International Journal of Industrial Engineering & Production Research

December 2012, Volume 23, Number 4
pp. 291-299

http://IJIEPR.iust.ac.ir/

ISSN: 2008-4889

Phase II Logistic Profile Monitoring

A. Saghaei*, M. Rezazadeh-Saghaei, R. Noorossana & M. Dorri

Abbas. Saghaei, *Industrial Engineering Department, Science and Research Branch, Islamic Azad University, Tehran, Iran*
Msryam Rezazadeh-Saghaei, *Parsian Quality and Productivity Research Center, Tehran, Iran.*
Rasoul Noorossana, *Industrial Engineering Department, Iran University of Science and Technology, Tehran,*
Mehdi Dorri, *Industrial Engineering Department, Islamic Azad University-South Tehran Branch, Tehran, Iran.*

KEYWORDS

Average run length,
Binary response,
Logistic regression model,
Profile monitoring,
Statistical process control

ABSTRACT

In many industrial and non-industrial applications the quality of a process or product is characterized by a relationship between a response variable and one or more explanatory variables. This relationship is referred to as profile. In the past decade, profile monitoring has been extensively studied under the normal response variable, but it has paid a little attention to the profile with the non-normal response variable. In this paper, the focus is especially on the binary response followed by the bernoulli distribution due to its application in many fields of science and engineering. Some methods have been suggested to monitor such profiles in phase I, the modeling phase; however, no method has been proposed for monitoring them in phase II, the detecting phase. In this paper, two methods are proposed for phase II logistic profile monitoring. The first method is a combination of two exponentially weighted moving average (EWMA) control charts for mean and variance monitoring of the residuals defined in logistic regression models and the second method is a multivariate T^2 chart to monitor model parameters. The simulation study is done to investigate the performance of the methods.

1. Introduction

The vast number of research studies on statistical process control (SPC) and particularly charting techniques demonstrates their importance in quality improvement for today's competitive industries. The quality of a process or product is characterized by univariate or multivariate quality characteristics. However, sometimes, a relationship between a response variable and one or more explanatory variables, referred to as profile, characterizes the quality of a process or product in a better way. For profile monitoring, one can measure the value of the response variable along with the corresponding values of one or more explanatory variables in order to evaluate the stability of profile relationship. The profile monitoring includes two phases. In phase I, the purpose is to evaluate the stability of a process and to estimate its parameters. In phase II, it is desirable to detect any change in the process parameters and variance of the profile as soon as possible. There are many studies on profile monitoring in which the response variable of interest follows the normal distribution in both phases I and II. Many authors including Mestek et al. [17], Stover and Brill [24], Lawless et al. [13], Kang and Albin [9], Mahmoud and Woodall [15], Woodall et al. [31], Wang and Tsung [27], Gupta et al. [6], Woodall [30], Zou et al. [33], and Jensen and Birch [7] have presented real-world example in which profile is applicable.

Mestek et al. [17], Stover and Brill [24], Mahmoud and Woodall [15] and Mahmoud et al. [14] all have focused

* **Corresponding author: Abbas. Saghaei**
Email: *. a.saghaei@srbiau.ac.ir*
Paper first received July. 05, 2012, and in revised form Oct. 9, 2012.

on phase I profile monitoring. Kang and Albin [9] proposed two methods for monitoring the linear profiles in both phases I and II: one is a multivariate T^2 chart for monitoring the model parameters and the other is a combination of the exponentially weighted moving average (EWMA) control chart and the range (R) control chart for monitoring the mean and variance of errors, respectively.

Kim et al. [12] coded the explanatory variable values in order to change the average to zero and make the model parameters independent. Then, they applied three EWMA charts to monitor a simple linear profile for detecting a shift in the intercept, slop and error variance.

They showed that their proposed method is superior to the one recommended by Kang and Albin [9]. There are also other papers investigating the phase II simple linear profile monitoring. Noorossana et al. [22] recommended a method based on the combination of multivariate cumulative sum (MCUSUM) chart and R chart. They showed that their proposed method detects small shifts in the intercept and slop more quickly. However, the performance of this method is not better than the previous methods in detecting error variance shifts.

In Gupta et al. [6] the performance of the method in Kim et al. [12] was compared with a method developed by Croakin and Varner [3]. In Zou et al. [34] a control chart based on change point model was proposed. Noorossana and Amiri [21] explained a method that applied MCUSUM chart and chi-square chart simultaneously. This method has the better performance in detecting shifts in error variance compared to the MCUSUM/R method. In Niaki et al. [20] a new method was recommended based on a generalized linear statistical model along with an R chart.

In Zou et al. [33], a multivariate exponentially weighted moving average (MEWMA) control chart was proposed based on the likelihood ratio statistics for monitoring the general linear profiles in phase II. Zou et al. [35] proposed a self starting Phase II control chart for monitoring the linear profiles based on the recursive residuals when the process parameters are not known. Woodall [30] reviewed the research on the use of control charts for profile monitoring. Kazemzadeh et al. [10, 11] proposed methods for monitoring the polynomial profile in phases I and II, respectively. Saghaei et al. [23] proposed three cumulative sum (CUSUM) control charts in order to monitor a shift in the parameters of the simple linear profile.

