Cluster-Based Profile Monitoring in Phase I Analysis

Yajuan Chen

Department of Statistics

Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0439, USA.

Jeffrey B. Birch

Department of Statistics

Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0439, USA.

Abstract

An innovative profile monitoring methodology is introduced for Phase I analysis. The proposed technique, which is referred to as the *cluster-based profile monitoring method,* incorporates a cluster analysis phase to aid in determining if non conforming profiles are present in the historical data set (HDS). To cluster the profiles, the proposed method first replaces the data for each profile with an estimated profile curve, using some appropriate regression method, and clusters the profiles based on their estimated parameter vectors. This cluster phase then yields a main cluster which contains more than half of the profiles. The initial estimated population average (PA) parameters are obtained by fitting a linear mixed model to those profiles in the main cluster. In-control profiles, determined using the Hotelling's T^2 statistic, that are not contained in the initial main cluster are iteratively added to the main cluster and the mixed model is used to update the estimated PA parameters. A simulated example and Monte Carlo results demonstrate the performance advantage of this proposed method over a current non-cluster based method with respect to more accurate estimates of the PA parameters and better classification performance in determining those profiles from an in-control process from those from an out-of-control process in Phase I.

Key Words: Mixed Models; Outliers; Quality control; Robust Profile Monitoring; T^2 Statistic.

Introduction

In Phase I profile monitoring analysis, one goal is to distinguish between those profiles from the in-control process (called the normal profiles) in the HDS from those profiles from the out-of-control process (called the outlying profiles). The outlying profiles are usually removed and the remaining normal profiles are used to compute the statistics needed for establishing the in-control limits used in Phase II analysis.

To detect the abnormal profiles in Phase I of the profile monitoring process, several authors, including Kang and Albin (2000), Kim, et al. (2003) and Mahmoud and Woodall (2004) utilized the Hotelling's T^2 statistic to determine abnormal profiles based on the estimated regression parameters. Nonlinear and nonparametric profile applications were studied by Jin and Shi (1999), Walker and Wright (2002), Gupta, et al. (2006), Ding, et al. (2006), Williams, et al. (2007a), Williams, et al. (2007b) and Hung, et al. (2012). Jensen, et al. (2008), Jensen and Birch (2009) and Qiu, et al. (2010) proposed the use of mixed models to monitor the profiles in order to account for the correlation structure within profiles. Based on mixed models, Jensen, et al. (2008) and Jensen and Birch (2009) proposed detecting abnormal profiles by comparing each estimated profile specific (PS) curve to the estimated population average (PA) curve using the T^2 statistic. Jensen, et al. (2008) proposed the use of the T^2 statistic to determine abnormal profiles in the parametric mixed model and showed the equivalence between this approach and using the T^2 statistic based on the estimated best linear predictors (eblups) of each profile. This is the method that will be utilized in our paper to illustrate a non-cluster based method. It needs to be pointed out, however, that any reputable profile monitoring method for Phase I analysis could be used to illustrate a non-cluster based method. The proposed cluster based method would then be adjusted accordingly to account for this other method. It is believed that the advantages of clustering

demonstrated in the example and in the Monte Carlo study would still be present regardless of the type of non-cluster based method used.

The performance of a Phase I analysis method can be measured in terms of the method's ability to correctly identify the presence of abnormal profiles in the HDS. An important criterion used to measure the success of a Phase I method at detecting an unstable process is the probability of signal (POS), the probability of detecting at least one outlying profile in the HDS. However, one problem with many methods discussed above is that the estimated PA profile is based on averaging the fits of all the profiles, including any profiles from the out-of-control process. Thus, the estimated PA profile will be "pulled" in the direction of the profiles from the out-of-control process resulting in a biased estimate of the true PA profile. Additionally, the corresponding variance-covariance matrix, needed for computing the T^2 statistic for each estimated PS curve, will be similarly distorted. Consequently, the T^2 statistics can be misleading and the in-control limits used in Phase I will be less able properly separate those profiles belonging to the in-control process from those belonging to the out-of-control process. Further, the performance of previous methods is measured by using the POS, which only measures the ability of detecting the presence of outlying profiles in the HDS. However, the POS does not indicate whether the classification of profiles into the two categories of normal and outlying is correctly specified.

A new profile monitoring method, referred to as the cluster based profile monitoring method, is proposed to obtain T^2 statistics that are robust to outlying profiles in Phase I. Also, a classification table is identified that suggests other performance metrics, in addition to the POS, be used to evaluate a method's ability to properly classify profiles into the normal and outlying categories.

A simple example below gives a comparison of the proposed cluster based method to the existing non-cluster based method of Jensen, et al. (2008). In this example, it assumed that there are total 12

profiles in the HDS where nine are from the in-control process while the other three are from the out-of-control process. The normal profiles were generated from the linear mixed model (LMM)

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_2 + b_{1i})x_{ij} + (\beta_3 + b_{2i})x_{ij}^{\ 2} + \varepsilon_{ij},\ i = 1, 2, ..., m_1,\ j = 1, 2, ..., n, \quad (1)$$

and the outlying profiles were generated via the LMM as

$$y_{ij} = (\beta_0' + b_{0i}) + (\beta_1' + b_{1i})x_{ij} + (\beta_2' + b_{2i})x_{ij}^{\ 2} + \varepsilon_{ij},\ \ i = m_1 + 1, ..., m,\ \ j = 1, 2, ..., n, \quad (2)$$

where the random effects are defined as

$$\begin{pmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{pmatrix} \sim MN\left(\mathbf{0}\ , \begin{bmatrix} \sigma_0^2 & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & 0 & \sigma_2^2 \end{bmatrix} \right),$$

$$\boldsymbol{\varepsilon} \sim MN\left(0,\ \ \sigma^2 \boldsymbol{I}\right),$$

(here MN represents the multivariate normal distribution) and with fixed effects $\boldsymbol{\beta}^T = (12.5,\ \ -7,\ \ 2)$ for the normal profiles and $\boldsymbol{\beta}'^T = (21.875,\ \ -14.5,\ \ 3.5)$ for the outlying profiles. Additionally, $m_1 = 9, m = 12$ and $\sigma_0^2 = \sigma_1^2 = \sigma_2^2 = 0.5$ and $\sigma^2 = 4.$ Thus, profiles 10, 11, and 12 are outlying profiles.

