Computers & Industrial Engineering 94 (2016) 178–187

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Phase II monitoring of generalized linear profiles using weighted likelihood ratio charts

Dequan Qi [a], Zhaojun Wang [b], Xuemin Zi [c], Zhonghua Li [b,*]

[a] Jilin Medical University, Jilin 132013, PR China
[b] Institute of Statistics and LPMC, Nankai University, Tianjin 300071, PR China
[c] School of Science, Tianjin University of Technology and Education, Tianjin 300222, PR China

ARTICLE INFO

Article history:
Received 23 June 2015
Received in revised form 27 January 2016
Accepted 30 January 2016
Available online 13 February 2016

Keywords:
Exponential weighted moving average
Generalized linear model
Profile monitoring
Statistical process control
Weighted likelihood ratio test

ABSTRACT

In recent years, effective profile monitoring for discrete response variables, such as binary, multinomial, ordinal or Poisson variables, has increasingly attracted interest of researchers in the area of statistical process control. Such quality characteristics are often modeled as special cases of generalized linear models. The objective of this paper is to try to provide a unified framework for Phase II monitoring of generalized linear profiles of which the explanatory variables can be fixed design or random arbitrary design. To this end, a new control chart is developed based on the weighted likelihood ratio test, and it can be readily extended to other generalized profiles or profiles with random predictors if the likelihood function can be obtained. Numerical results and illustrative example show that the proposed control chart has satisfactory in-control run length distribution and stands out at early detection.

1. Introduction

Statistical profile monitoring has increasingly attracted researchers' attention in the area of statistical process control. Early reviews of work in profile monitoring include Woodall, Spitzner, Montgomery, and Gupta (2004) and Woodall (2007), and a recent comprehensive review Woodall and Montgomery (2014) recommend Noorossana, Saghaei, and Amiri (2011) for a more up-to-date overview as the chapters in this book were written by some of the leading researchers in profile monitoring. For profile monitoring, one group of monitoring methods are interested in the case that the response variables are continuous (e.g., Huwang, Wang, Xue, & Zou, 2014; Li & Wang, 2010; Zou, Ning, & Tsung, 2012). At the meantime, it is also quite common to deal with profile monitoring with discrete response variables. As far as we know, the pioneering work is Yeh, Huwang, and Li (2009). Some recent work, such as Amiri, Koosha, and Azhdari (2011), Noorossana, Aminnayeri, and Izadbakhsh (2013), Noorossana, Saghaei, Izadbakhsh, and Aghababaei (2013) and Soleymanian, Khedmati, and Mahlooji (2013), focused on profile monitoring whose response variables are Poisson, ordinal, multinomial and binary variables, respectively.

In the case of discrete response variables, the quality characteristics are often modeled as special cases of generalized linear models (GLM). Amiri, Koosha, Azhdari, and Wang (2015) and Shadman, Mahlooji, Yeh, and Zou (2015) provided a unified framework for Phase I control of generalized linear profiles. Besides the GLM, other types of models have also been used to represent profiles, such as simple linear regression (e.g., Aly, Mahmoud, & Woodall, 2015; Noorossana, Eyvazian, & Vaghefi, 2010; Zhang, Li, & Wang, 2009), nonlinear regression (e.g., Chang & Yadama, 2010; Paynabar, Jin, & Pacella, 2013), multiple regression (e.g., Eyvazian, Noorossana, Saghaei, & Amiri, 2011; Mahmoud, Saad, & El Shaer, 2015), nonparametric regression (e.g., Chuang, Hung, Tsai, & Yang, 2013; Qiu, Zou, & Wang, 2010), mixed models (e.g., Jensen & Birch, 2009; Koosha & Amiri, 2013), and wavelet models (e.g., Chicken, Pignatiello, & Simpson, 2009; Lee, Hur, Kim, & Wilson, 2012). All of the afore-mentioned research, however, only consider the case in which the explanatory variables are fixed from profile to profile. Shang, Tsung, and Zou (2011) provided an aluminium electrolytic capacitor example to illustrate the case in which different profiles often have random explanatory variables and these variables require careful monitoring as well. The major objective of this paper is to try to provide a unified framework for Phase II monitoring of generalized linear profiles of which the explanatory variables can be fixed design or random arbitrary design from profile to profile (the monitoring of the explanatory variables is not concerned). In Phase II, we are interested in

* Corresponding author.
E-mail address: zli@nankai.edu.cn (Z. Li).

http://dx.doi.org/10.1016/j.cie.2016.01.022
0360-8352/

detecting shifts in the model parameters as quickly as possible, while in Phase I, the purpose is to check the quality of historical data and to obtain accurate estimates of the model parameters.

In this paper, we developed a new control chart for generalized linear profile monitoring, which is based on the weighted likelihood ratio test (WLRT). Our proposed approach can be readily extended to other general profiles or profiles with random predictors if the likelihood function can be obtained. Other likelihood ratio test (LRT) based approaches can be found in Shang et al. (2011), Noorossana, Saghaei, et al. (2013) and Soleymanian et al. (2013). The exponentially weighted moving average (EWMA)-GLM control chart proposed by Shang et al. (2011) made use of all available profile samples up to the current time for estimating parameters, and different profiles are weighted as in an EWMA chart. Nevertheless, we found that the EWMA-GLM control chart has very large short-run false alarms, which renders this chart less useful and unacceptable in practice. Shewhart-type control charts (LRT) were proposed by Noorossana, Saghaei, et al. (2013) and Soleymanian et al. (2013). Another EWMA-type control chart (LRT-EWMA) was proposed by Soleymanian et al. (2013). It is shown that the Shewhart-type LRT control charts perform better at detecting large shifts, while the LRT-EWMA control charts perform better at detecting small to medium shifts. However, compared with our WLRT chart, the LRT-EWMA control chart was not as efficient due to the reason that it only used the current profile samples for estimating parameters, and thus the estimators would have considerably large bias and variance. Numerical results show that our proposed WLRT control chart has satisfactory in-control (IC) run length (RL) distribution and stands out at early detection, where RL is the number of points that must be plotted before a point indicates an out-of-control (OC) condition (Montgomery, 2013).

Now we summarize some abbreviated expressions used in this paper for easy reference.

IC	in-control
OC	out-of-control
RL	run length
ARL	average run length
SDRL	standard deviation of the run length
RMI	relative mean index
CED	conditional expected delay
EWMA	exponentially weighted moving average
MEWMA	multivariate exponentially weighted moving average
GLM	generalized linear models
LRT	likelihood ratio test
WLRT	weighted likelihood ratio test

The remainder of this paper is organized as follows. Our proposed methodology is described in detail in Section 2, including the statistical model and WLRT control chart. Section 3 is devoted to comparing the performance of five methods: WLRT, EWMA-GLM (Shang et al., 2011), LRT (Noorossana, Saghaei, et al., 2013; Soleymanian et al., 2013), LRT-EWMA (Soleymanian et al., 2013) and multivariate EWMA (MEWMA) (Soleymanian et al., 2013) charts. An illustrative example is given in Section 4. Section 5 concludes this paper and gives further discussion. The algorithm for obtaining the maximum weighted likelihood estimator is summarized in Appendix A.

2. The proposed WLRT scheme

In this section, we closely follow the notation and formulation used in Dobson (2002) to briefly discuss the generalized linear profiles. We assume that the observations are independent within and between profiles.

