Scientia Iranica E (2021) 28(1), 557–571

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
http://scientiairanica.sharif.edu

Research Note

Phase II monitoring of generalized linear profiles under different types of changes

S. Hajifar* **and H. Mahlooji**

Department of Industrial Engineering, Sharif University of Technology, Tehran, P.O. Box 11155-9414, Iran.

Received 18 April 2018; received in revised form 28 December 2018; accepted 2 June 2019

KEYWORDS
Change-point;
Control charts;
MEWMA;
RST;
T^2.

Abstract. Various control charts have been proposed to monitor generalized linear profiles in Phase II. However, the robustness of the proposed methods in detecting different types and especially different directions of changes is not well-studied in the literature. In real-world applications, different kinds of change such as drift and multiple changes are likely to occur, which can be isotonic (increasing) or antitonic (decreasing). This paper studies the robustness of the Rao Score Test (RST) method, T^2, and Multivariate Exponential Weighted Moving Average (MEWMA) in different types, drift and multiple, and directions of changes. The RST method also benefits from a change-point detection approach whose performance is studied as well. According to the results, generally, the RST method shows a better performance in detecting different types of changes. Moreover, the performance of the RST method is robust to the direction of the change, while T^2 and MEWMA are not ARL-unbiased and show different performances under isotonic and antitonic changes. Therefore, to address this issue, a bias-reduced estimator is proposed for use in T^2. The results demonstrate that the proposed control chart outperforms T^2 and is less biased than T^2. Finally, a real-world problem is presented in which the aforementioned methods are applied to real data.

1. Introduction

A comprehensive review of the relevant literature indicates that there has been much interest in the area of profile monitoring within the past decade. A profile is the functional relationship between a response variable and one or more predictor variables. Profile monitoring is the use of control charts to monitor this functional relationship. Profiles can be utilized in a wide variety of manufacturing and service areas to monitor product or process performance over time. For instance, Kang and Albin [1] studied the case of aspartame, which is an artificial sweetener. They also considered a mass flow controller, in which using profile monitoring was more advantageous than a single measurement over time.

Studies in the field of profile monitoring can be categorized by the types of the profiles being monitored and types of the change-points that occur in the process. The type of a profile being monitored can be simple linear, nonlinear/multiple/polynomial regression, and nonparametric. For a review of the studies in this area one can refer to Noorossana et al. [2]. The most important types of changes are single step, drift, and multiple step. Most research efforts have focused on single step change, but in real applications, sometimes it is not practical to model changes as a simple single step change. For instance, in chemical applications or real tear and wear examples in which change occurs

*. *Corresponding author. Tel.: +17166032446*
E-mail addresses: Sahand.hajifar7@gmail.com (S. Hajifar); Mahlooji@sharif.edu (H. Mahlooji)

doi: 10.24200/sci.2019.50783.1863

gradually, they can be modeled better by a drift model. Niaki and Khedmati [3] studied the change time of a multivariate binomial process for step and drift types of change. Kazemzadeh et al. [4] considered change-point estimation of multivariate linear profiles under linear drift. Korkas and Fryzlewicz [5] proposed a Wild Binary Segmentation (WBS) method to detect the number and location of multiple change-points in the second order structure of a time series. For a review of change-point estimation methods, one can refer to Atashgar [6] and Aminikhanghahi and Cook [7].

All the aforementioned research work can be classified into Phase I control and Phase II monitoring. According to Woodall et al. [8], in the Phase I analysis, practitioners use a set of historical process data to evaluate process stability over time. Once anomalous observations are omitted, the in-control process is modeled and unknown parameters are estimated. In the Phase II analysis, online data is utilized to quickly identify shifts in the process.

There are some limiting assumptions in most of the aforementioned control schemes. For instance, they mostly assume that response variables are continuous, while in many real examples, the response variable may be discrete. For example, the response variable may be a Poisson random variable (the number of defects per item) or a binomial random variable. These limitations in earlier methods encouraged the authors of this paper to use a more general model, including a wide variety of distributions and link functions. Generalized linear models (GLMs) looked appropriate for this purpose.

Qiu [9], for the first time, introduced the GLM to statistical process control. Yeh et al. [10] developed five methods based on the Hotelling T^2 statistic and by assuming that the response variable is binary, and by applying them to Phase I control, compared the performance of the proposed methods. Paynabar et al. [11] presented Phase I risk-adjusted control charts which are based on the likelihood ratio test approach extracted from the change-point model. Soleymanian et al. [12] presented four control charts, including Hotelling T^2, MEWMA, a Likelihood Ratio Test and LRT/EWMA, to monitor a binary response variable in Phase II. Thereafter, the performances of the proposed methods were evaluated by an Average Run Length (ARL) measure. According to simulation results, all methods perform quite well. Moreover, the performance of the methods improves as the size of the samples increases. According to the results, the Multivariate Exponential Weighted Moving Average (MEWMA) method performs better than others when the changes in the parameters of logistic regression are small or moderate in size, and LRT/EWMA performs better than MEWMA when the changes are large in size. Furthermore, the LRT/EWMA control chart performs somewhat better than T^2 and Likelihood Ratio Test (LRT) control charts for every size of change and every size of sample. Koosha and Amiri [13] showed that ignoring autocorrelation between observations in different levels of a binary profile may lead to misleading results in evaluating the performance of different monitoring methods. Afterwards, they evaluated the performance of five T^2 control charts, proposed by Yeh et al. [10], under step and drift types of change assuming that autocorrelation exists. Their results demonstrated that T_I^2 is the best method, as already shown by Yeh et al. [10]. Amiri et al. [14] assessed the performance of five methods proposed by Yeh et al. [10] in Phase I monitoring of Gamma profiles. The gamma parameter estimation process was another contribution of this research. According to their results, T_I^2 performs better than other methods in detecting step change and drift. Nevertheless, T_R^2 performs better than other methods when the step change or drift occurs in both regression parameters simultaneously. Noorossana et al. [15] evaluated four methods in Phase II monitoring of polytomous profiles. Their results demonstrate that the MEWMA method performs better in detecting small size changes. Moreover, they showed that MEWMA and χ^2 are susceptible to the size of samples, whereas EWMA/R and χ^2/EWMA are less affected by the size of samples and roughly give the same results for different sample sizes. Amiri et al. [16] modified three methods which include T^2, the likelihood ratio test and F to make them applicable in GLM regression profiles following Poisson distribution. Simulation results show that the likelihood ratio test method shows the best performance in detecting changes and the F method performs better than others in detecting outliers. Qi et al. [17] presented a method to monitor generalized linear profiles via weighted likelihood ratio charts. Simulation results show that their proposed method shows a better performance compared to some existing methods. Sogandi and Amiri [18] proposed a maximum likelihood estimator to monitor GLM based regression profiles under step change and drift. Simulation results show that the proposed method performs better than others in detecting small and moderate changes. Khedmati and Niaki [19] developed an approach based on the U statistic for Phase II monitoring of generalized linear profiles and removed the impact of between-profile autocorrelation of error terms.

One attractive and practical topic in the process monitoring area is the effective estimation of the time at which a change has occurred in the process. The change-point approach provides experts with a diagnostic tool which can help to identify the time of change in process. This tool simplifies the identification of the root causes. Maximum likelihood estimation, clustering analysis, and artificial neural networks are examples of the common change-point identification methods. Noorossana et al. [20] used the maximum likelihood method to detect step changes. Perry et al. [21] utilized the maximum likelihood estimation to estimate a change-point when a linear change occurs in the process. Noorossana and Shadman [22] estimated the time of a monotonic change through a maximum likelihood estimator. Alaeddini et al. [23] presented a clustering approach to identify the change-point of the process. Atashgar and Noorossana [24] took advantage of artificial neural networks to identify the change-point in a process. Ahmadzadeh [25] used an artificial neural network to estimate the change-point for multivariate control charts.

Some researchers have studied change-point estimation in the profile monitoring content. In Phase II, Zou et al. [26] utilized the likelihood ratio method to identify the time of change when a step shift takes place in the mean of a linear profile. Moreover, Paynabar et al. [27] developed a change-point estimation approach for monitoring multivariate profiles. They applied their presented method in a real-world problem related to a forging process. Simulation results show that their method performs better in estimating the change-point in comparison with some other methods. Maleki et al. [28] proposed two maximum likelihood estimators to identify the real time of step changes and drifts in Phase II monitoring of binary profiles, where within-profile autocorrelation exists.

