Nonparametric Profile Monitoring By Mixed Effects Modeling

Peihua Qiu[1], Changliang Zou[2] and Zhaojun Wang[2]

[1]School of Statistics, University of Minnesota

[2]LPMC and Department of Statistics, Nankai University, China

Abstract

In some applications, quality of a process is characterized by the functional relationship between a response variable and one or more explanatory variables. Profile monitoring is for checking the stability of this relationship over time. Control charts for monitoring nonparametric profiles are useful when the relationship is too complicated to be described parametrically. Most existing control charts in the literature are for monitoring parametric profiles. They require the assumption that within-profile measurements are independent of each other, which is often invalid in practice. This paper focuses on nonparametric profile monitoring when within-profile data are correlated. A novel control chart is suggested, which incorporates local linear kernel smoothing into the exponentially weighted moving average (EWMA) control scheme. In this method, within-profile correlation is described by a nonparametric mixed-effects model. Our proposed control chart is fast to compute and convenient to use. Numerical examples show that it works well in various cases. Some technical details are provided in an appendix available online as supplemental materials.

Key Words: EWMA; Local Linear Kernel Smoothing; Nonparametric Mixed-Effects Models; Phase II; Profile Monitoring; Statistical Process Control.

1 Introduction

In certain applications, quality of a process is characterized by the functional relationship between a response variable and one or more explanatory variables. At each sampling stage, one observes a set of data points of these variables that can be represented by a curve (or,

profile). Profile monitoring is mainly for checking the stability of this relationship over time based on observed profile data. In some applications (e.g., certain calibration applications), profiles can be described reasonably well by a linear regression model. But, in some others, more flexible models are necessary for properly describing profiles. This paper focuses on nonparametric profile monitoring when within-profile data are correlated.

As described by Woodall (2000), statistical process control (SPC) can generally be divided into two phases. In Phase I, a set of process data is gathered and analyzed. Any unusual "patterns" in the data lead to adjustments and fine tuning of the process. Once all such assignable causes are accounted for, we are left with a clean set of data, gathered under stable operating conditions and illustrative of the actual process performance. This dataset, which is referred to as the in-control (IC) dataset hereafter, is then used for estimating certain IC parameters of the process. In Phase II SPC, the estimated IC process parameters are used, and the major goal of this phase is to detect any change in the profiles. Performance of a Phase II SPC procedure is often measured by the average run length (ARL), which is the average number of time points needed for the procedure to signal a change in profiles. The IC ARL value of the procedure is often controled at a given level. Then, the procedure performs better if its out-of-control (OC) ARL is shorter, when detecting a specific profile change. In the literature, most SPC control charts are for Phase II process monitoring which is also the focus of the current paper.

In recent years, Phase II profile monitoring has drawn much attention from statisticians. Early research on this topic focuses on linear profile monitoring. See, for instance, Kang and Albin (2000), Kim *et al.* (2003), Mahmoud and Woodall (2004), Zou *et al.* (2006), and Mahmoud *et al.* (2007), among several others. Zou *et al.* (2007) and Kazemzadeh *et al.* (2007) consider cases when profiles can be described well by multiple and/or polynomial regression models. Some recent research concerns nonlinear profile monitoring. For instance, Williams *et al.* (2007a, 2007b) suggest three general approaches to nonlinear profile monitoring and use these approaches for monitoring nonlinear dose-response profiles. Colosimo and Pacella (2007) propose methods for monitoring roundness profiles of some manufactured items. Lada *et al.* (2002) and Ding *et al.* (2006) investigate a general class of nonlinear profiles, using

techniques such as dimension-reduction, wavelet transformations, and independent component analysis. Zou *et al.* (2008) discuss profile monitoring using nonparametric regression methods. A nice overview on profile monitoring can be found in Woodall *et al.* (2004).

In the literature, most existing profile monitoring control charts (e.g., Zou *et al.* 2008) require a fundamental assumption that observations within a profile are independent of each other, which is often invalid in applications. In practice, within-profile data are usually spatially or serially correlated. For instance, within-profile data of the vertical-density profiles (VDPs) considered by Walker and Wright (2002) and Williams *et al.* (2007b) are spatially correlated, since the density measurements are taken in intervals that are close to each other along the vertical depth of a particle board. As another example, within-profile data in the deep reactive ion etching (DRIE) example considered by Zou *et al.* (2007) exhibit obvious serial correlation over time. As demonstrated in the following sections, when within-profile correlation is present, proper setup of the profile model becomes challenging, and estimation of certain IC process parameters becomes difficult as well. If it is ignored in profile monitoring, then the IC and OC properties of the related control charts would be adversely affected.

In the literature, we have not found any existing research on Phase II nonparametric profile monitoring in cases when within-profile data are correlated. Recent articles by Jensen *et al.* (2008) and Jensen and Birch (2009) discuss linear and nonlinear profile monitoring in Phase I analysis, using linear and nonlinear mixed-effects modeling (cf., e.g., Laird and Ware 1982). Their methods can accommodate certain within-profile correlation. But, besides the fundamental difference between the Phase I linear/nonlinear profile monitoring considered in their papers and the Phase II nonparametric profile monitoring considered here, their approaches assume that both the fixed and random effects terms in their models follow certain parametric models, and that the covariance matrix of the random errors also follows a parametric form, such as the autoregressive or compound symmetry form. While parametric methods are useful in certain applications, questions will always arise about adequacy of these parametric model assumptions and about potential impact of model mis-specifications on profile monitoring performance. See Hart(1997) for related discussion. In addition, design

points are assumed to be deterministic (i.e., non-random) in Jensen *et al.* (2008) and Jensen and Birch (2009), and unchanged from profile to profile in Jensen *et al.*. (2008). In practice, however, different profiles often have different design points (i.e., the so-called unbalanced design cases). In some cases, they might even be random (i.e., the random design cases). Phase II profile monitoring in such cases is particularly challenging, which is also discussed in this paper.

To properly describe within-profile correlation, we propose to use a nonparametric mixed-effects model (cf., e.g., Shi *et al.* 1996; Rice and Wu 2001; Wu and Zhang 2002), which allows a flexible variance-covariance structure. Based on estimated variance structure from an IC dataset, we propose a novel Phase II control chart for monitoring nonparametric profiles, which can accommodate within-profile correlation and arbitrary design. Our proposed control chart is based on local linear kernel smoothing of profile data and on the EWMA weighting scheme as well. It incorporates properly both the exponential weights used in the EWMA scheme at different time points and the heteroscedasticity of observations within each profile into the local linear kernel smoother. Numerical results show that this approach performs well in various cases.

Our proposed control chart is described in detail in Section 2. Its numerical performance is investigated in Section 3. In Section 4, we apply this method to a real dataset from a manufacturing process of aluminium electrolytic capacitors. Several remarks conclude the article in Section 5. Some technical details are provided in an appendix, which is available online as supplementary materials.

2 Methodology

This section is organized in five parts. In Section 2.1, nonparametric mixed-effects modeling of an IC dataset is introduced. Its model estimation is discussed in Section 2.2. In Section 2.3, a new Phase II nonparametric profile control chart is proposed, which can accommodate within-profile correlation and arbitrary design. Certain computational issues are addressed in

Section 2.4. Some practical guidelines regarding design and implementation of the proposed control chart are given in Section 2.5.

2.1 Nonparametric Mixed-Effects Modeling

The Phase II nonparametric profile control chart proposed in this paper does not require IC process parameters to be known. Instead, it estimates the related IC process parameters from an IC dataset, using nonparametric mixed-effects modeling. In the literature, mixed-effects modeling is often used in longitudinal data analysis (cf., e.g., Laird and Ware 1982; Diggle *et al.* 1994). It has become a major tool for accommodating possible correlation among observed data. Nonparametric mixed-effects (NME) modeling for analyzing longitudinal data has been discussed by several authors, including Shi *et al.* (1996) and Rice and Wu (2001). Here, we follow this framework for modeling within-profile correlation of an IC dataset. In what follows, we use the term "profile" throughout; but, it should be noted that, in the literature on mixed-effects modeling, it is often referred to as "cluster" or "subject." To simplify presentation, we choose to discuss cases with a single covariate here; this discussion can be easily generalized to cases with multiple covariates. In the IC dataset, assume that there are m profiles and the i-th profile has n_i observations, for $i = 1, 2, \ldots, m$. Then, the NME model can be written as

$$y_{ij} = g(x_{ij}) + f_i(x_{ij}) + \varepsilon_{ij}, \text{ for } j = 1, 2, \ldots, n_i, \ i = 1, 2, \ldots, m, \tag{1}$$

where g is the population profile function (i.e., the fixed-effects term), f_i is the random-effects term describing the variation of the i-th individual profile from g, $\{x_{ij}, y_{ij}\}_{j=1}^{n_i}$ is the sample collected for the i-th profile, and ε_{ij}s are i.i.d. random errors with mean 0 and variance σ^2. In model (1), it is routinely assumed that the random-effects term f_i and the errors ε_{ij} are independent of each other, and f_i is a realization of a mean 0 process with a common covariance function

$$\gamma(x_1, x_2) = E[f_i(x_1) f_i(x_2)].$$

Without loss of generality, we further assume that $x_{ij} \in [0, 1]$, for all i and j.

Model (1) is fairly flexible. It includes many common correlation structures as special cases. For instance, if $f_i(x_{ij}) = \alpha_i$ and α_i is a mean 0 random variable, then within-profile correlation would have the compound symmetry form. If $\text{Corr}(f_i(x_1), f_i(x_2)) = \rho(|x_1 - x_2|; \alpha)$, for some correlation function ρ and a coefficient α, then the correlation structure includes the nonhomogeneous Ornstein-Uhlenbeck process and the Gaussian correlation model (cf., Zhang *et al.* 1998). When the design points are equally spaced and unchanged among different profiles, this model can also be used for describing the autoregressive correlation structure. Because of its flexibility, model (1) requires a relatively large set of IC profiles for model estimation and calibration, compared to its parametric counterparts. Thanks to fast progress in sensor and information technology, automatic data acquisition becomes increasingly common in industry. Consequently, a large amount of IC data is often available, and model (1) allows us to make use of such data without imposing a parametric model form.