2.1. The statistical model

At any time point t, for the ith profile, our statistical model has three components:

1. Response variables $\widetilde{Y}_i = (Y_{i1}, \ldots, Y_{iN})^T$ share the same distribution from the exponential family with a canonical form,
$$f(y_{ij};\theta_{ij}) = \exp[y_{ij}b(\theta_{ij}) + c(\theta_{ij}) + d(y_{ij})],\ i = 1,\ldots,t,\ j = 1,\ldots,N,$$
where $b(\cdot)$, $c(\cdot)$ and $d(\cdot)$ are known functions and θ_{ij}'s are the parameters of the exponential family of distributions.
2. Explanatory variables
$$\widetilde{X}_i = \begin{pmatrix} X_{i1}^T \\ \vdots \\ X_{iN}^T \end{pmatrix} = \begin{pmatrix} x_{i11} & \ldots & x_{i1p} \\ \vdots & & \vdots \\ x_{iN1} & \ldots & x_{iNp} \end{pmatrix},$$
where $X_{ij}^T = (x_{ij1}, \ldots, x_{ijp}),\ i = 1,\ldots,t,\ j = 1,\ldots,N$, can be combined linearly with a coefficient vector $\beta = (\beta_1, \ldots, \beta_p)^T$ (where $p < N$) to form the linear predictor $\eta_{ij} = X_{ij}^T\beta$.
3. A monotone link function $g(\cdot)$ such that
$$g(\mu_{ij}) = \eta_{ij} = X_{ij}^T\beta,\ i = 1,\ldots,t,\ j = 1,\ldots,N,$$
where $\mu_{ij} = E(Y_{ij})$.

Here, the explanatory variables $\widetilde{X}_i$ can be fixed design or random design from profile to profile. We suppose β changes from β_{IC} to another unknown value β_{OC} immediately after an unknown time point τ, which suffices to test the following hypotheses
$$\begin{cases} H_0 : \beta = \beta_{IC}, \\ H_1 : \beta \neq \beta_{IC}, \end{cases}$$
at each time point. Note that β_{IC} can be assumed known for Phase II monitoring.

2.2. Some existing work

From Dobson (2002), we know that, for the ith profile, the log-likelihood function is
$$l_i(\beta) = \sum_{j=1}^{N}[y_{ij}b(\theta_{ij}) + c(\theta_{ij}) + d(y_{ij})].$$
To obtain the maximum likelihood estimator of β, we can use the following estimating equation
$$\mathbf{b}_i^{(m)} = \mathbf{b}_i^{(m-1)} + [\mathfrak{J}_i^{(m-1)}]^{-1}U_i^{(m-1)},$$
where $\mathbf{b}_i^{(m)}$ is the vector of estimates of β at the mth iteration, $[\mathfrak{J}_i^{(m-1)}]^{-1}$ is the inverse of the information matrix, $U_i^{(m-1)}$ is the vector of score. When the difference between successive approximations $\mathbf{b}_i^{(m-1)}$ and $\mathbf{b}_i^{(m)}$ is sufficiently small, $\mathbf{b}_i^{(m)}$ is taken as $\hat{\beta}$ (maximum likelihood estimator of β).

Now we briefly review the LRT, LRT-EWMA and MEWMA control charts, which were proposed by Soleymanian et al. (2013) to monitor binary response profiles in Phase II. In fact, the LRT monitoring statistic can be expressed as
$$LRT_i = 2[l_i(\hat{\beta}_i) - l_i(\beta_{IC})],\ i = 1,2,\ldots.$$

Soleymanian et al. (2013) first normalized the values of LRT_i (here, termed NL_i), and then calculated the statistic of LRT-EWMA control chart by

$$LE_i = \lambda NL_i + (1-\lambda)LE_{i-1},\ i = 1,2,\ldots.$$

where λ represents the smoothing parameter and $LE_0 = 0$.

The MEWMA monitoring statistic can be calculated using three steps. We first calculate the following variable

$$Z_i = (\widetilde{X}_i^T W \widetilde{X}_i)^{1/2}(\hat{\beta}_i - \beta_{IC}),$$

where W is an $n \times n$ diagonal matrix (see details in Soleymanian et al., 2013). Then, we calculate the following statistic

$$E_i = \lambda Z_i + (1-\lambda)E_{i-1},\ i = 1,2,\ldots.$$

Finally, the MEWMA monitoring statistic can be calculated by

$$M_i = E_i^T E_i.$$

We will leave the brief review of EWMA-GLM (Shang et al., 2011) control chart in the next subsection to emphasize the differences of it and our proposed control chart.

2.3. The WLRT control chart

Similar to Qi, Li, Zi, and Wang (in press), up to time point t, the weighted-log-likelihood function can be derived as

$$wl_t(\beta) = \sum_{i=0}^{t} w_i l_i(\beta) = \sum_{i=0}^{t} w_i \left\{ \sum_{j=1}^{N} [y_{ij} b(\theta_{ij}) + c(\theta_{ij}) + d(y_{ij})] \right\}, \quad (1)$$

where the weights $w_0 = (1-\lambda)^t$, $w_i = \lambda(1-\lambda)^{t-i}$, $i = 1,\ldots t$, and $\lambda \in (0,1)$ is a smoothing parameter. Here, the observations of $(\widetilde{X}_0, \widetilde{Y}_0)$ can be viewed as pseudo "sample", which are chosen from the IC dataset. We can obtain the $(\widetilde{X}_0, \widetilde{Y}_0)$ from Phase I study or by simulation such that the difference between $\hat{\beta}_0$ (maximum likelihood estimator of β) and β_{IC} is small. In Sections 3 and 4, we obtain the $(\widetilde{X}_0, \widetilde{Y}_0)$ by simulation, and the random seeds are chosen as 8417 and 123, respectively. Then, we can calculate $l_0(\beta)$ based on $(\widetilde{X}_0, \widetilde{Y}_0)$.

Including the weight w_0 and the observations of $\widetilde{Y}_0$ in Eq. (1) has its own merit:

- It ensures that all of the weights sum to one.
- It confirms that the IC run length distribution of our chart proposed below is satisfactory.

Obviously, $wl_t(\beta)$ makes full use of all available samples up to the current time point t, and the more recent samples receive more weight. An analogous idea, which does not include w_0 and the pseudo "sample", has been used by Shang et al. (2011) for binary profile monitoring. Shang et al. (2011) expanded the WLRT statistics to asymptotically equivalent Wald-type charting statistics using standard Taylor's expansion.

Given the value of λ, we can express the WLRT statistic as

$$W_t = 2[wl_t(\hat{\beta}_t) - wl_t(\beta_{IC})], \quad (2)$$

where $\hat{\beta}_t = \arg\max_\beta wl_t(\beta)$ is the maximum weighted likelihood estimator of β, which can be obtained using the algorithm shown in Appendix A. When the WLRT statistic in Eq. (2) is larger than a prespecified upper control limit, we can declare the model parameter β has deviated from the nominal value, which means the process is OC.

Now we summarize the implementation of our proposed WLRT scheme for profile monitoring as follows:

Table 1
k values such that $\lambda(1-\lambda)^k < \varepsilon$.

λ	ε						
	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
0.05	122	167	211	256	301	346	391
0.1	66	88	110	132	153	175	197
0.2	35	45	55	66	76	86	96

1. Obtain the upper control limit for the WLRT control chart by the bisection searching algorithms to achieve the desired IC average run length (ARL). The ARL is the average number of points that must be plotted before a point indicates an OC condition (Montgomery, 2013).
2. Begin monitoring the profiles in Phase II. After obtaining the new observations, we calculate the monitoring statistics W_t using Eq. (2), and then plot them on the control chart until W_t is larger than the upper control limit.
3. After detecting the shift, we identify and remove the root causes, and then monitor the profiles continuously.

It is worth more detailed explanations that, the proposed WLRT control chart used all the profile data upper to the time t (including the IC and OC profile data), but different from the LRT-EWMA control chart, the WLRT chart only estimated one $\hat{\beta}_t$ rather than estimating $\hat{\beta}_1,\ldots,\hat{\beta}_t$ for different profiles. The LRT-EWMA control chart uses the current profile samples for estimating parameters, while the WLRT control chart gives more weight to more recent samples, which ensures that there is no over-reliance on the most recent data. Let k be a sufficiently large integer such that $\lambda(1-\lambda)^k$ close to zero. As t increase, the weights $w_1, w_2, \ldots$ will be close to zero sequentially. In fact, when t is sufficiently large, we only use the most recent k sets of OC profile data to estimate $\hat{\beta}_t$, which ensures that $\hat{\beta}_t$ is close to β_{OC}.

