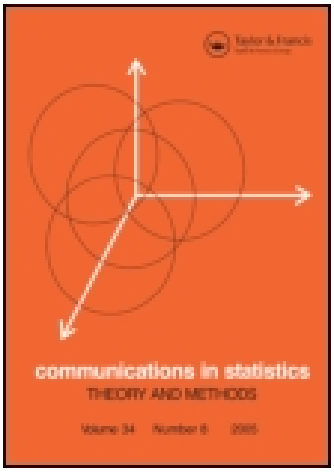

Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Diagnosis Aids in Multivariate Multiple Linear Regression Profiles Monitoring

Amirhossein Amiri[a], Abbas Saghaei[b], Mohammad Mohseni[c] & Yaser Zerehsaz[d]

[a] Industrial Engineering Department, Faculty of Engineering, Shahed University, Tehran, Iran

[b] Industrial Engineering Department, Islamic Azad University-Science and Research Branch, Tehran, Iran

[c] Industrial Engineering Department, Islamic Azad University-Qazvin Branch, Qazvin, Iran

[d] Industrial Engineering Department, Islamic Azad University-South Tehran Branch, Tehran, Iran

Accepted author version posted online: 10 Oct 2013.

To cite this article: Amirhossein Amiri, Abbas Saghaei, Mohammad Mohseni & Yaser Zerehsaz (2014) Diagnosis Aids in Multivariate Multiple Linear Regression Profiles Monitoring, Communications in Statistics - Theory and Methods, 43:14, 3057-3079

To link to this article: http://dx.doi.org/10.1080/03610926.2012.694543

Communications in Statistics—Theory and Methods, 43: 3057–3079, 2014

ISSN: 0361-0926 print / 1532-415X online
DOI: 10.1080/03610926.2012.694543

Diagnosis Aids in Multivariate Multiple Linear Regression Profiles Monitoring

AMIRHOSSEIN AMIRI,[1] ABBAS SAGHAEI,[2] MOHAMMAD MOHSENI,[3] AND YASER ZEREHSAZ[4]

[1]Industrial Engineering Department, Faculty of Engineering, Shahed University, Tehran, Iran
[2]Industrial Engineering Department, Islamic Azad University-Science and Research Branch, Tehran, Iran
[3]Industrial Engineering Department, Islamic Azad University-Qazvin Branch, Qazvin, Iran
[4]Industrial Engineering Department, Islamic Azad University-South Tehran Branch, Tehran, Iran

Diagnosis aids in addition to detecting the out-of-control state is an important issue in multivariate multiple linear regression profiles monitoring; because a large number of parameters and profiles in this structure are involved. In this paper, we specifically concentrate on identification of profile(s) and parameter(s) which have changed during the process in multivariate multiple linear regression profiles structure in Phase II. We demonstrate the effectiveness of our proposed approaches through Monte Carlo simulations and a real case study in terms of accuracy percent.

Keywords Diagnosis aids; Hoteling's T^2 statistic; Multivariate multiple linear regression profile; Phase II; Statistical process control.

Mathematics Subject Classification Primary 62J05; Secondary 62J20.

1. Introduction

In many practical applications of statistical process control (SPC), the quality of a product or process is characterized by multiple measurements constituting a line or curve that is referred to as a profile. The applications of profile monitoring have been introduced by several authors such as Kang and Albin (2000), Mahmoud and Woodall (2004), Woodall et al. (2004), Wang and Tsung (2005), Montgomery (2005), Woodall (2007), Zou et al. (2007), and Amiri et al. (2010). Different types of linear regression profiles such as simple linear regression, multiple and polynomial regression profiles have been investigated by many authors. Kang and Albin (2000), Kim et al. (2003), Gupta et al. (2006), Zou et al.

Received March 2, 2011; Accepted May 14, 2012.
Address correspondence to Amirhossein Amiri, Industrial Engineering Department, Faculty of Engineering, Shahed University, Tehran, Iran. E-mail: amiri@shahed.ac.ir
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lsta.

(2007), Zhang et al. (2009), and Saghaei et al. (2009) have studied Phase II monitoring of simple linear profiles. Phase I monitoring of simple linear profiles has also been investigated by researchers such as Mestek et al. (1994), Stover and Brill (1998), Kang and Albin (2000), Mahmoud and Woodall (2004), and Mahmoud et al. (2007). Multiple linear regression and polynomial profiles have been studied by Zou et al. (2007), Mahmoud (2008), and Kazemzadeh et al. (2008, 2009). In some situations, quality of a process or a product can be effectively characterized by two or more multiple linear regression profiles in which response variables are correlated referred to as multivariate multiple linear regression profiles. Noorossana et al. (2010a) and Eyvazian et al. (2011) have studied Phase I and Phase II multivariate multiple linear regression profiles, respectively. Moreover, multivariate simple linear profiles monitoring in Phase II has also been studied by Noorossana et al. (2010b).

In addition, diagnosing the out-of-control parameters is an important issue in multivariate statistical process control to help practitioners to identify the root causes and assignable ones quickly. The diagnosis has a deep literature in the area of multivariate SPC. Alt (1985) suggested the use of Bonferroni limits for this purpose. Hayter and Tsui (1994) extended the idea of Bonferroni-type control limits by giving a procedure for exact simultaneous control intervals for each of the variable means, using simulation. Jackson (1991) discussed the use of an elliptical control region for the process with two quality characteristics. An extension of the elliptical control region as a solution to the interpretation problem is given by Chua and Montgomery (1992). Today the use of T^2 decomposition proposed by Mason et al. (1995) is considered as the most valuable. The main idea of this method is to decompose the T^2 statistic into independent terms, each of which reflects the contribution of an individual variable. The problem with this method is that the decomposition of the T^2 statistic into p independent T^2 components is not unique. Thus, Mason et al. (1997) give an appropriate computing scheme that can greatly reduce the computational effort. The methodologies of Murphy (1987), Doganaksoy et al. (1991), and Runger et al. (1996) are special cases of Mason's et al. (1995) partitioning of T^2 statistic. Doganaksoy et al. (1991) classified the components of an observation vector by providing the relative contribution of the signal as a function of the univariate criteria of the Student's t-statistic. Hawkins (1993) and Wade and Woodall (1993) developed a T^2 statistic diagnosis technique using the adjusted regression of individual variables to detect the signal. Runger et al. (1996) proposed the use of different metric distances to decide which variables have shifted. Jackson (1991) proposed the use of principal components for monitoring a multivariate process. Since the principal components are uncorrelated, they may provide some insight into the nature of the out-of-control condition and then lead to the examination of particular original observations. Maravelakis et al. (2002) proposed a new method based on principal components analysis and developing theoretical control limits for identification of the out-of-control variable. Niaki and Abbasi (2005) proposed an artificial neural network based model to diagnose faults in out-of-control conditions when Shewhart-type multivariate control charts are used. Hafezi and Shahriari (2009) have proposed a method to identify the quality characteristics responsible for out-of-control signal. For this purpose, they improved the sensitivity of Hoteling's T^2 control chart by defining double warning limits. Zhu and Jiang (2009) suggested an adaptive multivariate T^2 control chart which included Hoteling's T^2 chart and the M chart to dynamically integrate pre-test information in multivariate process monitoring.

Diagnosing the out-of-control samples and parameters in the area of profile monitoring have also been considered by some authors. This issue is particularly important in profile monitoring, where the large number of parameters are involved. Zou et al. (2007) discussed the diagnosis of a general linear profile in Phase II to identify the out-of-control parameters. Mahmoud et al. (2007) proposed a segmenting approach of the likelihood ratio statistic

to determine the sources of profile variation in Phase I monitoring of simple linear profile structure. Mahmoud (2008) introduced an approach that can be used to find the out-of-control samples and to determine the source of profile variation based on F-statistic in the analysis of Phase I multiple linear regression profiles. Noorossana et al. (2010a) have developed a diagnostic scheme to find the out-of-control samples based on Wilk's lambda in Phase I of multivariate multiple linear regression profiles structure.

Although several methods have been developed to monitor multivariate multiple linear regression profiles, to the best of our knowledge, there is no research about diagnosing the parameters and profiles responsible for out-of-control signal in multivariate multiple linear regression profiles monitoring. Hence, it is essential to develop an approach to identify out-of-control signal sources, outlying profiles and parameters. In this paper, we specially focus on diagnosing of outlying profiles and out-of-control parameters in multivariate multiple linear regression profile structure in Phase II. This paper is classified as follows. In Sec. 2 we introduce the model and its assumptions. Diagnosis aids for finding the out-of-control profiles and parameters in multivariate multiple linear profiles are developed in Sec. 3. In Sec. 4, the performances of the proposed methods are evaluated by simulation studies in terms of accuracy percent. In order to demonstrate how the proposed methodology works, we provided a real calibration application at Body Shop of an automotive industrial group in Sec. 5. Our concluding remarks are given in the final section.

2. Model and Assumptions

In this section, the multivariate multiple linear regression structure is introduced and corresponding assumptions are expressed.

In multivariate multiple linear regression model, multivariate refers to the dependent variables and multiple pertains to the independent variables. In this case, several Y's $(y_1, y_2, \ldots, y_p)$ are measured corresponding to each set of X's $\left(x_1, x_2, \ldots, x_q\right)$.

