Phase-II monitoring in multichannel profile observations

Haojie Ren[1], Nan Chen[2*] and Zhaojun Wang[1]
[1] *Institute of Statistics and LPMC, Nankai University, China*
[2] *Department of Industrial & Systems Engineering, National University of Singapore, Singapore*

Abstract

Process monitoring and fault diagnosis using profile data remain an important problem in statistical process control (SPC). It is particularly challenging with multichannel profiles. This article proposes a novel scheme for Phase II monitoring of multichannel profiles. The proposed method integrates the exponentially weighted moving average (EWMA) scheme with multichannel functional principal component analysis (MFPCA) to obtain the charting statistics. Their control limits can be effectively computed by the Markov chain method. Moreover, we provide a natural diagnostic procedure to locate the possible change point once the chart is out-of-control. Our proposed method is demonstrated to be effective and efficient through simulation results and an industrial case study from a multi-operation forging process.

Keywords: EWMA; Functional data analysis; Multichannel functional component analysis; Markov chain method; Statistical process control

1 Introduction

Because of advanced sensing technologies in complex manufacturing systems, large amounts of sophisticated and real-time data become available. These data provide unprecedented opportunities to monitor and improve the product quality or process efficiency, yet they also pose significant challenges to conventional statistical process control (SPC) methods. Different from univariate or multivariate quality characteristics commonly used in the literature (see, e.g., Montgomery 2007, Qiu 2013), in many SPC applications, their quality is

*Corresponding author: isecn@nus.edu.sg

characterized by the functional relationship between a response variable and one or more explanatory variables. This functional relationship is often referred to as a profile, and the SPC problem for profiles is known as profile monitoring and diagnosis. See Woodall et al. (2004), Woodall (2007) and Noorossana et al. (2011) for a nice overview.

Because of its importance, extensive research works focusing on profile monitoring have been recently studied. The literature can be divided into two categories. The first category characterizes the profile using parametric models, such as linear and nonlinear models. Then they monitor the parameters of the models, estimated from the profile data (see, e.g., Kang and Albin 2000, Kim et al. 2003, Zou et al. 2007, Jensen et al. 2008). The second category uses nonparametric models, including wavelet methods (Chicken et al. 2009) and nonparametric smoothing methods (Zou et al. 2008; Qiu et al. 2010) among many others, to characterize the profile. Then the profile monitoring is translated to the monitoring of features derived from the models, e.g., residuals or projection coefficients.

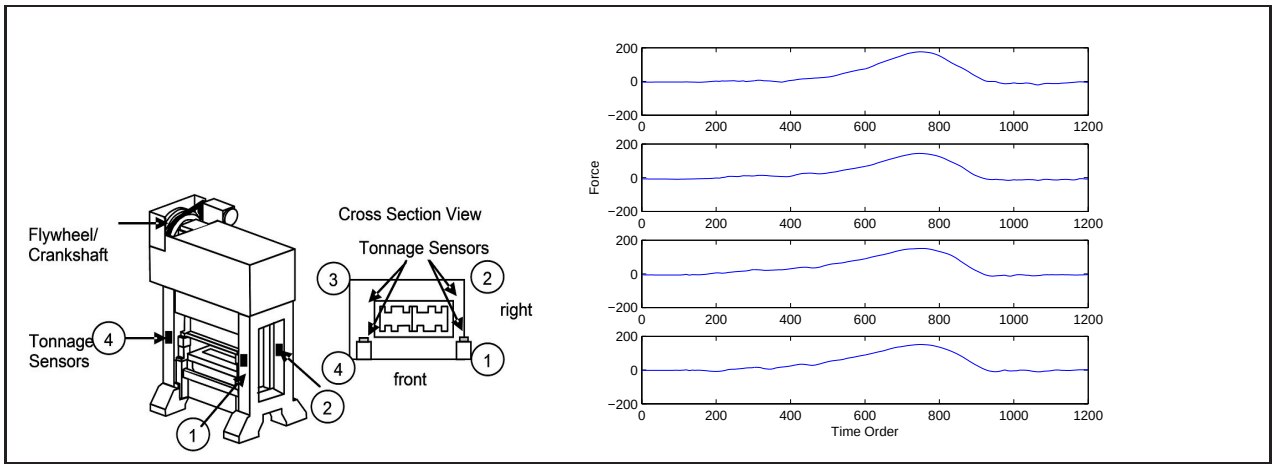

Figure 1: Left panel: A forging machine with four strain gages. Right panel: A sample of four-channel profiles.

Despite their significance, most existing profile monitoring methods are only applicable to a single profile data. In some industrial applications, however, product quality is characterized by profile data collected from multiple channels. For example, in the multi-operation forging process with transfer or progressive dies, as shown in the left panel of Figure 1, tonnage forces exerted on all dies are measured by four strain sensors mounted on four columns of the press. Each sensor records the tonnage force at a predefined equal sampling interval of a rotational crank angle. This results in multichannel profile data shown in the right panel of Figure 1. The four channels of profiles collectively indicate the quality of a particular

forging operation. See Section 4 for more details of this example.

Because of the inherent dependence among multichannel profiles, monitoring profile data in each single channel separately using aforementioned methods might not be effective. It is expected that by taking cross correlations among multichannel profiles into consideration, the profile monitoring could become more sensitive to a variety of shifts. However, research on the monitoring and diagnosis of general multichannel profiles is still limited. Among them, Noorossana et al. (2010) discussed multivariate linear profile monitoring in Phase I analysis, mainly based on the ordinary least square estimation. Zou et al. (2012) focused on the Phase II monitoring for multivariate linear profiles by using the LASSO-based multivariate SPC techniques. However, both approaches imposed linear assumptions on the profiles, and might not be generally applicable.

Among the research focusing on general multichannel profiles, Lei et al. (2010) combined multichannel profiles into a single profile by summing them up. Clearly, this approach does not utilize valuable cross-correlation information among channels. Chou et al. (2014) developed a process monitoring strategy for multiple correlated nonlinear profiles. They used B-splines to extract the mean of each profile, and monitored the combined coefficients of multichannel profiles. Another related approach is to combine multichannel profiles into a high-dimensional vector, and to apply dimension reduction methods, such as the principal component analysis (PCA) to extract features and construct monitoring statistics. Such vectorization and dimension reduction is also referred to as *Vectorized-PCA (VPCA)*. However, VPCA breaks the correlation structure in the original data and loses some useful representations that can be obtained in the original form. Moreover, it also suffers from "the curse of dimensionality" because it ignores the smoothness nature of functional profiles, making it less effective when the number of profiles or the number of sampling points is large. More recently, Paynabar et al. (2015) used a projection method called *Multivariate Functional Principal Component Analysis (MFPCA)* to extract informative features from multichannel profiles and to incorporate theses features into a change-point model for Phase-I analysis.

The main goal of this paper is to develop a Phase II monitoring approach for multichannel profiles. The proposed approach applies the functional principal components to obtain a set of extracted features that can be effectively used to characterize process variations and to construct an exponentially weighted moving average chart (EWMA). It possesses a good property that the IC average run length can be calculated via a one-dimensional Markov

chain model so that the control limit can be easily obtained. The effectiveness of the proposed method is shown via both simulation studies and a real-world case study in a multi-operation forging process. The remainder of the paper is organized as follows. Section 2 presents the proposed method for monitoring multichannel profiles in detail. Section 3 evaluates the performance of our proposed scheme and compares it with some alternatives through simulations. Section 4 revisits the industrial example in Figure 1 to illustrate the step-by-step implementation of the proposed approach. Section 5 concludes the article with several remarks.

2 Methodology

The Phase II monitoring and diagnosis procedure developed in this paper consists of five parts. In Section 2.1, the MFPCA is reviewed. The MFPCA is applied to transform the functional data, which has infinite dimension in nature, into a few components that can represent the main structure of multichannel profiles well. Then, these multivariate functional principal components are used to construct an EWMA control chart, as derived in Section 2.2. Section 2.3 discusses the practical guidelines for its design and implementation, especially the parameter estimations in Phase I analysis. A diagnostic approach is proposed to identify the change point of the process in Section 2.4. At last, an extension of the proposed method is discussed in Section 2.5.

2.1 A Brief Review of MFPCA

Motivated by the forging process example, to be specific, we assume $\boldsymbol{X}(u) = \{X_1(u), \ldots, X_p(u)\}$ is the p-channel profiles, following the model

$$\boldsymbol{X}(u) = \boldsymbol{\mu}(u) + \boldsymbol{Y}(u), \tag{1}$$

where $\boldsymbol{\mu}(u)$ denotes the p-dimensional mean function and $\boldsymbol{Y}(u)$ is the stochastic error with $E[\boldsymbol{Y}(u)] = \boldsymbol{0}$, u is the index, typically standing for time or space. We let $u \in \mathcal{T} = [0, 1]$ without loss of generality. The profile $\boldsymbol{X}(u)$ has an infinite dimension in nature, making it difficult to monitor directly. The MFPCA is designed to reduce its dimensionality and extracting a few major and representative features for process monitoring. Here we briefly review the MFPCA and introduce necessary notations.

