University of South Florida

Digital Commons @ University of South Florida

USF Tampa Graduate Theses and Dissertations

USF Graduate Theses and Dissertations

2005

Multichannel Functional Data Decomposition and Monitoring

Hani Kababji
University of South Florida

Follow this and additional works at: https://digitalcommons.usf.edu/etd

Part of the American Studies Commons

Multichannel Functional Data Decomposition and Monitoring

by

Hani Kababji

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Industrial Engineering
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Major Professor: Qiang Huang, Ph.D.
Jose L. Zayas-Castro, Ph.D.
Tapas Das, Ph.D.

Date of Approval:
November 10, 2005

Keywords: Variation decomposition, Change detection, Faulty condition,
Discrimination, False alarm

Table of Contents

List of Figures

Multichannel Functional Data Decomposition and Monitoring

Hani Kababji

Hani Kababji

Abstract

With current advances in sensors and information technology, online measurements of process variables become increasingly accessible for process control and monitoring. Such measurements may take the shape of curves rather than scalar values. The term Multichannel Functional Data (MFD) is used to represent the observations of multiple process variables in the shape of curves. Generally MFD contains rich information about processes. The challenge of process control in MFD is that Statistical Process Control (SPC) is not directly applicable. Furthermore, there is no systematic approach to interpret the complex variation in MFD. In this research, our objective is to develop an approach to systematically analyze the complex variation in MFD for process change detection and process faulty condition discrimination.

The main contributions of this thesis are: MFD decomposition, process change detection, and process faulty condition discrimination. We decomposed MFD into global and local components. The approach reveals global and local variations that are due to global signal shifts and local variations. Global variation was extracted using weighted spline smoothing technique, whereas, local variation was obtained by subtracting the

global variation from original signals. Weights were obtained using the local moving average of the generalized residuals.

The proposed approach helps in process change detection and process faulty condition discrimination based on further MFD analysis using Principal Curve Regression (PCuR) Test. For process change detection, global variation component was used in the PCuR test. In-control global data sets were used as training data to detect process change that is due to global and local variation. On the other hand, for faulty condition discrimination purpose, local variation component was used in the PCuR test. In-control local variation data sets were used as training data in the PCuR test; therefore, process faulty condition that is due to local variations remains in control, whereas, process faulty condition that is due to global shifts appears as random out of control points in the PCuR test.

We applied our approach on real life forging data sets. A simulation study was also conducted to verify the approach and results are promising for wide applications.

Chapter 1. Introduction

This chapter presents a brief introduction and motivation of this research. The research objectives and an overview of thesis work are also introduced.

1.1 Multichannel Functional Data (MFD) and Statistical Process Control

This section presents a brief description about MFD and its corresponding challenges in statistical process control. An example is also provided to illustrate complexity in analyzing MFD.

With current advances in sensors and information technology, online measurements for process variables become increasingly usable for industrial process monitoring and control. Process characteristics may be in the shape of multiple curves since process performance might be affected by multiple functional process variables. For instance, many industrial processes are time dependent; therefore, such process variables, which are monitored among time, take the shape of functional data.

Generally, functional process variables contain rich information about monitored processes. Process characteristics track and describe the behavior of a process response with respect to a desired predictor, such as: process time, change in distance, etc....

Therefore, multiple sensors are commonly installed to collect data of functional process variables. The term Multichannel Functional Data (MFD) is used to represent observations of multiple function process variables.

