Profile Monitoring with Binary Data and Random Predictors

YANFEN SHANG and FUGEE TSUNG

Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

CHANGLIANG ZOU

Nankai University, Tianjin, China

The purpose of profile monitoring is to check the stability over time of relationships between response variables and one or more explanatory variables. In many applications, categorical response variables are common and a generalized linear model is usually utilized to model this kind of profile for quality improvement. In practice, different profiles often have random explanatory variables and these variables require careful monitoring as well. Statistical process control is important and challenging for monitoring profiles in such situations. A novel control chart is proposed by integrating an exponentially weighted moving-average scheme and a likelihood-ratio test for the parameters of a logistic regression model. This new scheme not only monitors the functional relationship of the profile but also the mean of the explanatory variables. The proposed chart has reasonable computational and implementation complexity and is efficient in detecting shifts. The simulation results show that it performs better than the standard benchmarks in the literatures for the array of simulation examples that we consider. A real example from the electronic industries is used to illustrate the implementation of the proposed approach.

Key Words: Categorical Response; Exponential Weighted Moving Average; Generalized Linear Model; Profile Monitoring; Random Explanatory Variables; Statistical Process Control.

Introduction

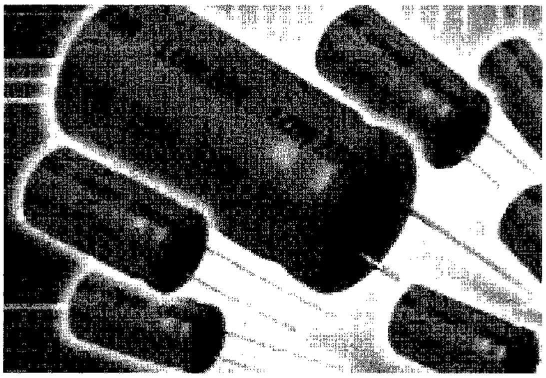
STATISTICAL PROCESS CONTROL (SPC) schemes have been widely applied in various industries. In most applications, the quality of a process can be characterized by the distribution of a single variable or multiple variables, and a variety of univariate and multivariate control schemes have been developed to monitor the process. However, in some applications, the quality of a process must be characterized by a functional relationship between the response variable and one or more explanatory variables, in addition to the distribution of the variables. Therefore, studies on profile monitoring have proliferated in recent years. An extensive discussion of research problems on this topic was given by Woodall et al. (2004).

Studies focusing on simple linear profiles have been particularly common, for instance, Kang and Albin (2000), Kim et al. (2003), Mahmoud and Woodall (2004), Zou et al. (2006; 2007b), among several others. Multiple and polynomial regression profile models were considered by Zou et al. (2007a), Kazemzadeh et al. (2008), Mahmoud (2008), Jensen et al. (2008), and Jensen and Birch (2009). Nonlinear profile models were investigated by Williams et al. (2007). Recently, nonparametric profile monitoring for general profiles has also attracted much attention. The reader is referred to Zou et al. (2008, 2009) and Qiu et al. (2010) for Phase II methods based on nonparametric regression and to Ding et al. (2006), Colosimo et al. (2008), Chicken et al. (2009), and

Ms. Shang is a PhD student in the Department of Industrial Engineering and Logistics Management. Her email address is shangyf@ust.hk.

Dr. Tsung is a Professor in the Department of Industrial Engineering and Logistics Management. He is a Fellow of ASQ. His email address is season@ust.hk.

Dr. Zou is an Assistant Professor in LPMC and Department of Statistics. His email address is chlzou@yahoo.com.cn. Dr. Zou is the author to whom all corresondence should be addressed.



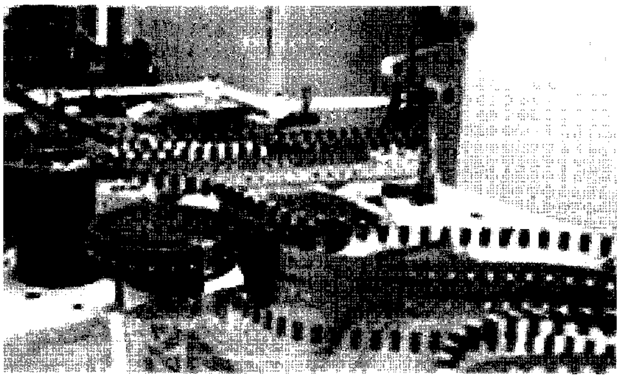


FIGURE 1. Aluminium Electrolytic Capacitors.

Zhang and Albin (2009) for procedures using various dimension-reduction techniques, such as wavelet transformations and independent component analysis. A recent review of the literature was given by Woodall (2007).

All the above-mentioned control schemes for monitoring linear and/or nonlinear profiles require the fundamental assumption that the response variables are continuous. However, due to practical restrictions, e.g., time, cost, or intrinsic characteristics of the variables, often qualitative response variables are easily collected for on-line monitoring. For instance, on some production lines, each item is inspected and classified as conforming or nonconforming, according to some predefined specification on its quality characteristic. Similarly, a service level can also be assessed as satisfactory or unsatisfactory. In such situations, the observed qualitative responses are typically related to some quantitative predictor variables. The profile to be investigated is therefore a functional relationship between a binary (or binomial) response variable and one or more continuous predictor variables. One real example is the manufacture of an aluminium electrolytic capacitor (AEC), which is displayed in Figure 1. The quality of the AEC is inspected based on the specifications of some variables, and the inspection output is "pass" or "fail". Besides the binary output, the quality of AECs is related to the levels of some predictor variables, e.g., the dissipation factor and the leakage current. The functional relationship between the binary output and these continuous predictor variables is critical to maintaining process and product quality and should be monitored (we investigate this example further in the next section), as we propose in this paper.

With respect to monitoring categorical data, besides conventional charts, such as p and np charts, various types of charting schemes have been developed, such as Steiner (1998), Reynolds and Stoumbos (2000), and Somerville et al. (2002), etc. However, regarding profile-monitoring schemes for categorical data, few studies have been conducted. Steiner et al. (2000) proposed risk-adjusted cumulative-sum charts to monitor surgical performance, which are applicable to a logistic regression model, but these charts were purposefully developed for surgical outcomes that follow Bernoulli distributions and assumed the parameter's value after shifts occurred was known, i.e., a known odds ratio. Gurevich and Vexler (2005) also considered changes in parameters of logistic regression models, but they studied how to estimate a change point. In the literature, we have not found any research mainly focusing on Phase II profile monitoring in cases where the response variables are categorical. Yeh et al. (2009) proposed Phase I profile-monitoring schemes for binary responses that could be represented by a logistic regression model. They modeled the relationship between the binary response and explanatory variables using a logistic regression model and studied how to extend the classical T^2-chart for monitoring profiles with continuous data to logistic regression profiles.

In Phase I, a set of process data is gathered and analyzed. Any unusual "patterns" in the data lead to adjustments and fine tuning of the process. Once all such assignable causes are accounted for, we are left with a clean set of data, gathered under stable operating conditions and illustrative of the in-control (IC) process performance. This dataset, which is referred to as the IC dataset, is then used for estimating certain IC parameters of the process. In Phase II SPC, the estimated IC process parameters are used, and the major issue of this phase is fast detection of shifts in the profiles. Besides the fundamental difference between the Phase I profile monitoring considered in their paper and the Phase II monitoring considered here, their approach assumes that the explanatory variables are fixed from profile to profile. These assumptions are (approximately) valid in certain calibration applications in manufacturing industries. In other applications, however, they may be invalid (Qiu and Zou (2009)). Specifically, when data acquisition adopts a random-design scheme, design points within a profile would be independently and identically distributed (i.i.d.) random variables from a given distribution (Qiu and Zou (2009)). In a random-design scheme, the values of predictors in each profile sample would be different. Therefore, the existing schemes may not be applicable in such a case, and how to efficiently use the data from a random design scheme in the Phase II stage needs attention. Moreover, in such situations, the explanatory variables' observations themselves require careful monitoring and control, along with the profile monitoring. This issue is unique to the random-design profile situation. Phase II profile monitoring in such cases is particularly challenging for fast detection of shifts and is the focus of this paper.