Similar to conventional principal components analysis, which is based on the covariance matrix, the MFPCA is based on the covariance function of $\boldsymbol{Y}(u)$. Let $\mathbf{c}(u,u') = E\{\langle \boldsymbol{Y}(u), \boldsymbol{Y}(u')\rangle\}$ be the "overall covariance" function of vector-valued random function $\boldsymbol{Y}(\cdot)$ for $u, u' \in \mathcal{T}$, where we use a straightforward definition of the inner product between multi-dimensional functions, i.e., $\langle \boldsymbol{f}, \boldsymbol{g}\rangle = \sum_{j=1}^{p} f_j g_j$ introduced by Dubin and Müller (2005). Then, the spectral decomposition of $\mathbf{c}(u,u')$ is given by

$$\mathbf{c}(u,u') = \sum_{k=1}^{\infty} \lambda_k v_k(u) v_k(u').$$

Here λ_k and $v_k(\cdot)$ denote the eigenvalues and eigenfunctions, respectively. They satisfy $\int_0^1 \mathbf{c}(u,u')v_k(u')du' = \lambda_k v_k(u), \quad k = 1, 2, \ldots$. Based on this eigen decomposition, an integrable vector-valued random function $\boldsymbol{Y}(u)$ can be represented by the set of eigenfunctions as $\boldsymbol{Y}(u) = \sum_{1\le k<\infty} \boldsymbol{\xi}_k v_k(u)$, where the sequences $\boldsymbol{\xi}_k = \int_0^1 \boldsymbol{Y}(u)v_k(u)dt, k = 1, 2, \ldots$ are i.i.d. p-variate random variables with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_k$. It is easy to show that $\sigma_{kj}^2 = E[\xi_{kj}^2]$, where ξ_{kj} is the jth component of $\boldsymbol{\xi}_k$, and $\lambda_k = \sum_{j=1}^{p} \sigma_{kj}^2$. Note that in practice, only a few eigenvalues and eigenfunctions are required to capture the important variations of a sample of multichannel profiles.

The MFPCA model discussed above assumes that all functions from the p channels share a common set of eigenfunctions, and their inter-dependence are essentially described by the correlations among components of $\boldsymbol{\xi}_k$. This assumption is reasonable when profiles from different channels exhibit similar patterns like those in Figure 1, and thus is not restrictive in many applications. In particular, when $p = 1$, MFPCA reduces to the widely used univariate FPCA as discussed in Ramsay and Silverman (2005). In this situation, ξ_k is a random variable with mean 0 and variance λ_k instead of a random vector with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_k$. Therefore, our method also applies to single-channel profile monitoring. See some simulation results when $p = 1$ in Section 3.

2.2 Control Scheme for Monitoring Multichannel Profiles

Let us consider the Phase II monitoring for multichannel profiles. We assume the profile observations are collected sequentially according to the following model

$$\boldsymbol{X}_i(u) = \begin{cases} \boldsymbol{\mu}(u) + \boldsymbol{Y}_i(u), & \text{for} \quad i = 1, \ldots, \tau, \\ \boldsymbol{\mu}(u) + \boldsymbol{\delta}(u) + \boldsymbol{Y}_i(u), & \text{for} \quad i = \tau+1, \ldots, \end{cases} \tag{2}$$

where τ is the unknown change point and $\boldsymbol{\mu}(u)$ is the in-control mean function of profiles. We would like to focus on detecting any change $\boldsymbol{\delta}(u)$ in the mean function as quickly as possible. Following the convention in the literature, in Phase-II analysis, the in-control parameters including the mean function $\boldsymbol{\mu}(u)$, eigenfunction $v_k(u)$ and covariance matrix $\boldsymbol{\Sigma}_k$ are assumed to be known; or equivalently, the in-control data set in Phase I is sufficient to estimate the model well. Some discussions on the estimation of IC parameters are presented in Section 2.3.

Without loss of generality, we assume the in-control mean function $\boldsymbol{\mu}(u) = \mathbf{0}$. Otherwise, we can transform $\boldsymbol{X}_i(u)$ to $\boldsymbol{X}_i(u) - \boldsymbol{\mu}(u)$ to make the assumption valid. In the light of MFPCA, for on-line collected observations $\boldsymbol{X}_i(u), i = 1, 2, \ldots$, we can reduce their dimension and extract the crucial information by projecting them on the principal functions. The projections (or called PC-scores) corresponding to the largest d eigenvalues are obtained by

$$\boldsymbol{\xi}_{ik} = \int_0^1 \boldsymbol{X}_i(u) v_k(u) du, k = 1, \ldots, d.$$

If the multichannel profile sample $\boldsymbol{X}_i(u)$ is from the in-control process, $\boldsymbol{\xi}_{ik}$ are uncorrelated with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_k$. In particular, if $\boldsymbol{X}_i(u)$ follows a Gaussian process, the PC-scores $\boldsymbol{\xi}_{ik} \overset{i.i.d}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma}_k)$. With above projections, one can see $\boldsymbol{\xi}_{ik}, k = 1, \ldots, d$ are ideal indicators, which have relatively low dimension yet are capable of capturing the variation information of the multichannel profile. As a result, we can define an EWMA sequence based on $\boldsymbol{\xi}_{ik}$ as

$$\boldsymbol{\eta}_{ik} = (1 - w)\boldsymbol{\eta}_{i-1,k} + w\boldsymbol{\xi}_{ik}, k = 1, \ldots, d, \tag{3}$$

where the initial vector $\boldsymbol{\eta}_{0k}$ is usually taken to be $\mathbf{0}$ of p dimension, and $0 < w \leq 1$ is the smoothing constant. The proposed control chart triggers a signal if

$$Q_i = \frac{2 - w}{w} \sum_{1 \leq k \leq d} \boldsymbol{\eta}_{ik}^{\mathsf{T}} \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\eta}_{ik} > L, \tag{4}$$

where $L > 0$ is the control limit chosen to achieve a specified IC average run length (ARL_0). Note that the weighted average in (3) reflects that: the more recent observations offer more information for identifying the change and thus have higher weights. Hereinafter, this chart is referred to as principal components based EWMA chart, or PCEWMA for brevity. In what follows, we summarize some useful properties of PCEWMA chart:

(1) When the process is IC, under some mild conditions, the asymptotic distribution of the PCEWMA statistics Q_i is χ^2_{pd} (as $w \to 0$). The proof of this result is straightforward based on the proof of Proposition 3 in Zou and Tsung (2011) and is omitted here.

(2) The sequence $\{Q_1, \ldots, Q_i\}$ forms a Markov chain if the underlying distribution of $\boldsymbol{\xi}_{ik}$ follows the class of elliptical distributions.

These results are particularly useful in determining the control limit L, whose details are presented in the Appendix. To sketch the idea, we represent the charting statistics as

$$Q_i = \boldsymbol{\zeta}_i^{\mathsf{T}} \tilde{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\zeta}_i,$$

where $\boldsymbol{\zeta}_i = (\boldsymbol{\eta}_{i1}^{\mathsf{T}}, \ldots, \boldsymbol{\eta}_{id}^{\mathsf{T}})^{\mathsf{T}}$ and $\tilde{\boldsymbol{\Sigma}} = \frac{w}{2-w}\mathrm{diag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_d)$. Under this representation, $\boldsymbol{\zeta}_i$ is a pd-dimension vector, while the form of the statistics Q_i is analogous to that of MEWMA statistics with dimension pd used by Zou et al. (2007). That is, similar to MEWMA, we can easily prove that PCEWMA statistics Q_i forms a Markov chain. As a result, the IC ARL of PCEWMA can be approximated by using Markov chain method which is developed by Runger and Prabhu (1996). Table 1 provides the control limits of PCEWMA chart for commonly used combinations of w, d and IC ARLs under $p = 4$. The limits are obtained using a Markov chain with $m = 200$ transition states, and all the parameters are assumed known. We have conducted simulations to verify the accuracy of the Markov chain approximation, and the results are very satisfactory as long as $m > 50$. The simulation results shown in the next section demonstrate that the IC run-length performance of PCEWMA is quite robust under various process distributions and therefore, the control limits tabulated in Table 1 can be used in practical applications.

Like the MEWMA chart, the w in PCEWMA should be chosen to balance the robustness to non-normality and the detection ability to various shift magnitudes (c.f., Stoumbos and Sullivan 2002). In general, a smaller w leads to a quicker detection of smaller shifts (c.f., e.g., Lucas and Saccucci 1990; Prabhu and Runger 1997). This statement is still valid for PCEWMA. Based on the simulation results in Section 3, we suggest choosing $w \in [0.05, 0.2]$, which has satisfactory performance in practice.