Let us assume that for the kth random sample collected over time that there are n observations for independent variables and for each independent one we have set of p values of response variables. So for kth sample we have n observations that they are given as $(x_{1i}, x_{2i}, \ldots, x_{qi}, y_{i1k}, y_{i2k}, \ldots y_{ipk})$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots$, where p is the number of response variables and q is the number of explanatory variables. When the process is in statistical control, the model which relates the response variables with explanatory variables is a multivariate multiple linear regression and can be given as follows:

$$\mathbf{Y}_k = \mathbf{XB} + \mathbf{E}_k, k = 1, 2, \ldots. \tag{1}$$

We can express the model in the following form:

$$\begin{bmatrix} y_{11k} & y_{12k} & \cdots & y_{1pk} \\ y_{21k} & y_{22k} & \cdots & y_{2pk} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1k} & y_{n2k} & \cdots & y_{npk} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & \cdots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{bmatrix} \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{bmatrix}$$
$$+ \begin{bmatrix} \varepsilon_{11k} & \varepsilon_{12k} & \cdots & \varepsilon_{1pk} \\ \varepsilon_{21k} & \varepsilon_{22k} & \cdots & \varepsilon_{2pk} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1k} & \varepsilon_{n2k} & \cdots & \varepsilon_{npk} \end{bmatrix}, \tag{2}$$

where $\mathbf{Y}_k = (y_{1k}, y_{2k,\ldots,} y_{pk})^T$ represents a matrix $(n \times p)$ of response variables for kth sample and $\mathbf{X}$ is a matrix $(n \times (q+1))$ of independent variables and $\mathbf{B}$ is representative of a matrix $((q+1) \times p))$ of regression parameters. Under in-control condition, it is assumed that $\mathbf{E}_k$ is a matrix $(n \times p)$ of error terms and the vector of error terms has a p-variate normal distribution with mean vector $\mathbf{0}$ and $p \times p$ covariance matrix equals to

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}. \tag{3}$$

For the kth random sample, the ordinary least squares estimator of the matrix $\mathbf{B}$ is given by Eq. (4). For more information see Rencher (2002),

$$\hat{\mathbf{B}}_k = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}_k, \tag{4}$$

and finally, during this study we assume that the process covariance matrix, $\mathbf{\Sigma}$, is stable over time.

3. Proposed Diagnostic Methods

In this section, two diagnostic methods to identify outlying profile(s) and a method to determine the out-of-control parameter(s) are proposed. Both proposed methods for finding the out-of-control profiles are based on a T^2 control chart with warning limits proposed by Hafezi and Shahriari (2009) in Phase I monitoring of multivariate process control. The method is extended to Phase II monitoring of multivariate multiple linear regression profiles in our paper. In the first method, the sample mean of observations in each profile is used as a quality characteristic in the extended T^2-based method in Phase II. The second method uses the sample mean of residuals in each profile obtained by fitting the multivariate structure on the dataset. In the method proposed for finding the out-of-control parameters, exponentially weighted moving average (EWMA) statistic are used as inputs of the extended T^2 control chart.

3.1. Diagnosis Methods to Identify Out-of-Control Profiles

Since there is a multivariate structure, the importance of diagnosis of out-of-control profiles is vital. So for this purpose the following methods are proposed:

3.1.1. T^2 *control chart with warning limits based on sample mean of observations.* Hafezi and Shahriari (2009) introduced a T^2 control chart with warning limits for phase I monitoring of multivariate quality characteristics. Note that the warning limits are included to enhance the detection power of the T^2 control chart. This method; moreover, can be used to diagnose the responsible quality characteristic for the out-of-control condition based on a step-by-step algorithm. To be in lined with our structure, we developed this control methodology to be used with the aim of online monitoring of multivariate multiple linear profiles. The key goal of this paper is designing of a diagnosis strategy to detect the actual changed parameter as soon as a signal is triggered by the control chart. As stated earlier, Hafezi and Shahriari (2009) used a T^2 control chart with double warning limits. Three run rules are used in this

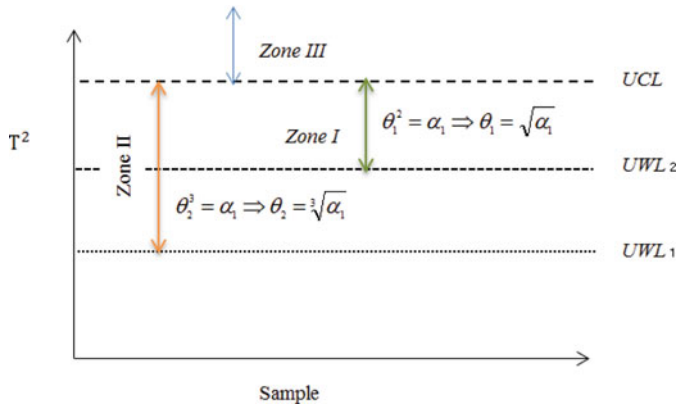

Figure 1. A T^2 control chart with double warning limits.

control chart including one sample out-of-control limit, two sequential samples between the second warning limit and upper control limit, and three sequential samples between the first warning limit and upper control limit. We assume the equal probability of Type I error of α_1 for each of these sensitizing rules (see Fig. 1).

If we assume the probability of locating one sample between the UWL_2 and UCL equal to θ_1, then we have the following equation:

$$\theta_1^2 = \alpha_1 \Rightarrow \theta_1 = \sqrt{\alpha_1}. \tag{5}$$

To determine the value of UWL_2, the relevant probability of Type I error in whole region above UWL_2 (α_2) should be calculated. Hence, $\alpha_2 = \theta_1 + \alpha_1$. Since the T^2 statistic follows a Chi-square distribution in Phase II, UWL_2 is the percentile point of Chi-square distribution with p degrees of freedom and the probability of α_2. Consequently, we have

$$UWL_2 = \chi^2_{p,\alpha_2}. \tag{6}$$

If we assume the probability of locating one sample between the first warning limit (UWL_1) and UCL equal to θ_2, then we have the following equation:

$$\theta_2^3 = \alpha_1 \Rightarrow \theta_2 = \sqrt[3]{\alpha_1}, \tag{7}$$

and to determine the value of UWL_1, the relevant probability of Type I error in whole region above UWL_1 (α_3) should be computed. So, $\alpha_3 = \theta_2 + \alpha_1$ and based on the explanations for UWL_2, the value of UWL_1 is calculated by Eq. (8).

$$UWL_1 = \chi^2_{p,\alpha_3}. \tag{8}$$

In Eqs. (6) and (8), p is the number of quality characteristics used in T^2 Hoteling statistic.

In our first proposed strategy, the T^2 chart has been constructed based on the sample mean of observations of each profile. In this method T^2 statistic is decomposed into three main terms and paired combinations of quality characteristics are used to diagnose the out-of-control profiles. It is remarkable that, the suggested algorithm of Hafezi and Shahriari (2009) reduces the dimension of problem rather than Mason et al. (1995) method. In this method, the paired combinations of quality characteristics are constituted and then by

choosing each paired the following procedure is performed. Let us assume that Cth and Dth profile quality characteristics are selected as a paired combination. The value of T^2 statistic for this pair in out-of-control sample (lth sample) is as follows:

$$T_l^2 = \left[\, \bar{y}_{Cl} - \mu_{\bar{y}_C} \;\; \bar{y}_{Dl} - \mu_{\bar{y}_D} \,\right] \mathbf{\Sigma}^{-1}_{(\bar{y}_C, \bar{y}_D)} \begin{bmatrix} \bar{y}_{Cl} & -\mu_{\bar{y}_C} \\ \bar{y}_{Dl} & -\mu_{\bar{y}_D} \end{bmatrix}, \tag{9}$$

where $y_{Cl} = [y_{1Cl}, y_{2Cl}, \ldots, y_{nCl}]^T$, $y_{Dl} = [y_{1Dl}, y_{2Dl}, \ldots, y_{nDl}]^T$, $\bar{y}_{Cl} = \frac{1}{n}\sum_{i=1}^{n} y_{iCl}$, $\bar{y}_{Dl} = \frac{1}{n}\sum_{i=1}^{n} y_{iDl}$, $\mu_{\bar{y}_C} = \beta_{0C} + \beta_{1C}\bar{x}_1 + \beta_{2C}\bar{x}_2 + \ldots + \beta_{qC}\bar{x}_q$, $\mu_{\bar{y}_D} = \beta_{0D} + \beta_{1D}\bar{x}_1 + \beta_{2D}\bar{x}_2 + \ldots + \beta_{qD}\bar{x}_q$, $\bar{x}_q = \frac{1}{n}\sum_{i=1}^{n} x_{iq}$, and

$$\mathbf{\Sigma}_{(\bar{y}_C, \bar{y}_D)} = \begin{bmatrix} \sigma^2_{\bar{y}_C} & \sigma_{\bar{y}_C, \bar{y}_D} \\ \sigma_{\bar{y}_C, \bar{y}_D} & \sigma^2_{\bar{y}_D} \end{bmatrix}. \tag{10}$$

The elements of the matrix $\mathbf{\Sigma}_{\bar{y}_C, \bar{y}_D}$ are $\sigma^2_{\bar{y}_C} = \sigma^2_{y_C}/n$, $\sigma^2_{\bar{y}_D} = \sigma^2_{y_D}/n$, and $\sigma_{\bar{y}_C, \bar{y}_D}$ is computed as follows:

$$\rho_{C,D} = \frac{\sigma_{\bar{y}_C, \bar{y}_D}}{\sqrt{\frac{\sigma^2_{y_C}}{n}} \times \sqrt{\frac{\sigma^2_{y_D}}{n}}} \Rightarrow \rho_{C,D}\left(\sqrt{\frac{\sigma^2_{y_C}}{n}} \times \sqrt{\frac{\sigma^2_{y_D}}{n}}\right) = \sigma_{\bar{y}_C, \bar{y}_D} \Rightarrow \frac{\left(\sqrt{\sigma^2_{y_C} . \sigma^2_{y_D}}\right)\rho_{C,D}}{n} = \sigma_{\bar{y}_C, \bar{y}_D}, \tag{11}$$

where $\rho_{C,D}$ is correlation coefficient between Cth and Dth sample mean of observations for kth sample, also $\sigma^2_{y_C}$ and $\sigma^2_{y_D}$ are the variances of Cth and Dth sample correspondingly.