They compared the performance of their proposed method with the other existing methods and showed that their proposed method has remarkable performance in detecting a broad range of different kinds of model parameter shifts. Sometimes, the profile relationship can be represented by more complicated models than the linear one. Jin and Shi [8], Brill [2], Walker and Wright [26], Ding et al. [4], Williams et al. [28, 29], Moguerza et al. [18], Vaghefi et al. [25] and Jensen and Birch [7] have studied and investigated nonlinear profile monitoring. In many applications, the response variable can be a discrete variable. In particular, a binary response can be applied for classifying the products as defective or non-defective following the bernoulli or the binomial distribution. In these cases, a logistic regression model can be applied for characterizing the profile relationship which is called logistic profile in this paper. Yeh et al. [32] studied and extended profile monitoring under a logistic regression model in phase I and discussed the T^2 chart based on five different estimates of variance matrix of the model parameter. So far, no method has been proposed for monitoring the logistic profile in phase II. In this paper, two methods are proposed for monitoring the phase II logistic profile. The rest of this paper is outlined as follows:

The logistic profile is described in the next section. The proposed methods for monitoring the logistic profile are described in Section 3. Simulation studies and performance comparisons of proposed methods are presented in Section 4. Two practical examples are presented in Section 5. In the final section, simulation results are discussed and concluding remarks are presented.

2 Logistic Profile

Suppose that *j*th random sample is collected over time in phase II when the process is in statistical control. There are a set of observations $\{\boldsymbol{x}_i, y_{ij}\}_{i=1}^{n}$, each set of which consists of *p* independent regressor variables denoted by $\boldsymbol{x}_i = (x_{i1}, x_{i2}, ..., x_{ip})^T$, and one binary response variable denoted by y_{ij}. It is assumed that each y_{ij} follows the bernoulli distribution with $E\left(y_{ij}\right) = \mu_i = \pi_i$, where $\pi_i = \pi(\boldsymbol{x}_i)$ represents the probability of the bernoulli process as a function of $\boldsymbol{x}_i$. As mentioned above, it is assumed that π_i is a function of $\boldsymbol{x}_i$ which can be represented by a logistic regression model as $E\left(y_{ij}\right) = \pi\left(\boldsymbol{x}_i\right) = f\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_0\right)$, where $\boldsymbol{\beta}_0 = (\beta_{01}, \beta_{02}, ..., \beta_{0p})^T$ is the vector of the model parameters. Note that β_{01} is the intercept of the model, it means that $x_{i1} = 1$. There are various functions to represent the relationship between independent variables and response variables in a logistic model; a well-known of which is logit function. If this function is used, the logistic model would be written as:

$$\pi\left(\boldsymbol{x}_i\right) = 1/\left(1 + exp\left(-\boldsymbol{x}_i^T \boldsymbol{\beta}_0\right)\right) \tag{1}$$

where $\boldsymbol{x}_i^T \boldsymbol{\beta}_0$ is called linear predictor. But sometimes, in real cases, repeated observations or trials are made at each level of the $\boldsymbol{x}$ variable. Let y_{ij} represent the number of 1's observed for the ith observation in jth sample, and m_i represent the number of trials at each observation. Then, the logistic model becomes:

$$E\left(y_{ij}\right) = m_i \pi\left(\boldsymbol{x}_i\right) = m_i / (1+exp\left(-\boldsymbol{x}_i^T \boldsymbol{\beta}_0\right)) \tag{2}$$

The estimates of the linear predictor parameters from sample j can be obtained using the maximum likelihood method which is the theoretical basis for the parameter estimation in the logistic models. Since the observations are independent and follow the bernoulli distribution, the likelihood function will be formed in its usual manner. Because of the equality of the maximum likelihood estimations and weighted least squares estimation, Myers et al. [19] proposed using an iteratively reweighted least squares (IRWLS) method to solve the score equation in the estimation procedure of unknown logistic parameters. The log likelihood function of n independent y_{ij} s is expressed by:

$$ln\, L\left(\boldsymbol{y}_j : \boldsymbol{\beta}_j\right) = \sum_{i=1}^{n}\left[y_{ij}\, ln(\frac{\pi_j\left(\boldsymbol{x}_i\right)}{1-\pi_j\left(\boldsymbol{x}_i\right)})\right] + \sum_{i=1}^{n} m_i\, ln(1-\pi_j\left(\boldsymbol{x}_i\right)) \tag{3}$$

From Eq. 1, the term $ln(\frac{\pi_j\left(\boldsymbol{x}_i\right)}{1-\pi_j\left(\boldsymbol{x}_i\right)})$ is given by $\boldsymbol{x}_i^T \boldsymbol{\beta}_j$. Accordingly, the log likelihood function in Eq. 3 can be written as:

$$ln\, L\left(\boldsymbol{y}_j : \boldsymbol{\beta}_j\right) = \boldsymbol{\beta}_j^T \boldsymbol{X}\, \boldsymbol{y}_j - \sum_{i=1}^{n} m_i\, ln(1+exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_j\right)) \tag{4}$$

where $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)^T$ is an $n * p$ matrix of regressors.