The maximum likelihood framework comprises three general tests: Wald, Likelihood Ratio, and Rao Score Test (RST). There are some differences between these three tests. In order to carry out a likelihood ratio test, one must estimate both models under the null and alternative hypotheses. The RST requires estimation of the model subject to the null hypothesis, while the Wald test needs to estimate the model based on the alternative hypothesis.

In previous research less effort has been devoted to study the susceptibility of the proposed methods to type, and, especially, the direction and increase or decrease of the change. Most of the proposed methods use Iterative Weighted Least Square (IWLS) as an estimator, which is a biased estimator [29]. Therefore, methods that use IWLS as an estimator are expected to be susceptible to the direction of the change, and they are not ARL-unbiased, while the RST method, which does not use an estimator, is robust to this issue. This paper studies this problem and proposes a bias-reduced estimator to be used in T^2, which satisfactorily attempts to solve this problem.

2. Methodology

2.1. RST method

Suppose that profile samples are collected over time and the jth random profile is comprised of n observations. There are a set of observations $\{(\mathbf{x}_{ij}, y_{ij}), i = 1, 2, ..., n\}$, in which y_{ij} is the ith response observation in the jth profile and $\mathbf{x}_{ij}$ is a vector consisting of q predictor variables ($\mathbf{x}_{ij} = (x_{ij1}, x_{ij2}, ..., x_{ijq})$). In this research, it is assumed that predictor variables in each profile are known and constant over time. Moreover, it is assumed that the relationship between the response variable and predictor variables can be adequately modeled by GLMs, which means:

1. Response variables are from an exponential family with a canonical form.
2. Linear combination of predictor variables with the coefficient vector $\boldsymbol{\beta}_j$ is as follows:

$$\eta_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta}_j, \quad (1)$$

$$\boldsymbol{\beta}_j = (\beta_{j1}, \beta_{j2}, ..., \beta_{jq}). \quad (2)$$

3. A monotone link function g exists that connects the mean of the response variable μ_{ij} to the linear predictor:

$$g(\mu_{ij}) = \eta_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta}_j, \quad (3)$$

$$\mu_{ij} = E(y_{ij}). \quad (4)$$

Most times, as in this research, x_{ij1} is set to be 1 for all i and j, therefore, $\boldsymbol{\beta}_{j1}$ is the intercept of the model. As mentioned earlier, the RST method is discussed in Shadman et al. [30] in detail but the following steps give a summary of the method:

1. For $j = 1, 2, 3, ..., t$, we have:

$$\mathbf{w}_j = \mathbf{w}_{j-1} + \mathbf{X}^T \mathbf{y}_j. \quad (5)$$

In Eq. (5), t is the current time of monitoring, $\mathbf{w}_0$ is a q dimensional $\mathbf{0}$ vector, $\mathbf{X}$ is an $n \times q$ regressor matrix and $\mathbf{y}_j$ is an n-variate response vector.

2. For $k = 1, 2, ..., t-1$, one has:

$$\mathbf{U}_0(k,t) = \mathbf{w}_t - \mathbf{w}_k - (t-k)\mathbf{X}^T \boldsymbol{\mu}_0. \quad (6)$$

In Eq. (6) $\boldsymbol{\mu}_0$ is an n-dimensional in-control mean vector of response variable.

And:

$$R_{k,t} = \mathbf{U}_0^T(k,t) \times \left(\frac{1}{t-k}\right) \mathbf{J}_0^{-1} \times \mathbf{U}_0(k,t). \quad (7)$$

In Eq. (7), $\mathbf{J}_0^{-1}$ is an in-control variance-covariance matrix.

3. The statistic of the method equals:

$$R_{\max,t} = \max_{\max(0,t-window) \leq k < t}(R_{k,t}). \quad (8)$$

In Eq. (8), *window* is a rather newly adapted notion which is used because as t becomes very large, it is difficult to record all the past data and find the maximum. Choosing a *window* is a trade-off between recording less data and having a good performance for the control chart. Theoretically, having an infinite *window* (recording all of the data) should show the best performance. However, it can be shown that in most control charts if the *window* is not too small, it will show a performance close to the method without using *window*.

If $R_{\max,t} \geq h_t$, then control chart signals and the tth profile is out of control. h_t is the control limit for the tth profile and the method to determine this limit will be described later. The change-point is estimated by:

$$\widehat{\tau} = \arg\max_{\max(0,t-window \leq k<t)}(R_{k,t}). \tag{9}$$

In most research efforts the control limits are constant for every t but Margavio et al. [31] showed that this approach could result in variations of false alarms over time. Therefore, in this research, the following conditional probability is used, which generates a constant alarm rate for every t:

$$\Pr(R_{\max,t} > h_t | R_{\max,i} \leq h_i, 1 \leq i < t) = \alpha. \tag{10}$$

Since Eq. (10) is not easily tractable, simulation is used to determine the sequence of control limits [32]. Similar to Shadman et al. [30], it is assumed that after the 100th profile, the control limit converges to a constant control limit, and the control limit is set equal to the control limit of the 100th profile. In order to estimate control limits using simulation, an 80000*100 matrix was generated, in which the ith row is an in-control simulation of the following vector:

$$\mathbf{V}_i = [R_{\max,1}, R_{\max,2}, ..., R_{\max,100}]. \tag{11}$$

In order to set ARL at 200, the 0.995th quantile of the first column in the matrix was calculated and determined as h_1. After that, for elements of column 1 which are greater than h_1, the relevant row was removed from the matrix. A similar procedure was applied to the second and other columns in order to estimate all the control limits.

2.2. Bias-reduced T^2

Firth [29] proposed an approach called bias reduction of maximum likelihood estimates. In the fitting process of the logistic model sometimes one or more parameter estimates diverge to ± infinity. This phenomenon is called separation which is ideally solved by the Firth procedure [33]. This method is built on the generic iteration proposed in Kosmidis and Firth [34] to solve the bias-reducing adjusted score equations. According to Kosmidis and Firth [34] this estimator has two advantages: 1) It is a second order unbiased estimator and has smaller variance compared to Maximum Likelihood Estimation (MLE), 2) The resulted estimates and their standard errors are finite. A detailed description of this estimator is given in Heinze and Schemper [35]. As mentioned earlier, most methods that are proposed to monitor generalized linear profiles are biased due to using a biased estimator. Therefore, to deal with this problem, the T^2 method proposed by Yeh et al. [10] was enhanced using the estimator proposed by Firth [29] instead of the usual IWLS method.

3. Performance comparisons

This section is devoted to evaluate performances of the RST, T^2, MEWMA, and proposes using the bias reduced method under different types and direction of changes. This comparison is done using Monte Carlo simulation and the ARL measure is used to evaluate the methods. The RST method also benefits from a change-point estimation approach, whose performance is studied as well.

3.1. Hotelling T^2

A Hotelling T^2 control chart was used by Kang and Albin [1] for monitoring simple linear profiles. Yeh et al. [10] developed five Hotelling T^2 control charts to monitor binary profiles in Phase I. The T^2 statistic for the jth profile is calculated as:

$$T_j^2 = (\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0)^T \boldsymbol{\Sigma}_0^{-1} (\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0). \tag{12}$$

The above statistic can be rewritten as:

$$T_j^2 = (\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0)^T \mathbf{X^T W X} (\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0), \tag{13}$$

where $\hat{\boldsymbol{\beta}}_{\boldsymbol{j}}$ is the estimated vector of parameters via iterative weighted least square algorithm, $\boldsymbol{\beta}_0$ is the in-control vector of parameters and $\boldsymbol{\Sigma}_0$ is the in-control variance-covariance matrix. In Eq. (13), $\mathbf{W}$ is a $N \times N$ diagonal matrix in which diagonal elements are calculated as follows:

$$w_{ii} = \frac{1}{Var(y_i)} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\eta}_i}\right)^2. \tag{14}$$

As long as this statistic is less than the upper control limit, it is assumed that the process is in-control, but, when $T_j^2 >$ UCL, it is assumed that the process is out of control. The upper control limit is estimated according to a given ARL.