2.2 Estimation of the NME Model

In this part, we discuss estimation of the IC g, γ and σ^2 (cf., expression (1)) from an IC dataset. These quantities will be used in constructing a Phase II nonparametric profile control chart when within-profile correlation is present and can be described by the NME model (1) (see related discussion in Section 2.3 below).

In the literature, there are some existing discussions about statistical analysis of correlated data under various settings and assumptions, including those in Altman (1990), Hart (1991), Hoover *et al.* (1998), Wang (1998), Zhang *et al.* (1998), Fan and Zhang (2000), Lin and Carroll (2000), and many others. Wu and Zhang (2002) propose a method for estimating model (1) by combining linear mixed-effects (LME) modeling and local linear kernel smoothing (cf., Fan and Gijbels 1996). They demonstrate that their estimator of g, which is referred to as LLME, is often more efficient than certain alternative estimators in terms of the mean squared errors. Furthermore, by their approach, it is fairly easy to obtain consistent estimators of γ and σ^2, which is important for constructing a Phase II control chart in the current study. For these reasons, we adopt Wu and Zhang's method here, which

is briefly described below.

For a given point $s \in [0,1]$, LLMEs of $g(s)$ and $f_i(s)$ are obtained by minimizing the following penalized, negative log, local linear kernel likelihood function:

$$\sum_{i=1}^{m}\left\{\frac{1}{\sigma^2}\sum_{j=1}^{n_i}[y_{ij}-\boldsymbol{z}_{ij}^T(\boldsymbol{\beta}+\boldsymbol{\alpha}_i)]^2K_h\left(x_{ij}-s\right)+\boldsymbol{\alpha}_i^T\mathbf{D}^{-1}\boldsymbol{\alpha}_i+\ln|\mathbf{D}|+n_i\ln(\sigma^2)\right\}, \tag{2}$$

where $K_h(\cdot)=K(\cdot/h)/h$, K is a symmetric density kernel function, h is a bandwidth, $\boldsymbol{z}_{ij}^T=(1,x_{ij}-s)$, $\boldsymbol{\beta}$ is a deterministic two-dimensional coefficient vector, and $\boldsymbol{\alpha}_i\sim(\mathbf{0},\mathbf{D})$ is a two-dimensional vector of the random effects. Minimization of (2) can be accomplished by the following iterative procedure:

Step 1. Set the initial values for $\mathbf{D}$ and σ^2, denoted as $\mathbf{D}_{(\mathbf{0})}$ and $\sigma^2_{(0)}$.

Step 2. At the k-th iteration, for $k\geq 0$, compute estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}_i$ by solving the so-called mixed-model equation (cf., Davidian and Giltinan 1995; Wu and Zhang 2002), and the resulting estimates are

$$\widehat{\boldsymbol{\beta}}^{(k)}=\{\sum_{i=1}^{m}\boldsymbol{Z}_i^T\boldsymbol{\Sigma}_i\boldsymbol{Z}_i\}^{-1}\{\sum_{i=1}^{m}\boldsymbol{Z}_i^T\boldsymbol{\Sigma}_i\boldsymbol{y}_i\} \tag{3}$$

$$\widehat{\boldsymbol{\alpha}}_i^{(k)}=\{\boldsymbol{Z}_i^T\mathbf{K}_i\boldsymbol{Z}_i+\sigma^2_{(k)}[\mathbf{D}_{(k)}]^{-1}\}^{-1}\boldsymbol{Z}_i^T\mathbf{K}_i(\boldsymbol{y}_i-\boldsymbol{Z}_i\widehat{\boldsymbol{\beta}}^{(k)}), \tag{4}$$

where $\boldsymbol{Z}_i=(\boldsymbol{z}_{i1},\ldots,\boldsymbol{z}_{in_i})^T$, $\boldsymbol{y}_i=(y_{i1},\ldots,y_{in_i})^T$, $\boldsymbol{\Sigma}_i=(\boldsymbol{Z}_i\mathbf{D}_{(k)}\boldsymbol{Z}_i^T+\sigma^2_{(k)}\mathbf{K}_i^{-1})^{-1}$ and $\mathbf{K}_i=\mathrm{diag}\{K_h(x_{i1}-s),\ldots,K_h(x_{in_i}-s)\}$.

Step 3. Based on $\widehat{\boldsymbol{\beta}}^{(k)}$ and $\widehat{\boldsymbol{\alpha}}_i^{(k)}$, update the estimates of $\mathbf{D}$ and σ^2 by

$$\mathbf{D}_{(k+1)}=\frac{1}{m}\sum_{i=1}^{m}\widehat{\boldsymbol{\alpha}}_i^{(k)}[\widehat{\boldsymbol{\alpha}}_i^{(k)}]^T \tag{5}$$

$$\sigma^2_{(k+1)}=\frac{1}{m}\sum_{i=1}^{m}\frac{1}{n_i}[\boldsymbol{y}_i-\boldsymbol{Z}_i(\widehat{\boldsymbol{\beta}}^{(k)}+\widehat{\boldsymbol{\alpha}}_i^{(k)})]^T\mathbf{K}_i[\boldsymbol{y}_i-\boldsymbol{Z}_i(\widehat{\boldsymbol{\beta}}^{(k)}+\widehat{\boldsymbol{\alpha}}_i^{(k)})]. \tag{6}$$

Step 4. Repeat Steps 2-3 until the following condition is satisfied:

$$\|\mathbf{D}_{(l)}-\mathbf{D}_{(l-1)})\|_1\,\big/\|\mathbf{D}_{(l-1)}\|_1\leq\epsilon,$$

where ϵ is a pre-specified small positive number (e.g., $\epsilon=10^{-4}$), and $||\mathbf{A}||_1$ denotes the sum of absolute values of all elements of $\mathbf{A}$. Then, the algorithm stops at the l-th iteration.

Note that, in Step 4, we use the relative error of the successive estimates of $\mathbf{D}$ in the convergence criterion. In fact, other estimates can also be used for this purpose. We use $\mathbf{D}$ here because our simulation shows that it gives good results in various cases. As a side note, similar to estimation of LME models, nonconvergence of the above iterative procedure can occasionally happen, although we found that the frequency of nonconvergence is negligible in all our simulation studies, except certain extreme cases such as the ones when m or n_is are too small. To reduce the frequency of nonconvergence, it is suggested in the literature to use good initial values for $\mathbf{D}$ and σ^2. A simple but effective method is to set $\mathbf{D}_{(0)}$ to be the identity matrix and

$$\sigma^2_{(0)} = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n_i}\sum_{j=1}^{n_i}[y_{ij} - \widehat{g}^{(P)}(x_{ij})]^2,$$

where $\widehat{g}^{(P)}(x_{ij})$ is the standard local linear kernel estimator constructed from the pooled data (cf., Hoover *et al.* 1998).

After obtaining estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}_i$ using the above algorithm, we can define

$$\widehat{g}(s) = \boldsymbol{e}_1^T\widehat{\boldsymbol{\beta}}(s), \quad \widehat{f}_i(s) = \boldsymbol{e}_1^T\widehat{\boldsymbol{\alpha}}_i(s),$$

$$\widehat{\gamma}(s_1, s_2) = \frac{1}{m}\sum_{i=1}^{m}\widehat{f}_i(s_1)\widehat{f}_i(s_2), \text{ for any } s_1, s_2 \in [0,1], \tag{7}$$

where $\boldsymbol{e}_1 = (1,0)^T$. Note that the variance estimator from the above iterative procedure depends on s. Since σ^2 is a population parameter that does not depend on s, we suggest estimating it by

$$\widehat{\sigma}^2 = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n_i}\sum_{j=1}^{n_i}[y_{ij} - \widehat{g}(x_{ij}) - \widehat{f}_i(x_{ij})]^2, \tag{8}$$

which is similar to the nonparametric estimator proposed by Hall and Marron (1990).

The following proposition investigates the asymptotic properties of the one-step estimators of g, γ and σ^2. For given initial values $\mathbf{D}_{(0)}$ and $\sigma^2_{(0)}$, the one-step estimators are those calculated by (3), (7) and (8) when $k = 1$.

Proposition 1 *Under Conditions (C1)–(C6), (C8)-I and (C9) given in Appendix A, for any points $s_1, s_2 \in [0,1]$, we have*

(i) $\widehat{g}(s_1) = g(s_1)\{1 + O_p[m^{-\frac{1}{2}} + O(h^2)]\}$;

(ii) $\widehat{\gamma}(s_1, s_2) = \gamma(s_1, s_2)\{1 + O_p[h^2 + (nh)^{-\frac{1}{2}} + m^{-\frac{1}{2}} + (mnh^3)^{-\frac{1}{2}}]\}$; *and*

(iii) $\widehat{\sigma}^2 = \sigma^2\{1 + O_p[h^2 + (nh)^{-\frac{1}{2}} + m^{-\frac{1}{2}} + (mnh^3)^{-\frac{1}{2}}]\}$,
where Condition (C6) assumes that $n_i \sim n$, *for all* i.

Result (i) of Proposition 1 is a direct conclusion of Theorem 1 in Wu and Zhang (2002), except that certain conditions have been modified. The other two results establish the consistency of the estimators of γ and σ^2, which is important for the Phase II profile monitoring problem discussed in the following sections.