The 12 true profiles, based on the actual parameter values and random effects, are plotted in Figure 1.1 where the blue curves represent the normal profiles while the red curves represent the outlying profiles. It is difficult to distinguish the three outlying profiles from the normal profiles by looking only at the plot.

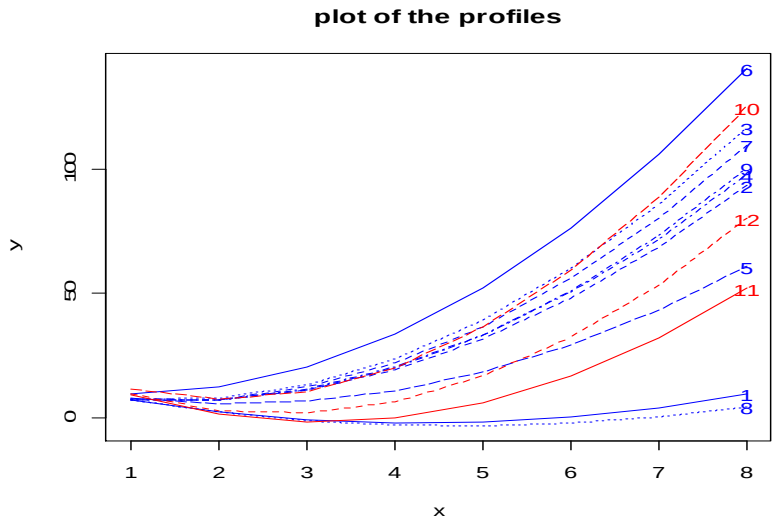

Figure 1.1: The plot of 12 true profiles

Using the T^2 statistic, both the existing non-cluster based method and the proposed cluster based method signaled, indicating that both methods detected a change in the process. However, the non-cluster based method signaled due to misclassifying the 6th profile as the outlying profile. The cluster based method, on the other hand, correctly classified the 10th, 11th and 12th profiles as outlying profiles and classified the other nine profiles as normal profiles. The estimates of the PA parameters from the non-cluster based method (Jensen, et al. (2008)) are $\hat{\boldsymbol{\beta}}^T = (20.081,\ -14.214,\ 2.737)$ while the estimates of the PA parameters from the proposed method are $\hat{\boldsymbol{\beta}}^T = (14.486,\ -7.764,\ 2.027)$. Compared to the true PA parameters, $\boldsymbol{\beta}^T = (12.5,\ -7,\ 2)$, the estimates of the non-cluster based method (Jensen, et al. (2008)) are severely distorted while the proposed method provided PA estimates much closer to the true values, as expected. This example will be illustrated in greater detail in section 3.2.

Review of the non-cluster based method

In the HDS, the representation of the i^{th} profile using the linear mixed model (LMM) (Laird and Ware (1982)) is

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, 2, \cdots, m \quad j = 1, 2, \cdots, n_i, \tag{3}$$

where $\boldsymbol{y}_i$ is the $n_i \times 1$ response vector for the i^{th} profile, $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ are $n_i \times p$ and $n_i \times q$, respectively, matrices of explanatory variables, $\boldsymbol{b}_i$ is a $q \times 1$ vector of random effects for the i^{th} profile with $\boldsymbol{b}_i \sim MN(\boldsymbol{0}, \boldsymbol{G})$ and $\boldsymbol{G}$ is a $q \times q$ covariance matrix. $\boldsymbol{\varepsilon}_i$ is the random error term for the i^{th} profile with $\boldsymbol{\varepsilon}_i \sim MN(\boldsymbol{0}, \boldsymbol{R}_i)$. For more details of the LMM, see Schabenberger and Pierce (2002), Seber and Wild (2003) , and Demidenko (2004). The convenient way to derive an estimator of $\boldsymbol{\beta}$ is to stack the responses and the model matrices for the m individual profiles. Let $\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_m \end{pmatrix}$, $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_m \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_m \end{pmatrix}$, $n = \sum_{i=1}^{m} n_i$, and $\boldsymbol{Z}$ is the $n \times mq$ block diagonal matrix with $\boldsymbol{Z}_i$ along each diagonal, $\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_1 & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & \boldsymbol{Z}_m \end{bmatrix}$.

The above model can be written as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{b} + \boldsymbol{\varepsilon}. \tag{4}$$

With the stack equation above, the corresponding distributions for $\boldsymbol{b}$ and $\boldsymbol{\varepsilon}$ can be written as

$$\boldsymbol{b} \sim MN(\boldsymbol{0}, \boldsymbol{G}), \tag{5}$$

$$\boldsymbol{\varepsilon} \sim MN(\boldsymbol{0}, \boldsymbol{R}), \tag{6}$$

where $\boldsymbol{R} = diag\left(\boldsymbol{R}_i\right),$ and the conditional and marginal distributions for $\boldsymbol{y}$ are

$$\boldsymbol{y} \mid \boldsymbol{b} \sim MN\left(\boldsymbol{X\beta} + \boldsymbol{ZB}, \boldsymbol{R}\right), \tag{7}$$

And

$$\boldsymbol{y} \sim MN\left(\boldsymbol{X\beta}, \boldsymbol{\Sigma}\right), \tag{8}$$

where

$$\boldsymbol{\Sigma} = \text{var}\left(\boldsymbol{y}\right) = \boldsymbol{ZGZ}^T + \boldsymbol{R}.$$

We denote by $\hat{\boldsymbol{\beta}}_{LMM}$ the estimator for the PA parameter vector for the fixed effects and denote by $\hat{\boldsymbol{b}}_i$ the eblups of the random effects for the i^{th} profile. Then it can be shown that (Schabenberger and Pierce (2002))

$$\hat{\boldsymbol{\beta}}_{LMM} = \left(\boldsymbol{X}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{y}, \tag{9}$$

$$\hat{\boldsymbol{b}}_i = \boldsymbol{GZ}_i^T \boldsymbol{\Sigma}_i^{-1} \left(\boldsymbol{y} - \boldsymbol{X}_i \hat{\boldsymbol{\beta}}\right). \tag{10}$$

