Rejoinder

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

We thank the editor Professor David Steinberg for organizing this stimulating discussion. We are also grateful to all the discussants for their constructive comments about the method proposed in our paper (denoted as QZW hereafter). Profile monitoring is a relatively new research area; but, it has a profound application background (cf., Wang and Tsung 2005, Woodall et al. 2004). Due to the fact that the data structure in profile monitoring is much more complicated than that in conventional process monitoring problems, it is also a challenging task. In QZW, we try to apply some recent statistical tools developed in some other research areas, including longitudinal data and functional data analysis (e.g., Fitzmaurice et al. 2008, Liang and Zeger 2002), to the area of profile monitoring. As pointed out by the discussants, there are still some issues in our proposed method that need to be addressed in future research. In the next several parts, we provide our thoughts about some main issues raised by the discussants.

1 Phase I and Phase II Profile Monitoring

Our paper focuses on Phase II profile monitoring. Instead of assuming the in-control (IC) profile mean function and other parameters and functions (i.e., g, γ, and σ^2) to be known, we assume that there is an IC dataset from which g, γ, and σ^2 can be estimated. In practice, however, it is still a challenging task to obtain the IC dataset. In that regard, we agree with Woodall, Birch and Du completely, and think that much future research is required on Phase I analysis of profile data.

In the limited literature on Phase I analysis of profile data, Jensen et al. (2008) and Jensen and Birch (2009) have developed procedures for monitoring linear and nonlinear profiles. Their nonlinear profile monitoring procedure in Jensen and Birch (2009) can be easily generalized to the nonparametric setup, by using the nonparametric mixed-effects modeling described in Section 2.2 of QZW. Let

$$T_i^2 = \widehat{\boldsymbol{f}}_i^T \widehat{\boldsymbol{\Sigma}}^{-1} \widehat{\boldsymbol{f}}_i, \tag{1}$$

where $\widehat{\boldsymbol{f}}_i = (\widehat{f}_i(s_1), \ldots, \widehat{f}_i(s_{n_0}))^T$, $\{s_1, s_2, \ldots, s_{n_0}\}$ are n_0 given points in the design interval of x (cf., model (1) in QZW), and $\widehat{\boldsymbol{\Sigma}}$ is an estimated covariance matrix of $\widehat{\boldsymbol{f}}_i$. Then, the ith profile is an outlier if T_i^2 is larger than a threshold value. In the AEC data example discussed in Section 4 of QZW, the first 96 profiles are treated as an IC dataset. We agree with Woodall, Birch and Du that, in practice, it still needs to be checked whether there are any outliers in this data. To this end, we apply the above Phase I outlier detection procedure to this IC dataset, with $\widehat{\boldsymbol{\Sigma}}$ chosen to be the successive difference estimator, as recommended by Jensen and Birch (2009) for detecting sustained step shifts in profiles. Namely,

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} (\widehat{\boldsymbol{f}}_{i+1} - \widehat{\boldsymbol{f}}_i)(\widehat{\boldsymbol{f}}_{i+1} - \widehat{\boldsymbol{f}}_i)^T.$$

Figure 1 presents T_i^2, for $1 \leq i \leq 96$, along with a control limit corresponding to the significance level of 0.05 that is computed by a bootstrap procedure as follows. Each time we draw 96 $\widehat{f}_i$s with replacement from $\{\widehat{f}_i, 1 \leq i \leq 96\}$ that are computed beforehand by the procedure described in Section 2.2 of QZW. Then, a bootstrap version of $\widehat{\boldsymbol{\Sigma}}$ is computed from the resampled $\widehat{f}_i$s, and 96 bootstrap observations of T_i^2 are computed from the resampled $\widehat{f}_i$s and the corresponding bootstrap version of $\widehat{\boldsymbol{\Sigma}}$. This process is repeated 10,000 times, and the control limit is defined to be the 95th percentile of all bootstrap observations of T_i^2 computed. From the plot, it can be seen that no outliers are detected by this procedure.