2.3 Phase II Nonparametric Profile Monitoring

In this part, we present a Phase II nonparametric profile monitoring scheme in the general case when within-profile data might be correlated and the design points within and between profiles are arbitrary. This is a challenging task due to the following two major reasons. First, because the within-profile data might be correlated, estimation of the profile function g involves a considerable amount of computation if the NME modeling is also used in Phase II SPC, as described in Section 2.2. However, a good on-line control chart should maintain a reasonable efficiency while being effective in detecting profile shifts. Second, in cases when the design points $\boldsymbol{X}_i = \{x_{i1}, x_{i2}, \ldots, x_{in_i}\}$ are unchanged from profile to profile, one method that comes to mind is to first average observed responses y_{ij}s across individual profiles and then detect potential profile shifts using a generalized likelihood ratio test statistic (c.f., Fan *et al.* 2001). This idea can not be applied to the current problem directly because the response is observed at different design points in different profiles. One immediate alternative is to estimate g from individual profile data at a given set of points in $[0, 1]$. But the resulting estimates would be inefficient since they are constructed from individual profile data instead of from all observed data.

To overcome the above difficulties, at any point $s \in [0, 1]$, we consider the following local

weighted negative log likelihood:

$$WL(a,b;s,\lambda,t)=\sum_{i=1}^{t}\sum_{j=1}^{n_i}[y_{ij}-a-b(x_{ij}-s)]^2K_h\left(x_{ij}-s\right)(1-\lambda)^{t-i}/\nu^2(x_{ij}),$$

where λ is a weighting parameter and $\nu^2(x)=\gamma(x,x)+\sigma^2$ is the variance function of the response. Note that $WL(a,b;s,\lambda,t)$ combines the exponential weighting scheme used in EWMA at different time points through the term $(1-\lambda)^{t-i}$ and the local linear kernel smoothing procedure (cf., Fan and Gijbels 1996). At the same time, it takes into account the heteroscedasticity of observations by using $\nu^2(x_{ij})$. Then, the local linear kernel estimator of $g(s)$, defined as the solution to a of the minimization problem $\min_{a,b} WL(a,b;s,\lambda,t)$, has the expression

$$\widehat{g}_{t,h,\lambda}(s)=\sum_{i=1}^{t}\sum_{j=1}^{n_i}U_{ij}^{(t,h,\lambda)}(s)y_{ij}\Bigg/\sum_{i=1}^{t}\sum_{j=1}^{n_i}U_{ij}^{(t,h,\lambda)}(s), \tag{9}$$

where

$$\begin{aligned}U_{ij}^{(t,h,\lambda)}(s)&=\frac{(1-\lambda)^{t-i}K_h(x_{ij}-s)}{\nu^2(x_{ij})}\left[m_2^{(t,h,\lambda)}(s)-(x_{ij}-s)m_1^{(t,h,\lambda)}(s)\right],\\ m_l^{(t,h,\lambda)}(s)&=\sum_{i=1}^{t}(1-\lambda)^{t-i}\sum_{j=1}^{n_i}(x_{ij}-s)^lK_h(x_{ij}-s)/\nu^2(x_{ij}),\quad l=0,1,2.\end{aligned} \tag{10}$$

Note that $m_0^{(t,h,\lambda)}(s)$ is not used in (9); but, it will be used in Section 2.4 below.

From (9) and (10), we can see that $\widehat{g}_{t,h,\lambda}(s)$ makes use of all available observations up to the current time point t, and different profiles are weighted as in a conventional EWMA chart (i.e., more recent profiles get more weight and the weight changes exponentially over time). When $\lambda=0$ (i.e., all profiles receive equal weight), the resulting estimator is similar to the local linear generalized estimating equations (GEE) estimator considered in Lin and Carroll (2000). The GEE estimator can accommodate within-profile correlation without specifying the correlation structure (it uses the so-called independent working correlation matrix). Under certain mild conditions, Lin and Carroll show that it is asymptotically the best estimator. Although Wu and Zhang (2002) demonstrate that their LLME estimator performs better in certain cases, especially when within-profile correlation is strong, this latter estimator involves a considerable amount of computation, and may not be feasible for

Phase II profile monitoring which is an on-line sequential procedure. As a comparison, the estimator (9) has an explicit formula, and the related computation is relatively fast.

Following the convention in Phase II analysis, we assume that the IC regression function, denoted as g_0, and the variance function $\nu^2(\cdot)$ are both known. In practice, they need to be estimated from an IC dataset, as described in Section 2.2. Let $\xi_{ij} = [y_{ij} - g_0(x_{ij})]$, for all i and j, and $\widehat{\xi}_{t,h,\lambda}(s)$ be the estimator defined in (9) after y_{ij} are replaced by ξ_{ij}. Then, the IC distribution of $\widehat{\xi}_{t,h,\lambda}(s)$ does not depend on g_0, and the original testing problem with $H_0 : g = g_0$ versus $H_1 : g \neq g_0$, which is associated with the profile monitoring problem, is changed to the one with $H_0 : g = 0$ versus $H_1 : g \neq 0$. Consequently, the IC distribution of the proposed control chart defined below and all quantities related to this distribution (e.g., the control limit L) do not depend on g_0 either, which would simplify the design and implementation of our proposed control chart.

When the process is IC, $|\widehat{\xi}_{t,h,\lambda}(s)|$ should be small. So, a natural statistic that can be used for SPC would be

$$T_{t,h,\lambda} = c_{0,t,\lambda} \int \frac{[\widehat{\xi}_{t,h,\lambda}(s)]^2}{\nu^2(s)} \Gamma_1(s) ds,$$

where

$$c_{t_0,t_1,\lambda} = a^2_{t_0,t_1,\lambda}/b_{t_0,t_1,\lambda}, \quad a_{t_0,t_1,\lambda} = \sum_{i=t_0+1}^{t_1} (1-\lambda)^{t_1-i} n_i, \quad b_{t_0,t_1,\lambda} = \sum_{i=t_0+1}^{t_1} (1-\lambda)^{2(t_1-i)} n_i,$$

and Γ_1 is some pre-specified density function. In the expression of $T_{t,h,\lambda}$, quantities $c_{0,t,\lambda}$ and $\nu(\cdot)$ are used for unifying its asymptotic variance. See Theorem 1 below and its proof in the appendix for details. In practice, we suggest using the following discretized version:

$$T_{t,h,\lambda} \approx \frac{c_{0,t,\lambda}}{n_0} \sum_{k=1}^{n_0} \frac{\left[\widehat{\xi}_{t,h,\lambda}(s_k)\right]^2}{\nu^2(s_k)}, \tag{11}$$

where $\{s_k, k = 1, \ldots, n_0\}$ are some i.i.d. random numbers generated from Γ_1. Then, the chart triggers a signal if

$$T_{t,h,\lambda} > L,$$

where $L > 0$ is a control limit chosen to achieve a specific IC ARL, denoted as ARL_0. Hereafter, this chart is referred to as the mixed-effects nonparametric profile control (MENPC) chart.

In Phase II SPC, it is a convention that the IC distribution of the process measurements y_{ij}s is assumed known. Then, the control limit L can be searched by simulation based on this distribution. In practice, the IC distribution is often unknown. Instead, we usually have a quite large IC dataset. In such cases, L can be searched by a resampling algorithm, briefly described below. In each simulation run, we resample the IC dataset by randomly choosing a sequence of profiles with replacement. The sequence of profiles is sequentially chosen until a signal of shift is triggered by chart MENPC. Then, an actual ARL_0 value is computed based on B simulation runs, and L is searched by matching the actual ARL_0 value to the nominal value. In our numerical examples in Section 3, B is chosen 10,000.

It should be pointed out that it is computationally faster to use points s_k rather than the original design points x_{ij} in (11). As shown in Section 2.4 below, $T_{t,h,\lambda}$ can be calculated in a recursive manner when s_k are used in (11), and this recursive feature would be lost if x_{ij} are used. Further, from theoretical properties of $T_{t,h,\lambda}$ given in Theorem 2 below and from our numerical results, selection of $\{s_k, k = 1, 2, \ldots, n_0\}$ has little effect on the performance of the MENPC chart, as long as n_0 is not too small. See related discussion in Section 2.5 about practical guidelines on selection of certain procedure parameters. In the special case when design points $\boldsymbol{X}_i$ are unchanged for different profiles, we could use $\boldsymbol{X}_i$ directly (instead of $\{s_k, k = 1, 2, \ldots, n_0\}$) when computing the charting statistic.

Next, we give some asymptotic properties of the charting statistic $T_{t,h,\lambda}$, which could justify the performance of the MENPC chart to certain degree and shed some light on practical design of the chart as well. Theorem 1 below gives the asymptotic null distribution of $T_{t,h,\lambda}$, where design points x_{ij}s in each IC profile are assumed to be i.i.d. with a density Γ_2.

