Jihyun Kim
Research Institute for Information and Communication Technology,
Korea University,
Seoul, Korea

Qiang Huang[1]
Department of Industrial and Management Systems Engineering,
The University of South Florida,
Tampa, FL 33620
e-mail: huangq@eng.usf.edu

Jianjun Shi
Department of Industrial and Operations Engineering,
The University of Michigan,
Ann Arbor, MI 48109

Tzyy-Shuh Chang
OG Technologies, Inc.,
58 Parkland Plaza Suite 200,
Ann Arbor, MI 48103

Online Multichannel Forging Tonnage Monitoring and Fault Pattern Discrimination Using Principal Curve

Due to the late response to process condition changes, forging processes are normally exposed to a large number of defective products. To achieve online process monitoring, multichannel tonnage signals are often collected from the forging press. The tonnage signals contain significant amount of real time information regarding the product and the process conditions. In this paper, a methodology is developed to detect profile changes of multichannel tonnage signals for forging process monitoring and to classify fault patterns. The changes include global or local profile deviations, which correspond to deviations of a whole process cycle or process segment(s) within a cycle, respectively. The principal curve method is used to conduct feature extraction and discrimination of tonnage signals. The developed methodology is demonstrated with industry data from a crankshaft forging processes. [DOI: 10.1115/1.2193552]

1 Introduction

Forging is a complex process, typically performing several steps of plastic deformations with many process variables affecting the product quality. The current online monitoring practice in forging industry relies heavily on human visual inspection and mechanical gauges. Online sensor readings are used to make simple go/no go judgments rather than assess the accurate process conditions. Furthermore, it is prevalent in most cases to wait until the end of the forging process to perform an inspection and acquire the quality of the forged product. In addition, the inspection themselves are time consuming and infrequent. Such process monitoring and quality inspection practices lead inevitably to long time delays and often cause a significant increase of defective products. Therefore, there is a high demand to develop an online monitoring method that can be used to ascertain the final product quality effectively.

Among many process variables involved with the forging process, tonnage signals contain rich process information that is directly related with the forging force required to deform materials. They are nonstationary waveform signals. In general, deviations from the in-control process condition lead to changes in tonnage signal profiles. Furthermore, in the case of multiple tonnage sensors mounted onto a forging press, the location difference enables some channels to be more sensitive to certain process changes than the others, thereby making every tonnage signal vital in understanding the forging process condition. (A further discussion is given in Sec. 2.1 for the necessity of studying multichannel signals.) By developing a methodology to distinguish changes in the profiles of multichannel signals concurrently, tonnage signals can be used more effectively to monitor forging processes and also lead to a proper diagnosis of the process.

In the literature, most research works have been focusing on an analysis of either a tonnage signal from a single channel or the averaged/summed signal from multichannels. Usually features are extracted from tonnage signals for monitoring [1–5] or diagnosis [2–4,6]. For instance, a Shewhart control chart was developed to monitor peak tonnage value or other features extracted from a tonnage signal [1–3]. The principal Component Analysis (PCA) was used to extract key features from the tonnage signal for stamping process monitoring and diagnosis [4,7] and forging process monitoring [5]. Wavelets provided another efficient approach in the analysis of tonnage signal. The Haar transform was applied to the stamping tonnage signal to identify and monitor fault sensitive coefficients [3] and wavelet analysis was further exploited by incorporating the segmentation method [6] to determine wavelet coefficients sensitive to specific segments enabling more efficient data compression [8]. The previous approaches, however, are limited when dealing with multichannel tonnage signals. As mentioned before, sensor signals from different locations might respond to process changes differently. It is necessary to develop a new methodology by considering multichannel signals simultaneously.

The concept of a principal curve was first introduced by Hastie and Stuetzle in 1989 [9]. Principal curve is a smooth and one-dimensional curve that passes through the middle of a p-dimensional dataset in an orthogonal sense. It is a nonparametric generalization of a linear principal component. (Fig. 1 illustrates a principal curve obtained from a two-dimensional dataset with a parabolic relation.) The main advantages of using a principal curve are: (1) There is no distribution assumption about the data. Therefore, functional datasets, like tonnage signals, can be readily analyzed. (2) The dimension of the dataset is not a constraint for principal curve analysis. (3) The principal curve can be treated as one key feature or pattern extracted from p-dimensional data. Global or local changes of the curve indicate the change in the process.

Various research on principal curve applications have been reported, e.g., repositioning the misplaced magnets that keep the particle beams focused [9], identifying floes and their outlines in satellite images cases [10], and finding skeletons of hand-written character templates [11]. However, very limited work has been done to apply principal curve in process monitoring [12–14]. In those works, several nonlinear principal components (NLPCA), in addition to the principal curve or the first nonlinear principal com-

[1]Author to whom correspondence should be addressed.

Contributed by the Manufacturing Engineering Division of ASME for publication in the JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING. Manuscript received December 2, 2003; final manuscript received September 2, 2005. Review conducted by C. J. Li.

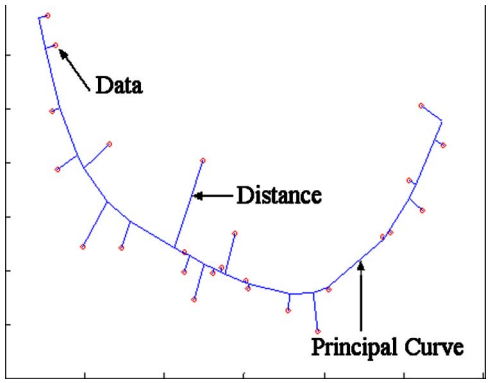


Fig. 1 Principal curve

ponent, were found by neural networks. The Squared Prediction Error (SPE) chart and associated score plots [15], which were developed for linear PCA, were borrowed for process monitoring [13,14]. Although the analogy between PCA and NLPCA is interesting, there are some open issues such as: (1) the structure of the neural network needs to be properly designed. It is time consuming to train a neural network. (2) Theoretical validity on the Chi-Square distribution assumption for SPE is not provided in the case of NLPCA.