For example, forging process is equipped with sensors to measure tonnage signal as process variable. Since process faults may occur at different locations in the workpiece, multiple sensors are installed at different workpiece locations. Consequently, multiple channels are extracted from process sensors measuring tonnage signals as a function of time for every sample. See Figure (1.1, 1.2) [1]. In this example, four channels were extracted measuring the same process characteristic i.e. tonnage signal. However, in every channel the corresponding tonnage signal is considered as independent variable since tonnage signals are extracted from different sensors. Moreover, the tonnage values vary among different locations in the workpiece due to process physics. On the other hand, other processes and applications may contain extracted channels with completely different process characteristic.
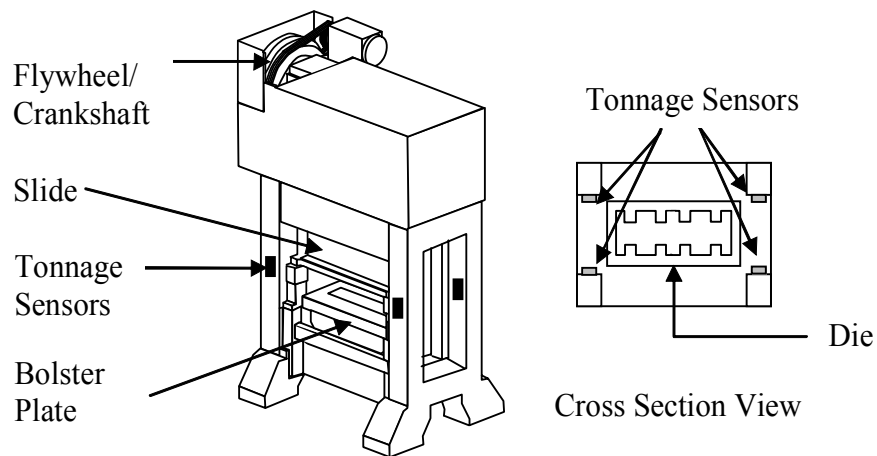
Figure 1.1 Forging Press, Die, and Installed Sensors

2

The importance of monitoring MFD in such processes lies behind the fact that MFD monitoring provides an effective way in controlling the process since it monitors different process variables or multiple realizations of a process variable rather than just monitoring the average or the sum of sensed process variables. Therefore, more valuable information about a process can be obtained.

The process cycle, which is the time interval needed for a part to be completely manufactured, in such processes is repeatable among samples on each part under the assumption that the samples of a particular channel are identical and independently distributed. Although this assumption eases controlling MFD, many challenges still exist while monitoring sensed MFD.
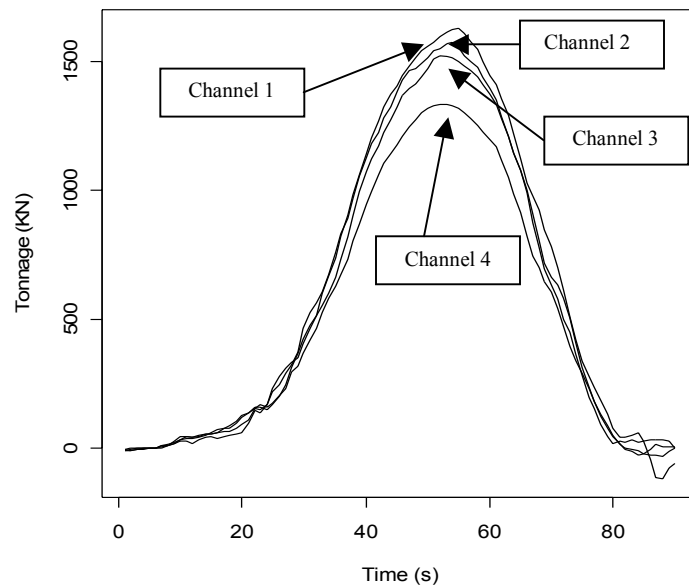
Figure 1.2 MFD of One Part (Sample) in Forging Process

3

MFD is complex by its nature because of the existence of more than one variable in the shape of functional data which makes it hard to interpret. Besides, no systematic approach exists for analyzing and interpreting MFD. Statistical Process Control (SPC) is not directly applicable to this type of process characteristics because of the complexity in the variation of MFD [2, 3]. For example, the mean and the variance are not sufficient to represent process variations unlike vector process characteristics where it is easier to extract variations using process parameters.