3.2. MEWMA method

MEWMA which was proposed by Zou et al. [26], is also applicable in monitoring generalized linear profiles. Soleymanian et al. [12] and Noorossana et al. [15] used this method for monitoring generalized linear profiles.

The MEWMA statistic for the jth profile is calculated as:

$$Q_j = \mathbf{Z}_j^T \mathbf{\Sigma}_{z_j}^{-1} \mathbf{Z}_j. \tag{15}$$

In Eq. (15), $\mathbf{Z}_j$ is a q-dimensional vector which is equal to:

$$\mathbf{Z}_j = \lambda(\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0) + (1-\lambda)\mathbf{Z}_{j-1}. \tag{16}$$

$\mathbf{\Sigma}_{\mathbf{Z_j}}$ is the asymptotic variance-covariance matrix of $\mathbf{Z}_j$ which is calculated as follows :

$$\mathbf{\Sigma}_{Z_j} = \frac{\lambda}{2-\lambda}\mathbf{J}_0^{-1}, \tag{17}$$

$\hat{\boldsymbol{\beta}}_j$ is the maximum likelihood estimator of $\boldsymbol{\beta}_\mathbf{j}$ which is calculated by the iterative weighted least square algorithm, λ is weighting parameter and $\mathbf{Z_0} = \mathbf{0}$. This control chart gives a signal when $Q_j > L_{MEWMA}$ and L_{MEWMA}, the MEWMA upper control limit, is simulated to give a known ARL, which is 200.

3.3. Binomial profiles

To simplify the problem, suppose that there is only one predictor variable ($q = 2$) and $y_{ij} \sim$ Binomial (m, π_{ij}) in which m is the number of observations and π_{ij} is the probability of success for the jth profile and ith observation. The link function is considered as:

$$g(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 \mathbf{x}_{ij}. \tag{18}$$

Eq. (18) can be simplified as:

$$\pi_{ij} = \frac{\exp(\mathbf{x}_{ij}^T\boldsymbol{\beta})}{1+\exp(\mathbf{x}_{ij}^T\boldsymbol{\beta})} = \frac{\exp(\eta_{ij})}{1+\exp(\eta_{ij})}. \tag{19}$$

In-control parameters are assumed as $\boldsymbol{\beta}_0 = (\beta_{00}, \beta_{10})^T = (-2.8, 1)^T$. Predictor variables are 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1 and $n = 10$. $m = 30$ is set and control charts are simulated for three different time windows: 20, 50, 100.

3.3.1. Isotonic drift

The change-point is assumed to be $\tau = 30$, this means that for $j = 1, 2, ..., 30$ the profiles are in-control and parameters are as $\boldsymbol{\beta}_0 = (-2.8, 1)^T$. After the 30th profile for $j = 31, 32, ...$ the process undergoes a change which is modeled as:

$$\boldsymbol{\beta}_1 = (\beta_{00} + (j-30)\times\delta_1\times\sigma_1, \beta_{10} + (j-30) \times\delta_2\times\sigma_2), \tag{20}$$

σ_1 and σ_2 are the standard deviation of intercept and slope, respectively, which are calculated as:

$$\mathbf{\Sigma}_0 = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1} = \begin{bmatrix} 0.2186 & -0.2936 \\ -0.2936 & 0.4771 \end{bmatrix}. \tag{21}$$

Therefore, standard deviations are calculated as $\sigma_1 = 0.4676$ and $\sigma_2 = 0.6907$, and the performance of the method is evaluated by ARL and estimated change-point criteria for different δ_1 and δ_2. Results are given in Tables 1 and 2. In this research, three cases are simulated: 1) change in intercept; 2) change in slope; and 3) change in intercept and slope, simultaneously. However, for the sake of brevity, simulation results are presented only for case 1. In Tables 1 and 3, the numbers stand for ARL's and the numbers inside parentheses are standard deviations of run lengths. All the results are determined based on 10000 iterations.

3.3.2. Isotonic multiple step change

In the case of isotonic multiple step change, two step changes occur in the process at $\tau_1 = 30$ and $\tau_2 = 35$. The monitoring is done in Phase II and, according to the literature, the aim is to detect the change as soon as possible and estimate the change-point. Let δ_{ij} control the size of change in the ith parameter and jth step. The model of change is as:

$$\boldsymbol{\beta}_0 = (-2.8, 1), \quad j = 1, 2, ..., 30, \tag{22}$$

$$\boldsymbol{\beta}_1 = (\beta_{00} + \delta_{11}\times\sigma_1, \beta_{10} + \delta_{21}\times\sigma_2), \quad j = 31, 32, ..., 35, \tag{23}$$

$$\boldsymbol{\beta}_2 = (\beta_{00} + \delta_{12}\times\sigma_1, \beta_{10} + \delta_{22}\times\sigma_2), \quad j = 36, 37, ... \tag{24}$$

Standard deviations of intercept and slope are calculated as shown in the previous section and the number of simulation replications is set at 10000. Moreover, since the performance of Hotelling T^2 was very poor, its results are not given for this case. Results are shown in Tables 3 and 4. In Table 4, which is related to the change-point, the first change has been assumed as the change-point.

As discussed before, to simplify the problem, $q = 2$ and $n = 10$ are considered. In order to see if the RST method is still efficient for larger values of q and n, both parameters ($q = 4$ and $n = 20$) are doubled and the simulations are run. In this case, the link function, in-control and out of control parameters are defined by:

$$g(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + \beta_3 x_{ij}^3, \tag{25}$$

$$\boldsymbol{\beta}_0 = (\beta_{00}, \beta_{10}, \beta_{20}, \beta_{30}) = (-2.8, 1, 2, 3)^T, \quad j = 1, 2, ..., 30, \tag{26}$$

$$\boldsymbol{\beta}_0 = (\beta_{00} - (j-30)\times\delta_1\times\sigma_1, \beta_{10}, \beta_{20}, \beta_{30}) = (-2.8, 1, 2, 3)^T, \quad j = 31, 32, ... \tag{27}$$

Table 5 demonstrates that RST still outperforms T^2.

Table 1. Average Run Length (ARL) for isotonic drift in intercept of binomial profile.

δ_1	RST (Window = 20)	RST (Window = 50)	RST (Window = 100)	Hotelling T^2	MEWMA $\lambda = 0.05$	MEWMA $\lambda = 0.2$	MEWMA $\lambda = 0.4$
0.001	95.8035 (63.06388)	93.1838 (59.71108)	90.5993 (57.3664)	367.0953 (333.1151)	165.4975 (85.42436)	200.363 (117.1865)	250.1403 (165.8289)
0.005	42.7568 (20.78963)	41.3498 (19.76968)	41.3548 (19.54215)	163.8714 (68.62347)	61.8955 (19.43502)	72.2004 (24.62263)	93.2443 (32.86928)
0.01	28.5414 (12.36723)	27.9485 (12.08684)	27.9149 (11.93115)	97.516 (30.76698)	40.3651 (10.5095)	43.9174 (12.56879)	55.6392 (15.826645)
0.05	10.5216 (3.699964)	10.4708 (3.803413)	10.4925 (3.724882)	25.8201 (5.102287)	15.8617 (2.7960)	14.1949 (2.848072)	15.9725 (3.325168)
0.1	6.8208 (2.221821)	6.8206 (2.238262)	6.8363 (2.237611)	14.2792 (2.521239)	10.8136 (1.7164)	9.0160 (1.582259)	9.4987 (1.787232)
0.5	2.4638 (0.689122)	2.4664 (0.68518)	2.4648 (0.688739)	3.6675 (0.656311)	4.6202 (0.6744)	3.4937 (0.564589)	3.212 (0.527879)
1	1.6466 (0.481569)	1.6480 (0.479475)	1.6487 (0.479467)	2.0833 (0.362162)	3.2505 (0.5266)	2.4007 (0.500739)	2.0708 (0.302634)
1.5	1.1592 (0.365862)	1.1638 (0.370094)	1.1567 (0.363518)	1.6528 (0.47608)	2.7079 (0.4604)	1.97 (0.194165)	1.8744 (0.331398)
2	1.0062 (0.078496)	1.0057 (0.075283)	1.0052 (0.071923)	1.121 (0.326127)	2.2599 (0.4554)	1.8458 (0.36114)	1.5147 (0.499784)
2.5	1 (0)	1 (0)	1 (0)	1.0021 (0.045778)	2.0029 (0.2247)	1.5922 (0.491426)	1.1359 (0.342682)
3	1 (0)	1 (0)	1 (0)	1 (0)	1.9484 (0.22167)	1.2788 (0.448409)	1.0114 (0.10616)

Table 2. Change-point estimation (isotonic drift in intercept of binomial profile).