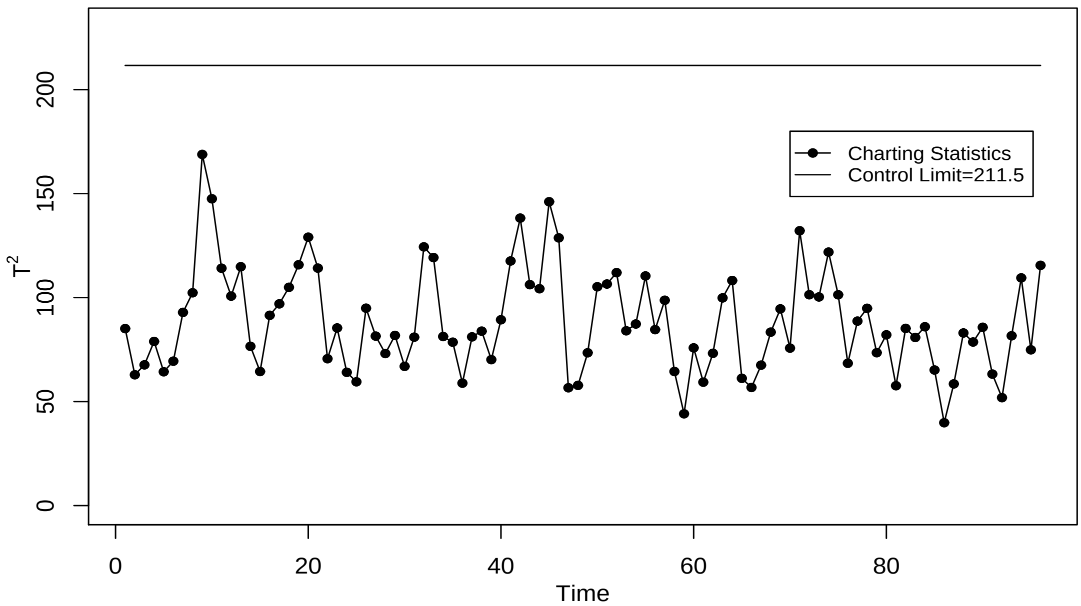

Figure 1: Phase I T^2 control chart defined in (1) for monitoring the first 96 profiles of the AEC dataset.

It should be pointed out that, for detecting outliers in Phase I profile monitoring, the T^2 control chart defined in (1) may not be the most powerful one. In the context of longitudinal data analysis, a similar issue has been investigated by Fung et al. (2002) who proposed certain influence diagnostics and outlier detection procedures using semiparametric mixed-effects modeling. For Phase I monitoring of nonparametric profiles, discussion about *nonparametric covariance analysis* and *comparison of multiple curves* in the context of nonparametric regression testing (cf., Dette and Neumeyer 2001, Neumeyer and Dette 2003, Young and Bowman 1995) might also be relevant.

We appreciate the comment made by Chipman, MacKay and Steiner that the absence of random-effects terms in our Phase II modeling (cf., the expression of $WL(a, b; s, \lambda, t)$ in the second paragraph of Section 2.3 in QZW) may affect the efficiency of our proposed Phase II profile monitoring chart. Woodall, Birch and Du raise a similar issue and they ask why we use the method by Wu and Zhang (2002) in Phase I modeling and the different method by Lin and Carroll (2000) in Phase II profile monitoring. As explained in Section 2.3 of QZW, the major reason for us to use two different methods in Phase I and Phase II analysis is that the computation involved in the iterative algorithm of the method by Wu and Zhang (2002) is quite substantial. For Phase I analysis in which the sample size is fixed, that algorithm is still feasible. But, for online Phase II monitoring, this method is cumbersome and may not be feasible. It is true that, by using the method of Lin and Carroll (2000) in Phase II analysis (cf., (9) and (10) in QZW), it appears that only the heteroscedasticity of the within-profile observations is taken care of and the within-profile correlation has not been fully accommodated. However, according to Lin and Carroll (2000), under some regularity conditions, it would not have much of an effect on the estimated profile mean function to only accommodate the heteroscedasticity properly without specifying the complete correlation structure of the within-profile observations. To further investigate this issue, we run a simulation in cases when the IC model (II), the out-of-control (OC) models (i) and (ii), and $\lambda = 0.1$ are considered (cf., Section 3 of QZW for their definitions). In this example, besides our proposed chart MENPC, we also consider the chart constructed as follows. Let $\widehat{g}^*_{t,h,\lambda}(s)$ be the estimator of $g(s)$, obtained by the algorithm described in Section 2.2 of QZW, except that expression (2) in QZW is replaced by