To alleviate the computation burden, when $t \leqslant k$, we make use of all available samples up to the current time point t to estimate $\hat{\beta}_t$ and calculate W_t. Otherwise, we only use the most recent k sets of sample profile observations, say the observations of $(\widetilde{X}_i, \widetilde{Y}_i)$, $i = t-k+1,\ldots,t$, to estimate $\hat{\beta}_t$ and calculate W_t. It is worth pointing out that, if $\widetilde{X}_{t-k+1} = \cdots = \widetilde{X}_t$, then $\sum_{i=t-k+1}^{t} w_i \mathfrak{I}_i = [1-(1-\lambda)^k]\mathfrak{I}_t$. Some k values such that $\lambda(1-\lambda)^k < \varepsilon$ for given ε are given in Table 1. We will choose the small positive value ε as 10^{-7} for simplicity in the next section.

3. Performance comparisons

In this section, we compare the proposed WLRT control chart with four alternative methods, EWMA-GLM (Shang et al., 2011), LRT (Noorossana, Saghaei, et al., 2013; Soleymanian et al., 2013), LRT-EWMA (Soleymanian et al., 2013) and MEWMA (Soleymanian et al., 2013), to demonstrate the effectiveness of our approach.

Following Amiri et al. (2015) and Shadman et al. (2015), we focus on the Poisson profile in this Section. Similar to Shadman et al. (2015), we assume that the Poisson profiles are as follows:

a. Response variables Y_{ij}'s are independent Poisson random variables, $j = 1,\ldots,10$.
b. Explanatory variables X_{ij} such that
   - when the design points are fixed

$$(X_{i1},\ X_{i2},\ \cdots,\ X_{i10}) = \begin{pmatrix} x_{i11} & x_{i21} & \cdots & x_{i10,1} \\ x_{i12} & x_{i22} & \cdots & x_{i10,2} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0.1 & 0.2 & \cdots & 1.0 \end{pmatrix}.$$

Table 2
IC comparisons.

	h	ARL_0	SDRL	$Q(.10)$	Median	$Q(.90)$	F_{30}
EWMA-GLM0.05	9.90756	370	520	2	154	1040	0.355
EWMA-GLM0.2	11.60040	370	409	4	239	922	0.183
LRT	11.89143	370	369	38	258	850	0.083
LRT-EWMA0.05	0.68820	370	348	63	261	813	0.022
LRT-EWMA0.2	1.47790	370	370	43	256	850	0.069
MEWMA0.05	0.31650	370	338	57	268	826	0.030
MEWMA0.2	1.55000	370	353	43	265	839	0.068
WLRT0.05	0.22710	370	369	47	255	853	0.055
WLRT0.2	1.22170	370	371	40	257	845	0.075

- Otherwise, for the ith profile, nine different design points randomly take values in the above equation.

c. The log link function such that $g(\mu_{ij}) = \log(\mu_{ij}) = X_{ij}^T\beta$, where $\mu_{ij} = E(Y_{ij})$.

Assume further that the IC parameters β_{IC} is $(1,1)^T$, while the OC profile parameter at ith sample profile is equal to

$$\beta_i = \begin{cases} \beta_{IC}, i = 1, \ldots, \tau, \\ \beta_{OC} = \beta_{IC} + \Delta, i = \tau + 1, \ldots, \end{cases}$$

where $\Delta = (\delta_1\sigma_1, \delta_2\sigma_2)^T$, $\delta_1 \neq 0$ or $\delta_2 \neq 0$, and $\sigma_1 = 0.35181$, $\sigma_2 = 0.50947$ are the standard deviation of the maximum likelihood estimator of the profile parameters.

3.1. Comparisons when design points are fixed

For a relatively fair comparison, we adjust the control limits of different charts to make their IC ARL (termed ARL_0) as close as 370 by convention. In comparison of various candidate control charts, ARL is very important and also popular used criterion (Li, Zou, Gong, & Wang, 2014). When the process is IC, a chart with a larger ARL_0 indicates a lower false alarm rate than other charts. When the process is OC, a chart with a smaller OC ARL (termed ARL_1) indicates a better detection ability of process shifts than other charts. Hereafter, we use the notation h to denote the control limit coefficients, and obtain all results in this section based on 5000 replications. A Fortran program is also available from the authors upon request.

We first study IC performance comparison. Zhou, Zou, Wang, and Jiang (2012) pointed out that the IC run length distribution is considered to be satisfactory if it is close to the geometric distribution (Hawkins & Olwell, 1998) or more generally its variation is less than that of a geometric distribution. We use notation $Q(.10)$ and $Q(.90)$ to respectively denote the 10th and 90th percentile of the marginal distribution of the run length. We also study the false-alarm rate for the first 30 observations, $F_{30} = Pr_{IC}(\text{RL} \leqslant 30)$, where $Pr_{IC}(\text{RL} \leqslant 30)$ denotes the probability of run length being less than or equal to 30 when the process is IC. Note that when the run length distribution is geometric, the standard deviation of the run length (SDRL) should be approximately equal to ARL_0, and $Q(.10)$, Median, $Q(.90)$ and F_{30} are about 38, 256, 850 and 0.080 respectively. The IC comparison results are shown in Table 2. Fig. 1 presents the cumulative distribution function of IC run-length distributions of the different charts considered when $t \leqslant 100$.

Theoretically, the IC run length distribution of a Shewhart-type chart is the geometric distribution. The IC run length distribution of an EWMA-type chart with larger smoothing parameter will be closer to the geometric distribution. It is obvious from Table 2 and Fig. 1 that the IC run length distribution of the LRT chart is the most close to the geometric distribution. In addition, the IC performances of the EWMA-GLM, LRT-EWMA, MEWMA and WLRT

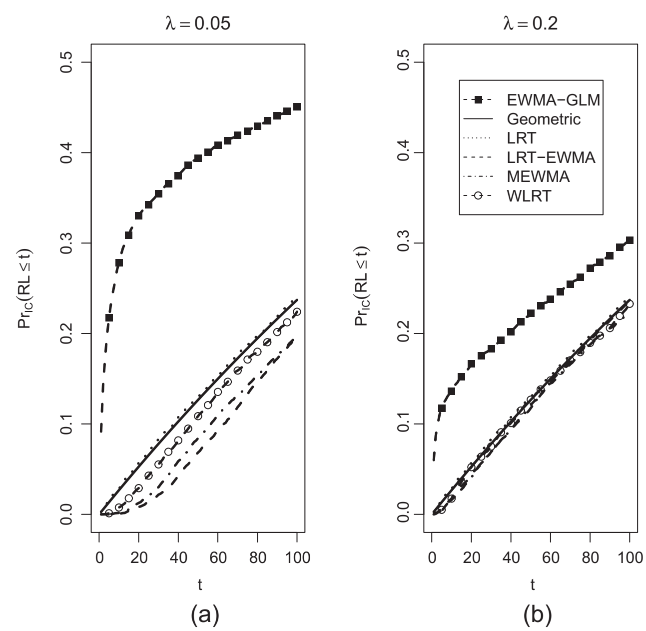

Fig. 1. The in-control cumulative distribution function curves for the Poisson profiles along with Geometric distribution (with expectation 370).

charts depend on the smoothing parameter, i.e., charts with larger parameters perform better. These findings are consistent with the literature. We can also find that the EWMA-GLM control chart has very large short-run false alarms. For example, F_{30} can be as large as 0.183 when $\lambda = 0.2$, and 0.355 when $\lambda = 0.05$. Consequently, the EWMA-GLM control chart is not acceptable in terms of run length distribution because excessive false alarms at early runs will make the detection results unreliable. Moreover, the probabilities of very long runs would decrease, which will lead to the EWMA-GLM control chart having quite small ARL_1 compared to the LRT, LRT-EWMA, MEWMA and WLRT charts. However, this "advantage" is mainly due to very large short-run false alarms, which is consistent with Zhou et al. (2012). Following Zhou et al. (2012), we will also consider the "true" detection capability as another criterion for the OC performance comparison.

