

Construction of An Efficient Multivariate Dynamic Screening System

Jun Li^a and Peihua Qiu^b

^aDepartment of Statistics, University of California at Riverside

^bDepartment of Biostatistics, University of Florida

Abstract

Recently, multivariate dynamic screening system (MDySS) has been proposed in the literature for monitoring processes whose in-control distributions change over time. MDySS has broad applications, ranging from monitoring of dynamic engineering systems, such as nuclear reactors, airplanes and other durable goods, to early disease detection. Conventional construction of MDySS is appropriate only in cases when sampling rate of observations is the same in Phase-I and Phase-II process monitoring, and the process observations are independent. In practice, these assumptions are rarely valid. In this paper, we propose a new construction of the MDySS method, which is shown to be more reliable than the conventional construction in various cases. The new construction is demonstrated using a real-data from the SHARe Framingham Heart Study of the National Heart, Lung and Blood Institute.

Key Words: Correlation; Dynamic screening; Efficiency; Multivariate distribution; Nonparametric procedures; Unequal sampling intervals.

1 Introduction

Statistical process control (SPC) charts provide a major tool for monitoring production processes.¹⁻³ A basic assumption to use conventional SPC charts is that the process distribution is unchanged when the process is in-control (IC). This assumption is often invalid when monitoring dynamic systems, such as nuclear reactors, airplanes and other durable goods, in which distributions of the quality variables would change over time even when they are IC. This paper aims to develop an effective method for process monitoring of such dynamic systems.

In the SPC literature, four types of control charts have been proposed, including the Shewhart,⁴ CUSUM,⁵ EWMA,⁶ and change-point detection⁷ charts. Numerous generalizations and modifications of these four types of control charts have been suggested (see, for example, Chapters 3-7

of Qiu³). Most of these charts and their variants are constructed based on the assumption that the process observations are independently and identically distributed as a normal distribution. There are some existing methods in the area of nonparametric (or distribution-free) SPC to address cases when the normality assumption is invalid.^{8–14} There are also some methods to address cases when process observations are correlated and follow some parametric time series models.^{15–17} The assumption about identical distribution was addressed recently by Qiu and Xiang¹⁸ where we proposed a so-called dynamic screen system (DySS) to handle cases when the IC process distribution changes over time. Its multivariate version was discussed in Qiu and Xiang¹⁹ and the related control scheme was called multivariate DySS (MDySS). However, both DySS and MDySS are mainly designed for cases when the process observations are independent and the sampling rates in Phase I and Phase II SPC are the same. In cases when these assumptions are invalid, these approaches may not be reliable.

In this paper, we propose a novel construction of the MDySS method to handle cases when the process observations are not independent and identically normally distributed. The new construction decorrelates the observed data and makes their IC distribution closer to normal as well before process monitoring. As such, the modified MDySS method would be more efficient than the original one, especially when the true process distribution is very skewed and process observations are highly correlated. The modified method will be described in detail in Section 2. Some numerical results are presented in Section 3. A real-data application is discussed in Section 4. Some remarks conclude the article in Section 5.

2 Methodology

Similar to many SPC procedures, our proposed MDySS method consists of two steps. In the first step, a flexible multivariate nonparametric model is used to estimate the regular multivariate longitudinal pattern from an IC dataset. To this end, the recent model fitting procedure proposed by Xiang, Qiu and Pu²⁰ is used, and it is briefly described in Subsection 2.1. Based on the fitted multivariate nonparametric model, we develop a new MDySS procedure for online monitoring of

the multivariate longitudinal pattern of a new subject, which is described in detail in Subsection 2.2.

2.1 Estimation of regular multivariate longitudinal pattern

Assume that we have m well-functioning subjects in an IC dataset. For each subject, q performance variables that we are interested in monitoring are measured over time. Let $\mathbf{y}_i(t_{ij}) = (y_{i1}(t_{ij}), y_{i2}(t_{ij}), \dots, y_{iq}(t_{ij}))^T$ denote the q -dimensional observation collected from the i -th subject at the j -th time point $t_{ij} \in [0, T]$. We use the following multivariate nonparametric longitudinal model to describe such data:

$$\mathbf{y}_i(t_{ij}) = \boldsymbol{\mu}(t_{ij}) + \boldsymbol{\epsilon}_i(t_{ij}), \quad i = 1, \dots, m, j = 1, \dots, n_i, \quad (1)$$

where $\boldsymbol{\mu}(t_{ij}) = (\mu_1(t_{ij}), \mu_2(t_{ij}), \dots, \mu_q(t_{ij}))^T$ is the population mean vector of $\mathbf{y}_i(t_{ij})$, and $\boldsymbol{\epsilon}_i(t_{ij}) = (\epsilon_{i1}(t_{ij}), \epsilon_{i2}(t_{ij}), \dots, \epsilon_{iq}(t_{ij}))^T$ is the q -dimensional error term with the covariance matrix function $\Sigma(s, t) = \text{Cov}(\boldsymbol{\epsilon}_i(s), \boldsymbol{\epsilon}_i(t))$ for any $s, t \in [0, T]$. By the estimation procedure in Xiang, Qiu and Pu,²⁰ we can obtain estimates of $\boldsymbol{\mu}(t)$ and $\Sigma(s, t)$ by the algorithm below.

Step 1. Use the local linear kernel smoothing procedure to obtain an initial estimate of $\mu_k(\cdot)$ ($k = 1, \dots, q$). More specifically, define

$$X_i = \begin{pmatrix} 1 & (t_{i1} - t) \\ \vdots & \vdots \\ 1 & (t_{in_i} - t) \end{pmatrix}_{n_i \times 2},$$

and $K_{ih} = \text{diag}(K((t_{i1} - t)/h), \dots, K((t_{in_i} - t)/h))/h$, where $K(\cdot)$ is a kernel function and $h > 0$ is a bandwidth. Let $J_i = \text{diag}(I_{\{|t_{i1} - t| \leq h\}}, \dots, I_{\{|t_{in_i} - t| \leq h\}})$, where $I_{\{A\}}$ is the indicator function and takes the value of 1 if A is true and 0 otherwise. Then the initial local linear

kernel estimator of $\mu_k(t)$, $k = 1, \dots, q$, is given by

$$\tilde{\mu}_k(t) = \mathbf{e}_1^T \left(\sum_{i=1}^m X_i^T W_i X_i \right)^{-1} \left(\sum_{i=1}^m X_i^T W_i \mathbf{y}_{ik} \right),$$

where $\mathbf{e}_1 = (1, 0)^T$, $W_i = K_{ih}^{1/2} (J_i I_{n_i} J_i)^{-1} K_{ih}^{1/2}$, I_{n_i} is the n_i -dimensional identity matrix, and $\mathbf{y}_{ik} = (y_{ik}(t_{i1}), y_{ik}(t_{i2}), \dots, y_{ik}(t_{in_i}))^T$.

Step 2. Based on the initial estimate $\tilde{\mu}_k(t)$, calculate the residuals

$$\tilde{\epsilon}_{ik}(t_{ij}) = y_{ik}(t_{ij}) - \tilde{\mu}_k(t_{ij}), \quad i = 1, \dots, m, j = 1, \dots, n_i, k = 1, \dots, q.$$

Let $\tilde{\epsilon}_i(t_{ij}) = (\tilde{\epsilon}_{i1}(t_{ij}), \tilde{\epsilon}_{i2}(t_{ij}), \dots, \tilde{\epsilon}_{iq}(t_{ij}))^T$. An initial estimator of $\Sigma(s, t)$ for $s \leq t$ is obtained by using the following kernel estimator,

$$\tilde{\Sigma}(s, t) = \frac{\sum_{i=1}^m \sum_{j_1, j_2=1, \dots, n_i, j_1 \leq j_2} \tilde{\epsilon}_i(t_{ij_1}) \tilde{\epsilon}_i(t_{ij_2})^T I_{\{t_{ij_2} - t_{ij_1} = t - s\}} K\left(\frac{t_{ij_1} - s}{h}\right)}{\sum_{i=1}^m \sum_{j_1, j_2=1, \dots, n_i, j_1 \leq j_2} I_{\{t_{ij_2} - t_{ij_1} = t - s\}} K\left(\frac{t_{ij_1} - s}{h}\right)}.$$

Since $\Sigma(s, t) = \Sigma(t, s)^T$, estimates of $\Sigma(s, t)$ for $s > t$ can be defined accordingly.

Step 3. Based on $\tilde{\Sigma}(s, t)$ from Step 2, we obtain an updated estimator of $\boldsymbol{\mu}(\cdot)$. More specifically, denote a diagonal matrix with the $[j + (k-1)J]$ -th diagonal element being a_{jk} by $\text{diag}\{a_{jk}, j = 1, \dots, J, k = 1, \dots, q\}$. Let $H = (h_1, \dots, h_q)^T$ be the bandwidth vector with h_k being the bandwidth for the k -th component ($k = 1, \dots, q$). Then, define

$$K_{iH} = \text{diag}\{K((t_{ij} - t)/h_k), j = 1, \dots, n_i, k = 1, \dots, q\},$$

$$J_{iH} = \text{diag}\{I_{\{|t_{ij} - t| \leq h_k\}}, j = 1, \dots, n_i, k = 1, \dots, q\}.$$

Let $\mathbf{Y}_i = (y_{i1}(t_{i1}), \dots, y_{i1}(t_{in_i}), y_{i2}(t_{i1}), \dots, y_{i2}(t_{in_i}), \dots, y_{iq}(t_{i1}), \dots, y_{iq}(t_{in_i}))^T$ and $V_i = \text{Cov}(\mathbf{Y}_i)$.

