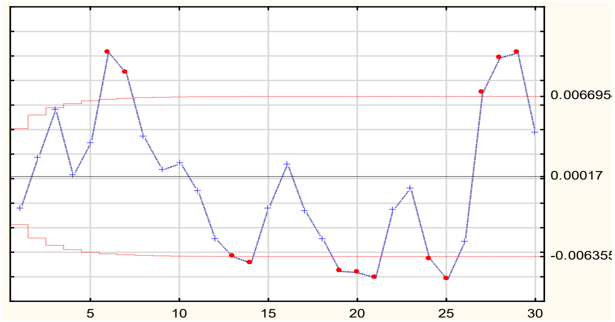

Figure 15 EWMA control chart of IP with $n = 10$ and $m = 30$

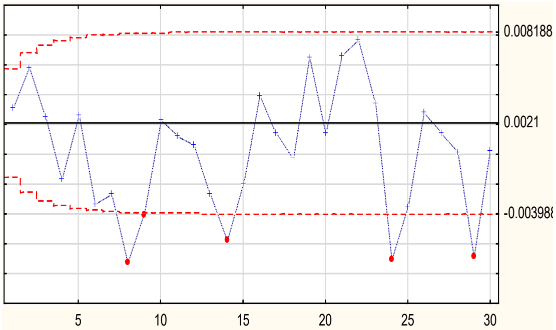

Figure 16 EWMA control chart of LP with $n = 10$ and $m = 30$

Using identity link function, Figure 13 shows that 17% OC (lower) and 17% OC (upper) of all samples and Figure 15 shows that 23% OC (lower) and 17% OC (upper) of all samples. But, using log link function, Figure 14 shows that 20% OC (lower) of all samples and Figure 16 shows that 17% OC (lower) of all samples.

In addition, the values of RARL, Table 2 show that the performance of deviance residuals with identity link function (i.e., RARL= one) is less than the RALR value of Pearson residuals (i.e., RARL= 1.000038) and the same result was achieved with the RARL values using log link function, the RARL values of deviance and Pearson residuals are 0.999965 and 1.000038 respectively.

And finally, from Figures 17 and 18 it can be concluded that, with $m = 30$ there were the best values for EWMA of deviance residuals using log link function.

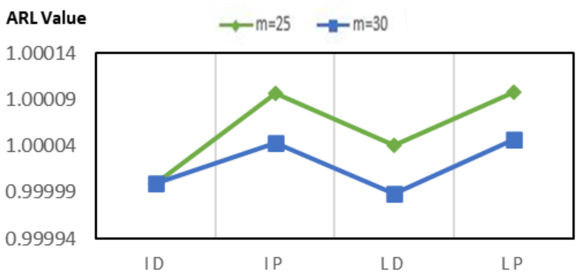

Figure 17 RARL values with $n = 5$

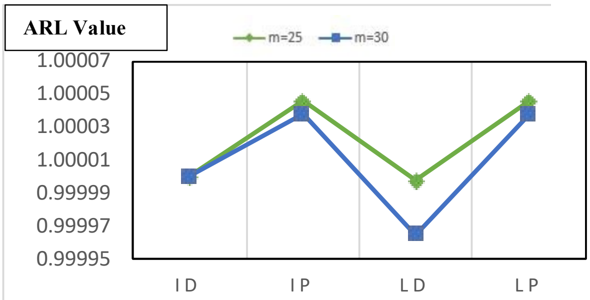

Figure 18 RARL values with $n = 10$

Also, using identity and log link functions, the values of RARL of deviance residuals are usually less than the RARL values of Pearson residuals.

6.4. General conclusion of EWMA control charts

Using identity function, EWMA results of deviance residuals were identical with EWMA results of Pearson in approximately 50% of cases. With the other 50% of cases number, OC samples increases in most cases of deviance residuals. But, by using log function, EWMA results of deviance residuals were identical with EWMA results of Pearson in approximately 25% of cases. With the other 75% of cases number of OC samples increases in most cases using deviance residuals.

Therefore, with all sample sizes, the identity link function gave OC samples more than log link function. And, with number of samples equals 25, and by increasing the sample size, the percentages of OC samples increased which is matching with the theoretical, using identity link function. This was not achieved with a sample count of 30, which was not expected.