2.3 Guidelines for Design and Implementation

In Section 2.2, it is assumed that the IC parameters, i.e., the mean function $\boldsymbol{\mu}(u)$, the eigenfunctions $v_k(u)$ and the covariance matrix $\boldsymbol{\Sigma}_k$ are known or, equivalently, that they are

Table 1: The control limits of the PCEWMA chart for IC ARL=200, 370 and 500 with various values of d and w under $p = 4$

IC ARL	w	$d=2$	$d=3$	$d=4$	$d=5$	$d=7$	$d=10$	$d=15$	$d=20$
200	0.3	21.396	27.707	33.643	39.343	50.286	65.988	91.074	115.363
	0.2	20.867	27.130	33.025	38.689	49.567	65.185	90.154	114.400
	0.1	19.541	25.659	31.430	36.983	47.666	63.068	87.762	111.250
	0.05	17.713	23.587	29.143	34.475	44.569	58.722	80.991	101.946
370	0.3	23.131	29.634	35.734	41.579	52.775	68.798	94.326	118.992
	0.2	22.678	29.147	35.217	41.036	52.184	68.145	93.575	118.200
	0.1	21.515	27.874	33.850	39.583	50.575	66.334	91.445	115.124
	0.05	19.894	26.064	31.845	37.315	47.382	61.057	82.457	102.840
500	0.3	23.958	30.548	36.722	42.634	53.945	70.115	95.845	120.682
	0.2	23.536	30.098	36.247	42.135	53.405	69.521	95.159	119.947
	0.1	22.439	28.904	34.969	40.782	51.907	67.802	92.976	116.560
	0.05	20.903	27.191	33.041	38.514	48.437	61.840	82.909	103.104

well estimated from a sufficiently large reference dataset. This section provides guidelines on how to obtain the reliable estimations of these parameters for Phase II analysis.

Intuitively, given the IC sample size m_0, the mean function $\boldsymbol{\mu}(u)$ can be estimated by $\bar{\boldsymbol{X}}(u) = m_0^{-1}\sum_{i=1}^{m_0}\boldsymbol{X}_i(u)$. Further, we can get the estimation of the covariance function $\mathbf{c}(u,u') = E\{\langle \boldsymbol{Y}_i(u), \boldsymbol{Y}_i(u')\rangle\}$ by the method of moments as

$$\tilde{\mathbf{c}}(u,u') = \frac{1}{m_0}\sum_{i=1}^{m_0}\langle\{\boldsymbol{X}_i(u) - \bar{\boldsymbol{X}}(u)\}, \{\boldsymbol{X}_i(u') - \bar{\boldsymbol{X}}(u')\}\rangle. \tag{5}$$

Consequently, the corresponding estimators of eigenvalues λ_k and eigenfunctions $v_k(\cdot)$ denoted by $\tilde{\lambda}_k$ and $\tilde{v}_k(\cdot)$, respectively, are calculated by decomposing

$$\int_0^1 \tilde{\mathbf{c}}(u,u')\tilde{v}_k(u')du' = \tilde{\lambda}_k\tilde{v}_k(u), \quad u \in \mathcal{T}, \quad k = 1,2,\ldots.$$

Furthermore, the covariance matrix of $\boldsymbol{\xi}_{ik}$, i.e., $\boldsymbol{\Sigma}_k$ can be estimated by

$$\tilde{\boldsymbol{\Sigma}}_k = \frac{1}{m_0}\sum_{i=1}^{m_0}\int_0^1\{\boldsymbol{X}_i(u) - \bar{\boldsymbol{X}}(u)\}\tilde{v}_k(u)du\int_0^1\{\boldsymbol{X}_i(u) - \bar{\boldsymbol{X}}(u)\}^{\mathsf{T}}\tilde{v}_k(u)du.$$

Under some mild conditions, the foregoing estimators are consistent.

In practice, a functional profile is always measured at an ordered dense grid of points within an interval of finite length (Ramsay and Silverman 2005). That is, each observation

$\boldsymbol{X}_i(u)$, is only observed at points $u_{ij} : j = 1, \cdots, n_i$. If u_{ij} are the same across different observations, i.e., $u_{ij} = u_j$ and $n_i = n$, we can solve the MFPCA by applying traditional PCA to the $n \times p$ matrix of all observed data. To be more specific, instead of the the covariance function $\tilde{\mathbf{c}}(u, u')$ obtained in (5), we can obtain the covariance matrix $\tilde{\mathbf{C}} = (\tilde{C}_{hl})_{n\times n}$, which are discrete evaluations of $\tilde{\mathbf{c}}(u, u')$ at n points $u_1, \ldots, u_n$. Its (h, l)th component can be estimated by

$$\tilde{C}_{hl} = \frac{1}{m_0} \sum_{i=1}^{m_0} \sum_{j=1}^{p} \{X_{ij}(u_h) - \bar{X}_j(u_h)\}\{X_{ij}(u_l) - \bar{X}_j(u_l)\}, \tag{6}$$

where $\bar{X}_j(u_l) = m_0^{-1} \sum_{i=1}^{m_0} X_{ij}(u_l)$. Then we can decompose $\tilde{\mathbf{C}}$ directly using PCA and obtain the corresponding estimators of eigenvalue $\tilde{\lambda}_k$ and eigenfunctions $\tilde{v}_k(u)$ evaluated at $u_1, \ldots, u_n$. Hence, the (h, l)th element of $\tilde{\boldsymbol{\Sigma}}_k$, $\tilde{\sigma}_{khl}$, can be estimated by

$$\tilde{\sigma}_{khl} = \frac{1}{m_0} \sum_{i=1}^{m_0} \left[\sum_{j=1}^{n} \{X_{ih}(u_j) - \bar{X}_h(u_j)\}\tilde{v}_k(u_j) \sum_{j=1}^{n} \{X_{il}(u_j) - \bar{X}_l(u_j)\}\tilde{v}_k(u_j) \right]. \tag{7}$$

If the sampling grid is sparse or the grid points are unequally spaced, one can smooth the individual profile data first by using any smoothing techniques such as splines or kernel regression methods, and then apply the method above on the predicted (interpolated) values at an equally-spaced grid of points. See Ramsay and Silverman (2005) for some detailed discussion. In what follows, we provide some guidelines on the implementation of the methods in practice.

On determining Phase I sample size m_0. It should be pointed out that when IC sample size m_0 is not large (say, $m_0 \leq 2000$; see Table 3 in Section 3), there would be considerable uncertainty in the parameter estimation, which in turn would distort the IC run length distribution of PCEWMA control chart. From the results in Section 3, we can see that, as long as $m_0 \geq 2000$, the ARL_0 values are quite stable in various cases. Hence, we suggest collecting at least 2000 IC samples before Phase II process monitoring. When a sufficiently large reference dataset is unavailable, self-starting methods that handle sequential monitoring by simultaneously updating parameter estimates and checking for OC signals have been developed accordingly for conventional SPC methods (see, e.g., Hawkins 2007, 2008). The further studies, including thorough discussion of the effect of m_0 on the PCEWMA scheme and the corresponding self-starting charts, are subjects of future research.

On choosing d. Note that d is the number of eigenfunctions $v_k(u)$ used for projection, or equivalently the number of principal components considered. In Phase II monitoring, we also

assume d is unchanged and has been well estimated in Phase I. There are many approaches proposed in the literature, with respect to the choice of d. One approach, which is adopted in our numerical study, is to determine d based on the percentage of total variation explained by the extracted PC-scores. Another approach is to use the pseudo Akaike information criterion (AIC) and the cross-validation procedure (see Yao et al. 2005) to determine d.

On grid points n. In practice each profile is measured at a set of grid points. The number and positions of the grid points for Phase I and Phase II should be properly chosen to describe the functional curve well. We strongly suggest choosing the same number and positions for Phase I and Phase II samples, i.e., $u_{ij} = u_j, j = 1, \ldots, n$. In this case, the PC-scores $\boldsymbol{\xi}_{ik}$ can be directly approximated by $\boldsymbol{\xi}_{ik} \approx \sum_{j=1}^{n} \boldsymbol{X}_i(u_j)v_k(u_j), k = 1, \ldots, d$. Based on extensive simulations, we recommend the number n of grid points in each multichannel profile should not be fewer than 50. Engineers can also take the related engineering knowledge into consideration, such as the smoothness and variation of the profile curves, to determine the appropriate n.

2.4 Diagnostic Aids in Multichannel Profile Monitoring

In SPC, besides quickly detecting a process change, it is also critical to determine when the change occurs after an OC signal is triggered. A diagnostic aid to locate the possible change point and to isolate the type of changes in the profile can help an engineer quickly and easily identify the root cause of the problem. In this section, we discuss the diagnosis of multichannel profiles.