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

where $\lambda \in [0, 1]$ is a weighting parameter. Obviously, the above expression combines different profiles for Phase II monitoring using the EWMA weighting scheme. Then, a charting statistic can be constructed in a similar way to (11) in QZW, after $\{y_{ij}\}$ are replaced by $\{\xi_{ij} = y_{ij} - g_0(x_{ij})\}$ in the above expression. This control chart is denoted as MENPC1, and it is based on the method by Wu and Zhang (2002) in both the Phase I and Phase II analysis. With all the procedure parameters chosen in the same way as those in Table 2 of QZW, the OC ARL values of the charts MENPC and MENPC1 are presented in Table 1. From the table, it can be seen that the two charts perform similarly in all cases considered, and the chart MENPC1 is slightly better in cases with OC Model (ii).

Table 1: OC ARL comparison of the charts MENPC and MENPC1 when ARL$_0$=200, $n = 20$, $n_0 = 40$, $\lambda = 0.1$ and IC model (II) is used.

	OC Model (i)		OC Model (ii)	
θ	MENPC	MENPC1	MENPC	MENPC1
0.20	130 (1.36)	134 (1.22)	85.3 (0.83)	84.8 (0.86)
0.30	80.5 (0.78)	77.2 (0.80)	40.5 (0.32)	37.4 (0.33)
0.40	48.6 (0.42)	46.5 (0.44)	22.3 (0.15)	20.6 (0.16)
0.60	20.7 (0.13)	19.8 (0.11)	10.6 (0.05)	10.1 (0.05)
0.80	12.1 (0.06)	11.8 (0.07)	6.81 (0.03)	6.62 (0.03)
1.20	6.64 (0.02)	6.60 (0.03)	4.06 (0.02)	3.95 (0.02)
1.60	4.60 (0.02)	4.64 (0.02)	2.93 (0.01)	2.88 (0.01)
2.00	3.51 (0.01)	3.54 (0.01)	2.33 (0.01)	2.33 (0.01)
2.40	2.88 (0.01)	2.85 (0.01)	1.95 (0.01)	1.96 (0.01)

Regarding the initial estimator $\sigma^2_{(0)}$ used in the iterative algorithm in Section 2.2 of QZW, Woodall, Birch and Du suggest to replace $\widehat{g}^{(P)}(x_{ij})$ in the formula for $\sigma^2_{(0)}$ given in the paragraph immediately before expression (7) of QZW by $\widehat{g}_i(x_{ij})$ where $\widehat{g}_i$ is the local linear kernel estimator of g that is constructed from the ith profile data alone. We have tried this idea in some numerical examples and find that the modified initial estimator is indeed better.

2 Temporal Autocorrelation

In QZW, we only consider possible correlation among within-profile observations, and assume that observations between profiles are independent of each other. We appreciate the comment made by Apley that temporal autocorrelation among profiles collected at consecutive time points might also be common in practice. We agree with Apley completely on this issue, and believe that it is an important future research problem to propose profile monitoring charts that can accommodate both within-profile and between-profile correlation. A number of useful references on handling temporal autocorrelation in some conventional process monitoring problems have been cited in Apley's discussion. Another relevant paper, by Noorossana et al. (2008), tries to handle autocorrelated linear profiles using certain time series models. The method proposed in that paper has the potential to be generalized for handling autocorrelated nonparametric profiles, which needs to be further studied.

By the suggestion of Apley, in the AEC data example, we present the estimated profiles $\widehat{g}(s) + \widehat{f}_i(s)$ over i, for $1 \leq i \leq 96$, at two specific positions $s = x_{10}^*$ and $s = x_{30}^*$ in the two panels of Figure 2. We also compute the lag-1 and lag-2 autocorrelations of the time series shown in each panel. They are 0.104 and 0.058 in the case of panel (a), and 0.154 and 0.074 in the case of panel (b). From the plots and the computed autocorrelation values, we can see that temporal autocorrelation is not evident in this data, which can be explained by the fact that all the profiles in this example are actually collected over a relatively long time period.