The estimator of V_i can then be obtained from $\tilde{\Sigma}(s, t)$ as

$$\hat{V}_i = \begin{pmatrix} \tilde{\sigma}_{11}(t_{i1}, t_{i1}) & \cdots & \tilde{\sigma}_{11}(t_{i1}, t_{in_i}) & \cdots & \tilde{\sigma}_{1q}(t_{i1}, t_{i1}) & \cdots & \tilde{\sigma}_{1q}(t_{i1}, t_{in_i}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\sigma}_{11}(t_{in_i}, t_{i1}) & \cdots & \tilde{\sigma}_{11}(t_{in_i}, t_{in_i}) & \cdots & \tilde{\sigma}_{1q}(t_{in_i}, t_{i1}) & \cdots & \tilde{\sigma}_{1q}(t_{in_i}, t_{in_i}) \\ \tilde{\sigma}_{21}(t_{i1}, t_{i1}) & \cdots & \tilde{\sigma}_{21}(t_{i1}, t_{in_i}) & \cdots & \tilde{\sigma}_{2q}(t_{i1}, t_{i1}) & \cdots & \tilde{\sigma}_{2q}(t_{i1}, t_{in_i}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\sigma}_{21}(t_{in_i}, t_{i1}) & \cdots & \tilde{\sigma}_{21}(t_{in_i}, t_{in_i}) & \cdots & \tilde{\sigma}_{2q}(t_{in_i}, t_{i1}) & \cdots & \tilde{\sigma}_{2q}(t_{in_i}, t_{in_i}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\sigma}_{q1}(t_{i1}, t_{i1}) & \cdots & \tilde{\sigma}_{q1}(t_{i1}, t_{in_i}) & \cdots & \tilde{\sigma}_{qq}(t_{i1}, t_{i1}) & \cdots & \tilde{\sigma}_{qq}(t_{i1}, t_{in_i}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\sigma}_{q1}(t_{in_i}, t_{i1}) & \cdots & \tilde{\sigma}_{q1}(t_{in_i}, t_{in_i}) & \cdots & \tilde{\sigma}_{qq}(t_{in_i}, t_{i1}) & \cdots & \tilde{\sigma}_{qq}(t_{in_i}, t_{in_i}) \end{pmatrix}$$

where $\tilde{\sigma}_{l_1, l_2}(s, t)$ is the (l_1, l_2) -th element of $\tilde{\Sigma}(s, t)$. Define

$$W_{iH} = K_{iH}^{1/2} (J_{iH} \hat{V}_i J_{iH})^{-1} K_{iH}^{1/2}.$$

Then, the updated and final estimate of $\boldsymbol{\mu}(t)$ is

$$\hat{\boldsymbol{\mu}}(t) = \left\{ \left[\sum_{i=1}^m (I_q \otimes X_i)^T W_{iH} (I_q \otimes X_i) \right]^{-1} \left[\sum_{i=1}^m (I_q \otimes X_i)^T W_{iH} \mathbf{Y}_i \right] \right\}^T (I_q \otimes \mathbf{e}_1), \quad (2)$$

where I_q is the q -dimensional identity matrix, and \otimes denotes the Kronecker product.

Step 4. Based on the updated estimate of $\boldsymbol{\mu}(\cdot)$ from Step 3, we can update the residuals by

$$\hat{\epsilon}_{ik}(t_{ij}) = y_{ik}(t_{ij}) - \hat{\mu}_k(t_{ij}), \quad i = 1, \dots, m, \quad j = 1, \dots, n_i, \quad k = 1, \dots, q,$$

and obtain the updated estimate of $\Sigma(s, t)$ as

$$\hat{\Sigma}(s, t) = \frac{\sum_{i=1}^m \sum_{j_1, j_2=1, \dots, n_i, j_1 \leq j_2} \hat{\epsilon}_i(t_{ij_1}) \hat{\epsilon}_i(t_{ij_2})^T I_{\{t_{ij_2} - t_{ij_1} = t - s\}} K\left(\frac{t_{ij_1} - s}{h}\right)}{\sum_{i=1}^m \sum_{j_1, j_2=1, \dots, n_i, j_1 \leq j_2} I_{\{t_{ij_2} - t_{ij_1} = t - s\}} K\left(\frac{t_{ij_1} - s}{h}\right)},$$

where $\widehat{\boldsymbol{\epsilon}}_i(t_{ij}) = (\widehat{\epsilon}_{i1}(t_{ij}), \widehat{\epsilon}_{i2}(t_{ij}), \dots, \widehat{\epsilon}_{iq}(t_{ij}))^T$.

In all our numerical examples presented later, we use $K(t) = 0.75(1 - t^2)I_{\{|t| \leq 1\}}$, and all bandwidths are selected separately by the conventional cross-validation method.

2.2 Online monitoring of multivariate longitudinal data

Now, we want to online monitor the performance of a new subject. Assume that her/his performance variables are observed at $\{t_j^*, j = 1, 2, \dots\}$ and the observations are $\{\mathbf{y}(t_j^*), j = 1, 2, \dots\}$. Our major goal is to decide at each time point t_j^* whether the new subject's observations follow the regular multivariate longitudinal pattern, based on all available observations collected so far. If not, then a signal should be given as early as possible. So, when the new subject is IC, $\{\mathbf{y}(t_j^*), j = 1, 2, \dots\}$ should follow model (1), namely,

$$\mathbf{y}(t_j^*) = \boldsymbol{\mu}(t_j^*) + \boldsymbol{\epsilon}(t_j^*), \quad j = 1, 2, \dots \quad (3)$$

There are many multivariate SPC procedures in the existing literature²¹. However, most of these SPC procedures require that the observations are uncorrelated over time. Therefore, they cannot be applied directly to the current problem because longitudinal association is allowed in (3). The MDySS method in Qiu and Xiang¹⁹ ignores the longitudinal association in its construction. In this paper, we try to sequentially decorrelate the observed data $\{\mathbf{y}(t_j^*)\}$ first each time when a new observation is obtained, and then apply a standard multivariate SPC procedure to the decorrelated data. The sequential data decorrelation procedure is described below.

2.2.1 Sequential decorrelation procedure

To facilitate the exposition, we treat $\boldsymbol{\mu}(t)$ and $\Sigma(s, t)$ as known. In reality, they should be replaced by their respective estimates $\widehat{\boldsymbol{\mu}}(t)$ and $\widehat{\Sigma}(s, t)$ described in the previous section. We start from the first observation $\mathbf{y}(t_1^*)$. We first compute $\boldsymbol{\epsilon}(t_1^*) = \mathbf{y}(t_1^*) - \boldsymbol{\mu}(t_1^*)$. Then, the covariance matrix of

$\boldsymbol{\epsilon}(t_1^*)$ is $A_{11} = \Sigma_{11} = \Sigma(t_1^*, t_1^*)$. Therefore, we can define its standardized vector by

$$\boldsymbol{e}^*(t_1^*) = \Sigma_{11}^{-1/2} \boldsymbol{\epsilon}(t_1^*).$$

For the second observation $\boldsymbol{y}(t_2^*)$, we can also calculate $\boldsymbol{\epsilon}(t_2^*) = \boldsymbol{y}(t_2^*) - \boldsymbol{\mu}(t_2^*)$. Now, $\boldsymbol{\epsilon}(t_2^*)$ is correlated with $\boldsymbol{\epsilon}(t_1^*)$, and the covariance matrix of $\boldsymbol{\epsilon}_2 = (\boldsymbol{\epsilon}(t_1^*)^T, \boldsymbol{\epsilon}(t_2^*)^T)^T$ is given by $A_{22} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}$, where $\Sigma_{12} = \Sigma(t_1^*, t_2^*)$, $\Sigma_{22} = \Sigma(t_2^*, t_2^*)$. If we define $\Phi_2 = \begin{pmatrix} I_q & \mathbf{0} \\ -\Sigma_{12}^T \Sigma_{11}^{-1} & I_q \end{pmatrix}$ and $D_2 = \begin{pmatrix} \Sigma_{11} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22.1} \end{pmatrix} = \text{diag}\{\Sigma_{11}, \Sigma_{22.1}\}$, where $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12}$, we have

$$\Phi_2 A_{22} \Phi_2^T = D_2.$$

This motivates us to define $\boldsymbol{e}_2 = \Phi_2 \boldsymbol{\epsilon}_2 = (\boldsymbol{e}(t_1^*)^T, \boldsymbol{e}(t_2^*)^T)^T$, where

$$\begin{aligned} \boldsymbol{e}(t_1^*) &= \boldsymbol{\epsilon}(t_1^*), \\ \boldsymbol{e}(t_2^*) &= -\Sigma_{12}^T \Sigma_{11}^{-1} \boldsymbol{\epsilon}(t_1^*) + \boldsymbol{\epsilon}(t_2^*). \end{aligned}$$

Since $\text{Cov}(\boldsymbol{e}_2) = D_2$, $\boldsymbol{e}(t_1^*)$ and $\boldsymbol{e}(t_2^*)$ are uncorrelated. Noticing that the standardized vector defined at the first observation time is $\boldsymbol{e}^*(t_1^*) = \Sigma_{11}^{-1/2} \boldsymbol{e}(t_1^*)$, we can define the standardized vector of the second observation by

$$\boldsymbol{e}^*(t_2^*) = \Sigma_{22.1}^{-1/2} \boldsymbol{e}(t_2^*) = \Sigma_{22.1}^{-1/2} (-\Sigma_{12}^T \Sigma_{11}^{-1} \boldsymbol{\epsilon}(t_1^*) + \boldsymbol{\epsilon}(t_2^*)).$$

Then $\boldsymbol{e}^*(t_1^*)$ and $\boldsymbol{e}^*(t_2^*)$ are uncorrelated and both have covariance matrices equal to I_q .