Then, we study the OC performance comparison. For the zero state (shift occurs at $\tau = 0$), we compare the ARL_1, the "true" detection capability and the relative mean index (RMI). The "true" detection capability of a chart is reflected by the quantity γ_t, where

$$\gamma_t = Pr_{OC}(\text{RL} \leqslant t) - Pr_{IC}(\text{RL} \leqslant t).$$

Here, γ_t is a reasonable index for OC comparison given that the RL distributions of some charts are far away from geometric, and a control chart with a larger value of γ_t is considered better (Zhou et al., 2012). In order to assess the overall performance of different

Table 3
Comparisons of ARL_1 for the Poisson profiles ($\tau = 0$).

(δ_1, δ_2)	LRT	EWMA-GLM		LRT-EWMA		MEWMA		WLRT	
		$\lambda = 0.05$	0.2	0.05	0.2	0.05	0.2	0.05	0.2
(0.2,0)	201(199)	18.9(19.8)	67.2(71.3)	125(102)	153(151)	95.5(72.3)	365(366)	26.4(17.4)	44.8(39.6)
(0,0.2)	202(209)	18.3(19.3)	64.3(67.8)	124(100)	154(152)	79.2(55.7)	265(255)	26.1(17.3)	45.1(39.5)
(0,0.25)	151(152)	11.8(11.5)	35.0(35.8)	81.9(61.7)	102(99.7)	45.5(25.9)	130(123)	18.4(10.6)	27.5(22.4)
(0.31,0)	106(107)	8.31(7.22)	20.0(19.4)	52.4(35.2)	62.9(58.5)	31.6(15.7)	70.8(61.6)	13.5(6.82)	17.5(13.2)
(0.2,0.2)	64.0(65.7)	5.55(4.28)	10.5(9.21)	31.5(18.2)	33.2(29.3)	20.1(7.71)	28.7(21.2)	9.74(4.19)	10.8(7.05)
(0.5,0)	33.9(33.8)	3.77(2.62)	5.90(4.49)	18.4(8.87)	16.3(12.5)	13.9(4.20)	14.1(8.24)	7.10(2.63)	6.89(3.65)
(0.32,0.32)	16.1(15.4)	2.65(1.62)	3.62(2.44)	11.3(4.87)	8.47(5.55)	10.1(2.50)	8.29(3.63)	5.34(1.73)	4.76(2.15)
(0,0.7)	10.6(10.2)	2.20(1.27)	2.90(1.77)	8.76(3.56)	6.24(3.73)	8.67(1.89)	6.58(2.49)	4.63(1.38)	3.98(1.63)
(0.44,0.44)	5.30(4.72)	1.69(0.86)	2.06(1.11)	5.78(2.22)	3.89(2.08)	6.84(1.30)	4.77(1.45)	3.65(0.99)	3.02(1.07
(0.59,0.59)	2.03(1.45)	1.21(0.44)	1.35(0.57)	3.23(1.13)	2.10(0.97)	4.96(0.76)	3.21(0.74)	2.61(0.64)	2.09(0.62)
(1.0,1.0)	1.01(0.08)	1.00(0.02)	1.00(0.03)	1.22(0.41)	1.02(0.15)	3.00(0.21)	1.99(0.16)	1.49(0.50)	1.07(0.26)
RMI	6.679	0.00	0.993	3.795	3.920	2.932	5.882	0.779	0.930

NOTE: Standard deviations are in parentheses.

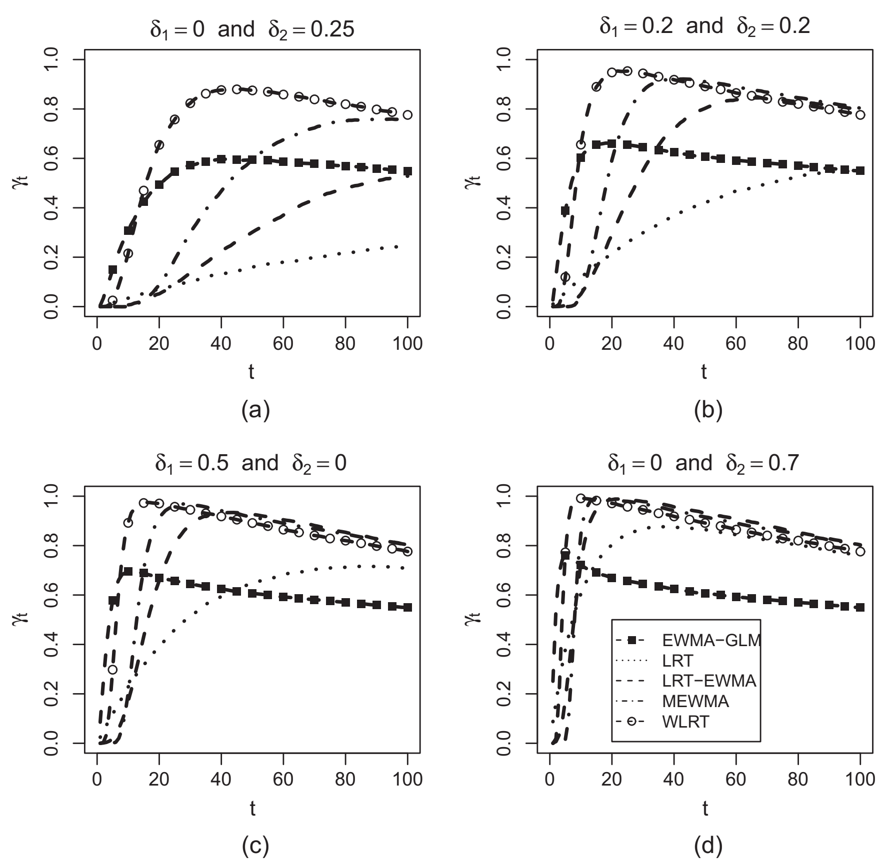

Fig. 2. The "true" detection capability for the Poisson profiles ($\lambda = 0.05$). The legend in the last plot is applicable for all the others.

charts, we compare the RMI values. The RMI index of a control chart, suggested by Han and Tsung (2006), is defined as

$$RMI = \frac{1}{M}\sum_{l=1}^{M}\frac{ARL_{\Delta l} - MARL_{\Delta l}}{MARL_{\Delta l}},$$

where M is the total number of shifts considered, $ARL_{\Delta l}$ is the ARL_1 of the given control chart when detecting a parameter shift of magnitude Δl, and $MARL_{\Delta l}$ is the smallest among all ARL_1 values of the charts considered when detecting the shift Δl. A control chart with a smaller RMI value is considered better in its overall performance (Zhou et al., 2012). As for the steady state, we compare the conditional expected delay (CED) (Kenett & Zacks, 1998; Lee & Jun, 2012) as the detection ability depends on the time point of the change (Sonesson & Bock, 2003). The CED is defined by

$$CED = E[RL - \tau | RL > \tau].$$

A control chart with a smaller CED value is considered better than another one. The comparisons of ARL_1 and RMI values are reported in Table 3.

From Table 3, we can see that the EWMA-GLM ($\lambda = 0.05$) chart outperforms other competitors considering the overall performance. Additionally, the LRT control chart performs better at detecting large shifts, while the LRT-EWMA, MEWMA and WLRT control charts perform better at detecting small to medium shifts.

Table 4
The ARL of WLRT chart with different k ($\tau = 0$).