Our purpose in this paper is to develop a methodology to detect profile changes of multichannel tonnage signals for forging process monitoring and to classify fault patterns. The changes include global or local profile deviations, which correspond to deviations of a whole process cycle or process segment(s) within a cycle. Real datasets of multichannel tonnage signals that are collected under both in-control and faulty process conditions are investigated. Treated as a one key feature or pattern extracted from in-control tonnage signals, a principal curve is used to benchmark and classify signals from faulty conditions.

The remaining paper is organized as follows: in Sec. 2 we introduce a background of forging processes and principal curves. The proposed methodology is presented in Sec. 3. Three industry datasets are used to demonstrate the methodology in Sec. 4. A summary and future works are summarized in Sec. 5.

2 Background of Forging and Principal Curve

2.1 Forging Process and Tonnage Signals. Forging process typically consists of several consecutive steps, such as upsetting, pre-forming, main forging, punching, trimming, and precision forging. Among these steps, the majority of plastic deformation occurs at the main forging press and, consequently, it is of a main concern. In general, the forging press is equipped with several strain gauges to monitor tonnage signals. A forging press with four strain gauges mounted is shown in Fig. 2.

Each of the tonnage sensors records one cycle of tonnage signal per every stroke. Every tonnage signal contains rich information regarding process conditions. An example of tonnage signals recorded from four strain gauges is shown in Fig. 3. In the figure, the x axis is the crank angle of the main shaft of the press and the y axis is the tonnage measured. As it can be seen, the four signals are different from each other in both the profile and the magnitude. This is due to the difference in the location of each sensor, resulting in different sensitivity to each phase of the forging process. Thus, restricting the scope of analysis to a specific signal or the average of the signals can generally result in losing process information. For instance, there are two types of die setup errors in forging processes, i.e., die parallelism error and die off-center on the bed of the press. Die parallelism error occurs when the upper and lower dies are not parallel with each other. When it occurs, the two dies will contact at one corner first. This will be reflected as an earlier increase of one tonnage signal than the others. Die off-center error occurs when a die is not positioned at the center of the bed of a press due to die setup error or material feeding problems. It will generate unbalanced forces among four columns. Under both cases, only by monitoring four channels of signals can we detect the process changes and conduct further investigation of root causes.

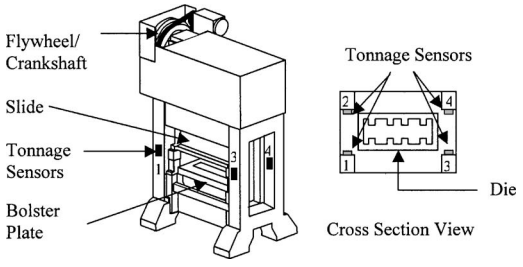


Fig. 2 Forging press and tonnage sensor locations

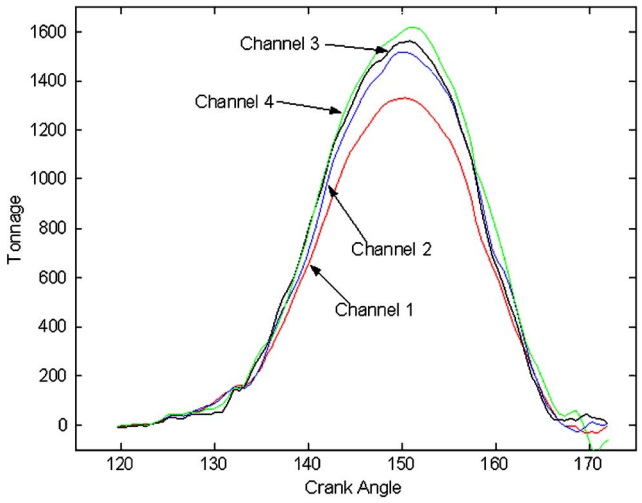


Fig. 3 Tonnage signals from four tonnage sensors

2.2 Introduction to Principal Curves. Principal curves are nonparametric, nonlinear generalizations of the first principal components. A brief introduction of principal curve is as follows.

Let $\mathbf{X}$ be a random vector in p-dimensional space $\mathbf{R}^p$ with density h and finite second moments. Let $\mathbf{x}$ denote the sample vector and assume that n data samples are available. Let $\mathbf{f}$ denote an infinitely differentiable unit-speed curve in $\mathbf{R}^p$ parametrized by $t \in \mathbf{R}^1$ [9,16], a closed interval, that does not intersect itself. The projection index $t_{\mathbf{f}}: \mathbf{R}^p - \mathbf{R}^1$ is defined as $t_{\mathbf{f}}(\mathbf{x}) = \sup_t\{t: \|\mathbf{x}-\mathbf{f}(t)\| = \inf_t\|\mathbf{x}-\mathbf{f}(t)\|\}$, where $t_{\mathbf{f}}(\mathbf{x})$ of $\mathbf{x}$ is the largest value of t for which $\mathbf{f}(t)$ is closest to $\mathbf{x}$. The curve $\mathbf{f}$ is defined as a principal curve of h if $E[\mathbf{X}|t_{\mathbf{f}}(\mathbf{X}) = t] = \mathbf{f}(t)$ for almost every t, i.e., each point on the principal curve is the average of all points projecting onto it. Note that $\mathbf{f}$ has no fixed parametric form and is defined by an ordered list of points in $\mathbf{R}^p$ using t. Next, let $d(\mathbf{x},\mathbf{f})$ denote the usual Euclidean distance from a point $\mathbf{x}$ to its projection on $\mathbf{f}$: $d(\mathbf{x},\mathbf{f}) = \|\mathbf{x}-\mathbf{f}[t_{\mathbf{f}}(\mathbf{x})]\|$; where $\|\cdot\|$ denotes the Euclidean norm in $\mathbf{R}^p$. The expectation of the sum of the squared error is defined as $D^2(h,\mathbf{f}) = E_h d^2(\mathbf{X},\mathbf{f})$.