Similarly, for the third observation $\boldsymbol{y}(t_3^*)$, we calculate $\boldsymbol{\epsilon}(t_3^*) = \boldsymbol{y}(t_3^*) - \boldsymbol{\mu}(t_3^*)$. The covariance matrix of $\boldsymbol{\epsilon}_3 = (\boldsymbol{\epsilon}_2^T, \boldsymbol{\epsilon}(t_3^*)^T)^T$ is given by $A_{33} = \begin{pmatrix} A_{22} & \Sigma_{23} \\ \Sigma_{23}^T & \Sigma_{33} \end{pmatrix}$, where $\Sigma_{23} = \text{Cov}(\boldsymbol{\epsilon}_2, \boldsymbol{\epsilon}(t_3^*)) =$

$(\Sigma(t_1^*, t_3^*)^T, \Sigma(t_2^*, t_3^*)^T)^T$, $\Sigma_{33} = \text{Cov}(\boldsymbol{\epsilon}(t_3^*)) = \Sigma(t_3^*, t_3^*)$. It can be shown that

$$\Phi_3 A_{33} \Phi_3^T = D_3,$$

where $\Phi_3 = \begin{pmatrix} \Phi_2 & \mathbf{0} \\ -\Sigma_{23}^T A_{22}^{-1} & I_q \end{pmatrix}$, and $D_3 = \text{diag}\{\Sigma_{11}, \Sigma_{22 \cdot 1}, \Sigma_{33 \cdot 2}\}$, $\Sigma_{33 \cdot 2} = \Sigma_{33} - \Sigma_{23}^T A_{22}^{-1} \Sigma_{23}$. Therefore, if we define

$$\boldsymbol{e}(t_3^*) = -\Sigma_{23}^T A_{22}^{-1} \boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}(t_3^*),$$

then $\boldsymbol{e}_3 = (\boldsymbol{e}_2^T, \boldsymbol{e}(t_3^*)^T)^T = \Phi_3 \boldsymbol{\epsilon}_3$ and $\text{Cov}(\boldsymbol{e}_3) = D_3$, which implies that $\boldsymbol{e}(t_3^*)$ is uncorrelated with $\boldsymbol{e}(t_1^*)$ and $\boldsymbol{e}(t_2^*)$. The standardized vector of the third observation is then defined as

$$\boldsymbol{e}^*(t_3^*) = \Sigma_{33 \cdot 2}^{-1/2} \boldsymbol{e}(t_3^*) = \Sigma_{33 \cdot 2}^{-1/2} (-\Sigma_{23}^T A_{22}^{-1} \boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}(t_3^*)),$$

which is uncorrelated with $\boldsymbol{e}^*(t_1^*)$ and $\boldsymbol{e}^*(t_2^*)$ and has covariance matrix equal to I_q .

Following the same procedure we can define the standardized vector sequentially after a new observation is collected for the new subject. More specifically, at the j -th observation time, we calculate $\boldsymbol{\epsilon}(t_j^*) = \boldsymbol{y}(t_j^*) - \boldsymbol{\mu}(t_j^*)$. Then, the covariance matrix of $\boldsymbol{\epsilon}_j = (\boldsymbol{\epsilon}_{j-1}^T, \boldsymbol{\epsilon}(t_j^*)^T)^T$ is $A_{jj} = \begin{pmatrix} A_{j-1, j-1} & \Sigma_{j-1, j} \\ \Sigma_{j-1, j}^T & \Sigma_{jj} \end{pmatrix}$, where $\boldsymbol{\epsilon}_{j-1}$ is the residual vector from the first $(j-1)$ -th observations, $A_{j-1, j-1} = \text{Cov}(\boldsymbol{\epsilon}_{j-1})$, $\Sigma_{j-1, j} = \text{Cov}(\boldsymbol{\epsilon}_{j-1}, \boldsymbol{\epsilon}(t_j^*)) = (\Sigma(t_1^*, t_j^*)^T, \dots, \Sigma(t_{j-1}^*, t_j^*)^T)^T$, $\Sigma_{jj} = \text{Cov}(\boldsymbol{\epsilon}(t_j^*)) = \Sigma(t_j^*, t_j^*)$. It can be shown that

$$\Phi_j A_{jj} \Phi_j^T = D_j,$$

where $\Phi_j = \begin{pmatrix} \Phi_{j-1} & \mathbf{0} \\ -\Sigma_{j-1, j}^T A_{j-1, j-1}^{-1} & I_q \end{pmatrix}$, $D_j = \text{diag}\{\Sigma_{11}, \Sigma_{22 \cdot 1}, \dots, \Sigma_{jj \cdot j-1}\}$ with $\Sigma_{jj \cdot j-1} = \Sigma_{jj} - \Sigma_{j-1, j}^T A_{j-1, j-1}^{-1} \Sigma_{j-1, j}$. Therefore, if we define

$$\boldsymbol{e}(t_j^*) = -\Sigma_{j-1, j}^T A_{j-1, j-1}^{-1} \boldsymbol{\epsilon}_{j-1} + \boldsymbol{\epsilon}(t_j^*),$$

$\mathbf{e}_j = (\mathbf{e}(t_1^*)^T, \dots, \mathbf{e}(t_j^*)^T)^T = (\mathbf{e}_{j-1}^T, \mathbf{e}(t_j^*)^T)^T = \begin{pmatrix} \Phi_{j-1} & \mathbf{0} \\ -\Sigma_{j-1,j}^T A_{j-1,j-1}^{-1} & I_q \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_{j-1} \\ \boldsymbol{\epsilon}(t_j^*) \end{pmatrix} = \Phi_j \boldsymbol{\epsilon}_j$ and $\text{Cov}(\mathbf{e}_j) = D_j$. This implies that, if we define the standardized vector of the j -th observation by

$$\mathbf{e}^*(t_j^*) = \Sigma_{jj \cdot j-1}^{-1/2} \mathbf{e}(t_j^*) = \Sigma_{jj \cdot j-1}^{-1/2} \left(-\Sigma_{j-1,j}^T A_{j-1,j-1}^{-1} \boldsymbol{\epsilon}_{j-1} + \boldsymbol{\epsilon}(t_j^*) \right),$$

$\mathbf{e}^*(t_j^*)$ is uncorrelated with $\mathbf{e}^*(t_1^*), \dots, \mathbf{e}^*(t_{j-1}^*)$ and has covariance matrix equal to I_q . So, by the above sequential data transformation, we transform the original data $\{\mathbf{y}(t_1^*), \dots, \mathbf{y}(t_j^*)\}$, which are longitudinally correlated, to the sequence of uncorrelated data $\{\mathbf{e}^*(t_1^*), \dots, \mathbf{e}^*(t_j^*)\}$, and each vector $\mathbf{e}(t_{j'}^*)$, for $j' = 1, \dots, j$, has mean $\mathbf{0}$ and covariance matrix I_q .

Remark 1 *In the above sequential decorrelation procedure, at the $(j+1)$ -th observation, we need to calculate A_{jj}^{-1} . To avoid the difficulty associated with finding the inverse of high-dimensional matrices, we can use the following recursive formula:*

$$A_{jj}^{-1} = \begin{pmatrix} A_{j-1,j-1}^{-1} + A_{j-1,j-1}^{-1} \Sigma_{j-1,j} \Sigma_{jj \cdot j-1}^{-1} \Sigma_{j-1,j}^T A_{j-1,j-1}^{-1}, & -A_{j-1,j-1}^{-1} \Sigma_{j-1,j} \Sigma_{jj \cdot j-1}^{-1} \\ -\Sigma_{jj \cdot j-1}^{-1} \Sigma_{j-1,j}^T A_{j-1,j-1}^{-1}, & \Sigma_{jj \cdot j-1}^{-1} \end{pmatrix}.$$

This result simplifies the computation greatly.

2.2.2 Online monitoring procedure

Because $\mathbf{e}^*(t_j^*)$ is a linear combination of observations $\mathbf{y}(t_1^*), \mathbf{y}(t_2^*), \dots, \mathbf{y}(t_j^*)$, its distribution would be asymptotically normal. Therefore, the series $\{\mathbf{e}^*(t_j^*), j = 1, 2, \dots\}$ would be asymptotically i.i.d. with the common distribution to be the standard multivariate normal distribution $N_p(\mathbf{0}, I_q)$. In such cases, many standard multivariate control charts can be applied. As an example, in this paper we use the multivariate EWMA chart proposed by Lowry et al.²² to demonstrate our proposed MDySS method. More specifically, we define the charting statistic to be

$$\mathbf{S}_j = \lambda \mathbf{e}^*(t_j^*) + (1 - \lambda) \mathbf{S}_{j-1}, \text{ for } j \geq 1, \quad (4)$$

where $\mathbf{S}_0 = \mathbf{0}$ and $\lambda \in (0, 1]$ is a weighting parameter, and the chart gives an signal when

$$\frac{2 - \lambda}{\lambda} \mathbf{S}_j^T \mathbf{S}_j > l,$$

where $l > 0$ is a control limit.