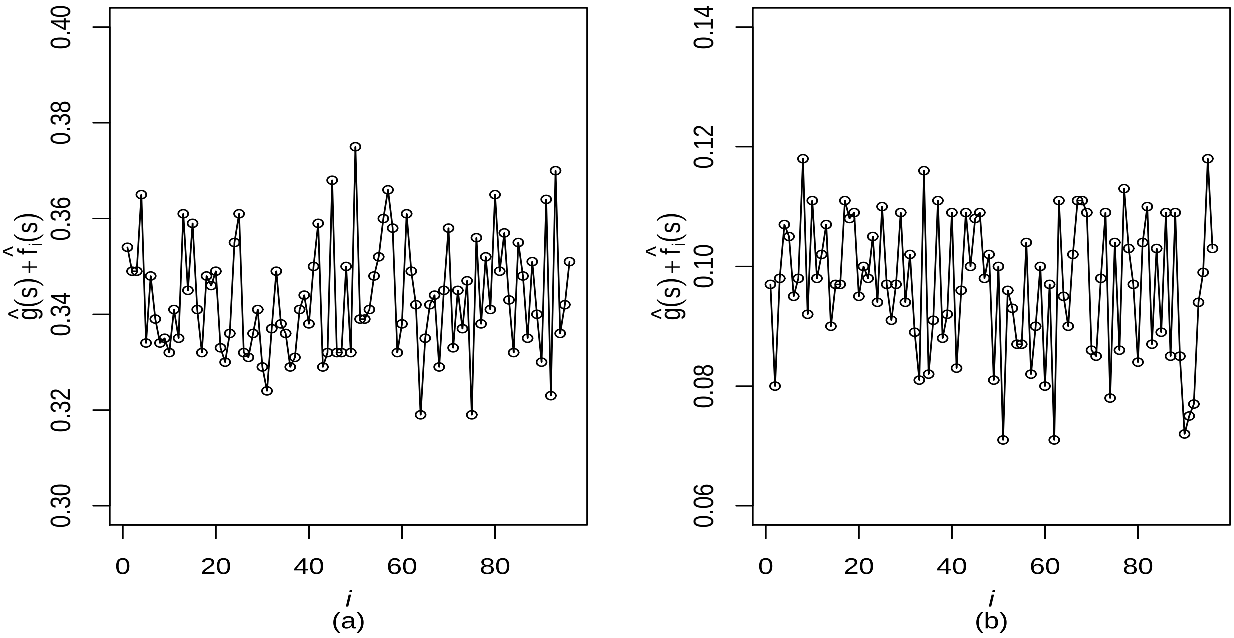

Figure 2: Plots of the estimated profiles $\widehat{g}(s) + \widehat{f}_i(s)$ over i, for $1 \leq i \leq 96$, at two specific positions $s = x_{10}^*$ (plot (a)) and $s = x_{30}^*$ (plot (b)).

In his discussion, Apley proposes two possible approaches for accommodating between-profile correlation. One is the Markov bootstrap procedure and the other one is the block bootstrap procedure. He thinks that the block bootstrap procedure might be more appropriate to use for monitoring profiles because of the relatively complicated structure of the profile data. In the AEC example, we compute the control limit using the block bootstrap procedure with the bootstrap sample size 10,000 and the block size 9 which is about 10% of the IC data. The computed control limit is 19.37. Compared to the control limit 18.24 reported in QZW, this control limit is marginally larger and it does not change the signal time at the 112-th time point (cf., Figure 3 in QZW).