To obtain the maximum likelihood estimates, the log likelihood in Eq. 4 must be maximized with respect to $\boldsymbol{\beta}_j$. After differentiating Eq. 4 with respect to $\boldsymbol{\beta}_j$, the score equation of:

$$\boldsymbol{X}^T\left(\boldsymbol{y}_j - \boldsymbol{\mu}_j\right) = 0 \tag{5}$$

can be obtained, where $\boldsymbol{\mu}_j = (\mu_{1j}, \mu_{2j}, \ldots, \mu_{nj})^T$ and $\mu_{ij} = m_i \pi_j\left(\boldsymbol{x}_i\right)$. Solving the score equation is nontrivial since $\boldsymbol{\mu}_j$ is a nonlinear function of $\boldsymbol{\beta}_j$. The model parameters can be estimated by applying the Newton-Raphson procedure which is expressed as IRWLS method. This procedure can be initialized the arbitrary value of $\boldsymbol{\beta}$ denoted by $\hat{\boldsymbol{\beta}}^0$, and then it can be continued using the following Newton-Raphson updating rule:

$$\hat{\boldsymbol{\beta}}^{t+1} = \hat{\boldsymbol{\beta}}^t + \left(\boldsymbol{X}^T \boldsymbol{W}^t \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T\left(\boldsymbol{y} - \boldsymbol{\mu}^t\right) \tag{6}$$

The estimate of model parameters updates until $\left\|\hat{\boldsymbol{\beta}}^{t+1} - \hat{\boldsymbol{\beta}}^t\right\| / \left\|\hat{\boldsymbol{\beta}}^t\right\| \leq \alpha$, which $\|.\|$ calculates the Euclidean norm and α is chosen to be a sufficiently small constant (e.g. $\alpha = 10^{-5}$).

McCullagh and Nelder [16] showed that when $\hat{\boldsymbol{\beta}}$ is the maximum likelihood estimate of the logistic model parameters and n or for a fixed n each m_i is large enough, $\hat{\boldsymbol{\beta}}$ follows a multivariate normal distribution asymptotically with $E\left(\hat{\boldsymbol{\beta}}\right) = \boldsymbol{\beta}_0$ and $Var\left(\boldsymbol{\beta}\right) = (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1}$, where $\boldsymbol{W}$ is a diagonal matrix with $m_i \pi\left(\boldsymbol{x}_i\right)(1-\pi\left(\boldsymbol{x}_i\right))$ as the ith diagonal element.

Like any other linear and non-linear models, the analysis of residuals is important in a logistic model. Residuals are ordinarily shown as the difference between the reference model and sample profile. This kind of residuals is only useful in the case of homogeneous variance. For example, in the linear and non-linear normal profiles, the mentioned residual variables are independent and follow the normal distribution with mean zero and constant variance σ^2. But in the logistic model, because of its dependency on the mean value, the variance values of the observations (which are the same as the residual variance) change for the different levels of $\boldsymbol{x}$ variable. Therefore, to avoid this problem, other kinds of residuals have been proposed.

McCullagh and Nelder [16] mentioned the Pearson and Anscombe residuals that can be used as standard normal residuals for logistic models similar to the ordinary residuals employed in the standard regression. The Pearson residual is defined as follow:

$$R_{ij}^p = \frac{y_{ij} - \mu_i}{\sqrt{Var(y_{ij})}}, \tag{7}$$

where $\mu_i = m_i \pi\left(\boldsymbol{x}_i\right)$ is the predicted value of response variables in the ith level of c variable and $Var\left(y_{ij}\right) = m_i \pi\left(\boldsymbol{x}_i\right)(1-\pi\left(\boldsymbol{x}_i\right))$ in the logistic models.

The Anscombe residual is computed by the following expression:

$$R_{ij}^a = \frac{A\left(y_{ij}\right) - A(\mu_i)}{A'(\mu_i)\sqrt{Var(y_{ij})}} \tag{8}$$

where the function $A(.)$ is given by:

$$A(.) = \int \frac{d\mu}{Var^{1/3}(\mu)} \tag{9}$$

For the binomial distribution, the Anscombe residual takes the form of:

$$R_{ij}^{a} = \sqrt{m_i}\,\frac{B\left(\frac{2}{3},\frac{2}{3}\right)\left(I_{\frac{y_{ij}}{m_i}}\left(\frac{2}{3},\frac{2}{3}\right) - I_{\pi_i}\left(\frac{2}{3},\frac{2}{3}\right)\right)}{(\pi_i(1-\pi_i))^{1/6}}, \tag{10}$$

where $B(a,b)$ and $I_x(a,b)$ are calculated through:

$$B(a,b) = (\Gamma(a)\Gamma(b))/\Gamma(a+b), I_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1}dt \tag{11}$$

3 Proposed Methods

3.1 EWMA2 Method

The difference between observed and predicted values for a given level of independent variables follows a normal distribution with the mean of zero and a constant variance, $e_{tj} \approx N(\mathbf{0}, \sigma_j)$, in a profile with the normal response variable.

In the logistic profile, ordinary residuals have the mean of zero and a variable variance, $e_{tj} \approx (\mathbf{0}, m_j \pi_{ij}(\mathbf{1} - \pi_{ij}))$; therefore, using some residuals defined in the logistic regression model such as Pearson and Anscombe residuals is proposed. In the first method, two EWMA charts are proposed for the phase II monitoring of logistic profiles; one is to monitor the mean of residuals like the EWMA/R method in Kang and Albine [9] and another is to monitor the variance of residuals based on the inverse normal transformation proposed by Acosta-Mejia et al. [1]. If MSE_j represents the variance of the *j*th sample and:

$$P_{\sigma_j} = \Phi^{-1}\left(F_{\chi_n^2}\left(n\,MSE_j / \sigma_0^2\right)\right), \tag{12}$$

then $nMSEj / \sigma_0^2$ and P_{σ_j} would follow a chi-square distribution with n degrees of freedom and a standard normal distribution, respectively. In Eq. 12, $F_{\chi_n^2}(.)$ is the cumulative distribution function (cdf) for the chi-square distribution with n degrees of freedom and $\Phi(.)$ is the cdf for the standard normal distribution. An increase (decrease) in the standard deviation of residuals results in an increase (decrease) in the mean of P_{σ_j}.