2.2.3 Determining the control limit

Because the observation times may not be equally spaced in most longitudinal data studies, following Qiu and Xiang¹⁸, we use average time to the signal (*ATS*) when evaluating the performance of control charts in this paper. In particular, when calculating *ATS*, the time to signal is expressed in terms of the basic time unit ω , where ω is the largest time unit that all unequally spaced observation times are its integer multiples. We also note that the IC data usually have a fixed time frame. Without loss of generality, we assume that the domain of the observation times $\{t_{ij}\}$ for the IC data is $[0, 1]$. Therefore, the mean vector function and the covariance matrix function estimated from the IC data are only appropriate for use within the domain $[0, 1]$. As a consequence, when using these estimates in our proposed MDySS procedure to monitor the observations from a new subject, even if our MDySS procedure has not signaled when we reach the last observation of the individual within the domain $[0, 1]$, we cannot continue to monitor. When this happens, we can only know that the actual time to the signal is greater than $1/\omega$, and we would not know the exact time to the signal. To make the calculation of *ATS* still feasible in this situation, we let the time to the signal to be $1/\omega$ when our MDySS procedure has not signaled at the last observation of the individual subject within the domain $[0, 1]$. In other words, the time to the signal is right truncated at $1/\omega$.

Based on the above method to calculate *ATS*, we next describe how to determine the control limit l such that the IC *ATS* (denoted by ATS_0) of our MDySS procedure in (4) is controlled at a desired level. From Section 2.2.1, we know that $\{e^*(t_j^*), j = 1, 2, \dots\}$ in (4) are uncorrelated. If it is reasonable to assume that the error term $\epsilon(\cdot)$ in (1) follows a normal distribution, then $\{e^*(t_j^*), j = 1, 2, \dots\}$ are a sequence of i.i.d. standard multivariate normal random variables if the observations

are IC. Therefore, determining the control limit l in our proposed MDySS procedure can be achieved by generating data from the standard multivariate normal distribution as $\{\mathbf{e}^*(t_j^*), j = 1, 2, \dots\}$ and finding l to obtain the desired ATS_0 for any given λ through a bi-section search. The bi-section search algorithm is given below:

Step 1. For a given control limit l , we simulate 10,000 IC sample paths. In each of the IC sample paths, $\{\mathbf{e}^*(t_j^*), j = 1, 2, \dots\}$ are generated from the standard multivariate normal distribution. The corresponding ATS_0^l for the given l is determined by averaging out the time to a signal in the basic time unit ω from these 10,000 sample paths. Based on this approach, we first find l_1 such that $ATS_0^{l_1} < ATS_0$, and l_2 such that $ATS_0^{l_2} > ATS_0$, where we use the notation ATS_0^l to denote the ATS_0 value when the control limit l is used.

Step 2. Find $ATS_0^{l_3}$ where l_3 is the midpoint of l_1 and l_2 .

Step 3. If $ATS_0^{l_3} < ATS_0$, assign $l_1 = l_3$. If $ATS_0^{l_3} > ATS_0$, assign $l_2 = l_3$;

Step 4. Repeat Steps 2 and 3 until $ATS_0^{l_3}$ is sufficiently close to ATS_0 ;

Step 5. Use l_3 as the control limit.

As mentioned earlier, the observation times in most longitudinal data studies may not be equally spaced. We specify the distribution of the observation times by the sampling rate d , which is defined to be the number of observation times every 10 basic time units here. The calculated control limits l using the above bi-section search algorithm for various values of d and λ in our proposed MDySS procedure (4) with $ATS_0 = 20$ or 50 are given in Table 1.

Table 1: Control limits l for different d and λ values when observations are normally distributed.

λ	$ATS_0 = 20$		$ATS_0 = 50$	
	$d = 2$	$d = 5$	$d = 2$	$d = 5$
0.05	1.368	3.061	3.239	6.100
0.1	2.415	4.786	5.088	8.300
0.2	3.899	6.644	7.116	10.150

If the distribution of the error term $\epsilon(t)$ in (1) is unknown but the number of observations for

each subject is large, we can still use the above method to determine the control limit, since $\mathbf{e}^*(t_j^*)$ is simply a linear combination of observations and its distribution is asymptotically normal. If the distribution of the error term $\boldsymbol{\epsilon}(t)$ is unknown and the number of observations for each subject is small, $\mathbf{e}^*(t_j^*)$ might not be asymptotically normally distributed. In this case, we can use a bootstrap procedure to determine the control limit as follows. Among the m well-functioning subjects in the IC dataset, we first use the observations from m_1 ($m_1 < m$) subjects to fit the multivariate nonparametric model (1) and obtain the estimates of $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\boldsymbol{\Sigma}}(s, t)$, as discussed in Section 2.1. Based on these $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\boldsymbol{\Sigma}}(s, t)$, we apply the sequential decorrelation procedure described in Section 2.2.1 to the remaining $m - m_1$ subjects in the IC dataset and obtain the decorrelated residuals $\{\mathbf{e}^*(t_{ij}), i = m_1 + 1, m_1 + 2, \dots, m, j = 1, \dots, n_i\}$. Those residuals are asymptotically uncorrelated and can be used to approximate the distribution of $\mathbf{e}^*(t_j^*)$ in our proposed MDySS procedure in (4). To determine the control limit l to achieve the desired ATS_0 for any given λ and d , we can then follow the same bi-section search algorithm as described above except that, in each of the IC sample paths, $\{\mathbf{e}^*(t_j^*), j = 1, 2, \dots\}$ are now drawn randomly with replacement from $\{\mathbf{e}^*(t_{ij}), i = m_1 + 1, m_1 + 2, \dots, m, j = 1, \dots, n_i\}$.

As a summary, our proposed sequential decorrelation procedure transforms the original data $\{\mathbf{y}(t_j^*), j = 1, 2, \dots\}$, which might be longitudinally correlated, to the sequence of uncorrelated data $\{\mathbf{e}^*(t_j^*), j = 1, 2, \dots\}$. The distribution of the transformed observations should be close to normal in many cases because each of them is a linear combination of the original data. So, the EWMA chart (4) should be effective in such cases. In some special cases, the distribution of the transformed observations could be quite different from a normal distribution. In such cases, we have suggested a bootstrap procedure to determine the control limit of the chart (4). So, the chart should still be reliable. A potential improvement in the latter cases is to consider using a multivariate nonparametric control chart (cf., Chapter 9 of Qiu³), instead of the multivariate EWMA chart (4), for monitoring $\mathbf{e}^*(t_j^*)$. We leave this issue to our future research.

3 Numerical Study

In this section, we report some simulation results to demonstrate the performance of our proposed MDySS procedure. In our simulations, we follow Qiu and Xiang¹⁹ and choose the IC mean function $\boldsymbol{\mu}(t)$ to be

$$\boldsymbol{\mu}(t) = (0, t, 1 + 0.2t + 0.3t^2, 1 - \exp(-10t), \cos(t))^T, \quad t \in [0, 1],$$

and the error term $\boldsymbol{\epsilon}(t)$ to be $\boldsymbol{\epsilon}(t) = B(t)\boldsymbol{\epsilon}^*(t)$, for $t \in [0, 1]$, where $B(t) = \text{diag}(1, \exp(t), 1/(1+t), 2, \log(t+5))$, and $\boldsymbol{\epsilon}^*(t)$ is generated from the following 5-dimensional AR(1) model

$$\boldsymbol{\epsilon}^*(t) = 0.5\boldsymbol{\epsilon}^*(t - \omega) + \boldsymbol{e}(t).$$

In the above multivariate AR(1) model, ω is the basic time unit, and $\boldsymbol{e}(t)$ is a white noise process with mean $\mathbf{0}$ and covariance matrix

$$\Sigma_0 = \begin{pmatrix} 1 & 0.8 & 0.8^2 & 0.8^3 & 0.8^4 \\ 0.8 & 1 & 0.8 & 0.8^2 & 0.8^3 \\ 0.8^2 & 0.8 & 1 & 0.8 & 0.8^2 \\ 0.8^3 & 0.8^2 & 0.8 & 1 & 0.8 \\ 0.8^4 & 0.8^3 & 0.8^2 & 0.8 & 1 \end{pmatrix}.$$

The corresponding IC covariance matrix function $\Sigma(s, t)$ is

$$\begin{aligned} \Sigma(s, t) &= \frac{4}{3} \times 0.5^{(s-t)/\omega} \times \text{diag}(1, \exp(s), 1/(1+s), 2, \log(s+5)) \\ &\quad \times \Sigma_0 \times \text{diag}(1, \exp(t), 1/(1+t), 2, \log(t+5)), \quad t \in [0, 1]. \end{aligned}$$

As described earlier, the observation times for each new subject is specified by the sampling rate d . Throughout all the simulation studies presented in this section we choose the basic time unit ω to be 0.01, and d to be 2 or 5. For our proposed MDySS procedure (4) to a new subject's observations, we set the nominal ATS_0 value at 20 or 50, and the weighting parameter λ at 0.05,

0.1 or 0.2.