The algorithm for constructing principal curves starts with some prior summaries, such as the usual principal-component line. The curve in successive iteration is a local average of the p-dimensional points, where the definition of local is based on the distance in arclengths of the projections of the points onto the curve found in the previous iteration. The algorithm to obtain the

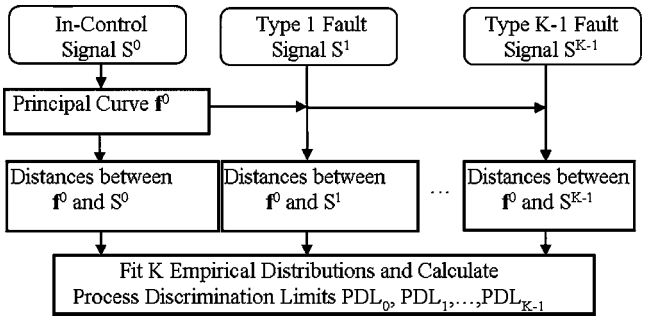


Fig. 4 Flow chart of the proposed methodology

principal curve is summarized as follows.

Step 0. Set $\mathbf{f}^{(0)}(t)=\bar{\mathbf{x}}+\mathbf{a}t$, where $\bar{\mathbf{x}}$ is the mean of the data set $\mathbf{x}$ and $\mathbf{a}$ is the first linear principal component of density function h. Set $t^{(0)}(\mathbf{x})=t_{\mathbf{f}^{(0)}}(\mathbf{x})$ and $j=1$.

Step 1. Define $\mathbf{f}^{(j)}(t)=E[\mathbf{X}|t_{\mathbf{f}^{(j-1)}}(\mathbf{X})=t]$.

Step 2. Set $t^{(j)}(\mathbf{x})=t_{\mathbf{f}^{(j)}}(\mathbf{x})$ for $\forall \mathbf{x}\in h$, where

$$t_{\mathbf{f}^{(j)}}(\mathbf{x})=\sup_t\{t:\|\mathbf{x}-\mathbf{f}^{(j)}(t)\|=\inf_t\|\mathbf{x}-\mathbf{f}^{(j)}(t)\|\}$$

Step 3. Evaluate $D^2(h,\mathbf{f}^{(j)})=E_{t^{(j)}}E\{\|\mathbf{X}-\mathbf{f}[t^{(j)}(\mathbf{X})]\|^2|t^{(j)}(\mathbf{X})\}$.

If $|D^2(h,\mathbf{f}^{(j-1)})-D^2(h,\mathbf{f}^{(j)})|/D^2(h,\mathbf{f}^{(j-1)})<\varepsilon$, then stop.

Otherwise, let $j=j+1$ and go to Step 1.

The algorithm for fitting the principal curve from datasets, rather than from a density function h, follows similar procedures [9].

3 Multichannel Tonnage Signals Monitoring and Fault Pattern Discrimination

Based on design specification and process knowledge, historical tonnage signals are assumed to be qualitatively classified into K types of data sets, i.e., $S^i(i=0,1,\ldots,K-1)$, where $i=0$ denotes in-control signals and $i=1,2,\ldots,K-1$ are signals with type i fault. The first research issue is how to classify a newly observed tonnage signal $\mathbf{X}_{\text{new}}$ into the given K types of datasets. In Sec. 3.1, a method to discover differences in overall tonnage profiles among the K datasets is proposed. Since different segments of tonnage signals correspond to different phases of plastic deformation, changes of overall tonnage profiles may be caused by specific deformation phases. In addition, certain deformation phase is more sensitive to small tonnage deviation than the others in terms of quality impact. Therefore, the second research issue, addressed in Sec. 3.2, is to discover at which stage(s) or process segment(s) the tonnage profiles start to differ from each other.

3.1 Overall Tonnage Profile Monitoring and Fault Pattern Discrimination. One of the fundamental assumptions when applying the proposed method is that if a process is under normal condition, the relationship among multichannel signals remains unchanged in a statistical sense. Any deviations from the forging process would result in a disruption to the relationship among the tonnage signals. The relationship is captured by the principal curve of tonnage signals.

Based on the assumption, the proposed methodology follows three steps (Fig. 4):

Step 1. Find principal curve $\mathbf{f}^0$ from in-control signal S^0

Step 2. Calculate the point-to-point distances between $\mathbf{f}^0$ and S^0, S^1, S^2, …, S^{K-1}.

Step 3. Fit K empirical distributions and find process discrimination limits PDL_0, $\text{PDL}_1,\ldots,\text{PDL}_{K-1}$.

Before discussing detailed procedures in each step, some notations are introduced first. Let $\mathbf{X}$ be a random vector representing tonnage sensor readings from p channels. Within each cycle of producing one part, n consecutive sensor readings are observed for each channel. Let $\boldsymbol{\Psi}$ be the $n\times p$ matrix recording all the sensor readings $\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\}$ within a cycle. One $\boldsymbol{\Psi}$ is treated as one *sample* of observation. Denote by $\boldsymbol{\Psi}_i^k$ the ith sample in S^K with $i=1,2,\ldots,N_{S^k}$ and $k=0,1,\ldots,K-1$, where N_{S^k} is the sample size of S^K. Define $\bar{\boldsymbol{\Psi}}^k$ as $\bar{\boldsymbol{\Psi}}^k=(1/N_{S^k})\Sigma_{i=1}^{N_{S^k}}\boldsymbol{\Psi}_i^k$, i.e., the sample mean of S^k $\{\bar{\mathbf{x}}_1^k,\bar{\mathbf{x}}_2^k,\ldots,\bar{\mathbf{x}}_n^k\}$, where $\bar{\mathbf{x}}_j^k=1/N_{S^k}\Sigma_{i=1}^{N_{S^k}}\mathbf{x}_{ji}^k$, $j=1,\ldots,n$. The principal curve $\mathbf{f}(t)$ is a vector of p functions of single variable t, i.e., $\mathbf{f}(t)=[f_1(t),f_2(t),\ldots,f_p(t)]^T$, where $f_1(t),f_2(t),\ldots,f_p(t)$ are coordinate functions in p-dimensional space. Under dataset $\boldsymbol{\Psi}$, $\mathbf{f}(t)$ is a polygonal curve of n vertices by connecting pairs of consecutive projection points $[\mathbf{f}(t_i),\mathbf{f}(t_{i+1})]$, $i=1,2,\ldots,n-1$. $\mathbf{f}(t)$ can also be described by an $n\times p$ matrix $\mathbf{F}$ as follows:

$$\mathbf{F}=\begin{bmatrix} f_1[t_f(\mathbf{x}_1)] & f_2[t_f(\mathbf{x}_1)] & \cdots & f_p[t_f(\mathbf{x}_1)] \\ f_1[t_f(\mathbf{x}_2)] & f_2[t_f(\mathbf{x}_2)] & \cdots & f_p[t_f(\mathbf{x}_2)] \\ \vdots & \vdots & \ddots & \vdots \\ f_1[t_f(\mathbf{x}_n)] & f_2[t_f(\mathbf{x}_n)] & \cdots & f_p[t_f(\mathbf{x}_n)] \end{bmatrix} \tag{1}$$

where $\{f_1[t_f(\mathbf{x}_i)],f_2[t_f(\mathbf{x}_i)],\ldots,f_p[t_f(\mathbf{x}_i)]$ or $\mathbf{f}[t_f(\mathbf{x}_i)]^T$ is the projection of point $\mathbf{x}_i$ onto curve $\mathbf{f}(t)$. The relationship between project indices $\{t_1,t_2,\ldots,t_n\}$ and $\{t_f(\mathbf{x}_1),t_f(\mathbf{x}_2),\ldots,t_f(\mathbf{x}_n)\}$ is that t_i is the ith smallest values in $\{t_f(\mathbf{X}_1),t_f(\mathbf{X}_2),\ldots,t_f(\mathbf{X}_n)\}$.

- *Step 1. Find principal curve $\mathbf{f}^0$ from in-control signal S^0*

  By implementing the algorithms in [9], the principal curve $\mathbf{f}^0$ is extracted from $\bar{\boldsymbol{\Psi}}^0$. The corresponding $\mathbf{F}$ matrix, defined by (1), is denoted by $\mathbf{F}^0$.

  Here is a remark on using $\bar{\boldsymbol{\Psi}}^0$ to find principal curve $\mathbf{f}^0$.

  R1 (**Remark 1**): Another way to extract $\mathbf{f}^0$ from S^0 is to find principal curves for each $\boldsymbol{\Psi}_i^0$ in $S^0(i=1,2,\ldots,N_{S^0})$. Then treat the average of N_{S^0} principal curves as $\mathbf{f}^0$. The two curves converge to the true principal curve with probability 1 when N_{S^0} goes to infinity. A short proof of R1 is given in the Appendix.
- *Step 2. Calculate the point-to-point distances between $\mathbf{f}^0$ and S^0, S^1, S^2, …, S^{K-1}*

  The distances are calculated between $\mathbf{f}^0$ and each $\boldsymbol{\Psi}_i^k$ in $S^k(k=0,1,\ldots,K-1;i=1,2,\ldots,N_{S^k})$. For the data point $\mathbf{x}_i^k$ in $\boldsymbol{\Psi}_i^k$, its distance to $\mathbf{f}^0$ is

$$d(\mathbf{x}_i^k,\mathbf{f}^0)=\|\mathbf{x}_i^k-\mathbf{f}^0[t_f(\mathbf{x}_i^k)]\| \tag{2}$$

  where $\mathbf{f}^0[t_f(\mathbf{x}_i^k)]^T$ is the ith row in $\mathbf{F}^0$. The overall squared distance between $\boldsymbol{\Psi}_i^k$ and $\mathbf{f}^0$ is

$$D^2(\boldsymbol{\Psi}_i^k,\mathbf{f}^0)=\sum_{i=1}^{N_{S^k}}d^2(\mathbf{x}_i^k,\mathbf{f}^0) \tag{3}$$

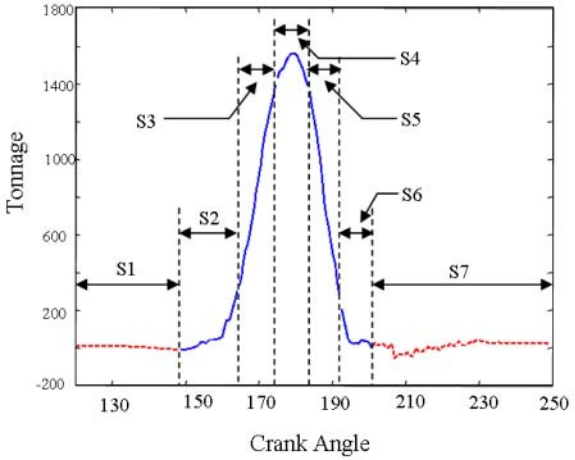
  If there are profile changes in $\boldsymbol{\Psi}_i^k(k=1,\ldots,K-1)$, the deviations from $\mathbf{f}^0$ can be measured by $D^2(\boldsymbol{\Psi}_i^k,\mathbf{f}^0)$. This is the rationale used to develop discrimination limits in step 3.
- *Step 3. Fit empirical distributions for each process condition S^K and find corresponding process discrimination limit* $\text{PDL}_k(k=0,1,\ldots,K-1)$

For each data set $S^k(k=0,\ldots,K-1)$, empirical cumulative distribution function CDF_k is fitted from $D^2(\boldsymbol{\Psi}_i^k,\mathbf{f}^0)$'s, $i=1,2,\ldots,N_{S^k}$. Discrimination limit PDL_k is defined as

$$\text{PDL}_k=\text{CDF}_k^{-1}(1-\alpha) \tag{4}$$

i.e., PDL_k is the $(1-\alpha)$ percentile of CDF_k, where α is called the type I error in Statistical Process Control.