δ_1	Window = 20		Window = 50		Window = 100	
	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)
0.001	119.8958	62.18153	112.6677	56.93061	104.8658	52.83577
0.005	65.0683	20.78963	57.9071	19.16977	56.3100	19.54215
0.01	50.4597	12.13221	46.1548	12.08684	45.8965	11.93115
0.05	34.7407	5.126818	34.1740	6.237445	34.2766	6.211899
0.1	32.4584	3.981491	32.2473	4.606446	32.2701	4.622764
0.5	30.3311	2.385932	30.2677	2.60765	30.2682	2.660013
1	30.0730	1.917517	30.0787	2.032512	30.0470	0.479467
1.5	29.9599	1.270548	29.9419	1.366793	29.9429	1.333731
2	29.9426	0.603246	29.9327	0.663303	29.9402	0.600187
2.5	29.9817	0.280651	29.9715	0.378798	29.9761	0.325467
3	29.9931	0.158279	29.9890	0.21466	29.9931	0.150507

3.4. Poisson profiles

Similar to a binomial example, suppose that there is only one predictor variable ($q = 2$) and $y_{ij} \sim$ Poisson (μ_{ij}), in which μ_{ij} is the mean for the jth profile and ith observation. Link function is considered as:

$$g(\mu_{ij}) = \log(\mu_{ij}) = \beta_0 + \beta_1 x_{ij}. \quad (28)$$

Eq. (28) can be simplified as:

$$\mu_{ij} = \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta}) = \exp(\eta_{ij}). \quad (29)$$

In-control parameters are assumed as $\boldsymbol{\beta}_0 = (\beta_{00}, \beta_{10})^T = (3, 2)^T$. Predictor variables are 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1 and, therefore, $n = 10$. Simulations are run for three different time windows: 20, 50,100.

3.4.1. Antitonic drift

Change-point is assumed to be $\tau = 30$, that is, for $j = 1, 2, ..., 30$ profiles are in-control and parameters are as $\boldsymbol{\beta}_0 = (3, 2)^T$. After the 30th profile for $j = 31, 32, ...,$ the process undergoes a change which is modeled as:

Table 3. Average Run Length (ARL) for isotonic multiple change in intercept of binomial profile.

$(\delta_{11}, \delta_{12})$	RST (Window = 20)	RST (Window = 50)	RST (Window = 100)	MEWMA $\lambda = 0.05$	MEWMA $\lambda = 0.2$	MEWMA $\lambda = 0.4$
(0.1,0.2)	35.4825 (30.38748)	31.3287 (24.69028)	30.4503 (22.42146)	87.2043 (63.5371)	249.3337 (242.0212)	622.2976 (610.2426)
(0.1,0.3)	19.1274 (12.88038)	18.1492 (10.95447)	18.1014 (10.83084)	37.2727 (17.7532)	67.8792 (57.2167)	292.2946 (289.1756)
(0.1,0.4)	13.0761 (6.754236)	12.9803 (6.47031)	12.9902 (6.524163)	24.3403 (8.6288)	28.7672 (18.4522)	92.6143 (86.0084)
(0.1,0.5)	10.3039 (4.444361)	10.3657 (4.428178)	10.3195 (4.430533)	18.7677 (5.4528)	17.7730 (8.1244)	37.0998 (29.2889)
(0.1,0.6)	8.8622 (3.246754)	8.8922 (3.218288)	8.8538 (3.21612)	15.7622 (3.9346)	13.3857 (4.4544)	20.2450 (12.4547)
(0.1,0.7)	7.9498 (2.50533)	7.9072 (2.630359)	7.9525 (2.543628)	13.8507 (3.0414)	11.2853 (2.9839)	13.8790 (6.313)
(0.1,0.8)	7.2955 (2.109545)	7.3186 (2.158447)	7.2930 (2.122346)	12.4965 (2.5694)	10.0001 (2.1770)	10.9394 (3.5871)
(0.1,0.9)	6.8901 (1.855915)	6.8799 (1.87512)	6.8723 (1.900051)	11.5568 (2.1714)	9.1583 (1.7149)	9.4649 (2.4113)
(0.1,1)	6.5446 (1.659406)	6.5754 (1.693669)	6.5639 (1.6764)	10.8136 (1.9273)	8.5933 (1.4045)	8.5968 (1.7331)
(0.1,1.1)	6.3227 (1.524849)	6.3221 (1.541347)	6.3354 (1.528367)	10.2322 (1.6928)	8.1681 (1.2038)	7.9762 (1.3479)

Table 4. Change-point estimation (isotonic multiple change in intercept of binomial profile).

$(\delta_{11}, \delta_{12})$	Window = 20		Window = 50		Window = 100	
	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)
(0.1,0.2)	57.7286	30.29054	47.1392	23.52195	43.2806	20.43896
(0.1,0.3)	40.8859	12.58369	36.4500	10.64466	36.1211	10.52653
(0.1,0.4)	35.7408	6.588734	34.4359	7.410512	34.3527	7.343862
(0.1,0.5)	34.3025	4.85477	33.7665	6.049163	33.7434	5.901471
(0.1,0.6)	33.9021	4.26712	33.7055	5.062467	33.5491	5.172503
(0.1,0.7)	33.8740	3.775781	33.5166	4.771658	33.6375	4.522731
(0.1,0.8)	33.7937	3.636776	33.5818	4.35579	33.5708	4.374127
(0.1,0.9)	33.8383	3.450703	33.6934	4.021442	33.6544	4.11227
(0.1,1)	33.8549	3.389638	33.7509	3.845237	33.6924	3.985923
(0.1,1.1)	33.9336	3.27249	33.8121	3.671974	33.8363	3.73067

$$\boldsymbol{\beta}_1 = (\beta_{00} - (j-30) \times \delta_1 \times \sigma_1, \beta_{10} - (j-30) \times \delta_2 \times \sigma_2), \quad (30)$$

σ_1 and σ_2 are the standard deviation of intercept and slope, respectively, which are calculated as:

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1} = \begin{bmatrix} 0.0117 & -0.0146 \\ -0.0146 & 0.0207 \end{bmatrix}. \quad (31)$$

Therefore, $\sigma_1 = 0.1082$ and $\sigma_2 = 0.1440$ and the performance of the methods are evaluated by ARL and estimated change-point criterion for different δ_1 and δ_2. All simulation results are based on 10000 replications. Results are given in Tables 6 and 7.

3.4.2. Antitonic multiple change

In antitonic multiple change, two step changes occur in the process at $\tau_1 = 30$ and $\tau_2 = 35$. It was allowed that δ_{ij} control the size of change in the ith parameter and jth step. The model of the change is as:

Table 5. Result for effect of number of predictors and observations (standard deviations are given in parenthesis).

δ_1	RST		T^2
	ARL	**Change-point**	**ARL**
0.001	81.8060 (45.6467)	94.1230 (37.4829)	165.7940 (137.5765)
0.005	32.1920 (12.5176)	45.5820 (12.5241)	107.7130 (68.9080)
0.01	21.6800 (7.6661)	38.6630 (8.9933)	79.7070 (43.4161)
0.05	8.0310 (2.3966)	32.2180 (4.3323)	30.9970 (13.2006)
0.1	5.3520 (1.4430)	31.1470 (3.1638)	19.7210 (7.3898)
0.5	1.9090 (0.3935)	29.8060 (2.5642)	6.5260 (2.1391)
1	1.1440 (0.3513)	29.8800 (1.4421)	4.0500 (1.2685)
1.5	1 (0)	29.9560 (0.4902)	3.0420 (0.9743)
2	1 (0)	29.9870 (0.1512)	2.4830 (0.8357)
2.5	1 (0)	29.9810 (0.3045)	2.1580 (0.7011)
3	1 (0)	29.9960 (0.0774)	1.8900 (0.6483)

Table 6. Average Run Length (ARL) for antitonic drift in intercept of Poisson profile.