3 Alternative Charting Statistics

Besides our proposed charting statistic $T_{t,h,\lambda}$ defined by (11) in QZW, the discussants propose two alternative charting statistics. For the purpose of detecting a shift in the covariance function $\gamma(x_1, x_2)$ (cf., its definition in Section 2.1 of QZW), Apley suggests using the charting statistic

$$\widetilde{T}_t^{(1)} = \sum_{i=0}^{t-1} (1-\rho)^i T_{t-i,h,1},$$

where $\rho \in [0, 1]$ denotes an EWMA weighting parameter. From the definition of $T_{t,h,\lambda}$, we can see that $T_{t-i,h,1}$ is a quadratic measure of the difference between the estimated profile mean function from the $(t-i)$th profile data alone and the IC profile mean function g_0. Therefore, $\widetilde{T}_t^{(1)}$ which is an EWMA statistic constructed from $\{T_{t-i,h,1}\}$ tries to combine information from different profiles about the difference between individual profiles and g_0. In their discussion, Woodall, Birch and Du propose this charting statistic as well.

The statistic $\widetilde{T}_t^{(1)}$ is natural to use. As a matter of fact, we also thought of it at the begining of this research project. It was finally given up and replaced by $T_{t,h,\lambda}$ for the following reason. The estimator of g from a single profile is relatively noisy, especially when the profile contains only a small number of observations. The relatively large variability of such estimators of g would be inherited by $\widetilde{T}_t^{(1)}$ and make it less sensitive to shifts in the profile mean function. As a comparison, $\widehat{g}_{t,h,\lambda}(s)$ defined in (9) of QZW is obtained from multiple profiles through the EWMA weighting scheme. Its variability is therefore smaller than the variability of $\widehat{g}_{t,h,1}(s)$ which is constructed from the tth profile alone. Consequently, the control chart based on $T_{t,h,\lambda}$ is expected to be more powerful for detecting a shift in the profile mean function, compared to the chart based on $\widetilde{T}_t^{(1)}$.

Instead of $T_{t,h,\lambda}$, Chipman, MacKay and Steiner think that it is more convenient to use the charting statistic

$$\widetilde{T}_t^{(2)} = \sum_{i=0}^{t-1} (1-\rho)^i \left(\mathbf{y}_i - \mathbf{g}_{0,i}\right)^T \Sigma_i^{-1} \left(\mathbf{y}_i - \mathbf{g}_{0,i}\right),$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})^T$, and $\mathbf{g}_{0,i} = (g_0(x_{i1}), g_0(x_{i2}), \ldots, g_0(x_{in_i}))^T$. Obviously, $\widetilde{T}_t^{(2)}$ is an EWMA statistic constructed from quadratic discrepancies between $\mathbf{y}_i$ and $\mathbf{g}_{0,i}$. From its construction, we think that $\widetilde{T}_t^{(2)}$ would not be as sensitive to possible profile mean shifts as $T_{t,h,\lambda}$, because of the large variability in $\mathbf{y}_i - \mathbf{g}_{0,i}$. However, this chart might be good for detecting shifts in the covariance function of $\gamma(x_1, x_2)$, because $\mathbf{y}_i - \mathbf{g}_{0,i}$ is just $\mathbf{f}_i + \boldsymbol{\varepsilon}_i$ when the process is IC, where $\mathbf{f}_i = (f_i(x_{i1}), f_i(x_{i2}), \ldots, f_i(x_{in_i}))^T$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in_i})^T$.

To investigate the performance of the alternative charts based on $\widetilde{T}_t^{(1)}$ and $\widetilde{T}_t^{(2)}$, which are denoted as ALT$^{(1)}$ and ALT$^{(2)}$, respectively, we consider the following example, where the IC models (II) and (III) and the OC model (i) used in QZW are considered. In addition, we consider the following OC model in which the shift is in variances:

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

where θ^* is a constant. In the charts MENPC, ALT$^{(1)}$ and ALT$^{(2)}$, $\lambda = \rho = 0.1$ and all other parameters are chosen to be the same as those used in the example of Table 2 of QZW. The OC ARL values of the three charts are presented in Table 2.