3.1.1 Pearson Residual-Based Method

EWMA2 method based on Pearson residuals, $EWMA^p2$, uses the following $EWMA_{M,j}^p$ and $EWMA_{E,j}^p$ statistics for monitoring the mean and variance of Pearson residuals, respectively.

$$EWMA_{M,j}^p = \theta \bar{R}_j^p + (1-\theta) EWMA_{M,j-1}^p, \tag{13}$$

$$EWMA_{E,j}^p = \theta P_{\sigma_j} + (1-\theta) EWMA_{E,j-1}^p, \tag{14}$$

where $\bar{R}_j^p = \sum_{i=1}^{n} R_{ij}^p / n$, $EWMA_{M,0}^p = EWMA_{E,0}^p = 0$, and θ is a smoothing constant, $0 < \theta \le 1$. The upper and lower control limits for the charts are given by:

$$UCL_M^p = L_M^p \sqrt{\theta / ((2-\theta)n)}, LCL_M^p = -UCL_M^p, \tag{15}$$

$$UCL_E^p = L_E^p \sqrt{\theta / (2-\theta)}, LCL_E^p = -UCL_E^p. \tag{16}$$

The multipliers $L_M^p (>0)$ and $L_E^p (>0)$ are chosen in order to give a specified in-control ARL.

3.1.2 Anscombe Residual-Based Method

EWMA2 method based on Anscombe residuals, $EWMA^a2$, uses the following $EWMA_{M,j}^a$ and $EWMA_{E,j}^a$ statistics for monitoring the mean and variance of Anscombe residuals, respectively.

$$EWMA_{M,j}^a = \theta \bar{R}_j^a + (1-\theta) EWMA_{M,j-1}^a \tag{17}$$

$$EWMA_{E,j}^a = \theta P_{\sigma_j} + (1-\theta) EWMA_{E,j-1}^a, \tag{18}$$

where $\bar{R}_j^a = \sum_{i=1}^{n} R_{ij}^a / n$, $EWMA_{M,0}^a = EWMA_{E,0}^a = 0$, and θ is a smoothing constant, $0 < \theta \le 1$. The upper and lower control limits for the charts are given by:

$$UCL_M^a = L_M^a \sqrt{\theta / ((2-\theta)n)}, LCL_M^a = -UCL_M^a, \tag{19}$$

$$UCL_E^a = L_E^a \sqrt{\theta(2-\theta)}, LCL_E^a = -UCL_E^a. \tag{20}$$

The multipliers $L_M^a (>0)$ and $L_E^a (>0)$ are chosen in order to give a specified in-control ARL.

3.2 T^2 method

The model parameters, $\boldsymbol{\beta}_j$, follow a multivariate normal distribution asymptotically with $E(\boldsymbol{\beta}_j) = \boldsymbol{\beta}_0$

and $Var\left(\boldsymbol{\beta}_j\right) = (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1}$, (see Section 2). The following T^2 statistic is proposed for monitoring the logistic profile in phase II:

$$T_j^2 = (\boldsymbol{\beta}_j - \boldsymbol{\beta}_0)^T (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})(\boldsymbol{\beta}_j - \boldsymbol{\beta}_0) \quad (21)$$

T_j^2 follows a chi-square distribution with p degrees of freedom. The UCL for the given False Alarm Rate α is $UCL_{T^2} = \chi^2_\alpha(p)$.

4. Simulation Studies

In this section, a simple logistic model is considered for comparing the performance of the proposed methods used by Yeh et al. [32] in their simulation study. In this model, the probability parameter equals the following equation:

$$\pi\left(\boldsymbol{x}_i\right) = 1/(1+exp(-\beta_1 + \beta_2 \boldsymbol{x}_i)) = 1/(1+exp\left(-(3+2\boldsymbol{x}_i)\right)) \quad (22)$$

The performance of the proposed control charts was investigated based on the values of out-of-control ARL under positive and negative shifts from β_1 to $\beta_1 \pm \gamma * \sigma_{\beta_1}$ and β_2 to $\beta_2 \pm \gamma * \sigma_{\beta_2}$. The values of response variable were measured along with the corresponding values of one explanatory variable, $\boldsymbol{x}_i = (.1, .2, .3, ..., .9)$. The values of $\boldsymbol{x}_i$ can be replaced by the transformed logarithmic values. This transforming is often done and is effective when the range of values is quite large (see Finney (1950)). Therefore, the matrix $\boldsymbol{X}$ was obtained:

$$\boldsymbol{X} = \begin{pmatrix} 1 & \dots & 1 \\ log(.1) & \dots & log(.9) \end{pmatrix}$$

It was also assumed that each independent variable level had m replications and the performance of proposed control charts was investigated for different values of m for $m = 30, 60, 100$. Ten thousand vectors of $\boldsymbol{y}$ were drawn from the binomial distribution in order to compute the ARL values.

The smoothing constants in EWMA2 method were set to 0.2. The combination of two EWMA control charts had an overall in-control ARL of roughly 200 and each of the two charts had the in-control ARL of approximately 390. The UCL_{T^2} was also set in order to give an in-control ARL of roughly 200.

The values of UCL^p_M, UCL^p_E, UCL^a_M, UCL^a_E and UCL_{T^2} are summarized in Table 1. The simulation results under positive and negative shifts are illustrated in Tables 2-5.