In this paper, we utilize a logistic regression model, a type of generalized linear model (GLM), to represent the functional relationship between the binary response and explanatory variables, which are assumed random with specified in-control distributions. Under this premise, a control scheme is proposed based on exponentially weighted moving average (EWMA) process-control schemes. This control scheme is able to simultaneously monitor for shifts in the logistic regression parameters as well as in the means of the explanatory variables. The remainder of this paper is organized as follows: We elaborate on the AEC example to motivate this research in the next section. After that, our proposed methodology is described in detail. The average run length (ARL) performance is then thoroughly investigated via Monte Carlo simulation. Following that, the motivating example, which has a profile that is fit well by a logistic regression model, is used to illustrate the implementation of the proposed approach step by step. Finally, several remarks conclude the article. Certain technical details are provided in the Appendix.

A Motivating Example

We use an example taken from the manufacture of an AEC (provided by ENW Electronics Ltd.; see Figure 1) to motivate this research. During the process, raw materials, including anode aluminum foil, cathode aluminum foil, guiding pin, electrolyte sheet, plastic cover, aluminum shell, and plastic tube, are transformed into AECs with given specifications. The quality of the unfinished AEC products (or capacitor elements) in terms of appearance and functional performance is inspected by sampling. The inspection result will either be "pass" or "fail". During the process, some important characteristics in the specification of AECs, such as leakage current (LC) and dissipation factor (DF), are automatically measured by an electronic device at some given measuring voltage, frequency, and temperature.

The number of defective capacitors in a certain sample of size n, denoted by y, is an obvious quality measurement. The current industrial practice is usually to monitor the mean change of this variable. However, as mentioned above, in the on-line process, the DF and LC are also collected for each capacitor. These two variables are usually randomly distributed and both affect the defective rate of AECs to a certain extent. In this example, $n = 1$ and $y = 1$ or 0. If we denote the predictor variables DF and LC by x_1 and x_2, the sample data is comprised of (y, x_1, x_2). The relationship between y and x_1, x_2 can be modeled as a logistic regression model with binary response,

$$\text{logit}(p) = \alpha + \beta_1 x_1 + \beta_2 x_2,$$

where p is the defect rate and it is assumed that $y \sim \text{Bernoulli}(p)$. To estimate the model parameters, α, β_1, and β_2, a dataset of size N, $\{y_i, x_{1i}, x_{2i}\}_{i=1}^{N}$ are required. The changes in the mean of x_1 and x_2 indicate the changes of DF and LC values of products, and the changes in α, β_1 and β_2 indicate that the relationship between the defect rate and the DF and LC of products changes, which indicates that special causes may have occurred. For example, cosmetic defects of AECs may induce a shift of the functional relationship. Therefore, jointly monitoring the relationship between y and $\{x_1, x_2\}$ and the mean of $\{x_1, x_2\}$ may give more complete information for effective monitoring and diagnosis and may result in better quality improvement. In the remainder of this paper, we propose an SPC scheme to monitor such a profile in Phase II and give a step-by-step demonstration of how to implement the proposed scheme in practice in a later section.

Methodology

Profile Model and Assumptions

In this subsection, we describe the modeling of the profile with a GLM regression model and random predictor variables. We then elaborate on the binary (binomial) responses and the corresponding logistic regression as they are relevant to our application and as an example to illustrate our method. The extension to general cases will be discussed in the final section.

Assume that, for the jth ($j \geq 1$) random profile sample collected over time, we have the observations $(\tilde{\mathbf{X}}_j, \mathbf{y}_j)$, where $\mathbf{y}_j = (y_{j1}, \ldots, y_{jN})$ is an N-variate response vector and $\tilde{\mathbf{X}}_j$ is an $N \times q$ regressor matrix. N is the sample size of profiles, which is consistent with the work of Kang and Albin (2000), Kim et al. (2003), and Zou et al. (2007a). It is assumed that the process observations are collected over time and follow the profile model

$$\operatorname{logit}(p_{ji}) = \alpha_j + \mathbf{x}_{ji}^{\mathrm{T}}\boldsymbol{\beta}_j, \qquad (1)$$

$i = 1, \cdots, N$, $j = 1, \ldots, \tau, \tau + 1, \ldots$, where τ is an unknown change point, y_{ji} is the ith response observation on the jth random profile, $\mathbf{x}_{ji}^{\mathrm{T}}$ denotes the ith row of $\tilde{\mathbf{X}}_j$ (such as x_1 and x_2 in the above example), α_j is the intercept parameter, and $\boldsymbol{\beta}_j = (\beta_{1j}, \ldots, \beta_{qj})^{\mathrm{T}}$ is a q-dimensional parameter vector. Here y_{ji} is assumed to be drawn from a binomial (or Bernoulli) distribution, with parameter p_{ji}, denoted $y_{ji} \sim \text{Binomial}(n_{ji}, p_{ji})$, where n_{ji} is the sample size for the ith observation of the jth profile. Note that, in the AEC example, $n_{ji} = 1$ and y_{ji} is a binary response. p_{ji} represents the ith defect rate of the products in the jth profile sample. Typically, when one group or batch of products is generated according to a particular setting of predictor variables, such as temperature and pressure (represented by $\mathbf{x}_{ji}$), n_{ji} would be greater than one; when the setting of predictor variables is unique to one product, as that in the motivating example, $n_{ji} = 1$. In addition, we also assume $\mathbf{x}_{ji} \sim N_q(\boldsymbol{\mu}_j, \boldsymbol{\Sigma})$ in this paper.

It is assumed that, after an unknown change point τ, there is a change in the intercept and/or coefficient and/or the mean vector of explanatory variables. Say

$$\alpha_j = \alpha_{(0)},\ \boldsymbol{\beta}_j = \boldsymbol{\beta}_{(0)},\ \boldsymbol{\mu}_j = \boldsymbol{\mu}_{(0)} \quad \text{for } j \leq \tau,$$
$$\alpha_j = \alpha_{(1)},\ \boldsymbol{\beta}_j = \boldsymbol{\beta}_{(1)},\ \boldsymbol{\mu}_j = \boldsymbol{\mu}_{(1)} \quad \text{for } j > \tau,$$

and $\alpha_{(0)} \neq \alpha_{(1)}$ and/or $\boldsymbol{\beta}_{(0)} \neq \boldsymbol{\beta}_{(1)}$ and/or $\boldsymbol{\mu}_{(0)} \neq \boldsymbol{\mu}_{(1)}$. Here we shall assume that $N > q + 1$, which is not restrictive and can be easily satisfied in practical applications.

The maximum likelihood estimations (MLEs) of the model parameters $\boldsymbol{\xi} = (\alpha, \boldsymbol{\beta}^{\mathrm{T}})^{\mathrm{T}}$ can be obtained via the standard GLM procedure by using the iterative weighted least square (IWLS) method. Details on how to obtain the MLE $\widehat{\boldsymbol{\xi}}$ are presented in Appendix A. The index "j" is suppressed here for ease of exposition. It can be seen that, under the IC model, $\widehat{\boldsymbol{\xi}} \mid \mathbf{X}$ asymptotically follows the multivariate normal distribution

$$\widehat{\boldsymbol{\xi}} \mid \mathbf{X} \xrightarrow{\mathcal{L}} N_{q+1}\left(\boldsymbol{\xi}_0, (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\right),$$

in which $\boldsymbol{\xi}_0 = (\alpha_{(0)}, \boldsymbol{\beta}_{(0)}^{\mathrm{T}})^{\mathrm{T}}$, $\mathbf{X} = (\ddot{\mathbf{x}}_1, \ldots, \ddot{\mathbf{x}}_N)^{\mathrm{T}}$ is an $N \times (q+1)$ matrix and $\ddot{\mathbf{x}}_i = (1, \mathbf{x}_i^{\mathrm{T}})^{\mathrm{T}}$, and $\mathbf{W} = \operatorname{diag}\{w_1, \ldots, w_N\}$ denotes the GLM weight functions, where $w_i = [n_i p_i(1 - p_i)]$.