Theorem 1 *Assume that the process is IC and that Conditions (C1)–(C7) given in Appendix A all hold. Then, we have the following results.*

(i) If $n_i h$ is bounded, for each i, and Condition (C8)-II in Appendix A holds, then

$$\left(T_{t,h,\lambda} - \tilde{\mu}_h\right)/\tilde{\sigma}_h \xrightarrow{\mathcal{L}} N(0,1),$$

where

$$\tilde{\mu}_h = \frac{\int [K(u)]^2 du}{h} \int \frac{\Gamma_1(x)}{\Gamma_2(x)} dx, \quad \tilde{\sigma}_h^2 = \frac{2\int [K * K(u)]^2 du}{h} \int \frac{\Gamma_1^2(x)}{\Gamma_2^2(x)} dx.$$

(ii) If $n_i h \to \infty$, for each i, and Conditions (C8)-III and (C10) in Appendix A hold, then

$$\frac{1}{d_{0,t,\lambda}} T_{t,h,\lambda} \overset{D}{\sim} \frac{1}{n_0} \boldsymbol{\zeta}^T \boldsymbol{\zeta},$$

where $\overset{D}{\sim}$ denotes asymptotic equivalence in distribution, $d_{t_0,t,\lambda} = \sum_{i=t_0+1}^{t_1} (1-\lambda)^{2(t_1-i)} n_i^2 / b_{t_0,t,\lambda}$, and $\boldsymbol{\zeta}$ is an n_0-dimensional multivariate normal random vector with mean $\mathbf{0}$ and covariance matrix

$$\boldsymbol{\Omega} = \begin{pmatrix} \frac{\gamma(s_1,s_1)}{\nu^2(s_1)} & \cdots & \frac{\gamma(s_1,s_{n_0})}{\nu(s_1)\nu(s_{n_0})} \\ \vdots & \ddots & \vdots \\ \frac{\gamma(s_{n_0},s_1)}{\nu(s_{n_0})\nu(s_1)} & \cdots & \frac{\gamma(s_{n_0},s_{n_0})}{\nu^2(s_{n_0})} \end{pmatrix}.$$

From Theorem 1-(i), we can see that $T_{t,h,\lambda}$ is asymptotically independent of the nuisance parameters $\gamma(\cdot,\cdot)$ and σ^2. The condition that $n_i h$ is bounded for each i is satisfied when n_i is finite and bounded for each i, which is often the case in practice. When $n_i h \to \infty$ for each i, the within-profile correlation would play an important role in the expansion of the variance of $T_{t,h,\lambda}$, which leads to a different asymptotic distribution, as described in Theorem 1-(ii). In such situations, it seems desirable to incorporate the covariance matrix $\boldsymbol{\Omega}$ into the test statistic. However, $\boldsymbol{\Omega}$ may not be positive definite in certain cases (e.g., the case of compound symmetry correlation). Therefore, it is not obvious how to do so, which is left to our future research.

The next theorem investigates the asymptotic behavior of $T_{t,h,\lambda}$ under the OC model

$$y_{ij} = \begin{cases} g_0(x_{ij}) + f_i(x_{ij}) + \varepsilon_{ij}, & \text{if} \quad 1 \le i \le \tau \\ g_1(x_{ij}) + f_i(x_{ij}) + \varepsilon_{ij}, & \text{if} \quad i > \tau \end{cases} \tag{12}$$

where τ is an unknown change point, and $g_1(x) = g_0(x) + \delta(x)$ is the unknown OC profile function. In the theorem, we use the following notations:

$$\zeta_\delta = \int \left[\delta(u) + \frac{h^2\eta_1}{2}\delta''(u)\right]^2 \frac{\Gamma_1(u)}{\nu^2(u)}du, \quad \eta_1 = \int K(t)t^2dt,$$
$$\zeta_1 = \int \delta^2(u)\frac{\Gamma_1(u)\gamma(u,u)}{\nu^2(u)}du, \quad \zeta_2 = \int [\delta''(u)]^2\Gamma_1(u)du.$$

Theorem 2 *Under Conditions (C1)-(C7) given in Appendix A and the extra condition that $\zeta_2 < M$ for some constant $M > 0$, we have*

(i) If $n_i h$ is bounded for each i, $c_{0,t,\lambda} nh\zeta_1 \to 0$, and Condition (C8)-IV in Appendix A holds, then $(T_{t,h,\lambda} - \tilde{\mu}_h - c_{0,t,\lambda}\zeta_\delta)/\tilde{\sigma}_h \xrightarrow{\mathcal{L}} N(0,1)$;

(ii) If $n_i h$ is bounded for each i, $\zeta_2 \to 0$, and Condition (C8)-IV in Appendix A holds, then $T_{t,h,\lambda}$ has nontrivial power (i.e., greater than the nominal level) when $\delta \propto c_{0,t,\lambda}^{-4/9}$ and $h = O(c_{0,t,\lambda}^{-2/9})$.

(iii) If $n_i h \to \infty$ for each i, and Conditions (C8)-III and (C10) in Appendix A hold, then $\frac{1}{d_{0,t,\lambda}} T_{t,h,\lambda} \overset{D}{\sim} \frac{1}{n_0}\boldsymbol{\zeta}^T\boldsymbol{\zeta}$, where $\boldsymbol{\zeta}$ is an n_0-dimensional multivariate normal random vector with mean $\boldsymbol{\delta} = [\delta(s_1), \ldots, \delta(s_{n_0})]^T$ and covariance matrix $\boldsymbol{\Omega}$.

2.4 Some computational issues

Although computing power has improved dramatically and it is computationally trivial to do nonparametric function estimation for individual profiles, for on-line process monitoring, which generally handles a large amount of profiles, fast implementation is still important and some computational issues deserve our careful examination. For the proposed chart, computing the test statistic $T_{t,h,\lambda}$ by formulas (9)–(11) requires a considerable amount of computing time and a substantial amount of storage space as well to save all past profile observations. In this part, we provide updating formulas for computing $T_{t,h,\lambda}$, which can

greatly simplify the computation and reduce the storage requirement. Let

$$\tilde{m}_l^{(t,h)}(s) = \sum_{j=1}^{n_k} (x_{tj} - s)^l K_h(x_{tj} - s)/\nu^2(x_{tj}), \quad l = 0, 1, 2,$$

$$\tilde{q}_l^{(t,h)}(s) = \sum_{j=1}^{n_k} (x_{tj} - s)^l K_h(x_{tj} - s) y_{tj}/\nu^2(x_{tj}), \quad l = 0, 1.$$

Then, $m_l^{(t,h,\lambda)}(s)$ in (10) can be recursively updated by

$$m_l^{(t,h,\lambda)}(s) = (1-\lambda) m_l^{(t-1,h,\lambda)}(s) + \tilde{m}_l^{(t,h)}(s), \quad l = 0, 1, 2,$$

where $m_l^{(0,h,\lambda)}(s) = 0$, for $l = 0, 1, 2$. Let $q_l^{(t,h,\lambda)}(s)$, for $l = 0, 1$, be two working functions defined by the recursive formula

$$q_l^{(t,h,\lambda)}(s) = (1-\lambda) q_l^{(t-1,h,\lambda)}(s) + \tilde{q}_l^{(t,h)}(s), \quad l = 0, 1,$$

where $q_l^{(0,h,\lambda)}(s) = 0$, for $l = 0, 1$. Then, we have

$$\widehat{g}_{t,h,\lambda}(s) = \left[M^{(t,h,\lambda)}\right]^{-1} \left\{ (1-\lambda)^2 M^{(t-1,h,\lambda)} \widehat{g}_{t-1,h,\lambda} + \left[\tilde{q}_0^{(t,h)} m_2^{(t,h,\lambda)} - \tilde{q}_1^{(t,h)} m_1^{(t,h,\lambda)}\right] \right.$$
$$\left. + (1-\lambda) \left[q_0^{(t-1,h,\lambda)} \tilde{m}_2^{(t,h)} - q_1^{(t-1,h,\lambda)} \tilde{m}_1^{(t,h)} \right] \right\}, \tag{13}$$

where $M^{(t,h,\lambda)}(s) = m_2^{(t,h,\lambda)}(s) m_0^{(t,h,\lambda)}(s) - [m_0^{(t,h,\lambda)}(s)]^2$. On the right hand side of the above equation, dependence on s in each function is not made explicit in notation for simplicity, which should not cause any confusion.

Using the above updating formulas, implementation of the MENPC chart can be briefly described as follows. At time point t, we first compute quantities $\tilde{m}_l^{(t,h)}(s)$, for $l = 0, 1, 2$, and $\tilde{q}_l^{(t,h)}(s)$, for $l = 0, 1$, at n_0 pre-determined s locations (see related discussion in Sections 2.3 and 2.5 about selection of $\{s_k, k = 1, \ldots, n_0\}$). Then, $m_l^{(0,h,\lambda)}(s_k)$, for $l = 0, 1, 2$, and $q_l^{(0,h,\lambda)}(s_k)$, for $l = 0, 1$, are updated by the above formulas. Finally, $\widehat{g}_{t,h,\lambda}(s)$ is computed from (13), and the test statistic $T_{t,h,\lambda}$ is computed by $\widehat{g}_{t,h,\lambda}(s)$, after y_{ij} is replaced by ξ_{ij}. This algorithm only requires $O(n_0 n_i h)$ operations for monitoring the i-th profile, which is the same order as the computation involved in conventional local linear kernel smoothing. If n_i and n_0 are both large, we could further decrease the computation to the order of $O(n_i h)$, by using the updating algorithm proposed by Seifert *et al.* (1994). See Fan and Marron (1994) for a similar algorithm. Obviously, using the proposed updating formulas, computer storage does not grow sequentially with time t.

2.5 Certain practical guidelines

On the sizes of m and n_i: The number of IC profiles should be large enough to generate accurate estimates of IC g, γ, and σ^2. This has become a less significant issue nowadays because a large amount of IC data is often available due to advances in data acquisition techniques. Empirically speaking, to attain desirable IC distributional properties, we recommend using IC data with $n_i \geq 20$ and $m \geq 500$, although more systematic future research is required to determine the size of a necessary IC dataset.

On choosing the bandwidth: In estimation of the NME model (1) by the iterative procedure described in Section 2.2 for Phase I analysis, people often use data-driven bandwidth selection techniques, such as the least squares cross-validation (CV) and the generalized cross-validation (GCV) procedures. Wu and Zhang (2002) propose a CV method by combining leave-one-subject-out and leave-one-point-out CV schemes. Their study shows that this method can effectively track estimates of both g and f_i. We adopt it in our numerical analysis of the IC data. With respect to Phase II on-line profile monitoring, like many other smoothing-based tests, performance of the MENPC chart depends on selection of the bandwidth parameter h used in (9). Optimal selection of h remains an open problem in this area, and it is widely recognized that optimal h for nonparametric curve estimation is generally not optimal for testing (cf., e.g., Hart 1997). A uniformly most powerful test usually does not exist due to the fact that nonparametric regression functions have infinite dimensions. We suggest using the following empirical bandwidth formula,

$$h_E = \begin{cases} c_1 n^{-\frac{1}{5}} \left(\sum_{i=1}^{n} (x_i - \bar{x})^2 / n \right)^{\frac{1}{2}} & \text{for balanced design} \\ c_2 [\tilde{n}(2-\lambda)/\lambda]^{-1/5} \sqrt{\text{Var}(x)} & \text{for random design,} \end{cases} \tag{14}$$

where $\bar{x} = \sum_{i=1}^{n} x_i$ is the mean of the n design points in the balanced design case, $\tilde{n}$ and $\text{Var}(x)$ are the average number of design points and the variance of design points within a profile, respectively, in the random design case, which can be estimated from the IC data, and c_1 and c_2 are two constants. Empirically, c_1 and c_2 can be any values in the interval [1.0,2.0]. By (14), a smaller bandwidth is suggested for the random design case because the actual number of observations used in the MENPC chart at each time point is about $c_{0,t,\lambda}$

in such cases which is roughly $\tilde{n}(2-\lambda)/\lambda$.

