**Statistical Process Control for Multistage Processes with Non-repeating Cyclic Profiles**

Wenmeng Tian[1], Ran Jin[1], Tingting Huang[2] and Jaime A. Camelio[1]

[1]Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061, USA

[2]School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China

## Abstract

In many manufacturing processes, process data are observed in the form of time-based profiles, which may contain rich information for process monitoring and fault diagnosis. Most approaches currently available in profile monitoring focus on single-stage processes or multistage processes with repeating cyclic profiles. However, a number of manufacturing operations are performed in multiple stages, where non-repeating profiles are generated. For example, in a broaching process, non-repeating cyclic force profiles are generated by the interaction between each cutting tooth and the workpiece. This paper presents a process monitoring method based on Partial Least Squares (PLS) regression models, where PLS regression models are used to characterize the correlation between consecutive stages. Instead of monitoring the non-repeating profiles directly, the residual profiles from the PLS models are monitored. A Group Exponentially Weighted Moving Average (GEWMA) control chart is adopted to detect both global and local shifts. The performance of the proposed method is compared with conventional methods in a simulation study. Finally, a case study of a hexagonal broaching process is used to illustrate the effectiveness of the proposed methodology in process monitoring and fault diagnosis.

**Keywords:** Broaching, cyclic signals, EWMA control chart, multistage manufacturing processes, partial least squares regression

## 1. Introduction

In a data-rich manufacturing environment, high-density data of process conditions are continuously collected over time. These process data generate profiles, which are widely used for process monitoring

and fault diagnosis. In general, the objective of process monitoring is to detect any process change as soon as it occurs to prevent quality losses.

Profile monitoring has received considerable attention in the Statistical Process Control (SPC) literature, and it has been used in various applications including calibration process (Kang and Albin, 2000), healthcare and public health surveillance (Woodall, 2006), and a lot of manufacturing process such as turning (Colosimo *et al.*, 2008), welding, stamping (Jin and Shi, 2001), and semiconductor manufacturing (Jin and Liu, 2013). In profile monitoring, process quality is characterized by the relationship between the response variable and the explanatory variables (Kim *et al.*, 2003; and Woodall *et al*., 2004). Approaches have been developed for monitoring linear profiles, including Kang and Albin (2000); Kim *et al.* (2003); and Zou *et al.* (2006). Also, as a growing number of process variables demonstrate nonlinear relationships with each other, extensive research efforts have been focused on nonlinear profile monitoring. Current nonlinear profile monitoring schemes include parametric (Ding *et al.*, 2006; and Jensen and Birch, 2009) and nonparametric approaches (Zou *et al*., 2008; Qiu *et al.*, 2010; Paynabar and Jin, 2011), most of which have been summarized in Noorossana *et al*. (2011).

In recent years, a special class of nonlinear profiles, called cyclic or cycle-based signals, has been studied. Cyclic signals usually refer to signals collected from repeating operations, such as stamping processes (Jin and Shi, 1999a, 2000, 2001; Zhou *et al.*, 2006) and forging processes (Zhou and Jin, 2005; Zhou *et al.*, 2005; Wang *et al.*, 2009; and Yang and Jin, 2012). As the profiles are obtained from repeating operations, they are presumed to follow the same or similar statistical distribution under normal operating conditions. To analyze these profiles, current approaches include signal compression based on wavelet transformation and denoising (Jin and Shi, 1999a), principal components analysis and clustering methods (Zhou and Jin, 2005), principal curve method (Kim *et al.*, 2006), and sparse component analysis method (Yang and Jin, 2012). In terms of SPC, some techniques that have been considered include directionally variant control chart systems for both known and unknown fault detection (Zhou *et al.*, 2005), $T^2$ control chart based on selected levels of wavelet coefficients (Zhou *et al.*, 2006), process monitoring based on global and local variations in multichannel functional data (Wang *et al.*, 2009),

multichannel profile monitoring and diagnosis based on uncorrelated multilinear principal component analysis (Paynabar *et al*., 2013; and Paynabar *et al*., 2015), and automatic process monitoring technique based on recurrence plot methods (Zhou *et al*. 2015).

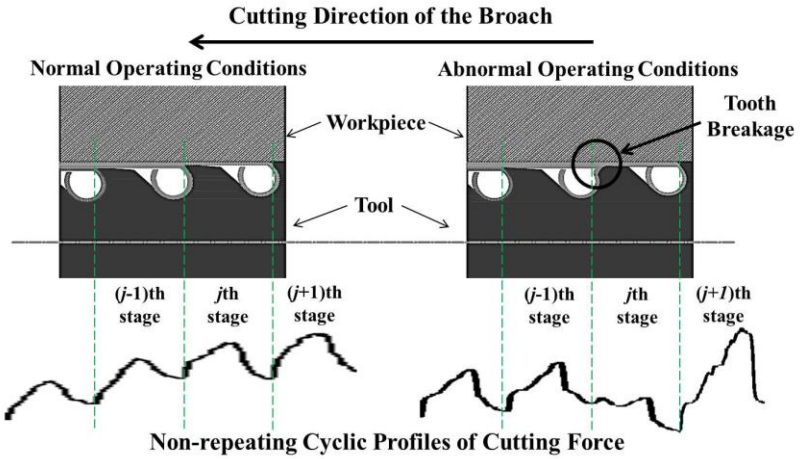


**Fig. 1.** Multistage manufacturing process with non-repeating cyclic profile outputs

In summary, most approaches in profile monitoring focus mainly on the single-stage processes with profile outputs or multistage processes with repeating cyclic profiles. However, some manufacturing operations consist of multiple stages and the conditions of different stages are characterized as non-repeating cyclic profiles. For example, in a broaching process, the desired contour in a part is sequentially shaped through material removal by multiple teeth. The performance of each tooth or each set of teeth can be reflected by the cutting force collected over time, i.e., a cutting force profile (Axinte and Nabil, 2003; and Shi *et al.*, 2007). Each tooth or each set of teeth is considered as a stage in this process. The cutting force profiles of the downstream tooth or set of teeth are largely affected by amount of material removed by previous teeth and the condition of the currently cutting tooth (Robertson *et al.*, 2013). As illustrated in Fig. 1, when a tooth breaks at the *j*th stage on the right panel, the cutting force profile of the *j*th stage becomes smaller than it should be under normal conditions, due to the changed dimension of the tooth, and the uncompleted material removal left by the *j*th stage is accomplished by the (*j*+1)th stage. Thus, the cutting force profile of the (*j*+1)th stage becomes larger than it is under normal conditions. In Fig. 2, four broaching force profiles from different operating conditions are shown. Due to the engagement and disengagement of the multiple teeth on the broach, the profiles collected under normal operating

conditions demonstrate cyclic patterns. In addition, given that the material removal rates for different teeth are based on different designed geometries of the teeth, the cutting force profiles at different teeth do not follow the same or similar statistical distribution. Moreover, there are global and local shifts which can both possibly occur in the process. Therefore, the repeating cyclic profile monitoring methods in the literature may not be effective for such a process like broaching.

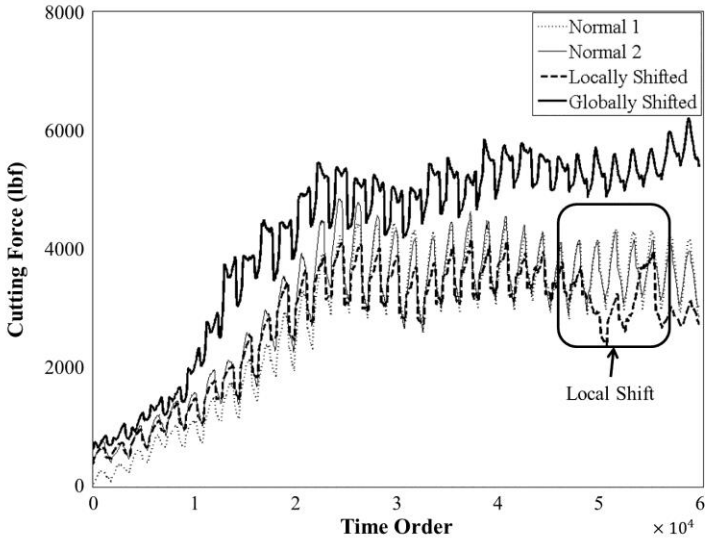


**Fig. 2.** Cutting force profiles collected from broaching process

Even though this paper is motivated by broaching process with non-repeating cyclic profiles, it should be noted that non-repeating cyclic profiles are present in various multistage manufacturing processes, including temperature change over time profiles collected in fruit drying processes (Ho *et al.*, 2002), CaS level in a $SO_2$ reduction chemical conversion processes (Sohn and Kim, 2002), and many other machining operations such as end milling (Sutherland and DeVor, 1986).

In this paper, profiles collected from individual stages are called *non-repeating cyclic profiles*. It should be noted that the well-studied repeating cyclic operations can be regarded as special cases for the non-repeating cyclic profiles. Additionally, the profiles at all stages are usually continuously collected in temporal order, which results in one profile for each multistage process. The profile containing the information of all the stages in the entire multistage process is called the *original profile* of the process.

For example, Fig. 2 demonstrates four original profiles collected from the broaching process under different operating conditions.

The objective of this paper is to detect a shift in multistage manufacturing processes with non-repeating cyclic profiles, and this paper focuses on simultaneously detecting two types of process mean shifts, as illustrated in Fig. 2, namely the global shifts and the local shifts. The global shifts indicate the process changes resulting in mean shifts occurring at all the stages in the same direction, such as misalignment of the broach tool or large pilot hole in the workpiece, while the local shifts are the process changes leading to mean shifts of the profiles at only one or its adjacent stages, which indicates some change of distribution in material removed by those adjacent teeth, such as wear or breakage of a tooth.