Equation (9) is extended again in the following way (see Appendix A for proof):

$$\left(\frac{\bar{y}_{Cl} - \mu_{\bar{y}_C}}{\sigma^2_{\bar{y}_C}}\right)^2 + \left(\frac{\bar{y}_{Dl} - \mu_{\bar{y}_D}}{\sigma^2_{\bar{y}_D}}\right)^2 - 2\left(\frac{\bar{y}_{Cl} - \mu_{\bar{y}_C}}{\sigma^2_{\bar{y}_C}}\right)\left(\frac{\bar{y}_{Dl} - \mu_{\bar{y}_D}}{\sigma^2_{\bar{y}_D}}\right) \times \rho_{(C,D)} = T_l^2\left(1 - \rho^2_{(C,D)l}\right). \tag{12}$$

The first and second terms in the above equation, $(\frac{\bar{y}_{Cl} - \mu_{\bar{y}_C}}{\sigma^2_{\bar{y}_C}})^2$ and $(\frac{\bar{y}_{Dl} - \mu_{\bar{y}_D}}{\sigma^2_{\bar{y}_D}})^2$ equal to value of the T^2 statistic when only Cth and Dth profile quality characteristics are used, respectively, and $2(\frac{\bar{y}_{Cl} - \mu_{\bar{y}_C}}{\sigma^2_{\bar{y}_C}})(\frac{\bar{y}_{Dl} - \mu_{\bar{y}_D}}{\sigma^2_{\bar{y}_D}}) \times \rho_{(C,D)}$ represents the reciprocal effects of Cth and Dth profile quality characteristics for sample l when both quality characteristics are considered and it is denoted by $T^2_{C,D}$. According to the above equation the T^2 statistic can be decomposed into three terms as follows:

$$T^2_{Cl} + T^2_{Dl} + T^2_{(C,D)l} = T_l^2(1 - \rho^2_{(C,D)}). \tag{13}$$

By decomposing the T_l^2 statistic into three terms, the effects of Cth and Dth profile quality characteristics and their effects on each other easily can be distinguished leading to diagnosing the out-of-control profiles. As explained in Sec. 1, Hafezi and Shahriari (2009) proposed a diagnosis procedure in Phase I monitoring of multivariate process control.

In the proposed method by Hafezi and Shahriari (2009), following steps are implemented for identification:

Step 1: For each sample calculate T^2 statistic by equation $T_k^2 = (\bar{\mathbf{y}}_k - \boldsymbol{\mu}_{\bar{\mathbf{y}}_k})\boldsymbol{\Sigma}_{\bar{\mathbf{y}}_k}^{-1}(\bar{\mathbf{y}}_k - \boldsymbol{\mu}_{\bar{\mathbf{y}}_k})^T$. If there are any out-of-control samples or one of the out-of-control conditions for T^2 control chart with warning limits holds, we will go to next step.

Step 2: The value of T_k^2 statistic is calculated for each paired combination of quality characteristics by using Eq. (9). Then each part in the left side of Eq. (12) is calculated and by dividing each three terms by the right part of the equation, the percentage of each term in T^2 for out-of-control sample is estimated.

Step 3: The values of q_1 and q_2 are calculated by Eq. (14). Simultaneously, the value of $\rho_{\min}$ is set by analyzer. The $\rho_{\min}$ is the minimum value of correlation coefficient between two quality characteristics leading to significant interaction between them.

$$q_1 = UWL_1 = \chi^2_{2,\alpha_3} \text{ and } q_2 = \frac{2\rho_{\min}}{q_1} \times 100, \tag{14}$$

where q_1 and q_2 are considered to signify the values of $T^2_{C,D}$ and T^2 for analyzing, respectively.

Step 4: After calculating the percentage of $T^2_{(C,D)l}$, T^2_{Cl}, T^2_{Dl} in T_l^2 statistic in Step 2 and according to the values ofq_2, q_1, $\rho_{\min}$, $\rho_{(C,D)}$, the following conditions are analyzed:

- If $T^2_{(C,D)l} \geq 0$ and
  - If $|\rho_{(C,D)}| < \rho_{\min}$, $T^2_{(C,D)l} > q_2$, $T_l^2 > q_1$, it is concluded that both quality characteristics have changed and the effect of correlation is not significant.
  - If $|\rho_{(C,D)}| > \rho_{\min}$, $T^2_{(C,D)l} > q_2$, $T_l^2 > q_1$, both quality characteristics have been changed by the effect of correlation between them. In this case, we can identify which quality characteristic leads to the out-of-control state by comparing T^2_{Cl} and T^2_{Dl}. The one with the larger magnitude represents the actual changed parameter.
  - If $0 \leq T^2_{(C,D)l} \leq q_2$, $T_l^2 > q_1$, one of the quality characteristic with greater T_l^2 has changed.
- If $T^2_{(C,D)l} < 0$ and
  - If $|\rho_{(C,D)}| < \rho_{\min}$, $T^2_{(C,D)l} < -q_2$, $T_l^2 > q_1/2$, both quality characteristics have changed. In this case the effect of correlation is not significant.
  - If $|\rho_{(C,D)}| > \rho_{\min}$, $T^2_{(C,D)l} < -q_2$, $T_l^2 > q_1/2$, both quality characteristics have been changed by the effect of correlation between them.
  - If $-q \leq T^2_{(C,D)l} < 0$, $T_l^2 > q_1$, one of the quality characteristic with greater T_l^2 has changed.
- If none of the above-mentioned conditions are satisfied, the paired quality characteristics are not responsible for out-of-control state.

3.1.2. T^2 *control chart based on the EWMA statistic.* In this method, EWMA statistic are computed based on average residuals. Then, the T^2 statistic will be formed applying these EWMA statistic. Since there are p dependent profiles, the inputs of the T^2 statistic are p EWMA statistic that each of them is computed based on the average of each profile's residuals.

Since there are n different levels for each independent variable, the residuals of each sample are computed by

$$\mathbf{Y}_k - \hat{\mathbf{Y}}_k = \mathbf{E}_k, \quad k = 1, 2, \ldots, \tag{15}$$

where $\hat{\mathbf{Y}}_k = \mathbf{X}\hat{\mathbf{B}}$ is the estimated matrix of observations.

Then, the average of the residuals for each sample is calculated by Eq. (16) as follows:

$$\bar{\boldsymbol{e}}_k = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{e}_{ik}, \tag{16}$$

where e_{ik} is the vector of residuals regarding to all p profiles in kth sample and $\bar{e}_k$ is the average vector of error terms for kth sample.

Finally, the T^2 statistic is computed as follows:

$$T^2_{Z_k} = z_k^T \boldsymbol{\Sigma}^{-1}_{Z_k} z_k, \tag{17}$$

where z_k is given by

$$z_k = \omega \bar{\boldsymbol{e}}_k + (1-\omega)\, z_{k-1}, \tag{18}$$

and the variance covariance matrix (Σ_z) is defined by

$$\Sigma_z = \frac{\omega}{2-\omega}\Sigma_{\bar{e}}, \tag{19}$$

where $\boldsymbol{\Sigma}_{\bar{e}}$ is given by the following equation:

$$\boldsymbol{\Sigma}_{\bar{e}} = \begin{bmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} & \cdots & \boldsymbol{S}_{1p} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} & \cdots & \boldsymbol{S}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{S}_{p1} & \boldsymbol{S}_{p2} & \cdots & \boldsymbol{S}_{pp} \end{bmatrix}. \tag{20}$$

S_{ii} is variance of the residuals of ith profile which equals to $\sigma_{ii}/_{n}$. S_{ij} is covariance between residuals of ith and jth profiles.

The parameter ω ($0 < \omega \leq 1$) is the smoothing parameter, z_k is a multivariate normal random vector, and z_0 is $p \times 1$ vector of zeros. The upper control limit and also UWL_1 and UWL_2 for the statistic in Eq. (17) are chosen to achieve a specified in-control ARL.

Now the mentioned steps in Sec. 3.1.1 are repeated except that Eq. (12) is replaced by Eq. (21) (see Appendix B for proof). This equation shows the value of T^2 statistic for a pair of profiles in an out-of-control sample (lth sample),

$$\left(\frac{Z_{Cl}}{S_{Cl}}\right)^2 + \left(\frac{Z_{Dl}}{S_{Dl}}\right)^2 - 2\left(\frac{Z_{Cl}}{S_{Cl}}\right)\left(\frac{Z_{Dl}}{S_{Dl}}\right) \times \rho_{(y_C, y_D)} = T_l^2 \frac{\left(1-\rho^2_{(y_C, y_D)}\right)(2-\omega)}{\omega}, \tag{21}$$

where the first and second terms of left side in the above equation, $(\frac{Z_{Cl}}{S_{Cl}})^2$ and $(\frac{Z_{Dl}}{S_{Dl}})^2$ equal to value of T^2 when only Cth and Dth profile quality characteristics are available, respectively, and $2(\frac{Z_{Cl}}{S_{Cl}})(\frac{Z_{Dl}}{S_{Dl}}) \times \rho_{(y_C y_D)}$ represents the reciprocal effects of Cth and Dth profile quality characteristics when both characteristics are considered denoted by $T^2_{(C,D)l}$. According to the above equation, the T^2 statistic can be decomposed into three terms as follows:

$$T^2_{Cl} + T^2_{Dl} + T^2_{(C,D)l} = T_l^2 \frac{\left(1-\rho^2_{(y_C, y_D)}\right)(2-\omega)}{\omega}. \tag{22}$$

3.2. Diagnosis Methods to Identify Out-of-Control Parameters

Since in practice the multivariate multiple linear regression profiles may involve a huge number of parameters, after finding the outlying profiles, it is critical to diagnose which parameters in profile have shifted.

3.2.1. T^2 *control chart based on the EWMA statistic.* This method is similar to the diagnosis method proposed in Sec. 3.1.1 and it is an extension of method proposed by Eyvazian et al. (2011).