Note, $\boldsymbol{\Sigma}$ here is usually unknown and needs to be estimated first. The most commonly used estimators for $\boldsymbol{\Sigma}$ include the maximum likelihood estimator (MLE) and the restricted maximum likelihood estimator (REMLE) (Ruppert, et al. (2003)). By substituting the estimates $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{G}}$, the parameter estimates and eblups can be obtained. Subsequently, the estimated parameter vector and eblups for the i^{th} profile are

$$\hat{\boldsymbol{\beta}}_i = \hat{\boldsymbol{\beta}}_{LMM} + \hat{\boldsymbol{b}}_i^*, \tag{11}$$

where $\hat{\boldsymbol{\beta}}_{LMM} = \left(\boldsymbol{X}^T \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{y}$ and $\hat{\boldsymbol{b}}_i^*$ is a $p \times 1$ vector containing $\hat{\boldsymbol{b}}_i$ for the columns of $\boldsymbol{Z}_i$ that are equal to the columns of $\boldsymbol{X}_i$ and zero otherwise. Consequently, $\hat{\boldsymbol{b}}_i = \hat{\boldsymbol{b}}_i^*$ if $\boldsymbol{X}_i = \boldsymbol{Z}_i$. The estimated fits for PS_i curve and for the PA curve are expressed as

$$\hat{\boldsymbol{y}}_{PS,i} = \boldsymbol{X}_i \hat{\boldsymbol{\beta}}_i = \boldsymbol{X}_i \hat{\boldsymbol{\beta}}_{LMM} + \boldsymbol{Z}_i \hat{\boldsymbol{b}}_i, \tag{12}$$

and

$$\hat{\boldsymbol{y}}_{PA} = \boldsymbol{X}_i \hat{\boldsymbol{\beta}}_{LMM}. \tag{13}$$

Jensen, et al. (2008) proposed a parametric approach to determine the unusual profiles based on the distance of the estimated parameter vector from the center of the group of estimated parameter vectors. They introduced a formula for the T^2 statistic based on comparing $\hat{\boldsymbol{\beta}}_i$ to the sample mean of $\hat{\boldsymbol{\beta}}_i$, $\hat{\boldsymbol{\beta}}_{LMM}$. The T^2 statistic for the i^{th} estimated PS curve is defined as

$$T_i^2 = \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{LMM}\right)^T \hat{\boldsymbol{V}}^{-1} \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{LMM}\right), \tag{14}$$

where $\hat{\boldsymbol{V}}$ is the estimated variance covariance matrix of $\hat{\boldsymbol{\beta}}_i$. The successive difference estimator, $\hat{\boldsymbol{V}}_D$, first introduced by Hawkins and Merriam (1974) is preferred here. Sullivan and Woodall (1996) showed that $\hat{\boldsymbol{V}}_D$ is effective in detecting sustained step changes in the process that may occur in Phase I data. The successive difference estimator of $\boldsymbol{V}$ is

$$\hat{\boldsymbol{V}}_D = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} \left(\hat{\boldsymbol{\beta}}_{i+1} - \hat{\boldsymbol{\beta}}_i\right)^T \left(\hat{\boldsymbol{\beta}}_{i+1} - \hat{\boldsymbol{\beta}}_i\right). \tag{15}$$

Jensen, et al. (2008) showed that the distribution of T^2 follows asymptotically a chi-squared distribution with p degrees of freedom for large samples, where p is the number of estimated

parameters. Since $\sum_{i=1}^{m} \hat{\boldsymbol{b}}_i = \boldsymbol{0}$, it follows that (Jensen, et al. (2008)) the above formulas can be written equivalently as

$$T_i^2 = \hat{\boldsymbol{b}}_i^T \hat{\boldsymbol{V}}_D^{-1} \hat{\boldsymbol{b}}_i,$$

and

$$\hat{\boldsymbol{V}}_D = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} \left(\hat{\boldsymbol{b}}_{i+1} - \hat{\boldsymbol{b}}_i\right)^T \left(\hat{\boldsymbol{b}}_{i+1} - \hat{\boldsymbol{b}}_i\right).$$

The Cluster Based Method

The proposed cluster based profile monitoring method is designed to provide a procedure that is robust to outlying profiles for the Phase I profile monitoring process. The main idea is to first cluster the profiles to obtain a set of initial main cluster profiles with similar shapes. A cluster based method has been used previously in the robust regression context to cluster n independent $p \times 1$ vectors by Lawrence (2003). Jobe and Pokojovy (2009) also proposed a cluster based method for use with multivariate control charts. However, clustering in the profile monitoring context is more complex than clustering data points in that the goal now is to cluster estimated curves involving intra-profile correlated data. A general discussion of the method is outlined below followed by a more detailed discussion.

The first step is to fit a curve, by some appropriate method, to each of m independent $n_i \times 1$ profiles (vectors) where the data within each profile are likely to be correlated. The proposed method thus allows each estimated profile to be represented by a vector of estimated model parameters. After each profile is represented with a parameter vector, the estimated variance-covariance matrix estimator, $\hat{\boldsymbol{V}}$, can be calculated by using the successive difference of the estimated parameter vectors. The second step is to

calculate the similarity matrix $\boldsymbol{S}$ based on the estimated parameter vectors and $\hat{\boldsymbol{V}}$. Then, an appropriate cluster method is used to cluster each profile based on $\boldsymbol{S}$.

To obtain a tight, compact sphere of similar profiles, hierarchical clustering with complete linkage is performed until an initial main cluster of more than half of the profiles is formed. After obtaining an initial main cluster, denoted by C_{main}, the profiles in C_{main} can be used to obtain an initial estimate of PA. This estimated curve can be used with the previously estimated variance-covariance matrix, $\hat{\boldsymbol{V}}$, to calculate the T^2 statistics for the profiles not in C_{main}. The profiles which have in-control T^2 statistics are then added to C_{main} to obtain a new set of profiles, denoted as C_{new}. Then, the mixed model approach is used to update the estimate of the PA profile from the profiles in C_{new}. Repeat the above procedure of updating C_{new} by adding the profiles not in C_{new} until either the smallest T^2 statistic for the remaining profiles outside of C_{new} is beyond the control limits or all the profiles have been added to C_{new}. Upon completion of the algorithm, those profiles contained in C_{new} are labeled as "in-control profiles" and those not included in C_{new} are labeled as "outlying profiles". The proposed algorithm is now outlined in detail.