Tab. 1. Values of upper control limits

Replication Number	Control Limit				
	UCL^a_M	UCL^a_E	UCL^p_M	UCL^p_E	UCL_{T^2}
m=100	0.3513	0.9787	0.3395	0.9633	11.3700
m=60	0.3613	1.0133	0.3396	0.9667	11.9600
m=30	0.3903	1.1083	0.3398	0.9700	13.4532

91, pp. 68-83

Tab. 2. ARL comparisons under positive shifts from β_1 to $\beta_1 + \gamma * \sigma_{\beta_1}$

Replication Number	Proposed method	γ 0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0
m=100	$\gamma * \sigma_{\beta_1}$	0.0541	0.1082	0.1624	0.2165	0.2706	0.3247	0.3789	0.433	0.4871	0.5412
	T^2	88.2419	27.8901	9.8345	4.2019	**2.3291**	**1.5234**	**1.2069**	**1.0715**	**1.02**	**1.0046**
	$EWMA^p 2$	26.8163	7.8607	4.3971	3.131	2.4893	2.1201	1.9003	1.7017	1.4879	1.2807
	$EWMA^a 2$	**22.0075**	**7.303**	**4.2523**	**3.0516**	2.4543	2.0931	1.862	1.6716	1.4529	1.2668
m=60	$\gamma * \sigma_{\beta_1}$	0.0699	0.1397	0.2096	0.2795	0.3494	0.4192	0.4891	0.559	0.6289	0.6987
	T^2	93.1249	30.4103	10.592	4.5811	2.4979	**1.6071**	**1.2429**	**1.0886**	**1.0269**	**1.0068**
	$EWMA^p 2$	27.0174	7.858	4.4449	3.1726	2.5211	2.139	1.9303	1.7443	1.5477	1.3476
	$EWMA^a 2$	**21.9159**	**7.2989**	**4.3506**	**3.1155**	**2.4889**	2.1224	1.9055	1.719	1.5239	1.3294
m=30	$\gamma * \sigma_{\beta_1}$	0.0988	0.1976	0.2964	0.3953	0.4941	0.5929	0.6917	0.7905	0.8893	0.9882
	T^2	98.7887	34.3091	13.0109	5.418	2.8838	**1.8216**	**1.3514**	**1.1475**	**1.0462**	**1.0135**
	$EWMA^p 2$	27.5923	8.1341	4.5924	**3.2369**	**2.5824**	2.2129	1.9893	1.8296	1.657	1.4544
	$EWMA^a 2$	**22.0881**	**7.6445**	**4.5084**	3.268	2.6203	2.243	2.0081	1.8566	1.6886	1.5094

Tab. 3. ARL comparisons under negative shifts from β_1 to $\beta_1 - \gamma * \sigma_{\beta_1}$

Replication Number	Proposed method	γ 0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0
m=100	$\gamma * \sigma_{\beta_1}$	0.0541	0.1082	0.1624	0.2165	0.2706	0.3247	0.3789	0.433	0.4871	0.5412
	T^2	121.2224	36.7556	11.7589	4.6285	2.3806	**1.5139**	**1.1845**	**1.0521**	**1.0111**	**1.0023**
	$EWMA^p 2$	**25.0241**	**7.373**	**4.1079**	**2.8983**	**2.2724**	1.9179	1.6603	1.4026	1.1967	1.0586
	$EWMA^a 2$	36.4613	8.768	4.5503	3.118	2.4082	2.034	1.7854	1.5396	1.2926	1.116
m=60	$\gamma * \sigma_{\beta_1}$	0.0699	0.1397	0.2096	0.2795	0.3494	0.4192	0.4891	0.559	0.6289	0.6987
	T^2	134.8965	44.9237	14.1218	5.3678	2.5762	**1.609**	**1.2119**	**1.0661**	**1.0149**	**1.0017**
	$EWMA^p 2$	**24.7476**	**7.219**	**4.0371**	**2.8544**	**2.2497**	1.8916	1.625	1.3729	1.1677	1.051
	$EWMA^a 2$	44.615	9.3759	4.7786	3.2174	2.4826	2.0674	1.8099	1.5658	1.307	1.1305
m=30	$\gamma * \sigma_{\beta_1}$	0.0988	0.1976	0.2964	0.3953	0.4941	0.5929	0.6917	0.7905	0.8893	0.9882
	T^2	159.8154	61.272	20.1896	7.2575	3.3014	1.8482	**1.319**	**1.094**	**1.0239**	**1.0044**
	$EWMA^p 2$	**24.1641**	**7.2091**	**3.9689**	**2.78**	**2.1847**	**1.8431**	1.5596	1.306	1.1259	1.0384
	$EWMA^a 2$	72.7151	11.7829	5.4231	3.5265	2.6517	2.179	1.9064	1.6688	1.3891	1.1851

Tab. 4. ARL comparisons under positive shifts from β_2 to $\beta_2 + \gamma * \sigma_{\beta_2}$