Identifying the location of the change point is the critical step in our diagnostic procedure. Because the change point separates the OC profiles from IC profiles, locating the change point can make an accurate inference about features of a particular change. To identify the change point, a distance based principal components is used. Here we assume that an OC signal is triggered at the mth profile, by the PCEWMA chart. Our suggested estimator of the change point, τ, is given by

$$\hat{\tau} = \arg \max_{1 \le l < m} D(l, m). \tag{8}$$

where $D(l, m) = \sum_{1 \le k \le d} \hat{\boldsymbol{\zeta}}_{lk}^T \boldsymbol{\Sigma}_k^{-1} \hat{\boldsymbol{\zeta}}_{lk}$ is the sample distance between two groups separated at the lth sample, and

$$\hat{\boldsymbol{\zeta}}_{lk} = \sqrt{\frac{l(m-1)}{m}} \int_0^1 \left\{ \frac{1}{l} \sum_{i=1}^{l} \boldsymbol{X}_i(u) - \frac{1}{m-l} \sum_{i=l+1}^{m} \boldsymbol{X}_i(u) \right\} v_k(u) du.$$

Intuitively, if l is the change point, the distance $D(l, m)$ should be large. The foregoing formulation is similar to that of Zou et al. (2008), which used a nonparametric maximum likelihood estimator. Under certain conditions, we can establish the consistency of the change point $\hat{\tau}$, i.e., $|\hat{\tau} - \tau| = O_p(1)$.

After detecting the change point, an engineer may care about the following problems: which channel(s) of the profiles changes; what the OC mean function looks like, and how great the difference between IC and OC models is. Using existing testing methods to address all the issues seems inappropriate and infeasible. A closely related work is Zou et al. (2011), which constructed a BIC criterion by generalizing the best-subset searching procedure. However, since few OC samples are available before the chart signals an alarm, the effectiveness of their method can hardly be guaranteed. We suggest that engineers plot the functional curves of the average of $(m - \hat{\tau})$ multichannel profile samples and compare it with the IC profile together, to offer a visual and practical interpretation of the aforementioned problems.

2.5 Extension of PCEWMA

Here, we discuss a possible extension of the PCEWMA chart to demonstrate its versatility. In practice, shifts in the mean function will influence the correlation among the channels even if $\boldsymbol{\Sigma}_k$ remains at its IC value. To detect the changes in the covariance matrix, we can extend the multivariate exponentially weighted moving covariance matrix (MEWMC) proposed by Hawkins and Maboudou (2008) to monitor the PC-scores $\boldsymbol{\xi}_{ik}$ to construct control chart. When the process is in control, the PC-scores are random vectors with $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_k$. In particular, they follow the multinormal distribution when $\boldsymbol{X}_i(u)$ follows a Gaussian process. As a result, we can build an EWMA sequence based on $\boldsymbol{\xi}_{ik}$ by

$$\boldsymbol{S}_{ik} = (1 - w)\boldsymbol{S}_{i-1,k} + w\boldsymbol{\xi}_{ik}\boldsymbol{\xi}_{ik}^{\top}, \quad k = 1, \ldots, d$$

with $\boldsymbol{S}_{0k} = \boldsymbol{\Sigma}_k$. When the process is in control, $E(\boldsymbol{S}_{ik}) = \boldsymbol{\Sigma}_k$. Here, we apply the likelihood ratio statistic to monitor the stability of the process using the statistics

$$T_i = \sum_{k=1}^{d} \{\mathrm{tr}(\boldsymbol{S}_{ik}) - \log|\boldsymbol{S}_{ik}| - p\}$$

The proposed control chart for covariance matrix triggers a signal if $T_i > L_c$, where the control limit L_c is chosen to achieve a specified IC ARL. It is noting that to obtain the appropriate value of L_c is not an easy task, because the statistics T_i does not have a known

standard distribution. Generally, we need to obtain L_c by simulations given design parameters such as w, IC ARL.

It is often a good practice to simultaneously monitor the changes in both location and scatter. We suggest constructing the EWMA scheme for covariance matrix in conjunction with a PCEWMA to detect multichannel profile changes.

3 Numerical Performance Comparison

In this section, we evaluate the performance of this new scheme PCEWMA in detecting the shift of multichannel profile model (2) through ARL comparisons. It is challenging to compare the proposed method with alternative methods, since there is no obvious comparable method in the literature. To evaluate the performance of the proposed PCEWMA, we follow three related approaches for comparisons.

The first approach to monitor multichannel profiles is to stack up profiles from each channel and transform them into a high-dimensional vector. One could then use the PCA to the resulting vector and extract features to build a EWMA control chart. We refer this method as Vectorized-PCA EWMA (VPEWMA). There are several issues in using VPCA for the analysis of multichannel profiles. As mentioned before, this method breaks the correlation structure in the original data, and loses the useful representations that can be obtained in the original form. In addition, the efficiency of VPEWMA decreases since too many grid points increase the dimension and accumulate large stochastic noise. See also Paynabar et al. (2013) for related discussions. Another FPCA-based method is to use standard FPCA on each individual channel to construct EWMA statistics, and to sum them up to obtain the monitoring statistics. This method is referred to as sIPEWMA. However, this approach fails to consider the correlation among the multichannel profiles. We also consider the nonparametric method NEWMA (Zou et al. 2008), which was proposed for single-channel functional profile. To adapt it to multichannel profile data, one could apply NEWMA on profile from each channel and sum up their statistics. This method is referred to as mNEWMA for brevity. It should be emphasized that NEWMA method assumed the random noises are i.i.d normal at every grid point. This assumption neglects the continuity of the functional curve, and might not be reasonable in practice. Furthermore, it does not consider the correlation among profile channels either.

For simplicity and consistency with the literature, the change point is assumed to be

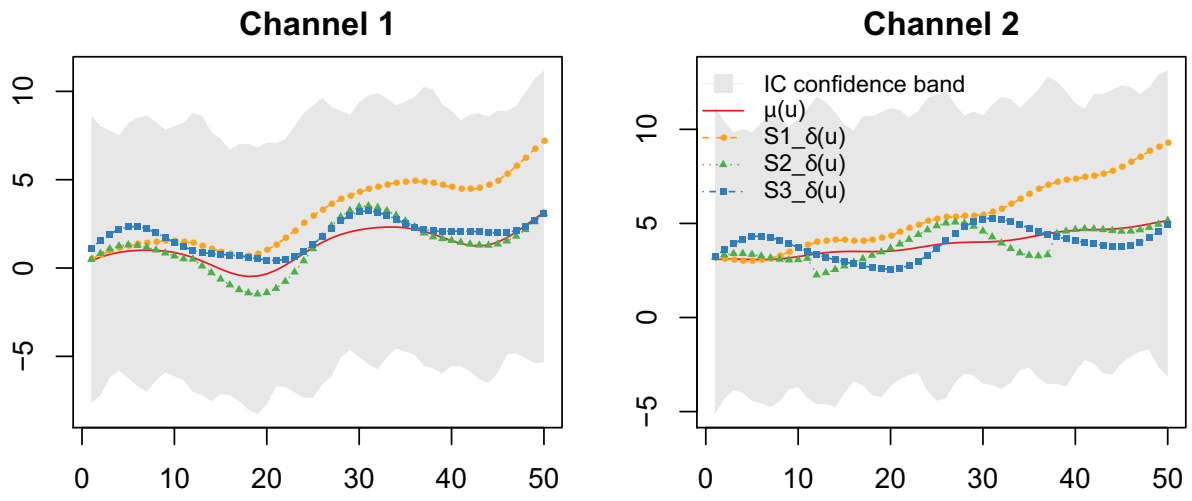

Figure 2: The averages of 1000 curves of the first two changed channels generated from IC Model and three OC Scenarios along with $\gamma = 1$, where the gray band is 95% IC confidence band.

$\tau = 0$ for OC scenarios, and only cases with IC $\text{ARL}_0 = 200$ are reported here. All the results in this section are evaluated with 1000 replications. The number of profile channels is set to $p = 4$ firstly. Following the general IC model $\boldsymbol{X}_i(u) = \boldsymbol{\mu}(u) + \boldsymbol{Y}_i(u)$, where $u \in [0, 1]$, the IC mean functions are $\mu_1(u) = u + 2u^2 + \sin(4\pi u)$, $\mu_2(u) = 2u + 3\exp(-u)$ and $\mu_3(u) = \mu_4(u) = 0$. We consider $\boldsymbol{Y}_i(u) = \sum_{k=1}^{4} \boldsymbol{\xi}_{ik} \upsilon_k(u)$, where $\upsilon_k(u)$'s are the first four non-constant Fourier basis functions with a base period of 0.5, and $\boldsymbol{\xi}_{ik}$ are p-dimensional multivariate normally distributed vectors with mean zero and covariance $(\boldsymbol{\Sigma}_k)_{ij} = k(\sigma)^{|i-j|}$. The default $\sigma = 0.8$. In multichannel profile data, the size of covariance $\boldsymbol{\Sigma}_k$ corresponds to the inter-relationship among the p profile channels.

The number and variety of OC models are too large to allow an all-encompassing and comprehensive comparison. We only choose three representative shift scenarios for illustration, where γ is the magnitude of the shift. In all following OC scenarios, we set $\delta_3(u) = \delta_4(u) = 0$.

- *Scenario 1:* $\delta_1(u) = \gamma(3u + u^2)$, $\delta_2(u) = \gamma(u + 3u^2)$ for $u \in [0, 1]$.
- *Scenario 2:* $\delta_1(u) = \gamma \sin(4\pi u)$, $\delta_2(u) = \gamma \cos(4\pi u)$ for $u \in [1/4, 3/4]$.
- *Scenario 3:* $\delta_1(u) = \gamma exp(-u)$, $\delta_2(u) = \gamma \sin(4\pi u)$ for $u \in [0, 1]$.