At first the matrix $\hat{\mathbf{B}}_k$ is rewritten as a $((q+1)\,p) \times 1$ multivariate normal random vector denoted by $\hat{\boldsymbol{\beta}}_k$:

$$(\hat{\beta}_{01k}, \hat{\beta}_{11k}, \ldots, \hat{\beta}_{q1k}, \hat{\beta}_{12k}, \ldots, \hat{\beta}_{q2k}, \ldots, \hat{\beta}_{0pk}, \hat{\beta}_{1pk}, \ldots, \hat{\beta}_{qpk})^T = \hat{\boldsymbol{\beta}}_k. \tag{23}$$

When the process is in-control, the expected value and covariance matrix for $\hat{\boldsymbol{\beta}}_k$ is given by

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_K} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \cdots & \boldsymbol{\Sigma}_{1p} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \cdots & \boldsymbol{\Sigma}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{p1} & \boldsymbol{\Sigma}_{p2} & \cdots & \boldsymbol{\Sigma}_{pp} \end{bmatrix}, \tag{24}$$

$$E(\hat{\boldsymbol{\beta}}_k) = (\beta_{01}, \beta_{11}, \ldots, \beta_{q1}, \beta_{02}, \beta_{12}, \ldots, \beta_{q2}, \ldots, \beta_{0p}, \beta_{1p}, \ldots, \beta_{qp})^T, \tag{25}$$

where $(\beta_{01}, \beta_{11}, \ldots, \beta_{q1}, \beta_{02}, \beta_{12}, \ldots, \beta_{q2}, \ldots, \beta_{0p}, \beta_{1p}, \ldots, \beta_{qp})^T$ is denoted by $\boldsymbol{\beta}$. Eyvazian et al. (2011) have shown that $\boldsymbol{\Sigma}_{hj}$ is a $(q+1) \times (q+1)$ matrix equals to $[X^T X]^{-1}\sigma_{hj}$ where σ_{hj} denotes the hjth element of the covariance matrix $\boldsymbol{\Sigma}$ in Eq. (3).

To identify out-of-control parameters, the following proposed statistic is used. For ith and jth parameters, z_{li} and z_{lj} are defined as follows:

$$\begin{aligned} z_{li} &= \omega(\hat{\beta}_i - \beta_i) + (1-\omega)z_{li-1} \\ z_{lj} &= \omega(\hat{\beta}_j - \beta_j) + (1-\omega)z_{lj-1}. \end{aligned} \tag{26}$$

So the T^2 statistic is computed by Eq. (27) as follows:

$$T_l^2 = [z_{li} - \mu_{z_{li}} \quad z_{lj} - \mu_{z_{lj}}]\boldsymbol{\Sigma}^{-1}_{z_{l(i,j)}} \begin{bmatrix} z_{li} - \mu_{z_{li}} \\ z_{lj} - \mu_{z_{lj}} \end{bmatrix}, \mu_{z_{li}} = \mu_{z_{lj}} = 0, \tag{27}$$

where

$$\boldsymbol{\Sigma}_{z_{l(i,j)}} = \frac{\lambda}{2-\lambda}\boldsymbol{\Sigma}_{\hat{\beta}_{l(i,j)}}. \tag{28}$$

Hence,

$$\left(\frac{Z_{li}}{S_{li}}\right)^2 + \left(\frac{Z_{lj}}{S_{lj}}\right)^2 - 2\left(\frac{Z_{li}}{S_{li}}\right)\left(\frac{Z_{lj}}{S_{lj}}\right) \times \rho_{(i,j)} = T_l^2 \frac{\left(1-\rho^2_{(i,j)}\right)(2-\omega)}{\omega}, \tag{29}$$

(the proof is given in Appendix B)
where

$$\rho_{(i,j)} = \sigma_{ij} \big/ \sqrt{\sigma_{ii}\sigma_{jj}}, \tag{30}$$

and σ_{ij} denotes the elements of $\mathbf{\Sigma}_{hj}$ matrix.

This method is repeated $\binom{p(q+1)}{2}$ times to diagnose the out-of-control parameters.

4. Performance Analysis

In this section, the performances of the proposed methods are evaluated by simulation studies in terms of accuracy percent criteria. It means that we are computing the number of iterations in which the algorithm finds out-of-control characteristics correctly in simulation runs when an out-of-control situation occurs. For this purpose, we set the probability of each run rules equal to 0.0017. Without loss of generality, in all the EWMA statistic, the value of smoothing constant, ω, is set to 0.2 as generally used in the literature. Our simulation studies have been performed using 20,000 replications. The underlying multivariate profile model considered in this paper is

$$\begin{aligned} y_1 &= 3 + 2x_1 + x_2 + \varepsilon_1, \\ y_2 &= 2 + x_1 + x_2 + \varepsilon_2. \end{aligned}$$

The pair of observations (2,1), (4,2), (6,3), (8,2) are considered as the values for independent variables x_1 and x_2. The vector of error terms $(\varepsilon_1, \varepsilon_2)$ follows a bivariate normal random variable with mean vector zero and known covariance matrix,

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

where $\sigma_1^2 = 1$ and $\sigma_2^2 = 1$. To investigate the effect of correlation between profiles, different values of ρ, namely, $\rho = 0.1$, 0.5, and 0.9 are used in our simulation studies for individual shifts and also the minimum value of correlation coefficient between two quality characteristics, $\rho_{\min}$ is set equal to 0.1, 0.55, and 0.9. For simultaneous shifts, the correlation between response variables is set equal to 0.5.

4.1. Investigation of Simulation Results for Out-of-Control Profile Identification

To investigate the results of diagnosis methods for out-of-control profiles the following parameters are set:

$$\begin{aligned} \theta_2^3 &= 0.0017 \rightarrow \theta_2 = 0.1193, \\ \alpha_{UWL_1} &= \alpha_3 = 0.0017 + 0.1193 = 0.121. \end{aligned}$$

So, the value of UWL_1 is calculated based on Eq. (13) as follows:

$$UWL_1 = \chi^2_{p,\alpha_3} = \chi^2_{2,0.121} = 4.2231.$$

UWL_2 based on Eq. (11) and UCL are calculated by

$$\begin{aligned} \theta_1^2 &= 0.0017 \rightarrow \theta_1 = 0.04123, \\ \alpha_{UWL_2} &= \alpha_2 = 0.0017 + 0.04123 = 0.04293, \\ UWL_2 &= \chi^2_{p,\alpha_2} = \chi^2_{2,0.04293} = 6.2963, \\ UCL &= \chi^2_{p,\alpha_1} = \chi^2_{2,0.0017} = 12.7543. \end{aligned}$$

4.1.1. T^2 *control chart with warning limits.* Let us set

$$\bar{y}_1 = \beta_{01} + \beta_{11}\bar{x}_1 + \beta_{21}\bar{x}_2 + \bar{\varepsilon}_1,$$
$$\bar{y}_2 = \beta_{02} + \beta_{12}\bar{x}_1 + \beta_{22}\bar{x}_2 + \bar{\varepsilon}_2,$$

$$\mu_{\bar{y}_1} = E\left[\bar{y}_1\right] = \beta_{01} + \beta_{11}\bar{x}_1 + \beta_{21}\bar{x}_2,$$
$$\mu_{\bar{y}_2} = E\left[\bar{y}_2\right] = \beta_{02} + \beta_{12}\bar{x}_1 + \beta_{22}\bar{x}_2.$$

The results of the simulation studies to diagnose the outlying profiles based on the example described former are given in Table 1. The simulations are done under different correlation coefficients. The values in Table 1 show the accuracy percent of the first proposed method to diagnose the outlying profiles under different situations.

Table 1 shows that in individual shifts when the values of ρ and $\rho_{\min}$ are close together the performance of the proposed method is worse than the other conditions in which this difference (ρ and $\rho_{\min}$) becomes lager. So the effect of $\rho_{\min}$ depends on the value of ρ. Moreover, the results show that when there is much difference between the values of ρ and $\rho_{\min}$, the performance in smaller value of ρ is better. Finally, as the magnitude of shift increases, in all situations, the accuracy percent improves.

As shown in Table 2, increasing the magnitude of the shift improves the performance of the method. Moreover, by comparison of results in Tables 1 and 2, it can be concluded that the performance of the proposed method under simultaneous shifts is better than the individual shifts when $\rho_{\min} = 0.55$.

4.1.2. T^2 *control chart based on the EWMA statistic.* Table 3 shows that in individual shifts, if the difference between values of ρ and $\rho_{\min}$ becomes larger, the performance of this method improves. In addition, the performance of the method under smaller ρ is better.

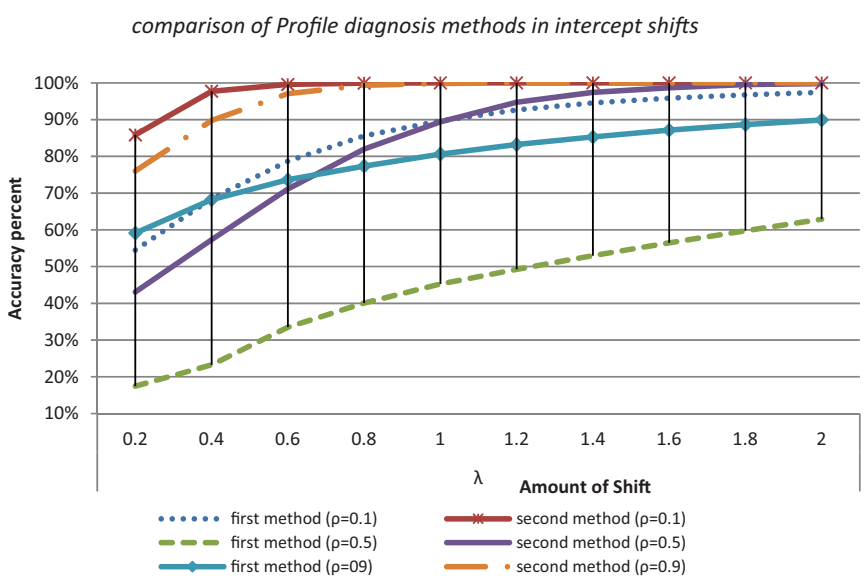

Figure 2. Comparison of proposed methods for identification of out-of-control profiles in individual intercept shifts ($\rho_{\min} = 0.55$).