We first study the IC performance of the proposed method in cases when the IC mean vector function $\boldsymbol{\mu}(t)$ and the IC covariance matrix function $\Sigma(s, t)$ are both known. Using the control limits reported in Table 1, we apply our proposed MDySS procedure (4) to a new subject's observations generated from the above IC model and obtain the IC ATS value. This is repeated 1,000 times, and the average of the 1,000 IC ATS values is computed as the simulated ATS_0 value of our proposed MDySS procedure. We repeat this simulation 100 times. Table 2 shows the means of the simulated ATS_0 values along with their corresponding standard errors (in the parentheses) under different settings. From the table, it can be seen that the simulated ATS_0 values are close to the nominal levels in all cases considered, which indicates that our sequential decorrelation procedure can successfully decorrelate the observed data because the control limits used are based on the assumption that observations are longitudinally uncorrelated.

Table 2: Simulated ATS_0 for different d and λ values when $\boldsymbol{\mu}(t)$ and $\Sigma(s, t)$ are both assumed known.

λ	$ATS_0 = 20$		$ATS_0 = 50$	
	$d = 2$	$d = 5$	$d = 2$	$d = 5$
0.05	19.972(0.038)	19.990(0.040)	49.941(0.095)	50.099(0.096)
0.1	19.937(0.043)	19.908(0.042)	49.957(0.089)	50.228(0.106)
0.2	20.043(0.044)	19.890(0.043)	50.090(0.088)	49.848(0.104)

We next study the OC performance of the proposed MDySS procedure, and apply it to the simulated data of 1,000 new subjects whose mean vector function shifts to $\boldsymbol{\mu}_1(t) = \boldsymbol{\mu}(t) + \delta \cdot (1, 1, 1, 1, 1)^T$, which represents a step shift of size δ in all 5 components at the initial observation time. We choose δ to be 0.25, 0.5, 0.75, or 1. The simulated ATS_1 values of our proposed MDySS procedure under different settings are presented in Table 3 along with their standard errors (in parentheses), based on 100 replications. As can be seen from the table, our MDySS procedure has a decent detection power.

As a comparison, we also apply the procedure proposed by Qiu and Xiang¹⁹ to this example. That procedure uses the same MEWMA statistic as in (4), except that the $\mathbf{e}^*(t_j^*)$ are simply the residuals standardized by the estimated covariance matrices. When the original data are correlated,

Table 3: Simulated ATS_1 for different d and λ values when $\boldsymbol{\mu}(t)$ and $\Sigma(s, t)$ are both known.

d	ATS_0	λ	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1$
2	20	0.05	18.882(0.035)	16.728(0.028)	14.510(0.021)	12.556(0.019)
		0.1	18.781(0.033)	16.508(0.029)	15.978(0.025)	12.229(0.020)
		0.2	18.680(0.035)	16.234(0.034)	15.516(0.025)	11.736(0.021)
	50	0.05	43.070(0.063)	33.488(0.044)	29.286(0.033)	22.151(0.023)
		0.1	42.753(0.071)	32.630(0.048)	28.160(0.033)	21.032(0.024)
		0.2	42.934(0.079)	32.153(0.053)	27.077(0.037)	19.871(0.029)
5	20	0.05	18.330(0.033)	15.322(0.025)	12.645(0.018)	10.571(0.015)
		0.1	18.167(0.037)	14.986(0.027)	13.349(0.021)	10.108(0.017)
		0.2	18.064(0.040)	14.755(0.028)	12.893(0.021)	9.643(0.018)
	50	0.05	39.802(0.067)	28.417(0.039)	22.993(0.027)	17.245(0.021)
		0.1	39.910(0.079)	27.771(0.046)	21.911(0.030)	16.076(0.021)
		0.2	40.784(0.092)	28.247(0.059)	21.586(0.037)	15.417(0.024)

$e^*(t_j^*)$ are still correlated. If we ignore such autocorrelation and pretend the transformed data $e^*(t_j^*)$ to be independent, then the actual ATS_0 values of that procedure in cases considered in the above example are presented in Table 4. From the table, it can be seen that the simulated ATS_0 values are quite far away from the nominal ATS_0 values, indicating that the procedure is unreliable in this scenario. To overcome this limitation, Qiu and Xiang also proposed a block bootstrap procedure to determine the control limit of the MEWMA chart so that the underlying data autocorrelation can be accommodated. The results corresponding to Tables 2 and 3 are presented in Table 5, where the

Table 4: Simulated ATS_0 of the procedure proposed by Qiu and Xiang¹⁹ when the autocorrelation in the observed data is ignored.

λ	$ATS_0 = 20$		$ATS_0 = 50$	
	$d = 2$	$d = 5$	$d = 2$	$d = 5$
0.05	17.207(0.031)	11.970(0.024)	41.224(0.066)	25.041(0.049)
0.1	17.050(0.032)	11.332(0.024)	40.314(0.072)	22.466(0.051)
0.2	16.930(0.034)	10.704(0.025)	39.435(0.084)	20.389(0.052)

column labeled by “ $\delta = 0$ ” gives the IC results and the remaining columns give OC results. From the table, it can be seen that the chart is quite reliable in this scenario because the simulated ATS_0 values are quite close to the nominal values. By comparing results in this table and those in Table 3, we can also see that the method by Qiu and Xiang¹⁹ using the block bootstrap procedure has

similar shift detection power. However, if we compare the standard errors presented in parentheses in the two tables, we can see that the ones in Table 5 are about 2-3 times larger than the ones in Table 3. Because the results in these tables are based on 100 replicated simulations, the standard deviations of the related methods are 10 times the standard errors presented in the tables. So, for instance, in the case when $d = 2$, $ATS_0 = 20$, $\lambda = 0.05$ and $\delta = 0.25$, the simulated ATS_0 values of the method by Qiu and Xiang¹⁹ can vary in the interval (17.041, 21.001) (i.e., $(\bar{x} - 2s, \bar{x} + 2s)$), while the simulated ATS_0 values of the current method can only vary in the interval (18.182, 19.582). In this sense, the current method is much more reliable.

Table 5: Simulated ATS for different d and λ values for Qiu and Xiang's block bootstrap EWMA procedure.

d	ATS_0	λ	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1$
2	20	0.05	20.062(0.114)	19.021(0.099)	16.805(0.074)	14.555(0.058)	12.625(0.045)
		0.1	20.043(0.109)	18.913(0.095)	16.603(0.072)	15.880(0.059)	12.315(0.041)
		0.2	19.990(0.127)	18.736(0.110)	16.267(0.082)	15.353(0.067)	11.791(0.045)
	50	0.05	50.431(0.231)	43.552(0.179)	33.733(0.112)	29.198(0.081)	22.267(0.064)
		0.1	50.424(0.236)	43.213(0.186)	32.869(0.110)	28.057(0.078)	21.111(0.059)
		0.2	50.526(0.237)	43.263(0.203)	32.264(0.125)	26.854(0.084)	19.813(0.061)
5	20	0.05	20.063(0.111)	18.374(0.090)	15.377(0.063)	12.715(0.047)	10.654(0.036)
		0.1	20.083(0.114)	18.264(0.094)	15.054(0.066)	13.100(0.047)	10.169(0.034)
		0.2	20.081(0.128)	18.169(0.104)	14.802(0.072)	12.617(0.052)	9.668(0.037)
	50	0.05	50.240(0.223)	39.952(0.159)	28.520(0.092)	22.709(0.060)	17.372(0.046)
		0.1	49.940(0.264)	39.774(0.192)	27.723(0.108)	21.493(0.068)	16.074(0.052)
		0.2	49.636(0.268)	40.643(0.229)	28.057(0.124)	21.035(0.077)	15.295(0.056)

It should also be pointed out that the method by Qiu and Xiang¹⁹ is more computationally intensive, compared to the current procedure, because the former needs to use the block bootstrap procedure to determine its control limit. Furthermore, Qiu and Xiang's procedure requires that the sampling scheme in the IC data collected at the end of Phase I SPC should be exactly the same as the sampling scheme in the Phase II data. To show the importance of this sampling scheme requirement for Qiu and Xiang's procedure, we carry out a simulation using different sampling schemes for IC data and Phase II data. The sampling rate d for the IC data and the Phase II data are denoted by d_1 and d_2 , respectively. The simulated ATS_0 values for Qiu and Xiang's procedure for different d_1 and d_2 are shown in Table 6. As we can see from the table, even when

d_1 and d_2 are slightly different, the simulated ATS_0 values can be quite off from the nominal levels. In contrast, our proposed method does not have this sampling scheme requirement, because the transformed data $\{\mathbf{e}^*(t_j^*)\}$ used in (4) are uncorrelated and thus we can always adopt the same sampling scheme as that in the Phase II data when designing the control chart. See Section 2.2.2 about the details of control chart design. This property makes our method potentially more applicable in real applications, since the IC data collected at the end of Phase I SPC and the Phase-II data could have different sampling schemes in many situations. Our real-data application discussed in Section 4 presents one such example.

Table 6: Simulated ATS_0 for different d_1 and d_2 values for Qiu and Xiang’s block-bootstrap MEWMA procedure.