From Table 2, we can see that chart ALT$^{(1)}$ does not work well in all cases considered in this example, in comparison with the other two charts. As explained earlier, this chart would not be efficient for detecting profile mean shifts, which is confirmed here. From the table, it seems that this chart is not good for detecting shifts in variances either. This latter result is not surprising because the quantity $T_{t-i,h,1}$ that it uses does not contain much information about the covariance function $\gamma(x_1, x_2)$. Chart ALT$^{(2)}$, on the other hand, does perform reasonably well for detecting shifts in variances. But it is not powerful for detecting small to moderate profile mean shifts, as expected. Our proposed chart MENPC is designed for detecting profile mean shifts. So, it performs reasonably well in cases with OC model (i), especially when the mean shift is small or moderate (i.e., θ value in the table is between 0.20 and 1.20). However, this chart does not perform well for detecting shifts in variances. So, in practice, if shifts in both mean and variances are our concern, then we probably want to use the charts MENPC and ALT$^{(2)}$ simultaneously.

Table 2: OC ARL values of the charts MENPC, $ALT^{(1)}$ and $ALT^{(2)}$ in cases when ARL_0=200, $n = 20$, $n_0 = 40$ and $\lambda = \rho = 0.1$.

	OC Model (i)				OC model defined in (2)			
	θ	MENPC	$ALT^{(1)}$	$ALT^{(2)}$	θ^*	MENPC	$ALT^{(1)}$	$ALT^{(2)}$
IC model (II)	0.20	130	197	196	0.05	173	183	74.1
	0.30	80.5	196	165	0.10	156	177	33.1
	0.40	48.6	195	142	0.15	134	174	16.6
	0.60	20.7	191	82.5	0.20	129	160	9.66
	0.80	12.1	192	45.8	0.30	101	156	4.26
	1.20	6.64	194	13.2	0.40	77.6	138	2.40
	1.60	4.60	197	4.42	0.50	62.3	124	1.77
	2.00	3.51	190	2.10	0.75	35.7	111	1.17
	2.40	2.88	185	1.37	1.00	21.9	92.4	1.05
IC model (III)	0.20	131	199	186	0.05	179	187	75.9
	0.30	81.0	197	153	0.10	157	182	34.7
	0.40	48.1	197	119	0.15	142	162	16.5
	0.60	21.4	196	68.4	0.20	138	168	9.83
	0.80	12.4	195	34.9	0.30	106	146	4.31
	1.20	6.59	193	8.77	0.40	82.7	140	2.45
	1.60	4.51	193	3.02	0.50	65.4	127	1.80
	2.00	3.43	192	1.60	0.75	34.5	111	1.17
	2.40	2.81	190	1.17	1.00	21.4	97.5	1.04

4 Are Asymptotic Results Relevant?

We appreciate the comment made by Apley regarding the asymptotic results included in QZW, and we agree with him completely that readers should not assign too much significance to Theorem 1 and other asymptotic results in QZW when designing the proposed control chart. Generally speaking, asymptotic results are valid only when the sample size is large. In reality, the related sample size is always finite. Therefore, asymptotic results are always a certain distance away from reality, and that distance depends on the sample size and how all the conditions and assumptions required by the asymptotic results are satisfied in a practical situation. In SPC, this is especially true because whenever we get a signal of shift from a control chart, the process (e.g., a production line) should be stopped immediately for people to find the root cause of the shift and then make certain appropriate adjustments of the process. Therefore, the sample size is hardly large in such cases. This may be the reason why asymptotic results are not often seen in the SPC literature.