A further issue in analyzing MFD is late process change detection. Since some manufacturing processes contain high throughput time, late process change detection may result in large number of defective parts. Therefore, an online scheme for such processes monitoring is of crucial need. The online detection allows corrective actions to be taken before parts are completely manufactured. This lead to better process performance with less quality cost.

Process faulty condition discrimination is a concurrent challenge with process change detection. For instance, in forging process different types of process faulty conditions may occur due to different causes, such as: deformity in counterweight, and incorrect die setup. The discrimination among different process conditions needs deep understanding of process physics and further analysis in MFD to pinpoint a key feature that helps in discriminating different process conditions.

1.2 Objective and Contribution

The objective of this research is to systematically analyze MFD for process change detection and process faulty condition discrimination. In this study, we propose a new scheme for interpreting and monitoring MFD. The approach provides a new way of interpreting variations in MFD. We decompose original sensed MFD into global and local components to capture global and local variations in the sensed data based on previous process knowledge. Global patterns can reveal process global and local shifts. However, rapid local variations and the time at which they occur can be noticed by analyzing the local variation profile taking into account the correlation among time.

Global variation is extracted using weighted spline smoothing applied on the original data, whereas, local variation is obtained by subtracting global pattern from the original data. Consequently, two data sets can be obtained providing rich information about the variation patterns in the original signals.

For process monitoring purposes, the data sets obtained is further used as an input for Principal Curve Regression (PCuR) test [2]. Using global component, process faulty conditions can be detected. On the other hand, process faulty condition discrimination can be achieved by using local variation as input in PCuR test.

1.3 Overview of the Research

The thesis is organized as follows: Chapter 2 presents a review of the literature related to multivariate process control, sensor fusion, spline smoothing technique, and approaches in MFD processing and analysis. Chapter 3 summarizes the problem being analyzed in this research. Chapter 4 focuses on global and local variations extraction. Chapter 5 presents the grouping technique conducted to choose training data sets and PCuR model application. In chapter 6 a simulation study is presented to verify the application of the methodology proposed. Finally, chapter 7 consists of the summary, the conclusion of this research, and some suggestions for future research. The figure below illustrates the flow chart of the thesis contents.
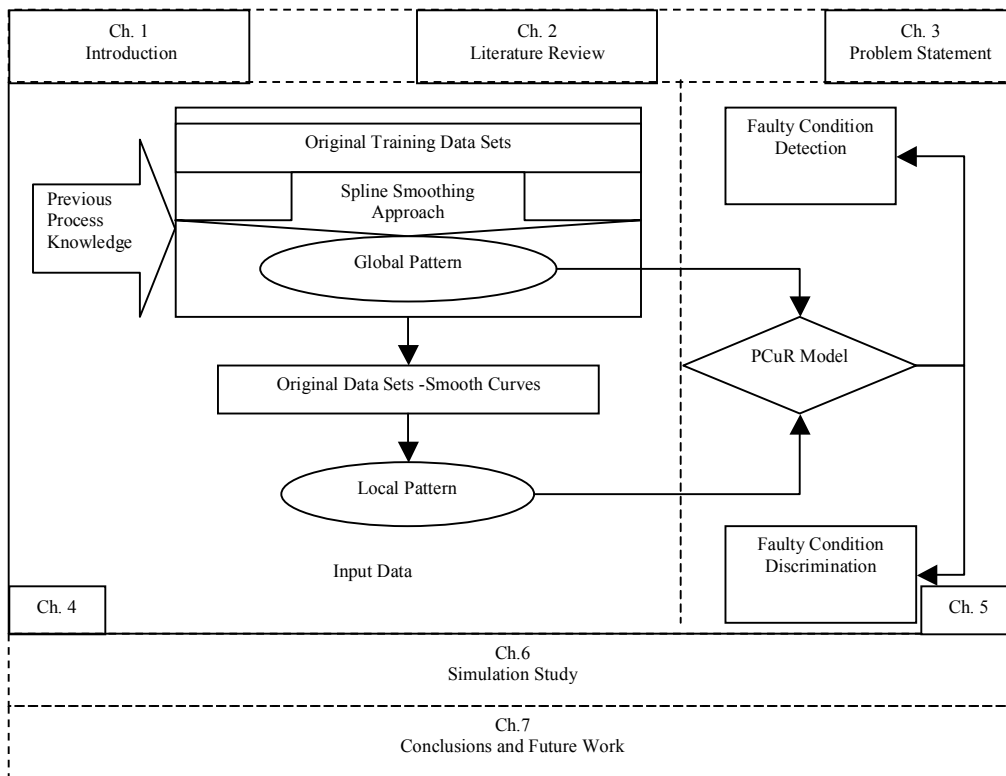
Figure 1.3 Thesis Flow Chart