All the processes are sampled on a grid of $n = 50$ equispaced points in $\mathcal{T} = [0, 1]$. The averages of 1000 curves generated from true IC models and three OC curves with $(\gamma = 1)$ are depicted in Figure 2. We can observe that the average curves could capture the pattern

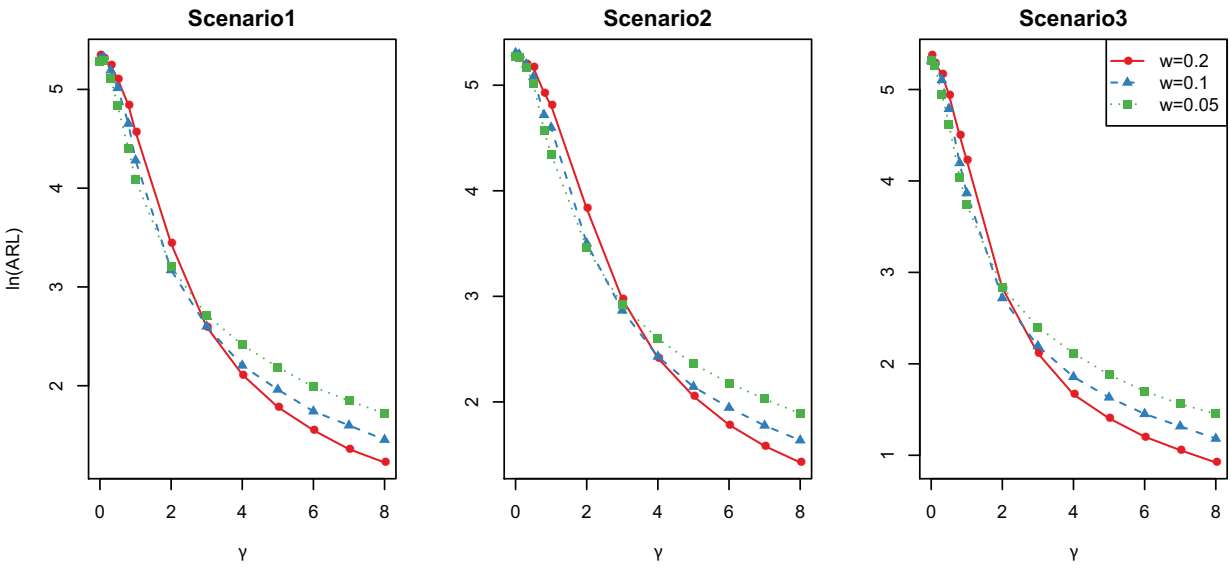

Figure 3: OC ARL comparison of PCEWMA scheme with different smoothing parameters w under Scenario 1-3.

of true functions. It should be emphasized that it is difficult to obtain control limits for VPEWMA, mNEWMA and sIPEWMA methods for a given ARL_0. Therefore, all control limits are obtained through simulations, which makes them inapplicable in practice. For mNEWMA chart, the bandwidth h, is chosen from (10) in Zou et al. (2008) with $c = 1.0$.

Before we compare the OC ARLs of our method with the others, we first study the effect of the smoothing parameter w on the performance of PCEWMA. In Figure 3, the OC ARL of PCEWMA under three scenarios with different smoothing parameters $w = 0.2, 0.1$ and 0.05 are compared. Form this figure, PCEWMA scheme with $w \leq 0.2$ performs well under different models. In particular, the PCEWMA scheme with larger w is superior to the ones with smaller w in detecting large shifts, while the PCEWMA scheme with smaller w is better than the ones with larger w in detecting small shifts. This result is consistent with the analysis in Section 2.

Since VPEWMA scheme breaks the natural correlation among profiles from different channels, while mNEWMA and sIPEWMA schemes entirely ignore the correlation, it is important to study the influence of correlations on the OC performance. Figure 4 summarizes the ARL values with high-level ($\sigma = 0.8$) and low-level ($\sigma = 0.4$) correlation in the covariance matrix $(\mathbf{\Sigma}_k)_{ij} = k(\sigma)^{|i-j|}$. From Figure 4, all plots indicate that the proposed PCEWMA scheme has superior efficiency with both strong and weak between-profile correlations. In addition, when σ becomes higher, both mNEWMA and sIPEWMA detect the shifts slower.

VPEWMA scheme has the similar property like mNEWMA and sIPEWMA in Scenario 1 and 2. In contrast, our PCEWMA performs better with higher σ, because our model consider the correlation information among channels effectively.

Then, we turn to Table 2, which compares OC ARL values of the proposed PCEWMA scheme with VPEWMA, mNEWMA and sIPEWMA in various shift sizes when the smoothing parameter w equals 0.2. Besides the ARLs, the corresponding standard deviation of the run lengths (SDRL) are also included in Table 2 to give a broader picture of the run-length distribution. Table 2 shows that our PCEWMA chart has superior efficiency in most cases. Further, we also can find the advantage of PCEWMA becomes larger in detecting larger shift. On the other hand, VPEWMA and sIPEWMA outperform the mNEWMA chart because these two PCA-based methods captures some major information of multichannel profile to some extent. Conversely, the mNEWMA chart is inefficient in detecting shifts in all scenarios. This is because besides ignoring the correlation among channels, mNEWMA chart assumes that the stochastic noise at every observation point is independent from each other. Their direct comparisons with $w = 0.1, \sigma = 0.8$ are summarized in Figure 5, from which we can see that PCEWMA significantly outperforms others in efficiency and sensitivity, especially for large shifts.

As discussed earlier, PCEWMA can also work in monitoring single-channel profiles, i.e. $p = 1$. In this situation, our proposed PCEWMA is equivalent to VPEWMA and sIPEWMA. To demonstrate, we use the first channel profile samples under the previous simulation models to compare the performance of the proposed PCEWMA scheme with NEWMA scheme (see Zou et al. 2008). Figure 6 compares the OC ARLs of PCEWMA and NEWMA charts under different scenarios with smoothing parameter $w = 0.2$ and 0.1. It clearly shows that PCEWMA detects shift faster in all occasions than NEWMA, because of the independent error assumption of NEWMA.

In all the simulations above, it is assumed that the IC parameters are known or, equivalently, that they are estimated from a sufficiently large reference dataset. Here, we study the performance of PCEWMA when this assumption is violated. Table 3 shows the IC ARLs and SDRLs of PCEWMA under various smoothing parameters w when $(\boldsymbol{\mu}(u), \boldsymbol{\Sigma}_k, v_k(u))$ are estimated from IC dataset with various historical sample sizes, m_0. In each simulation, a sample of m_0 is firstly generated and the IC parameters are estimated based on this sample. Then, an independent sequence of p multichannel profiles is generated and the correspond-