Rearrange PDL_0, $\text{PDL}_1,\ldots,\text{PDL}_{K-1}$ as an increasing order as



	From	To
Seg.1	120°	148°
Seg.2	149°	164°
Seg.3	165°	174°
Seg.4	174°	183°
Seg.5	184°	192°
Seg.6	193°	201°
Seg.7	202°	250°

Fig. 5 Segmentation of the tonnage signal and the segment intervals

$\text{PDL}_{(0)}, \text{PDL}_{(1)}, \ldots, \text{PDL}_{(K-1)}$. The corresponding fault types are called Fault Type (0), Fault Type (1),…, Fault Type $(K-1)$. Since $\text{PDL}_0 = \text{PDL}_{(0)}$ by the fact that $\mathbf{f}^0$ is fitted from in control datasets, Fault Type (0) represents that the signal is in control. For observation $\boldsymbol{\Psi}_{\text{new}}$, plot $D^2(\boldsymbol{\Psi}_{\text{new}}, \mathbf{f}^0)$ on a control chart, and judgment is made based on the following rule:

$$\boldsymbol{\Psi}_{\text{new}} \text{ is } \begin{cases} \text{in-control}, & \text{if } D^2(\boldsymbol{\Psi}_{\text{new}}, \mathbf{f}^0) \leq \text{PCL}_0; \\ \text{Type (1) fault}, & \text{if } \text{PCL}_0 \leq D^2(\boldsymbol{\Psi}_{\text{new}}, \mathbf{f}^0) \leq \text{PCL}_{(1)}; \\ \vdots & \\ \text{Type } (k) \text{ fault}, & \text{if } \text{PCL}_{(k-1)} \leq D^2(\boldsymbol{\Psi}_{\text{new}}, \mathbf{f}^0) \leq \text{PCL}_{(k)}; \\ \vdots & \\ \text{Type } (K-1) \text{ fault}, & \text{if } \text{PCL}_{(K-2)} \leq D^2(\boldsymbol{\Psi}_{\text{new}}, \mathbf{f}^0) \leq \text{PCL}_{(K-1)}; \\ \text{Unkown fault}, & \text{if } D^2(\boldsymbol{\Psi}_{\text{new}}, \mathbf{f}^0) \geq \text{PCL}_{(K-1)}. \end{cases} \tag{5}$$

Comments on Step 3 are given as follows:

R2 The decision rule in (5) works efficiently only if (a) the interval between $\text{PDL}_{(k)}$ and $\text{PDL}_{(k-1)}$ is large enough to accommodate the corresponding Fault Type (k) dataset $(k=0,1,\ldots,K-1, \text{PDL}_{-1}=0)$ and (b) the interval only contains one type of fault data, namely Fault Type (k). To overcome the constraints and increase the separability of the process conditions for the procedure, a reduction in the scope of the dataset is suggested. Instead of using the overall tonnage signal, the signal is divided into smaller segments. Applying the proposed methodology on each of the segments can result in identifying segment(s) that have larger separability than the case using a global signal.

3.2 Localization of Faulty Process Segments Within a Forging Cycle. There are two physical facts supporting the study of the individual segments of tonnage signals for the purpose of increasing separability of the process conditions: (1) Different segments of tonnage signals correspond to different phases of plastic deformation. The process dataset detected to deviate from the normal condition using overall profile monitoring may have deformities in only certain deformation phases. Reducing the scope to each of the localized segments can aid in identifying the cause of abnormal process condition. (2) In terms of quality impact, a certain deformation phase is more sensitive to small tonnage deviation than the others. Focusing on a specific segment of the tonnage signal can increase the sensitivity of detecting relatively small but important deviation. Our objective in this subsection is to discover at which stage(s) or process segment(s) the K tonnage datasets differ from each other. The proposed approach is as follows:

(1) Tonnage Signal Segmentation: The tonnage signal is partitioned into several segments based on the understanding of the plastic deformation phases. Thus, the signal $\boldsymbol{\Psi}$ or $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is divided into γ segments.
(2) Profile Monitoring and Discrimination for Each Segment: Apply the three-step procedure presented in Sec. 3.1 to each segment. This results in obtaining γ sets of discrimination limits specific to each segment.

The set of discrimination limits obtained can be used to identify the abnormal forging phase(s). The exact location(s) of the deformity combined with the engineering knowledge can lead to a more detailed fault discrimination of the process.

For this paper, the segmentation of tonnage signals proposed in Koh et al. [6] is used. The forging tonnage signal is divided into seven segments with respect to the plastic deformation of the process. The segmented tonnage signal along with the break points of the segmentation is shown in Fig. 5. Among the seven segments, active plastic deformation occurs between segments 2 and 6. Therefore, concentrating on these five segments will suffice in demonstrating the effectiveness of the proposed methodology.

4 Case Studies With Real Tonnage Signals

The tonnage signals used in the paper are collected from the main press of a crankshaft forging plant. Four strain gauges are installed on each of the four columns of the forging press to obtain the tonnage signals. The tonnage signals along with the appropriate crank angles (between 120 deg and 250 deg) are obtained with the frequency of 224 different angles per stroke.

From the main press, tonnage signals that correspond to three types of process conditions are obtained. They are a nominal dataset and two types of faulty process conditions that correspond to deformity in counterweight (Abnormal Process 1) and incorrect die setup (Abnormal Process 2). The deformity in counterweight was caused by die wear, especially in the counterweight region, and resulted in early die change. The process condition is ascertained based on the plant specification, quality measurements, and expert knowledge. Two cases of tonnage signal analysis will be shown in the paper.

As stated previously, the most common method used for tonnage monitoring is using peak values or maximum tonnage as the key feature. Figure 6 shows the maximum values for the three process conditions. Although Abnormal Process 2 condition is dis-

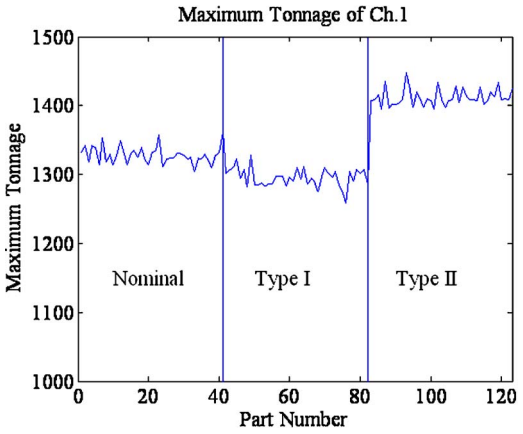


Fig. 6 Tonnage signal monitoring based on peak values

tinguishable from the other two process conditions, the distinction between the normal condition and Abnormal Process 1 is not apparent.