Control Schemes for Monitoring the Profile Model (1)

In this section, we propose a control scheme based on the profile model (1), in which the $(q+1)$-variate parameter vector $\boldsymbol{\xi}$ and the q-variate mean vector $\boldsymbol{\mu}$ can be simultaneously monitored. Recall the model (1) and associated notation. The joint log likelihood of $(\tilde{\mathbf{X}}_j, \mathbf{y}_j)$ can be expressed as (see Appendix B for details)

$$\begin{aligned} l_j = \sum_{i=1}^{N} \log \mathrm{C}_{n_{ji}}^{y_{ji}} &+ y_{ji}(\alpha_j + \mathbf{x}_{ji}^{\mathrm{T}}\boldsymbol{\beta}_j) \\ &- n_{ji} \log\left[1 + \exp\{(\alpha_j + \mathbf{x}_{ji}^{\mathrm{T}}\boldsymbol{\beta}_j)\}\right] \\ &- \frac{1}{2}\log|2\pi\boldsymbol{\Sigma}| \\ &- \frac{1}{2}(\mathbf{x}_{ji} - \boldsymbol{\mu}_j)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{ji} - \boldsymbol{\mu}_j). \end{aligned} \qquad (2)$$

The MLEs of the profile parameters and the mean of explanatory variables based on Equation (2), defined as $(\tilde{\boldsymbol{\xi}}_j, \tilde{\boldsymbol{\mu}}_j) = \operatorname{argmax}_{\boldsymbol{\xi},\boldsymbol{\mu}} l_j$, can then be obtained. It is straightforward to see that the MLE of $\boldsymbol{\mu}$ is $\tilde{\boldsymbol{\mu}}_j = \sum_{i=1}^{N} \mathbf{x}_{ji}/N$ and $\tilde{\boldsymbol{\xi}}_j$ can be obtained via the procedure in Appendix A.

Based on the MLEs, one straightforward on-line detection method is to use the current profile estimates to construct two charts for parameters $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$, respectively. However, each chart has a statistic that must be updated and plotted and has a control limit and type I error to be decided. Therefore, the setup of a scheme including one more chart is complicated and difficult (Zou et al. (2007a)). Another straightforward method is to construct a single

Shewhart-type T^2 chart. However, this would be very inefficient for detecting moderate and small changes because it completely ignores the profile samples. As an alternative, we may consider the EWMA chart, as in Kim et al. (2003), Zou et al. (2007a), etc. A natural idea is to first obtain estimates of $(\boldsymbol{\xi}, \boldsymbol{\mu})$ for each profile and then apply the multivariate EWMA chart (Lowry et al. (1992)) to those estimates. However, this naive approach may not be efficient either, because only N random explanatory observations are used for estimating parameters in individual profiles, and thus the estimators would have considerable bias and variance.

Alternatively, in order to monitor the functional relationship and the mean of the explanatory variables efficiently, we propose a new scheme for monitoring the profile, based on the exponentially weighted joint log likelihood at time t,

$$\begin{aligned} l_{t,\lambda}(\boldsymbol{\xi}, \boldsymbol{\mu}) = \lambda \sum_{j=1}^{t} (1-\lambda)^{(t-j)} \\ \times \sum_{i=1}^{N} \log \mathrm{C}_{n_{ji}}^{y_{ji}} + y_{ji}(\alpha + \mathbf{x}_{ji}^{\mathrm{T}}\boldsymbol{\beta}) \\ - n_{ji} \log\left[1 + \exp\{(\alpha + \mathbf{x}_{ji}^{\mathrm{T}}\boldsymbol{\beta})\}\right] \\ - \frac{1}{2}\log|2\pi\boldsymbol{\Sigma}| \\ - \frac{1}{2}(\mathbf{x}_{ji} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{ji} - \boldsymbol{\mu}), \quad (3) \end{aligned}$$

where λ is a weighting parameter. Obviously, the $l_{t,\lambda}(\boldsymbol{\xi}, \boldsymbol{\mu})$ in Equation (3) makes use of all available profile samples up to the current time t, and different profiles are weighted as in an EWMA chart (i.e., more recent profiles are given more weight and the weight changes exponentially over time). Then the maximum weighted likelihood estimator (MWLE), defined as $(\widehat{\boldsymbol{\xi}}_t, \widehat{\boldsymbol{\mu}}_t) = \mathrm{argmax}_{\boldsymbol{\xi},\boldsymbol{\mu}}\, l_{t,\lambda}(\boldsymbol{\xi}, \boldsymbol{\mu})$, can be obtaind via the IWLS method (see Appendix C).

After obtaining the MWLE $(\widehat{\boldsymbol{\xi}}_t, \widehat{\boldsymbol{\mu}}_t)$, the charting statistics is defined as

$$\begin{aligned} lr_t = (\widehat{\boldsymbol{\xi}}_t - \boldsymbol{\xi}_0)^{\mathrm{T}}\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_t}^{-1}(\widehat{\boldsymbol{\xi}}_t - \boldsymbol{\xi}_0) \\ + \frac{N(2-\lambda)}{\lambda}(\mathbf{E}_t - \boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{E}_t - \boldsymbol{\mu}_0), \quad (4) \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_t} &= \frac{\lambda}{2-\lambda}(\widehat{\mathbf{X}}_t^{\mathrm{T}}\widehat{\mathbf{W}}_t\widehat{\mathbf{X}}_t)^{-1}, \\ \widehat{\mathbf{X}}_t &= (\mathbf{X}_1^{\mathrm{T}}, \ldots, \mathbf{X}_t^{\mathrm{T}})^{\mathrm{T}}, \\ \widehat{\mathbf{z}}_t &= (\mathbf{z}_1^{\mathrm{T}}, \ldots, \mathbf{z}_t^{\mathrm{T}})^{\mathrm{T}}, \\ \widehat{\mathbf{W}}_t &= \mathrm{diag}\{\widehat{\mathbf{w}}_1, \ldots, \widehat{\mathbf{w}}_t\} \\ \widehat{\mathbf{w}}_j &= \mathrm{diag}\{\widehat{w}_{j1}, \ldots, \widehat{w}_{jN}\}, \\ \widehat{w}_{ji} &= \lambda(1-\lambda)^{t-j} n_{ji} p_{ji}(1-p_{ji}), \\ \mathbf{E}_t &= \lambda\overline{\mathbf{x}}_t + (1-\lambda)\mathbf{E}_{t-1}, \quad t = 1, 2, \ldots, \end{aligned}$$

$\mathbf{E}_0 = \boldsymbol{\mu}_0$ is the starting vector, and $\overline{\mathbf{x}}_t = \sum_{i=1}^{N} \mathbf{x}_{ti}/N$. The chart signals when $lr_t > L_M$, where L_M is the control limit according to a specific IC ARL, ARL_0. This charting statistic is an approximation of the likelihood-ratio test statistic based on the weighted likelihood function (3). Details of the derivation are presented in Appendix C. After detecting the shift, the hypothesis testing methods can be used to diagnose where the shift occurs. The detailed diagnosis scheme is not considered in this paper but certainly deserves future research. Hereafter, this chart is referred to as the EWMA-GLM control chart.

Performance Assessment

In this section, we investigate the performance of this new scheme (EWMA-GLM) in detecting shifts in profile parameters and the mean of random explanatory variables through Monte Carlo simulations. It is challenging to compare the proposed method with alternative methods because there is no obviously comparable alternative method in the literature. Here, we consider the Shewhart-type T^2 scheme mentioned at the beginning of the previous subsection. To be specific, we define the charting statistic as

$$\begin{aligned} T_{St}^2 = (\tilde{\boldsymbol{\xi}}_t - \boldsymbol{\xi}_0)^{\mathrm{T}}\boldsymbol{\Sigma}_{\tilde{\boldsymbol{\xi}}_t}^{-1}(\tilde{\boldsymbol{\xi}}_t - \boldsymbol{\xi}_0) \\ + N(\overline{\mathbf{x}}_t - \boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}_t - \boldsymbol{\mu}_0), \quad t = 1, 2, \ldots \end{aligned} \quad (5)$$

where $\tilde{\boldsymbol{\xi}}_t$ is the MLE obtained as in Appendix A, $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{\xi}}_t} = (\mathbf{X}_t^{\mathrm{T}}\mathbf{W}_t\mathbf{X}_t)^{-1}$, and $\mathbf{X}_t^{\mathrm{T}}$, and $\mathbf{W}_t$ are the corresponding matrices defined at the beginning of the previous section for the tth profile sample. The chart signals when $T_{St}^2 > L_S$, where L_S is the control limit chosen to achieve a specific ARL_0. We call this chart the Shewhart-GLM chart.