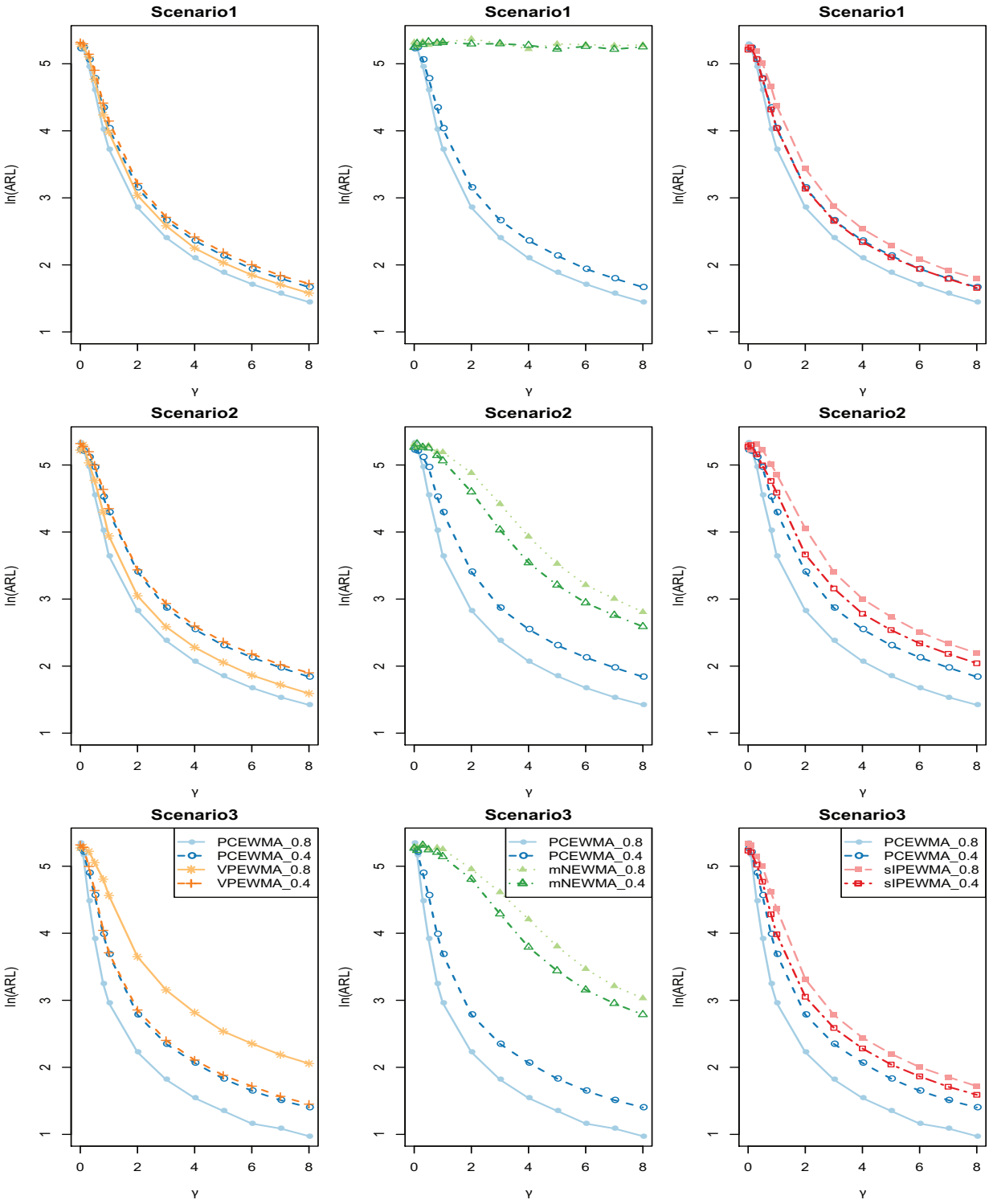

Figure 4: ARL values for PCEWMA vs VPEWMA (first column), PCEWMA vs mNEWMA (second column) and PCEWMA vs sIPEWMA (third column) with $\sigma = 0.8$ and 0.4, $w = 0.05$ under Scenario 1-3.

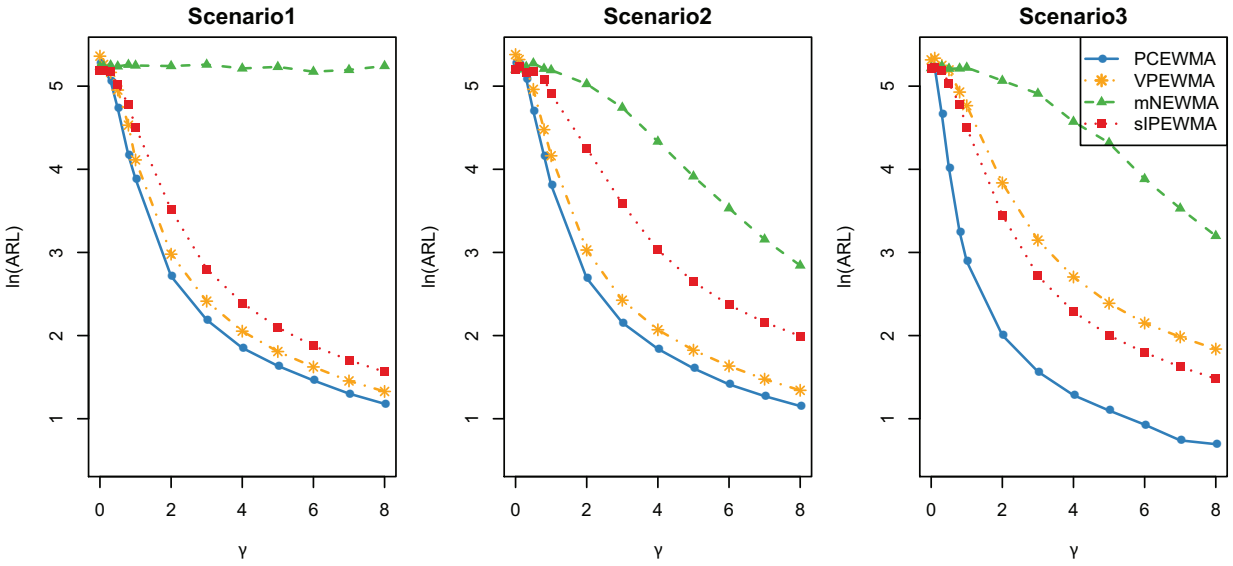

Figure 5: OC ARL values for PCEWMA, VPEWMA, mNEWMA and sIPEWMA with $\sigma = 0.8$ and $w = 0.1$ under Scenario 1-3.

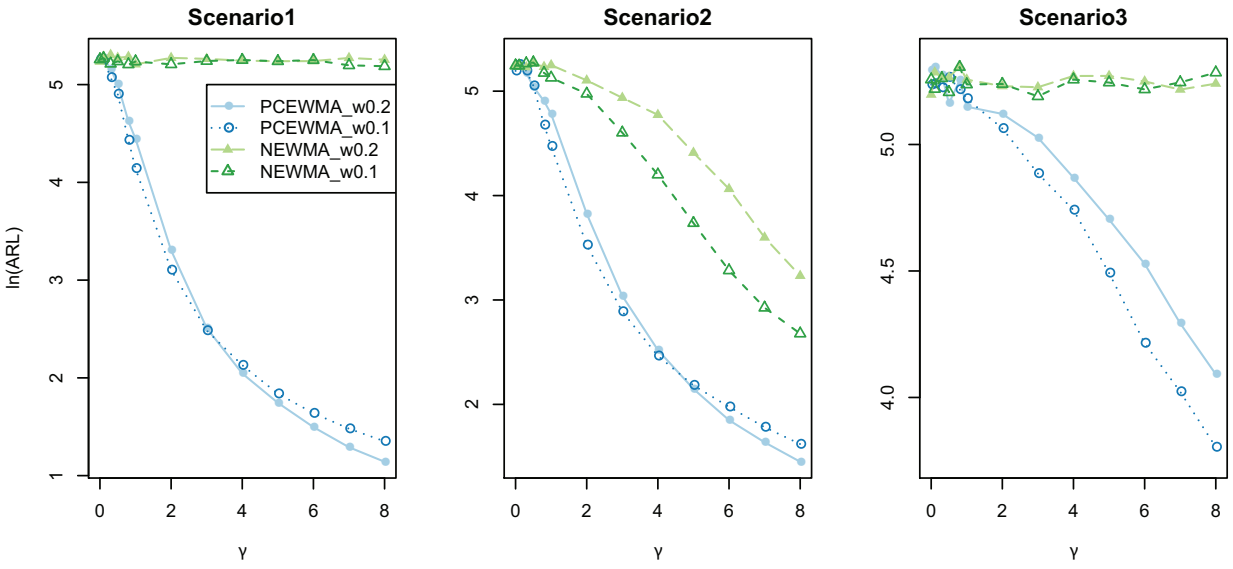

Figure 6: OC ARL values for PCEWMA and NEWMA with $p = 1$ and $w = 0.2$ under Scenario 1-3.

Table 2: OC ARL and SDRL values for PCEWMA, VPEWMA, mNEWMA and sIPEWMA with various shifts of size γ, $w = 0.2$ and $\text{ARL}_0 = 200$ under Scenario 1-3.

		PCEWMA		VPEWMA		mNEWMA		sIPEWMA	
Scenario	γ	ARL	SDRL	ARL	SDRL	ARL	SDRL	ARL	SDRL
(1)	0	198	176	191	172	204	184	203	182
	0.1	193	169	200	177	210	191	188	171
	0.3	172	168	191	174	203	185	189	179
	0.5	138	126	153	142	197	179	171	158
	0.8	88.7	80.0	108	102	206	188	146	138
	1	65.8	58.3	80.6	73.8	203	190	123	112
	2	16.7	10.6	23.3	17.3	194	178	49.6	41.5
	4	5.3	1.7	6.6	2.5	196	180	11.4	6.1
(2)	0	199	186	204	189	208	188	199	177
	0.1	190	172	206	185	203	189	191	173
	0.3	166	153	183	173	215	193	196	181
	0.5	143	133	165	159	204	187	193	178
	0.8	89.1	88.7	104	98.7	203	183	187	175
	1	65.8	58.8	87.2	77.9	202	186	161	150
	2	16.2	10.2	25.2	19.0	185	172	105	101
	4	5.2	1.5	6.9	2.7	125	118	29.4	23.2
(3)	0	198	177	197	179	199	177	193	177
	0.1	190	173	197	185	208	195	198	182
	0.3	126	116	186	172	214	193	191	176
	0.5	76.1	72.7	173	166	205	193	169	160
	0.8	32.5	25.7	149	142	204	191	148	144
	1	21.1	13.7	139	132	191	182	127	120
	2	6.2	2.2	61.2	52.2	187	177	49.3	40.0
	4	2.8	0.6	16.0	9.7	141	130	10.2	4.6

ing run lengths are obtained. From Table 3, we can find that when the IC sample size is relatively small, the actual IC ARLs and SDRLs of the PCEWMA chart are both quite far away from the nominal value 200; and as the IC sample size increases, the actual results are closer to their nominal values. In this paper, we use $m_0 = 50000$ IC samples to estimate the IC parameters for the PCEWMA scheme in aforementioned simulations.