4.1 Case 1 Overall Forging Profile Monitoring and Fault Discrimination. The methodology is applied to segments 2 through 6 for the three types of process conditions. There are 90 consecutive sensor readings in between. Thus, $\mathbf{\Psi}^k$ is a 90×4 matrix with $K=3$. Given 51 observations for each process conditions, the first 41 observations for each process type ($N_{S^k}=41$ for all S^k, $k=0,1,2$) are used as a training set to construct respective discrimination limits and the next 10 observations are used for validation purposes. First, the sample mean of S^0, i.e., $\mathbf{\Psi}^0$, is obtained by $(1/41)\Sigma_{i=1}^{41}\mathbf{\Psi}_i^0$. Using $\mathbf{\Psi}^0$, the principal curve $\mathbf{f}^0$ is extracted and $\mathbf{F}^0$ is obtained (a portion of $\mathbf{F}^0$ is shown in Table 1).

Next, point to point distance between $\mathbf{f}^0$ and each $\mathbf{\Psi}_i^k$ in S^0, S^1, S^2 is obtained. Table 2 lists $D^2(\mathbf{\Psi}_i^k,\mathbf{f}^0)$ for $i=1$ to 10.

Three empirical distributions are fitted to the overall squared distances. The CDF plots are depicted in Fig. 7. From $D^2(\mathbf{\Psi}_i^k,\mathbf{f}^0)$, three process discrimination limits $\text{PDL}_{(0)}$, $\text{PDL}_{(1)}$, $\text{PDL}_{(2)}$ are obtained.

The discrimination limits in Table 3 are obtained from the respective distributions using $\alpha=5\%$. $D^2(\mathbf{\Psi}_i^k,\mathbf{f}^0)$ and the discrimination limits are depicted in Fig. 8. Using the obtained limits, the process conditions of new process data can be ascertained.

Next, $\mathbf{X}_{\text{new}}$ with ten observations is used for validation purposes. The point to point distances between the new dataset and $\mathbf{f}^0$ are obtained. The overall squared distance is obtained and the data is plotted on the chart (Fig. 8). It can be seen that the proposed methodology provides a valuable means of process monitoring

Table 1 Principal curve F^0 (the first ten projections)

	p			
n	1	2	3	4
1	−8.933	−8.461	−4.527	−15.745
2	−6.702	−6.218	−2.523	−13.048
3	−4.339	−3.842	−0.400	−10.193
4	−2.420	−1.911	1.325	−7.877
5	−0.328	0.194	3.206	−5.352
6	−1.181	−0.664	2.440	−6.381
7	2.598	3.140	5.841	−1.823
8	8.602	9.189	11.254	5.406
9	16.569	17.231	18.460	14.979
10	29.199	30.018	29.955	30.103

Table 2 Squared distance $D^2(\Psi_i^k,\mathbf{f}^0)$ (first ten observations)

	k		
i	1	2	3
1	1.490e+05	4.916e+05	1.457e+06
2	1.854e+05	3.906e+05	1.443e+06
3	1.515e+05	3.681e+05	1.755e+06
4	2.050e+05	3.324e+05	1.030e+06
5	1.748e+05	5.460e+05	2.542e+06
6	1.608e+05	5.125e+05	1.247e+06
7	2.691e+05	8.663e+05	1.468e+06
8	1.609e+05	3.503e+05	1.462e+06
9	1.453e+05	7.943e+05	1.341e+06
10	1.660e+05	8.016e+05	1.417e+06

and faulty discrimination.

R3 The proposed methodology does not impose any distribution assumption on the dataset. Therefore, any type of dataset can be used for the analysis, making the methodology robust. However, due to the use of empirical distribution when constructing $(1-\alpha)\%$ discrimination limits, the discrimination chart is inevitably exposed to $\alpha\%$ outliers.

4.2 Case 2 Localization of Faulty Process Segments Within a Forging Cycle. The tonnage dataset is partitioned into respective segments. For each of the five segments, segment 2 to 6, the proposed three-step procedure is applied. Empirical distributions are fitted to the squared distances. CDF is employed to obtain respective discrimination limits. These are listed in Table 4.

The discrimination limits for each of the segments are depicted in Fig. 9. The new validation dataset is plotted in the same figure.

R4 The plots can be divided into three groups. In the first group (segment 2), Abnormal Process 1 is more distinguishable from the others. This is due to the nature of the process fault for major deviation occurs in segment 2. In the second group (segments 3 and 4), Abnormal Process 2 shows a clear divergence from the others, showing its big contribution in process deviation. There exists no apparent distinction between Abnormal Processes 1 and 2 in group 3 (segments 5 and 6). This indicates that although a

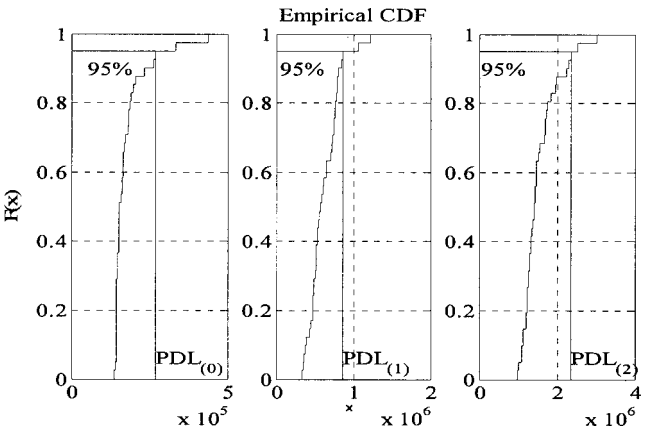


Fig. 7 CDF plots of overall squared distances

Table 3 Discrimination limits for overall signals

Process type	Discrimination limit
$\text{PDL}_{(0)}$	2.691E+05
$\text{PDL}_{(1)}$	8.663E+05
$\text{PDL}_{(2)}$	2.349E+06

Table 4 Segmentation limits for Segments 2 through 6

	Discrimination limits				
Process type	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6
Nominal	6228.106	43197.364	61688.457	50870.812	10817.298
Fault type (1)	57254.453	95255.969	296197.861	323587.877	49888.582
Fault type (2)	30327.458	599245.539	949595.989	676685.646	77734.651