In the monitoring of a non-repeating cyclic profile, a common baseline distribution cannot be identified for profiles at multiple stages. Instead, the correlation between profiles from consecutive stages should be modeled and monitored. In the literature, such modeling and monitoring methods have been widely used in multistage manufacturing processes, including cause-selecting charts (Zhang, 1984, 1985, and 1992), and regression adjustment approaches (Hawkins, 1991 and 1993). Furthermore, Jin and Shi (1999b) proposed the use of linear state space models to characterize variation propagation between multiple stages. Zantek *et al.* (2006) proposed to use simultaneous CUSUM charts to monitor multiple prediction errors at the same time. A comprehensive review of the approaches and extensions related to multistage process monitoring is given by Tsung *et al*. (2008). Furthermore, Xiang and Tsung (2008) proposed an approach to convert the multistage process into a multi-stream process composed of the standardized One-Step ahead Forecast Errors (OSFEs) at all the stages. They adopted the group control charts proposed by Nelson (1986) to monitor a multi-stream process, which was defined as a process with several streams of outputs. Zou and Tsung (2008) developed a directional MEWMA scheme based on generalized likelihood ratio tests for multistage process monitoring. Jin and Liu (2013) proposed a control charting system to use regression tree models for serial-parallel multistage manufacturing processes. Zhang *et al*. (2015) proposed a Phase I analysis method for multivariate profile data based on functional regression adjustment and functional principal component analysis. All these methods have good

capability in identifying the potentially shifted stage. However, their methods experienced some limitations to monitor the non-repeating cyclic profiles. For example, the global shifts cannot be effectively detected by the OSFEs obtained from consecutive stages. The relationship between consecutive stages will not change when all stages experience the mean shifts with similar magnitudes in the same direction.

In this paper, a Partial Least Squares (PLS) regression model is used to characterize the relationship between the profiles from consecutive stages and thus standardized OSFEs can be obtained. Then, multiple streams are monitored simultaneously under the assumption that the OSFEs are identically distributed. The multi-stream process is comprised of the following: 1) the streams of the OSFEs which are used to detect local mean shifts in the process; 2) one stream of the global mean of the original profile, which is used for global mean shift detection. Then, a Group Exponentially Weighted Moving Average (GEWMA) control chart is used to monitor the extracted multiple streams simultaneously. In the proposed methodology, when a shift is detected, the potential root cause can be identified. This includes determining whether the shift occurs locally or globally, and locating any locally shifted stage.

The remainder of the paper is organized as follows. In Section 2, the multi-stream process extracted from the multistage process is introduced, and a multistage modeling approach based on PLS regression to obtain the OSFEs is proposed. In Section 3, a GEWMA monitoring scheme and the diagnostic approach for the proposed chart is introduced. Simulation studies are used to evaluate the performance of the process monitoring and diagnosis in Section 4. A case study of a broaching process monitoring is presented in Section 5. Finally, Section 6 draws conclusions and presents potential future research topics.

## 2. Multi-stream Extraction from Original Profiles

In this section, the configuration of the multi-stream process extracted from the original profiles is introduced. PLS regression models are proposed to model the relationship between the profiles of consecutive stages.

### 2.1. *Multi-stream Processes*

Multiple streams can be extracted from original profiles to isolate the global and local information of different stages in the process. There are quite a few methods including Wavelet Decomposition and Hilbert Huang Transformation that can extract global information from the original profile (Burrus *et al.*, 1997; Wu *et al.*, 2007). However, due to the smoothing concern in most of these transformations, the extracted global trend will always take local shift information from the original profile, which makes it difficult to determine whether the shift is a global change or a local one at a specific location. Therefore, the global mean of the original profile is used to minimize the effect of local change on the global trend information.

Considering that the total number of stages in a process is $q$, the potential streams available will include one stream of the global mean of the original profile and ($q$-1) streams obtained from the OSFEs of consecutive stages. Not only can this configuration make full use of the information contained in the original profiles, but it can also ensure that the extracted process streams have explicit engineering explanations. Each process stream has its one-to-one correspondence to one type of shift or one specific location of the local shift. The global stream is responsible for detecting global shifts; while the ($q$-1) OSFEs are responsible for detecting the local shifts at the corresponding stage.

### 2.2. *Profile Segmentation*

The original profile will be segmented such that each segment represents the output from one stage. The segmentation can provide fault diagnostic guidelines to locate the shifted stage when a shift is detected. It can be observed in Fig. 1 that each stage starts with a sharp increase of the force as a new broach tooth engages with the workpiece and ends with a sharp decrease of the force as a tooth disengages (Axinte *et al*., 2004; and Klocke *et al*., 2012). Thus, one local minimum point indicates the landmark between adjacent stages. Even if there is a local shift, the local minimum point still exists as it represents the engagement of a new tooth to the workpiece. Therefore, the original profiles can be segmented by the valleys in the profiles based on the process knowledge.

Given the sampling rate of the specific sample, the predetermined cutting speed, and the pitch of the broach, the profile valleys can be found easily by searching the minimum point in the potential scope which comes from the tool displacement relative to the workpiece and takes into account the process uncertainty, such as geometry deviation of the tooth, and sensory measurement error in the system. Denote the $i$th original profile as $\boldsymbol{C_i(t)}$ ($i$=1,2,…,$m$) which can be segmented into $q$ profiles as $\boldsymbol{X_{i,j}(t)}$ ($i$=1,2,..,$m$; $j$=1,2,…,$q$). The procedure is illustrated in Fig. 3. After segmentation, each profile only has one maximum value and characterizes the operating condition of its corresponding stage, as shown in Fig. 3. Due to the uncertainty in the process, the segmented profiles may not have the same dimension after segmentation. In that case, linear interpolation can be used to adjust the dimensions of different stages to the same value.

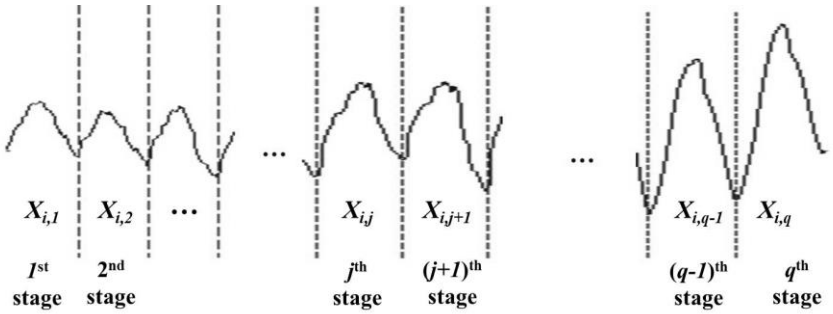


**Fig. 3.** Profile segmentation

### 2.3. *PLS Modeling*

The objective of PLS modeling is to remove the heterogeneity among the stages to obtain multiple homogeneous process streams under the normal manufacturing conditions. To model the relationships between $\boldsymbol{X_{\cdot,j}}$ and $\boldsymbol{X_{\cdot,j+1}}$ using PLS regression, the underlying assumption is that the major variation pattern in $\boldsymbol{X_{\cdot,j+1}}$ can be described with a small number of Principal Components (PCs) extracted from $\boldsymbol{X_{\cdot,j}}$.

In the PLS modeling approach, the profile of each stage is considered as a multivariate vector. A series of PLS regression models are estimated to describe the correlation between each pair of consecutive profiles. As shown in Fig. 4, the covariance between $\boldsymbol{X_{\cdot,j}}$ and $\boldsymbol{X_{\cdot,j+1}}$ are maximized by extracting $A_j$ PCs from these two matrices that could explain a predefined percentage of variance in the response matrix $\boldsymbol{X_{\cdot,j+1}}$. Then the OSFEs of these models are assumed to follow the same distribution

under the normal operating conditions. As illustrated in Equation (1), $\boldsymbol{\varepsilon_{i,j+1}}$ is the OSFE of the ($j$+1)th stage in the $i$th observation ($i$=1,2,…,$m$; $j$=1,2,…,$q$-1).

$$\boldsymbol{X_{i,j+1}} = f_j(\boldsymbol{X_{i,j}}) + \boldsymbol{\varepsilon_{i,j+1}}, \tag{1}$$

where $\boldsymbol{X_{i,j}}$ and $\boldsymbol{X_{i,j+1}}$ denote the profiles collected from the $j$th and ($j$+1)th stages in the $i$th observation. For each pair of consecutive stages $\boldsymbol{X_{\cdot,j}}$ and $\boldsymbol{X_{\cdot,j+1}}$ ($j$=1,2,…,$q$-1), one PLS regression model is estimated based on a fixed set of $m$ samples. The PLS regression models can be obtained using the following procedures (Höskuldsson, 1988).