λ	$ATS_0 = 20$		$ATS_0 = 50$	
	$d_1 = 2, d_2 = 3$	$d_1 = 5, d_2 = 6$	$d_1 = 2, d_2 = 3$	$d_1 = 5, d_2 = 6$
0.05	12.311(0.034)	15.384(0.053)	30.605(0.090)	38.570(0.145)
0.1	12.244(0.042)	15.263(0.057)	30.452(0.100)	38.170(0.171)
0.2	12.265(0.044)	15.262(0.058)	30.459(0.122)	38.279(0.182)

Next, we consider cases when the IC mean vector function $\boldsymbol{\mu}(t)$ and the IC covariance matrix function $\Sigma(s, t)$ are both unknown and they need to be estimated from an IC data. We assume that we have $m = 1,000$ subjects in the IC data, their observations are generated from the IC mean vector function $\boldsymbol{\mu}(t)$ and the IC covariance matrix function $\Sigma(s, t)$ described at the beginning of the section. We apply the estimation procedure described in Section 2.1 to the IC data, and obtain the estimated IC mean vector function and IC covariance matrix function, $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\Sigma}(s, t)$. By using those $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\Sigma}(s, t)$, we apply the proposed MDySS procedure (4) to the 1,000 new subjects generated from the mean vector function $\boldsymbol{\mu}_1(t) = \boldsymbol{\mu}(t) + \delta \cdot (1, 1, 1, 1, 1)^T$, as in the previous simulation example. The average time to signal from these 1,000 new subjects is the simulated ATS value from one single IC dataset. We then repeat this simulation with 100 different IC datasets. Table 7 shows the averaged ATS_0 (in the column labeled by “ $\delta = 0$ ”) and ATS_1 values along with their standard errors (in the parentheses) under different settings, based on the 100 replicated simulations. As we can see from the table, all the values here are close to their counterparts

in Tables 2 and 3 when $\boldsymbol{\mu}(t)$ and $\Sigma(s, t)$ are assumed known. This example indicates that the estimation procedure described in Section 2.1 can provide reliable estimates of $\boldsymbol{\mu}(t)$ and $\Sigma(s, t)$.

Table 7: Simulated ATS_0 and ATS_1 values for different d and λ values when $\boldsymbol{\mu}(t)$ and $\Sigma(s, t)$ are both unknown.

d	ATS_0	λ	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1$
2	20	0.05	19.988(0.047)	18.886(0.046)	16.513(0.036)	14.091(0.029)	12.002(0.024)
		0.1	19.974(0.047)	18.800(0.047)	16.298(0.035)	15.821(0.032)	11.677(0.025)
		0.2	20.005(0.055)	18.711(0.049)	16.025(0.038)	15.359(0.032)	11.189(0.024)
	50	0.05	49.757(0.094)	43.173(0.077)	33.166(0.061)	29.026(0.047)	21.124(0.036)
		0.1	49.694(0.104)	42.855(0.084)	32.350(0.062)	27.897(0.049)	19.999(0.037)
		0.2	49.863(0.115)	43.045(0.098)	31.900(0.069)	26.765(0.052)	18.757(0.038)
5	20	0.05	19.968(0.040)	18.322(0.036)	15.230(0.028)	12.476(0.024)	10.362(0.020)
		0.1	19.897(0.047)	18.132(0.042)	14.890(0.029)	13.273(0.024)	9.874(0.019)
		0.2	19.776(0.053)	17.940(0.047)	14.603(0.031)	12.776(0.025)	9.378(0.020)
	50	0.05	49.390(0.106)	39.576(0.065)	28.247(0.049)	22.876(0.037)	16.898(0.029)
		0.1	49.359(0.106)	39.509(0.074)	27.597(0.054)	21.782(0.040)	15.699(0.031)
		0.2	48.765(0.111)	40.159(0.087)	27.975(0.064)	21.355(0.046)	15.023(0.035)

In Section 2.2.3, we proposed a bootstrap procedure to determine the control limit when the distribution of the error term is unknown and the number of observations for each subject is small. In the following, we report some simulation results to evaluate the performance of this bootstrap procedure. More specifically, in each simulation run, we first generate an IC dataset with $m = 1,000$ subjects from the same IC model described above expect that the $\boldsymbol{e}(t)$ are now i.i.d. multivariate t random vectors with mean $\mathbf{0}$, degrees of freedom 4, and the covariance matrix Σ_0 (see its definition at the beinning of this section). The corresponding IC covariance matrix function $\Sigma(s, t)$ is

$$\begin{aligned} \Sigma(s, t) &= \frac{8}{3} \times 0.5^{(s-t)/\omega} \times \text{diag}(1, \exp(s), 1/(1+s), 2, \log(s+5)) \\ &\quad \times \Sigma_0 \times \text{diag}(1, \exp(t), 1/(1+t), 2, \log(t+5)), \quad t \in [0, 1]. \end{aligned}$$

We then use the data from the first 800 subjects to obtain the estimates of the IC mean vector function and covariance matrix function, $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\Sigma}(s, t)$. Based on these $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\Sigma}(s, t)$, we sequentially decorrelate the residuals from the remaining 200 subjects. The control limit is then determined using the bootstrap resamples of those decorrelated residuals as described in Section

2.2.3. We apply our MDySS procedure with this control limit to the 1,000 new subjects generated from the mean vector function $\boldsymbol{\mu}_1(t) = \boldsymbol{\mu}(t) + \delta \cdot (1, 1, 1, 1, 1)^T$. The average time to the signal from these 1,000 new subjects is the simulated ATS . We repeat this simulation 100 times. Table 8 shows the averaged ATS_0 and ATS_1 values along with their standard errors (in the parentheses) under different settings, based on the 100 simulations. As we can see from the table, all the ATS_0 values are within 10% of the nominal ATS_0 values, which indicates the validity of our bootstrap procedure. When δ increases, the ATS_1 value decreases, which indicates a good detection power of our bootstrap procedure.

Table 8: Simulated ATS values for different d and λ values when the bootstrap procedure is used in determining the control limit.

d	ATS_0	λ	$\delta = 0$	$\delta = 0.5$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
2	20	0.05	21.379(0.056)	18.474(0.047)	14.119(0.036)	10.952(0.028)	8.778(0.023)
		0.1	21.562(0.061)	18.473(0.048)	13.892(0.037)	13.643(0.029)	8.462(0.022)
		0.2	21.827(0.063)	18.543(0.052)	13.642(0.039)	13.104(0.028)	7.984(0.022)
	50	0.05	53.299(0.144)	40.477(0.110)	27.102(0.070)	23.561(0.052)	15.489(0.037)
		0.1	53.625(0.150)	40.717(0.122)	26.473(0.073)	22.525(0.053)	14.513(0.038)
		0.2	53.913(0.165)	42.493(0.148)	26.805(0.091)	21.751(0.062)	13.564(0.042)
5	20	0.05	21.123(0.059)	17.491(0.049)	12.797(0.038)	9.639(0.028)	7.600(0.022)
		0.1	21.296(0.066)	17.456(0.054)	12.542(0.039)	10.855(0.029)	7.184(0.023)
		0.2	21.512(0.069)	17.920(0.060)	12.740(0.043)	10.611(0.031)	6.872(0.024)
	50	0.05	52.456(0.153)	35.903(0.099)	22.842(0.063)	18.077(0.045)	12.390(0.035)
		0.1	52.503(0.159)	37.256(0.112)	22.838(0.072)	17.354(0.047)	11.530(0.036)
		0.2	52.219(0.157)	41.664(0.142)	25.994(0.099)	18.457(0.064)	11.779(0.047)

4 A Real Application

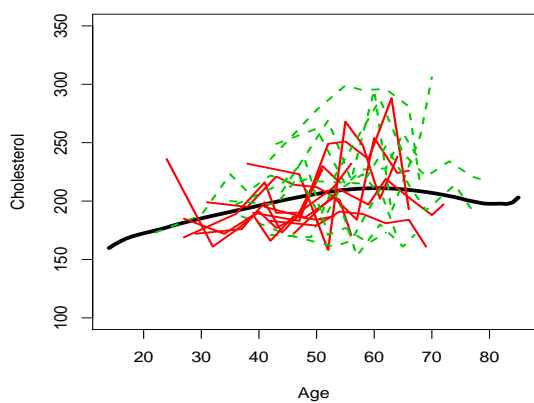
In this section we use a data set from the SHARe Framingham Heart Study of the National Heart, Lung and Blood Institute to demonstrate the application of our proposed MDySS procedure. In the SHARe Framingham Heart Study, there are 1006 patients, and each of the patients was followed 7 times. During each clinic visit, four medical indices (the total cholesterol level (mg/100ml), the glucose level (mg/100ml), the diastolic blood pressure (mmHg), and the systolic blood pressure (mmHg)) of the patients were recorded. Among the 1006 patients, 979 did not have any strokes

during the study, and the remaining 27 patients had at least one stroke. Figure 1 shows the four medical index values from 10 randomly selected non-stroke patients (thin solid red line) and 10 randomly selected stroke patients (dashed green line) over the 7 clinic visits. It seems that the stroke patients have higher values in all four medical indices than those non-stroke patients.