On choosing λ: Traditionally, a larger λ leads to a quicker detection of larger shifts (cf., e.g., Lucas and Saccucci 1990). However, in the mixed-effects modeling, efficient estimation of the common profile function g requires use of observations across a number of different profiles, due mainly to the existence of random effects. From Theorems 1 and 2, we can see that the effective number of profiles used in the MENPC chart at each time point is asymptotically $(2-\lambda)/\lambda$. So, to estimate g properly, $(2-\lambda)h/\lambda$ should be large enough (cf., Condition (C8)-II in Appendix A). Consequently, λ cannot be chosen too large. Otherwise, even for a shift of large magnitude, $T_{t,h,\lambda}$ may not be able to detect it quickly, due to a large bias in estimating g. Empirically, we suggest choosing $\lambda \in [0.02, 0.1]$ if h_E in (14) is used.

On choosing $\{s_k, k = 1, 2, \ldots, n_0\}$: Based on our numerical experience, selection of $\{s_k, k = 1, 2, \ldots, n_0\}$ does not affect performance of the MENPC chart much, as long as n_0 is not too small and s_ks cover all the key parts of g_0 (e.g., peaks/valleys or oscillating regions) well. In our numerical examples presented in Section 3, we find that results would hardly change when $n_0 \geq 40$.

3 A Simulation Study

We present some simulation results in this section regarding the numerical performance of the proposed Phase II nonparametric profile monitoring chart MENPC. Throughout this section, the kernel function is chosen to be the Epanechnikov kernel function $K(x) = 0.75(1-x^2)I(-1 \leq x \leq 1)$, which is commonly used in the local smoothing literature due to some of its optimality properties. See Chapter 2 of Fan and Gijbels (1996) for related discussion. The IC ARL is fixed at 200. The error distribution is assumed to be standard Normal. For simplicity, we assume that $n_i = n = 20$ for all i, $x_{ij} \sim \text{Uniform}(0, 1)$, for $j = 1, \ldots, n$, $s_k = (k - 0.5)/n_0$, for $k = 1, \ldots, n_0$, and $n_0 = 40$. All ARL values reported in this section are averages of 10,000 replicated simulations. In addition, as suggested by Hawkins and Olwell (1998), here we focus on the steady-state OC ARL behavior of the chart, and assume

that shifts can only occur after time $\tau = 30$. When computing the OC ARL values, any simulation run in which a signal occurs before the $(\tau + 1)$th profile would be ignored.

It is challenging to compare the proposed method with alternative methods, since there is no obvious comparable method in the literature. Here, we first consider the control chart based on fixed-effects modeling for monitoring nonparametric profiles as an alternative method, denoted as FENPC. In this approach, f_i in model (1) is assumed to be zero, and consequently $\nu^2(x) = \sigma^2$ is used in the construction of $T_{t,h,\lambda}$ (cf., expressions (9)–(11)). Note that the FENPC chart can be regarded as a generalization of the NEWMA chart by Zou *et al.* (2008); the latter assumes that design points in different profiles are deterministic and unchanged from one profile to another while the former can handle arbitrary designs. By comparing the MENPC chart with the FENPC chart, we can see what will happen if within-profile correlation exists but is ignored. Following the recommendations in Section 2.5, for both charts, we use $h = 1.5[n(2-\lambda)/\lambda]^{-1/5}\sqrt{\mathrm{Var}(x)}$ in Phase II SPC, and use the CV method by Wu and Zhang (2002) for choosing bandwidths in modeling the IC data.

First, we study the possible effect of within-profile correlation on the IC run-length distributions of the two charts in the following four cases:

$$\begin{aligned} &\text{(I)} : f_i(x_{ij}) = 0; \quad \text{(II)} : f_i(x_{ij}) = b\alpha_i x_{ij}; \\ &\text{(III)} : f_i(x_{ij}) = b\alpha_i \cos(2\pi x_{ij}); \quad \text{(IV)} : [f_i(x_{i1}), \ldots, f_i(x_{in})]^T \sim b \cdot \mathrm{MN}_n(\mathbf{0}, \mathbf{\Sigma}), \end{aligned}$$

where $\alpha_i, i = 1, 2, \ldots,$ are independent standard normal random variables, $\mathrm{MN}_n(\mathbf{0}, \mathbf{\Sigma})$ denotes the n-dimensional multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$, and b is a constant. In all cases, we assume that $g_0(\cdot) = 0$. Obviously, in case (I), there is no within-profile correlation. In case (II), the random component $f_i(x)$ is a linear function of the covariate x. In case (III), it is a cosine function of the covariate x. In case (IV), the random component vector has a joint Normal distribution. In this case, we further assume that $\mathbf{\Sigma} = (\rho_{jk})$ and $\rho_{jk} = 0.2^{|x_{ij} - x_{ik}|}$, for $j, k = 1, \ldots, n$. In each case, a large IC sample with $m = 500$ and $n = 200$ is generated. By using estimated $\gamma(\cdot, \cdot)$ and σ^2 from this IC data, the control limits of the two control charts are computed, as described in Section 2.3. Then, their IC ARLs and the corresponding standard deviations of the run length, denoted

Table 1: IC ARL and SDRL values of the MENPC and FENPC charts.

		Model (I)		Model (II)		Model (III)		Model (IV)	
	b	ARL_0	$SDRL_0$	ARL_0	$SDRL_0$	ARL_0	$SDRL_0$	ARL_0	$SDRL_0$
	0.25	205	203	196	197	198	199	206	208
MENPC	0.50	205	203	201	200	195	194	208	204
	1.00	205	203	193	190	194	194	206	205
	0.25	199	197	110	109	170	172	38.3	34.0
FENPC	0.50	199	197	29.8	29.2	105	104	21.5	20.0
	1.00	199	197	8.48	8.29	35.5	34.6	15.1	14.2

as ARL_0 and $SDRL_0$ respectively, are summarized in Table 1. From the table, it can be seen that, in case (I) when the assumption of within-profile independence is valid, the ARLs and SDRLs of both charts are close to their nominal values 200, as expected. However, in cases (II)-(IV) when within-profile correlation is substantial, the FENPC chart has large biases in both ARL_0 and $SDRL_0$, especially when b is large. As a comparison, our proposed MENPC chart still performs well in these cases.

Next, we investigate the OC performance of the two control charts. The following two representative OC models are considered here:

$$\text{(i)} : g_1(x) = 2\theta(x - 0.5); \quad \text{(ii)} : g_1(x) = \theta \sin(2\pi(x - 0.5)).$$

In case (i), $\delta(x) = g_1(x) - g_0(x)$ is a straight line; it oscillates much in case (ii). The parameter θ controls the shift magnitude. For each control chart, two λ values 0.1 and 0.2 are used. With each λ value, the bandwidth (14) is used. In this comparison, we pretend that the IC model is known exactly, and the control limit of the FENPC chart is adjusted to attain the desired IC ARL value 200. Therefore, the difference between the MENPC and FENPC charts in this comparison is mainly in whether or not the within-profile correlation is taken into account in Phase II SPC. Obviously, for IC models (I) and (IV), $\nu^2(x)$ is independent of x; thus, the two charts would be equivalent in such cases. For this reason, Table 2 presents the OC ARL values of the two charts for IC models (II) and (III) only.

From Table 2, we can have the following results. First, the MENPC chart outperforms

Table 2: OC ARL comparison of the MENPC and FENPC charts when ARL_0=200, $n = 20$, $n_0 = 40$ and $\lambda = 0.1$ or 0.2.