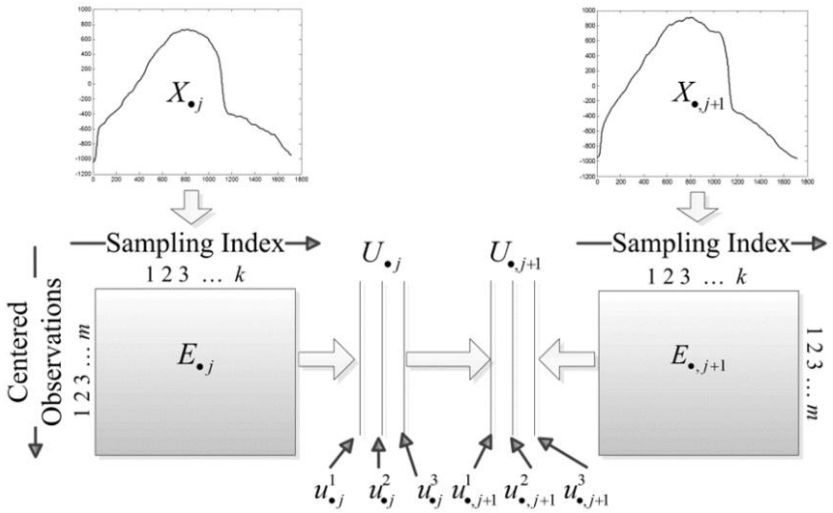


**Fig. 4.** An illustration of PLS regression model between the profiles at the consecutive stages (redrawn from Höskuldsson, 1988)

*Step 1* The estimated mean of $\boldsymbol{X_{\cdot,j}}$ and $\boldsymbol{X_{\cdot,j+1}}$ are subtracted from $\boldsymbol{X_{\cdot,j}}$ and $\boldsymbol{X_{\cdot,j+1}}$ to obtain $\boldsymbol{E_{\cdot,j}}$ and $\boldsymbol{E_{\cdot,j+1}}$, where $\boldsymbol{X_{\cdot,j}}$ is an $m \times k$ matrix containing $m$ samples of the $j$th stage. The centering is implemented by

$$\boldsymbol{E_{\cdot,j}} = \boldsymbol{X_{\cdot,j}} - \mathbf{1}\boldsymbol{\mu_j}', \tag{2}$$

where $\boldsymbol{\mu_j}$ is a $k \times 1$ vector of the column means for matrix $\boldsymbol{X_{\cdot,j}}$, and $\mathbf{1}$ is an $m \times 1$ vector of ones.

*Step 2* $A_j$ principal components (PCs) are extracted one by one from $\boldsymbol{E_{\cdot,j}}$ and $\boldsymbol{E_{\cdot,j+1}}$, respectively, to exceed the total percentage of explained variance in $\boldsymbol{E_{\cdot,j+1}}$, denoted as $p_m$. The PCs are extracted based on the following iterations:

*Step 2.0* Initialize $\boldsymbol{E}_{\cdot j}^{\boldsymbol{0}} = \boldsymbol{E}_{\cdot j}$, $\boldsymbol{E}_{\cdot j+1}^{\boldsymbol{0}} = \boldsymbol{E}_{\cdot j+1}$,and $\boldsymbol{u}_{\cdot j+1}^{\boldsymbol{0}}$= first column of $\boldsymbol{E}_{\cdot j+1}^{\boldsymbol{0}}$;

For $a = 1{:}\, A_j$

Extract the PCs $\boldsymbol{u}_{\cdot j}^{\boldsymbol{a}}$ and $\boldsymbol{u}_{\cdot j+1}^{\boldsymbol{a}}$ from $\boldsymbol{E}_{\cdot j}^{\boldsymbol{a-1}}$ and $\boldsymbol{E}_{\cdot j+1}^{\boldsymbol{a-1}}$ by the following sub-steps

*Step 2.1* Calculate the coefficient vector $\boldsymbol{q}_{\boldsymbol{j}}^{\boldsymbol{a}}$ of the *a*th PC for $\boldsymbol{E}_{\cdot j}^{\boldsymbol{a-1}}$ by $\boldsymbol{q}_{\boldsymbol{j}}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j}^{\boldsymbol{a-1}}\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a-1}}/(\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a-1}\prime}\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a-1}})$, where $\boldsymbol{q}_{\boldsymbol{j}}^{\boldsymbol{a}}$ is usually called the weight vector for the *a*th PC, and then scale $\boldsymbol{q}_{\boldsymbol{j}}^{\boldsymbol{a}}$ to be length of one

*Step 2.2* Calculate the *a*th PC $\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}$ for the *j*th stage by $\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j}^{\boldsymbol{a-1}}\boldsymbol{q}_{\boldsymbol{j}}^{\boldsymbol{a}}$

*Step 2.3* Calculate the coefficient vector $\boldsymbol{q}_{\boldsymbol{j+1}}^{\boldsymbol{a}}$ of the *a*th PC for $\boldsymbol{E}_{\cdot j+1}^{\boldsymbol{a-1}}$ by $\boldsymbol{q}_{\boldsymbol{j+1}}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j+1}^{\boldsymbol{a-1}}\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}/(\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}\prime}\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}})$ and then scale $\boldsymbol{q}_{\boldsymbol{j+1}}^{\boldsymbol{a}}$ to be length of one

*Step 2.4* Calculate the *a*th PC $\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a}}$ for the (*j*+1)th stage by $\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j+1}^{\boldsymbol{a-1}}\boldsymbol{q}_{\boldsymbol{j+1}}^{\boldsymbol{a}}$

*Step 2.5* The loadings vectors of the *a*th PC for the *j*th and (*j*+1)th stage can be estimated by least squares $\boldsymbol{f}_{\boldsymbol{j}}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j}^{\boldsymbol{a-1}\prime}\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}/\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}\prime}\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}$ and $\boldsymbol{f}_{\boldsymbol{j+1}}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j+1}^{\boldsymbol{a-1}\prime}\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a}}/\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a}\prime}\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a}}$

*Step 2.6* Regress $\boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a}}$ on $\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}$ to estimate the regression coefficients for the *a*th PC by $\boldsymbol{B}_{\boldsymbol{j}}^{\boldsymbol{a}} = \boldsymbol{u}_{\boldsymbol{j+1}}^{\boldsymbol{a}\prime}\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}/(\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}\prime}\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}})$

*Step 2.7* Subtract the extracted PCs to obtain new residuals by $\boldsymbol{E}_{\cdot j}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j}^{\boldsymbol{a-1}} - \boldsymbol{u}_{\cdot j}^{\boldsymbol{a}}\boldsymbol{f}_{\boldsymbol{j}}^{\boldsymbol{a}\prime}$, and $\boldsymbol{E}_{\cdot j+1}^{\boldsymbol{a}} = \boldsymbol{E}_{\cdot j+1}^{\boldsymbol{a-1}} - \boldsymbol{B}_{\boldsymbol{j}}^{\boldsymbol{a}}\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}\boldsymbol{q}_{\boldsymbol{j+1}}^{\boldsymbol{a}\prime}$

End

The extracted PCs can explain a large proportion of the variance. $\boldsymbol{Q}_{\boldsymbol{j}}$ is a $k \times A_j$ transformation matrix containing all the $\boldsymbol{q}_{\boldsymbol{j}}^{\boldsymbol{a}}$'s ($a = 1,2,\ldots,A_j$), and $\boldsymbol{U}_{\cdot \boldsymbol{j}}$ is the extracted principal components containing all the $\boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{a}}$'s ($a = 1,2,\ldots,A_j$).

$$\boldsymbol{U}_{\cdot \boldsymbol{j}} = \boldsymbol{E}_{\cdot \boldsymbol{j}}\boldsymbol{Q}_{\boldsymbol{j}}\,. \tag{3}$$

After extracting $A_j$ PCs, there is still a small portion of variance that has not been modeled, which is regarded as noise. That component is calculated by

$$\boldsymbol{E}_{\cdot j}^{\boldsymbol{A}} = \boldsymbol{E}_{\cdot j} - \boldsymbol{U}_{\cdot j}\boldsymbol{F}_{\cdot j}', \tag{4}$$

where $\boldsymbol{F}_{\cdot j}$ is usually named as the loading matrices and it contains all the $\boldsymbol{f}_{\boldsymbol{j}}^{\boldsymbol{a}}$'s ($a = 1,2,\dots,A_j$).

*Step 3* Based on the iterations in Step 2, a multiple linear regression model can be obtained to characterize the correlation between the PCs of the consecutive stages, where $\boldsymbol{B}_{\boldsymbol{j}}$ is the $A_j \times A_j$ diagonal coefficient matrix with $B_j^a$'s at the diagonal, and $\boldsymbol{H}_{\cdot \boldsymbol{j+1}}$ is the residual matrix,

$$\boldsymbol{U}_{\cdot \boldsymbol{j+1}} = \boldsymbol{U}_{\cdot \boldsymbol{j}}\boldsymbol{B}_{\boldsymbol{j}} + \boldsymbol{H}_{\cdot \boldsymbol{j+1}}. \tag{5}$$

*Step 4* Given Equations (2), (3), and (5), Equation (4) can be rewritten and rearranged to obtain Equation (6), where $\boldsymbol{b}_{\boldsymbol{j}}$ is the $k \times k$ transformation matrix, $\boldsymbol{L}_{\boldsymbol{j}}$ is the intercept term, and $\boldsymbol{\varepsilon}_{\cdot \boldsymbol{j+1}}$ is the $m \times k$ residual matrix.

$$\boldsymbol{X}_{\cdot \boldsymbol{j+1}} = \boldsymbol{L}_{\boldsymbol{j}} + \boldsymbol{X}_{\cdot \boldsymbol{j}}\boldsymbol{b}_{\boldsymbol{j}} + \boldsymbol{\varepsilon}_{\cdot \boldsymbol{j+1}}, \tag{6}$$

where $\boldsymbol{b}_{\boldsymbol{j}} = \boldsymbol{Q}_{\boldsymbol{j}}\boldsymbol{B}_{\boldsymbol{j}}\boldsymbol{F}'_{\cdot \boldsymbol{j+1}}$, $\boldsymbol{L}_{\boldsymbol{j}} = \mathbf{1}\boldsymbol{\mu}'_{\boldsymbol{j+1}} - \mathbf{1}\boldsymbol{\mu}'_{\boldsymbol{j}}\boldsymbol{Q}_{\boldsymbol{j}}\boldsymbol{B}_{\boldsymbol{j}}\boldsymbol{F}'_{\cdot \boldsymbol{j+1}}$, and $\boldsymbol{\varepsilon}_{\cdot \boldsymbol{j+1}} = \boldsymbol{E}_{\cdot \boldsymbol{j+1}}^{\boldsymbol{A}} - \boldsymbol{H}_{\cdot \boldsymbol{j+1}}\boldsymbol{F}_{\cdot \boldsymbol{j+1}}{}'$.