Table 1

The results (accuracy percent) of individual shifts in the intercept (a) and the slope (b) of the first profile

	$\rho_{\min} = 0.9$			$\rho_{\min} = 0.55$			$\rho_{\min} = 0.1$		
λ	$\rho = 0.1$	$\rho = 0.5$	$\rho = 0.9$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 0.9$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 0.9$
(a) Intercept shifts from β_0 to $\beta_0 + \lambda\sigma$ in the first profile									
0.2	75.93%	48.89%	49.09%	54.49%	17.44%	59.11%	37.27%	53.19%	56.10%
0.4	76.19%	74.44%	50.31%	68.53%	23.25%	68.21%	51.33%	59.29%	73.12%
0.6	82.19%	81.26%	52.91%	78.67%	33.50%	73.64%	51.38%	73.25%	82.52%
0.8	88.94%	89.58%	56.27%	85.59%	40.08%	77.35%	57.30%	74.62%	90.42%
1	93.55%	91.55%	60.80%	89.68%	45.28%	80.63%	58.04%	77.07%	94.24%
1.2	95.48%	93.74%	61.94%	92.64%	49.24%	83.22%	64.36%	77.91%	95.88%
1.4	97.53%	95.57%	64.72%	94.54%	53.00%	85.27%	64.76%	79.57%	98.26%
1.6	98.73%	97.29%	66.43%	95.83%	56.48%	87.15%	65.71%	80.31%	98.84%
1.8	99.57%	98.66%	68.19%	96.74%	59.76%	88.65%	67.72%	82.16%	99.51%
2	99.69%	99.06%	68.35%	97.43%	62.89%	89.93%	68.78%	82.69%	99.77%
(b) Slope shifts from β_1 to $\beta_1 + \beta\sigma$ in the first profile									
0.025	43.75%	58.33%	40.39%	55.46%	15.25%	63.99%	25.49%	39.13%	56.67%
0.05	51.22%	63.46%	52.54%	63.58%	21.10%	64.57%	43.23%	40.74%	75.68%
0.075	70.91%	75.00%	60.31%	68.00%	25.00%	67.56%	49.51%	56.04%	76.32%
0.1	72.92%	77.46%	60.35%	74.71%	30.76%	70.69%	57.06%	67.81%	83.75%
0.125	85.90%	82.93%	62.68%	79.05%	34.73%	73.16%	59.16%	73.95%	85.44%
0.15	86.22%	85.90%	65.63%	82.89%	38.52%	75.46%	63.46%	75.60%	88.66%
0.175	90.44%	88.52%	69.02%	86.12%	41.40%	77.62%	65.75%	77.87%	91.66%
0.2	93.13%	92.10%	71.57%	88.87%	44.12%	79.66%	65.67%	78.45%	93.66%
0.225	94.33%	93.05%	74.29%	90.92%	47.14%	81.46%	68.20%	80.90%	96.05%
0.25	96.58%	94.88%	76.23%	92.51%	49.36%	82.99%	69.62%	82.04%	96.86%

Table 2
The results (accuracy percent) of simultaneous shifts in intercepts of both profiles

Intercept shifts from β_{01} to $\beta_{01}+\lambda\sigma$ in the first profile and β_{02} to $\beta_{02}+\lambda\sigma$ in the second profile										
Λ	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2
$\rho_{\min}=0.9$	54.71%	72.25%	79.08%	85.315	89.71%	92.54%	94.47%	95.76%	96.69%	97.36%
$\rho_{\min}=0.55$	73.45%	80.99%	86.68%	91.52%	94.22%	96.07%	97.24%	97.92%	98.42%	98.75%

According to Table 3, the performance of proposed method in case of $\rho = 0.1$ and $\rho_{\min} = 0.9$ is better than the other cases.

It is obvious that this method is powerful in simultaneous shifts when $\rho_{\min} = 0.55$ and from medium shifts the accuracy percent equals to 100%. Tables 3 and 4 show that the sensitivity of T^2 control chart, especially in small shifts when $\rho_{\min} = 0.55$, has been improved rather than the first method for diagnosing outlying profiles.

As shown in Figs. 2 and 3, the performance of the second method based on residual approach in diagnosis of out-of-control profile is more powerful than the first one in individual shifts of the intercept and the slope in one profile. Moreover, the performance of second one in diagnosis of out-of-control profile in intercept shifts is better than its performance in slope shifts. Also when the difference between ρ and $\rho_{\min}$ increases, the performance of both methods in medium and large shifts becomes better and in this situation the smaller ρ leads to better performance. Finally, in simultaneous shifts, Fig. 4 obviously shows that the second diagnosis method performs better than the first one.

4.2. Investigation of Simulation Results for Out-of-Control Parameters Identification

In this simulation study there are two profiles each containing two independent variables, so there are 6 parameters. The value of *UCL* equals to

$$UCL = 18.547.$$

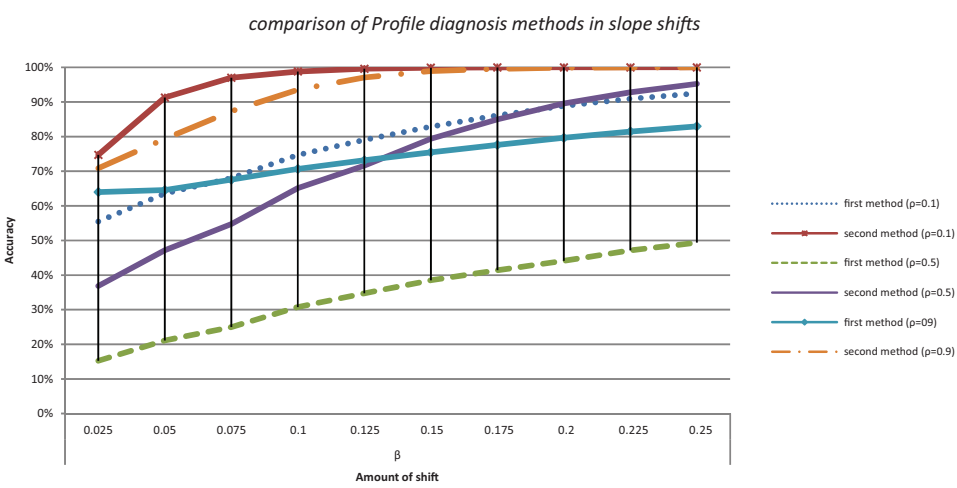

Figure 3. Comparison of proposed methods for identification of out-of-control profiles in individual slope shifts ($\rho_{\min} = 0.55$).

Table 3

The results (accuracy percent) of individual shifts in the intercept (a) and the slope (b) of the first profile

	$\rho_{\min} = 0.9$			$\rho_{\min} = 0.55$			$\rho_{\min} = 0.1$		
λ	$\rho = 0.1$	$\rho = 0.5$	$\rho = 0.9$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 0.9$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 0.9$
(a) Intercept shifts from β_0 to $\beta_0 + \lambda\sigma$ in the first profile									
0.2	86.98%	60.23%	30.38%	85.84%	43.08%	75.98%	39.24%	84.49%	78.52%
0.4	97.87%	80.26%	49.55%	97.73%	57.31%	89.80%	52.41%	95.46%	89.11%
0.6	99.64%	89.56%	65.80%	99.57%	71.11%	97.06%	63.05%	98.40%	96.40%
0.8	99.94%	96.55%	79.31%	99.91%	81.98%	99.34%	76.54%	99.69%	99.24%
1	100%	98.84%	87.68%	100%	89.41%	99.84%	86.63%	99.99%	99.94%
1.2	100%	100%	91.52%	100%	94.71%	99.98%	93.17%	100%	99.99%
1.4	100%	100%	95.14%	100%	97.44%	100%	95.80%	100%	100%
1.6	100%	100%	97.26%	100%	98.70%	100%	97.87%	100%	100%
1.8	100%	100%	98.64%	100%	99.53%	100%	99.40%	100%	100%
2	100%	100%	99.12%	100%	99.78%	100%	99.76%	100%	100%
(b) Slope shifts from β_1 to $\beta_1 + \beta\sigma$ in the first profile									
0.025	75.65%	54.02%	23.35%	74.73%	36.89%	70.89%	35.14%	80.68%	70.11%
0.05	91.99%	71.23%	33.19%	91.28%	47.17%	79.38%	36.20%	90.44%	81.83%
0.075	97.33%	80.65%	46.57%	97.02%	54.74%	87.31%	49.65%	95.25%	87.65%
0.1	99.03%	87.03%	57.48%	98.83%	65.14%	93.57%	59.54%	97.48%	93.79%
0.125	99.67%	90.85%	67.54%	99.61%	71.61%	97.10%	66.85%	98.65%	98.11%
0.15	99.95%	94.97%	75.74%	99.89%	79.38%	98.95%	73.62%	99.50%	98.70%
0.175	99.99%	97.26%	82.28%	99.98%	84.97%	99.55%	80.60%	99.82%	99.68%
0.2	100%	98.86%	86.08%	100%	89.60%	99.92%	86.40%	99.99%	99.89%
0.225	100%	99.32%	90.09%	100%	92.82%	99.99%	89.35%	100%	99.96%
0.25	100%	100%	93%	100%	95%	100%	93.55%	100%	100%

Table 4
The results (accuracy percent) of simultaneous shifts in intercepts of both profiles

	Intercept shifts from β_{01} to $\beta_{01}+\lambda\sigma$ in the first profile and β_{02} to $\beta_{02}+\lambda\sigma$ in the second profile									
Λ	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2
$\rho_{\min}=0.9$	87.65%	97.47%	99.59%	99.98%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
$\rho_{\min}=0.55$	92.81%	99.18%	99.99%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

It is important to be mentioned that the value of q_1 in all methods are equal to the value of UWL_1 when two quality characteristics are considered.

4.2.1. T^2 *control chart based on the EWMA statistic.* Following tables are the results of simulation studies for T^2 control chart based on EWMA statistic under different values of ρ and $\rho_{\min}$.
According to Tables 5 and 6, the performance of proposed method when $\rho_{\min}=0.1$ is better than the other values of $\rho_{\min}$ especially in medium and large shifts and also whenever the value of $\rho_{\min}$ and ρ are far from each other the results will be better.