However, if we check the asymptotic results and their associated conditions and assumptions carefully, then we can still get some helpful information about the related SPC procedure. For instance, Apley already provides a thorough explanation why Theorem 1 in QZW concludes that $T_{t,h,\lambda}$ is independent of $\gamma(x_1, x_2)$ and σ^2 when $n_i h$ is bounded for each i and when other conditions hold. We agree with Apley that this result should not be used directly for choosing the control limit of our proposed chart. One major reason is that this result describes the IC behavior of the charting statistic $T_{t,h,\lambda}$ only; it does not take any profile mean shift into account. More specifically, the result holds when h tends to 0 and when other conditions are valid. But, a too small h would not be appropriate to use for detecting a profile mean shift effectively. On the other hand, this result together with the result (ii) of Theorem 1 in QZW also implies that h should be chosen small if it is desirable to have a chart that is less affected by the correlation among within-profile observations. In the case when $n_i h$ are large, the result (ii) reveals how the asymptotic distribution of $T_{t,h,\lambda}$ depends on $\gamma(x_1, x_2)$, which might be helpful in future research to modify $T_{t,h,\lambda}$ such that the modified version would incorporate the correlation function $\gamma(x_1, x_2)$ more effectively. As another example, according to result (i) of Theorem 2 in QZW, after the profile mean function changes from $g_0(x)$ to $g_1(x)$, the asymptotic distribution of the charting statistic $T_{t,h,\lambda}$ would depend on $\delta(x) = g_1(x) - g_0(x)$ and $\delta''(x)$. Therefore, if the curvature of $\delta(x)$ is bigger, then the corresponding shift is easier to detect, which has been confirmed in the numerical examples presented in Section 3 of QZW. See, for instance, Table 2 in QZW, where the curvature of $\delta(x)$ is much larger with OC model (ii) than the curvature of $\delta(x)$ with OC model (i).

5 Generalizations to Multivariate Cases

We appreciate Tsung's comments on possible generalization of our proposed method to multivariate cases. He provides a nice description about several potential applications of multivariate profile monitoring and about some related research. We believe that his discussion provides us a strong motivation to study profile monitoring in multivariate cases in future research. In their discussion, Woodall, Birch and Du also provide some comments on this topic, and they think that other smoothing techniques such as the smoothing spline ANOVA might be more convenient to use, compared to the local polynomial kernel smoothing used in QZW. Frankly, we do not have much experience in multivariate cases, but would still like to share with readers some of our initial impressions described below.

Multivariate profile monitoring can be roughly classified into the following three categories:

(i) each profile has one response and multiple covariates,

(ii) each profile has multiple responses and one covariate, and

(iii) each profile has multiple responses and multiple covariates.

For category (i), some semi-parametric modeling methods might be useful to describe the complicated high-dimensional profiles. For instance, the single-index and partial linear models (cf., Ruppert et al. 2003), which has a relatively simple interpretation of the effect of each covariate on the response, might be appropriate in certain cases for describing multivariate profiles. As an example, one type of multivariate profile monitoring problem can be described using partial linear modeling as follows.

$$y_{ij} = \begin{cases} g_0(t_{ij}) + \mathbf{X}_i\boldsymbol{\beta} + f_i(t_{ij}) + \varepsilon_{ij}, \text{ for } j = 1, 2, \ldots, n_i, i = 1, \ldots, \tau, \\ g_1(t_{ij}) + \mathbf{X}_i\boldsymbol{\beta} + f_i(t_{ij}) + \varepsilon_{ij}, \text{ for } j = 1, 2, \ldots, n_i, i = \tau + 1, \ldots, \end{cases}$$

where t denotes a univariate covariate that has a nonparametric relationship with the response y, $\mathbf{X}$ denotes multiple covariates that affect y linearly, τ is an unknown change-point, $\boldsymbol{\beta}$ is a coefficient vector, and other quantities are the same as those used in QZW. This model describes cases when the nonparametric relationship between y and t has a shift at τ. Obviously, similar models can be formulated for cases when the linear relationship between y and $\mathbf{X}$ has a shift. By combining existing semi-parametric model estimation methods and process control schemes, we believe that appropriate control charts can be constructed for monitoring such multivariate profiles in a way that is similar to the chart MENPC.

For category (ii), assume that we have p responses, and the observed IC data are from the following multivariate nonparametric mixed-effects model:

$$\boldsymbol{y}_{ij} = \boldsymbol{g}(x_{ij}) + \boldsymbol{f}_i(x_{ij}) + \boldsymbol{\varepsilon}_{ij}, \text{ for } j = 1, 2, \ldots, n_i, i = 1, \ldots, m,$$

where $\boldsymbol{g}(x) = (g_1(x), \ldots, g_p(x))^T$ is the fixed-effects term, $\boldsymbol{f}_i(x) = (f_{i1}(x), \ldots, f_{ip}(x))^T$ is the random-effects term, $\boldsymbol{y}_{ij} = (y_{ij1}, \ldots, y_{ijp})^T$, and $\text{Cov}(\boldsymbol{\varepsilon}_{ij}) = \boldsymbol{\Sigma}$. For a given point $s \in [0, 1]$, $\boldsymbol{g}(s)$ and $\boldsymbol{f}_i(s)$ can be estimated by minimizing the following penalized local linear likelihood function which is similar to expression (2) in QZW.