The computing algorithm used for PLS regression modeling is the Nonlinear Iterative Partial Least Squares (NIPALS) algorithm (Wold, 1966). For more information about PLS regression models and the computing algorithms, see Geladi and Kowalski (1986) and Wold *et al*. (2001).

The number of stages ($q$) is determined by how many non-repeating cycles in the process. For example, in a broaching process, the number of stages is defined as the total number of teeth on the broaching tool. The value of the parameter $A_j$ could be regarded as the tuning parameter in the proposed method. In the literature, there are quite a few methods proposed to select the number of components extracted, among which the methods of minimizing the Predictive Residual Sum of Squares (PRESS) based on cross validation (Wold *et al*., 2001; Rosipal and Krämer, 2006) are the most popular. However, in the proposed method, the balance among the accuracies of the models between different stages is more of concern than optimizing the modeling accuracy of every model individually. Therefore, the $A_j$ value for

the PLS regression model between the $j$th and ($j$+1)th stages has been selected based on the training dataset by searching for the smallest number of components which could exceed a predefined percentage, i.e., 95%, of the total variance explained in the profiles at the ($j$+1)th stage. Additionally, the ARL performance of the proposed method given different percentages of total variance explained $p_m$ are summarized in the Supplemental Material A, which illustrates high robustness in the proposed control charting system.

The required number of reference samples ($m$) depends on the dimension of the data and the correlation within the data. For strongly correlated data sets, the number of reference samples $m$ can be smaller than the dimension of the data yet a reasonably robust estimation can still be obtained (Chun and Keleş, 2010).

Based on the reference data set, ($q$-1) PLS regression models can be estimated. For future observations, ($q$-1) OSFEs can be calculated by subtracting the predicted profiles from the observed profiles. The OSFEs can be used to characterize the working condition of each stage, which can be monitored simultaneously as ($q$-1) streams in a control charting scheme.

## 3. GEMWA Monitoring and Diagnostic Scheme for Global and Local Shift Detection

Given multiple streams extracted from original profiles, a Group Exponentially Weighted Moving Average (GEWMA) monitoring scheme is adopted to monitor the different streams simultaneously and detect global and local shifts.

### 3.1. *GEWMA Monitoring Scheme for Simultaneous Global and Local Shift Detection*

To use the GEWMA chart to monitor a multi-stream process, the streams should follow two assumptions. First, each stream should have the same mean and same variation. Second, the distribution of each single stream should be "approximately normally distributed" (Nelson, 1986). To satisfy the first assumption, the multiple streams extracted in Section 2 can be scaled to have the mean and variation based on their distribution. For the second assumption, the normality assumptions of the multiple streams will be

validated later in the case study based on the real data set. Furthermore, the EWMA chart used in the proposed method is usually quite robust to non-normal distributions (Stoumbos and Sullivan, 2002).

The GEWMA monitoring statistics based on the global mean and the OSFEs are calculated to detect the global and local shifts. First, the deviations from nominal $\boldsymbol{\varepsilon_{i,j}}(i = 1,2,\dots,m, and\ j = 1,2,\dots,q)$ are calculated as

$$\begin{cases} \varepsilon_{ij} = \overline{c_i} - \bar{\bar{c}}, & \text{if } j = 1 \\ \boldsymbol{\varepsilon_{ij}} = \boldsymbol{X_{ij}} - \widehat{\boldsymbol{X_{ij}}}, & \text{if } j = 2,3,\dots,q \end{cases}, \tag{7}$$

where $\varepsilon_{i1}$ denotes the deviation from nominal for the global mean of the original profiles, $\overline{c_i}$ denotes the global mean of the *i*th original profile, and $\bar{\bar{c}}$ is the average global mean obtained from the reference samples. Then the GEWMA statistics can be calculated based on the Equations (8) and (9) in which the $\boldsymbol{\varepsilon_{i,j}}$'s of the same stream are exponentially weighted across the sequential observations, and the multivariate EWMA statistic can be calculated based on the method proposed by Lowry *et al.* (1992). That is

$$\begin{cases} Z_{i1} = \lambda\varepsilon_{i1} + (1-\lambda)Z_{i-1,1}, & \text{if } j = 1 \\ \boldsymbol{Z_{ij}} = \lambda\boldsymbol{\varepsilon_{ij}} + (1-\lambda)\boldsymbol{Z_{i-1,j}}, & \text{if } j = 2,3,\dots,q \end{cases}, \tag{8}$$

where $0 < \lambda < 1$ is the smoothing parameter in traditional EWMA charts.

$$T_{ij}^2 = \begin{cases} Z_{i1}^2/\sigma_{Z1}^2, & \text{if } j = 1 \\ \boldsymbol{Z_{ij}'}(\sum_{\boldsymbol{Z}}^{\boldsymbol{j}})^{-1}\boldsymbol{Z_{ij}}, & \text{if } j = 2,3,\dots,q \end{cases}, \tag{9}$$

where

$$\begin{cases} \sigma_{Z1}^2 = \frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]\sigma_{\varepsilon 1}^2, & \text{if } j = 1 \\ \sum_{\boldsymbol{Z}}^{\boldsymbol{j}} = \frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]\sum_{\boldsymbol{\varepsilon}}^{\boldsymbol{j}}, & \text{if } j = 2,3,\dots,q \end{cases}. \tag{10}$$

Here $\sigma_{\varepsilon 1}^2$ denotes the variance of the deviations from the nominal of the global mean, and it is estimated based on the $\varepsilon_{i1}$'s of the historical data; while $\sum_{\boldsymbol{\varepsilon}}^{\boldsymbol{j}}$ represents the estimated variance covariance matrix of the OSFEs of the *j*th stage, and it is estimated using the method based on successive differences, which is recommended by Sullivan and Woodall (1995).

As the $T_{ij}^2$ statistic approximately follows a $\chi^2$ distribution with degree of freedom value $p_j$ and

$$p_j = \begin{cases} 1, \text{if } j = 1 \\ k, \text{if } j = 2,3,\dots,q \end{cases}, \tag{11}$$

the $T_{ij}^2$ statistics have different mean and variances given various $p_j$ values. To deal with the dimensional inconsistency in all the streams, the adjustment suggested in Xiang and Tsung (2008) is performed as

$$\tilde{T}_{ij}^2 = \frac{T_{ij}^2 - p_j}{\sqrt{2p_j}}. \tag{12}$$

Therefore, the $\tilde{T}_{ij}^2$ statistics of all the streams have the same mean of zero and the same variance of one. The GEMWA statistics of the *i*th observation can thus be defined as the largest $\tilde{T}_{ij}^2$ statistics of all the streams, which is

$$MZ_i = max_{1 \le j \le q}\left(\tilde{T}_{ij}^2\right), \tag{13}$$

for *i*=1,2,…,*m*. The control limit of the chart is given by

$$h = L\frac{\lambda}{2-\lambda}, \tag{14}$$

where $L$ and $h$ are design parameters of the chart. Given a desired $ARL_0$ and a predefined $\lambda$ value, the value of $L$ and $h$ can be obtained by simulation. The control chart signals when an $MZ_i$ value exceeds the control limit $h$, indicating a shift is detected in the process.

In addition, although the non-repeating cyclic profiles are not assumed to be identically distributed, it is assumed that the monitoring statistics of the OSFEs at all the stages approximately follow the same distribution after the adjustment based on the dimension of the OSFEs. Therefore, after the calculation of Equation (12), the streams derived from the OSFEs can be monitored with one group chart while the global stream can be monitored separately if its dimension varies significantly from the other streams.

### 3.2. *The Diagnostic Approach of the Proposed Control Chart*

Once a process shift is detected, the fault diagnostic approach is used to determine what type of shift has occurred. If there is a local shift detected, the diagnostic approach will further locate the shifted stage. As illustrated in Equation (13), the maximum of all the $\tilde{T}_{ij}^2$ statistics is used as the monitoring statistics.

Xiang and Tsung (2008) suggested that the shifted stage should be $\text{argmax}_{1\le j\le q}(\tilde{T}_{ij}^2)$ for the *i*th observation. Therefore, the diagnostic operation can be performed by two steps:

*Step 1* Determine if there is a global shift occurring by checking if $\tilde{T}_{i1}^2 > h$ is true. If the inequality holds, then a global shift occurs in the process.

*Step 2* Determine if there is a local shift by checking if $\tilde{T}_{ij}^2 > h$ $(j = 2,3 \dots, q)$, and if any, identify the shifted stage as $\text{argmax}_{2\le j\le q}(\tilde{T}_{ij}^2)$.

Based on the monitoring scheme and fault diagnostic approach proposed in this section, detection of global and local shifts and fault diagnosis can be achieved simultaneously.

## 4. Performance Analysis using Simulation

The performance of the process monitoring and diagnosis is evaluated using a simulation study. Following similar patterns of the real faults and force signals from the broaching process, original profiles are generated by summing up a signal component of global trend and a component with non-repeating cyclic patterns.

In the simulation, the number of stages in the process is $q$=8. The in-control $ARL_0$ are adjusted to 370 and the value of the smoothing constant $\lambda$ used is 0.2. Here $\lambda$ is a user specified parameter. If historical data has higher weight, the value of $\lambda$ should be smaller, or vice versa. The out-of-control signals have three types of shifts:

(1) Global shift: the global component has a mean shift as $\boldsymbol{\mu_\epsilon} = \delta \times diag(\sum_{G0})^{1/2}$, where $\delta$ and $\sum_{G0}$ are the magnitude of the mean shift and variance covariance matrix of the global trend of the in-control profiles, respectively.