δ_1	**RST (Window = 20)**	**RST (Window = 50)**	**RST (Window = 100)**	**Hotelling T^2**	**MEWMA $\lambda = 0.05$**	**MEWMA $\lambda = 0.2$**	**MEWMA $\lambda = 0.4$**
0.001	102.1630 (65.0406)	97.6786 (59.8645)	91.6726 (55.2262)	122.5580 (84.5574)	77.8625 (41.9650)	90.4725 (52.4263)	100.5746 (61.7167)
0.005	42.6454 (19.5507)	40.6931 (18.3308)	40.4171 (17.9133)	59.3012 (29.7561)	35.8102 (13.9485)	39.1006 (16.3133)	44.7450 (19.7063)
0.01	27.8681 (11.4363)	27.3876 (10.9084)	27.2422 (10.9074)	39.6646 (17.1881)	24.9405 (8.4750)	25.9056 (9.6675)	29.3069 (11.4395)
0.05	10.4254 (3.4140)	10.4066 (3.5638)	10.4094 (3.4250)	13.5452 (4.5513)	10.8028 (2.8830)	9.8563 (2.7702)	10.3102 (3.1317)
0.1	6.8028 (2.0960)	6.8258 (2.1055)	6.8451 (2.1027)	8.2083 (2.5387)	7.5457 (1.8817)	6.6082 (1.6707)	6.6275 (1.8213)
0.5	2.5733 (0.6901)	2.5679 (0.6824)	2.5857 (0.6782)	2.5374 (0.7294)	3.3113 (0.7840)	2.6995 (0.6152)	2.4756 (0.6093)
1	1.7318 (0.4570)	1.7407 (0.4546)	1.7514 (0.4454)	1.5783 (0.5029)	2.3713 (0.5807)	1.8991 (0.3951)	1.7460 (0.4413)
1.5	1.2753 (0.4467)	1.2788 (0.4484)	1.2843 (0.4511)	1.1369 (0.3438)	1.9288 (0.4349)	1.6055 (0.4888)	1.3342 (0.4717)
2	1.0302 (0.1711)	1.0314 (0.1744)	1.0330 (0.1786)	1.0066 (0.0810)	1.7440 (0.4410)	1.2933 (0.4553)	1.0721 (0.2587)
2.5	1.0007 (0.0264)	1.0008 (0.0283)	1.0014 (0.0374)	1 (0)	1.5820 (0.4933)	1.0849 (0.2787)	1.0061 (0.0779)
3	1 (0)	1.0001 (0.0100)	1 (0)	1 (0)	1.4059 (0.4911)	1.0136 (0.1158)	1 (0)

$$\boldsymbol{\beta}_0 = (-2.8, 1), \quad j = 1, 2, ..., 30, \tag{32}$$

$$\boldsymbol{\beta}_1 = (\beta_{00} - \delta_{11} \times \sigma_1, \beta_{10} - \delta_{21} \times \sigma_2),$$
$$j = 31, 32, ..., 35, \tag{33}$$

$$\boldsymbol{\beta}_2 = (\beta_{00} - \delta_{12} \times \sigma_1, \beta_{10} - \delta_{22} \times \sigma_2),$$
$$j = 36, 37, ... \tag{34}$$

Standard deviations of intercept and slope are calculated as shown in subsection 3.4.1 and simulation results are based on 10000 replications. Results are given in Tables 8 and 9, in which the first change has been assumed as the change-point.

3.5. A summary of simulation results

3.5.1. Isotonic drift

In the isotonic drift case results agree with those results presented by Shadman et al. [30] for isotonic single step change. The RST method outperforms MEWMA and T^2 in all cases and a much better performance is noticed for small changes. When the changes are very

Table 7. Change-point estimation (antitonic drift in intercept of Poisson profile).

δ_1	Window = 20		Window = 50		Window = 100	
	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)	$\bar{\hat{\tau}}$	s.e ($\hat{\tau}$)
0.001	124.1868	63.3619	111.9734	55.0070	98.2496	48.1345
0.005	62.7816	18.3835	53.4607	17.6813	51.3865	18.8041
0.01	47.8178	11.1994	43.1857	12.4164	42.5046	13.0578
0.05	33.6446	4.9518	32.9300	6.3049	32.9654	6.2887
0.1	31.6196	3.9092	31.3930	4.8558	31.3608	4.8428
0.5	30.0230	2.4248	29.9161	2.9887	29.9218	2.9734
1	29.8152	2.0916	29.7412	2.4162	29.7918	2.3062
1.5	29.8052	1.6822	29.7457	2.0124	29.7564	2.0091
2	29.8415	1.2015	29.8212	1.3430	29.8537	1.1899
2.5	29.9039	0.8001	29.8973	0.8731	29.9180	0.7137
3	29.9455	0.607396	29.9545	0.5001	29.9454	0.5292

Table 8. Average Run Length (ARL) for antitonic multiple change in intercept of Poisson profile.

$(\delta_{11}, \delta_{12})$	RST (Window = 20)	RST (Window = 50)	RST (Window = 100)	Hotelling T^2	MEWMA $\lambda = 0.05$	MEWMA $\lambda = 0.2$	MEWMA $\lambda = 0.4$
(0.1,0.2)	33.0810 (27.5027)	28.6003 (19.9843)	27.6637 (18.2534)	76.9551 (74.3105)	21.2364 (11.5879)	25.8082 (19.5560)	36.7855 (31.4660)
(0.1,0.3)	17.3729 (10.3399)	16.6281 (8.8489)	16.4788 (8.7580)	46.0258 (42.7766)	14.8859 (6.2077)	15.1887 (8.2943)	19.6177 (13.7646)
(0.1,0.4)	12.2378 (5.4792)	12.2963 (5.3827)	12.1655 (5.3349)	28.5567 (23.9756)	12.0527 (4.1468)	11.3892 (4.6050)	12.9013 (6.8018)
(0.1,0.5)	10.0138 (3.7808)	10.0327 (3.8308)	10.0225 (3.7681)	19.3053 (14.5890)	10.6090 (3.2261)	9.5737 (3.0303)	10.0674 (4.0634)
(0.1,0.6)	8.7051 (2.9399)	8.7580 (2.9375)	8.7518 (2.8576)	13.8968 (9.0027)	9.6099 (2.5577)	8.5186 (2.2459)	8.6840 (2.6887)
(0.1,0.7)	7.9212 (2.3745)	7.9552 (2.4096)	7.9992 (2.3449)	10.7440 (5.6346)	8.9306 (2.1335)	7.9273 (1.7415)	7.8477 (2.0036)
(0.1,0.8)	7.3439 (2.0637)	7.3938 (2.0668)	7.3812 (2.0530)	8.9828 (3.8050)	8.4001 (1.8876)	7.4613 (1.4778)	7.3397 (1.5578)
(0.1,0.9)	6.9886 (1.8006)	7.0202 (1.7799)	7.0362 (1.7580)	7.8999 (2.6537)	8.0562 (1.6717)	7.1758 (1.2929)	6.9960 (1.3255)
(0.1,1)	6.7054 (1.6333)	6.6985 (1.6996)	6.7278 (1.6435)	7.2182 (1.9965)	7.7799 (1.5107)	6.9349 (1.1899)	6.7302 (1.1539)
(0.1,1.1)	6.4822 (1.5443)	6.5006 (1.4901)	6.4893 (1.5672)	6.7191 (1.4712)	7.5433 (1.4035)	6.7422 (1.1216)	6.5400 (1.0511)

small, the out of control ARL's related to MEWMA and T^2 methods are even larger than in-control ARL (200), which shows these control charts are not ARL-unbiased.

Results of MEWMA for different values of λ show that MEWMA with a small λ performs better in detecting small changes and MEWMA with a large λ performs better in detecting large changes. Comparing MEWMA with T^2 shows that MEWMA performs better in detecting small changes while T^2 performs better in detecting large changes.

Different time windows give close results for moderate and large changes but this similarity is caused because the change occurs early. For small changes, the effect of the time window is relatively discernible and increasing the time window improves the performance of the control chart. Nevertheless, the time window of 50 gives satisfactory results. Moreover, the change-point estimator estimates the change-point close to the real change-point for moderate and large changes.