$$\sum_{i=1}^{m}\Bigg\{ \quad \mathrm{tr}\{[\boldsymbol{Y}_i - \boldsymbol{Z}_i(\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\alpha}}_i)]\boldsymbol{\Sigma}^{-1}[\boldsymbol{Y}_i - \boldsymbol{Z}_i(\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\alpha}}_i)]^T\mathbf{K}_i\} \\ +[\mathrm{vec}(\tilde{\boldsymbol{\alpha}}_i)]^T\mathbf{D}^{-1}\mathrm{vec}(\tilde{\boldsymbol{\alpha}}_i) + \ln|\mathbf{D}| + n_i\ln|\boldsymbol{\Sigma}|\Bigg\},$$

where

$$\boldsymbol{Y}_i = (\boldsymbol{y}_{i1},\ldots,\boldsymbol{y}_{in_i})^T, \quad \tilde{\boldsymbol{\beta}} = (\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_p), \quad \tilde{\boldsymbol{\alpha}}_i = (\boldsymbol{\alpha}_{i1},\ldots,\boldsymbol{\alpha}_{ip});$$
$$\boldsymbol{Z}_i = (\boldsymbol{z}_{i1},\ldots,\boldsymbol{z}_{in_i})^T, \quad \boldsymbol{z}_{ij}^T = (1, x_{ij} - s),$$

each $\boldsymbol{\beta}_j$ is a deterministic two-dimensional coefficient vector, each $\boldsymbol{\alpha}_{ij}$ is a two-dimensional vector of the random effects with mean $\mathbf{0}$ and covariance $\mathbf{D}_j$, $\mathbf{D} = \mathrm{diag}\{\mathbf{D}_1,\ldots,\mathbf{D}_p\}$, and $\mathbf{K}_i$ are defined in (4) of QWZ. Then, a similar iterative algorithm to the one described in Section 2.2 of QZW can be developed for Phase I model estimation. The local weighted negative log likelihood estimation and a corresponding Phase II control chart can also be developed in a similar way to that described in Section 2.3 of QZW.

Category (iii) is much more complicated than the previous two categories. Probably certain appropriate combinations of the models described above can handle some special cases. This is an important and challenging area and serious research will be required to develop effective monitoring schemes.

Acknowledgments: This research is supported in part by the grant DMS-0721204 from NSF of USA and the grants 10771107 and 10711120448 from NNSF of China.

Additional References

Dette, H., and Neumeyer, N. (2001), "Nonparametric Analysis of Covariance," *The Annals of Statistics*, **29**, 1361–1400.

Fung, W., Zhu, Z., Wei, B., and He, X. (2002), "Influence Diagnostics and Outlier Tests for Semiparametric Mixed Models," *Journal of Royal Statistical Society (Series B)*, **64**, 565–579.

Fitzmaurice, G., Davidian, M., Verbeke, G., and Molenberghs, G. (2008), *Longitudinal Data Analysis*, London: Chapman & Hall/CRC.

Liang, K.Y., and Zeger, S. (2002), *Analysis of Longitudinal Data (2nd Ed.)*, New York: Oxford University Press.

Neumeyer, N., and Dette, H. (2003), "Nonparametric comparison of regression curves: an empirical process approach," *The Annals of Statistics*, **31**, 880–920.

Noorossana, R., Amiri, A., and Soleimani, P. (2008), "On the Monitoring of Autocorrelated Linear Profiles," *Communications in Statistics-Theory and Methods*, **37**, 425–442.

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

Wang, K., and Tsung, F. (2005), "Using profile monitoring techniques for a data-rich environment with huge sample sizes," *Quality and Reliability Engineering International*, **21**, 677–688.

Young, S. G., and Bowman, A. W. (1995), "Non-Parametric Analysis of Covariance," *Biometrics*, **51,** 920–931.