(δ_1, δ_2)	$\lambda = 0.05$			$\lambda = 0.2$		
	$k = 122$	256	391	35	66	96
(0,0)	368.68(366.19)	369.86(368.53)	369.86(368.53)	369.05(370.83)	369.82(371.28)	369.82(371.28)
(0.2,0)	26.354(17.418)	26.354(17.418)	26.354(17.421)	44.795(39.551)	44.795(39.551)	44.788(39.546)
(0,0.2)	26.060(17.260)	26.060(17.260)	26.060(17.260)	45.047(39.441)	45.064(39.475)	45.064(39.475)
(0,0.25)	18.362(10.631)	18.362(10.631)	18.362(10.631)	27.450(22.351)	27.455(22.368)	27.455(22.368)
(0.31,0)	13.481(6.815)	13.481(6.815)	13.481(6.814)	17.489(13.197)	17.489(13.197)	17.489(13.197)
(0.2,0.2)	9.735(4.192)	9.735(4.192)	9.735(4.191)	10.842(7.047)	10.842(7.047)	10.842(7.047)
(0.5,0)	7.098(2.630)	7.098(2.630)	7.098(2.630)	6.887(3.648)	6.887(3.648)	6.889(3.648)
(0.32,0.32)	5.339(1.725)	5.339(1.725)	5.339(1.725)	4.761(2.148)	4.761(2.148)	4.761(2.148)
(0,0.7)	4.632(1.381)	4.632(1.381)	4.632(1.381)	3.982(1.628)	3.982(1.628)	3.981(1.628)
(0.44,0.44)	3.647(0.993)	3.647(0.993)	3.647(0.993)	3.020(1.069)	3.020(1.069)	3.020(1.069)
(0.59,0.59)	2.614(0.642)	2.614(0.642)	2.614(0.642)	2.091(0.620)	2.091(0.620)	2.091(0.620)
(1.0,1.0)	1.489(0.500)	1.489(0.500)	1.489(0.500)	1.071(0.257)	1.071(0.257)	1.071(0.257)

NOTE: h is same as Table 2, standard deviations are in parentheses.

Table 5
Comparisons of CEDs for the Poisson profiles ($\tau = 50$).

(δ_1, δ_2)	LRT	EWMA-GLM		LRT-EWMA		MEWMA		WLRT	
		$\lambda = 0.05$	0.2	0.05	0.2	0.05	0.2	0.05	0.2
(0.2,0)	201(200)	35.4(23.3)	77.4(72.1)	107(100)	151(151)	98.7(72.9)	364(367)	29.7(19.4)	44.8(40.4)
(0,0.2)	201(208)	35.0(23.3)	74.3(67.5)	107(98.9)	151(151)	80.1(57.8)	261(253)	29.3(19.1)	44.3(40.5)
(0,0.25)	152(152)	24.3(14.0)	42.1(37.2)	68.8(61.4)	98.6(99.3)	47.7(27.6)	127(120)	21.0(11.7)	27.7(22.8)
(0.31,0)	106(108)	17.7(8.92)	25.2(20.2)	43.0(36.4)	60.4(57.8)	33.9(16.4)	69.6(61.8)	15.9(8.14)	17.9(13.7)
(0.2,0.2)	64.2(65.1)	12.7(5.69)	14.3(9.78)	24.2(17.8)	31.5(29.1)	21.9(8.48)	28.4(21.5)	11.5(5.37)	11.2(7.46)
(0.5,0)	34.1(34.5)	9.25(3.87)	8.72(4.82)	13.5(8.79)	15.1(12.5)	15.5(4.98)	14.4(8.59)	8.50(3.65)	7.23(3.94)
(0.32,0.32)	16.5(16.2)	6.93(2.68)	5.71(2.66)	8.07(4.71)	7.86(5.76)	11.2(3.34)	8.58(3.81)	6.47(2.57)	4.99(2.32)
(0,0.7)	10.8(10.5)	5.94(2.23)	4.71(2.02)	6.20(3.53)	5.67(3.85)	9.49(2.69)	6.76(2.69)	5.55(2.16)	4.15(1.85)
(0.44,0.44)	5.28(4.68)	4.71(1.66)	3.53(1.37)	4.07(2.12)	3.42(2.03)	7.66(1.99)	4.97(1.65)	4.42(1.64)	3.13(1.24)
(0.59,0.59)	2.03(1.45)	3.36(1.13)	2.39(0.79)	2.30(1.06)	1.89(0.92)	5.55(1.34)	3.33(0.91)	3.15(1.10)	2.18(0.75)
(1.0,1.0)	1.01(0.08)	1.86(0.57)	1.28(0.45)	1.06(0.24)	1.01(0.12)	3.31(0.75)	1.94(0.45)	1.76(0.57)	1.19(0.39)
RMI	3.332	0.365	0.559	1.176	1.697	1.523	3.067	0.242	0.163

NOTE: Standard deviations are in parentheses.

We can also find that the performance of the LRT-EWMA, MEWMA and WLRT control charts depend on the smoothing parameter, i.e., charts with smaller parameter λ perform better for detecting small shifts, while those with larger parameter λ perform better for detecting larger shifts. Fig. 2 presents the "true" detection capability γ_t of the different charts considered when $t \leqslant 100$. We can see that, when t is small, the EWMA-GLM chart outperforms the other four charts in the sense that its γ_t curve increases much faster. But this advantage diminishes quickly as t becomes large due to its very large false alarms. In general, the WLRT chart performs better than the MEWMA chart, and the MEWMA chart performs better than the LRT-EWMA and LRT charts. We can also find that, the LRT chart performs worst at detecting small and medium shifts.

Recall that we recommend to use the most recent k sets of profile observations when $t > k$. Table 4 shows that, when the integer k is sufficiently large, it has little effect on the performance of the WLRT chart. Table 5 provides the comparison results of CEDs. We discard any series in which a signal occurs before the $(\tau + 1)$th observation. This action coincides with the proposals presented by Zhou et al. (2012) and Hawkins and Olwell (1998). We only present the CED's results when $\tau = 50$ for illustration purpose, and a similar conclusion holds for other cases. It is clear that the performance of the WLRT chart is satisfactory, especially when the shifts are small.

3.2. Comparisons when design points are not fixed

In this subsection, we consider the case in which the explanatory variables are not fixed from profile to profile. Note we will, here, not focus on the monitoring of the explanatory variables themselves. If it is concerned instead, we need change the weighted-log-likelihood function in Eq. (1) correspondingly. To generate the values of the explanatory variables, we first generate an integer j from a discrete uniform distribution over the integers from 1 to 10. Then, we delete the corresponding jth design point X_{ij} from the ten design points $X_{i1}, X_{i2}, \ldots, X_{i10}$. In this way, we get nine different design points. By similar ways, we get other number of different design points. Here, we use the same control limits as those in Table 2. We only present the OC comparison results when $\tau = 0$ and $\tau = 50$ in Tables 6 and 7 respectively for illustration purpose. We find that the performance of the WLRT chart with parameter $\lambda = 0.05$ is still satisfactory, especially when the shifts are small.

Finally, we consider the effects of the number and values of design points on our WLRT control chart. The ARL performances of WLRT chart depending on different number of design points are given in Table 8. The first 50 design points used in Table 8 based on 1 simulation run are shown in Fig. 3. From Table 8, the performance is better when the number of design points is larger.

4. Illustrative example

In this section, we adopt and extend the multinomial logistic regression model discussed by Goeman and le Cessie (2006) as an illustrative example. In the production processes, no product is created quite the same as the others due to the machine equipment, material, environment, operator, and some other reasons

Table 6
Comparisons of ARL_1 when the design points are not fixed ($\tau = 0$).