		$\lambda = 0.1$			
		OC Model (i)		OC Model (ii)	
IC model	θ	MENPC	FENPC	MENPC	FENPC
	0.20	130 (1.36)	139 (1.48)	85.3 (0.83)	100 (0.98)
	0.30	80.5 (0.78)	98.0 (0.99)	40.5 (0.32)	52.2 (0.46)
	0.40	48.6 (0.42)	62.6 (0.59)	22.3 (0.15)	29.0 (0.21)
(II)	0.60	20.7 (0.13)	28.4 (0.20)	10.6 (0.05)	13.1 (0.06)
	0.80	12.1 (0.06)	16.0 (0.09)	6.81 (0.03)	8.57 (0.03)
	1.20	6.64 (0.02)	8.43 (0.03)	4.06 (0.02)	5.14 (0.02)
	1.60	4.60 (0.02)	5.82 (0.02)	2.93 (0.01)	3.71 (0.01)
	2.00	3.51 (0.01)	4.49 (0.01)	2.33 (0.01)	2.96 (0.01)
	2.40	2.88 (0.01)	3.68 (0.01)	1.95 (0.01)	2.50 (0.01)
	0.20	131 (1.38)	162 (1.73)	68.3 (0.64)	121 (1.25)
	0.30	81.0 (0.79)	121 (1.26)	31.2 (0.24)	65.7 (0.60)
	0.40	48.1 (0.42)	81.2 (0.76)	17.6 (0.11)	34.2 (0.25)
(III)	0.60	21.4 (0.14)	33.3 (0.24)	9.05 (0.04)	14.4 (0.06)
	0.80	12.4 (0.06)	17.7 (0.09)	6.02 (0.02)	9.14 (0.03)
	1.20	6.59 (0.03)	9.04 (0.03)	3.70 (0.01)	5.39 (0.02)
	1.60	4.51 (0.02)	6.10 (0.02)	2.68 (0.01)	3.92 (0.01)
	2.00	3.43 (0.01)	4.71 (0.01)	2.20 (0.01)	3.15 (0.01)
	2.40	2.81 (0.01)	3.85 (0.01)	1.85 (0.01)	2.65 (0.01)
		$\lambda = 0.2$			
		OC Model (i)		OC Model (ii)	
IC model	θ	MENPC	FENPC	MENPC	FENPC
	0.20	162 (1.70)	167 (1.83)	137 (1.51)	136 (1.43)
	0.30	128 (1.36)	131 (1.38)	85.3 (0.89)	87.2 (0.92)
	0.40	93.8 (0.97)	97.1 (1.02)	47.4 (0.46)	52.5 (0.51)
(II)	0.60	43.6 (0.40)	48.1 (0.47)	15.7 (0.11)	19.6 (0.15)
	0.80	19.5 (0.15)	24.4 (0.20)	7.94 (0.04)	9.81 (0.05)
	1.20	7.48 (0.03)	9.23 (0.05)	4.06 (0.01)	4.82 (0.02)
	1.60	4.59 (0.02)	5.49 (0.02)	2.79 (0.01)	3.28 (0.01)
	2.00	3.39 (0.01)	3.97 (0.01)	2.16 (0.01)	2.55 (0.01)
	2.40	2.70 (0.01)	3.15 (0.01)	1.84 (0.01)	2.11 (0.01)
	0.20	164 (1.77)	181 (1.94)	124 (1.32)	156 (1.66)
	0.30	133 (1.41)	156 (1.66)	69.2 (0.70)	113 (1.17)
	0.40	93.7 (0.98)	125 (1.33)	35.2 (0.32)	71.5 (0.71)
(III)	0.60	41.6 (0.38)	70.1 (0.72)	12.0 (0.07)	24.3 (0.19)
	0.80	18.8 (0.14)	32.0 (0.27)	6.69 (0.03)	10.7 (0.05)
	1.20	7.25 (0.03)	9.94 (0.05)	3.60 (0.01)	5.04 (0.02)
	1.60	4.39 (0.02)	5.73 (0.02)	2.53 (0.01)	3.42 (0.01)
	2.00	3.23 (0.01)	4.13 (0.01)	2.03 (0.01)	2.70 (0.01)
	2.40	2.60 (0.01)	3.27 (0.01)	1.73 (0.01)	2.21 (0.01)

NOTE: Standard errors are in parentheses.

the FENPC chart in all cases, which demonstrates the fact that the former chart is more effective because it explicitly incorporates within-profile heteroscedasticity of observations into its charting statistic. Second, the control charts with $\lambda = 0.2$ do not have satisfactory performance in most cases, compared to the charts with $\lambda = 0.1$. That is because the charts use about $[(2-\lambda)h/\lambda]n$ observations at each given point and the factor $(2-\lambda)h/\lambda = 1.37$ is fairly small in the case when $\lambda = 0.2$. Consequently, the charts estimate the regression function g with large bias and its ability to detect profile shifts is thus greatly reduced. This result confirms our recommendation in Section 2.5 that λ should be chosen smaller for monitoring profiles with within-profile correlation than for monitoring profiles with independent observations. By the way, further simulations (not reported here) also show that, when n_0 is chosen larger than 40, performance of either chart would not change much.

Next, we compare our proposed MENPC control chart, which is based on nonparametric mixed-effects modeling, with the control charts by Jensen *et al.* (2008) and Jensen and Birch (2009) that are based on linear and nonlinear mixed-effects modeling. It should be pointed out that both charts by Jensen and co-authors focus on Phase I SPC only. To compare with the proposed Phase II MENPC chart, they need to be modified for online sequential profile monitoring. With the nonlinear profile monitoring chart by Jensen and Birch (2009), this modification turns out to be difficult for the following reason. As pointed out by Zou *et al.* (2008) and Williams *et al.* (2007b), when nonlinear regression methods are used for constructing a control chart, non-convergence or slow convergence of numerical algorithms is often an issue, because Newton-Raphson iterative algorithms are routinely used in such cases to obtain parameter estimates. For Phase I analysis, this issue may not be serious as long as the initial values of the iterative algorithm are properly chosen. However, for Phase II online process monitoring where a large number of tests are performed, it is usually difficult to find proper initial values to guarantee the convergence of the iterative algorithm, especially when the profile model (instead of just the model parameters) changes after the process goes OC. For this reason, only the linear mixed-effects (LME) modelling approach by Jensen *et al.* (2008) is considered here, and it is modified for Phase II profile monitoring

as follows. Assume that the i-th profile data follow the LME model

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{X}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, 2, \ldots,$$

where $\boldsymbol{X}_i$ is the design matrix, $\boldsymbol{y}_i = (y_{i1}, \cdots, y_{in_i})^T$, $\boldsymbol{\beta}$ is the coefficient vector of the fixed-effects term, $\boldsymbol{b}_i \sim N_p(0, \mathbf{D})$ the coefficient vector of the random-effects term, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \cdots, \varepsilon_{in_i})^T$. Then, $\boldsymbol{\beta}$ can be estimated by the following weighted least squares estimator constructed from the i-th profile data:

$$\widehat{\boldsymbol{\beta}}_i = (\boldsymbol{X}_i^T\mathbf{V}_i^{-1}\boldsymbol{X}_i)^{-1}\boldsymbol{X}_i^T\mathbf{V}_i^{-1}\boldsymbol{y}_i,$$

where $\mathbf{V}_i = \boldsymbol{X}_i\mathbf{D}\boldsymbol{X}_i{}^T + \sigma^2\mathbf{I}$. Following the framework of the MEWMA chart by Zou *et al.* (2007), which is for online monitoring of general linear profiles using multivariate EWMA schemes, let us consider a sequence of EWMA working vectors

$$\boldsymbol{w}_i = (1-\lambda)\boldsymbol{w}_{i-1} + \lambda(\widehat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_0),$$

where $\boldsymbol{\beta}_0$ denotes the IC value of $\boldsymbol{\beta}$. Then, the control chart triggers a signal if

$$Q_i = \frac{2-\lambda}{\lambda}\boldsymbol{w}_i^T(\boldsymbol{X}_i^T\mathbf{V}_i^{-1}\boldsymbol{X}_i)\boldsymbol{w}_i > L,$$

where $L > 0$ is a control limit chosen to achieve a specific IC ARL. This chart is called the linear mixed-effect profile (LMEP) monitoring chart hereafter.

In the next example, we compare the LMEP and MENPC charts under the IC model:

$$y_{ij} = 1 + 2x_{ij} + 3x_{ij}^2 + \alpha_i x_{ij} + \varepsilon_{ij},$$

where α_i, for $i = 1, 2, \cdots$, are i.i.d. standard normal random variables. This model assumes that the fixed-effects part is a quadratic function of the predictor, and the random-effects part is a linear function. So, it is a LME model. In the simulation, n_i, λ, x_{ij}, and ε_{ij} are chosen or generated in the same way as that in the example of Tables 1 and 2. The following two OC models are considered here:

$$(1): y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_{ij}; \quad (2): y_{ij} = 1 + 2x_i + 3x_i^2 + \beta_4 \sin(2\pi\beta_5 x_i) + \varepsilon_{ij},$$

where βs are deterministic coefficients. Obviously, OC model (1) is a LME model, and OC model (2) is not a LME model. Six sets of values of βs are listed in Table 3, which correspond

to OC models that differ from the IC model with different degrees. The OC ARL values of the LMEP and MENPC charts are presented in Table 4. From the table, we can see that, even for OC model (1) which is a LME model, results of MENPC are compatible with results of LMEP. For OC model (2), MENPC outperforms LMEP uniformly.

Table 3: Six sets of parameters of the two OC models for comparing the MENPC and LMEP control charts

	Model (1)			Model (2)	
	β_0	β_1	β_2	β_4	β_5
(i)	1.3	2.0	3.0	0.1	1.5
(ii)	1.5	2.0	3.0	0.3	1.5
(iii)	1.0	2.3	3.0	0.5	1.5
(iv)	1.0	2.5	3.0	0.1	2.5
(v)	1.0	2.0	3.3	0.3	2.5
(vi)	1.0	2.0	3.5	0.5	2.5

Table 4: OC ARL comparison of the MENPC and LMEP charts when ARL_0=200, $n = 20$, $\lambda = 0.1$ or 0.2.

		OC Model (1)		OC Model (2)	
		MENPC	LMEP	MENPC	LMEP
	(i)	18.9 (0.13)	18.3 (0.12)	73.1 (0.69)	154 (1.61)
	(ii)	8.93 (0.04)	8.67 (0.04)	24.4 (0.15)	109 (1.15)
	(iii)	63.3 (0.60)	91.7 (0.91)	12.5 (0.05)	75.1 (0.77)
$\lambda = 0.1$	(iv)	29.0 (0.25)	44.1 (0.38)	107 (1.10)	166 (1.75)
	(v)	108 (1.08)	102 (1.03)	44.6 (0.35)	130 (1.38)
	(vi)	59.4 (0.54)	50.9 (0.46)	20.5 (0.10)	94.2 (0.98)
	(i)	25.0 (0.21)	22.0 (0.19)	127 (1.20)	149 (1.57)
	(ii)	9.90 (0.06)	8.54 (0.05)	59.4 (0.57)	102 (1.07)
	(iii)	76.0 (0.73)	113 (1.17)	23.3 (0.18)	64.5 (0.68)
$\lambda = 0.2$	(iv)	35.4 (0.35)	56.9 (0.57)	163 (1.74)	165 (1.75)
	(v)	121 (1.27)	123 (1.29)	114 (1.13)	124 (1.33)
	(vi)	69.6 (0.71)	67.5 (0.70)	68.8 (0.64)	87.9 (0.92)

NOTE: Standard errors are in parentheses.