We also conducted simulations with various combinations of ($\mathbf{\Sigma}_k$,$v_k(u)$,p,w) and other OC settings. These simulation results, not reported here but available from the authors upon request, also show the proposed PCEWMA chart works well. R codes for implementing the proposed procedure are available in the Supplementary Material.

Table 3: IC ARL and SDRL values of PCEWMA scheme with various Phase I sample sizes m_0.

	$w = 0.2$		$w = 0.1$		$w = 0.05$	
m_0	ARL	SDRL	ARL	SDRL	ARL	SDRL
50	56.2	50.8	70.5	60.3	86.6	69.6
100	105	100	117	105	130	109
200	135	125	143	129	157	134
300	152	144	163	145	167	141
400	164	151	169	154	174	149
500	170	157	170	150	179	151
750	182	166	185	164	185	158
1000	183	170	185	166	191	162
1500	185	170	190	168	189	160
2000	186	170	186	166	190	161
4000	196	178	196	175	196	167
10000	196	180	196	172	196	167
50000	200	181	197	176	197	168

4 Case study

We revisit the multi-operation forging process presented in Section 1 and use this example to illustrate how to implement the proposed method step by step in practice. In this process, a forging machine (shown in the left panel of Figure 1) is comprised of multiple dies, each assigned to perform one operation during a stroke. Tonnage forces exerted on all dies are measured by four strain gauge sensors mounted on four columns of the press. In each cycle of operation, four-channel tonnage profiles are recorded with length 1200 (See the right panel of Figure 1). To reduce the measurement noise of profile observations in each channel, we first apply the non-overlapping moving average function with the window size of 6. After smoothing the length of each profile channel becomes 200. A sample of 526 multichannel profiles was collected under different experimental settings. The sample includes 457 in-control profiles collected under the normal production condition, and a group of 69 OC (faulty) profiles. These OC profiles are measured when one part is missing in piercing station shown in the first panel of Figure 7. To illustrate, the average profiles of two groups are shown in the second panel of Figure 7. It is easily found that profile samples corresponding to piercing fault and the normal operation are very similar, thus difficult to monitor. More detailed discussion about this example can be found in Lei et al. (2010).

In this case study, we focus on constructing a Phase II monitoring method and thus, we

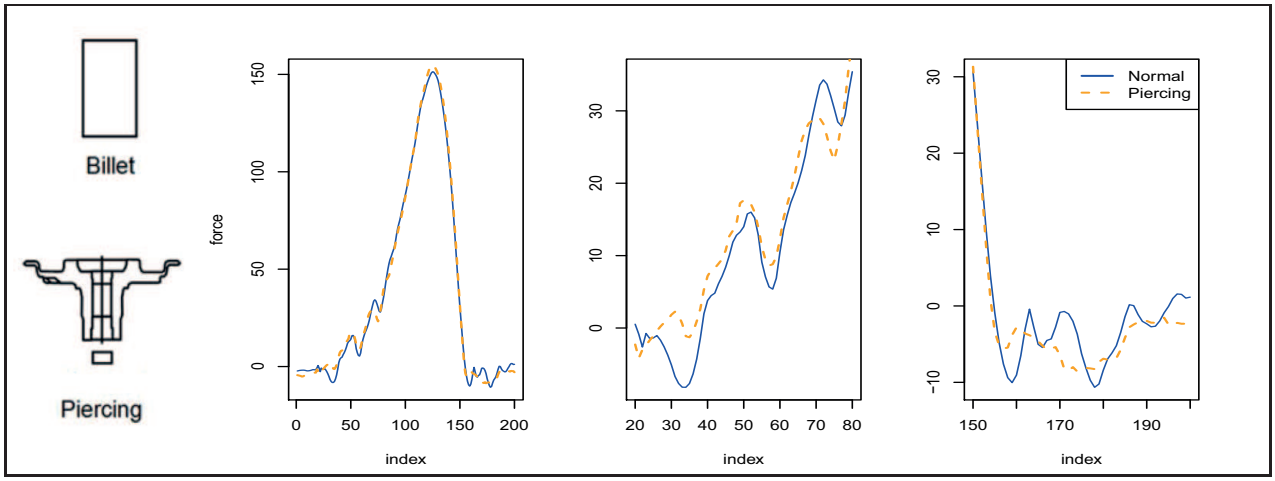

Figure 7: (1): Shape of workpieces at normal and piercing operation. (2): Average profiles of aggregated tonnage profiles for normal and piercing operations. (3): Average profiles zoomed-in between 20 and 80. (4): Average profiles zoomed-in between 150 and 200.

separate the IC profile samples to estimate the related parameters. Detailed implementation steps are as follows.

Step 1. Compute the related IC parameters. Here we apply the MFPCA to $m_0 = 457$ IC profile samples and calculate the mean function $\boldsymbol{\mu}(u)$ and the first $d = 16$ eigenfunctions $v_k(u)$ whose corresponding eigenvalues account for more than 85% of the profiles' total variation. Because profiles are measured at $n = 200$ discrete points, we apply formula (7) to estimate the corresponding covariance matrix $\boldsymbol{\Sigma}_k, k = 1, \ldots, d$.

Step 2. Choose the desired IC ARL and the smoothing constant, w. Determine the control limit, L, based on p, d, IC ARL and w. Given $p = 4$, $d = 16$, $ARL_0 = 200$ and $w = 0.05$, the L is found to be 85.28. Then we can construct the PCEWMA control chart as demonstrated in Figure 8.

Step 3. Start monitoring the process. To illustrate the validity of our method, we set the change point $\tau = 20$. That is the first 20 profiles are randomly sampled from the IC dataset, and subsequent profiles are randomly sampled from OC dataset. From Figure 8, we can see that the PCEWMA chart quickly detect the changes at the 24th sample and Q_i remains above the control limit.

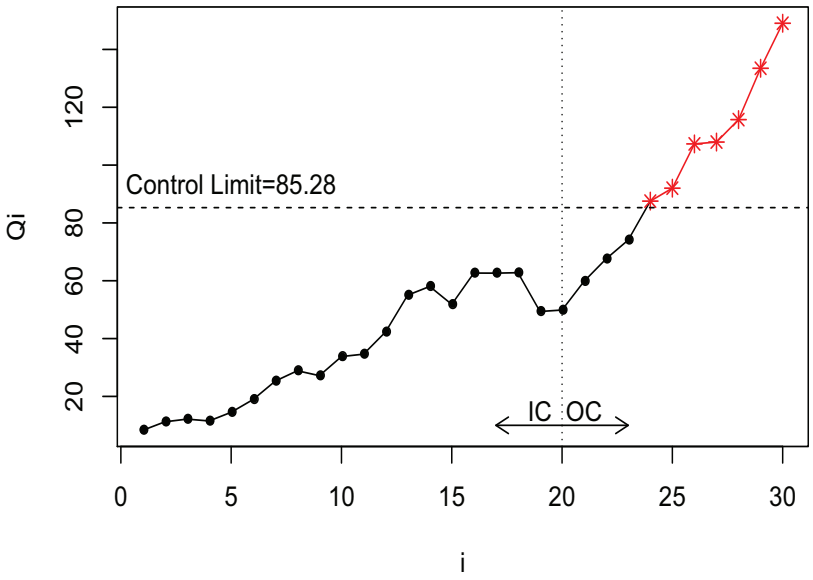

Figure 8: PCEWMA chart for the forging process.

5 Concluding Remarks

Phase II monitoring for multichannel profiles is a challenging problem and it has not been thoroughly investigated in the literature. We proposed an approach that combines the classical EWMA procedure with the MFPCA. Our approach fully incorporates the correlation information of multichannel profiles for monitoring. The proposed PCEWMA scheme can be easily designed and constructed, and it has satisfactory performance. We also suggested a Markov chain approach to determine the control limit of the proposed chart. A diagnostic procedure was then developed, which was capable of identifying the change-point after a process change was detected. In addition to the theoretical results, various simulation studies were conducted to validate the proposed methodology. Through simulation, we also showed that the proposed monitoring scheme PCEWMA outperformed conventional approaches in terms of the OC ARL for monitoring both multichannel and single-channel profile observations. We applied PCEWMA to multichannel tonnage profiles for Phase II monitoring in a multi-operation forging process to demonstrate the implementation of our proposed methodology in practice.

Extensions of the proposed MFPCA-based method to the non-normal or autocorrelated data could be a worthwhile and necessary future contribution to the existing literature (Capizzi and Masarotto 2008). In addition, our proposal only handles the case of "sustained" shifts in the process mean. Patterned, oscillatory shifts in mean as well as in variance may

happen in practice and thus it is worth investigation in Phase II profile monitoring for these types of changes.

Supplementary Material

The supplementary material contains R codes.