(2) Local wear: the *j*th stage has a mean shift as $\boldsymbol{\mu_\varepsilon} = \delta \times diag(\sum_{\boldsymbol{\varepsilon}}^{\boldsymbol{j}})^{1/2}$, where $\delta$ and $\sum_{\boldsymbol{\varepsilon}}^{\boldsymbol{j}}$ are the magnitude of the shift and the variance covariance matrix of the added noise, respectively.

(3) Local breakage: the *j*th stage has a mean shift as $\boldsymbol{\mu_\varepsilon} = -\delta \times diag(\sum_{\boldsymbol{\varepsilon}}^{\boldsymbol{j}})^{1/2}$; while its next adjacent stage has a mean shift as $\boldsymbol{\mu_\varepsilon} = \delta \times diag(\sum_{\boldsymbol{\varepsilon}}^{\boldsymbol{j+1}})^{1/2}$, where $\delta$ and $\sum_{\boldsymbol{\varepsilon}}^{\boldsymbol{j}}$ represent the magnitude of the shift and the variance covariance matrix of the added noise at the *j*th stage, respectively.

Such mean shift patterns are generated in the spirit of local wear and breakage from the real case study. The magnitudes of the mean shifts are set as $\delta = 0.1, 0.25, 0.5, 1, 1.5, 2, 2.5, 3$ for all three types of shifts. The ARL is obtained by averaging 10,000 run lengths. The steady-state performances of the charts are compared by first generating ten normal observations in each run and then generating observations with sustained global/local mean shift until the chart signals. Any premature signals during the first ten normal observations are discarded.

To deal with the singularity problem in the OSFEs' variance covariance matrix, PCs are extracted from the OSFEs, and the number of PCs is determined by the percentage of variance explained in the PCs, i.e., any PC which explains more than 0.05% of the total variance in the OSFE will be retained in the monitoring statistics. An excessively large threshold percentage will lead to loss of information in the shifts, while a very small threshold percentage cannot solve the singularity problem. Based on the training data set, the numbers of PCs extracted are listed in Table 1. It is clear that the numbers listed are significantly greater than the dimension of the global stream. Therefore, one one-dimensional EWMA chart is used to monitor the global trend while the GEWMA chart is used to monitor the streams of the OSFEs. The decision rule of the control chart system is that once either one of the two charts signals, the system signals. To achieve comparable in-control ARL with the benchmark methods, false alarm rates can be applied to the two charts based on their relative costs of false alarms of the two charts and potential loss due to slow detection of different process changes by Bonferroni correction (Wu *et al.*, 2004). In this study, the false alarm rate is evenly distributed to the one-dimensional EWMA chart and the GEWMA chart, and thus the control chart system has a combined in-control ARL of 370 approximately.

**Table 1.** Number of PCs Extracted from OSFEs

| OSFE | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| Number of PCs | 11 | 11 | 11 | 10 | 11 | 10 | 10 |

The numerical results include an Average Run Length (ARL) performance comparison with two benchmark methods; one is proposed by Zhou *et al.* (2006) and the other is proposed by Zou *et al.* (2008), with different preset parameter values, respectively. The two benchmark methods are selected because they are effective in detecting mean shifts in complicated profiles. In addition, an EWMA chart based on the features extracted from Zhou's method is used for the performance comparison. This is because the method proposed in Zhou *et al*. (2006) considers a Shewhart-type control chart, while both our proposed method and the method proposed by Zou *et al.* (2008) are EWMA-type control charts, which are usually more capable of detecting small shifts. Therefore, the EWMA chart is used to replace the Shewhart chart in Zhou's method for a fair comparison.

To compare the ARL performances, local wear and local breakage at Stages 2 and 7 are generated in the original profiles, and global shifts are also generated. The simulation results of the control chart performance are illustrated in Table 2. In this table, GEWMA denotes the proposed approach; Zhou *et al.*'s EWMA denotes the EWMA version of the approach proposed by Zhou *et al.* (2006), and the percentage of energy eliminated from the original profiles denoted as $Q$ is directly related to the number of wavelet coefficient levels $L$ involved in the control chart in Zhou *et al.* (2006). In their method, the number of wavelet coefficient levels $L$ used for constructing the monitoring statistics can vary from 1 to 6. Zou *et al.*'s EWMA denotes the method proposed by Zou *et al.* (2008), and the parameter $c$ denotes a tuning parameter to determine the bandwidth used to control the error in nonparametric profile representation. In the performance comparison, all the suggested choices of $c$ values are examined.

The best performing ARL values of all the tested methods for each scenario is bolded in Table 2. The proposed GEWMA method outperforms both benchmarking methods in most of the scenarios. In local shift detection, the proposed GEWMA scheme outperforms the two methods in all the tested cases. In global shift detection, our proposed method outperforms the two benchmark methods under most choices of their preset parameter, though when the $L$ value is as small as 1 or 2 the EWMA version of Zhou *et al.*'s method can detect global shifts more quickly. However, when the $L$ value is selected as 1 or 2, the local shift detection is significantly slower than our proposed method, especially when there is a local

breakage shift Zhou *et al.*'s EWMA chart can barely detect anything as their out-of-control ARLs are close to the in-control ARL.

**Table 2.** ARL comparison for various types of shifts

| | Shift Stage | $\delta$ | GEWMA | Zhou *et al.*'s EWMA | | | | | | Zou *et al.*'s EWMA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $L$=6 | $L$=5 | $L$=4 | $L$=3 | $L$=2 | $L$=1 | $c$=1.0 | $c$=1.5 | $c$=2.0 |
| In-control | - | 0 | 370.6 | 370.9 | 370.8 | 370.3 | 371.0 | 370.5 | 370.3 | 370.1 | 370.2 | 370.0 |
| Local wear | 2 | 0.1 | **187.7** | 311.7 | 324.9 | 313.7 | 303.8 | 340.9 | 363.7 | 365.6 | 364.4 | 368.0 |
| | | 0.25 | **13.9** | 121.1 | 182.1 | 193.9 | 194.4 | 284.6 | 350.2 | 370.3 | 359.6 | 359.6 |
| | | 0.5 | **4.0** | 21.6 | 49.3 | 72.8 | 85.7 | 189.8 | 294.8 | 337.4 | 323.9 | 335.2 |
| | | 1 | **1.9** | 5.3 | 9.6 | 15.7 | 21.8 | 80.3 | 171.7 | 256.4 | 182.2 | 263.6 |
| | | 1.5 | **1.3** | 3.1 | 5.0 | 7.4 | 10.0 | 39.7 | 93.6 | 156.3 | 71.6 | 172.4 |
| | | 2 | **1.0** | 2.3 | 3.4 | 4.7 | 6.2 | 22.3 | 55.3 | 79.6 | 29.7 | 97.6 |
| | | 2.5 | **1.0** | 1.9 | 2.6 | 3.5 | 4.5 | 14.6 | 35.4 | 37.5 | 15.2 | 50.7 |
| | | 3 | **1.0** | 1.7 | 2.2 | 2.9 | 3.5 | 10.5 | 24.3 | 19.8 | 9.9 | 26.3 |
| | 7 | 0.1 | **241.5** | 255.8 | 286.3 | 249.4 | 377.7 | 409.9 | 362.5 | 372.4 | 368.1 | 360.6 |
| | | 0.25 | **19.0** | 73.0 | 109.8 | 82.4 | 332.0 | 428.9 | 352.8 | 366.3 | 356.7 | 357.7 |
| | | 0.5 | **4.6** | 13.9 | 21.4 | 18.2 | 202.2 | 363.2 | 280.7 | 335.6 | 307.0 | 335.4 |
| | | 1 | **2.1** | 4.3 | 5.5 | 5.3 | 55.5 | 161.3 | 146.2 | 238.7 | 145.7 | 252.0 |
| | | 1.5 | **1.5** | 2.7 | 3.2 | 3.1 | 21.0 | 69.1 | 75.4 | 130.0 | 54.1 | 142.6 |
| | | 2 | **1.1** | 2.0 | 2.4 | 2.3 | 11.2 | 33.4 | 43.7 | 58.0 | 22.3 | 71.6 |
| | | 2.5 | **1.0** | 1.7 | 1.9 | 1.9 | 7.4 | 19.6 | 27.6 | 27.0 | 12.3 | 36.6 |
| | | 3 | **1.0** | 1.4 | 1.7 | 1.6 | 5.5 | 13.1 | 19.1 | 14.5 | 8.4 | 19.6 |
| Local breakage | 2 | 0.1 | **50.5** | 114.6 | 257.0 | 330.8 | 344.1 | 372.1 | 359.3 | 368.4 | 347.0 | 375.0 |
| | | 0.25 | **5.5** | 14.2 | 76.9 | 161.3 | 188.1 | 360.4 | 364.6 | 346.5 | 294.7 | 357.1 |
| | | 0.5 | **2.4** | 4.4 | 15.2 | 36.5 | 48.6 | 351.9 | 358.3 | 287.0 | 201.8 | 317.8 |
| | | 1 | **1.2** | 2.0 | 4.6 | 7.8 | 10.3 | 320.8 | 361.1 | 135.8 | 66.6 | 209.8 |
| | | 1.5 | **1.0** | 1.4 | 2.8 | 4.2 | 5.3 | 289.5 | 346.9 | 50.6 | 23.3 | 110.5 |
| | | 2 | **1.0** | 1.1 | 2.1 | 3.0 | 3.6 | 259.2 | 341.0 | 20.5 | 11.4 | 52.0 |
| | | 2.5 | **1.0** | 1.0 | 1.7 | 2.4 | 2.8 | 231.0 | 328.6 | 10.9 | 7.5 | 25.5 |
| | | 3 | **1.0** | 1.0 | 1.5 | 2.0 | 2.3 | 207.2 | 310.3 | 7.4 | 5.6 | 14.7 |
| | 7 | 0.1 | **106.2** | 162.2 | 200.0 | 162.6 | 361.3 | 373.3 | 360.5 | 374.7 | 374.9 | 369.9 |
| | | 0.25 | **8.3** | 20.9 | 31.3 | 24.7 | 363.0 | 376.0 | 360.6 | 358.9 | 373.4 | 365.0 |
| | | 0.5 | **3.0** | 5.3 | 7.0 | 6.3 | 368.9 | 379.8 | 359.2 | 313.9 | 307.3 | 344.1 |
| | | 1 | **1.6** | 2.3 | 2.7 | 2.6 | 370.8 | 377.7 | 361.7 | 178.9 | 132.3 | 261.4 |
| | | 1.5 | **1.0** | 1.7 | 1.9 | 1.8 | 370.6 | 371.3 | 354.8 | 69.0 | 36.1 | 154.0 |
| | | 2 | **1.0** | 1.2 | 1.5 | 1.4 | 361.4 | 368.6 | 352.0 | 23.0 | 12.9 | 75.8 |
| | | 2.5 | **1.0** | 1.0 | 1.1 | 1.1 | 367.2 | 375.8 | 354.0 | 10.6 | 7.9 | 32.8 |
| | | 3 | **1.0** | 1.0 | 1.0 | 1.0 | 362.7 | 372.9 | 357.4 | 7.0 | 5.8 | 15.9 |
| Global | All | 0.1 | 328.6 | 348.6 | 352.4 | 334.0 | 324.1 | 313.2 | **293.0** | 365.1 | 345.3 | 369.2 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.25 | 167.4 | 278.5 | 267.0 | 222.0 | 195.8 | 156.1 | **130.7** | 337.4 | 246.7 | 318.2 |
| 0.5 | 54.8 | 140.2 | 116.8 | 82.7 | 63.8 | 46.6 | **36.2** | 225.7 | 96.2 | 193.3 |
| 1 | 11.3 | 30.1 | 23.1 | 16.6 | 13.9 | 11.1 | **9.4** | 63.7 | 20.1 | 42.1 |
| 1.5 | 5.7 | 11.3 | 9.3 | 7.4 | 6.6 | 5.6 | **5.0** | 18.9 | 8.8 | 13.7 |
| 2 | 3.9 | 6.6 | 5.7 | 4.8 | 4.4 | 3.9 | **3.5** | 9.1 | 5.6 | 7.4 |
| 2.5 | 3.0 | 4.7 | 4.1 | 3.6 | 3.3 | 3.0 | **2.7** | 6.1 | 4.2 | 5.2 |
| 3 | 2.4 | 3.7 | 3.3 | 2.9 | 2.7 | 2.4 | **2.2** | 4.6 | 3.3 | 4.0 |