4 A Real-Data Application

In this section, we demonstrate the proposed methodology by applying it to a dateset from a manufacturing process of aluminium electrolytic capacitors (AECs). This process transforms

raw materials, such as anode aluminum foil, cathode aluminum foil, guiding pin, electrolyte sheet, plastic cover, aluminum shell and plastic tube, into AECs that are appropriate for use in low leakage circuits and are well adapted to a wide range of environmental temperatures. The whole manufacturing process consists of a sequence of operations, including clenching, rolling, soaking, assembly, cleaning, aging and classifying. Before packing, a careful quality monitoring step is required by sampling from a batch of products.

Regarding quality of AECs, the most important characteristic is dissipation factor (DF), which can be automatically measured by an electronic device. However, it is known that DF measurements would change significantly with environmental temperature, and there is a specific requirement about the adaptability of AECs to the temperature. In order to monitor the adaptability, engineers put a sampled AEC in a container. Then, the container's temperature is controlled, and the temperature is supposed to stably increase from -26^oF to 78^oF. In this process, measurements of DF and the actual temperature inside the container are taken at 53 equally spaced time points. The actual temperature inside the container is reported by a temperature sensor. So, for each sampled AEC, a set of 53 observations of the pair (temperature, DF), which corresponds to (x, y) in model (1), are obtained for monitoring the adaptability of the AEC to the temperature. Figure 1 shows three AEC profiles along with an NME estimate of the IC profile function (see related discussion below). It should be noted that the actual temperature inside a container would fluctuate around its nominal level at each observation time. Therefore, actual temperature readings of different containers at a given observation time are all different, although the differences are usually small. For this dataset, profile monitoring charts requiring deterministic and fixed design points in different profiles (e.g., the one by Zou *et al.* 2008) would be difficult to use.

The entire AEC dataset contains 144 profiles, and each profile has $n = 53$ observations. We use the first 96 profiles to calibrate the proposed model and the remaining ones to test the model. A calibration sample of this size might be smaller than one would like to fully determine the IC distribution, but suffices to illustrate the use of the method in a real-world setting. Since the DF measurements are taken in consecutive time intervals, the AEC data exhibit a considerable amount of positive serial autocorrelation, which is confirmed by our

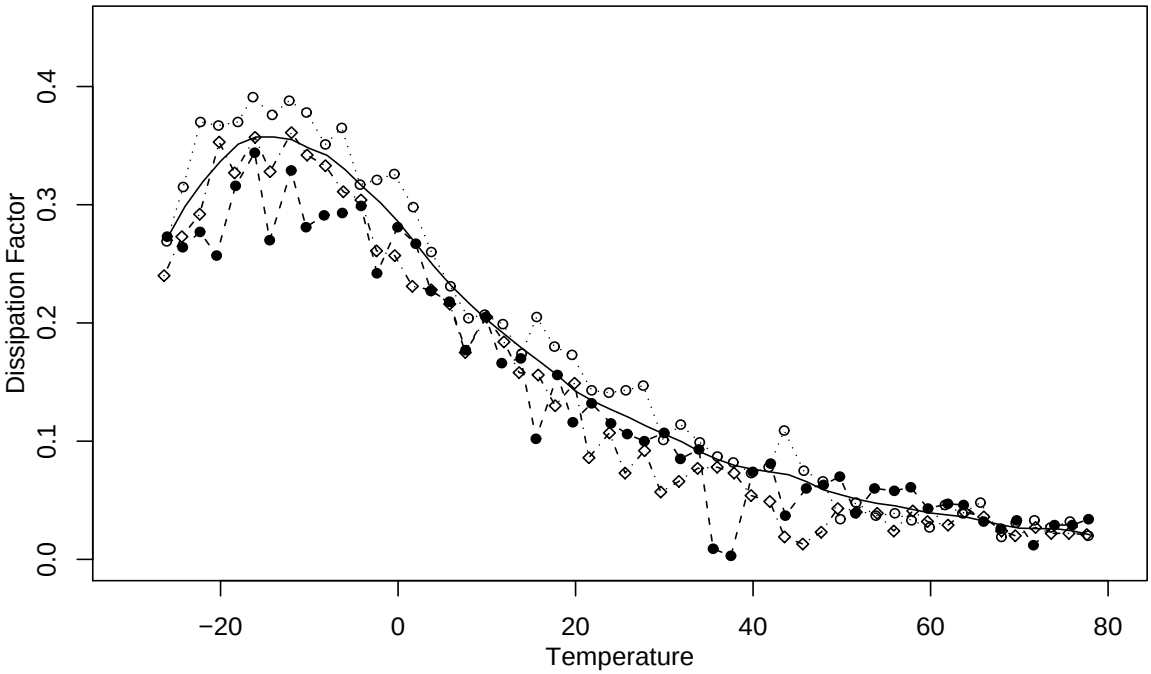

Figure 1: Three AEC profiles (lines connecting points with three different symbols) and the NME estimate (solid curve) of the IC profile function.

analysis described below.

We first fit model (1) to the calibration sample of the first 96 profiles of the data by the iterative procedure (3)–(6), using the suggested initial values of $\mathbf{D}$ and σ^2 given in Section 2.2 and the CV bandwidth selection procedure suggested by Wu and Zhang (2002). The resulting IC profile estimate $\widehat{g}$ is displayed in Figure 1 by the solid curve. From (7) and (8), we can also compute the estimated correlation of two observations of the response variable y at any two points s_1 and s_2 in the design interval

$$\widehat{\rho}(s_1, s_2) = \widehat{\gamma}(s_1, s_2)/\left[\widehat{\nu}(s_1)\widehat{\nu}(s_2)\right],$$

where $\widehat{\gamma}(s_1, s_2)$ is defined in (7), $\widehat{\nu}^2(s) = \widehat{\gamma}(s, s) + \widehat{\sigma}^2$ is the estimated variance of y at s, and $\widehat{\sigma}^2$ is defined in (8). Let $x_j^* = 2(j-1) - 26$, for $j = 1, 2, \ldots, 53$, be 53 equally spaced points in the design interval $[-26, 78]$, which denote the nominal temperature levels used in taking DF measurements of the sampled AECs. The estimated correlations $\widehat{\rho}(x_j^*, x_{j+1}^*)$, $\widehat{\rho}(x_j^*, x_{j+3}^*)$, $\widehat{\rho}(x_1^*, x_j^*)$, and $\widehat{\rho}(x_j^*, x_{53}^*)$, for $j = 1, 2, \ldots, 53$, are shown in Figure 2(a). From the plot, we can see that correlation within AEC profiles is substantial; thus, it should

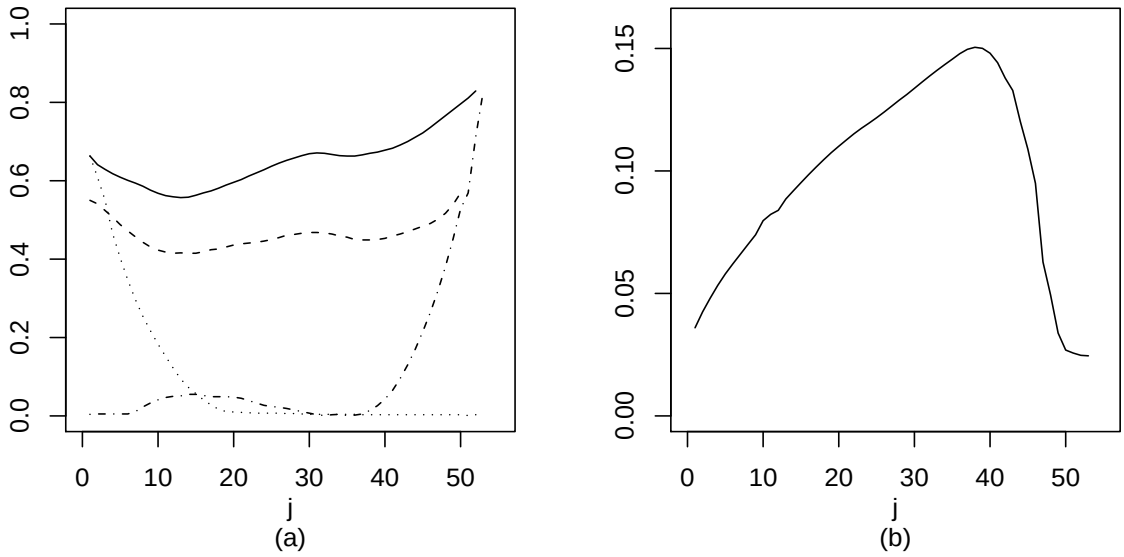

Figure 2: (a) Solid, dashed, dotted, and dash-dotted curves represent estimated within-profile correlations $\widehat{\rho}(x_j^*, x_{j+1}^*)$, $\widehat{\rho}(x_j^*, x_{j+3}^*)$, $\widehat{\rho}(x_1^*, x_j^*)$, and $\widehat{\rho}(x_j^*, x_{53}^*)$, for $j = 1, 2, \ldots, 53$, where $\{x_j^*,\ j = 1, 2, \ldots, 53\}$ are 53 equally spaced points in the design interval $[-26, 78]$. (b) Estimated standard deviation $\widehat{\nu}(x_j^*)$ of the response variable y at x_j^*, for $j = 1, 2, \ldots, 53$.

not be ignored. Figure 2(b) shows the estimated standard deviation $\widehat{\nu}(x_j^*)$ of the response variable y at x_j^*, for $j = 1, 2, \ldots, 53$, from which heteroscedasticity of the response variable y at different positions of x is clearly seen. Therefore, the proposed MENPC chart would be more appropriate to use in this case, compared to the FENPC chart discussed in the previous section which ignores the heteroscedasticity. In addition, we can obtain an estimate of the error standard deviation σ to be 0.016, by formula (8), which is much smaller than $\widehat{\nu}(x_j^*)$, especially when $j \in [12, 50]$. This result implies that the random-effects term in model (1) describes a substantial amount of random variation in the data.