3.5.2. Isotonic multiple step change

In isotonic multiple step change the superiority of the RST method over MEWMA is evident. Since the T^2 method's performance was poor, it was not used in simulations. As mentioned earlier, since the research

Table 9. Change-point estimation (antitonic multiple change in intercept of Poisson profile).

	Window = 20		Window = 50		Window = 100	
$(\delta_{11},\delta_{12})$	$\bar{\hat{\tau}}$	s.e $(\hat{\tau})$	$\bar{\hat{\tau}}$	s.e $(\hat{\tau})$	$\bar{\hat{\tau}}$	s.e $(\hat{\tau})$
(0.1,0.2)	52.9091	27.0831	40.5922	18.2721	37.5572	15.7719
(0.1,0.3)	37.5309	9.7858	33.6154	8.9670	33.5525	8.9150
(0.1,0.4)	33.9825	5.4301	32.8955	6.6706	32.8347	6.8298
(0.1,0.5)	33.4308	4.4555	32.7501	5.8495	32.6726	6.0940
(0.1,0.6)	33.2084	4.2335	32.8438	5.3553	32.8973	5.1945
(0.1,0.7)	33.3253	3.9451	33.0012	4.9768	33.0551	4.8684
(0.1,0.8)	33.3672	3.8647	33.1313	4.6572	33.1544	4.6130
(0.1,0.9)	33.4746	3.6598	33.2840	4.4724	33.3692	4.1998
(0.1,1)	33.6080	3.5141	33.3649	4.3120	33.4304	4.2077
(0.1,1.1)	33.6382	3.5171	33.5438	4.0952	33.4695	4.1614

is done in Phase II, the aim of the research is to detect the change as soon as possible and estimate the first change. This assumption agrees with Perry et al. [21]. In simulations, a constant size for the first change and 10 different sizes for the second change are assumed. According to the results, as the size of the second change increases, at first, the performance of the change-point estimation improves. However, for larger sizes of the second change beyond a point, the performance of the change-point estimation worsens, and the change-point is estimated closer to the second step change. It is obvious that for large sizes of the second step change, most of the time the estimator estimates the second step as the change-point.

3.5.3. Antitonic drift

In antitonic drift the RST method is not always superior to the other two methods. For small changes, MEWMA with a small λ performs better than the RST method. Moreover, for moderate size changes, sometimes the MEWMA method with a moderate λ performs better than the RST method. For large changes, T^2 performs better than the RST method. Similar to an isotonic case for large and moderate changes, the change-point estimation approach comes up with estimates close to the real change-point.

3.5.4. Antitonic multiple step change

In antitonic multiple step change, the RST method outperforms the other two methods most of the time. The RST method is always superior to the T^2 method but MEWMA with a small or moderate λ performs better than the RST method in detecting small changes. Similar to the isotonic multiple example, increasing the size of the second step change causes the change-point approach to estimate the second change as the change-point most of the time.

3.6. Bias-reduced T^2

According to the results, since T^2 and MEWMA use a biased estimator, the performance of these methods depends on the direction of the change, and they are not ARL-unbiased. Therefore, the Firth bias-reduced estimator was used as the estimator of the T^2 and the simulations were run based on the following model:

$$\beta_1 = (\beta_{00} \pm \delta_1 \times \sigma_1, \beta_{10}). \tag{35}$$

The only difference is that, here, both increasing and decreasing changes were used to see the results. The results show that the proposed method is significantly less biased than T^2. The results are given in Table 10 and Figure 1. According to the results, the proposed method shows less difference between isotonic change ARL and antitonic change ARL. Therefore, the proposed method is less biased than T^2.

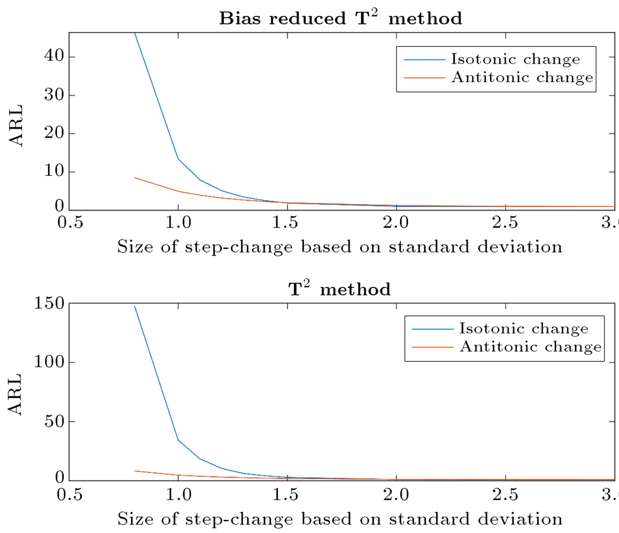

Figure 1. Comparison of the Bias reduced T^2 with T^2.

Table 10. Average Run Length (ARL) comparison of the proposed method with T^2.

δ_1	Bias reduced T^2		T^2	
	Isotonic	**Antitonic**	**Isotonic**	**Antitonic**
0.8	46.3708 (45.4409)	8.5361 (7.9815)	147.8055 (148.8448)	8.2820 (7.8097)
1	13.4137 (12.7153)	4.9672 (4.4272)	34.4333 (33.8360)	4.7895 (4.2527)
1.1	7.9368 (7.3582)	3.9600 (3.4593)	18.4573 (18.0038)	3.8633 (3.3482)
1.2	5.1139 (4.6667)	3.2107 (2.6734)	10.4202 (9.8837)	3.1227 (2.5878)
1.3	3.4750 (2.9379)	2.6992 (2.1582)	6.1981 (5.6435)	2.6422 (2.0693)
1.4	2.5291 (1.9569)	2.2879 (1.7030)	4.2043 (3.6496)	2.2458 (1.6594)
1.5	1.9047 (1.3017)	1.9858 (1.4171)	2.8906 (2.3483)	1.9450 (1.3494)
2	1.0534 (0.2374)	1.2604 (0.5634)	1.1296 (0.3821)	1.2444 (0.5460)
2.5	1.0004 (0.0200)	1.0545 (0.2382)	1.0014 (0.0374)	1.0579 (0.2453)
3	1 (0)	1.0093 (0.0960)	1 (0)	1.0097 (0.1010)

4. A real-data analysis

In this section, the RST method is used to monitor a real-world data example in Phase II. This example was used and monitored in Shadman et al. [30,36], in both Phase I and Phase II. This research, however, focuses on Phase II and drift. An instrument named a Dispergrader was used to get data. This instrument is used to assess the dispersion of carbon black filler in a rubber mix.

Evaluation of filler dispersion is very important in resin substances, especially in the tire industry. The amount of dispersion affects the quality of the product; and therefore, it is considered a parameter in quality control systems. In the literature, different methods are introduced to evaluate the dispersion. These methods are mostly subjective, time-consuming and costly. The Dispergrader, using a microscope, provides an appropriate method for measuring the dispersion of fillers.

In resin production filler pellets, which are mostly 1 milimeter in diameter, are broken down into aggregates which are mostly 1 micron in diameter to produce the desirable substance. In the process of production, this breaking down does not occur in all particles, and particles which are not broken down remain as agglomerates and can be considered as defects. These defects can highly affect mechanical features such as tensile strength, rupture, and fatigue.

In the Dispergrader, beams are sent out to the surface of the sample in such a way that the beams direction and surface meet at a 30-degree angle to each other. This equipment magnifies the image 100 times.

A sample curved rubber bar is put in the Dispergrader and beams are sent out on its surface. The beams reflected from agglomerates larger than 3 microns in diameter make white spots in the image. Afterwards, the number of white spots is recorded for each given spot size. Recording is done in such a way that sizes between 3 and 6 microns are recorded as 3, sizes between 6 and 9 microns are recorded as 6, and this approach is also used for larger spots. All the white spots which are 57 microns in diameter or greater are recorded as 57. The quality characteristic is defined as a profile in which agglomerate count is the response variable and the agglomerate diameter is the predictor variable.