(δ_1, δ_2)	LRT	EWMA-GLM		LRT-EWMA		MEWMA		WLRT	
		$\lambda = 0.05$	0.2	0.05	0.2	0.05	0.2	0.05	0.2
(0,0)	368(371)	364(520)	367(415)	363(333)	368(370)	335(301)	319(308)	377(368)	369(371)
(0.2,0)	213(216)	21.0(22.4)	78.2(84.4)	134(110)	164(159)	116(91.5)	460(458)	29.5(19.7)	51.0(45.2)
(0,0.2)	208(212)	20.3(22.1)	74.7(79.3)	133(110)	162(157)	95.1(72.5)	312(305)	29.2(19.3)	49.8(44.5)
(0,0.25)	157(158)	13.2(13.1)	40.9(42.6)	90.0(69.1)	111(106)	53.0(32.3)	161(157)	20.6(11.9)	31.2(25.8)
(0.31,0)	114(113)	9.11(8.16)	23.0(22.3)	58.4(40.6)	70.8(66.3)	35.8(18.7)	91.0(83.1)	15.0(7.67)	19.9(15.3)
(0.2,0.2)	71.0(70.9)	6.11(4.78)	12.0(10.8)	34.7(20.3)	38.2(33.7)	22.2(8.81)	34.8(27.0)	10.8(4.69)	12.1(8.06)
(0.5,0)	38.6(38.4)	4.13(2.95)	6.67(5.19)	20.5(10.4)	18.8(14.7)	15.1(4.82)	16.4(10.3)	7.87(2.98)	7.77(4.20)
(0.32,0.32)	18.7(18.2)	2.86(1.80)	4.06(2.82)	12.4(5.56)	9.75(6.65)	10.9(2.85)	9.24(4.39)	5.88(1.93)	5.32(2.45)
(0,0.7)	12.2(12.0)	2.39(1.41)	3.21(2.05)	9.62(3.99)	7.06(4.44)	9.32(2.13)	7.27(2.90)	5.13(1.58)	4.43(1.88)
(0.44,0.44)	6.09(5.42)	1.82(0.95)	2.26(1.25)	6.40(2.51)	4.34(2.42)	7.30(1.43)	5.21(1.67)	4.02(1.12)	3.32(1.20)
(0.59,0.59)	2.31(1.73)	1.27(0.51)	1.45(0.64)	3.52(1.27)	2.31(1.10)	5.26(0.83)	3.45(0.84)	2.87(0.71)	2.28(0.69)
(1.0,1.0)	1.01(0.12)	1.00(0.04)	1.00(0.05)	1.31(0.47)	1.04(0.20)	3.09(0.30)	2.02(0.16)	1.66(0.48)	1.16(0.37)
RMI	6.598	0.00	1.068	3.797	3.944	3.036	6.545	0.823	0.980

NOTE: h is same as Table 2, standard deviations are in parentheses.

Table 7
Comparisons of CEDs when the design points are not fixed ($\tau = 50$).

(δ_1, δ_2)	LRT	EWMA-GLM		LRT-EWMA		MEWMA		WLRT	
		$\lambda = 0.05$	0.2	0.05	0.2	0.05	0.2	0.05	0.2
(0.2,0)	212(213)	38.5(25.8)	88.9(85.1)	115(108)	162(158)	120(93.1)	459(461)	32.3(21.2)	50.8(45.8)
(0,0.2)	209(208)	38.4(25.6)	84.7(79.6)	115(109)	158(155)	94.6(71.2)	305(301)	31.9(20.9)	49.8(45.7)
(0,0.25)	159(161)	26.5(15.8)	48.8(43.6)	75.1(67.6)	108(106)	55.2(34.1)	161(158)	23.0(13.4)	31.2(26.0)
(0.31,0)	113(113)	19.2(9.96)	28.9(23.2)	48.1(40.6)	67.8(65.5)	39.1(19.8)	89.1(82.6)	17.0(8.97)	19.8(15.5)
(0.2,0.2)	70.2(70.0)	13.7(6.29)	16.2(11.6)	26.8(19.9)	36.2(33.2)	24.4(10.0)	35.0(27.7)	12.3(5.90)	12.3(8.38)
(0.5,0)	38.3(39.3)	9.90(4.14)	9.58(5.68)	15.2(10.3)	17.5(15.0)	17.0(5.79)	17.1(10.6)	9.00(3.90)	7.79(4.53)
(0.32,0.32)	18.9(18.6)	7.39(2.89)	6.26(3.08)	8.89(5.27)	8.93(6.67)	12.2(3.72)	9.55(4.61)	6.90(2.78)	5.34(2.60)
(0,0.7)	12.3(12.4)	6.35(2.39)	5.14(2.33)	6.83(3.91)	6.40(4.51)	10.2(2.95)	7.49(3.17)	5.88(2.32)	4.47(2.02)
(0.44,0.44)	6.03(5.42)	4.99(1.79)	3.80(1.51)	4.47(2.34)	3.89(2.42)	8.23(2.15)	5.34(1.87)	4.70(1.73)	3.32(1.34)
(0.59,0.59)	2.30(1.74)	3.56(1.17)	2.53(0.88)	2.49(1.18)	2.06(1.03)	5.93(1.43)	3.59(1.00)	3.31(1.16)	2.29(0.84)
(1.0,1.0)	1.02(0.12)	1.93(0.61)	1.36(0.49)	1.11(0.31)	1.03(0.17)	3.49(0.79)	2.04(0.47)	1.85(0.59)	1.27(0.44)
RMI	3.320	0.363	0.619	1.202	1.739	1.627	3.534	0.236	0.183

NOTE: h is same as Table 2, standard deviations are in parentheses.

Table 8
The ARL of WLRT chart with different number of design points.

(δ_1, δ_2)	The number of design points				
	10	9	8	7	6
(0,0)	370(369)	377(368)	373(358)	367(351)	383(348)
(0.2,0)	26.4(17.4)	29.5(19.7)	33.1(22.5)	38.6(25.9)	45.6(30.4)
(0,0.2)	26.1(17.3)	29.2(19.3)	33.2(22.5)	38.7(26.4)	46.0(30.8)
(0,0.25)	18.4(10.6)	20.6(11.9)	23.4(14.1)	27.3(16.1)	32.4(19.7)
(0.31,0)	13.5(6.82)	15.0(7.67)	17.0(8.90)	19.6(10.0)	23.3(12.4)
(0.2,0.2)	9.74(4.19)	10.8(4.69)	12.2(5.44)	14.0(6.26)	16.5(7.59)
(0.5,0)	7.10(2.03)	7.87(2.98)	8.84(3.39)	10.0(3.83)	11.8(4.67)
(0.32,0.32)	5.34(1.73)	5.88(1.93)	6.60(2.24)	7.52(2.54)	8.76(3.06)
(0,0.7)	4.63(1.38)	5.13(1.58)	5.70(1.82)	6.49(2.11)	7.57(2.53)
(0.44,0.44)	3.65(0.99)	4.02(1.12)	4.47(1.27)	5.05(1.46)	5.86(1.70)
(0.59,0.59)	2.61(0.64)	2.87(0.71)	3.17(0.80)	3.55(0.90)	4.11(1.04)
(1.0,1.0)	1.49(0.50)	1.66(0.48)	1.82(0.43)	1.98(0.43)	2.20(0.51)

NOTE: $\lambda = 0.05,\ \tau = 0,\ h$ is same as Table 2, standard deviations are in parentheses.

(Chen, Chang, & Chen, 2011). Instead of simply classifying qualities into conforming and non-conforming, products can be classified into several classes of quality. Details of the multinomial logistic regression are referred to Dobson (2002) and Hosmer, Lemeshow, and Sturdivant (2013), and please refer Böhning (1992) and Hasan, Zhiyu, and Mahani (2014) as for its algorithm.

Multinomial logistic regression is often used when the response variable is categorical, with more than two categories. Two variants exist: one for nominal and one for ordinal scale outcomes. Here, we consider only the nominal scale version. For ease of exposition, we will suppress the index "i" and "j" which were used in Section 2. Consider a response variable Y with four categories. Let $\pi_1, \ldots, \pi_4$ denote the respective probabilities, with $\pi_1 + \cdots + \pi_4 = 1$. We consider a case with three covariates as follows

$$\begin{cases} \log\left(\frac{\pi_2}{\pi_1}\right) = 2x_1 + \delta x_1^2, \\ \log\left(\frac{\pi_3}{\pi_1}\right) = 2x_2, \\ \log\left(\frac{\pi_4}{\pi_1}\right) = 2x_3, \end{cases}$$

where $x_1,\ x_2$ and x_3 each takes values -1, 0 and 1. At each time point t, we obtained a data set of 25 observations which were taken randomly from each of the $3^3 = 27$ possible combinations of the three covariate values.