Next, we construct the proposed MENPC chart for Phase II profile monitoring, using the estimated IC parameters computed from the IC data. As in the simulation study discussed in the previous section, the IC ARL is fixed at 200, and λ is chosen to be 0.1. For simplicity, we choose $n_0 = n = 53$ and $\{s_k, k = 1, 2, \ldots, n_0\}$ to be equally spaced in the design interval $[-26, 78]$ of the explanatory variable. All other parameters are chosen to be the same as those used in the example of Table 1. The control limit is computed to be 18.24 by simulation. The charting statistics $T_{t,h,\lambda}$, for $t = 97, \ldots, 144$, are shown in Figure 3 along with the control

limit, by the solid curve and solid horizontal line, respectively. The corresponding FENPC chart, using the same λ and h as those in the MENPC chart, is also presented in the figure along with its control limit 34.52, by the dashed curve and the dashed horizontal lines. From the plot, it can be seen that the MENPC chart gives a signal of profile shift around the 112-th time point, and remains above the control limit for several profiles until the 120-th profile. This result confirms a marked step-change which seems to have occurred around the 108-th profile. The process seems to have been adjusted around the 119-th profile; thus, the MENPC charting statistic goes back below its control limit afterward. As a comparison, the FENPC chart does not give a signal until the 118-th profile.

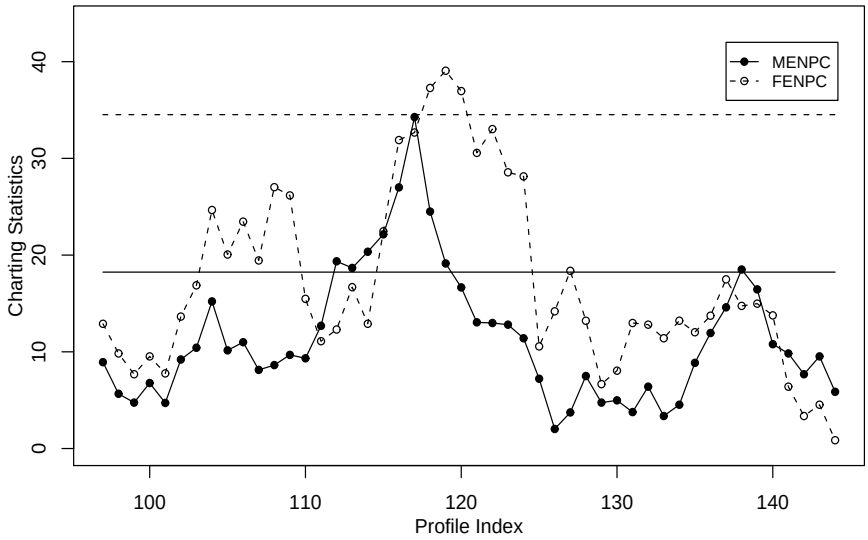

Figure 3: MENPC and FENPC control charts for monitoring the AEC process. The solid and dashed horizontal lines indicate their control limits, respectively.

5 Summary and Concluding Remarks

In this paper, we propose a Phase II control chart for monitoring nonparametric profiles. This chart is based on nonparametric mixed-effects modeling, local linear kernel smoothing, and EWMA process monitoring. It can accommodate within-profile correlation and arbitrary design. Numerical studies show that it is effective in detecting step profile shifts in various cases. Some numerical studies not reported in the paper show that it is also effective in

detecting certain drifts in profiles. The AEC example demonstrates that our method can be implemented conveniently in industrial applications.

As pointed out in Section 1, this paper focuses on Phase II profile monitoring only. It requires much future research to extend our method to Phase I analysis, in which detection of outliers and spikes would also be of interest, besides detection of step shifts in profiles. For Phase II profile monitoring, we only consider possible step shifts in the fixed-effects term g of model (1). In some cases, the variance-covariance structure of the profiles may also change over time. Such a change may or may not occur simultaneously with the shift in g. Online detection of possible changes in the variance-covariance structure of the profiles is not trivial, and we leave it for our future research. In addition, in some applications, we might be interested in monitoring a multivariate relationship between a response variable and several predictors over time. At this moment, we are not aware of any existing research on this topic, and we leave it to our future research to generalize the proposed control chart discussed in this paper to multivariate cases.

6 Supplemental Materials

supplement.pdf This pdf file provides certain technical details, including proofs of Proposition 1 and Theorems 1 and 2 in Section 2.

Acknowledgments: We thank the editor, the associate editor, and two referees for many constructive comments and suggestions, which greatly improved the quality of the paper. This research is supported in part by the grant DMS-0721204 from NSF of USA and the grant 10771107 from NNSF of China.

References

Altman, N. S. (1990), "Kernel Smoothing of Data With Correlated Errors," *Journal of the American Statistical Association*, 85, 749–759.

Colosimo, B. M., and Pacella, M. (2007), "On the Use of Principle Component Analysis to Identify Systematic Patterns in Roundness Profiles," *Quality and Reliability Engineering International*, 23, 707–725.

Davidian, M., and Giltinan, D.M. (1995), *Nonlinear Models for Repeated Measurement Data,* Chapman and Hall, London.

De Jong, P. (1987), "A Central Limit Theorem for Generalized Quadratic Forms," *Probability Theory and Related Fields*, 75, 261–277.

Diggle, P. J., Liang, K. Y., and Zeger, S. L. (1994), *Analysis of Longitudinal Data*, Oxford University Press, Oxford.

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

Fan, J., and Gijbels, I. (1996), *Local Polynomial Modeling and Its Applications*, Chapman and Hall, London.

Fan, J., and Marron, S. (1994), "Fast Implementation of Nonparametric Curve Estimators," *Journal of Computational and Graphical Statistics*, 3, 35–56.

Fan, J., Zhang, C., and Zhang, J. (2001), "Generalized Likelihood Ratio Statistics and Wilks Phenomenon," *The Annals of Statistics*, 29, 153–193.

Fan, J., and Zhang, J. (2000), "Two-Step Estimation of Functional Linear Models With Applications to Longitudinal Data," *Journal of the Royal Statistical Society B*, 62, 303–322.

Hall, P., and Marron, J. S. (1990), "On Variance Estimation in Nonparametric Regression," *Biometrika*, 77, 415–419.

Hart, J. D. (1991), "Kernel Regression Estimation with Time Series Errors," *Journal of the Royal Statistical Society B*, 53, 173–187.

Hart, J. D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*, Springer, New York.

Hawkins, D. M., and Olwell, D. H. (1998), *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Springer-Verlag.

Hoover, D. R., Rice, J. A., Wu, C. O., and Yang, L. P. (1998), "Nonparametric Smoothing Estimates of Time-Varying Coefficient Models with Longitudinal Data," *Biometrika*, 85, 809–822.

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

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.

Kazemzadeh, R. B., Noorossana, R., Amiri, A. (2007), "Phase I Monitoring of Polynomial Profiles," to appear in *Communications in Statistics: Theory and Methods*.

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

Lada, E. K., Lu, J. -C., and Wilson, J. R. (2002), "A Wavelet-Based Procedure for Process Fault Detection," *IEEE Transactions on Semiconductor Manufacturing*, 15, 79–90.

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

Lin, X., and Carroll, R. J. (2000), "Nonparametric Function Estimation for Clustered Data When the Predictor is Measured Without/With Error," *Journal of the American Statistical Association*, 95, 520–534.

Lucas, J. M., and Saccucci, M. S. (1990), "Exponentially Weighted Moving Average Control Scheme Properties and Enhancements," *Technometrics*, 32, 1–29.

Mahmoud, M. A., Parker, P. A., Woodall, W. H., and Hawkins, D. M. (2007), "A Change Point Method for Linear Profile Data," *Quality and Reliability Engineering International*, 23, 247–268.

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

Rice, J. A., and Wu, C. O. (2001), "Nonparametric Mixed Effects Models for Unequally Sampled Noisy Curves," *Biometrics*, 57, 253–259.

Seifert, B., Brockmann, M., Engel, J., and Gasser, T. (1994), "Fast Algorithms for Nonparametric Estimation," *Journal of Computational and Graphical Statistics*, 3, 192–213.

Shi, M., Weiss R. E., and Taylor, J. M. G. (1996), "An Analysis of Paediatric CD4 Counts for Acquired Immune Deficiency Syndrome Using Flexible Random Curves," *Applied Statistics*, 45, 151-163.

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

Wang, Y. (1998), "Smoothing Spline Models with Correlated Random Errors," *Journal of the American Statistical Association*, 93, 341–348.

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

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

Woodall, W. H. (2000), "Controversies and contradictions in statistical process control," *Journal of Quality Technology*, 32, 341–350.

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, 309–320.

Wu, H., and Zhang, J. (2002), "Local Polynomial Mixed-Effects Models for Longitudinal Data," *Journal of the American Statistical Association*, 97, 883–897.

Young, T. M., Winistorfer, P. M., and Wang, S. (1999), "Multivariate Control Charts of MDF and OSB Vertical Density Profile Attributes," *Forest Products Journal*, 49, 79–86.

Zhang, D., Lin, X., Raz, J., and Sowers, M. (1998), "Semiparametric Stochastic Mixed Models for Longitudinal Data," *Journal of the American Statistical Association*, 93, 710–719.

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

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

Zou, C., Zhang, Y., and Wang, Z. (2006), "Control Chart Based on Change-Point Model for Monitoring Linear Profiles," *IIE Transactions*, 38, 1093–1103.