The aim of Phase II is online monitoring of the Dispergrader. To start Phase II, a set of in-control historical data are needed to estimate parameters and control limits. For this purpose, in-control data used by Shadman et al. [30] is chosen. They assumed that the response variable can be adequately modeled by negative binomial distribution. They used a log function and assumed that there is a second order polynomial relationship between response variable and predictor variable:

$$Y \sim Negative\ Binomial(v = 1; p = 1/(1 + \mu/v)), \tag{36}$$

$$Log(\mu) = \beta_0 + \beta_1 X + \beta_2 X^2. \tag{37}$$

In this research, the authors used the same model and the approach of Ver Hoef and Boveng [37] to estimate parameters. Parameters were estimated as:

$$\beta_0^{in-control} = 6.0781,$$

$$\beta_1^{in-control} = -0.0104,$$

$$\beta_2^{in-control} = -0.0030.$$

Three different scenarios are considered, in which parameters of the model undergo isotonic drift. Scenario 1 is as follows:

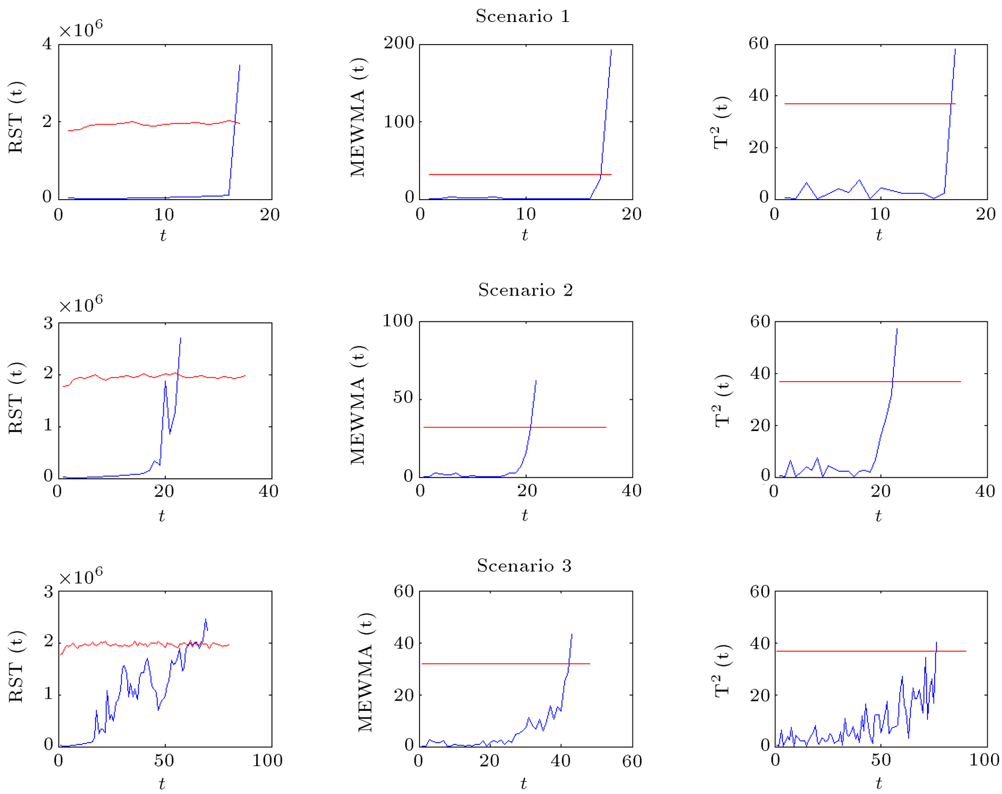

Figure 2. Control charts for real-data analysis.

$$\begin{cases}\beta_0=\beta_0^{in-control}, & t=1,2,...,16\\ \beta_0=\beta_0^{in-control}+0.5\times\sigma_0\times(t-16), & t=17,18,...\end{cases}\quad(38)$$

$$\begin{cases}\beta_1=\beta_1^{in-control}, & t=1,2,...,16\\ \beta_1=\beta_1^{in-control}+0.5\times\sigma_1\times(t-16), & t=17,18,...\end{cases}\quad(39)$$

$$\begin{cases}\beta_2=\beta_2^{in-control}, & t=1,2,...,16\\ \beta_2=\beta_2^{in-control}+0.5\times\sigma_2\times(t-16), & t=17,18,...\end{cases}\quad(40)$$

The models for scenario 2 and 3 are similar to scenario 1, and the only difference is that for scenario 2, the parameters which determine the slope of change are $\delta_0 = \delta_1 = \delta_2 = 0.05$. The same parameters for scenario 3 are 0.005.

Finally, a control chart based on the change-point approach and the RST statistic is implemented to monitor a negative binomial profile, as specified in the following steps:

1. A simulated control limit which is not constant over time is determined in a way that leads to a 200 in-control ARL;
2. After obtaining new observations, the $R_{\max,t}$ statistic is calculated and compared to the control limit. If it is below the control limit, new observations will be made, otherwise, control chart signals and the change-point are estimated.

Table 11. Alarm time for different scenarios in real-data analysis (CP is the estimated change-point).

	RST	**MEWMA**	T^2
Scenario 1	17 (CP = 16)	18	17
Scenario 2	23 (CP = 19)	22	23
Scenario 3	61 (CP = 47)	43	76

Three methods, RST, T^2 and MEWMA, are compared under the three mentioned scenarios. The proposed method, bias reduced T^2, is not used in real-data analysis for two following reasons:

1. Phase I of this example is done using IWLS which is a biased estimator. Therefore, using another estimator (bias reduced estimator) in Phase II might be misleading;
2. The bias reduced GLM package in *R* (brglm *R*) currently is developed only for binomial response GLMs.

The time window of the RST method is 100 and the λ of the MEWMA method is 0.2. The results and control charts are given in Figure 2 and Table 11, respectively.

5. Conclusions

The aim of this study is to evaluate the performance of the Rao Score Test (RST) method compared to

Multivariate Exponential Weighted Moving Avrage (MEWMA) and T^2 in monitoring generalized linear profiles in the presence of drift and multiple changes, which can be increasing or decreasing. It also attempts to solve the problem of being ARL-biased in control charts that use IWLS as the estimator.

In this research, the performances of the mentioned methods are evaluated in detecting isotonic drift and multiple changes in the parameters of the binomial profile, and antitonic drift and multiple changes in the parameters of the Poisson profile. For the case of increasing drift and multiple changes, the results are in agreement with Shadman et al. [30] and the RST method outperforms MEWMA and T^2. Furthermore, the change-point estimator matches the change-point closely to the real change-point in large and moderate drifts. For the case of multiple changes, in which the first step is constant and the second step can have ten different sizes, increasing the size of the second step, at first, makes the change-point estimation approach to estimate the change-point close to the real change-point. However, beyond a point, increasing the size of the second step makes the change-point estimation estimate the second step as the change-point most of the time. In decreasing changes, the RST is not superior to the other two methods in all cases. In terms of detecting a small decreasing drift, MEWMA with a small λ value performs better than the RST method. For a moderate change, sometimes MEWMA with moderate λ values shows a better performance, and for a large change, T^2 outperforms the RST method. In detecting small decreasing multiple changes, MEWMA with a small or moderate λ outperforms the RST method. The performance of the change-point estimator in estimating the change-point in the case of antitonic drift and multiple changes was similar to the case of isotonic change. Another important concern might be the effect of the number of predictors and observations on the performance of the RST method. In order to study this issue, The numbers of the predictors and observations for the binomial profile under drift were doubled. Simulation results show that the RST method still works well compared to T^2.

It was noticed that the performance of MEWMA and the T^2 method differs for the case of isotonic and antitonic change, and they are not ARL-unbiased. This happens because MEWMA and T^2 methods do not have the characteristic of being unbiased, which is a result of the biased estimator that they use. Therefore, to deal with this problem, the use of a bias reduced Generalized Linear Model (GLM) approach was proposed as an estimator in T^2. This estimator was proposed by Firth [29], but it had never been used in profile monitoring. The authors results show that the proposed method reduces the bias of the control chart satisfactorily.

A real world example of the resin industry is presented to show the implementation of the methods in Phase II. At first, through statistical analysis, a negative binomial profile, with a log function which relates the mean of the response variable to a second order polynomial of predictor variable, is chosen to model the data. Afterwards, three methods are used to monitor negative binomial profiles in Phase II and under drift. According to the results in the scenario with large drift, RST and T^2 show a similar performance and outperform MEWMA. In the scenario with moderate drift, three methods perform closely. In the scenario with small drift, MEWMA outperforms RST and RST outperforms T^2. As mentioned earlier, a bias reduced GLM package in R (brglm R) currently works only for binomial responses. Developing and testing the package for other applications and using it in profile monitoring could be considered for future study.