Chapter 2. Literature Review

This chapter presents a summary of previous work on multivariate process monitoring, sensor fusion, process monitoring with MFD, and spline smoothing technique.

2.1   Multivariate Process Monitoring and Control

Multivariate quality control provides procedures to monitor multiple variables simultaneously. It is inherently more complex than univariate SPC, but it may be a more realistic representation of the data since in the real world processes do not usually have only one variable that is measured independently of all other variables in a system. However, multivariate control charts are designed to handle multiple variables in the shape of scalar vectors.

Several articles found in the multivariate control charts literature present different models and control charts to monitor manufacturing processes. Lowry and Montgomery [4] and Alt [5] presented reviews of multivariate control charts. They introduced brief discussion of Hotteling, Cumulative Sum Charts (CUSUM), and Exponentially Weighted Moving Average Charts (EWMA).

For instance, Alloway and Raghavachari [6] explored Hotelling (T^2) charts to represent multivariate data. Hotteling charts were used to detect shifts in the mean. James Williams et al. [7] extended the application of Hotelling (T^2) charts to monitor the coefficients of nonlinear regression applied on data profiles. They identified outlying observations and step or ramp vector in the mean over time.

For the purpose of process variation detection, Hawkins [8] proposed a multivariate control scheme based on the linear regression of each variable of the remaining variables. He also extended his regression in [9] for variables in multivariate quality control to a cascade process. The regression adjusted variables were plotted on CUSUM charts for both location and scale variations. The performance of his approach depends on the nature of the process being monitored. Pignatello and Kasunic [10] introduced a multivariate CUSUM control chart (MCUSUM) to detect small deviations and shifts; they used simulation to compare the performance of this chart to the multivariate Shewart chart and multiple univariate Shewart and CUSUM chart. This work can not be applied to time related processes or functional data, since further analysis should take place to consider correlation among time.

Fuchs and Benjamini [10] proposed the multivariate profile charts (PM) for Statistical Quality Control (SQC) which provides a higher accuracy in retrieving quantitative data information. Profile chart is a symbolic scatterplot, where summaries of data for individual variables are presented by symbol, and global information about the

8

group is displayed by the location of the symbol in the scatterplot. It is a display for the univariate and multivariate statistics.

There has been increase in multivariate quality control studies that use principal components (PCs) to aid monitoring of multiple variables i.e. to reduce the number of variables to be studied and monitored. Jackson and Morris [12] introduced the use of PCs instead of the original variables for multivariate quality control charts. The main advantage of PCs is that they are independent and few of them explain the variability of the original variables.

Sparks et al. [13] used the Gabriel Biplot in multivariate process monitoring, which allows the detection of locations in variation and correlation structure accurately. The Biplot is a way to simplify the interpretation of the Principal Component Analysis of a data matrix. It is based on its singular value decomposition.