Another possible alternative method to compare the proposed method against is the naive EWMA chart mentioned in the previous subsection, which is described as follows:

$$\begin{aligned} T_{NEt}^2 = (\mathbf{E}_{\xi t} - \boldsymbol{\xi}_0)^{\mathrm{T}}\Sigma_{\mathbf{E}_{\xi t}}^{-1}(\mathbf{E}_{\xi t} - \boldsymbol{\xi}_0) \\ + \frac{N(2-\lambda)}{\lambda}(\mathbf{E}_t - \boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{E}_t - \boldsymbol{\mu}_0), \\ t = 1, 2, \ldots \end{aligned}$$

where

$$\Sigma_{\mathbf{E}_{\xi t}} = \lambda^2 \Sigma_{\tilde{\xi}_t} + (1-\lambda)^2 \Sigma_{\mathbf{E}_{\xi(t-1)}},$$
$$\mathbf{E}_{\xi t} = \lambda \tilde{\xi}_t + (1-\lambda)\mathbf{E}_{\xi(t-1)},$$

and $\Sigma_{\tilde{\xi}_t}$ is defined as for Equation (5). We refer to this naïve EWMA scheme as NEWMA-GLM hereafter.

The monitoring performance of the control schemes in this section is evaluated through ARL comparisons. Three cases are studied here: (1) the performance of EWMA-GLM is compared with that of Shewhart-GLM in detecting shifts in model parameters and explanatory variables, (2) the performance of EWMA-GLM is investigated for different smoothing parameters λ, (3) the performance of EWMA-GLM is compared with that of NEWMA-GLM in detecting shifts in different parameters. Without loss of generality, we assume the time τ at which the shifts initially occur is 40 and only consider the sustained shifts for the remaining samples after the changepoint τ. In this section, only the case of $\text{ARL}_0 = 200$ is considered. In addition, the underlying IC model considered here is model (1) with the two explanatory variables $\mu_0 = (0,0)^{\text{T}}$ and $\Sigma = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$. The other parameters are assumed to be $\beta = (1.0, 1.0)^{\text{T}}$ and $\alpha = -2.2$, which make the expectation of the defective rate p_i approximately equal to 0.1. For the parameter n_i in the binomial distribution and N in the profile, the values 30 and 20 are used, respectively. The control limits of different control schemes are obtained by simulation to approximately achieve the given IC ARL. The out-of-control (OC) ARL results for detecting different magnitudes of shifts in different parameters are evaluated, and all of the results are obtained by running 5,000 simulations.

We compare the OC ARLs of the proposed EWMA-GLM scheme with those of Shewhart-GLM for detecting shifts in α, β_1, and μ_1. The smoothing

TABLE 1. ARL Comparisons Between EWMA-GLM with $\lambda = 0.2$ and the Shewhart-GLM Scheme in Detecting Various Shifts

	δ in α						
Chart	0.05	0.06	0.07	0.10	0.20	0.30	0.50
EWMA-GLM	68.3	52.4	39.5	19.5	5.68	3.23	1.78
	(65.3)	(49.4)	(35.3)	(15.1)	(2.87)	(1.28)	(0.61)
Shewhart-GLM	126	113	100.8	65.6	16.1	4.79	1.31
	(125)	(112)	(101)	(64.2)	(15.5)	(4.22)	(0.64)
	δ in β_1						
Chart	0.15	0.20	0.25	0.30	0.50	1.00	1.50
EWMA-GLM	79.0	52.3	34.8	23.8	6.66	3.31	2.13
	(74.9)	(48.6)	(30.3)	(19.2)	(3.62)	(1.45)	(0.86)
Shewhart-GLM	168	146	119	98.6	24.2	5.27	1.85
	(170)	(146)	(119)	(97.1)	(23.5)	(4.66)	(1.23)
	δ in μ_1						
Chart	0.025	0.030	0.035	0.07	0.12	0.15	0.20
EWMA-GLM	85.2	66.4	50.8	13.3	5.44	4.04	2.88
	(80.4)	(62.4)	(47.1)	(9.10)	(2.51)	(1.65)	(1.02)
Shewhart-GLM	171	164	149	71.6	20.9	10.4	4.02
	(171)	(167)	(151)	(70.1)	(20.3)	(10.1)	(3.45)

Note: Values in parentheses are the standard deviations of ARLs.

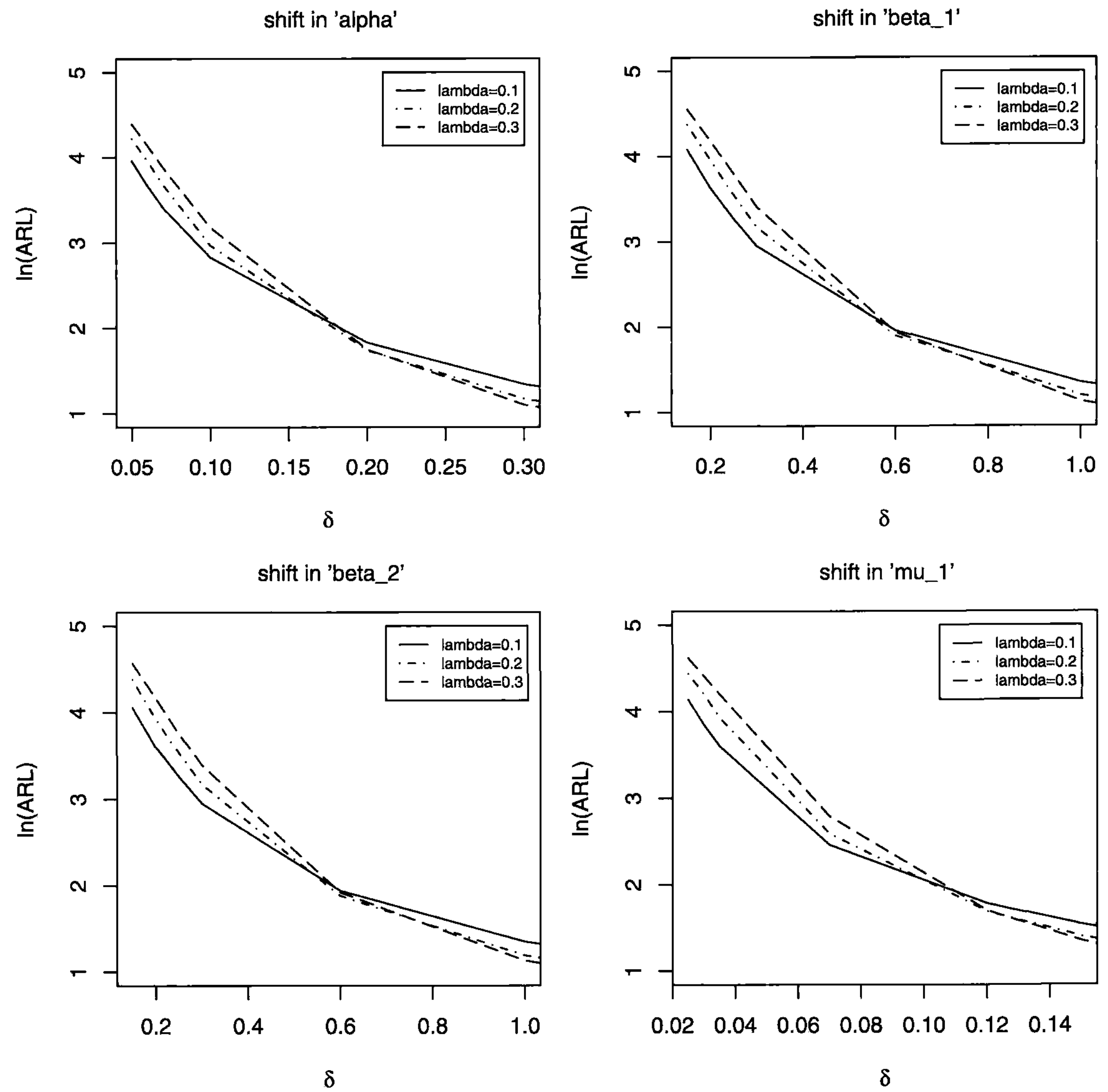


FIGURE 2. ARL Comparison of EWMA-GLM Scheme with Different Smoothing Weights λ. The solid, dash-dotted, and long-dashed lines represent the OC ARL curves of the EWMA-GLM charts with $\lambda = 0.1$, 0.2, and 0.3, respectively.

constant λ in EWMA-GLM is fixed to 0.2, as Zou et al. (2007a) have done. As shown in Table 1, our proposed EWMA-GLM control scheme performs much better than the Shewhart-GLM scheme in detecting small and moderate shifts δ in any parameter, while Shewhart-GLM has a slight advantage when the shifts are very large.