References

1. Kang, L. and Albin, S.L. "On-line monitoring when the process yields a linear profile", *Journal of Quality Technology*, **32**(4), pp. 418–426 (2000).

2. Noorossana, R., Saghaei, A., and Amiri, A., *Statistical Analysis of Profile Monitoring*, 1st Edn., In Wiley Series in Probability and Statistics (2011).

3. Niaki, S.T.A. and Khedmati, M. "Identifying the change time of multivariate binomial processes for step changes and drifts", *Journal of Industrial Engineering International*, **9**(1), p. 3 (2013).

4. Kazemzadeh, R.B., Noorossana, R., and Ayoubi, M. "Change point estimation of multivariate linear profiles under linear drift", *Communications in Statistics-Simulation and Computation*, **44**(6), pp. 1570–1599 (2015).

5. Korkas, K.K. and PryzlewiczV, P. "Multiple change-point detection for non-stationary time series using wild binary segmentation", *Statistica Sinica*, pp. 287–311 (2017).

6. Atashgar, K. "Identification of the change point: an overview", *The International Journal of Advanced Manufacturing Technology*, **64**(9–12), pp. 1663–1683 (2013).

7. Aminikhanghahi, S. and Cook, D.J. "A survey of methods for time series change point detection", *Knowledge and Information Systems*, **51**(2), pp. 339–367 (2017).

8. Woodall, W.H., Spitzner, D.J., Montgomery, D.C., and Gupta, S. "Using control charts to monitor process and product quality profiles", *Journal of Quality Technology*, **36**(3), pp. 309–320 (2004).

9. Qiu, P. "Distribution-free multivariate process control based on log-linear modeling", *IIE Transactions*, **40**(7), pp. 664–677 (2008).

10. Yeh, A.B., Huwang, L., and Li, Y.M. "Profile monitoring for a binary response", *IIE Transactions*, **41**(11), pp. 931–941 (2009).

11. Paynabar, K., Jin, J., and Yeh, A.B. "Phase I risk-adjusted control charts for monitoring surgical performance by considering categorical covariates", *Journal of Quality Technology*, **44**(1), pp. 39–53 (2012).

12. Soleymanian, M.E., Khedmati, M., and Mahlooji, H. "Phase II monitoring of binary response profiles", *Scientia Iranica, Transaction E, Industrial Engineering,* **20**(6), p. 2238 (2013).

13. Koosha, M. and Amiri, A. "Generalized linear mixed model for monitoring autocorrelated logistic regression profiles", *The International Journal of Advanced Manufacturing Technology*, **64**(1–4), pp. 487–495 (2013).

14. Amiri, A., Koosha, M., and Azhdari, A. "T2 based methods for monitoring gamma profiles", In *International Conference on Industrial Engineering and Operations Management*, pp. 580–585 (2012).

15. Noorossana, R., Aminnayeri, M., and Izadbakhsh, H. "Statistical monitoring of polytomous logistic profiles in phase II", *Scientia Iranica*, **20**(3), pp. 958–966 (2013).

16. Amiri, A., Koosha, M., Azhdari, A., and Wang, G. "Phase I monitoring of generalized linear model-based regression profiles", *Journal of Statistical Computation and Simulation*, **85**(14), pp. 2839–2859 (2015).

17. Qi, D., Wang, Z., Zi, X., and Li, Z. "Phase II monitoring of generalized linear profiles using weighted likelihood ratio charts", *Computers & Industrial Engineering*, **94**, pp. 178–187 (2016).

18. Sogandi, F. and Amiri, A. "Monotonic change point estimation of generalized linear model-based regression profiles", *Communications in Statistics-Simulation and Computation*, **46**(3), pp. 2207–2227 (2017).

19. Khedmati, M. and Niaki, S.T.A. "Phase II monitoring of general linear profiles in the presence of between-profile autocorrelation", *Quality and Reliability Engineering International*, **32**(2), pp. 443–452 (2016).

20. Noorossana, R., Saghaei, A., Paynabar, K., and Abdi, S. "Identifying the period of a step change in high-yield processes", *Quality and Reliability Engineering International*, **25**(7), pp. 875–883 (2009).

21. Perry, M.B., Pignatiello Jr, J.J., and Simpson, J.R. "Estimating the change point of a Poisson rate parameter with a linear trend disturbance", *Quality and Reliability Engineering International*, **22**(4), pp. 371–384 (2006).

22. Noorossana, R. and Shadman, A. "Estimating the change point of a normal process mean with a monotonic change", *Quality and Reliability Engineering International*, **25**(1), pp. 79–90 (2009).

23. Alaeddini, A., Ghazanfari, M., and Nayeri, M.A. "A hybrid fuzzy-statistical clustering approach for estimating the time of changes in fixed and variable sampling control charts", *Information Sciences*, **179**(11), pp. 1769–1784 (2009).

24. Atashgar, K. and Noorossana, R. "An integrating approach to root cause analysis of a bivariate mean vector with a linear trend disturbance", *The International Journal of Advanced Manufacturing Technology,* **52**(1–4), pp. 407–420 (2011).

25. Ahmadzadeh, F. "Change point detection with multivariate control charts by artificial neural network", *The International Journal of Advanced Manufacturing Technology*, **97**(9–12), pp. 3179–3190 (2018).

26. Zou, C., Zhou, C., Wang, Z., and Tsung, F. "A self-starting control chart for linear profiles", *Journal of Quality Technology*, **39**(4), pp. 364–375 (2007).

27. Paynabar, K., Zou, C., and Qiu, P. "A change-point approach for phase-I analysis in multivariate profile monitoring and diagnosis", *Technometrics*, **58**(2), pp. 191–204 (2016).

28. Maleki, M.R., Amiri, A., and Taheriyoun, A.R. "Identifying the time of step change and drift in phase II monitoring of autocorrelated logistic regression profiles", *Scientia Iranica, Transactions E, Industrial Engineering*, **25**(6), pp. 3654–3666 (2018).

29. Firth, D. "Bias reduction of maximum likelihood estimates", *Biometrika*, **80**(1), pp. 27–38 (1993).

30. Shadman, A., Zou, C., Mahlooji, H., and Yeh, A.B. "A change point method for Phase II monitoring of generalized linear profiles", *Communications in Statistics-Simulation and Computation,* **46**(1), pp. 559–578 (2017).

31. Margavio, T.M., Conerly, M.D., Woodall, W.H., and Drake, L.G. "Alarm rates for quality control charts", *Statistics & Probability Letters,* **24**(3), pp. 219–224 (1995).

32. Hawkins, D.M., Qiu, P., and Kang, C.W. "The change point model for statistical process control", *Journal of Quality Technology*, **35**(4), pp. 355–366 (2003).

33. Heinze, G. and Schemper, M. "A solution to the problem of separation in logistic regression", *Statistics in Medicine*, **21**(16), pp. 2409–2419 (2002).

34. Kosmidis, I. and Firth, D. "A generic algorithm for reducing bias in parametric estimation", *Electronic Journal of Statistics*, **4**, pp. 1097–1112 (2010).

35. Heinze, G. and Schemper, M. "A solution to the problem of separation in logistic regression", *Statistics in Medicine*, **21**(16), pp. 2409–2419 (2002).

36. Shadman, A., Mahlooji, H., Yeh, A.B., and Zou, C. "A change point method for monitoring generalized linear profiles in phase I", *Quality and Reliability Engineering International*, **31**(8), pp. 1367–1381 (2015).

37. Ver Hoef, J.M. and Boveng, P.L. "Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data?", *Ecology*, **88**(11), pp. 2766–2772 (2007).

Biographies

Sahand Hajifar holds an MS degree from Sharif University of Technology, Tehran, in Industrial Engineering. He received his BS degree in Industrial Engineering from Iran University of Science and Technology. His research interests are in the areas of applied statistics, statistical quality control and data mining.

Hashem Mahlooji is Professor of Industrial Engineering at Sharif University of Technology, Tehran, Iran. His research interests include simulations, Monte Carlo and probability and statistics.