Here, we extend the Newton–Raphson method in Hasan et al. (2014) to estimate the model parameters, and adjust the control limits of different charts to make their ARL_0 as close as 370 based on 5000 replicates. The first 20 profiles are generated from the IC ($\delta = 1$) normal operational condition and the remaining profiles are from the OC ($\delta = 1.6$) condition. The smoothing parameter λ is chosen as 0.1 for the LRT-EWMA and WLRT control charts. The LRT, LRT-EWMA and WLRT control charts are constructed in Fig. 4. From Fig. 4, we can see that the performance of the WLRT chart is satisfactory.

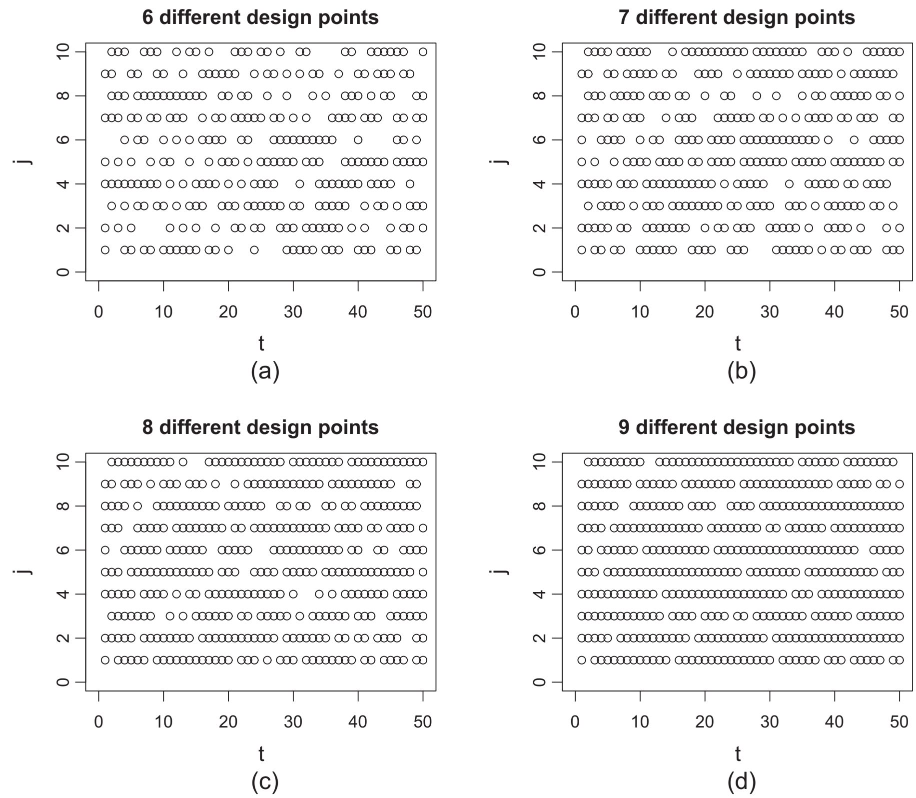

Fig. 3. The first 50 design points used in Table 8 based on 1 simulation run.

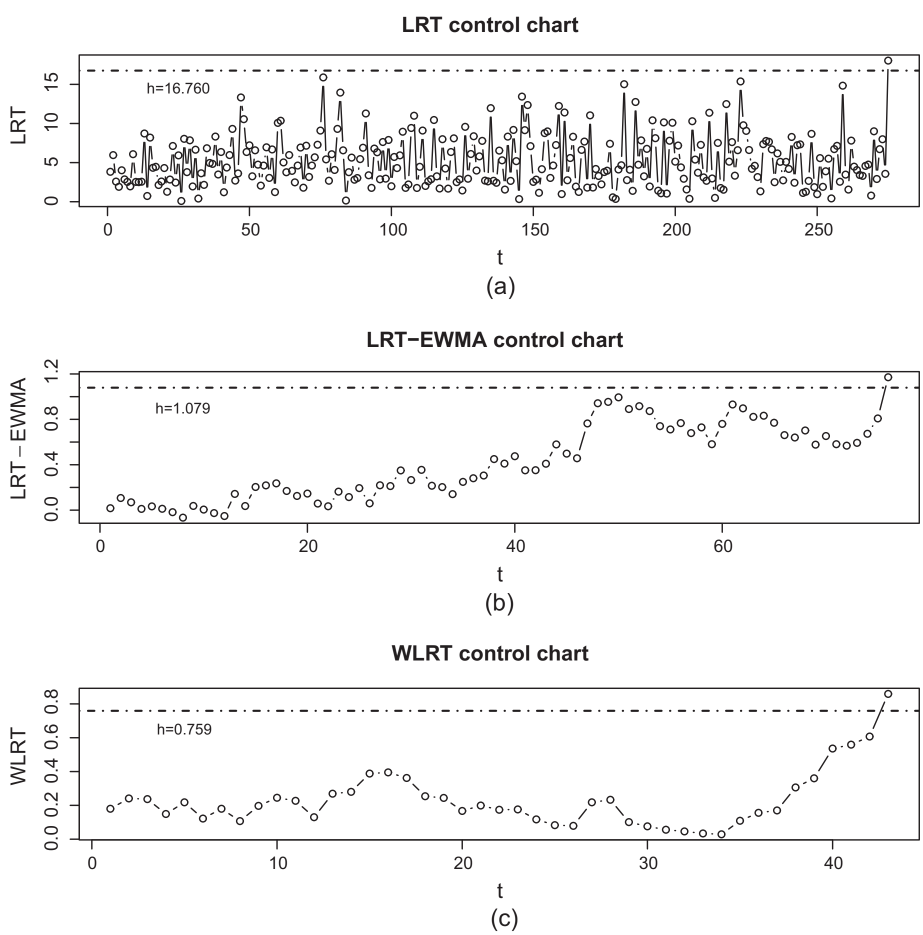

Fig. 4. The LRT, LRT-EWMA and WLRT control charts for the multinomial profiles.

5. Conclusion remarks

In this paper, we proposed a unified framework for Phase II monitoring of generalized linear profiles. In practical applications, it is not uncommon to encounter quality characteristics that are either count data or categorical in nature. Such quality characteristics are often modeled as special cases of generalized linear models. Thus, statistical process control monitoring is important and challenging for generalized linear profiles. The proposed control chart is essentially based on calculating the weighted log-likelihood ratio test statistics, which can be readily extended to other general profiles or profiles with random predictors if the likelihood function can be obtained. Numerical results show that the proposed control chart has satisfactory in-control run length distribution and stands out at early detection.

Our proposed scheme assumes that the observations are independent within and between profiles. The cases when observations are dependent, warrant further investigation.

Acknowledgements

The authors would like to thank the Editor, the Associate Editor and two anonymous referees for their insightful comments, that improve this paper greatly. This paper is supported by the National Natural Science Foundation of China Grants 11571191, 11401573, 11371202, 11431006, 11131002 and 11271205.

Appendix A

In this Appendix, we briefly introduce how to estimate $\hat{\beta}_t$, which is the maximum weighted likelihood estimator of β. Let $\mathfrak{I} = \sum_{i=0}^{t} w_i \mathfrak{I}_i$ and $U = \sum_{i=0}^{t} w_i U_i$. According to Dobson (2002), we can see, if $\widetilde{X}_0 = \cdots = \widetilde{X}_t$, then $\mathfrak{I}_0 = \cdots = \mathfrak{I}_t$, and then $\mathfrak{I} = \mathfrak{I}_1$. The proposed Newton–Raphson approximation for obtaining $\hat{\beta}_t$ proceeds as follows:

(1) Start with the initial values of β, denoted as $\beta^{(0)}$.
(2) Calculate $\mathfrak{I}^{(m)}$ and $U^{(m)}$, by using $\beta^{(m)}$ in the mth iteration.
(3) Update the estimation of β as follows:

$$\beta^{(m+1)} = \beta^{(m)} + [\mathfrak{I}^{(m)}]^{-1} U^{(m)}.$$

(4) Repeat steps (2) and (3) until adequate convergence is achieved as follows:

$$\|\beta^{(m)} - \beta^{(m-1)}\|_1 / \|\beta^{(m-1)}\|_1 \leqslant \epsilon,$$

where ϵ is a given small positive value (e.g., $\epsilon = 10^{-4}$) and $\|\beta\|_1$ denotes L_1 norm, that is, the sum of the absolute values of all elements of β. As such, $\hat{\beta}_t = \beta^{(m)}$ is the desired estimator of β.