Step 1

Represent each estimated profile by an estimated parameter vector (obtained using some appropriate method) and determine the $m \times p$ parameter matrix $\hat{\boldsymbol{B}}$. The i^{th} row of $\hat{\boldsymbol{B}}$, denoted by $p \times 1$ vector $\hat{\boldsymbol{\beta}}_i$, is defined as estimated parameter vector for the i^{th} profile. Use the successive difference estimator to obtain the estimated variance-covariance matrix, $\hat{\boldsymbol{V}}$, for $\hat{\boldsymbol{B}}$, as

$$\hat{\boldsymbol{V}} = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} \left(\hat{\boldsymbol{\beta}}_{i+1} - \hat{\boldsymbol{\beta}}_i \right)^T \left(\hat{\boldsymbol{\beta}}_{i+1} - \hat{\boldsymbol{\beta}}_i \right).$$

Step 2

Using $\hat{\boldsymbol{V}}$ obtained in step 1, compute a $m \times m$ similarity matrix $\boldsymbol{S}$,where the i, j entry is defined as

$$s_{ij} = \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_j\right)^T \hat{\boldsymbol{V}}^{-1} \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_j\right),$$

where $\hat{\boldsymbol{\beta}}_i$ and $\hat{\boldsymbol{\beta}}_j$ are i^{th} and j^{th} rows of $\hat{\boldsymbol{B}}$, respectively.

Step 3

Perform a cluster analysis on the given similarity matrix and use complete-linkage to obtain the clusters of $\hat{\boldsymbol{\beta}}_i$. The main cluster is defined as the first cluster that contains more than half of the profiles. Denote the indices of the main cluster as C_{main}. Stop the cluster process as soon as at least $\left[\frac{m}{2}\right]+1$ profiles are contained in the main cluster. Since new profiles may be added to C_{main} during the iteration process, we denote by C the main cluster at each iteration step. Thus at the end of step 3, $C = C_{main}$.

Step 4

Use the mixed model approach to estimate the PA profile for profiles in C, denoted as $\hat{\boldsymbol{\beta}}_{PA}$. For all profiles not contained in C, compute

$$T_i^2 = \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{PA}\right)^T \hat{\boldsymbol{V}}^{-1} \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{PA}\right),$$

where " i " denotes the i^{th} profile not contained in C and add the profiles which have $T_i^2 < \chi^2_{\left[1-\frac{\alpha}{m}\right], df=p}$ to C and obtain a new index set C_{new}.

Step 5

If the profiles in C_{new} are different from the profiles in C set $C = C_{new}$, and go back to step 4, otherwise set the final profiles in C_{new} as C_{final}.

Step 6

Use the mixed model approach to estimate the PA profile parameters for profiles in C_{final}. Denote this PA profile as $\hat{\boldsymbol{\beta}}_{CPA}$, the eblups for the i^{th} PS curve by $\hat{\boldsymbol{b}}_{C,i}$ and variance-covariance matrix, $\hat{\boldsymbol{V}}_C$. Here, the "C" in the subscript denotes that the estimates result from the cluster-based method.

A Detailed example

To aid in understanding the proposed algorithm, a detailed analysis of the example is now provided. Recall that the example in section 1 has nine profiles from the in-control process and three profiles from the out-of-control process. Figure 1.1 shows the true 12 profiles. Eight observations taken at the same equally spaced regressor values were randomly generated from the models for each of the 12 profiles. The data, connected by straight line segments for each profile, are displayed in Figure 3.1.

In Figure 3.1, it is easy to see that all profiles show a quadratic trend and it is reasonable to use the quadratic model to represent these profiles. The curves in red represent the outlying profiles, though, at this stage of the analysis, this fact is neither known nor clear from the plot that these three red curves are "different" from the nine blue curves.

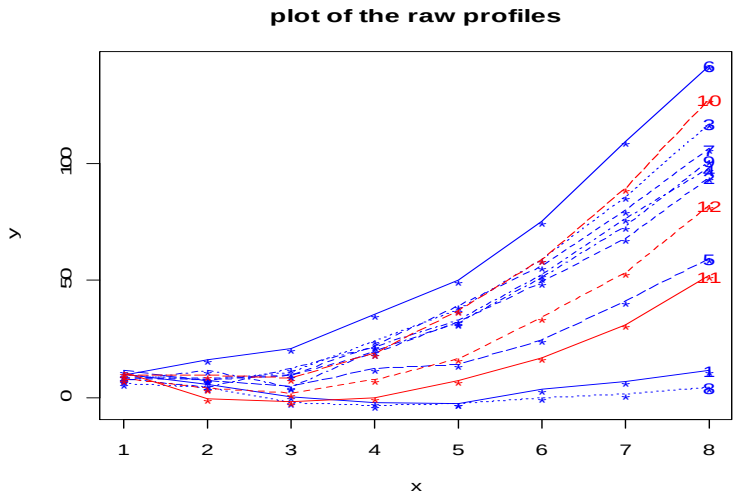

Figure 3.1: The plot of 12 observed profiles

Step 1

The parameters for each profile are estimated individually using the fixed effects quadratic model in one regressor and the method of least squares. The estimated parameters for each profile are listed in Table 3.1, with the last three columns representing the $\hat{\boldsymbol{B}}$ matrix. The estimated variance-covariance matrix, $\hat{\boldsymbol{V}}$, for the $\hat{\boldsymbol{\beta}}_i$ is computed using the successive difference estimator.

Table 3.1: 12×3 $\hat{\boldsymbol{B}}$ *matrix; the parameter estimates for* 12 *profiles*

Index of profiles	$\hat{\beta}_{0i}$	$\hat{\beta}_{1i}$	$\hat{\beta}_{2i}$
1	18.393	-9.171	1.055
2	13.14	-7.072	2.149
3	15.41	-9.214	2.748
4	9.743	-5.554	2.1
5	20.558	-10.704	1.941
6	15.127	-6.44	2.791
7	11.069	-6.338	2.299
8	12.029	-6.316	0.68
9	14.907	-9.068	2.488
10	21.645	-14.318	3.441
11	21.892	-14.832	2.324
12	20.081	-14.214	2.737

.

$$\hat{\boldsymbol{V}} = \begin{bmatrix} 4.503 & -4.285 & 0.387 \\ -4.285 & 5.035 & -0.494 \\ 0.387 & -0.494 & 0.492 \end{bmatrix}$$

Step 2

Using $\hat{\boldsymbol{V}}$ computed in step 1, obtain the similarity matrix $\boldsymbol{S}$, presented in table 3.2.

