# 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*

**Abstract**

Profile monitoring is for checking the stability of some functional relationships between response variables and one or more explanatory variables over time. In many applications, categorical response variables are common and the generalized linear model is usually utilized to model this kind of profiles for quality improvement. In practice, different profiles often have random covariates 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 the EWMA scheme and the likelihood ratio test based on logistic regression. This new scheme not only provides the ability to monitor the functional relationship of the profile but also detects the mean shift in explanatory variables. The proposed chart is fast to compute, easy to implement, and efficient in detecting shifts. The simulation results show that it performs almost always better than the standard benchmarks in the literatures. A real example from the electronic industries is used to illustrate the implementation of the proposed approach.

*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.

**Keywords:** Categorical Response; EWMA; GLM; Random Explanatory Variables; Profile Monitoring; 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 function or relationship between the response variable and one or more explanatory variables instead of the distribution of variables. Therefore, studies on profile monitoring have been popularly conducted in recent years. An extensive discussion of research problems on this topic has been given by Woodall et al. (2004).

Studies focused on simple linear profiles have been particularly prospering, 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 are considered by Zou et al. (2007a), Kazemzadeh et al. (2008), Mahmoud (2008), Jensen et al. (2008), and Jensen and Birch (2009). Nonlinear profile models are investigated by Williams et al. (2007). Recently, profile monitoring for general profile model has also attracted much attention. The reader is referred to Zou et al. (2008; 2009) and Qiu et al. (2010) for the Phase II methods based on nonparametric regression; Ding et al. (2006), Colosimo et al. (2008), Chicken et al. (2009) and 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 has been given by Woodall (2007).

All the above mentioned control schemes for monitoring linear and/or nonlinear profiles require the fundamental assumption that the measurements of response variables are continuous. However, due to some practical restrictions, e.g., time, cost or intrinsic characteristics of the variables, only qualitative response measurements rather than quantitative measurements can be directly or promptly collected for on-line monitoring. For instance, on a production line each item is inspected and classified as conforming or nonconforming, according to some predefined specification of 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 between a binary (or binomial) response variable and one or more continuous predictor variables (see the following motivating example in next section for detailed illustration).

With respect to the monitoring of 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 the profile monitoring schemes for categorical data, few studies have been conducted recently. In the literature, we have not found any research 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 the logistic regression model. They modeled the relationship between the binary response and explanatory variables by using the 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 actual process performance. This dataset, which is referred to as the in-control (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 goal of this phase is to detect any change 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 of the manufacturing industry. In some other applications, however, they may be invalid (Qiu and Zou 2009). Specifically, when data acquisition adopts the random design scheme, design points within a profile would be i.i.d. random variables from a given density (Qiu and Zou 2009). In the random design scheme, the values of predictors in each profile sample would be different. Therefore, the existing schemes may not be efficient in such a case, and how to efficiently use the data from random design scheme in the Phase II stage need to be studied. Moreover, in such situations, these covariates observations themselves require careful monitoring and control along with the monitoring of profile. This is a unique issue in such a random design profile problem. Phase II profile monitoring in such cases is particularly challenging, and is the focus of this paper.

In this paper, we utilize a generalized linear model (GLM), logistic regression model, to represent the function or relationship between the binary response and explanatory variables which are not deterministic and have given distributions. Under this premise, a control scheme is proposed based on exponential weighted moving average (EWMA) process control schemes. This control scheme is able to simultaneously monitor the parameters' shifts in the profile and mean shifts from explanatory variables. The remainder of this paper is organized as follows: We introduce an example from the electrolytic capacitor industry that motivates this research in the next section. After that, our proposed methodology is described in detail. And then its numerical performance is thoroughly investigated. Following that, the motivating example, which has a profile that fits a logistic regression model well, is used to illustrate the implementation of the proposed approach step by step. Finally, several remarks conclude the article. The technical details are provided in the Appendix.

# The motivating example: monitoring an aluminium electrolytic capacitors' manufacturing process

We use an example taken from the manufacturing process of an aluminium electrolytic capacitor (provided by ENW Electronics Ltd, see Figure 1) to illustrate the motivation for 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 aluminium electrolytic capacitors (AECs) with given specifications. The quality of the unfinished AEC products (or capacitor elements) in terms of appearance and functional performance will be 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 the Leakage Current (LC) and Dissipation Factor (DF), are automatically calibrated by an electronic device at some given measuring voltage, frequency and temperature.