For multivariate process change detection purposes, Apley and Shi [14] proposed a method based on Principal Component Analysis (PCA) to detect faults in a manufacturing process. They presented an example from the automotive industry, in which they determine which PCs are significant and analyze them to try to find the causes for the faults.

The previous work been done on control charts was applied on scalar data sets rather than functional non-stationary data or data with vectors that has correlation among

time. With current information technology and advanced data collection devices, manufacturing processes are monitored by sensors that reflects functional data in the shape of multiple curves to represent samples for process variable(s). This research studies the analysis and monitoring MFD taking time correlation into consideration.

2.2   Sensor Fusion

With current advances in technology and computational systems, more data can be easily collected from any controllable system. This relies on the fact that sensors ease the process of data collection and extraction. However, many challenges still exist when using sensors in industrial processes, such as: sensor allocation and sensor fusion problems.

The objective in previous research work in sensor fusion was mainly focused in determining optimal locations for sensors in a controllable system. Besides, other researchers proposed different approaches to extract and collect data from installed sensors [14-16]. In this research, our sensors related objective is to develop an approach to analyze the MFD that is collected from several allocated sensors, to detect process changes and to discriminate among different faulty conditions.

Process variation was lately used in sensor allocation problem, which has been studied in deep recently. For example, Y. Ding et al. [15] categorized traditional approaches in sensor optimization into two categories: multistation sensor allocation for the purpose of product inspection, and allocation of sensors for the purpose of variation

diagnosis but at a single measurement station. In their approach, based on the understanding of the mechanism of variation propagation, they developed a backward-propagation strategy to determine the locations of measurement stations and the minimum number of sensors needed to achieve full diagnosability by setting a minimization cost objective function.

In industrial processes, sensors may measure multiple realizations for the same process variable or different process variables in a given system. When sensors read more than one process variable, further analysis might take place to combine different outputs to generate multivariate characterization [15-16]. Many approaches could be used to analyze combinations of multiple variables, for example, principal component analysis reveals the variability among different variables, and correlation coefficients might reflect the relation among different monitored variables.

A common data fusion estimation approach for one variable system is to take the weighted average of various sensor realizations data to generate a composite fused value. Such approach may not lead to reliable measurements especially if one or more of the sensors are faulty [18].

Kalman filtering and extended Kalman filtering are another estimation approaches been used in this field. They are linear systems techniques which cannot be used if the process model is not available; however, they work well if the data is only corrupted by noise. Such approaches are very sensitive to process outliers. Moreover,

artificial intelligence approaches such as adaptive neural network and fuzzy logic require an extensive training of the system prior to performing the actual experiments.

Considering nonlinearity in sensor functions, Suranthiran et al. [18] developed a new framework to optimally fuse the multinonlinear sensor data based on the assumption that at any particular time instance. Only one measurement from the multi readings of the sensors can be used and a partial blending of signal is not possible. This was achieved by scheduling the sensors in the sense of one sensor reading at a time instance. This approach shows good result in reducing number of sensors to be used in processes; however, in this optimal approach the correlation among time in the functional data was not considered. Therefore, there is a need for an analytical approach that can gather MFD measurements from several sensors for process control purposes.

2.3   Process Monitoring With MFD

Different approaches were developed for process control purposes for MFD. Some approaches were developed for process variations detection, others were developed for process condition detection. However, there is no systematic approach exists that can interpret MFD and decompose it by analyzing the full process cycle for the purpose of process change detection and process faulty condition discrimination. This section presents some recent work on MFD and process monitoring.

For process variation detection, Mahmoud et al. [19] proposed a change point approach based on the segmented regression technique for testing the constancy of the

regression parameters in a linear profile data set. The advantages of their approach reveals more improved detection of sustained step changes in the process parameters and improved diagnostic tools to determine the sources of profile variation and the location(s) of the change point(s).