References

Aly, A. A., Mahmoud, M. A., & Woodall, W. H. (2015). A comparison of the performance of phase II simple linear profile control charts when parameters are estimated. *Communications in Statistics-Simulation and Computation, 44*(6), 1432–1440.
Amiri, A., Koosha, M., & Azhdari, A. (2011). Profile monitoring for Poisson responses. In *Industrial engineering and engineering management (IEEM), 2011 IEEE international conference on IEEE* (pp. 1481–1484).
Amiri, A., Koosha, M., Azhdari, A., & Wang, G. (2015). Phase I monitoring of generalized linear model-based regression profiles. *Journal of Statistical Computation and Simulation, 85*(14), 2839–2859.
Böhning, D. (1992). Multinomial logistic regression algorithm. *Annals of the Institute of Statistical Mathematics, 44*(1), 197–200.
Chang, S. I., & Yadama, S. (2010). Statistical process control for monitoring non-linear profiles using wavelet filtering and B-spline approximation. *International Journal of Production Research, 48*(4), 1049–1068.
Chen, L-H., Chang, F. M., & Chen, Y-L. (2011). The application of the multinomial control charts for inspection error. *International Journal of Industrial Engineering, 18*(5), 244–253.
Chicken, E., Pignatiello, J. J., & Simpson, J. (2009). Statistical process monitoring of nonlinear profiles using wavelets. *Journal of Quality Technology, 41*(2), 198–212.
Chuang, S. C., Hung, Y. C., Tsai, W. C., & Yang, S. F. (2013). A framework for nonparametric profile monitoring. *Computers & Industrial Engineering, 64*(1), 482–491.
Dobson, A. J. (2002). *An introduction to generalized linear models*. Boca Raton: Chapman & Hall/CRC.
Eyvazian, M., Noorossana, R., Saghaei, A., & Amiri, A. (2011). Phase II monitoring of multivariate multiple linear regression profiles. *Quality and Reliability Engineering International, 27*(3), 281–296.
Goeman, J. J., & le Cessie, S. (2006). A goodness-of-fit test for multinomial logistic regression. *Biometrics, 62*(4), 980–985.
Han, D., & Tsung, F. (2006). A reference-free cuscore chart for dynamic mean change detection and a unified framework for charting performance comparison. *Journal of the American Statistical Association, 101*(473), 368–386.
Hasan, A., Zhiyu, W., & Mahani, A. S. (2014). Fast estimation of multinomial logit models: R package mnlogit. Available from 1404.3177. Available at <http://arxiv.org/abs/1404.3177.pdf>.
Hawkins, D. M., & Olwell, D. H. (1998). *Cumulative sum charts and charting for quality improvement*. New York: Springer-Verlag.
Hosmer, D. W., Jr, Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons.
Huwang, L., Wang, Y. H. T., Xue, S., & Zou, C. (2014). Monitoring general linear profiles using simultaneous confidence sets schemes. *Computers & Industrial Engineering, 68*(1), 1–12.
Jensen, W. A., & Birch, J. B. (2009). Profile monitoring via nonlinear mixed model. *Journal of Quality Technology, 41*(1), 18–34.
Kenett, R. S., & Zacks, S. (1998). *Modern industrial statistics: Design and control of quality and reliability*. Pacific Grove, CA: Duxbury.
Koosha, M., & Amiri, A. (2013). Generalized linear mixed model for monitoring autocorrelated logistic regression profiles. *The International Journal of Advanced Manufacturing Technology, 64*(1), 487–495.
Lee, J., Hur, Y., Kim, S-H., & Wilson, J. R. (2012). Monitoring nonlinear profiles using a wavelet-based distribution-free CUSUM chart. *International Journal of Production Research, 50*(22), 6574–6594.
Lee, S-H., & Jun, C-H. (2012). A process monitoring scheme controlling false discovery rate. *Communications in Statistics-Simulation and Computation, 41*(10), 1912–1920.
Li, Z., & Wang, Z. (2010). An exponentially weighted moving average scheme with variable sampling intervals for monitoring linear profiles. *Computers & Industrial Engineering, 59*(4), 630–637.
Li, Z., Zou, C., Gong, Z., & Wang, Z. (2014). The computation of average run length and average time to signal: An overview. *Journal of Statistical Computation and Simulation, 84*(8), 1779–1802.
Mahmoud, M. A., Saad, A. E. N., & El Shaer, R. (2015). Phase II multiple linear regression profile with small sample sizes. *Quality and Reliability Engineering International, 31*(5), 851–861.
Montgomery, D. C. (2013). *Introduction to statistical quality control* (6th ed.). New York: John Wiley & Sons.
Noorossana, R., Aminnayeri, M., & Izadbakhsh, H. (2013). Statistical monitoring of polytomous logistic profiles in phase II. *Scientia Iranica E, 20*(3), 958–966.
Noorossana, R., Eyvazian, M., & Vaghefi, A. (2010). Phase II monitoring of multivariate simple linear profiles. *Computers & Industrial Engineering, 58*(4), 563–570.
Noorossana, R., Saghaei, A., & Amiri, A. (2011). *Statistical analysis of profile monitoring*. Hoboken, NJ: John Wiley & Sons, Inc..
Noorossana, R., Saghaei, A., Izadbakhsh, H., & Aghababaei, O. (2013). Monitoring multinomial log it profiles via log-linear models. *International Journal of Industrial Engineering & Production Research, 24*(2), 137–142.
Paynabar, K., Jin, J., & Pacella, M. (2013). Monitoring and diagnosis of multichannel nonlinear profile variations using uncorrelated multilinear principal component analysis. *IIE Transactions, 45*(11), 1235–1247.
Qi, D., Li, Z., Zi, X., & Wang, Z. (2015). Weighted likelihood ratio chart for statistical monitoring of queueing systems. *Quality Technology and Quantitative Management* (in press).
Qiu, P., Zou, C., & Wang, Z. (2010). Nonparametric profile monitoring by mixed effects modeling (with discussion). *Technometrics, 52*(3), 265–277.
Shadman, A., Mahlooji, H., Yeh, A. B., & Zou, C. (2015). A change point method for monitoring generalized linear profiles in phase I. *Quality and Reliability Engineering International, 31*(8), 1367–1381.
Shang, Y., Tsung, F., & Zou, C. (2011). Profile monitoring with binary data and random predictors. *Journal of Quality Technology, 43*(3), 196–208.
Soleymanian, M. E., Khedmati, M., & Mahlooji, H. (2013). Phase II monitoring of binary response profile. *Scientia Iranica E, 20*(6), 2238–2246.
Sonesson, C., & Bock, D. (2003). A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society, Series A, 166*(1), 5–21.
Woodall, W. H. (2007). Current research on profile monitoring. *Producão, 17*(3), 420–425.

Woodall, W. H., & Montgomery, D. C. (2014). Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology, 46*(1), 79–94.

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

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

Zhang, J., Li, Z., & Wang, Z. (2009). Control chart based on likelihood ratio for monitoring linear profiles. *Computational Statistics and Data Analysis, 53*(4), 1440–1448.

Zhou, Q., Zou, C., Wang, Z., & Jiang, W. (2012). Likelihood-based EWMA charts for monitoring Poisson count data with time-varying sample sizes. *Journal of American Statistical Association, 107*(499), 1049–1062.

Zou, C., Ning, X., & Tsung, F. (2012). LASSO-based multivariate linear profile monitoring. *Annals of Operations Research, 192*(1), 3–19.