Table 3.2: Similarity matrix using $s_{ij} = \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_j\right)^T \hat{\boldsymbol{V}} \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_j\right)$

s_{ij}	1	2	3	4	5	6	7	8	9	10	11	12
1	0	5.19	9	9.77	1.81	11.55	8.96	4.57	8.7	23.65	24.4	29.37
2	5.19	0	1.89	1.18	4.77	8.45	0.65	4.55	1.97	16.43	21.26	22.32
3	9	1.89	0	3.26	5.86	13.32	1.92	9.12	0.24	7.43	12.71	12.79
4	9.77	1.18	3.26	0	10.46	13.61	0.2	4.57	2.74	18.99	22.79	22.54
5	1.81	4.77	5.86	10.46	0	8.52	8.61	9.41	6.56	16.28	20.7	24.76
6	11.55	8.45	13.32	13.61	8.52	0	12.07	19.73	15.85	34.96	48.23	50.63
7	8.96	0.65	1.92	0.2	8.61	12.07	0	5.34	1.7	16.07	20.68	20.51
8	4.57	4.55	9.12	4.57	9.41	19.73	5.34	0	7.33	26.19	23.5	26.41
9	8.7	1.97	0.24	2.74	6.56	15.85	1.7	7.33	0	7.56	11.08	11.24
10	23.65	16.43	7.43	18.99	16.28	34.96	16.07	26.19	7.56	0	3.62	3.08
11	24.4	21.26	12.71	22.79	20.7	48.23	20.68	23.5	11.08	3.62	0	0.75
12	29.37	22.32	12.79	22.54	24.76	50.63	20.51	26.41	11.24	3.08	0.75	0

Step3

Perform the cluster analysis on the similarity matrix using complete-linkage. The cluster process is represented by a dendrogram in Figure 3.2. Since there are 12 profiles, the initial main cluster must consist of at least seven profiles. The dendrogram reveals that in the fifth step of the clustering process, two profiles (with indices 1 and 5) are added to a cluster containing six profiles, resulting in a new

cluster containing eight profiles. Since this is the first cluster formed with at least seven profiles, the initial main cluster will contain these eight profiles. This cluster step ends with an initial main cluster (seen on the right side of the dendrogram in Figure 3.2) and two minor clusters. The initial main cluster contains profiles 1-5, and 7-9. Using the profile index to represent each profile, the initial main cluster is defined as $C_{main} = \{1{:}5\ ,7{:}9\}$ and $C = C_{main} = \{1{:}5\ ,7{:}9\}$.

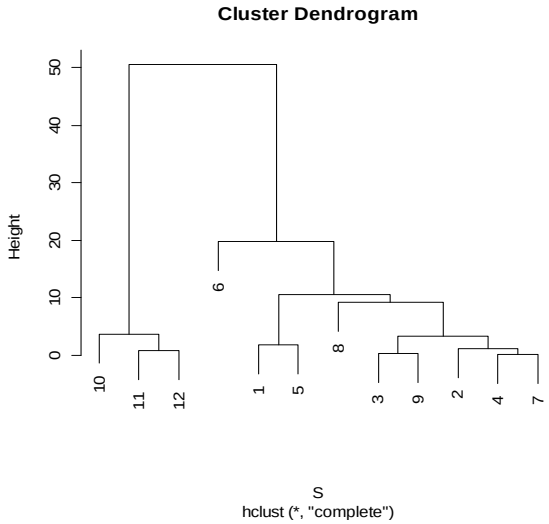

Figure 3.2: Dendrogram for clustering of example dataset.

Step 4

The LMM is used to obtain the PA parameter estimate $\hat{\boldsymbol{\beta}}_{PA}$ based on the profiles in C as

$$\hat{\boldsymbol{\beta}}_{PA}{}^{T} = (14.406,\quad -7.930,\quad 1.932),$$

and the T_i^2 statistics for those profiles not contained in C are displayed below

$i \notin C$	6	10	11	12
T_i^2	10.695	14.381	17.446	19.049

The cutoff value of T_i^2 is

$$cutoff = \chi^2_{\left[1-\frac{\alpha}{m}\right],df=3} = 13.229$$

Since the 6th profile has the T^2 statistics less than the cutoff, this profile is added to C to obtain $C_{new} = \{1{:}9\}$.

Step 5

Since $C \neq C_{new}$, set $C = C_{new} = \{1{:}9\}$ and repeat step 4 using the LMM. The updated $\hat{\boldsymbol{\beta}}_{PA}$ and the T^2 statistics are obtained as

$$\hat{\boldsymbol{\beta}}^T_{PA} = (14.486, \quad -7.764, \quad 2.027)$$

$i \notin C$	10	11	12
T_i^2	15.611	19.811	21.502

Since the T^2 statistics above show that no profile can be added, the algorithm stops here with $C_{final} = \{1{:}9\}$.

Step 6

All profiles in the final set C_{final} are used with the LMM model to estimate the PA parameter vector $\hat{\boldsymbol{\beta}}_{CPA}$, eblups $\hat{\boldsymbol{b}}_{i,C}$ and variance-covariance matrix $\hat{\boldsymbol{V}}_{\boldsymbol{C}}$ as

$$\hat{\boldsymbol{\beta}}^T_{PA} = (14.486, \quad -7.764, \quad 2.027).$$

The successive difference estimate $\hat{\boldsymbol{V}}_{\boldsymbol{C}}$ based on the ebulps is

$$\hat{\boldsymbol{V}}_{\boldsymbol{C}} = \begin{bmatrix} 2.110 & -0.969 & -0.643 \\ -0.969 & 0.619 & 0.209 \\ -0.643 & 0.209 & 0.462 \end{bmatrix}.$$

Table 3.3: eblups for the profiles in C_{final}

Index of profiles	$\widehat{b}_{0i}$	$\widehat{b}_{1i}$	$\widehat{b}_{2i}$
1	2.045	-0.735	-1.028
2	-0.524	0.271	0.164
3	-0.299	-0.354	0.586
4	-1.502	0.686	0.222
5	1.927	-0.933	-0.287
6	-1.27	0.499	0.358
7	0.474	-0.072	-1.19
8	-0.205	-0.44	0.347
9	-0.645	1.078	0.829

The example shows that the algorithm correctly identifies the three outlying profiles. In the cluster phase, the algorithm gives the initial main cluster of profiles as and two corresponding minor clusters and with and In the profile clustering process, the profile in the minor cluster is added to the initial main cluster while the profiles in are not added. This, of course, is the desired result. After correctly identifying the outlying profiles, the final PA profile and variance-covariance matrix were estimated by using the in-control profiles in C_{final}. The cluster phase shows that the 6th, 10th, 11th and 12th profiles in the two minor clusters do not behavior as similarly as other eight profiles in the initial main cluster.