Replication Number	Proposed method	γ 0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0
m=100	$\gamma * \sigma_{\beta_2}$	0.043	0.0861	0.1291	0.1722	0.2152	0.2583	0.3013	0.3444	0.3874	0.4304
	T^2	87.0811	27.2385	9.6438	4.2257	**2.2901**	**1.5342**	**1.2048**	**1.0746**	**1.0208**	**1.006**
	$EWMA^p 2$	**39.4069**	**11.4047**	**6.0744**	**4.0689**	3.1437	2.6027	2.246	2.0157	1.8254	1.6731
	$EWMA^a 2$	61.4691	14.4216	6.8278	4.4383	3.3381	2.7026	2.3443	2.0745	1.8881	1.7349
m=60	$\gamma * \sigma_{\beta_2}$	0.0556	0.1111	0.1667	0.2223	0.2779	0.3334	0.389	0.4446	0.5001	0.5557
	T^2	89.8738	29.1247	10.2943	4.5671	**2.4491**	**1.609**	**1.2456**	**1.096**	**1.0301**	**1.0063**
	$EWMA^p 2$	**40.1046**	**11.6227**	**6.0596**	**4.1451**	3.1717	2.6063	2.2757	2.0249	1.856	1.7056
	$EWMA^a 2$	76.6358	16.3345	7.443	4.7469	3.534	2.8561	2.4319	2.1517	1.9606	1.8024
m=30	$\gamma * \sigma_{\beta_2}$	0.0786	0.1572	0.2358	0.3144	0.3929	0.4715	0.5501	0.6287	0.7073	0.7859
	T^2	98.1847	33.048	12.416	5.2574	**2.9014**	**1.8133**	**1.3655**	**1.1515**	**1.0568**	**1.0177**
	$EWMA^p 2$	**39.8003**	**11.8934**	**6.1737**	**4.1841**	3.2403	2.6726	2.3143	2.0763	1.901	1.7504
	$EWMA^a 2$	138.6959	23.9356	9.3711	5.6084	4.0373	3.2106	2.6902	2.3483	2.1249	1.9748

Tab. 5. ARL comparisons under negative shifts from β_2 to $\beta_2 - \gamma * \sigma_{\beta_2}$

Replication Number	Proposed method	γ 0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0
m=100	$\gamma * \sigma_{\beta_2}$	0.043	0.0861	0.1291	0.1722	0.2152	0.2583	0.3013	0.3444	0.3874	0.4304
	T^2	126.7785	38.4535	12.1394	4.7103	**2.3901**	**1.5209**	**1.1696**	**1.0516**	**1.0113**	**1.0011**
	$EWMA^p 2$	39.5337	11.0306	5.7803	3.85	2.9479	2.4313	2.0983	1.8784	1.6754	1.4854
	$EWMA^a 2$	**32.8488**	**10.1375**	**5.5674**	**3.826**	2.9699	2.4486	2.1234	1.912	1.7265	1.5421
m=60	$\gamma * \sigma_{\beta_2}$	0.0556	0.1111	0.1667	0.2223	0.2779	0.3334	0.389	0.4446	0.5001	0.5557
	T^2	137.8975	46.076	14.2224	5.3506	**2.5755**	**1.5858**	**1.2109**	**1.0534**	**1.0116**	**1.0017**
	$EWMA^p 2$	39.6764	10.885	5.6571	**3.8333**	2.9064	2.4209	2.09	1.8717	1.6829	1.4664
	$EWMA^a 2$	**31.7437**	**10.0206**	**5.5568**	3.846	2.9951	2.4681	2.1471	1.9401	1.7645	1.5793
m=30	$\gamma * \sigma_{\beta_2}$	0.0786	0.1572	0.2358	0.3144	0.3929	0.4715	0.5501	0.6287	0.7073	0.7859
	T^2	166.2769	64.3396	20.7275	7.2055	3.2574	**1.8121**	**1.2885**	**1.088**	**1.021**	**1.0036**
	$EWMA^p 2$	39.8223	10.8734	**5.6287**	**3.8009**	**2.899**	2.3889	2.0821	1.8654	1.6733	1.4779
	$EWMA^a 2$	**31.3275**	**10.3276**	5.7776	4.0025	3.0995	2.5578	2.2211	2.0065	1.8454	1.6971

5. An Example

This section gives an example taken from a study of the press machine monitoring. The relationship between the percentage of defective products and the speed of a press machine can be modeled by a logistic profile.

One percent of products are defective when a special press machine works in normal speed. The speed increase (decrease) led to increase (decrease) the percentage of the defective products. For example, the percentage of defective products decreases by 0.5% when the speed of the press machine decreases by 1/4-fold. The following table gives more information on this relationship. The probability of defective product is the long-term average probability observed at the certain level of the speed for the samples of 100 products which is shown in Table 6.

Tab. 6. The probability of defective productive at the certain level of the machine speed

Speed of the press machine	probability of defective products
0.25	0.005
0.50	0.006
0.75	0.008
1.00	0.010
1.30	0.015
1.50	0.019
1.80	0.026
2.00	0.035

Based on the logistic model, the first-order model was fitted to data and the model parameters were estimated by the Newton-Raphson procedure. As a result, the probability of the bernoulli process (being defect or non-defect product) and the variance matrix were obtained:

$$\pi\left(\boldsymbol{x}_i\right)=1/(1+exp(\boldsymbol{x}_i^T\boldsymbol{\beta}))=1/(1+exp\left(5.702-1.174x_i\right))$$

$$\mathrm{Var}\left(\boldsymbol{\beta}\right)=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{X})^{-1}=\begin{pmatrix} .0621 & -.1658 \\ -.1658 & .6226 \end{pmatrix}$$

Notice that the matrix $\boldsymbol{X}$ was considered without logarithmic transformation of the $\boldsymbol{x}$ values. EWMA2 method based on Pearson residual selected for monitoring the model parameters over time in order to detect a shift in phase II control. The upper control limits for the $EWMA_M^p$, $EWMA_E^p$ charts were set to equal 2.87, 3.14, respectively, for obtaining an overall in-control ARL of roughly 200. When the sample *j* is collected, one can calculate the residual values and compute P_{σ_j} by Eq. 12. The *j*th plotting statistics are computed by Eqs. 14 and 15. Figure 1 shows the 127 plotted statistics on the $EWMA2^p$ control charts. The point 89 is out of control limits plotted on both the $EWMA_E^p$ and the $EWMA_M^p$ control charts. Analytical results showed that an assignable cause occurred on the 84th day and the process was out-of-control.