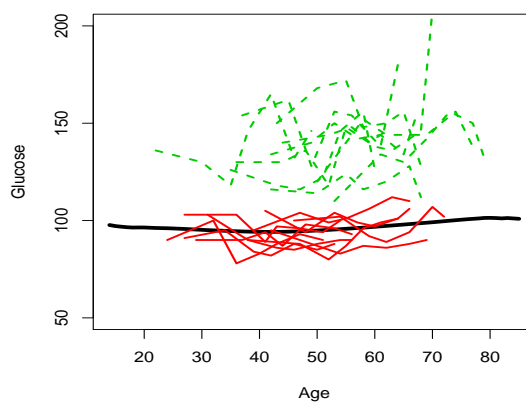
To develop a control chart that can simultaneously monitor a patient's four medical indices, we use the data of the 979 non-stroke patients as the IC data, and then apply the estimation procedure described in Section 2.1 to 800 of the non-stroke patients to obtain the estimated IC mean vector function $\hat{\boldsymbol{\mu}}(t)$ and the estimated IC covariance matrix function $\hat{\Sigma}(s, t)$. The estimated IC mean vector function $\hat{\boldsymbol{\mu}}(t)$ is shown as the dark solid curve in Figure 1. Based on $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\Sigma}(s, t)$, we apply our sequential decorrelation procedure to the remaining 179 non-stroke patients to obtain the decorrelated residuals. The control limit is then determined by the bootstrap procedure described in Section 2.2.3. Since each stroke patient has an average of 2.3 observations every 10 years, we choose $d = 2$ and the basic time unit is a year. We use $\lambda = 0.2$ and $ATS_0 = 20$ in our proposed MDySS procedure. The corresponding control limit l determined from our bootstrap procedure is 3.425. We then apply our MDySS procedure with this control limit to the 27 stroke patients, and all of them are detected to have mean shifts. The corresponding control charts of our MDySS procedure for the 27 stroke patients are shown in Figure 2. The signal times of the 27 stroke patients are reported in Table 9, and their average is 13.37 years.

Table 9: Signal times of the 27 stroke patients by our proposed MDySS procedure (ST1) and Qiu and Xiang's block bootstrap MEWMA procedure (ST2).

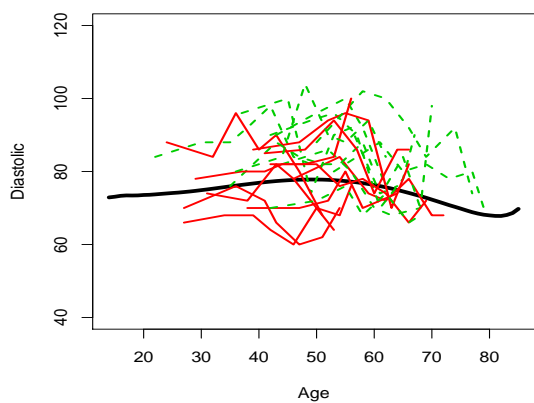
Patient ID	ST1	ST2	ID	ST1	ST2	ID	ST1	ST2
1	26	-	10	19	23	19	23	23
2	23	26	11	19	22	20	11	11
3	17	25	12	12	15	21	7	19
4	11	-	13	7	7	22	7	7
5	11	11	14	15	25	23	19	-
6	7	12	15	0	23	24	19	22
7	12	12	16	12	12	25	7	11
8	12	23	17	15	23	26	12	16
9	15	11	18	16	-	27	7	12



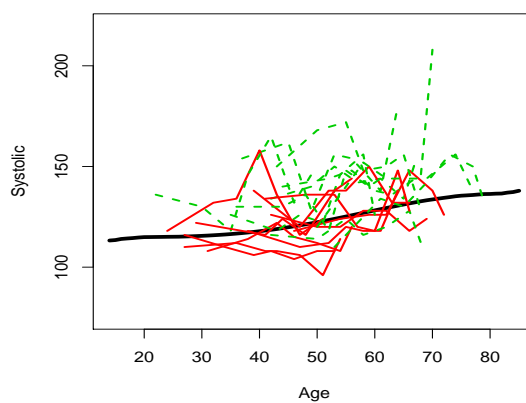
(a)



(b)



(c)



(d)

Figure 1: (a) Total cholesterol level (mg/100ml), (b) glucose level (mg/100ml), (c) diastolic blood pressure (mmHg), and (d) systolic blood pressure (mmHg) of 10 randomly selected non-stroke patients (thin solid red line) and 10 randomly selected stroke patients (dashed green line). The dark solid curve is the estimated IC mean function $\hat{\mu}(t)$ using the 979 non-stroke patients as the IC data.

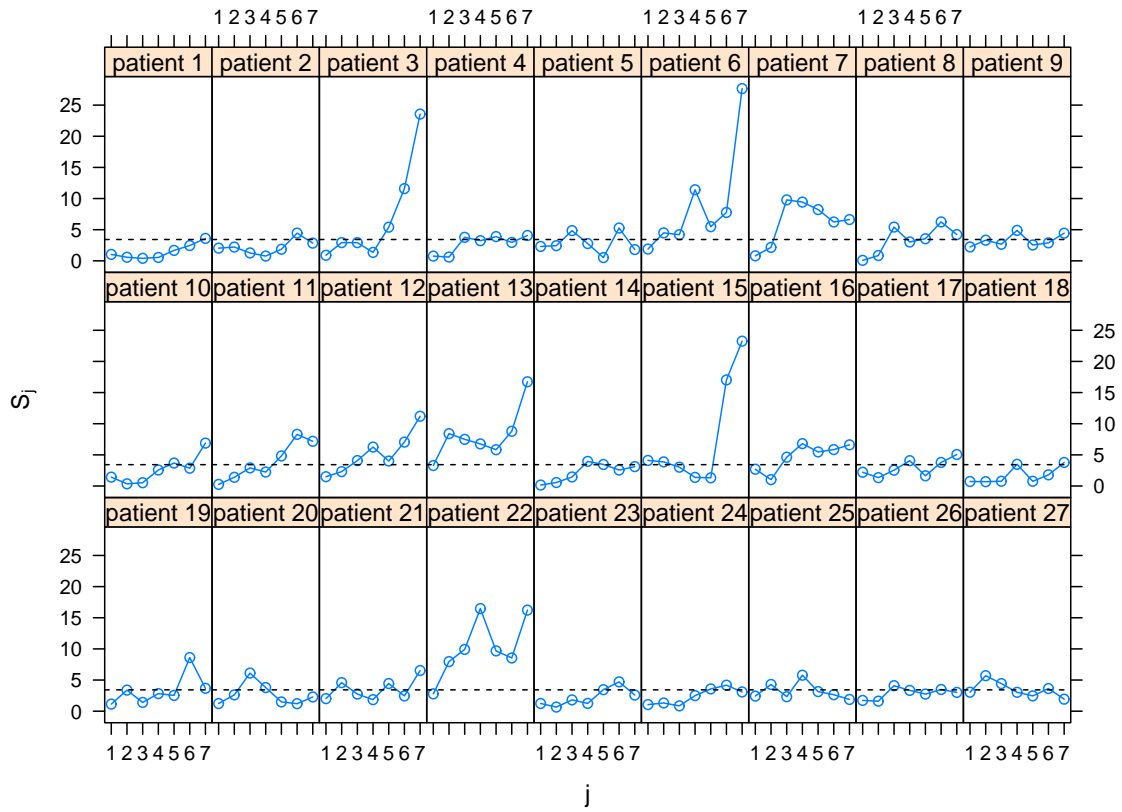


Figure 2: The proposed MEWMA chart for monitoring the 27 stroke patients. The dashed horizontal lines denote the control limit of $l = 3.425$.

As a comparison, we also apply Qiu and Xiang's block bootstrap MEWMA procedure to this data set. Its results are shown in Figure 3, and their procedure fails to detect 4 of 27 stroke patients. The signal times of the stroke patients detected by their procedure are also reported in Table 9, and their average is 17 years. From Table 9, we can also see that the signal times of Qiu and Xiang's procedure are mostly larger than the signal times of our proposed procedure.