We now study the effect of λ on the performance of EWMA-GLM. The OC ARL results of three EWMA-GLM schemes with different smoothing parameters are compared, i.e., $\lambda = 0.1$, 0.2, and 0.3. As shown in Figure 2, for those small and moderate shifts, the EWMA-GLM scheme with a smaller λ is superior to the one with a larger λ in detecting shifts in parameter α, while the EWMA-GLM scheme with a larger λ is better than the one with a smaller λ in detecting large shifts. This property is consistent with that of the classical EWMA schemes in the literature (Lucas and Saccucci (1990), Lowry et al. (1992)). Based on Figure 2 and other simulation results (available from the authors), the same conclusion can be reached when shifts occur in other parameters, e.g., β and/or μ.

Table 2 shows the comparison results between the EWMA-GLM scheme and the NEWMA-GLM scheme for different values of the smoothing param-

TABLE 2. ARL Comparisons Between EWMA-GLM and NEWMA-GLM Schemes with Different λ in Detecting Various Shifts

		$\lambda = 0.1$		$\lambda = 0.3$	
	δ	EWMA-GLM	NEWMA-GLM	EWMA-GLM	NEWMA-GLM
(i)	0.050	52.5 (46.6)	117 (111)	81.0 (79.5)	112 (112)
	0.060	39.1 (31.9)	84.4 (75.3)	62.7 (60.7)	90.2 (88.4)
	0.070	30.1 (23.2)	61.7 (51.6)	48.4 (46.0)	69.0 (65.7)
	0.100	16.9 (10.6)	29.3 (20.1)	24.2 (21.1)	33.8 (30.3)
	0.200	6.24 (2.79)	9.04 (3.74)	5.74 (3.38)	7.20 (4.30)
	0.300	3.82 (1.44)	5.40 (1.81)	3.01 (1.32)	3.61 (1.50)
	0.500	2.18 (0.73)	3.12 (0.89)	1.59 (0.57)	1.97 (0.62)
(ii)	0.150	59.2 (52.4)	77.8 (71.2)	94.8 (92.8)	109 (103)
	0.200	37.6 (30.2)	50.6 (43.3)	65.1 (61.6)	77.0 (72.3)
	0.250	26.2 (19.1)	35.0 (26.8)	44.7 (42.2)	53.1 (49.5)
	0.300	19.0 (12.5)	25.4 (17.7)	30.2 (27.1)	37.3 (33.8)
	0.600	7.11 (3.35)	8.85 (4.13)	6.98 (4.45)	8.47 (5.66)
	1.000	3.87 (1.62)	4.78 (1.81)	3.09 (1.46)	3.60 (1.61)
	1.500	2.56 (1.04)	3.14 (1.05)	1.94 (0.82)	2.22 (0.80)
(iii)	0.025	63.1 (57.3)	64.3 (56.9)	101 (97.2)	100 (98.0)
	0.030	47.0 (40.7)	49.5 (42.6)	82.7 (78.3)	79.7 (76.8)
	0.035	36.6 (29.5)	38.6 (31.3)	66.4 (62.9)	64.3 (61.5)
	0.070	11.7 (6.25)	12.4 (6.77)	16.2 (12.9)	15.7 (12.6)
	0.120	5.96 (2.45)	6.06 (2.49)	5.49 (2.98)	5.44 (2.94)
	0.150	4.66 (1.75)	4.70 (1.78)	3.87 (1.77)	3.82 (1.73)
	0.200	3.43 (1.18)	3.45 (1.19)	2.64 (0.99)	2.62 (0.98)

Note: Values in parentheses are the standard deviations of OC ARLs.

eter λ. Here we consider three types of shift settings: (i) a shift in α, (ii) a shift in β_1, and (iii) a shift in μ_1. As we can see from this table, our proposed EWMA-GLM performs much better than NEWMA-GLM in detecting any given magnitude of shift in parameters α and β. When the shift occurs in μ, the performance of the EWMA-GLM scheme is almost the same as that of the NEWMA-GLM scheme for the same value of λ.

The same shift settings as in Table 2 are also considered in Table 3. As shown in Table 3, the parameters N and n_i affect the performance of the chart in detecting shifts. A larger N or n_i allows the EWMA-GLM chart to perform better in detecting a shift in the model parameters α and β, and larger N improves its performance in detecting the shift in the mean of predictor variables if the magnitude of n_i is not too small. Therefore, the magnitude of N and n_i can be decided based on the particular problem. For a process with a high production rate, a large sample size can be used, and the proposed method will perform well. But for processes with a low production rate or low defect rates, only a small sample size can be used, and the proposed method may give inferior performance.

Although the performance of the EWMA-GLM chart declines as the parameters N and n_i become smaller, it always performs much better than the NEWMA-GLM chart in detecting the shift in parameters α and β and has almost the same performance as the NEWMA-GLM chart in detecting a shift in the mean of the predictor variables. In particula, when n_i is extremely small (e.g., $n_i = 1$), the EWMA-GLM chart outperforms the NEWMA-GLM chart by quite a substantial margin. The reason is that the NEMWA-GLM chart only uses data from the current

TABLE 3. ARL Comparisons Between EWMA-GLM and NEWMA-GLM Schemes Under Different N and n_i in Detecting Various Shifts ($\lambda = 0.2$)

		$(N, n_i) = (20, 30)$			$(N, n_i) = (10, 30)$	
	δ	EWMA-GLM	NWEMA-GLM	δ	EWMA-GLM	NWEMA-GLM
(i)	0.050	68.3 (65.3)	114 (113)	0.050	96.4 (94.1)	171 (164)
	0.060	52.4 (49.4)	87.4 (83.4)	0.060	79.2 (77.8)	151 (148)
	0.070	39.5 (35.3)	66.4 (61.2)	0.070	64.3 (61.4)	130 (127)
	0.100	19.5 (15.1)	31.0 (25.5)	0.100	35.7 (31.5)	77.1 (72.5)
	0.200	5.68 (2.87)	7.50 (3.78)	0.200	9.9 (6.40)	16.6 (11.4)
(ii)	0.150	79.0 (75.0)	96.0 (89.4)	0.150	118 (116)	140 (134)
	0.200	52.3 (48.6)	64.8 (59.2)	0.200	84.9 (81.7)	114 (113)
	0.250	34.9 (30.3)	44.3 (40.1)	0.250	60.8 (58.2)	89.7 (87.6)
	0.300	23.8 (19.2)	30.6 (25.3)	0.300	45.1 (41.8)	69.6 (66.6)
	0.600	6.66 (3.62)	8.22 (4.56)	0.600	11.9 (8.18)	18.8 (14.0)
(iii)	0.025	85.2 (80.4)	84.1 (79.8)	0.025	121 (114)	120 (116)
	0.030	66.4 (62.4)	65.7 (60.9)	0.030	103 (99.4)	99.8 (97.4)
	0.035	50.8 (47.1)	51.3 (47.5)	0.035	83.7 (83.3)	83.6 (81.8)
	0.070	13.3 (9.10)	13.2 (9.00)	0.070	26.6 (22.0)	25.4 (20.5)
	0.150	4.04 (1.65)	4.04 (1.63)	0.150	6.52 (3.29)	6.39 (3.21)
		$(N, n_i) = (20, 15)$			$(N, n_i) = (20, 1)$	
	δ	EWMA-GLM	NWEMA-GLM	δ	EWMA-GLM	NWEMA-GLM
(i)	0.050	96.7 (91.7)	172 (169)	0.100	118 (116)	223 (231)
	0.060	79.5 (74.5)	149 (148)	0.200	65.5 (62.9)	262 (265)
	0.070	65.7 (60.9)	127 (123)	0.300	38.2 (34.8)	319 (319)
	0.100	37.2 (33.8)	71.3 (65.9)	0.500	15.4 (12.4)	488 (482)
	0.200	9.72 (6.22)	15.0 (10.0)	0.800	6.38 (3.79)	830 (722)
(ii)	0.150	116 (111)	134 (131)	0.600	108 (104)	174 (176)
	0.200	86.5 (81.2)	104 (99.1)	0.800	78.8 (75.1)	168 (171)
	0.250	63.3 (58.6)	79.3 (74.7)	1.000	56.9 (52.3)	161 (162)
	0.300	46.3 (42.9)	60.2 (55.3)	1.500	25.2 (21.3)	136 (139)
	0.600	11.8 (7.95)	15.4 (10.7)	2.500	9.57 (5.93)	99.8 (104)
(iii)	0.025	85.3 (81.0)	84.0 (80.4)	0.025	87.5 (82.5)	175 (180)
	0.030	67.0 (62.9)	64.9 (61.2)	0.030	69.3 (66.80)	164 (169)
	0.035	51.8 (48.6)	51.0 (47.3)	0.035	53.7 (50.50)	152 (155)
	0.070	13.3 (9.16)	13.0 (8.86)	0.070	13.6 (9.44)	78.4 (75.5)
	0.150	4.05 (1.65)	4.02 (1.63)	0.150	4.01 (1.64)	11.3 (6.57)