Using simulation study on functional data, Jeong and Lu [3] proposed a methodology based on wavelet coefficients monitoring, to detect local shifts in processes. Their approach performed effectively against many types of process changes especially in detecting small shifts. This work has been applied on single channel functional data.

Jin and Shi [20] presented a statistical approach for feature preserving data compression of tonnage information using wavelets. Moreover, Jin and Shi [21] developed a feature extraction methodology using PCA to represent variation patterns of tonnage signals, and they considered the interaction among the variables using fractional factorial design of experiment. Also, they used waveform signals [22] without prior faulty knowledge to develop a diagnostic system to automatically detect faulty conditions.

Some efforts have been made in studying sensed MFD and investigating different process conditions. Kim et al. [1] studied the detection of faulty condition and faulty condition discrimination in forging process based on segmentation approach, the forging cycle was divided into segments and for each segment the distances of the original data points from their principle curve was used to produce an empirical distribution which

leads to faulty condition discrimination. They also proposed an approach, using the empirical distributions for the squared distances of data points from their principle curve, to detect global and local profile changes in multi-channel tonnage signals and to classify faulty patterns.

Koh et al. [23] have developed an approach to detect faulty conditions based on collected data and engineering knowledge in tonnage of stamping machine sensed signals. The approach consists of partitioning signals into segments to represent process phases, at which they identified a set of signal attributes to describe the faulty conditions. Also, Koh et al. [24] used Haar transformation to detect and isolate multiple fault conditions. By partitioning signals into disjoint segment, mutually exclusive Haar coefficients were used to isolate faults at each stage of the process.

Jin [25] proposed a new method to use the partitioned monitoring segments of press tonnage signals to monitor individual station conditions in multi operation stamping processes. For this purpose, she developed a generic signal segmentation principle. She also used Hotelling T^2 control charts with consideration of the interactions among stations. She demonstrated the analysis procedure by providing a real case study of doorknob stamping process.

Zhou et al. [26] proposed a new approach to utilize sensed data and fault pattern information in process monitoring. They developed a directionally variant control chart through effective combination of multivariate chart and univariate projection chart. They

also showed that adding univariate projection charts can improve the detection power for the pre-known process faults. They demonstrated their work by providing a case study of cycle-based tonnage monitoring of a forging process. Their approach showed capability in process faulty condition detection whether faulty conditions are pre-known or not; however, the approach could not provide faulty condition discrimination and signal variation analysis.

2.4 Principal Curve and Principal Curve Regression Test in Process Monitoring

Principal Curve was first introduced by Hastie and Stuetzle [27] in 1989 as a smooth line which is self consistent curve which means that each point of the curve is the average of all data that is projected onto it. This curve has many advantages such as no distribution assumption about data. The data can be readily analyzed without transformation as opposed to Non- Linear Principal Component Analysis (NLPCA). Besides, the application of this curve does not need any constraint on the dimension of the data as the curve feature maintains the shape of the original data set. Global and local changes in the curve indicate the change in the data.

Many authors studied further properties for the principal curve and proved its existence. For example, Duchamp and Stuezle [28] proved the existence of principal curves and they showed that principal curves are critical points of the expected squared distance from the data. Kégl et al. [29] introduced principal curves with fixed length with the proof of their existence and uniqueness.

The definition of the principal curve opened a wider vision for other researchers in the application of principal curves. For example, Banfield and Raftery [30] applied the principal curve to identify ice floes in satellite images and they reduce the estimation of bias that Hastie and Stuetzle [27] have introduced. In addition, LeBlance and Tibshirani [31] extended the principal curve concept to principal surfaces as they used it in multivariate regression splines study. Standford and Raftery [32] applied principal curves in clustering.

Chang and Ghosh [33] applied principal curves in feature extraction and pattern classification problems. In most of their experiments they noticed that closed principal curves gives better results than opened principal curves in pattern classification problem.