Moreover, the performances of the benchmark methods are very sensitive to the preset tuning parameters. For the method proposed in Zhou *et al.* (2006), as the *L* value gets larger the local shift detection power improves and the global shift detection power deteriorates. This can be explained by the fact that the detection power of either type of shift should be sacrificed for the other type, i.e., the optimal performance on detecting local and global shifts cannot be achieved based on the same choice of the *L* value. Furthermore, Zhou *et al.*'s method also demonstrates extremely high sensitivity when the shifted location varies. For example, given the preset parameter value $L$=3, satisfactory ARL performance can be achieved when all the tested global shifts and most of the local shifts occur. However, when there is a local breakage at Stage 7, ARL gets extremely large which indicates that the local breakage at Stage 7 is almost undetectable. For the approach proposed in Zou *et al.* (2008), the control chart achieves best performance mostly when the value of *c* is 1.5. However, the ARL performance is also not quite robust when the tuning parameter varies within the suggested range. All these results indicate that the performances of the benchmark methods are not robust enough to the preset parameter when monitoring the non-repeating cyclic profiles. The robustness of the proposed method has been studied, and the ARL performance with respect to different percentage values of total variance explained $p_m$ in the response matrices has been summarized in the Supplemental Material A.

Therefore, the proposed method gives best results for local shift detection and satisfactory performances for global shift detection. This method demonstrates better robustness when shifted location varies as shown in Table 2. The advantage of the proposed method in ARL performances lies in the

modeling of consecutive stages, which can reduce the total variance of the profiles by modeling their relationships with their preceding profiles.

To analyze the diagnostic performance, 10,000 simulation replicates are performed for the proposed approach. Table 3 summarizes the diagnostic accuracy. Motivated by the diagnostic performance metric in Zou and Tsung (2008), one can compare the probability that the global shift is identified when the control chart signals given a global shift. Furthermore, when there is a local shift, the probabilities that the estimated shifted location $\hat{\zeta}$ falls within a given interval, which centers around the actual shift location $\zeta$, are also summarized. For example, $\Pr(\hat{\zeta} = \zeta)$, $\Pr(|\hat{\zeta} - \zeta| \leq 1)$, and $\Pr(|\hat{\zeta} - \zeta| \leq 2)$ represent the probability of the estimated shifted stage index is exactly the actual shifted stage $\zeta$, the probability of the estimated shifted stage index is included in the interval from $\zeta - 1$ to $\zeta + 1$, and the probability of the estimated shifted stage index is included in the interval from $\zeta - 2$ to $\zeta + 2$, respectively. The three probabilities described above are denoted as $P_0$, $P_1$, and $P_2$.

It can be noted that the proposed method provides very precise diagnostic performance when there is a global shift. In addition, when there is a local shift, the probability that the estimated shift stage index is within the two adjacent neighbors of the actual shift location is close to 1 in most cases. It can be observed that there are a few factors that could affect the diagnostic performance, including the estimation accuracy of the regression models, the type of shift that occurs, and the magnitude of the mean shift. The relatively low diagnostic accuracy when there is a local wear at Stage 7 is due to the estimation of PLS regression model between Stage 7 and 8 is less accurate than the one between Stage 2 and 3. The diagnosis performance of the proposed method comes from the steps of profile segmentation and the modeling between consecutive stages.

**Table 3.** Diagnostic accuracy summary when using the GEWMA chart

| Shift type | Shift stage | $\delta$ | Proposed Method | | |
|---|---|---|---|---|---|
| | | | $P_0$ | $P_1$ | $P_2$ |
| Local wear | 2 | 0.1 | 0.5624 | 0.6500 | 0.6500 |
| | | 0.25 | 0.9637 | 0.9716 | 0.9716 |
| | | 0.5 | 0.9952 | 0.9985 | 0.9985 |
| | | 1 | 0.9990 | 0.9995 | 0.9995 |

| | | | | | |
|---|---|---|---|---|---|
| | | 1.5 | 0.9991 | 0.9998 | 0.9998 |
| | | 2 | 0.9999 | 1.0000 | 1.0000 |
| | | 2.5 | 1.0000 | 1.0000 | 1.0000 |
| | | 3 | 1.0000 | 1.0000 | 1.0000 |
| | 7 | 0.1 | 0.3007 | 0.4600 | 0.4600 |
| | | 0.25 | 0.6485 | 0.9597 | 0.9597 |
| | | 0.5 | 0.6745 | 0.9964 | 0.9964 |
| | | 1 | 0.7133 | 0.9989 | 0.9989 |
| | | 1.5 | 0.7456 | 0.9996 | 0.9996 |
| | | 2 | 0.7285 | 0.9998 | 0.9998 |
| | | 2.5 | 0.7853 | 1.0000 | 1.0000 |
| | | 3 | 0.8311 | 1.0000 | 1.0000 |
| Local breakage | 2 | 0.1 | 0.8681 | 0.9106 | 0.9106 |
| | | 0.25 | 0.9814 | 0.9916 | 0.9916 |
| | | 0.5 | 0.9952 | 0.9993 | 0.9994 |
| | | 1 | 0.9972 | 0.9998 | 0.9998 |
| | | 1.5 | 0.9999 | 1.0000 | 1.0000 |
| | | 2 | 1.0000 | 1.0000 | 1.0000 |
| | | 2.5 | 1.0000 | 1.0000 | 1.0000 |
| | | 3 | 1.0000 | 1.0000 | 1.0000 |
| | 7 | 0.1 | 0.6852 | 0.7666 | 0.7666 |
| | | 0.25 | 0.9309 | 0.9825 | 0.9825 |
| | | 0.5 | 0.9612 | 0.9986 | 0.9988 |
| | | 1 | 0.9798 | 0.9994 | 0.9995 |
| | | 1.5 | 0.9827 | 0.9999 | 0.9999 |
| | | 2 | 0.9969 | 1.0000 | 1.0000 |
| | | 2.5 | 0.9996 | 1.0000 | 1.0000 |
| | | 3 | 0.9999 | 1.0000 | 1.0000 |
| Global Shift | All | 0.1 | 0.5564 | - | - |
| | | 0.25 | 0.7359 | - | - |
| | | 0.5 | 0.8854 | - | - |
| | | 1 | 0.9327 | - | - |
| | | 1.5 | 0.9381 | - | - |
| | | 2 | 0.9341 | - | - |
| | | 2.5 | 0.9225 | - | - |
| | | 3 | 0.9139 | - | - |

## 5. Case Study

In this section, a push hexagonal broaching process is studied to further verify the effectiveness of the proposed method. The experimental setup is illustrated in Fig. 5. It replicates the setup of a real-world

broaching process with highly controlled displacement of the broach. This process considers 33 stages, corresponding to the 33 teeth on the hexagonal broach. During the cutting process, a real-time cutting force profile can be obtained using a force sensor embedded in the base of the tensile test machine. During the process, the base is moving upwards with a constant speed of one inch per minute.