The cluster based method, using the statistics in terms of the estimated PA profile from all eight normal profiles, correctly identified the 6th profile as a normal profile. The non-cluster based method, on the other hand, using the statistics in terms of the estimated PA profile from all normal and outlying profiles, misclassified the 6th profile as an outlying profile and the 10th, 11th, 12th profiles as normal profiles.

Monte Carlo Study

A Monte-Carlo study was performed in order to evaluate and compare the proposed cluster based method to the non-cluster based method.

Recall that the POS does not supply information about whether the classification of profiles into the two categories of normal and outlying is correctly specified. Each method's ability to make both correct classifications and incorrect classifications can be evaluated by computing the following performance characteristics: fraction correctly classified (FCC), sensitivity, specificity, false positive (FP) and false negative (FP). The definitions of these terms will be given below. After completing the Phase I analysis, the following classification table (Table 4.1) can be constructed.

Table 4.1: Classification table for Phase I analysis

Classified set / Actual set	Normal profiles	Outlying profiles
Normal profiles	A	B
Outlying profiles	C	D

In Table 4.1, "A" represents the number of normal profiles that are correctly identified as normal profiles and "D" represents the number of outlying profiles that are correctly identified as outlying profiles, respectively, after the Phase I analysis. "B" represents the number of profiles which are from the in-control process but mistakenly classified as outlying profiles while "C" represents the number of profiles which are from the out-of-control process but classified as normal profiles. With this table, the FCC can be defined as

$$FCC = \frac{A+D}{A+B+C+D}.$$

The sensitivity measures the ability of the classification method to identify the normal profiles among the normal profiles and it can be calculated as

$$Sensitivity = \frac{A}{A+B}.$$

The specificity, on the other hand, represents the ability to identify the outlying profiles among the outlying profiles which can be obtained as

$$Specificity = \frac{D}{C+D}.$$

FP is the fraction of actual outlying profiles that are incorrectly classified as normal profiles and FN is the fraction of actual normal profiles that are incorrectly classified as outlying profiles. FP and FN are computed as

$$FP = \frac{C}{A+C},$$

and

$$FN = \frac{B}{B+D}.$$

It is easy to show that all these metrics are bounded by 0 and 1, and that a method will perform well in Phase I analysis by achieving large values for FCC, sensitivity and specificity and small values for FP and FN.

A Monte-Carlo study is used to compare the non-cluster based method and the cluster based method using the performance metrics POS, FCC, sensitivity and specificity, FP, and FN. Also, the ability of each method to accurately estimate the PA parameters will be used to compare the methods. This Monte-Carlo study assumes the in-control profiles are randomly generated from the linear mixed model

$$y_{ij} = \beta_{0i} + \beta_{1i} x_{ij} + \beta_{2i} x_{ij}^{\ 2} + \varepsilon_{ij}, i = 1, 2, ..., m_1, \ j = 1, 2, ..., n. \qquad (16)$$

where

$$\beta_{0i} = \beta_2 \bar{x}^2 + b_{0i},$$

$$\beta_{1i} = \beta_1 - 2\beta_2 \bar{x} + b_{1i},$$

$$\beta_{2i} = \beta_2 \bar{x}^2 + b_{2i}.$$

Here, $\boldsymbol{\beta}^T = (\beta_0, \ \beta_1, \ \beta_2)$ represents the fixed parameters and $\boldsymbol{b}_i^T = (b_{0i}, \ b_{1i}, \ b_{2i})$ represents the random effects. Note, the corresponding PA parameter vector can also be written as $\boldsymbol{\beta}_{PA}^T = \left(\beta_2 \bar{x}^2, \ \beta_1 - 2\beta_2 \bar{x}, \ \beta_2 \bar{x}^2\right)$. Also, m_1 is the number of normal profiles, and $\bar{x} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}}{mn}$. Consequently, the PA profile can be written as

$$y_{PA,ij} = \beta_2 \bar{x}^2 + (\beta_1 - 2\beta_2 \bar{x}) x_{ij} + (\beta_2 \bar{x}^2) x_{ij}{}^2, i = 1, 2, ..., m_1, \ j = 1, 2, ..., n. \quad (17)$$

It is easy to show that the PA profile can be simplified as

$$y_{PA,ij} = \beta_1 x_{ij} + \beta_2 \left(x_{ij} - \bar{x}\right)^2, \ i = 1, 2, ..., m_1, \ j = 1, 2, ..., n. \quad (18)$$

The outlying profiles are also generated from the same form, but with

$$\beta_{0i} = (\beta_2 + shift) \bar{x}^2 + b_{0i},$$

$$\beta_{1i} = \beta_1 - 2(\beta_2 + shift) \bar{x} + b_{1i},$$

and

$$\beta_{2i} = (\beta_2 + shift) \bar{x}^2 + b_{2i},$$

and its corresponding PA profile is

$$\begin{aligned} y_{PA,ij} &= (\beta_2 + shift) \bar{x}^2 + \left[\beta_1 - 2(\beta_2 + shift) \bar{x}\right] x_{ij} + \left[(\beta_2 + shift) \bar{x}^2\right] x_{ij}{}^2, \\ & i = m_1 + 1, \ m_1 + 2, ..., m, \quad j = 1, 2, ..., n. \end{aligned} \quad (19)$$

Also, the above formula can be simplified as

$$y_{PA,ij} = \beta_1 x_{ij} + (\beta_2 + shift)\left(x_{ij} - \bar{x}_i\right)^2, \ i = m_1 + 1, m_1 + 2, ..., m, \ j = 1, 2, ..., n. \quad (20)$$

In above equations, it is assumed that

$$\begin{pmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{pmatrix} \sim MN\left(\mathbf{0} \;, \begin{bmatrix} \sigma_0^2 & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & 0 & \sigma_2^2 \end{bmatrix} \right),$$

$$\boldsymbol{\varepsilon} \sim MN\left(\boldsymbol{0},\quad \sigma^2 \boldsymbol{I}\right),$$

Here, $\sigma_0^2 = \sigma_1^2 = \sigma_2^2 = 0.5,\ \sigma^2 = 1,\ \beta_1 = 3,\ \beta_2 = 2$ and $x_{ij} = j,\ i = 1,2,...,m,\ j = 1,2,...,n.$ It is also assumed that $m_1 = 20,\ m = 30$ and $n = 10.$ The PA parameter vector for the in-control process is set at $\boldsymbol{\beta}_{PA}^T = \left(\beta_2 \bar{x}^2,\quad \beta_1 - 2\beta_2 \bar{x},\quad \beta_2 \bar{x}^2\right) = \left(60,\quad -19,\quad 2\right).$

In this Monte Carlo study, the shift values are set at (0.05, 0.075, 0.1, 0.125, 0.175, 0.2, 0.225, 0.25, 0.275, 0.3). For each value of the shift factor, the performance measures FCC, sensitivity, specificity, FN, FP and POS are averaged over 5,000 replications. The results are presented in Table 4.2.