The results of simultaneous shifts are also shown in Table 7. It shows that the performance of proposed method when $\rho_{\min}=0.1$ is better than the other value of $\rho_{\min}$.

According to Tables 5, 6, and 7 and Fig. 5, the performance of the method in individual shifts is better than simultaneous ones for diagnosing the out-of-control parameters.

As shown in Figs. 6 and 7 and Tables 5 and 6, the performance of proposed method improves in medium and large shifts by increasing the value of ρ.

If the practitioners are interested in indentifying the true out-of-control parameter (profile) in the single-shift case and there are only two quality characteristics, $\rho_{\min}$ should be selected far from the correlation coefficient ρ. The problem arises when there are more than two quality characteristics. In this case, we have more than one ρ thus the selection

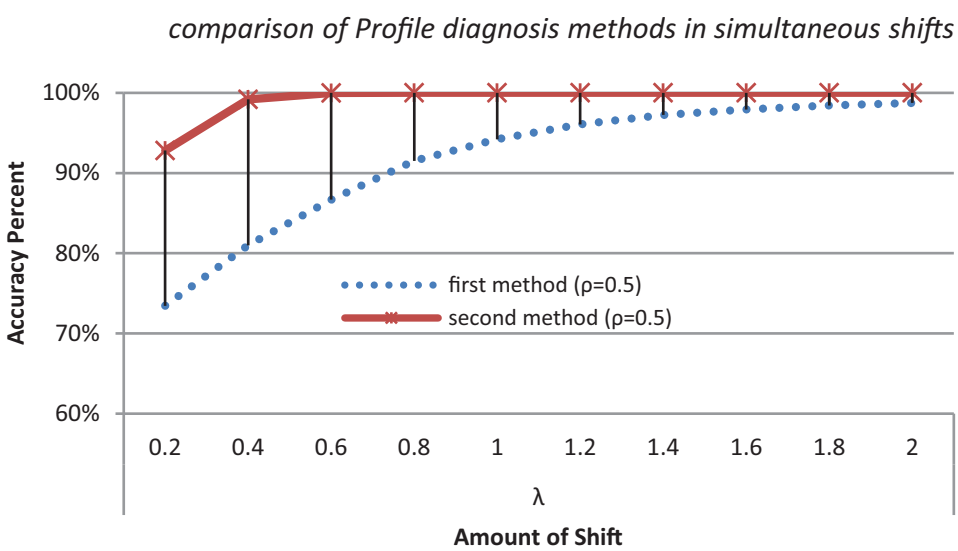

Figure 4. Comparison of proposed methods for identification of out-of-control profiles in simultaneous shifts ($\rho_{\min}=0.55$).

Table 5
The results (accuracy percent) of individual shifts of intercept in the first profile

Intercept shifts from β_{01} to $\beta_{01} + \lambda\sigma$ in the first profile									
	$\rho_{\min} = 0.9$			$\rho_{\min} = 0.55$			$\rho_{\min} = 0.1$		
λ	$\rho = 0.9$	$\rho = 0.5$	$\rho = 0.1$	$\rho = 0.9$	$\rho = 0.5$	$\rho = 0.1$	$\rho = 0.9$	$\rho = 0.5$	$\rho = 0.1$
0.2	30.82%	43.75%	47.31%	37.15%	42.00%	44.33%	43.31%	57.14%	58.51%
0.4	32.65%	43.98%	42.98%	43.69%	42.08%	46.80%	51.75%	55.49%	56.56%
0.6	49.10%	45.91%	64.80%	49.62%	48.70%	50.79%	67.96%	62.38%	61.71%
0.8	67.28%	56.95%	76.84%	68.90%	59.15%	61.84%	81.32%	69.94%	69.10%
1	79.06%	72.77%	84.80%	80.82%	73.26%	72.00%	86.40%	83.12%	77.01%
1.2	84.34%	83.84%	90.60%	82.57%	84.20%	85.02%	92.36%	91.70%	84.90%
1.4	91.24%	92.46%	94.19%	92.78%	92.80%	91.55%	95.54%	95.02%	90.96%
1.6	96.10%	96.42%	96.72%	96.06%	95.80%	96.00%	97.00%	98.14%	95.40%
1.8	97.84%	98.70%	99.27%	97.62%	98.41%	97.86%	98.80%	99.00%	97.92%
2	99.14%	99.38%	99.42%	99.17%	99.26%	99.14%	98.78%	99.66%	99.62%

Table 6
The results (accuracy percent) of individual shifts of intercept in the second profile

Intercept shifts from β_{02} to $\beta_{02} + \lambda\sigma$ in the second profile									
	$\rho_{\min} = 0.9$			$\rho_{\min} = 0.55$			$\rho_{\min} = 0.1$		
λ	$\rho = 0.9$	$\rho = 0.5$	$\rho = 0.1$	$\rho = 0.9$	$\rho = 0.5$	$\rho = 0.1$	$\rho = 0.9$	$\rho = 0.5$	$\rho = 0.1$
0.2	39.81%	50.33%	49.21%	47.55%	38.76%	47.04%	55.37%	42.95%	49.92%
0.4	54.57%	52.02%	50.86%	48.08%	41.68%	43.18%	59.15%	58.87%	56.80%
0.6	57.67%	53.94%	52.74%	65.55%	48.05%	50.63%	68.74%	62.21%	59.93%
0.8	61.13%	62.52%	65.92%	76.86%	58.07%	61.55%	80.38%	71.11%	69.47%
1	68.96%	70.52%	76.86%	85.84%	74.19%	73.39%	87.80%	82.91%	78.36%
1.2	76.82%	78.56%	82.99%	91.50%	85.27%	84.75%	92.70%	89.52%	87.29%
1.4	80.73%	82.57%	86.97%	94.15%	92.76%	92.32%	95.66%	93.82%	91.74%
1.6	84.90%	86.83%	90.42%	96.91%	94.36%	95.93%	96.90%	96.54%	96.48%
1.8	86.05%	88.00%	91.88%	98.04%	98.42%	97.57%	99.70%	99.12%	97.78%
2	87.21%	89.19%	92.55%	98.84%	98.44%	99.05%	99.92%	99.84%	98.10%

Table 7
The results of simultaneous shifts of intercept in the both profiles under value of $\rho = 0.5$

Intercept shifts from β_{01} to $\beta_{01} + \lambda\sigma$ in the first profile and β_{02} to $\beta_{02} + \lambda\sigma$ in the second profile										
λ	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2
$\rho_{\min} = 0.9$	31.55%	30.52%	34.75%	45.8%	60.31%	76.12%	87.45%	94.88%	98.05%	99.39%
$\rho_{\min} = 0.55$	42.64%	40.81%	39.04%	50.09%	64.98%	79.44%	89.76%	95.40%	98.34%	99.50%

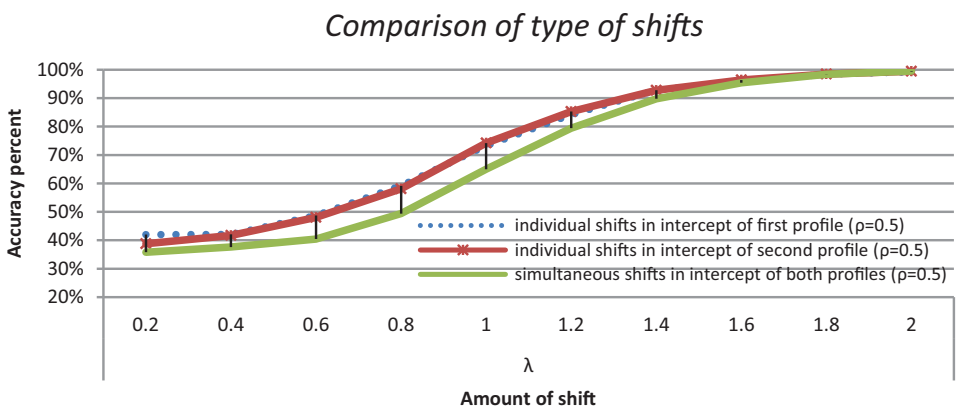

Figure 5. Performance comparison of proposed method for identification of out-of-control parameters in individual and simultaneous shifts.

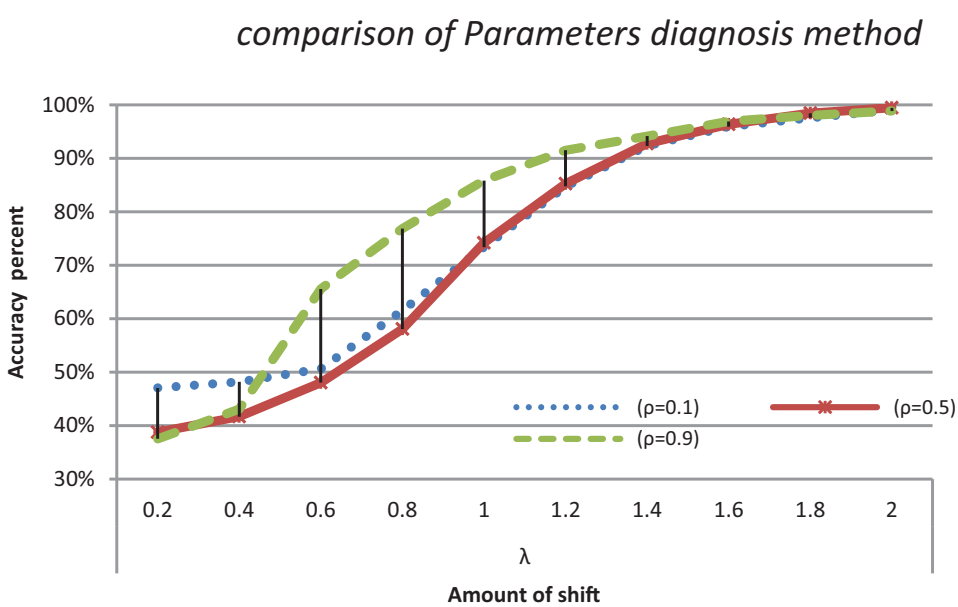

Figure 6. Comparison of proposed method for identification of out-of-control parameters in individual shifts of intercept in the first profile under different values of ρ.