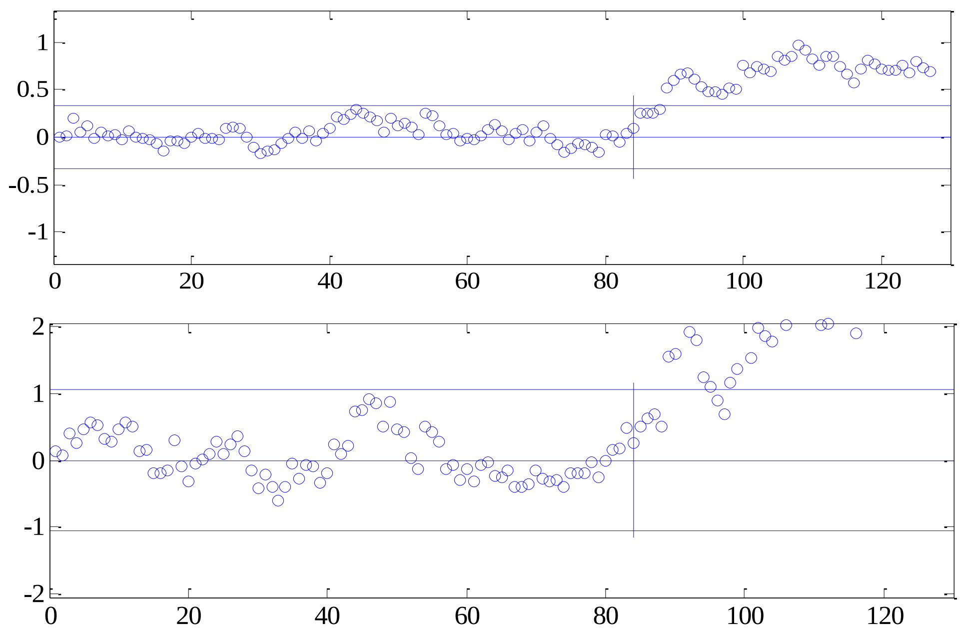

Fig. 1. The plots show the $EWMA2^p$ charts over a 127 days period time. The first point that signaled an out-of-control state is point 89 on both the $EWMA_M^p$ chart (upper) and the $EWMA_E^p$ chart (lower).

6 Conclusions

In this paper, two methods were proposed for phase II monitoring of profiles under the assumption that the response variable is binary. One is a combination of two EWMA control charts for the mean and variance monitoring of the residuals defined in logistic regression models. The details were represented for two different kinds of these residuals. The other method is a multivariate T^2 control chart for monitoring the logistic regression model parameters. The performance of the methods was compared in terms of ARL criterion.

The simulation studies showed that the performance of EWMA2 method is superior to T^2 method when the step shift is small. The T^2 method performs better than the EWMA2 method in the case of the large step shift; however, the performance of the methods is close to each other.

The simulations were run for different values of replication number of the bernoulli trials. The results showed that increasing the bernoulli trials number led to increasing the performance of two methods. Comparing the performance of two residuals showed that the EWMA2 method based on Pearson residual is robust; therefore, applying the Pearson residual is preferred in the cases which the replication number is small.

References

[1] Acosta-Mejia, C.A., Pignatiello, J.J., Rao, B.V., *"A Comparison of Control Charting Procedures for Monitoring Process Dispersion*," IIE Transactions, Vol. 31, 1999, pp. 569-579.

[2] Brill, R.V., *"A Case Study for Control Charting a Product Quality Measure that is a Continuous Function Over Time*," In Proceedings of the 45th Annual Fall Technical Conference, Toronto, Ontario, 2001.

[3] Croarkin, C., Varner, R., *"Measurement Assurance for Dimensional Measurements on Integrated-Circuit Photomasks,"* NBS Technical Note 1164, US Department of Commerce, Washington, DC. USA, 1982.

[4] Ding, Y., Zeng, L., Zhou, S., *"Phase I Analysis for Monitoring Nonlinear Profiles in Manufacturing Processes,"* J. Qual. Technol., Vol. 38, 2006, pp. 199–216.

[5] Finney, D.J., Probit Analysis, Cambridge University Press, Cambridge, 1950.

[6] Gupta, S., Montgomery, D.C., Woodall, W.H., *"Performance Evaluation of Two Methods for Online Monitoring of Linear Calibration Profiles*," International Journal of Production Research, Vol. 44, 2006, pp. 1927-1942.

[7] Jensen, W.A., Birch, J.B., *"Profile Monitoring Via Nonlinear Mixed Models* J. Qual. Technol., Vol. 41, 2009, pp. 18–34.

[8] Jin, J., Shi, J., "*Feature Preserving data Compression of Stamping Tonnage Information using Wavelets*," Technometrics, Vol. 41, 1999, pp. 327–339.

[9] Kang, L., Albin, S.L., "*On-Line Monitoring When the Process Yields a Linear Profile*," J. Qual. Technol., Vol. 32, 2000, pp. 418-426.

[10] Kazemzadeh, R.B., Noorossana, R., Amiri, A., "*Phase I Monitoring of Polynomial profiles*," Communications in Statistics Theory and Methods, Vol. 37, 2008, pp. 1671–1686.