Therefore, with the identity and log link functions-with any number of samples-and by increasing the sample size, the RARL values of deviance residuals are smaller than the RARL of Pearson residuals. And the values of RARL using log link function of deviance residuals are the smallest values in comparison with the other deviance residuals using identity link function and also the other Pearson residuals.

7. Recommendations and Future Research

We recommend that using deviance residuals with log link function seems more suitable than Pearson residuals. Further applications to the monitoring different types of residuals are a subject of ongoing research, for example to:

i. Estimate and assess the ARL for different generalized additive models to evaluate the accuracy of control charts in phase II analysis.

ii. Consider other residuals such as standardized residuals and working residuals.

iii. Consider other forms of the link function such as inverse link function.

iv. Using different measures of performance to evaluate control charts such as ATS and ANSS.

v. Consider other methods for fitting the models such as Bayesian approach and weighted least square method.

References

Areepong Y. Comparison of performance of residual control charts for trend stationary AR (p) processes. Int J Pure Appl Math. 2013; 85(3): 583-592.

Braimah OJ, Osanaiye PA, Omaku PE, Saheed YK, Eshimokhai SA. On the use of exponentially weighted moving average (Ewma) control chart in monitoring road traffic crashes. Int J Math Stat Invent. 2014; 2(5): 1-9.

Cordeiro GM, Simas AB. The distribution of Pearson residuals in generalized linear models. Comput Stat Data Anal. 2009; 53(9): 3397-3411.

Cuerv EC, Corrales M, Cifuentes MV, Zarate H. On gamma regression residuals. J Iran Stat Soc 2016; 15(1): 29-44.

Gipe GW. Using residual analysis to search for specification errors. Decis Sci. 1976; 7(1): 40-56.

Hastie T, Tibshirani R. Generalized additive models: some applications. J Am Stat Assoc. 1987; 82(398): 371-386.

Hastie TJ, Tibshirani RJ. Generalized additive models. Chapman and Hall/CRC; 1990.

Hunter JS. The exponentially weighted moving average. J Qual Technol. 1986; 18(4): 203-210.

Khan N, Yasmin T, Aslam M, Jun CH. On the performance of modified EWMA charts using resampling schemes. Oper Res Decis. 2018; 28(3): 29-43.

Knoth S. Accurate ARL calculation for EWMA control charts monitoring normal mean and variance simultaneously. Sequential Anal. 2007; 26(3): 251-263.

Lu CW, Reynolds Jr MR. EWMA control charts for monitoring the mean of autocorrelated processes. J Qual Technol. 1999; 31(2): 166-188.

Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. Technometrics. 1990; 32(1): 1-12.

Mahmoud MA. Simple linear profiles. In Noorossana R, Saghaei A, Amiri A (Eds.). Statistical Analysis of Profile Monitoring (ed., pp. 21-92). John Wiley & Sons; 2011.

Mahmoud MA. The performance of phase II simple linear profile approaches when parameters are estimated. Commun Stat Simul 2012; 41(10): 1816-1833.

Maravelakis PE, Castagliola P, Khoo MBC. Run length properties of run rules EWMA chart using integral equations. Qual Technol Quant Manag. 2019; 16(2): 129-139.

McCullagh P, Nelder JA. Generalized linear models. Chapman and Hall/CRC; 1989.

Messaoud A, Theis W, Hering F, Weihs C. Monitoring a drilling process using residual control charts. Qual Eng. 2008; 21(1): 1-9.

Pierce DA, Schafer DW. Residuals in generalized linear models. J Am Stat Assoc. 1986; 81(396): 977-986.

Pregibon D. Logistic regression diagnostics. Ann Statist. 1981; 9(4): 705-724.

Roberts SW. Control chart tests based on geometric moving averages. Technometrics. 1959; 239-250.

Shao YE, Lin Y. Applying residual control charts to identify the false alarms in a TFT-LCD manufacturing process. Appl Math Inf Sci. 2013; 7(4):1459-1464.

Swartzman G, Silverman E, Williamson N. Relating trends in Walleye Pollock (Theragra chalcogramma) abundance in the Bering Sea to environmental factors. Can J Fish Aquat Sci. 1995; 52(2): 369-380.

Woodall WH, Spitzner DJ, Montgomery DC, Gupta S. Using control charts to monitor process and product quality profiles. J Qual Technol. 2004; 36(3): 309-320.

Yee TW, Mitchell ND. Generalized additive models in plant ecology. J Veg Sci. 1991; 2(5): 587-602.