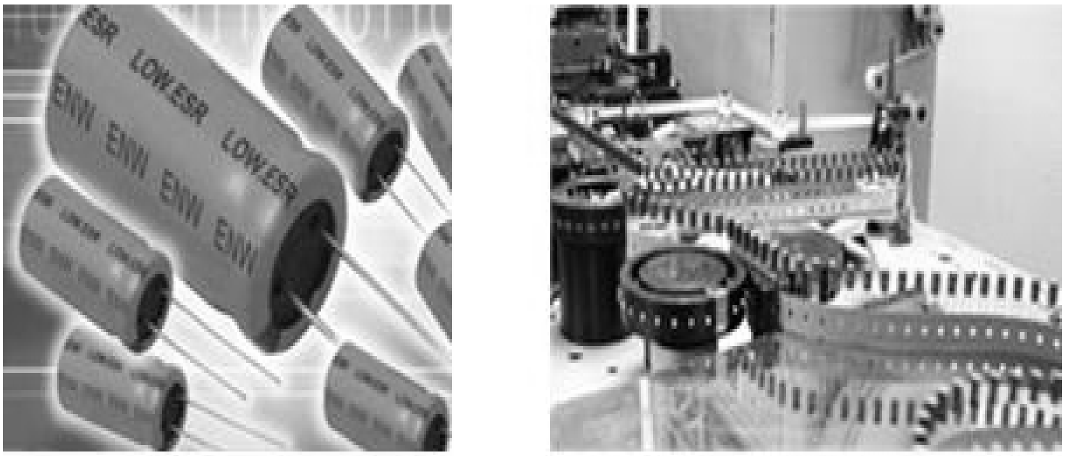



Figure 1: Aluminium Electrolytic Capacitors

The number of defective capacitors in a certain sample size $n$, denoted as $y$, 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, two variables, the DF and LC are also collected for each capacitor. These two variables are usually random distributed, and both of them affect the defective rate of AECs to a certain extent. Hence, in this example, $n = 1$ and $y = 1$ or 0. If we denote DF and LC as the predictor variables, $x_1$ and $x_2$, the dataset is collected as $(y, x_1, x_2)$. The relationship between $y$ and $x_1, x_2$ can be modeled as a classical GLM 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, $\alpha$, $\beta_1$ and $\beta_2$, 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 $\alpha$, $\beta_1$ and $\beta_2$ reflect that the relationship between the defect rate and the DF and LC of products changes, which indicates the special causes may occur. Therefore, jointly monitoring the relationship between $y$ and $x_1, x_2$ and the mean change of $x_1, x_2$ may give more complete information for effective monitoring and diagnosis and may result in better quality improvement. However, how to implement SPC monitoring for such a profile, to a problem which the methods currently available in the literatures apparently cannot help solve, still remains a challenge. In the remainder of this paper, we propose an SPC scheme to monitor such a profile 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. In what follows, we elaborate on binary (binomial) responses and the corresponding logistic regression because it is of greatest interest in our applications and suffices to illustrate our method. The extension to general cases will be discussed in the last section.

Assume that for the $j$th ($j \geq 1$) 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 from the following profile model

$$\text{logit}(p_{ji}) = \alpha_j + \mathbf{x}_{ji}^T \boldsymbol{\beta}_j, \quad i = 1, \cdots, N, \quad j = 1, \ldots, \tau, \tau + 1, \ldots, \tag{1}$$

where $\tau$ is the unknown change-point, $y_{ji}$ is the $i$th response observation of the $j$th random profile, $\mathbf{x}_{ji}^T$ denotes the $i$th row of $\tilde{\mathbf{X}}_j$ (such as the $x_1$ and $x_2$ in the above example), $\alpha_j$ is the intercept parameter, $\boldsymbol{\beta}_j = (\beta_{1j}, \ldots, \beta_{qj})^T$ is a $q$-dimensional coefficient vector. Here $y_{ji}$ is assumed to be drawn from a Binomial (or Bernoulli) distribution with the parameter $p_{ji}$, say $y_{ji} \sim \text{Binomial}(n_{ji}, p_{ji})$, where $n_{ji}$ is the sample size for the $i$th observation of the $j$th profile. Note that in the AEC example, $n_{ji} = 1$ and $y_{ji}$ is a binary response. $p_{ji}$ represents the $i$th defect rate of the products in the $j$th profile sample. Typically, when one group or batch of products are produced at a particular setting of predictor variables such as temperature and pressure (say $\mathbf{x}_{ji}$), $n_{ji}$ would be greater than one; when the setting of predictor variables are along with 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 supposed that after some unknown change-point $\tau$, there is a change in the intercept, and/or coefficient and/or the mean vector of covariates. Say,

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

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 easily be satisfied in practical applications.