Note: Values in parentheses are the standard deviations of OC ARLs.

profiles, which results in poor estimations due to information deficiency, while the EWMA-GLM chart pools the previous and current profiles' data. This is the major difference between the EWMA-GLM chart and the NEWMA-GLM and Shewhart-GLM charts. Therefore, based on the simulation results in Tables 1–3 and Figure 2, we conclude that our proposed EWMA-GLM scheme, which incorporates data from different time points with different weights, is always superior to the Shewhart-GLM scheme in detecting small and moderate shifts and also better than the traditional EWMA scheme (NEWMA-GLM) in detecting any magnitude of shifts occurring in the model parameters.

A Real-Data Application: The AEC Profile Monitoring Case Revisited

In this section, we use data from the AEC manufacturing process to demonstrate the implementation of our proposed EWMA-GLM scheme. Note that, in the modeling, LC measurements x_2 are replaced by x_2^*, which is equal to $x_2/10$. The model is rewritten as

$$\mathrm{logit}(p) = \alpha + \beta_1 x_1 + \beta_2 x_2^*.$$

Based on 200 historical observations of $\mathbf{y}$ and the predictor variable values $\mathbf{x}$ (available on request from the authors), the estimated parameters are $\alpha = -3.955$ and $(\beta_1, \beta_2) = (-2.049, 0.835)$. The estimated mean of the predictor variables is $(0.1027, 0.1066)$ and the estimated variance–covariance matrix Σ of the predictor variables is

$$\Sigma = \begin{pmatrix} 17.77 \times 10^{-4} & 7.33 \times 10^{-4} \\ 7.33 \times 10^{-4} & 52.71 \times 10^{-4} \end{pmatrix}.$$

Note that a calibration sample of this size might be smaller than needed to fully determine the IC distribution, but it suffices to illustrate the use of the method in a real-world setting.

Based on this estimated process model, we simulate new profiles. In each profile sample, we have $N = 100$ observations of $\mathbf{y}$ and the corresponding predictor variable vector $\mathbf{x}$. In addition, $n = 1$ in this example. The first 20 profiles are generated from the IC normal operational condition and the remaining profiles are from the OC condition. Two OC conditions are considered here to illustrate the implementation of the proposed chart: (1) a shift $\delta = 0.1$ in β_1, and (2) a shift $\delta = 0.05$ in the mean of x_1. The smoothing constant λ is set as 0.2. Our proposed

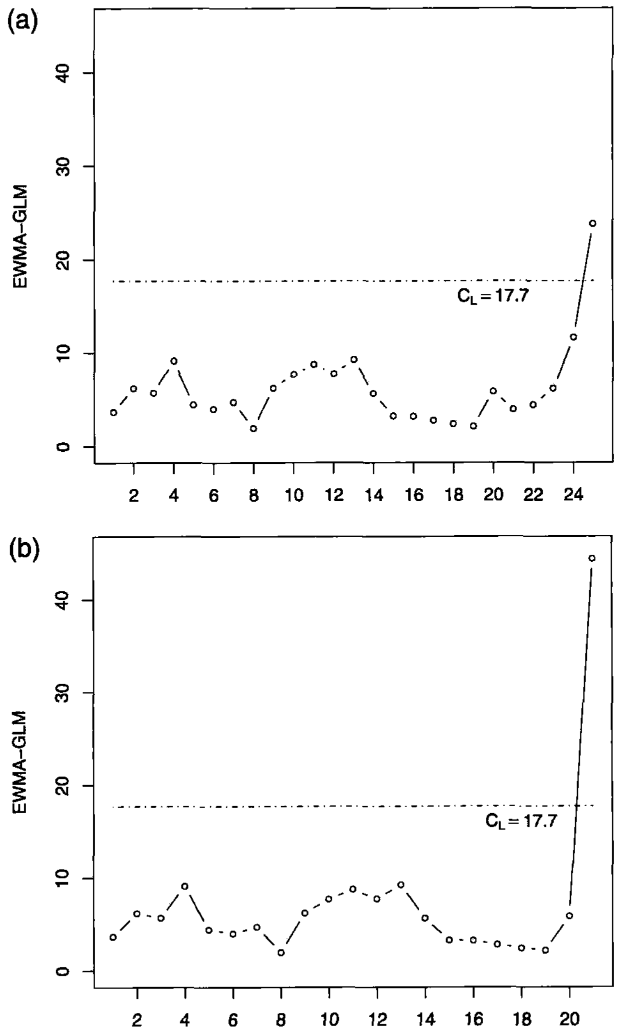


FIGURE 3. The EWMA-GLM Control Chart for the AEC Example. (a) A shift in β_1. (b) A shift in μ_1.

EWMA-GLM scheme for monitoring the profiles is implemented as follows:

1. Obtain the control limits L_M for the EWMA-GLM control chart by simulation to achieve the desired IC ARL. Here, we obtained the control limit CL = 17.7 for ARL_0=200 and then constructed the control chart in Figure 3.

2. Begin monitoring the profiles in Phase II. After obtaining the new observations, we calculate the control statistics using Equation (4), and then plot these control statistics on the control chart and compare them with the control limit. From Figure 3, we can see that the EWMA-GLM chart signals at the fifth OC profile for

the first case and at the first OC profile for the second case.

3. Identify and remove the root causes after detecting the shift, and then go back to step 1. Monitor the profiles continuously based on the revised control limit.

Conclusion

Statistical process control monitoring is important and challenging for profiles with categorical data and random predictor variables. In this paper, we used a type of GLM model to represent the functional relationship between a binary response variable and a set of random predictor variables. We proposed a novel control scheme, the EWMA-GLM, for monitoring such profiles. The EWMA-GLM scheme integrates an EWMA scheme and a logistic regression likelihood-ratio test. As shown by the simulation results in this paper, for the examples considered in this paper, the EWMA-GLM scheme almost always performed better than the Shewhart-GLM scheme and the NEWMA-GLM scheme, which are benchmarks for performance comparison based on existing research.

There are a number of issues not thoroughly addressed here that could be topics of future research. First, this paper focuses on Phase II monitoring only and presumes that the number of historical observations used for estimating the IC parameters is sufficiently large. In practical applications, the performance of the EWMA-GLM would be affected by the amount of data in the reference dataset (Jensen et al. (2006)). Thus, determination of required Phase I sample sizes to reduce the effects of estimated parameters and a general recommendation are needed. Second, this new control scheme is proposed based on a logistic regression model for the profiles. However, different types of categorical data, e.g., multinomial data, are not uncommon in many industries. Therefore, new control schemes for monitoring profiles with other types of categorical data are interesting topics for further research. This would be adapting the proposed weighted likelihood-ratio test to the general GLM model fitting. Moreover, our proposed scheme assumes that the observations are independent within and between profiles. When observations are dependent, this scheme will not be applicable. Therefore, how to develop new schemes for dealing with this correlation is another future research topic.