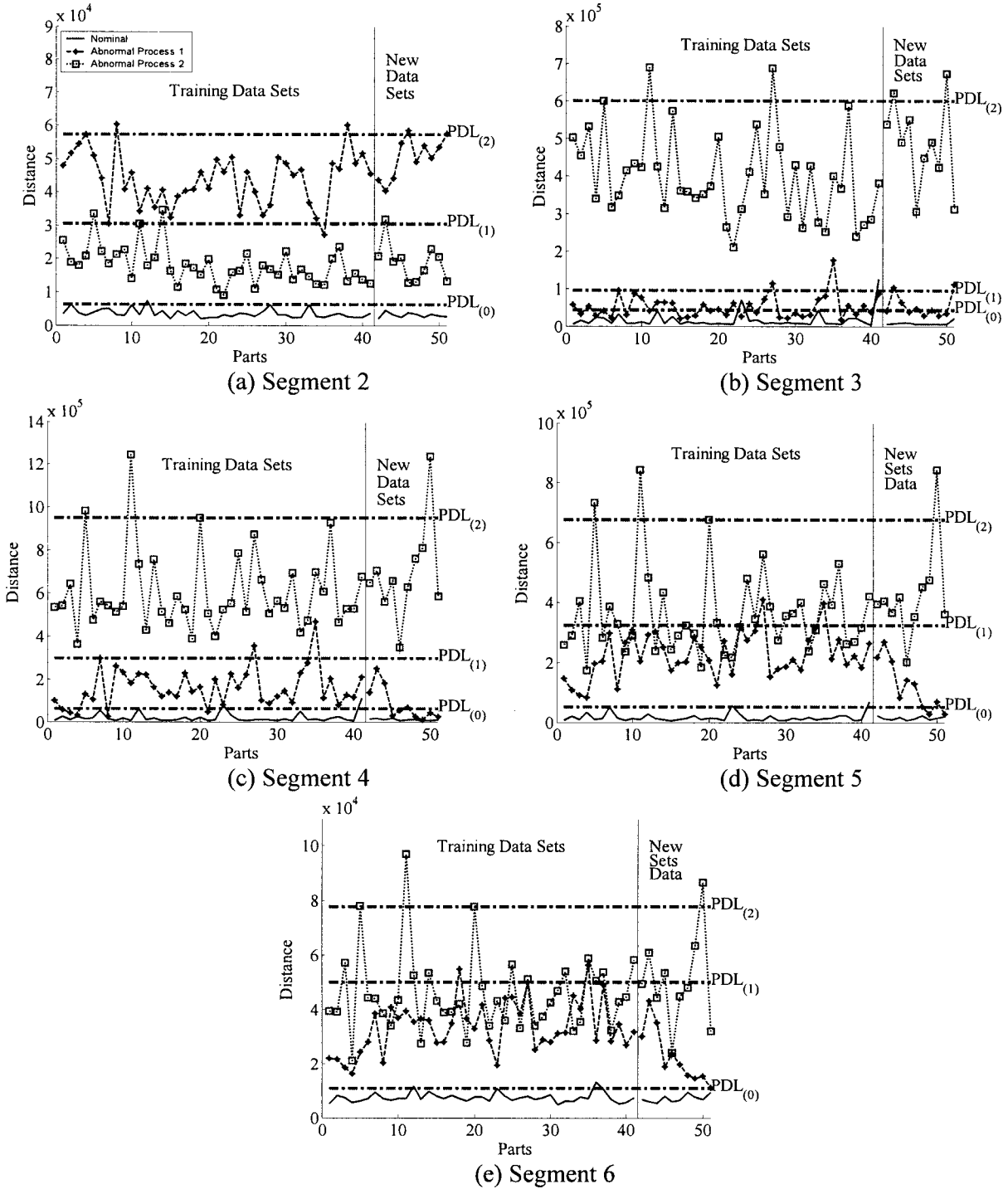


Fig. 9 Discrimination limits for segment 2 to segment 6 (*a*) segment 2; (*b*) Segment 3; (*c*) Segment 4; (*d*) Segment 5; (*e*) Segment 6

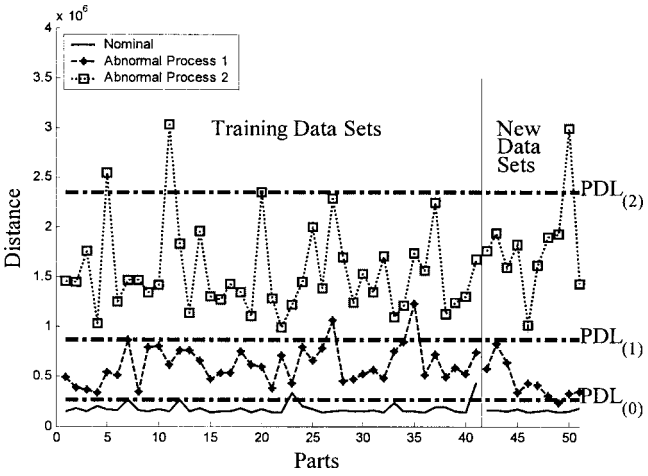


Fig. 8 Discrimination limits for case 1

deviation from the normal condition occurred, the results obtained from these two segments are not sufficient to distinguish the two faulty conditions. These results not only confirm that more sensitivity can be obtained when performing a localized analysis of the tonnage signal, but also provide information regarding the actual process condition and where distinction lies for each fault type conditions.

R5 In many cases, only a certain phase of the plastic deformation is of great concern to the process engineers. By identifying the segment(s) corresponding to the phase of interest and performing the three-step procedure on them can provide segment specific information as well as reduce the dimension of the dataset.

5 Conclusion

This paper proposed a methodology to detect profile changes of multichannel tonnage signals for forging process monitoring and to classify fault patterns. The principal curve is treated as one key feature or pattern extracted from in-control tonnage signals. This pattern is further being used to benchmark and classify signals from faulty conditions.

The case studies show that the comprehensive analysis of multichannel signals ensures the capture of the different aspects of the forging process for profile changes in both the whole process cycle as well as specific process segment(s) within a cycle. The segmentation study aimed to take a focused look on local profiles. Although segmentation of the signal is problem dependent, it is generally meaningful to conduct a segmentation study because process faults could occur at different stages within a cycle. Furthermore, the number of segments and their associated characteristics can be common in forming applications.