The maximum likelihood estimations (MLEs) of the model parameters $\boldsymbol{\xi} = (\alpha, \boldsymbol{\beta}^T)^T$ can be obtained via the standard GLM procedure by using iterative weighted least square (IWLS) method. Details of obtaining 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 IC model, $\widehat{\boldsymbol{\xi}}|\mathbf{X}$ asymptotically follows the multivariate normal distribution

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

in which $\boldsymbol{\xi}_0 = (\alpha_{(0)}, \boldsymbol{\beta}_{(0)}^T)^T$, $\mathbf{X} = (\acute{\boldsymbol{x}}_1, \ldots, \acute{\boldsymbol{x}}_N)^T$ is $N \times (q+1)$ matrix and $\acute{\boldsymbol{x}}_i = (1, \mathbf{x}_i^T)^T$ , and $\mathbf{W} = \text{diag}\{w_1, \ldots, w_N\}$ denote 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 model (1), in which $(q+1)$-variate parameter vector $\boldsymbol{\xi}$ and $q$-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):

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

Then, the MLEs of the profile parameters and the mean of explanatory variables based on Eq.(2) can be obtained, defined as $(\tilde{\boldsymbol{\xi}}_j, \tilde{\boldsymbol{\mu}}_j) = \arg\max_{\boldsymbol{\xi},\boldsymbol{\mu}} l_j$. 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 naive method that comes to mind for on-line detection is to use the current profile estimates to construct two charts for model parameter $\boldsymbol{\xi}$ 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 the scheme including one more charts is complicated and difficult (Zou et al. 2007a). Another naive method is to construct a single Shewhart-type $T^2$ chart. However, this would be very inefficient with moderate and small changes since 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, since only $N$ random explanatory observations are used for estimating parameters in individual profiles, and thus the estimators would have considerably large bias and variance.

Alternatively, in order to monitor the profile or the 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 $t$

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

where $\lambda$ is a weighting parameter. Obviously, the $l_{t,\lambda}(\boldsymbol{\xi}, \boldsymbol{\mu})$ in Eq.(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 have more weight and the weight changes exponentially over time). Then the maximum weighted likelihood estimator (MWLE), defined as $(\widehat{\boldsymbol{\xi}}_t, \widehat{\boldsymbol{\mu}}_t) = \arg\max_{\boldsymbol{\xi},\boldsymbol{\mu}} l_{t,\lambda}(\boldsymbol{\xi}, \boldsymbol{\mu})$, can be obtaind via IWLS method (see Appendix C).

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

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

where

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

$\mathbf{E}_0 = \boldsymbol{\mu}_0$ 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 average run length (ARL), $\text{ARL}_0$. This charting statistic is an approximation to the likelihood ratio test statistic based on weighted likelihood function (3). Details for 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 the shifts of profile parameters and the mean of random explanatory variables through Monte Carlo simulations. It is challenging to compare the proposed method with alternative methods, since there is no obvious comparable 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

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

where $\tilde{\boldsymbol{\xi}}_t$ is the MLE obtained as Appendix B, $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{\xi}}_t} = (\mathbf{X}_t^T \mathbf{W}_t \mathbf{X}_t)^{-1}$, and $\mathbf{X}_t^T$ and $\mathbf{W}_t$ are the corresponding matrices defined at the beginning of the previous section for the $t$th profile sample. The chart signals when $T_{St}^2 > L_S$, where $L_S$ is the control limit chosen to achieve a specific $\text{ARL}_0$. This chart will be called Shewhart-GLM chart.

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

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

where

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

where $\mathbf{\Sigma}_{\tilde{\boldsymbol{\xi}}t}$ is defined as for Eq.(5). We call this naive EWMA scheme NEWMA-GLM for abbreviation hereafter.