Lately, Huang [2] introduced the concept of Principal Curve Regression (PCuR) as a model based on the principal curve of data signals, the generated principal curve combine data using multivariate regression model. He applied this concept on tonnage signals. Using in control data the model can predict deviations in future observations.

2.5   Spline Smoothing

Before embarking to the literature survey of spline smoothing, we introduce a brief review about spline smoothing. Splines are drafting aids that draw smoothed curves across data points by interpolating data points that are called knots. In this research we use spline for data interpolating since splines are piece wise functions whose individual curves meet at knots. This occurs because usually in splines the first and the second

derivatives at the end of a polynomial, fitting an interval, are equal to the corresponding derivatives at the beginning of the following polynomial. This feature guarantees continuity among data knots [34].

Wegman and Wright [35] provided a detailed bibliographic review of previous work on spline smoothing. They emphasized that splines are interpolatory in nature. Besides, they demonstrated that the interpolation problem in splines is to fit curve through points in a plane. Besides, they introduced the cubic interpolating spline as a function with continuous derivatives up to and including order 2. They also categorized methods of spline smoothing according to the different ways of dealing wit noise into three categories: least square method, 100 percent confidence intervals methods, and regression splines method.

Spline smoothing is usually used because of its natural and flexible features in fitting data [36]. Using the spline regression, the smoothing parameter can be determined based on the need of the study and data interpretation in a way which represents the trade off between two competitive aims, good fitting and avoiding rapid fluctuation. Moreover, the amount of smoothing does not depend on the response values, instead, it depends on the design points and it copes well whether or not the design points are regularly spaced.

Eilers and Marx [36] presented a short review of B-splines. They showed connections to the familiar splines penalty on the integral of second squared derivative.

They also used nonparametric logistics regression, density estimation and scatterplot as examples.

In spline smoothing, the main fitting objective is to minimize the Residual Sum of Squares (RSS) and the local variation. A measure of the rapid local variability of a curve can be given by a roughness penalty such as the integrated squared second derivative. The modified sum of squares will be defined as follows [37]:

$$S(g) = \sum_{i=1}^{N} [y_i - G(x_i)]^2 + \lambda \int [G''(t)]^2 dt \qquad (2.1)$$

where: λ is the smoothing parameter, Y is the original data, and $G(t)$ is the fitted value. Therefore, two extreme cases might occur:

1) As $\lambda \to \infty$, the highest amount of smoothing occurs with the linear least squares fit with degree of freedom equals to 2.

2) As $\lambda \to 0$, the lowest amount of smoothing occurs with a curve that is similar to the original data since no basis functions were estimated and no degrees of freedom will be taken out from the system, thus the system will have a degree of freedom equals to the same number of design points.

The amount of smoothing can be controlled by either controlling the smoothing parameter or by specifying the degree of freedom. Choosing the smoothing parameter can be subjective by choosing the curve that fits the data subjectively with a corresponding smoothing parameter. On the other hand, a very well known proposed approach for an automatic choice of the smoother parameter is the cross validation method.

18

As any other linear operator, smoothed fitted curves can be represented as linear smoother:

$$\hat{G} = S_\lambda y \qquad (2.2)$$

where, y is the fit and S_λ is the linear operator and known as the smoother matrix.

The smoother matrix S_λ is symmetric positive $n \times n$ matrix, where n is the number of design points. The expression trace(S_λ) gives the number of basis functions or the number of parameters involved in fitting, thus the effective degree of freedom can be written as follows [38]:

$$df_\lambda = trace(S_\lambda) \qquad (2.3)$$

The smoother matrix diagonal values can be determined depending the on the design knots using the following equation [38]:

$$S_{\lambda i} = \lambda^{-1/4} n^{-3/4} 2^{-3/2} f(t_i)^{-3/4} \qquad (2.4)$$

where $f(t_i)$ is the standard normal density function of the design points.