The application of principal curve in online tonnage signal monitoring is shown to be both efficient and innovative. The proposed methodology ensures robustness through the use of an empirical distribution. It broadens the horizon of the application of the principal curve method to include a manufacturing field.

Focusing on online process monitoring and fault discrimination for forging processes, in this paper we do not investigate the correlation between tonnage signals and process physics. Establishing such a relationship, which is an ongoing research topic, is important for root cause identification. Future research efforts can be devoted to that area.

Acknowledgment

This work is partially supported by the NIST-ATP project 70NANBOH3014 (the Smartsmith project) and NSF DMI-0600066 grant.

Appendix

The principal curve obtained from the averaged data and the average of principal curves obtained from each data converge to the true principal curve with probability 1 when N_{S^0} goes to infinity.

Proof: Assume the true projection matrix is $\mathbf{A}_j$, which projects data $\mathbf{x}_j$ onto true principal curve $\mathbf{f}$, i.e.,

$$\mathbf{A}_j\mathbf{x}_j = \mathbf{f}[t_f(\mathbf{x}_j)] + \boldsymbol{\varepsilon}, \quad j = 1, \ldots, n \tag{A1}$$

where $\boldsymbol{\varepsilon}$ is noise term.

When fitting principal curves for $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_N$, there exist perturbations in $\mathbf{A}_j$, i.e.,

$$(\mathbf{A}_j + \Delta\mathbf{A}_{ji})\mathbf{x}_{ji} = \mathbf{f}[t_f(\mathbf{x}_{ji})] + \boldsymbol{\varepsilon}, \quad j = 1, \ldots, n; \quad i = 1, \ldots, N \tag{A2}$$

Since $\mathbf{A}_j\mathbf{x}_{ji} = \mathbf{f}[t_f(\mathbf{x}_j)]$, (2) can be rewritten as

$$\mathbf{f}[t_f(\mathbf{x}_j)] = \mathbf{f}[t_f(\mathbf{x}_{ji})] + \Delta\mathbf{A}_{ji}\mathbf{x}_{ji} + \boldsymbol{\varepsilon} \tag{A3}$$

The average of N projected points is close to $\mathbf{f}[t_f(\mathbf{x}_j)]$ when N is large.

$$\mathbf{f}[t_f(\mathbf{x}_j)] = \frac{1}{N}\sum_{i=1}^{N} f[t_f(\mathbf{x}_{ji})] + \frac{1}{N}\sum_{i=1}^{N} \Delta\mathbf{A}_{ji}\mathbf{x}_{ji} + \boldsymbol{\varepsilon} \tag{A4}$$

When fitting the principal curve from $\boldsymbol{\Psi}$,

$$[\mathbf{A} + \Delta\mathbf{A}_j]\bar{\mathbf{x}}_j = \mathbf{f}[t_f(\mathbf{x}_j)] + \Delta\mathbf{A}_j\bar{\mathbf{x}}_j + \boldsymbol{\varepsilon} \tag{A5}$$

which is also close to $\mathbf{f}[t_f(\mathbf{x}_j)]$ with large N is obtained.

The result is also true if there are small deviations in data $\mathbf{x}_{ji}$'s.

If the first principal component line of $\boldsymbol{\Psi}$ is a principal curve $\mathbf{f}$, $\mathbf{A}_j$ is the same for all j.

References

[1] Seem, J. E., and Knussmann, K. D., 1994, "Statistical Methods for On-Line Fault Detection in Press-Working Applications," Signature Technology Technical Report.
[2] Robbins, T., 1995, "Signature-Based Process Control & SPC Trending Evaluate Press Performance," Metal Forming, **50**, pp. 44–50.
[3] Koh, C. K. H., Shi, J., and Williams, W., 1995, "Tonnage Signature Analysis Using the Orthogonal (Harr) Transforms," NAMRI/SME Transactions, **23**, pp. 229–234.
[4] Jin, J., and Shi, J., 2000, "Automatic Feature Extraction for In-Process Diagnostic Performance Improvement," J. Intell. Manuf., **12**, pp. 267–268.
[5] Kim, J., Zhou, S., Shi, J., and Chang, T., 2000, "Forging Process Monitoring Through Multivariate Analysis of Tonnage Signals," *24th Forging Industry Technical Conference*, pp. 167–188.
[6] Koh, C. K.-H., Shi, J., Black, J. M., and Ni, J., 1996, "Tonnage Signature Analysis for Stamping Process," Trans. NAMRC/SME, **24**, pp. 193–198.
[7] Jin, J., and Shi, J., 2002, "Diagnostic Feature Extraction From Stamping Tonnage Signals Based on Design of Experiment," ASME J. Manuf. Sci. Eng., **122**, pp. 360–369.
[8] Jin, J., and Shi, J., 1999, "Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets," Technometrics, **41**, pp. 327–339.
[9] Hastie, T., and Stuetzle, W., 1989, "Principal Curves," J. Am. Stat. Assoc., **84**, pp. 502–516.
[10] Banfield, J. D., and Raftery, A. E., 1992, "Ice Floe Identification in Satellite Images Using Mathematical Morphology and Clustering About Principal Curves," J. Am. Stat. Assoc., **87**, pp. 7–16.
[11] Kégl, B., and Krzyzak, A., 2002, "Piecewise Linear Skeletonization Using Principal Curves," IEEE Trans. Pattern Anal. Mach. Intell., **24**, pp. 59–74.
[12] Kramer, M. A., 1991, "Nonlinear Principal Component Analysis Using Autoassociative Neural Networks," AIChE J., **37**, pp. 233–243.
[13] Dong, D., and McAvoy, T. J., 1995, "Nonlinear Principal Component Analysis—Based on Principal Curves and Neural Networks," Comput. Chem. Eng., **20**, pp. 65–78.
[14] Wilson, D. J. H., Irwin, G. W., and Lightbody, G., 1999, "RBF Principal Manifolds for Process Monitoring," IEEE Trans. Neural Netw., **10**, pp. 1424–1434.
[15] Nomikos, P., and MacGregor, J., 1995, "Multivariate SPC Charts for Monitoring Batch Processes," Technometrics, **37**, pp. 41–59.
[16] do Carmo Manfredo, P., 1976, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood, Cliffs, NJ.