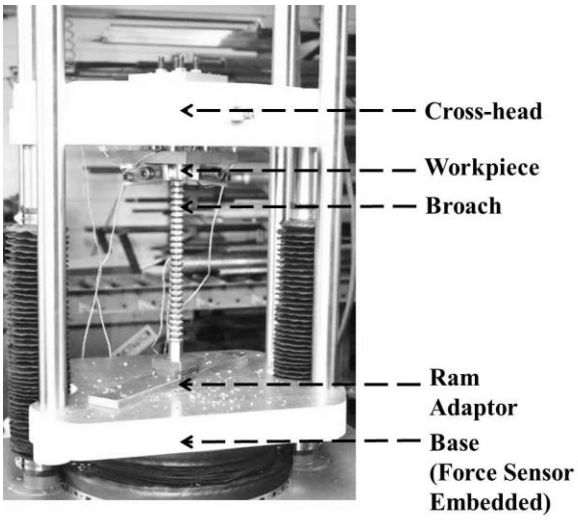


**Fig. 5.** Experiment setup (adopted from Robertson *et al.* (2013) with permission)

To illustrate the monitoring method, 20 force profiles were collected, including 15 normal runs and 5 abnormal runs. The first 10 runs were used as reference data set to estimate the model parameters. The rest of the runs were used for performance evaluation. The five abnormal runs were designed to represent the condition of tool wear on the 30th tooth of the broaching tool to test the monitoring performance. The segmented profiles were processed by linear interpolation, such that all segmented profiles had the same length (Davis, 1975). The proposed control charting system was implemented to determine if the increasing local wear could be detected. To deal with the dimensional inconsistency, one individual chart was used for the global trend stream while one group chart was used for the streams of the OSFEs.

The Phase II monitoring performance is presented in Fig. 6. The *MZ*s of the two charts are plotted against the observational order. On the left panel, the number beside each point indicates the stage index obtained from the fault diagnostic approach. The control limits for the individual chart $h_g$ and the group chart $h_l$ was obtained by using three standard deviations plus the mean of the monitoring statistics

obtained from the Phase I data. It is observed that the chart signaled as soon as there was a process change (the sixth observation in the chart), and the diagnostic stages always fell within the ±1 stage interval of the real shifted stage, while the individual chart doesn't signal.

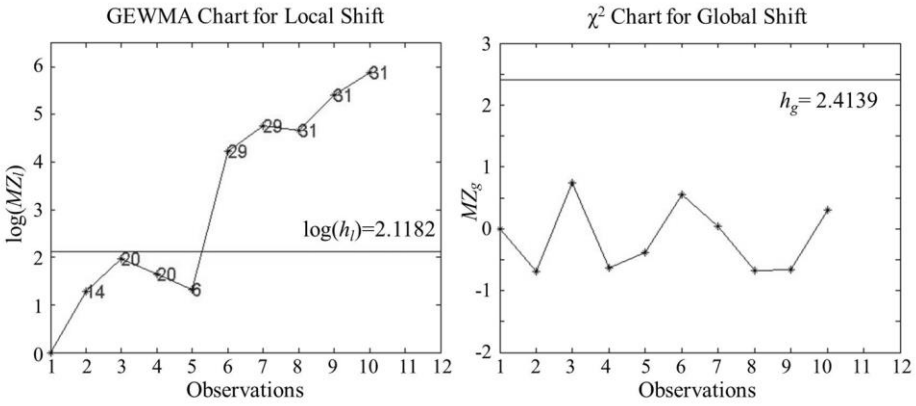


**Fig. 6.** The GEWMA chart for shift detection

Furthermore, Henze-Zirkler's multivariate normality test was performed to the extracted PCs of the OSFEs obtained from the PLS regression model at each stage (Trujillo-Ortiz *et al.*, 2007). In Fig. 7, the p-values calculated from all the stages (including the global mean component as the first stage) are plotted. There is no sufficient evidence to reject the null hypothesis that the OSFEs are normally distributed, since almost all the *p*-values associated to the tests are much larger than the predetermined significance level of 0.05. Therefore, the normality assumption is validated based on the real data, which is required for the multi-stream process monitoring.

The case study demonstrates the effectiveness of the proposed GEWMA monitoring and fault diagnostic approach in a real application. The early detection of process shifts can help reduce quality losses, and the precise shifted stage identification can lower the maintenance costs when a change is detected in the process.

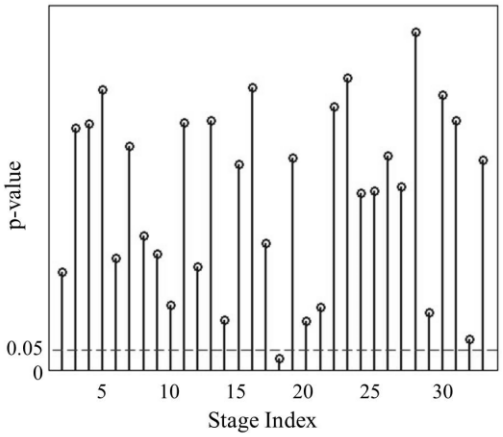


**Fig. 7.** The *p*-values of the Henze-Zirkler's multivariate normality tests

## 6. Conclusion and Future Work

Multistage processes with non-repeating cyclic profile outputs are widely encountered in various industries. However, current profile monitoring methodologies mainly focus on the profiles obtained from single-stage processes or multiple stages with repeating patterns. In this paper, a PLS regression method is proposed to model the relationship between profiles from consecutive stages. Thus, the distributional heterogeneity among the multiple stages can be removed and identically distributed OFSEs can be obtained. A Group EWMA control chart is proposed to detect global and local shifts simultaneously. An ARL simulation study shows that the proposed GEWMA chart outperforms previous methods in both detection power and robustness. In addition, the proposed diagnostic approach can accurately identify the shifted stage in the process. A case study involving a push hexagonal broaching process is used to demonstrate the effectiveness of the proposed monitoring and diagnostic methodology. Significant quality improvements are expected in multistage process with non-repeating cyclic output through proactive adjustment using more effective quality control tools.

A couple of interesting issues still remain open for future work. First, the proposed control chart is directionally invariant, which indicates that the detection power of a mean shift only depends on the shift magnitude (Montgomery, 2012). Directional variant monitoring schemes can be explored in future research by incorporating directional information to improve the monitoring performance for known

specific shift patterns. Second, the effect of signal denoising on the process monitoring performance should be studied, as data collected from real industrial settings is usually quite noisy. For example, adaptive denoising method can be used to optimize the modeling accuracy of the PLS models to address this issue. Third, the effect of estimation errors on the modeling between consecutive stages can be explored, as the accuracy of model estimation play an important role in the diagnostic performance for local shifts. Fourth, it is likely that there are different extents of wear on consecutive teeth in a broach tool. Additional sensing information, such as high definition images, can be used to capture and quantify the wear conditions. The monitoring performance should be further optimized considering the different wear extents. Finally, an online process monitoring and diagnostic scheme based on the non-repeating cyclic signals should also be explored.

**References**


Axinte, D. and Nabil, G. (2003) Tool condition monitoring in broaching. *Wear*, **254**(3), 370-382.

Axinte, D. A., N. Gindy, K. Fox and I. Unanue (2004) Process monitoring to assist the workpiece surface quality in machining. *International Journal of Machine Tools and Manufacture*, **44**(10): 1091-1108.

Burrus, C. S., Gopinath, R. A., and Guo, H. (1997) *Introduction to Wavelets and Wavelet transforms: A Primer*, Prentice Hall, New Jersey, NJ.

Chun, H., and Keleş, S. (2010) Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **72**(1), 3-25.

Colosimo, B. M., Semeraro, Q., and Pacella, M. (2008) Statistical process control for geometric specifications: On the monitoring of roundness profiles. *Journal of Quality Technology*, **40**(1), 1-18.

Cramér, H. (1999) *Mathematical Methods of Statistics*, Princeton University Press, New Jersey, NJ.

Davis, P. J. (1975) *Interpolation and Approximation*, Dover Publications, New York, NY.

Ding, Y., Zeng, L., and Zhou, S. (2006) Phase I analysis for monitoring nonlinear profiles in manufacturing processes. *Journal of Quality Technology*, **38**(3), 199-216.

Geladi, P. and Kowalski, B. R. (1986) Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, **185**, 1-17.

Hawkins, D. M. (1991) Multivariate quality control based on regression-adjusted variables. *Technometrics*, **33**, 61-75.

Hawkins, D. M. (1993) Regression adjustment for variables in multivariate quality control. *Journal of Quality Technology*, **25**, 170-182.

Ho, J. C., Chou, S. K., Chua, K. J., Mujumdar, A. S., and Hawlader, M. N. A. (2002) Analytical study of cyclic temperature drying: effect on drying kinetics and product quality. *Journal of Food Engineering*, **51**(1), 65-75.

Höskuldsson, A. (1988) PLS regression methods. *Journal of Chemometrics*, **2**(3), 211-228.

Jensen, W. A. and Birch, J. B. (2009) Profile monitoring via nonlinear mixed models. *Journal of Quality Technology*, **41**(1), 18-34.

Jin, R. and Liu, K. (2013) Multimode variation modeling and process monitoring for serial-parallel multistage manufacturing processes. *IIE Transactions*, 45(6), 617-629.

Jin, J. and Shi, J. (1999a) Feature-preserving data compression of stamping tonnage information using wavelets. *Technometrics*, **41**, 327-339.

Jin, J. and Shi, J. (1999b). State space modeling of sheet metal assembly for dimensional control. *Journal of Manufacturing Science and Engineering*, **121**(4), 756-762.

Jin, J. and Shi, J. (2000) Diagnostic feature extraction from stamping tonnage signals based on design of experiments. *ASME Transactions, Journal of Manufacturing Science and Engineering*, **122**(2), 360-369.

Jin, J. and Shi, J. (2001) Automatic feature extraction of waveform signals for in-process diagnostic performance improvement. *Journal of Intelligent Manufacturing*, **12**, 257-268.