Table 4.2 shows that when the shift is very small (shift less or equals 0.075), the non-cluster based method has a slightly larger POS than the cluster based method, but the cluster based method has superior performance based on the other criteria. For example, when the shift is 0.075, the cluster based method has FN=0.3922 while the non-cluster based method has FN= 0.4429. Also, the cluster based method has larger value of FCC, specificity and sensitivity with smaller FP when the shift is 0.075. When the shift is greater than 0.075, the cluster based method gives uniformly superior results compare to the non-cluster based method based on all performance criteria. For example, when the shift is equal to 0.2, the cluster based method has the FCC and FN equal 0.8234 and 0.003, respectively, while the non-cluster based method has the FCC and FN are equal to 0.7277 and 0.1176. Also, the POS of the cluster based method is 0.879 while the non-cluster based method is 0.823. Clearly the cluster based method is superior to the non- based method when a large number of very outlying profiles exist in the HDS.

Table 4.2: Average of performance metric based on Monte Carlo study (The top values are the results from the cluster based method and the bolded cells represent the best value)

Shift	FCC	Sensitivity	Specificity	FP	FN	POS
0.05	**0.6674**	**0.9981**	**0.0059**	**0.3324**	**0.3922**	0.0864
	0.667	0.9978	0.0055	0.3326	0.4429	**0.0904**
0.075	**0.6704**	**0.9978**	**0.0156**	**0.3303**	**0.2173**	0.1578
	0.6693	0.9974	0.0132	0.331	0.2814	**0.1594**
0.1	**0.6782**	**0.9978**	**0.0391**	**0.325**	**0.1016**	**0.2876**
	0.6731	0.9955	0.0282	0.328	0.2409	0.2812
0.125	**0.6948**	**0.9983**	**0.0879**	**0.3136**	**0.0381**	**0.4478**
	0.6805	0.9944	0.0528	0.3226	0.1749	0.4314
0.15	**0.7268**	**0.9986**	**0.1832**	**0.2903**	**0.0154**	**0.6396**
	0.6913	0.992	0.0899	0.3145	0.1518	0.5854
0.175	**0.7697**	**0.9992**	**0.3106**	**0.2565**	**0.005**	**0.7812**
	0.706	0.9902	0.1378	0.3033	0.1249	0.7236
0.2	**0.8234**	**0.9993**	**0.4716**	**0.2091**	**0.003**	**0.879**
	0.7227	0.9871	0.194	0.2899	0.1176	0.823
0.225	**0.8766**	**0.9995**	**0.6309**	**0.1559**	**0.0016**	**0.9438**
	0.7432	0.9854	0.2588	0.2733	0.1012	0.8968
0.25	**0.9219**	**0.9994**	**0.767**	**0.1044**	**0.0016**	**0.975**
	0.7627	0.9821	0.3241	0.256	0.0996	0.9336
0.275	**0.9548**	**0.9996**	**0.8654**	**0.0631**	**0.001**	**0.9896**
	0.7855	0.9806	0.3953	0.2357	0.0896	0.9698
0.3	**0.9749**	**0.9995**	**0.9256**	**0.0359**	**0.0011**	**0.9956**
	0.8052	0.9775	0.4604	0.2163	0.089	0.9806

The average estimated PA parameters were also calculated for each shift factor. Table 4.3 lists the results for both the cluster based method and the non-cluster based method.

Table 4.3: Average of PA parameter estimates based on a Monte Carlo study (*top values correspond to the cluster-based method; the bolded cells represent estimates closer to the true parameter values of* $\boldsymbol{\beta}_{PA}^{T} = (60, \quad -19, \quad 2)$)

Shift	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$se(\hat{\beta}_0)$	$se(\hat{\beta}_1)$	$se(\hat{\beta}_2)$
	60.9942	**-19.1802**	**2.0149**	0.00359	0.00226	0.00184
0.05	61.0026	-19.1814	2.0190	0.00354	0.00221	0.00182
	61.241	**-19.2674**	**2.0227**	0.00365	0.00228	0.00184
0.075	61.2574	-19.2736	2.0234	0.00353	0.00224	0.00183
	61.4596	**-19.3482**	**2.0308**	0.00407	0.00236	0.2812
0.1	61.5068	-19.3648	2.0357	0.00354	0.00221	0.00182
	61.6299	**-19.4085**	**2.036**	0.0048	0.00259	0.00187
0.125	61.7615	-19.4569	2.0401	0.00353	0.00224	0.00183
	61.6991	**-19.4370**	**2.0384**	0.00664	0.00311	0.00194
0.15	62.0110	-19.5481	2.0523	0.00354	0.00221	0.00182
	61.6867	**-19.4296**	**2.0372**	0.00879	0.00381	0.00199
0.175	62.2657	-19.6402	2.0568	0.00353	0.00224	0.00183
	61.5422	**-19.3761**	**2.0338**	0.01100	0.00450	0.00206
0.2	62.5151	-19.7314	2.0690	0.00354	0.00221	0.00182
	61.3216	**-19.2955**	**2.0262**	0.01187	0.00484	0.00213
0.225	62.7699	-19.8236	2.0734	0.00353	0.00224	0.00183
	61.0702	**-19.2083**	**2.0176**	0.01159	0.00474	0.00219
0.25	63.0193	-19.9148	2.0857	0.00354	0.00221	0.00182
	60.8742	**-19.1325**	**2.0108**	0.01024	0.00433	0.00221
0.275	63.2740	-20.0069	2.0901	0.00353	0.00224	0.00183
	60.7290	**-19.0814**	**2.0081**	0.00879	0.00391	0.00222
0.3	63.5235	-20.0981	2.1023	0.00354	0.00221	0.00182