Appendix A: Derivation of the MLE $\widehat{\boldsymbol{\xi}}$

The MLE of the model parameters $\boldsymbol{\xi} = (\alpha, \boldsymbol{\beta}^{\mathrm{T}})^{\mathrm{T}}$ can be obtained via the standard GLM procedure with the augmented dependent variable z_i, as briefly described in the following. Here we will suppress the index "j" for ease of exposition. Denote

$$z_i = \eta_i + (y_i - \mu_{yi})\frac{\partial \eta_i}{\partial \mu_{yi}} = \eta_i + \frac{y_i - \mu_{yi}}{n_i p_i(1-p_i)},$$

with $\eta_i = \alpha + \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}$, where $i = 1, \ldots, N$, η_i is defined as the linear predictor, and μ_{yi} is the mean of y_i, say, $n_i p_i$. Moreover, the GLM weight functions are denoted as $\mathbf{W} = \mathrm{diag}\{w_1, \ldots, w_N\}$, where $w_i = [n_i p_i(1-p_i)]$. Then the GLM augmented dependent-variable vector is written as

$$\mathbf{z} = \boldsymbol{\eta} + \mathbf{W}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y),$$

where $\mathbf{z} = (z_1, \ldots, z_N)^{\mathrm{T}}$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N)^{\mathrm{T}}$, and $\boldsymbol{\mu}_y = (\mu_{y1}, \ldots, \mu_{yN})^{\mathrm{T}}$. Let $\mathbf{X} = (\ddot{\mathbf{x}}_1, \ldots, \ddot{\mathbf{x}}_N)^{\mathrm{T}}$, which is an $N \times (q+1)$ matrix and $\ddot{\mathbf{x}}_i = (1, \mathbf{x}_i^{\mathrm{T}})^{\mathrm{T}}$. By McCullagh and Nelder (1989), the MLEs of model parameters $\widehat{\boldsymbol{\xi}}$ can be obtained by using the following iterative weighted least-square (IWLS):

1. Start with the initial values of $\widehat{\boldsymbol{\xi}}$, denoted as $\widehat{\boldsymbol{\xi}}^{(0)}$.
2. At the lth iteration, for $l \geq 0$, calculate $\mathbf{z}^{(l)}$ and $\mathbf{W}^{(l)}$ based on $\widehat{\boldsymbol{\xi}}^{(l)}$.
3. Update the estimation of $\boldsymbol{\xi}$ as follows:
$$\widehat{\boldsymbol{\xi}}^{(l+1)} = (\mathbf{X}^T\mathbf{W}^{(l)}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}^{(l)}\mathbf{z}^{(l)}.$$
4. Repeat steps 2 and 3 until the following condition is satisfied:
$$\|\widehat{\boldsymbol{\xi}}^{(l)} - \widehat{\boldsymbol{\xi}}^{(l-1)})\|_1 \Big/ \|\widehat{\boldsymbol{\xi}}^{(l-1)}\|_1 \leq \epsilon,$$
where ϵ is a given small positive value (e.g., $\epsilon = 10^{-4}$) and $||\boldsymbol{\xi}||_1$ denotes the sum of absolute values of all elements of $\boldsymbol{\xi}$. Then, the algorithm stops at the lth iteration.

Appendix B: Derivation of the Joint Log Likelihood

The joint log likelihood of $(\tilde{\mathbf{X}}_j, \mathbf{y}_j)$ is

$$\begin{aligned} l_j &= \log f(\mathbf{y}_j, \tilde{\mathbf{X}}_j) \\ &= \log f(\mathbf{y}_j \mid \tilde{\mathbf{X}}_j) f(\tilde{\mathbf{X}}_j) \\ &= \log f(\mathbf{y}_j \mid \tilde{\mathbf{X}}_j) + \log f(\tilde{\mathbf{X}}_j) \end{aligned}$$

and

$$\begin{aligned}
&\log f(\mathbf{y}_j \mid \tilde{\mathbf{X}}_j) \\
&= \log \prod_{i=1}^{N} \mathrm{C}_{n_{ji}}^{y_{ji}} {p_{ji}}^{y_{ji}} (1-p_{ji})^{n_{ji}-y_{ji}} \\
&= \sum_{i=1}^{N} \log \mathrm{C}_{n_{ji}}^{y_{ji}} + y_{ji} \log p_{ji} \\
&\qquad + (n_{ji} - y_{ji}) \log(1 - p_{ji}) \\
&= \sum_{i=1}^{N} \log \mathrm{C}_{n_{ji}}^{y_{ji}} + y_{ji}(\alpha_j + \mathbf{x}_{ji}^{\mathrm{T}} \boldsymbol{\beta}_j) \\
&\qquad - n_{ji} \log\left[1 + \exp\{(\alpha_j + \mathbf{x}_{ji}^{\mathrm{T}} \boldsymbol{\beta}_j)\}\right], \\
&\log f(\tilde{\mathbf{X}}_j) \\
&= -\frac{1}{2} \log |2\pi \boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{x}_{ji} - \boldsymbol{\mu}_j)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}_{ji} - \boldsymbol{\mu}_j).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
l_j = \sum_{i=1}^{N} &\log \mathrm{C}_{n_{ji}}^{y_{ji}} + y_{ji}(\alpha_j + \mathbf{x}_{ji}^{\mathrm{T}} \boldsymbol{\beta}_j) \\
&- n_{ji} \log\left[1 + \exp\{(\alpha_j + \mathbf{x}_{ji}^{\mathrm{T}} \boldsymbol{\beta}_j)\}\right] \\
&- \frac{1}{2} \log |2\pi \boldsymbol{\Sigma}| \\
&- \frac{1}{2} (\mathbf{x}_{ji} - \boldsymbol{\mu}_j)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}_{ji} - \boldsymbol{\mu}_j).
\end{aligned}$$

Appendix C: Obtain the EWMA-GLM Charting Statistic

The MWLEs of $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$ satisfy the following simultaneous score equations:

$$\partial l_{t,\lambda} / \partial \boldsymbol{\xi} = 0, \quad \partial l_{t,\lambda} / \partial \boldsymbol{\mu} = 0.$$

The MWLE of μ can be simply expressed as

$$\widehat{\boldsymbol{\mu}}_t = \sum_{j=1}^{t} \lambda (1-\lambda)^{t-j} \sum_{i=1}^{N} \mathbf{x}_{ji} / N.$$

On the other hand, the MWLE of $\boldsymbol{\xi}$ can be similarly obtained via GLM procedure with the augmented dependent variables via the following procedure. To alleviate the computation burden, we denote m as a sufficiently large integer to make $(1-\lambda)^m$ close to 0. Let $\widehat{\mathbf{X}}_t = (\mathbf{X}_{t-m+1}^{\mathrm{T}}, \ldots, \mathbf{X}_t^{\mathrm{T}})^{\mathrm{T}}$ be an $mN \times (q+1)$ matrix, which includes the most recent m sets of explanatory variable values, $\widehat{\mathbf{z}}_t = (\mathbf{z}_{t-m+1}^{\mathrm{T}}, \ldots, \mathbf{z}_t^{\mathrm{T}})^{\mathrm{T}}$ be mN-dimensional vector, and $\widehat{\mathbf{W}}_t = \mathrm{diag}\{\widehat{\mathbf{w}}_{t-m+1}, \ldots, \widehat{\mathbf{w}}_t\}$ be an $mN \times mN$ matrix. $\mathbf{X}_j$ and $\mathbf{z}_j$ are defined in a similar fashion to the notations in the above subsection, and $\widehat{\mathbf{w}}_j = \mathrm{diag}\{\widehat{w}_{j1}, \ldots, \widehat{w}_{jN}\}$, where $\widehat{w}_{ji} = \lambda(1-\lambda)^{t-j} n_{ji} p_{ji} (1-p_{ji})$. The MWLEs $\widehat{\boldsymbol{\xi}}_t$ can be immediately obtained by implementing the IWLS procedure in Appendix A, replacing $\mathbf{X}, \mathbf{W}, \mathbf{z}$ with $\widehat{\mathbf{X}}_t, \widehat{\mathbf{W}}_t, \widehat{\mathbf{z}}_t$.