Acknowledgments

The authors thank Professor Paynabar K. for sharing the dataset used in the case study.

Appendix: IC ARL Calibrations of PCEWMA via a Markov Chain Model

Without loss of generality, we assume that $\boldsymbol{Y}_i(u)$ is a Gaussian process, that is, $\boldsymbol{\xi}_{ik}$ are iid standard p-dimensional multinormal variables, which will facilitate the derivation of the transition probabilities as shown later. The Markov chain model can be regarded as an extension of Runger and Prabhu (1996) to PCEWMA. One can refer to Lucas and Saccucci (1990) and Runger and Prabhu (1996) for more details on the Markov chain approximation for the conventional EWMA and MEWMA charts. We only briefly describe the approximation method, but focus more on necessary formulas and modifications.

Following the representation of Q_i in Section 2, we denote $\boldsymbol{\vartheta}_i = (\boldsymbol{\xi}_{i1}^T, \ldots, \boldsymbol{\xi}_{id}^T)^T$, and $\boldsymbol{\vartheta}_i$ comes from a pd-dimensional multinormal distribution with mean 0 and covariance matrix $\frac{2-w}{w}\tilde{\boldsymbol{\Sigma}}$. Furthermore, we can easily obtain $\boldsymbol{\zeta}_i = (1-w)\boldsymbol{\zeta}_{i-1} + w\boldsymbol{\vartheta}_i$. Define the $(m+1)$ by $(m+1)$ transition probability matrix, $\mathbf{P} = (p_{ij})$, where the element p_{ij} is the transition probability from state i to j, and $(m+1)$ is the number of transition states. Denote $g = 2[Lw/(2-w)]^{1/2}/(2m+1)$. Now we have for $i = 0, 1, 2, \ldots$, where m and j are not equal to 0 and denote $\boldsymbol{e}_{pd}$ is the pd component unit vector $\boldsymbol{e}_{pd} = (1, 0, 0, \ldots, 0)^T$, that

$$\begin{aligned} p_{ij} &= Pr\{(j-0.5)g < \|w\boldsymbol{\vartheta}_t + (1-w)\boldsymbol{\zeta}_{t-1}\| < (j+0.5)g | \|\boldsymbol{\zeta}_{t-1}\| = ig\} \\ &= Pr\{(j-0.5)g < \|w\boldsymbol{\vartheta}_t + (1-w)ig\mathbf{U}\| < (j+0.5)g\} \\ &= Pr\{(j-0.5)g/w < \|\boldsymbol{\vartheta}_t + (1-w)ig\boldsymbol{e}_{pd}/w\| < (j+0.5)g/w\}, \end{aligned}$$

where we use the arguments from the proof of Proposition 1 in Runger and Prabhu (1996) that the distribution of $\boldsymbol{\zeta}_{t-1}$ given $\|\boldsymbol{\zeta}_{t-1}\| = ig$ is uniform on $S(\|\boldsymbol{\zeta}_{t-1}\|)$, say as $ig\mathbf{U}$. The

last equality comes from the fact that $\boldsymbol{\vartheta}_t$ and $\mathbf{U}$ are independent. Let $a = [(1-w)ig/w]^2$ and denote a non-central chi-squared random variable with pd degrees of freedom and non-centrality parameter a. Then

$$p_{ij} = Pr\{(j-0.5)^2 g^2/w^2 < \chi^2(pd, a) < (j+0.5)^2 g^2/w^2\},$$

For $j = 0$, we have $p_{i0} = Pr\{\chi^2(pd, a) < 0.5^2 g^2/w^2\}$.

Let $\mathbf{1}$ denote the $m+1$ vector of all ones, and let $\mathbf{I}$ denote the $m+1$ dimensional identity matrix. Let $\boldsymbol{e}_{m+1}$ be a $m+1$ vector with a 1 in the component that corresponds to the starting state of the chain and zeros elsewhere. Finally, the IC ARL can be evaluated by

$$ARL = \boldsymbol{e}_{m+1}^T (\mathbf{I}_{m+1} - \mathbf{P})^{-1} \mathbf{1}.$$

References:

Capizzi, G., and Masarotto, G. (2008), "Practical Design of Generalized Likelihood Ratio Control Charts for Autocorrelated Data," *Technometrics*, 50, 357–370.

Chicken, E., Pignatiello, J. J. Jr., and Simpson, J. (2009), "Statistical Process Monitoring of Nonlinear Profiles Using Wavelets," *Journal of Quality Technology*, 41, 198–212.

Chou, S. H., Chang, S. I., and Tsai, T. R. (2014), "On Monitoring of Multiple Non-linear Profiles," *International Journal of Production Research*, 52(11), 3209–3224.

Dubin, J. A., and Müller, H.-G, (2005), "Dynamical Correlation for Multivariate Longitudinal Data", *Journal of the American Statistical Association*, 100, 872–881.

Hawkins, D. M., and Maboudou-Tchao, E. M. (2007), "Self-Starting Multivariate Exponentially Weighted Moving Average Control Charting," *Technometrics*, 49, 199-?09.

Hawkins, D. M., and Maboudou-Tchao, E. M. (2008), "Multivariate Exponentially Weighted Moving Covariance Matrix," *Technometrics*, 50(2), 155–166.

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.

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(3), 317–328.

Lei, Y., Zhang Z. and Jin, J. (2010), "Automatic Tonnage Monitoring for Missing Part Detection in Multi-Operation Forging Processes," *ASME Transactions, Journal of Manufacturing Science and Engineering*, 132, 051010.1-10.

Lowry, C. A., Woodall, W. H., Champ, C. W., and Rigdon, S. E. (1992), "A Multivariate Exponentially Weighted Moving Average Control Chart," *Technometrics*, 34(1), 46–53.

Lucas, J. M., and Saccucci, M. S. (1990), "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," *Technometrics*, 32(1), 1–12.

Montgomery, D. C. (2007), *Introduction to Statistical Quality Control*, John Wiley & Sons, New York.

Noorossana, R., Eyvazian, M., Amiri, A., and Mahmoud, M. A. (2010), "Statistical Monitoring of Multivariate Multiple Linear Regression Profiles in Phase I with Calibration Application," *Quality and Reliability Engineering International*, 26(3), 291–303.

Noorossana, R., Saghaei, A., and Amiri, A. (2011), *Statistical Analysis of Profile Monitoring*, Wiley, New Jersey.

Paynabar, K., Jin, J., and Pacella, M. (2013), "Analysis of Multichannel Nonlinear Profiles Using Uncorrelated Multilinear Principal Component Analysis with Applications in Fault Detection and Diagnosis," *IIE Transactions*, 45, 1235–1247.

Paynabar, K., Qiu, P., and Zou, C. (2015), "A Change Point Approach for Phase-I Analysis in Multivariate Profile Monitoring and Diagnosis," *Technometrics*, accepted.

Prabhu, S. S., and Runger, G. C. (1997), "Designing a Multivariate EWMA Control Chart," *Journal of Quality Technology*, 29(1), 8–15.

Qiu, P. (2013), *Introduction to Statistical Process Control*, Boca Raton, FL: Chapman & Hall/CRC.

Qiu, P., Zou, C., and Wang, Z. (2010), "Nonparametric Profile Monitoring By Mixed Effect Modeling (with discussion)," *Technometrics*, 52, 265–277.

Ramsay, J. O., and Silverman, B. W. (2005), *Functional Data Analysis*, Springer, New York.

Runger, G. C., and Prabhu, S. S. (1996), "A Markov Chain Model for the Multivariate Exponentially Weighted Moving Averages Control Chart," *Journal of the American Statistical Association*, 91(436), 1701–1706.

Stoumbos, Z. G., and Sullivan, J. H. (2002), "Robustness to Non-normality of the Multivariate EWMA Control Chart," *Journal of Quality Technology*, 34(3), 260–276.

Woodall, W. H. (2007), "Current Research on Profile Monitoring," *Produc̃ao*, 17, 420–425.

Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004), "Using Control Charts to Monitor Process and Product Quality Profiles," *Journal of Quality Technology*, 36(3), 309–320.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005), "Functional Linear Regression Analysis for Longitudinal Data," *The Annals of Statistics*, 33, 2873–2903.

Zou, C. Jiang, W., and Tsung, F. (2011), "A LASSO-Based Diagnostic Framework for Multivariate Statistical Process Control", *Technometrics*, 53, 297–309.

Zou, C., Ning, X., and Tsung, F. (2012), "LASSO-Based Multivariate Linear Profile Monitoring," *Annals of Operations Research*, 192(1), 3–19.

Zou, C., and Tsung, F. (2011), "A Multivariate Sign EWMA Control Chart," *Technometrics*, 53(1), 84–97.

Zou, C., Tsung, F., and Wang, Z. (2008), "Monitoring Profiles Based on Nonparametric Regression Methods," *Technometrics*, 50(4), 512–526.

Zou, C., Tsung, F., and Wang, Z. (2007), "Monitoring General Linear Profiles Using Multivariate EWMA Schemes," *Technometrics*, 49, 395–408.