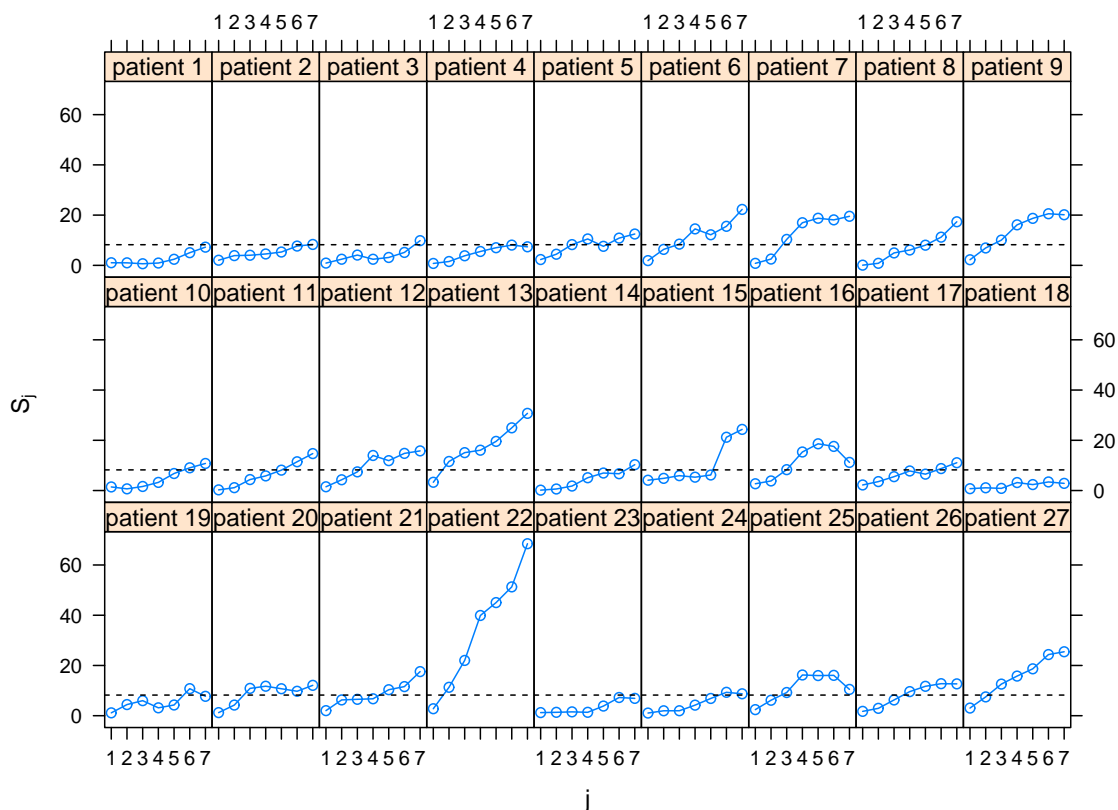


Figure 3: The MEWMA chart for monitoring the 27 stroke patients using the block bootstrap method proposed in Qiu and Xiang¹⁹. The dashed horizontal lines denote the control limit of $l = 8.215$.

The reason why Qiu and Xiang's block bootstrap MEWMA procedure does not perform well in this example might be due to the difference of observation times (ages) in non-stroke patients and stroke patients. To see this more clearly, we plot the histograms of the observation times (in ages) for the non-stroke patients and stroke patients in Figure 4. The average observation

times when the 7 measurements were taken from the non-stroke patients and stroke patients are reported in Table 10. As seen from those results, the stroke patients tend to be observed at later ages, compared to the non-stroke patients, indicating that the sampling scheme for the non-stroke patients (IC data) is quite different from that for the stroke patients (Phase II data). As shown in our simulation study presented in Section 3, to ensure the validity of Qiu and Xiang’s block bootstrap MEWMA procedure, the sampling scheme in the IC data has to be the same as the sampling scheme in the Phase II data. The difference of observation times in non-stroke patients and stroke patients might explain why Qiu and Xiang’s block bootstrap MEWMA procedure does not work well in this data set.

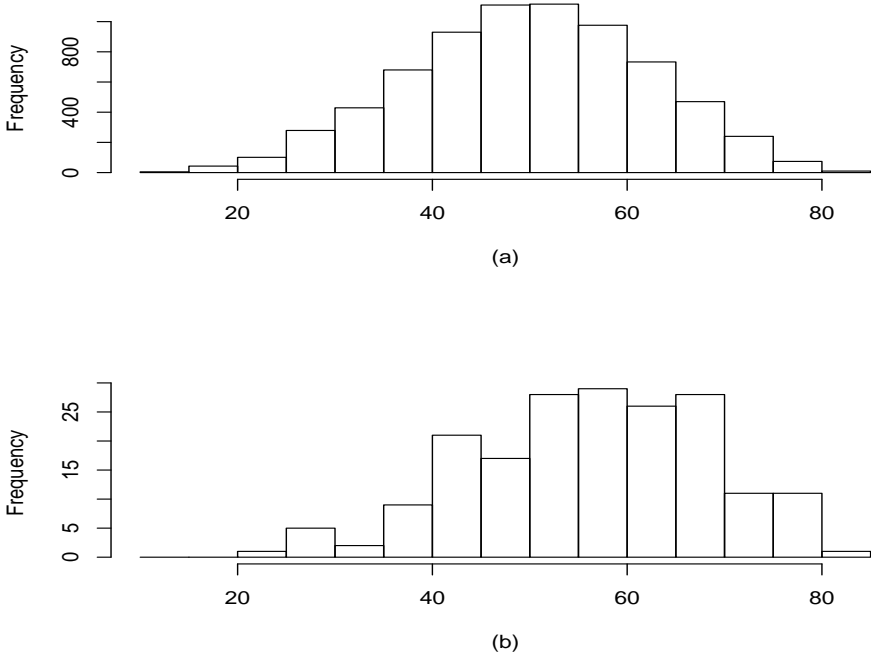


Figure 4: The histogram of the observation times (in ages) for (a) the non-stroke patients; (b) the stroke patients.

Table 10: The average times (in ages) when each of the seven measurements was taken from the non-stroke patients and stroke patients.

Measurement	1	2	3	4	5	6	7
Non-stroke patients	35.37	43.21	47.62	51.06	54.72	58.75	61.61
Stroke patients	41.89	49.81	53.93	57.59	61.15	65.37	67.93

5 Concluding Remarks

We have presented a new construction of the MDySS method for online monitoring of dynamic systems whose IC distributions change over time. The new construction has been shown to be more reliable in cases when the process observations are correlated and the sampling rates in Phase I and Phase II SPC are different. In Butte and Tang²³, a line-column approach was proposed to identify the OC components after a signal is given from the multivariate EWMA chart. We can similarly adopt their approach to identify which performance variables have caused the alarm in our online monitoring of dynamic systems. In this paper, the weighting parameter λ in the multivariate EWMA chart (4) is assumed fixed. This parameter can also be determined using an adaptive parameter selection procedure (see, for example, Section 5.4.2. of Qiu³). Furthermore, we can replace (4) by a self-starting control chart which keeps expanding the IC data each time when a Phase II observation is confirmed to be IC (Section 5.4.1. of Qiu³). Also, as mentioned at the end of Section 2, a nonparametric multivariate control chart can be considered for monitoring the decorrelated data. These possible improvements will be studied in a future research.

References

1. Hawkins DM, Olwell DH. *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Springer-Verlag, 1998.
2. Montgomery DC. *Introduction To Statistical Quality Control (6th edition)*, New York: John Wiley & Sons, 2009.
3. Qiu P. *Introduction to Statistical Process Control*, London: Chapman & Hall/CRC, 2014.

4. Shewhart WA. *Economic control of quality of manufactured product*, New York: D. Van Nostrand Company, 1931.
5. Page ES. Continuous inspection schemes. *Biometrika* 1954; **41**: 100-114.
6. Roberts SV. Control chart tests based on geometric moving averages. *Technometrics* 1959; **1**: 239-250.
7. Hawkins DM, Qiu P, Kang CW. The changepoint model for statistical process control. *Journal of Quality Technology* 2003; **35**: 355-366.
8. Chen N, Zi X, Zou C. A distribution-free multivariate control chart. *Technometrics* 2015; in press.
9. Bakir ST, Reynolds MR Jr. A nonparametric procedure for process control based on within group ranking. *Technometrics* 1979; **21**: 175-183.
10. Chakraborti S, van der Laan P, Bakir ST. Nonparametric control charts: an overview and some results. *Journal of Quality Technology* 2001; **33**: 304-315.
11. Li J. Nonparametric multivariate statistical process control charts: a hypothesis testing-based approach. *Journal of Nonparametric Statistics* 2015; **27**: 384-400.
12. Qiu P, Hawkins D. A rank based multivariate CUSUM procedure. *Technometrics* 2001; **43**: 120-132.
13. Qiu P. Distribution-free multivariate process control based on log-linear modeling. *IIE Transactions* 2008; **40**: 664-677.
14. Zou C, Tsung F. A multivariate sign EWMA control chart. *Technometrics* 2011; **53**: 84-97.
15. Apley DW, Lee HC. Design of exponentially weighted moving average control charts for autocorrelated processes with model uncertainty. *Technometrics* 2003; **45**: 187-198.
16. Lu CW, Reynolds MR Jr. Control charts for monitoring an autocorrelated process. *Journal of Quality Technology* 2001; **33**: 316-334.

17. Runger GC, Willemain TR. Batch means control charts for autocorrelated data. *IIE Transactions* 1996; **28**: 483-487.
18. Qiu P, Xiang D. Univariate Dynamic Screening System: An Approach For Identifying Individuals With Irregular Longitudinal Behavior. *Technometrics* 2014; **56**: 248-260.
19. Qiu P, Xiang D. Surveillance of cardiovascular diseases using a multivariate dynamic screening system. *Statistics in Medicine* 2015; **34**: 2204-2221.
20. Xiang D, Qiu P, Pu X. Local polynomial regression analysis of multivariate longitudinal data. *Statistica Sinica* 2013; **23**: 769-789.
21. Bersimis S, Psarakis S, Panaretos J. Multivariate statistical process control charts: An overview. *Quality and Reliability Engineering International* 2007; **23**: 517-543.
22. Lowry CA, Woodall WH, Champ CW, Rigdon SE. A multivariate exponentially weighted moving average control chart. *Technometrics* 1992; **34**: 46-53.
23. Bette VK, Tang LC. Multivariate charting techniques: a review and a line-column approach. *Quality and Reliability Engineering International* 2010; **26**: 443-451.