Table 1: ARL comparisons between EWMA-GLM with $\lambda = 0.2$ and Shewhart-GLM scheme in detecting various shifts

| | $\delta$ in $\alpha$ | | | | | | |
|---|---|---|---|---|---|---|---|
| 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) |
| | $\delta$ in $\beta_1$ | | | | | | |
| 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) |
| | $\delta$ in $\mu_1$ | | | | | | |
| 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

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 $\lambda$; 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 $\tau$ at which the shifts initially occur is 40 and only consider the sustained shifts for the remaining samples after the changepoint $\tau$. 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 $\boldsymbol{\mu}_0 = (0,0)^T$, and $\mathbf{\Sigma} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$. The other parameters are assumed to be $\boldsymbol{\beta} = (1.0, 1.0)^T$ and $\alpha = -2.2$ that make the expectation of 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 roughly 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 that of Shewhart-GLM for detecting shifts in $\alpha$, $\beta_1$ and $\mu_1$. The smoothing constant $\lambda$ in EWMA-GLM is fixed to 0.2 as in Zou et al. (2007a). As shown in Table 1, our proposed EWMA-GLM control scheme performs much better than the Shewhart-GLM scheme in detecting the small and moderate shifts $\delta$ in any parameter, while Shewhart-GLM has a slight advantage when shifts are very large.

Next, we study the effect of $\lambda$ 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 smaller $\lambda$ is superior to the one with larger $\lambda$ in terms of detecting shifts in parameter $\alpha$, while the EWMA-GLM scheme with larger $\lambda$ is better than the one with smaller $\lambda$ for detecting large shifts. This property is consistent with that of the classical EWMA schemes in the literatures (Lucas and Saccucci 1990; Lowry et al. 1992). Based on Figure 2 and other simulation results (available from authors), the same conclusion can be reached when the shift occurs in other parameters, e.g., $\boldsymbol{\beta}$ and/or $\boldsymbol{\mu}$.

Table 2 shows the comparison results between the EWMA-GLM scheme and the NEWMA-GLM scheme for different values of smoothing parameter $\lambda$. Here we consider three types of shift settings: (i) shift occurs in $\alpha$; (ii) shift occurs in $\beta_1$; (iii) shift occurs in $\mu_1$. As we can see from this table, our proposed EWMA-GLM performs much better than the NEWMA-GLM in detecting any given magnitude of shift in parameters $\alpha$ and $\boldsymbol{\beta}$. When the shift occurs in the mean of random variable, say $\boldsymbol{\mu}$, the performance of the EWMA-GLM scheme is almost the same as that of the NEWMA-GLM scheme for the same parameter $\lambda$.

Same with the shift settings in Table 2, those three types of settings 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 the shifts. Larger $N$ and/or $n_i$ makes the EWMA-GLM chart performing better