Kang, L. and Albin, S.L. (2000) On-line monitoring when the process yields a linear profile. *Journal of Quality Technology*, **32**, 418-426.

Kim, J., Huang, Q., Shi, J., and Chang, T. S. (2006) Online multichannel forging tonnage monitoring and fault pattern discrimination using principal curve. *ASME Transactions, Journal of Manufacturing Science and Engineering*, **128**(4), 944-950.

Kim, K., Mahmoud, M. A., and Woodall, W. H. (2003) On the monitoring of linear profiles. *Journal of Quality Technology*, **35**, 317-328.

Klocke, F., D. Veselovac, S. Gierlings and L. E. Tamayo (2012) Development of process monitoring strategies in broaching of nickel–based alloys. *Mechanics & Industry* **13**(01): 3-9.

Lowry, C. A., Woodall, W. H., Champ, C. W., and Rigdon, S. E. (1992) A multivariate exponentially weighted moving average control chart. *Technometrics*, **34**(1), 46-53.

Montgomery, D. C. (2012) *Introduction to Statistical Quality Control*, Seventh Edition, Wiley, New York, NY.

Nelson, L. S. (1986). Control chart for multiple stream processes. *Journal of Quality Technology*, **18**(4), 255-256.

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

Paynabar, K. and Jin, J. (2011) Characterization of non-linear profiles variations using mixed-effect models and wavelets. *IIE Transactions*, **43**(4), 275-290.

Paynabar, K., Jin, J., and Pacella M. (2013) Monitoring and diagnosis of multichannel nonlinear profile variations using uncorrelated multilinear principal component analysis. *IIE Transactions*, **45**, 1235-1247.

Paynabar, K., Qiu, P., and Zou, C. (2015) A Change Point Approach for Phase-I Analysis in Multivariate Profile Monitoring and Diagnosis. *Technometrics*, **58**(2), 191-204.

Qiu, P., Zou C., and Wang Z. (2010) Nonparametric profile monitoring by mixed effects modeling. *Technometrics*, **52**(3), 265-277.

Robertson, J., Rathinam, A., Wells, L., and Camelio, J. (2013) Statistical monitoring for broaching processes using energy features extracted from cutting force signatures. *Proceedings of NAMRI/SME*, **41**, 292-299.

Rosipal, R., and Krämer, N. (2006). Overview and recent advances in partial least squares. *In Subspace, Latent Structure and Feature Selection*, 34-51. Springer Berlin Heidelberg.

Shi, D., Axinte, D. A., and Gindy, N. N. (2007) Development of an online machining process monitoring system: a case study of the broaching process. *The International Journal of Advanced Manufacturing Technology*, **34**, 34-46.

Stoumbos, Z. G. and Sullivan, J. H. (2002) Robustness to non-normality of the multivariate EWMA control chart. *Journal of Quality Technology*, **34**(3), 260-276.

Sohn, H. Y. and Kim, B. S. (2002) A novel cyclic reaction system involving CaS and CaSO4 for converting sulfur dioxide to elemental sulfur without generating secondary pollutants. 1. Determination of process feasibility. *Industrial & Engineering Chemistry Research*, **41**(13), 3081-3086.

Sullivan, J. H. and Woodall, W. H. (1995) A comparison of multivariate control charts for individual observations. *Journal of Quality Technology*, **28**(4), 398-408.

Sutherland, J. W. and DeVor, R. E. (1986) An improved method for cutting force and surface error prediction in flexible end milling systems. *Journal of Engineering for Industry*, **108**(4), 269-279.

Trujillo-Ortiz, A., Hernandez-Walls, R., Barba-Rojo, K., and Cupul-Magana, L. (2007) HZmvntest: Henze-Zirkler's Multivariate Normality Test. A MATLAB file. URL http://www. mathworks. com/matlabcentral/fileexchange/loadFile. do.

Tsung, F., Li, Y., and Jin, M. (2008) Statistical process control for multistage manufacturing and service operations: a review and some extensions. *International Journal of Services Operations and Informatics*, **3**(2), 191-204.

Wang, H., Kababji, H., and Huang, Q. (2009) Monitoring global and local variations in multichannel functional data for manufacturing processes. *Journal of Manufacturing Systems*, **28**(1), 11-16.

Wold, H. (1966) Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, **1**, 391-420.

Wold, S., Sjostrom, M., and Eriksson, L. (2001) PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **58**, 109-130.

Woodall, W. H. (2006) The use of control charts in health-care and public-health surveillance. *Journal of Quality Technology*, **38**(2), 89-104.

Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004) Using control charts to monitor process and product quality profiles. *Journal of Quality Technology*, **36**(3), 309-320.

Wu, Z., Huang, N. E., Long, S. R., and Peng, C. K. (2007). On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proceedings of the National Academy of Sciences*, **104**(38), 14889-14894.

Wu, Z., Lam, Y.C., Zhang, S., and Shamsuzzaman, M. (2004) Optimization design of control chart systems. *IIE Transactions*, 36, 447-455.

Xiang, L. and Tsung, F. (2008) Statistical monitoring of multi-stage processes based on engineering models. *IIE Transactions*, **40**, 957-970.

Yang, Q. and Jin, J. (2012) Separation of individual operation signals from mixed sensor measurements. *IIE Transactions*, **44**(9), 780-792.

Zantek, P. F., Wright, G. P., and Plante, R. D. (2006) A self-starting procedure for monitoring process quality in multistage manufacturing systems. *IIE Transactions*, **38**(4), 293-308.

Zhang, G. (1984) A new type of control charts and a theory of diagnosis with control charts. *World Quality Congress Transactions*, 175-185.

Zhang, G. (1985) Cause-selecting control charts – a new type of quality control charts. *The QR Journal*, **12**, 221-225.

Zhang, G. (1992) Cause-selecting control chart and diagnosis, theory and practice. Aarhus School of Business. Department of Total Quality Management, Aarhus, Denmark.

Zhang, J., Ren, H., Yao, R., Zou, C., and Wang, Z. (2015) Phase I analysis of multivariate profiles based on regression adjustment. *Computers & Industrial Engineering*, **85**, 132-144.

Zhou, C., Liu, K., Zhang, X., Zhang, W., and Shi, J. (2015) An automatic process monitoring method using recurrence plot in progressive stamping processes. *IEEE Transactions on Automation Science and Engineering*, **99**, 1-10.

Zhou, S. and Jin, J. (2005) Automatic feature selection for unsupervised clustering of cycle-based signals in manufacturing processes. *IIE Transactions on Quality and Reliability*, **37**(6), 569-584.

Zhou, S., Jin, N., and Jin, J. (2005) Cycle-based signal monitoring using a directionally variant multivariate control chart system. *IIE Transactions on Quality and Reliability*, **37**(11), 971-982.

Zhou, S., Sun, B., and Shi, J. (2006) An SPC monitoring system for cycle-based waveform signals using Haar transform. *IEEE Transactions on Automation Science and Engineering*, **3**(1), 60-72.

Zou, C. and Tsung, F. (2008) Directional MEWMA schemes for multistage process monitoring and diagnosis. *Journal of Quality Technology*, **40**, 407-427.

Zou, C., Zhang, Y., and Wang, Z. (2006) A control chart based on a change-point model for monitoring linear profiles. *IIE Transactions*, **38**(12), 1093-1103.

Zou, C., Tsung, F., and Wang, Z. (2008) Monitoring profiles based on nonparametric regression methods. *Technometrics*, **50**(4), 512-526.

**Biographies**

Wenmeng Tian is currently a Ph.D. student in the Grado Department of Industrial and Systems Engineering at Virginia Tech. She received her bachelor's degree in Industrial Engineering and master's degree in Management Science and Engineering, both from Tianjin University, China. Her research interests focus on high density data modeling, monitoring, and prognostics in various manufacturing systems. She is a member of IIE.

Ran Jin is an Assistant Professor at the Grado Department of Industrial and Systems Engineering at Virginia Tech. He received his Ph.D. degree in Industrial Engineering from Georgia Tech, his master's degree in Industrial Engineering and in Statistics, both from the University of Michigan, and his bachelor's degree in Electronic Engineering from Tsinghua University. His research interests are in engineering driven data fusion for manufacturing system modeling and performance improvements, such as the integration of data mining methods and engineering domain knowledge for multistage system

modeling and variation reduction, and sensing, modeling, and optimization based on spatial correlated responses. He is a member of INFORMS, IIE, and ASME.

Tingting Huang is an Assistant Professor at the School of Reliability and Systems Engineering, Beihang University, Beijing, P. R. China. She got her Ph.D. degree from the School of Reliability and Systems Engineering, Beihang University in 2010. She worked as a postdoctoral researcher in the Department of Industrial Engineering, Tsinghua University in 2011. She got her master degree from the Department of Industrial and Systems Engineering, Virginia Tech in 2014. She was a visiting scholar in the Department of Industrial and Systems Engineering, Rutgers University, USA in 2008. Her research interests involve accelerated life testing, accelerated degradation testing and other reliability and environment testing technologies. Her recent work is on proportional hazards-odds model based accelerated degradation testing.

Jaime A. Camelio is currently the Rolls Royce Commonwealth Professor for Advanced Manufacturing in the Grado Department of Industrial and Systems Engineering at Virginia Tech. He obtained his B.S. and M.S. in Mechanical Engineering from the Catholic University of Chile in 1994 and 1995, respectively. In 2002, he received his Ph.D. from the University of Michigan. His professional experience includes working as a consultant in the Automotive/Operations Practice at A.T. Kearney Inc and as a Research Scientist in the Department of Mechanical Engineering at The University of Michigan, Ann Arbor. His research interests are in assembly systems, intelligent manufacturing, process monitoring and control, and cyber-physical security in manufacturing. He has authored or co-authored more than 70 technical papers and holds one patent.