Table 4.3 shows that both estimators have bias in parameter estimation compared to the true in-control PA parameters $\boldsymbol{\beta}_{PA}^{T} = (60, \quad -19, \quad 2)$ when there are large numbers of outlying profiles. However, the estimated PA parameters from the cluster based method have smaller bias than that from the non-cluster based method, especially when the shift is large. When the shift is small, both methods provide estimators with smaller bias. However, for the non-cluster based method, the bias is monotone increasing as the shift increases. For example, when the shift is 0.05, the non-cluster based method has

estimated PA parameters of $\hat{\boldsymbol{\beta}}_{PA}^{T} = (61.002, \quad -19.181, \quad 2.019)$, while when the shift is equal to 0.3,the non-cluster based method has estimated PA parameters of $\hat{\boldsymbol{\beta}}_{PA}^{T} = (63.524, \quad -20.098, \quad 2.102)$. The cluster based method, on the other hand, provides estimated PA parameters with smaller bias when the shift is very small or relatively large. For example, in Table 4.3, the cluster based method provides the estimate with the smallest bias when the shift equals 0.3 and the second smallest bias when the shift is equal to 0.05. In other words, the bias of the estimate from the cluster based method is increasing first when the shift increases and then is decreasing when the shift is larger than about 0.2.

These results are consistent with the result in Table 4.2 in that the cluster based method is superior to the non-cluster based method when a larger number of very outlying profiles are present in the HDS. Table 4.3 also shows that the estimates of both methods have very small standard errors based on 5,000 simulations. The non-cluster based method seems to have smaller standard errors. However, this result does not affect our final conclusion that the cluster based method is superior to the non-cluster based method since the differences between the standard errors are negligible compared to the size of the bias.

Conclusion and Future Work

The proposed profile monitoring methodology is a robust profile monitoring methodology for Phase I analysis. The goal is to improve upon the existing methods which can be distorted by the profiles from the out-of -control process. Specifically, the example shows that the cluster based method determined correctly that at least one outlying profile was contained in the HDS. In addition, the cluster based method provided accurate estimates of the PA parameters and also identified the in-control and out-of-control process correctly. The non-cluster based method, on the other hand, detected the presence of at least one outlying profile in the HDS but misclassified one normal profile as an outlying profile,

misclassified three outlying profiles as normal profiles, and provided biased estimates of the PA parameters.

The Monte Carlo study shows that the proposed method works uniformly better when there is a moderate or a large shift in the process. The proposed method not only had a larger POS, but also had a better performance regarding correct classification. Additionally, the proposed method gave more accurate estimates for the PA parameters.

The proposed method in the paper is illustrated for the case where the profiles are can be modeled using parametric regression techniques. However, as is often the case, nonparametric methods may be required to adequately model the profiles. The authors are currently studying applying the cluster based method when nonparametric regression methods are appropriate.

The proposed algorithm was programmed using R and the program is available from the authors upon request. The algorithm is surprisingly fast. For example, the case study required only a second to complete the method using a moderately equipped PC.

References

Demidenko, E. (2004), *Mixed Models : Theory and Applications*, Hoboken, N.J.: Wiley-Interscience.

Ding, Y., Zeng, L., and Zhou, S. (2006), "Phase I Analysis for Monitoring Nonlinear Profiles in Manufacturing Processes," *Journal of Quality Technology*, 38, 199-216.

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

Hawkins, D. M. and Merriam, D. F. (1974). "Zonation of Multivariate Sequences of Digitized Geologic Data," *Mathematical Geology*, 6:263–269.

Hung, Y.-C., Tsai, W.-C., Yang, S.-F., Chuang, S.-C., and Tseng, Y.-K. (2012), "Nonparametric Profile Monitoring in Multi-Dimensional Data Spaces," *Journal of Process Control*, 22, 397-403.

Jensen, W. A., Birch, J. B., and Woodall, W. H. (2008), "Monitoring Correlation within Linear Profiles Using Mixed Models," *Journal of Quality Technology*, 40, 167-183.

Jensen, W. A., and Birch, J. B. (2009), "Profile Monitoring Via Nonlinear Mixed Models," *Journal of Quality Technology*, 41, 18-34.

Jin, J. H., and Shi, J. J. (1999), "Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets," *Technometrics*, 41, 327-339.

Jobe, J. M., and Pokojovy, M. (2009), "A Multistep, Cluster-Based Multivariate Chart for Retrospective Monitoring of Individuals," *Journal of Quality Technology*, 41, 323-339.

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

Kim, K., Mahmoud, M. A., and Woodall, W. H. (2003), "On the Monitoring of Linear Profiles," *Journal of Quality Technology*, 35, 317-328.

Laird, N. M., and Ware, J. H. (1982), "Random-Effects Models for Longitudinal Data," *Biometrics*, 38, 963-974.

Lawrence, David E. (2003): "Cluster-Based Bounded Influence Regression," *Unpublished Ph.D. dissertation. Department of Statistics, Virginia Polytechnic Institute and State University Blacksburg, Virginia.*

Mahmoud, M. A., and Woodall, W. H. (2004), "Phase I Analysis of Linear Profiles with Calibration Applications," *Technometrics*, 46, 380-391.

Qiu, P., Zou, C., and Wang, Z. (2010), "Nonparametric Profile Monitoring by Mixed Effects Modeling," *Technometrics*, 52, 265-277.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003), *Semiparametric Regression*, Cambridge; New York: Cambridge University Press.

Schabenberger, O., and Pierce, F. J. (2002), *Contemporary Statistical Models for the Plant and Soil Sciences*, Boca Raton: CRC Press.

Seber, G. A. F., and Wild, C. J. (2003), *Nonlinear Regression*, Hoboken, N.J.: Wiley-Interscience.

Sullivan, J. H., and Woodall, W. H. (1996), "A Comparison of Multivariate Control Charts for Individual Observations," *Journal of Quality Technology*, 28, 398-408.

Walker, E., and Wright, S. P. (2002), "Comparing Curves Using Additive Models," *Journal of Quality Technology*, 34, 118-129.

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

Williams, J. D., Woodall, W. H., and Birch, J. B. (2007b), "Statistical Monitoring of Nonlinear Product and Process Quality Profiles," *Quality and Reliability Engineering International*, 23, 925-941.