[11] Kazemzadeh, R.B., Noorossana, R., Amiri, A., "*Monitoring Polynomial Profiles in Quality Control Applications*," The International Journal of Advanced Manufacturing Technology, Vol. 42, 2009, pp. 703–712.

[12] Kim, K., Mahmoud, M.A., Woodall, W.H., *"On the Monitoring of Linear Profiles,"* J. Qual. Technol., vol. 35, 2003, pp. 317-328.

[13] Lawless, J.F., MacKay, R.J., Robinson, J.A., "*Analysis of Variation Transmission in Manufacturing Processes-Part I*," J. Qual. Technol., Vol. 31, 1999, pp. 131-142.

[14] Mahmoud, M.A., Parker, P.A., Woodall, W.H., Hawkins, D.M., *"A Change Point Method for Linear Profile Data*," Quality and Reliability Engineering International, Vol. 23, 2007, pp. 247-268.

[15] Mahmoud, M.A., Woodall, W.H., *"Phase I Analysis of Linear Profiles with Calibration Applications*," Technometrics, Vol. 46, 2004, pp. 380-391.

[16] McCullagh, P., Nelder, J.A., *Generalized Linear Models*, Chapman and Hall, London, 1989.

[17] Mestek, O., Pavlik, J., Suchanek, M., *"Multivariate Control Charts: Control Charts for Calibration Curves*," Fresenius Journal of Analytical Chemistry, Vol. 350, 1994, pp. 344-351.

[18] Moguerza, J.M., Munoz, A., Psarakis, S., *"Monitoring Nonlinear Profiles using Support Vector Machines*," In Proceedings of the 12th Iberoamerican Congress on Pattern Recognition, Valparaiso, Chile, 2007.

[19] Myers, R.H., Montgomery, D.C., Vining, G.G., *"Generalized Linear Models with Application in Engineering and the Sciences*," John Wiley and Sons, New York, 2001.

[20] Niaki, STA., Abbasi, B., Arkat, J., *"A Generalized Linear Statistical Model Approach to Monitor Profiles*," Int. J. Eng. Trans. A Basics, Vol. 20, 2007, pp. 233–242.

[21] Noorossana, R., Amiri, A., "*Enhancement of Linear Profiles Monitoring in Phase II*," AmirKabir J. Sci. Technol., Vol. 18, pp. 19–27, 2007, in Farsi.

[22] Noorossana, R., Amiri, A., Vaghefi, S.A., Roghanian, E., *"Monitoring Quality Characteristics Using Linear Profile,"* In Proceedings of the 3rd International Industrial Engineering Conference, Tehran, Iran, 2004.

[23] Saghaei, A., Mehrjoo, M., Amiri, A., "*A CUSUM-Based Method for Monitoring Simple Linear Profiles*," The International Journal of Advanced Manufacturing Technology, Vol. 45, 2009, pp. 1252–1260.

[24] Stover, F.S., Brill, R.V., *"Statistical Quality Control Applied to Ion Chromatography Calibrations*," Journal of Chromatography A, Vol. 804, 1998, pp. 37–43.

[25] Vaghefi, A., Tajbakhsh,V., Noorossana, R., "*Phase II Monitoring of Nonlinear Profiles,"* Communications in Statistics–Theory and Methods, Vol. 38, 2009, pp. 1834–1851.

[26] Walker, E., Wright, S.P., "*Comparing Curves using Additive Models*," J. Qual. Technol., Vol. 34, 2002, pp. 118–129.

[27] Wang, K., Tsung, F., *"Using Profile Monitoring Techniques for a Data-Rich Environment with Huge Sample Size*," Quality and Reliability Engineering International, Vol. 21, 2005, pp. 677-688.

[28] Williams, J.D., Birch, J.B., Woodall, W.H., Ferry, N.M., *"Statistical Monitoring of Heteroscedastic Dose-Response Profiles from High-Throughput Screening*," Journal of Agricultural, Biological and Environmental Statistics, Vol. 12, 2007, pp. 216-235.

[29] Williams, J.D., Woodall, W.H., Birch, J.B., *"Statistical Monitoring of Nonlinear Product and Process Quality Profiles,"* Quality and Reliability Engineering International, Vol. 23, 2007, pp. 925–941.

[30] Woodall, W.H., *"Current Research on Profile Monitoring*," Rev. Producao, Vol. 17, 2007, pp. 420–425.

[31] Woodall, W.H., Spitzner, D.J., Montgomery, D.C., Gupta, S., *"Using Control Charts to Monitor Process and Product Quality Profiles*," J. Qual. Technol., Vol. 36, 2004, pp. 309–320.

[32] Yeh, A.B., Huwang, L., Li, Y.M., "*Profile Monitoring for a Binary Response*," IIE Transactions, Vol. 41, 2009, pp. 931– 941.

[33] Zou, C., Tsung, F., Wang, Z., "*Monitoring General Linear Profiles using Multivariate Exponentially Weighted Moving Average Schemes*," Technometrics, Vol. 49, 2007, pp. 395–408.

[34] Zou, C., Zhang, Y., Wang, Z., *"Control Chart Based on Change Point Model for Monitoring Linear Profiles*," IIE Transactions, Vol. 38, 2006, pp. 1093–1103.

[35] Zou, C, Zhou, C., Wang, Z., Tsung, F., "*A Self-Starting Control Chart for Linear Profiles,*" J. Qual. Technol., Vol. 39, 2007, pp. 364-375.