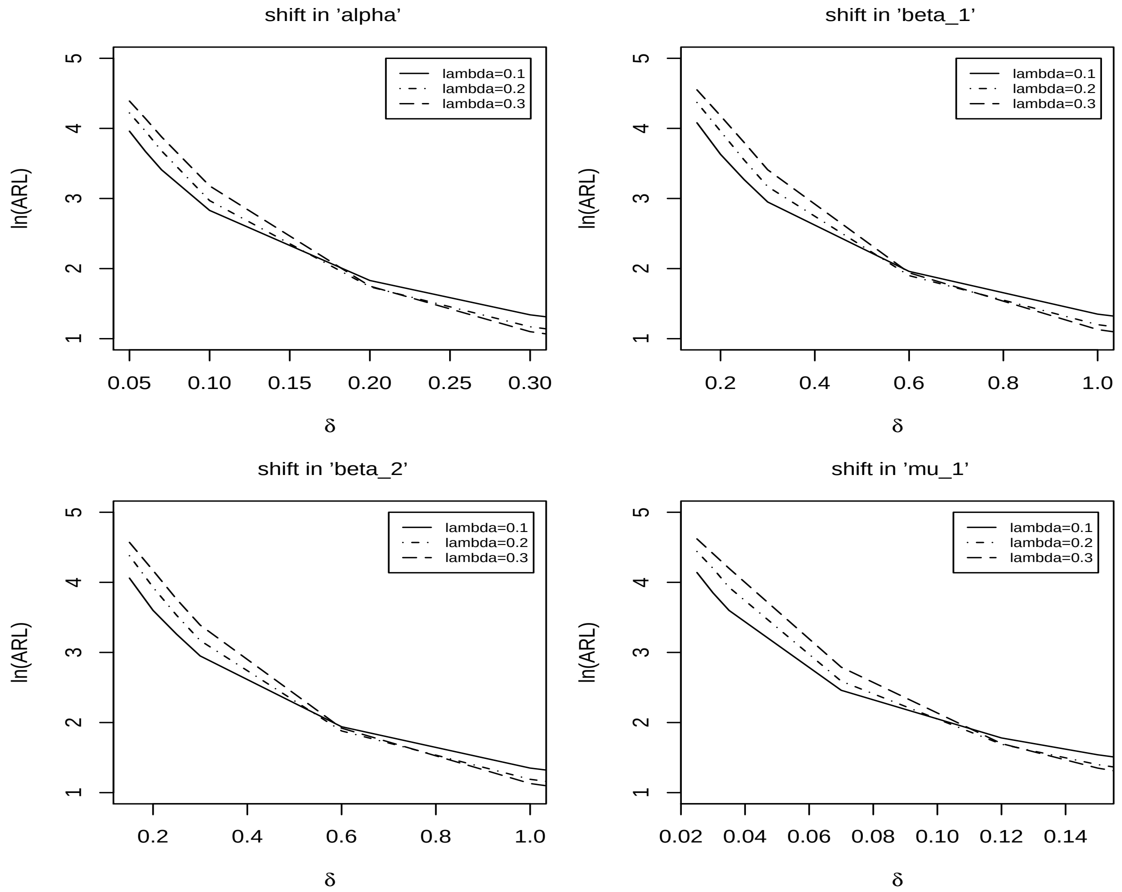



Figure 2: ARL comparison of EWMA-GLM scheme with different smoothing weights $\lambda$. The solid, dash-dotted and long-dashed lines represent the OC ARL curves of the EWMA-GLM charts with $\lambda = 0.1$, 0.2 and 0.3, respectively.

in detecting the shift of the model parameters $\alpha$ and $\beta$, and larger $N$ makes its performance better in detecting the shift in the mean of predictor varaibles as the magnitude of $n_i$ is not too small. Therefore, the magnitude of $N$ and $n_i$ can be decided based on the practical problem. If the industry has a high production rate, then large sample size can be used and the proposed method performs well, while if the industry has a low production rate and/or low defect rates, only small sample size can be used and the proposed method may have bad performance.

Although the performance of the EWMA-GLM chart declines as the parameters $N$ and $n_i$ are small, it always performs much better than the NEWMA-GLM in detecting the shift in parameters $\alpha$ and $\boldsymbol{\beta}$, and has almost the same performance as the NEWMA-GLM chart

Table 2: ARL comparisons between EWMA-GLM and NEWMA-GLM schemes with different $\lambda$ in detecting various shifts

| | | $\lambda = 0.1$ | | $\lambda = 0.3$ | |
|---|---|---|---|---|---|
| | $\delta$ | 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

in detecting the shift in the mean of the predictor variables. Especially when $n_i$ is very small (e.g. $n_i = 1$), the EWMA-GLM chart outperforms NEWMA-GLM chart by a quite substantial margin. The reason is that NEMWA-GLM chart only uses the current profiles' data which results in bad estimations based on the deficient information, while EWMA-GLM chart pools the previous and current profiles' data, which is the major difference between EWMA-GLM chart and 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 the data at different times with different weights, is always

Table 3: ARL comparisons between EWMA-GLM and NEWMA-GLM schemes under different $N$ and $n_i$ in detecting various shifts ($\lambda$=0.2)

| | | $(N, n_i) = (20, 30)$ | | | $(N, n_i) = (10, 30)$ | |
|---|---|---|---|---|---|---|
| | $\delta$ | EWMA-GLM | NWEMA-GLM | $\delta$ | 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) |
| | | $(N, n_i) = (20, 15)$ | | | $(N, n_i) = (20, 1)$ | |
| | $\delta$ | EWMA-GLM | NWEMA-GLM | $\delta$ | 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 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 the data from the aluminum electrolytic capacitors (AECs) manufacturing industry 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 equals to $x_2/10$. Then the model is rewritten as follows:

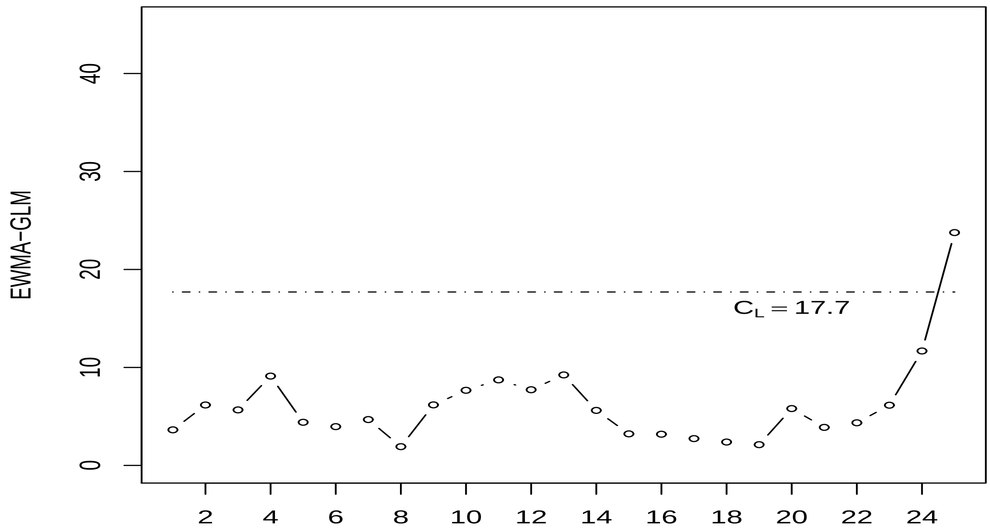
$$\mathrm{logit}(p) = \alpha + \beta_1 x_1 + \beta_2 x_2^*.$$

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

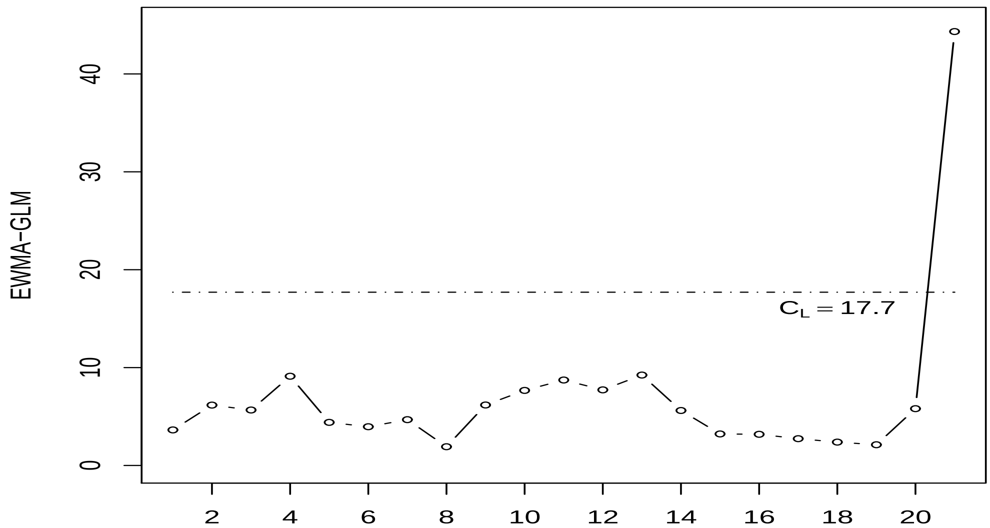
$$\Sigma = \begin{pmatrix} 17.77 \times 10^{-4} & 7.33 \times 10^{-4} \\ 7.33 \times 10^{-4} & 52.71 \times 10^{-4} \end{pmatrix}$$

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

Based on this estimated process model, we simulate new profiles, and in each profile sample, we have 100 observations $\boldsymbol{y}$ and 100 predictor variable vectors $\boldsymbol{x}$, which means $N = 100$ here. In addition, $n = 1$ in this example. The first 20 profiles are from in-control normal operational condition and the remaining profiles are from the OC condition. Two cases are considered here to illustrate the implementation of the proposed chart: 1) the shift $\delta = 0.1$ occurs in $\beta_1$; 2) the shift $\delta = 0.05$ occurs in the mean of $x_1$. The smoothing constant $\lambda$ is set as 0.2. We illustrate how to implement our proposed EWMA-GLM scheme to monitor the profiles:




(a)




(b)

Figure 3: The EWMA-GLM control chart for the AEC example: (a) shift in $\beta_1$ (b) shift in $\mu_1$

1. Obtain the control limits $L_M$ for the EWMA-GLM control chart by simulation to acheive the desired IC ARL. Here, we obtain the control limit CL = 17.7 for $ARL_0$=200 and then construct the control chart as seen in Figure 3.

2. Start monitoring the profiles. After obtaining the new observations, we calculate the control statistics with Eq.(4), and then plot these control statistics in 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 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 is important and challenging for monitoring profiles with categorical data and random predictor variables. In this paper, we used the GLM for modeling the relationship between the binary response variable and the random predictor variables. We proposed a novel control scheme, EWMA-GLM, for monitoring profiles with binary data and random explanatory variables. The EWMA-GLM scheme integrates the EWMA scheme and the logistic regression likelihood ratio test. As shown by the simulation results in this paper, the EWMA-GLM scheme performs almost always better than the Shewhart-GLM scheme and the NEWMA-GLM scheme, which are developed as the benchmark for performance comparison based on the 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 od historical observations used for estimating the IC parameters is sufficiently large. In practical applications, the performance of EWMA-GLM is 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. Secind, this new control scheme is proposed for the profile with logistic regression model. However, in real industries, different types of categorical data, e.g., multinomial data, exist. Therefore, new control schemes for monitoring profiles with other types of categorical data are interesting topics for further research. In fact, this amounts to adapting the general GLM model fitting to the proposed weighted likelihood ratio test. 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

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

The MLE of the model parameters $\boldsymbol{\xi} = (\alpha, \boldsymbol{\beta}^T)^T$ 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

$$\begin{aligned} z_i &= \eta_i + (y_i - \mu_{yi})\frac{\partial \eta_i}{\partial \mu_{yi}} = \eta_i + \frac{y_i - \mu_{yi}}{n_i p_i (1-p_i)}, \\ \eta_i &= \alpha + \mathbf{x}_i^T \boldsymbol{\beta}, \end{aligned}$$

where $i = 1, \ldots, N$, $\eta_i$ is defined as the linear predictor, and $\mu_{yi}$ 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 $\mathbf{z} = (z_1, \ldots, z_N)^T$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N)^T$, and $\boldsymbol{\mu}_y = (\mu_{y1}, \ldots, \mu_{yN})^T$. Let $\mathbf{X} = (\acute{\boldsymbol{x}}_1, \ldots, \acute{\boldsymbol{x}}_N)^T$, which is a $N \times (q+1)$ matrix and $\acute{\boldsymbol{x}}_i = (1, \mathbf{x}_i^T)^T$. 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 $l$th 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 $\epsilon$ 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 $l$th iteration.

**Appendix B: Derivation of the joint log-likelihood**

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

$$l_j = \log f(\mathbf{y}_j, \tilde{\mathbf{X}}_j) = \log f(\mathbf{y}_j|\tilde{\mathbf{X}}_j) f(\tilde{\mathbf{X}}_j) = \log f(\mathbf{y}_j|\tilde{\mathbf{X}}_j) + \log f(\tilde{\mathbf{X}}_j),$$

and

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

Thus, we have

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

**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 $\mu$ 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 $\widehat{\mathbf{X}}_t = (\mathbf{X}_{t-m+1}^T, \ldots, \mathbf{X}_t^T)^T$ be an $mN \times (q+1)$ matrix, which includes the most recent $m$ sets of explanatory variable values, $\widehat{\mathbf{z}}_t = (\mathbf{z}_{t-m+1}^T, \ldots, \mathbf{z}_t^T)^T$ 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

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

where

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

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

**References**:


Chicken, E., Pignatiello, J. J. JR., 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.

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.* Chapman and Hall, London.

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.* John Wiley & Sons, New York, NY.

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.

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 Binary Response". *IIE Transactions*, 41, pp. 931–941.

Zhang, H., and Albin, S. (2009)."Detecting Outliers in Complex Profiles using a $\chi^2$ 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.