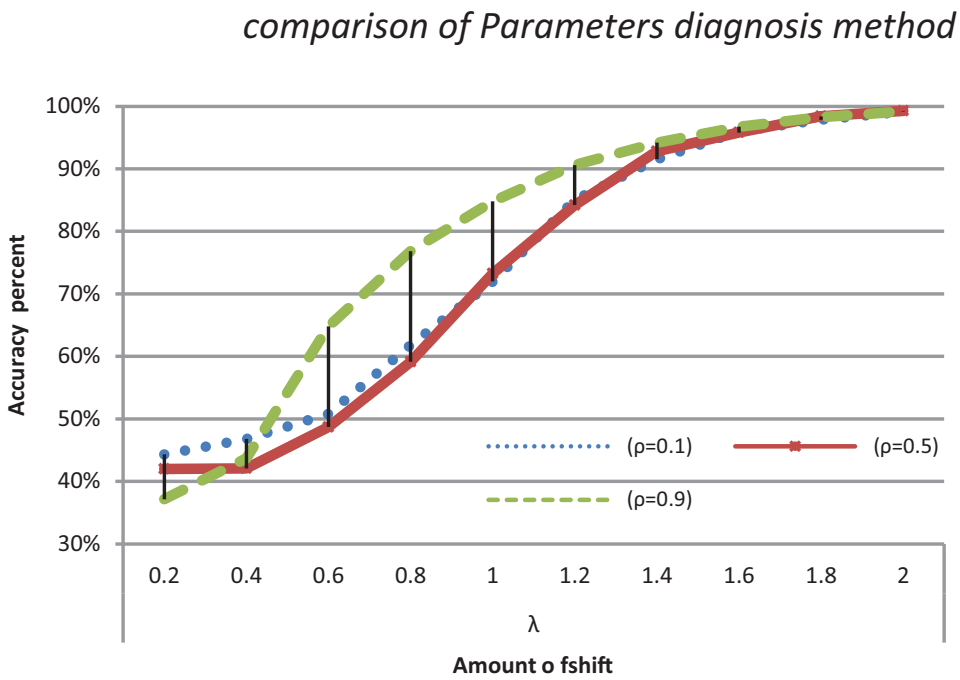

Figure 7. Comparison of proposed method for identification of out-of-control parameters in individual shifts of intercept in the second profile under different values of ρ.

of $\rho_{\min}$ is not straightforward. If the magnitudes of all correlation coefficients are close to each other, it is possible to set $\rho_{\min}$ far from these values. Otherwise, we suggest applying different amounts of $\rho_{\min}$ and see how the method works.

When we have simultaneous shifts in the parameters, we suggest choosing $\rho_{\min}$ close enough to ρ. In this case, the performance of the proposed method will definitely improve.

5. A Real Case Study

In this section, we illustrate how the proposed methods can be applied to the calibration case at Body Shop of an automotive industrial group discussed by Noorossana et al. (2010b). They studied Phase II monitoring of multivariate simple linear profiles to investigate calibration between desired force and the real force produced by a 1600-ton hydraulic press machine. An important input value in this press machine is the nominal force that should be exerted by cylinders on the metal plates to give the desired parts. Therefore, for each value of nominal force (explanatory variable) there are four real forces (response variables) collected from four cylinders of the press and measured by PLC.

The underlying model is

$$
\begin{aligned}
y_1 &= -8.5 + 0.87x + \varepsilon_1 \\
y_2 &= -5.8 + 0.95x + \varepsilon_2 \\
y_3 &= 3.2 + 1.04x + \varepsilon_3 \\
y_4 &= 13.6 + 1.09x + \varepsilon_4,
\end{aligned}
$$

Table 8
The percent of true diagnosis under the individual shifts of intercept in the first profile

	Intercept shifts from β_{01} to $\beta_{01}+\lambda\sigma$ in the first profile									
λ	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55
$\rho_{\min}=0.1$	74.23%	80.74%	86.08%	92.15%	96.42%	98.51%	99.48%	99.88%	99.96%	100%

where $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ is a multivariate random vector with mean vector zero and covariance matrix of

$$\mathbf{\Sigma} = \begin{bmatrix} 80 & 89.6 & 45.1 & 25.3 \\ 89.6 & 122.1 & 71.5 & 29.1 \\ 45.1 & 71.5 & 189 & -28.8 \\ 25.3 & 29.1 & -28.8 & 84.4 \end{bmatrix}.$$

Each sample consists of 11 values for nominal force as 50, 80, 110, 140, 170, 200, 230, 260, 290, 320, and 350 and there are four real force values (y_i) for each nominal force value.

To apply the proposed methods for identifying out-of-control profile(s),

UCL is calculated by

$$UCL = 14.8603.$$

The out-of-control state has been simulated by shifting the intercept parameter of the first profile. Table 8 shows the percent of true diagnosis by the proposed method with regard to the possible range of intercept shifts. It is notable that $\rho_{\min}$ is set equal to 0.1 for this example.

6. Conclusions

In this paper, we proposed three methods to diagnose out-of-control profiles and parameters in multivariate multiple linear regression profiles in Phase II. Two approaches were introduced to identify the out-of-control profiles. The first one was a method based on T^2 control chart with warning limits and the second one was combined residual approach and T^2 control chart with warning limits method. Our simulation studies showed that the second one is better than the first one under both individual and simultaneous shifts in the regression parameters. To identify out-of-control parameters the T^2 control chart with warning limits based on EWMA statistic method was introduced. Simulation studies showed the perfect performance of the proposed method in identifying the out-of-control parameters under both individual and simultaneous shifts. In addition, through simulation studies, we showed that the performance of proposed methods in both diagnosing out-of-control profiles and parameters is affected by values of ρ and $\rho_{\min}$. Moreover, a general guideline regarding the value of $\rho_{\min}$ is provided.

Acknowledgments

The authors are thankful for constructive and helpful comments by the anonymous referees which led to improvement in the paper.

References

Alt, F. B. (1985). *Multivariate Quality Control. The Encyclopedia of Statistical Sciences.* New York, NY: John Wiley & Sons.

Amiri, A., Jensen, W. A., Kazemzadeh, R. B. (2010). A case study on monitoring polynomial profiles in automotive industry. *Qual. Reliab. Eng. Int.* 26:509–520.

Chua, M. K., Montgomery, D. C. (1992). Investigation and characterization of a control scheme for multivariate quality control. *Qual. Reliab. Eng. Int.* 8:37–44.

Doganaksoy, N., Faltin, F. W., Tucker, W. T. (1991). Identification of out-of control characteristics in a multivariate manufacturing environment. *Commun. Stat. Theor. Methods*, 20:2775–2790.

Eyvazian, M., Noorossana, R., Saghaei, A., Amiri, A. (2011). Phase II monitoring of multivariate multiple linear regression profiles. *Qual. Reliab. Eng. Int.* 27:281–296.

Gupta, S., Montgomery, D. C., Woodall, W. H. (2006). Performance evaluation of two methods for online monitoring of linear calibration profiles. *Int. J. Prod. Res.* 44:1927–1942.

Hafezi, S., Shahriari, H. (2009). Interpretation of multivariate control chart based on paired decomposition of T^2. *Int. J. Ind. Eng. & Prod. Man.* 20:23–31.

Hawkins, D. M. (1993). Regression adjustment for variables in multivariate quality control. *J. Qual. Technol.* 25:170–182.

Hayter, A. J., Tsui, K. L. (1994). Identification and quantification in multivariate quality control problems. *J. Qual. Technol.* 26:197–208.

Jackson, J. E. (1991). *A User Guide to Principal Components.* New York, NY: John Wiley & Sons.

Kang, L., Albin, S. L. (2000). Online monitoring when the process yields a linear profile. *J. Qual. Technol.* 32:418–426.

Kazemzadeh, R. B., Noorossana, R., Amiri, A. (2008). Phase I monitoring of polynomial Profiles. *Commun. Stat. Theor. Methods* 37:1671–1686.

Kazemzadeh, R. B., Noorossana, R., Amiri, A. (2009). Monitoring polynomial profiles in quality control applications. *Int. J. Adv. Manufact. Technol.* 42:703–712.

Kim, K., Mahmoud, M. A., Woodall, W. H. (2003). On the monitoring of linear profiles. *J. Qual. Technol.* 35:317–328.

Mahmoud M. A. (2008). Phase I analysis of multiple linear regression profiles. *Commun. Stat. Simul. Comput.* 37:2106–2130.

Mahmoud, M. A., Parker, P. A., Woodall, W. H., Hawkins, D. M. (2007). A change point method for linear profile data. *Qual. Reliab. Eng. Int.* 23:247–268.

Mahmoud, M. A., Woodall, W. H. (2004). Phase I analysis of linear profiles with calibration applications. *Technometrics* 46:380–391.

Maravelakis, P. E., Bersimis, S., Panaretos, J., Psarakis, S. (2002). Identifying the out-of-control variable in a multivatiate control chart. *Commun. Stat. Theor. Methods* 31:2391–2408.

Mason, R. L., Tracy, N. D., Young, J. C. (1995). Decomposition of T^2 multivariate control chart interpretation. *J. Qual. Technol.* 27:99–108.

Mason, R. L., Tracy, N. D., Young, J. C. (1997). A practical approach for interpreting multivariate T2 control chart signals. *J. Qual. Technol.* 29:396–406.

Mestek, O., Pavlik, J., Suchanek, M. (1994). Multivariate control charts for calibration curves. *Fresenius J. Anal. Chem.* 350:344–351.

Murphy, B. J. (1987). Selecting out-of-control variables with T^2 multivariate quality procedures. *Statistician* 36:571–583.

Montgomery, D. C. (2005). *Introduction to Statistical Quality Control.* New York, NY: John Wiley & Sons.