After obtaining the MWLE $(\widehat{\boldsymbol{\xi}}_t, \widehat{\boldsymbol{\mu}}_t)$, the corresponding log-likelihood ratio test can be defined as

$$lr_t = -2[l_{t,\lambda}(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0) - l_{t,\lambda}(\widehat{\boldsymbol{\xi}}_t, \widehat{\boldsymbol{\mu}}_t)].$$

Using standard Taylor's expansion arguments of likelihood functions (Serfling (1980)), the expansion of lr_t leads to asymptotically equivalent Wald-type charting statistics,

$$\begin{aligned}
lr_t \approx\ & (\widehat{\boldsymbol{\xi}}_t - \boldsymbol{\xi}_0)^{\mathrm{T}} \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_t}^{-1} (\widehat{\boldsymbol{\xi}}_t - \boldsymbol{\xi}_0) \\
&+ \frac{N(2-\lambda)}{\lambda} (\mathbf{E}_t - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{E}_t - \boldsymbol{\mu}),
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_t} &= \frac{\lambda}{2-\lambda} (\widehat{\mathbf{X}}_t^{\mathrm{T}} \widehat{\mathbf{W}}_t \widehat{\mathbf{X}}_t)^{-1}, \\
\mathbf{E}_t &= \lambda \overline{\mathbf{x}}_t + (1-\lambda) \mathbf{E}_{t-1}, \quad t = 1, 2, \ldots,
\end{aligned}$$

$\mathbf{E}_0 = \boldsymbol{\mu}_0$ is the starting vector, and $\overline{\mathbf{x}}_t = \sum_{i=1}^{N} \mathbf{x}_{ti} / N$.

Acknowledgment

The authors thank the Editor and two anonymous referees for their many helpful comments that have resulted in significant improvements in the article. This research was supported by the RGC Competitive Earmarked Research Grant 620010 and NNSF of China Grants 70931004, 11001138, and 11071128. Zou also thanks the support of Nankai Young Grant 65010731.

References

CHICKEN, E.; PIGNATIELLO, JR., J. J.; and SIMPSON, J. (2009). "Statistical Process Monitoring of Nonlinear Profiles Using Wavelets". *Journal of Quality Technology* 41, pp. 198–212.

COLOSIMO, B. M.; PACELLA, M.; and SEMERARO, Q. (2008). "Statistical Process Control for Geometric Specifications: On the Monitoring of Roundness Profiles". *Journal of Quality Technology* 40, pp. 1–18.

DING, Y.; ZENG, L.; and ZHOU, S. (2006). "Phase I Analysis for Monitoring Nonlinear Profiles in Manufacturing Processes". *Journal of Quality Technology* 38, pp. 199–216.

GUREVICH, G. and VEXLER, A. (2005). "Change Point Problems in the Model of Logistic Regression". *Journal of Statistical Planning and Inference* 131, pp. 313–331.

JENSEN, W. A. and BIRCH, J. B. (2009). "Profile Monitoring via Nonlinear Mixed Models". *Journal of Quality Technology* 41, pp. 18–34.

JENSEN, W. A.; BIRCH, J. B.; and WOODALL, W. H. (2008). "Monitoring Correlation Within Linear Profiles Using Mixed Models". *Journal of Quality Technology* 40, pp. 167–183.

JENSEN, W. A.; JONES-FARMER, L. A.; CHAMP, C. W.; and WOODALL, W. H. (2006). "Effects of Parameter Estimation on Control Chart Properties: A Literature Review". *Journal of Quality Technology* 38, pp. 349–364.

KANG, L. and ALBIN, S. L. (2000). "On-Line Monitoring When the Process Yields a Linear Profile". *Journal of Quality Technology* 32, pp. 418–426.

KAZEMZADEH, R. B.; NOOROSSANA, R.; and AMIRI, A. (2008). "Phase I Monitoring of Polynomial Profiles". *Communications in Statistics: Theory and Methods* 37, pp. 1671–1686.

KIM, K.; MAHMOUD, M. A.; and WOODALL, W. H. (2003). "On the Monitoring of Linear Profiles". *Journal of Quality Technology* 35, pp. 317-328.

LOWRY, C. A.; WOODALL, W. H.; CHAMP, C. W.; and RIGDON, S. E. (1992). "Multivariate Exponentially Weighted Moving Average Control Chart". *Technometrics* 34, pp. 46–53.

LUCAS, J. M. and SACCUCCI, M. S. (1990). "Exponentially Weighted Moving Average Control Scheme Properties and Enhancements". *Technometrics* 32, pp. 1–29.

MAHMOUD, M. A. (2008). "Phase I Analysis of Multiple Linear Regression Profiles". *Communications in Statistics: Simulation and Computation* 37, pp. 2106–2130.

MAHMOUD, M. A. and WOODALL, W. H. (2004). "Phase I Analysis of Linear Profiles with Calibration Applications". *Technometrics* 46, pp. 380–391.

MCCULLAGH, P. and NELDER, J. A. (1989). *Generalized Linear Models.* London: Chapman and Hall

QIU, P. and ZOU, C. (2009). "Control Chart for Monitoring Nonparametric Profiles with Arbitrary Design". *Statistica Sinica*, to appear.

QIU, P.; ZOU, C.; and WANG, Z. (2010). "Nonparametric Profile Monitoring by Mixed Effects Modeling (with Discussions)". *Technometrics* 52, pp. 265–277.

REYNOLDS, M. R. and STOUMBOS, Z. G. (2000). "A General Approach to Modeling CUSUM Charts for a Proportion". *IIE Transactions* 32, pp. 515–535.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematicai Statistics.* New York, NY: John Wiley & Sons.

SOMERVILLE, S. E.; MONTGOMERY, D. C.; and RUNGER, G. C. (2002). "Filtering and Smoothing Methods for Mixed Particle Count Distributions". *International Journal of Production Research* 40, pp. 2991–3013.

STEINER, S. H. (1998). "Grouped Data Exponentially Weighted Moving Average Control Charts". *Applied Statistics* 47, pp. 203–216.

STEINER, S. H.; COOK, J. C.; FAREWELL, V. T.; and TREASURE, T. (2000). "Monitoring Surgical Performance Using Risk Adjusted Cumulative Sum Charts". *Biostatistics* 1, pp. 441–452.

WILLIAMS, J. D.; WOODALL, W. H.; and BIRCH, J. B. (2007)."Statistical Monitoring of Nonlinear Product and Process Quality Profiles". *Quality and Reliability Engineering International* 23, pp. 925-941.

WOODALL, W. H. (2007). "Current Research on Profile Monitoring". *Revista Producão* 17, pp. 420–425.

WOODALL, W. H.; SPITZNER, D. J.; MONTGOMERY, D. C.; and Gupta, S. (2004). "Using Control Charts to Monitor Process and Product Quality Profiles". *Journal of Quality Technology* 36, pp. 309–320.

YEH, A. B.; HUWANG, L.; and LI, Y. M. (2009). "Profile Monitoring for a Binary Response". *IIE Transactions* 41, pp. 931–941.

ZHANG, H. and ALBIN, S. (2009)."Detecting Outliers in Complex Profiles Using a χ^2 Control Chart Method". *IIE Transactions* 41, pp. 335–345.

ZOU, C.; QIU, P.; and HAWKINS, D. (2009). "Nonparametric Control Chart for Monitoring Profiles Using the Change Point Formulation". *Statistica Sinica* 19, pp. 1337–1357.

ZOU, C.; TSUNG, F.; and WANG, Z. (2007a). "Monitoring General Linear Profiles Using Multivariate EWMA Schemes". *Technometrics* 49, pp. 395–408.

ZOU, C.; TSUNG, F.; and WANG, Z. (2008). "Monitoring Profiles Based on Nonparametric Regression Methods". *Technometrics* 50, pp. 512–526.

ZOU, C.; ZHANG, Y.; and WANG, Z. (2006). "Control Chart Based on Change-Point Model for Monitoring Linear Profiles". *IIE Transactions* 38, pp. 1093–1103.

ZOU, C.; ZHOU, C.; WANG, Z.; and TSUNG, F. (2007b). "A Self-Starting Control Chart for Linear Profiles". *Journal of Quality Technology* 39, pp. 364-375.