Niaki, S. T., Abbasi, B. (2005). Fault diagnosis in multivariate control chart using artificial neural networks. *Qual. Reliab. Eng. Int.* 21:825–840.

Noorossana, R., Eyvazian, M., Amiri, A., Mahmoud, M. A. (2010a), Statistical monitoring of multivariate multiple linear regression profiles in Phase I with calibration application. *Qual. Reliab. Eng. Int.* 26:291–303.

Noorossana, R., Eyvazian, M., Vaghefi, A. (2010b), Phase II monitoring of multivariate simple linear profiles. *Comput. Ind. Eng.* 58:563–570.

Rencher, A. C. (2002). *Methods of Multivariate Analysis.* New York, NY: John Wiley & Sons, Inc.

Runger, G. C., Alt, F. B., Montgomery, D. C. (1996). Contributors to a multivariate statistical process control chart signal. *Commun. Stat. Theor. Methods* 25:2203–2213.

Saghaei, A., Mehrjoo, M., Amiri, A. (2009). A CUSUM based method for monitoring simple linear profiles. *Int. J. Adv. Manufact. Technol.* 45:1252–1260.

Stover, F. S., Brill, R. V. (1998). Statistical quality control applied to ion chromatography calibration. *J. Chrom. A* 804:37–43.

Wade, M. R., Woodall, W. H. (1993). A review and analysis of cause selecting control charts. *J. Qual. Technol.* 25:161–169.

Wang, K., Tsung, F. (2005). Using profile monitoring techniques for a data-rich environment with huge sample size. *Qual. Reliab. Eng. Int.* 21:677–688.

Woodall, W. H. (2007). Current research on profile monitoring. *Rev. Prod.* 17:420–425.

Woodall, W. H., Montgomery, D. C., Gupta, S. (2004). Using control charts to monitor process and product quality profiles. *J. Qual. Technol.* 36:309–320.

Zhang, J., Li, Z., Wang, Z. (2009). Control chart based on likelihood ratio for monitoring linear profiles. *Comput. Stat. Data Anal.* 53:1440–1448.

Zhu, Y., Jiang, W. (2009). An adaptive chart for multivariate process monitoring and diagnosis. *IIE Trans.* 41:1007–1018.

Zou, C., Tsung, F., Wang, Z. (2007). Monitoring general linear profiles using multivariate exponentially weighted moving average schemes. *Technometrics* 49:395–408.

Appendix A: T^2 Control Chart with Warning Limit to Diagnose the Out-of-Control Profiles

Assume that the Cth and Dth profiles for out-of-control sample (lth sample) are selected. The value of T^2 statistic for this pair of profiles is as follows:

$$T_l^2 = \begin{bmatrix} \bar{y}_{Cl} - \mu_{\bar{y}_C} & \bar{y}_{Dl} - \mu_{\bar{y}_D} \end{bmatrix} \boldsymbol{\Sigma}^{-1}_{(\bar{y}_C, \bar{y}_D)} \begin{bmatrix} \bar{y}_{Cl} - \mu_{\bar{y}_C} \\ \bar{y}_{Dl} - \mu_{\bar{y}_D} \end{bmatrix}, \tag{A1}$$

where

$$\boldsymbol{\Sigma}_{\bar{y}_C, \bar{y}_D} = \begin{bmatrix} \sigma^2_{\bar{y}_C} & \sigma_{\bar{y}_C, \bar{y}_D} \\ \sigma_{\bar{y}_C, \bar{y}_D} & \sigma^2_{\bar{y}_D} \end{bmatrix}. \tag{A2}$$

The correlation coefficient between Cth and Dth average of profiles is set to $\rho_{C,D}$ where equals to

$$\rho_{C,D} = \frac{\sigma_{\bar{y}_C, \bar{y}_D}}{\sqrt{\frac{\sigma^2_{\bar{y}_C}}{n}} \times \sqrt{\frac{\sigma^2_{\bar{y}_D}}{n}}} \Rightarrow \rho_{C,D} \left(\sqrt{\frac{\sigma^2_{\bar{y}_C}}{n}} \times \sqrt{\frac{\sigma^2_{\bar{y}_D}}{n}} \right) = \sigma_{\bar{y}_C, \bar{y}_D} \Rightarrow \frac{\left(\sqrt{\sigma^2_{\bar{y}_C} . \sigma^2_{\bar{y}_D}} \right) \rho_{C,D}}{n} = \sigma_{\bar{y}_C, \bar{y}_D}. \tag{A3}$$

Now by expanding Eq. (A1):

$$\begin{aligned} T_l^2 = {} & \frac{1}{\sigma^2_{\bar{y}_C} \sigma^2_{\bar{y}_D} (1 - \rho^2_{(C,D)})} \Big[(\bar{y}_{Cl} - \mu_{\bar{y}_C})^2 \sigma^2_{\bar{y}_D} - \rho_{(\bar{y}_C, \bar{y}_D)} \sigma_{\bar{y}_C} \sigma_{\bar{y}_D} \left(\bar{y}_{Dl} - \mu_{\bar{y}_D} \right) \\ & \times \sigma^2_{\bar{y}_C} \left(\bar{y}_{Dl} - \mu_{\bar{y}_D} \right)^2 - (\bar{y}_C l - \mu_{\bar{y}_C})^2 \rho_{(\bar{y}_C, \bar{y}_D)} \sigma_{\bar{y}_C} \sigma_{\bar{y}_D} \Big] \begin{bmatrix} \bar{y}_{Cl} - \mu_{\bar{y}_C} \\ \bar{y}_{Dl} - \mu_{\bar{y}_D} \end{bmatrix}. \end{aligned} \tag{A4}$$

Hence,

$$\left(\frac{\bar{y}_{Cl}-\mu_{\bar{y}_C}}{\sigma_{\bar{y}_C}}\right)^2 \times \frac{1}{1-\rho^2_{(C,D)}} + \left(\frac{\bar{y}_{Dl}-\mu_{\bar{y}_D}}{\sigma_{\bar{y}_D}}\right)^2 \times \frac{1}{1-\rho^2_{(C,D)}} - 2\left(\frac{\bar{y}_{Cl}-\mu_{\bar{y}_C}}{\sigma_{\bar{y}_C}}\right)$$
$$\times \left(\frac{\bar{y}_{Dl}-\mu_{\bar{y}_D}}{\sigma_{\bar{y}_D}}\right) \times \frac{\rho_{(C,D)}\sigma_{\bar{y}_C}\sigma_{\bar{y}_D}}{\left(1-\rho^2_{(C,D)}\right)\sigma_{\bar{y}_C}\sigma_{\bar{y}_D}} = T_l^2. \tag{A5}$$

Finally,

$$\left(\frac{\bar{y}_{Cl}-\mu_{\bar{y}_C}}{\sigma^2_{\bar{y}_C}}\right)^2 + \left(\frac{\bar{y}_{Dl}-\mu_{\bar{y}_D}}{\sigma^2_{\bar{y}_D}}\right)^2 - 2\left(\frac{\bar{y}_{Cl}-\mu_{\bar{y}_C}}{\sigma^2_{\bar{y}_C}}\right)\left(\frac{\bar{y}_{Dl}-\mu_{\bar{y}_D}}{\sigma^2_{\bar{y}_D}}\right)$$
$$\times\rho_{(C,D)} = T_l^2\left(1-\rho^2_{(C,D)}\right). \tag{A6}$$

Appendix B: T^2 Control Chart Based on the EWMA Statistic to Diagnose the Out-of-Control Profiles

Assume that the Cth and Dth profiles are selected. Then the average of each residual of profile is computed. So, the value of T^2 statistic for this pair in lth out-of-control sample is as follows:

$$T_l^2 = \left[\, z_{Cl}-\boldsymbol{\mu}_{Z_C} \quad z_{Dl}-\boldsymbol{\mu}_{Z_D} \,\right] \boldsymbol{\Sigma}^{-1}_{(z_C,z_D)} \begin{bmatrix} z_{Cl}-\boldsymbol{\mu}_{z_C} \\ z_{Dl}-\boldsymbol{\mu}_{z_D} \end{bmatrix}, \boldsymbol{\mu}_{z_C} = \boldsymbol{\mu}_{z_D} = 0, \tag{B1}$$

The variance covariance matrix equals to $\boldsymbol{\Sigma}_{(z_C,z_D)} = \frac{\omega}{2-\omega}\boldsymbol{\Sigma}_{(\bar{\boldsymbol{e}}_C,\bar{\boldsymbol{e}}_D)}$.

And the correlation coefficient between Cth and Dth profiles is set to $\rho_{C,D}$ where equals to

$$\rho_{C,D} = \frac{\sigma_{y_C,y_D}}{\sqrt{\frac{\sigma^2_{y_C}}{n}} \times \sqrt{\frac{\sigma^2_{y_D}}{n}}}. \tag{B2}$$

$$T_l^2 = \frac{\omega}{\sigma^2_{y_C}\sigma^2_{y_D}(1-\rho^2_{(C,D)})(2-\omega)}\left[(z_{Cl})^2\sigma^2_{y_D} - \rho_{(C,D)}\sigma_{y_C}\sigma_{y_D}(z_{Dl})\right.$$
$$\left.\times\,\sigma^2_{y_C}(z_{Dl})^2 - (z_{Cl})^2\rho_{(C,D)}\sigma_{y_C}\sigma_{y_D}\right]\left[\, z_{Cl}\ z_{Dl} \,\right]^T. \tag{B3}$$

By simplifying Eq. (B3), the T^2 statistic is given as follows:

$$\left(\frac{Z_{Cl}}{S_{Cl}}\right)^2 + \left(\frac{Z_{Dl}}{S_{Dl}}\right)^2 - 2\left(\frac{Z_{Cl}}{S_{Cl}}\right)\left(\frac{Z_{Dl}}{S_{Dl}}\right) \times \rho_{(C,D)} = T_l^2\frac{\left(1-\rho^2_{(C,D)}\right)(2-\omega)}{\omega}, \tag{B4}$$

where S_{ii} is variance of the residuals of ith profile which equals to $\sigma ii/_n$.