Choosing the smoothing parameter in least squares spline smoothing problem has been advocated by Wahba and Wold [39, 40], and Wahba [41-43]. Wahba introduced a measure for the goodness of fitting as the average squared error which is called the cross validation function. She also demonstrated that choosing the smoothing parameter is base on minimizing the cross validation function. The basic principle behind the cross validation approach is to leave the data points out one at a time and to choose the value of

the smoothing parameter which the missing data points are best predicted by the remainder if the data.

For example if G^{-i}_{λ} is the smoothing spline calculated from all the data pairs except (X_{t_i}, y_i) using λ smoothing parameter, the cross validation choice of λ is then the value of λ that minimizes the cross validation score XVS given by the following equation:

$$XVS(\lambda) = n^{-1} \sum \{y_i - G_{\lambda}^{-i}(t_i)\}^2 \qquad (2.5)$$

An easier computation of cross validation score is given in a standard argument in regression theory by Cook and Weisberg [44] and Craven and Wahba [45].The following equation represents their easier form of evaluating XVS:

$$XVS(\lambda) = n^{-1} \sum_{i=1}^{n} \frac{\{y_i - \hat{G}(t_i)\}^2}{\{1 - trace(S_{\lambda})\}^2} \qquad (2.6)$$

Caraven and Wahba also introduced a weighting technique to be implemented on the cross validation function to reflect unequally spaced data and called it the generalized cross validation function. This function was introduced by a simple matrix representation depending on the smoother matrix in the spline smoothing. Fleisher [46], Merz [47], and Paihua-Montes [48] provided a computer code of penalized least squares smoothing spline using the generalized cross validation function. Splines also have been introduced in the time series field. Wahba [43] developed the theory of periodic smoothing spline with application to spectral density estimation.

More sophisticated techniques for smooth spline smoother matrix and its role in curve fitting were discussed by Cook and Weisberg [44]. In their book they discuss approaches to calculate the smoother matrix and the cross validation function.

Many spline smoothing applications were studied by Silverman [38]. In his paper he promoted the applicability of non parametric regression using cubic splines. He also provided an inference region for curves as a confidence interval based on Bayesian formulation. Silverman also proposed a formulation for weighted smooth spline smoother matrix. Motivated from Craven and Wahba results, Silverman [49] also developed an approximation for the smoother matrix to quite reduce computational burden by providing a detailed mathematical justification to approximate the generalized cross validation score.

2.6   Summary

The most commonly used multivariate control charts are based on the Hotelling T^2 chart. Such control charts are dealing with scalar data rather than multiple curve functional data. This implies the necessity of an approach to analyze MFD for the purpose of process change detection and faulty condition discrimination.

Two major objectives were studied in the sensor fusion problem. They were mainly concentrated in the determination of optimal sensors location and the optimal extraction for sensed data. However, in this research we analyze the MFD sensed by

21

allocated sensors for the purpose of process change detection and faulty condition discrimination.

The majority of the work that has been done using principal curves shows the capability of the principle curve in summarizing the data since it has no distribution assumption, and the data can be readily analyzed without transformation as opposed to NLPCA approaches.

Using principal curve as a response in multivariate linear regression model as in the PCuR model shows the ability to detect deviations in multi-channel samples. Since principal curve feature maintains the shape of the original data; global and local changes in the data can be reflected on the curve.

Most of the work that has been done in faulty conditions diagnosis and monitoring has been applied on segments of channels signals. Some of the approaches depended on extracting empirical distributions from the data; others were depending on subjective judgment in pattern distinguishing. However, the need for an online solid approach for processes faulty conditions diagnosis offers opportunities for further research in this area. All the previous issues lead to the problem statement for this research, presented in the following chapter.

Chapter 3. Problem Statement

This chapter presents the problem definition in this research. Besides it gives a brief description about the approach being used to solve the problem. A summary of the main contributions that we have in this research is also provided.

This research takes account of the following issues: MFD interpretation, process change detection, and process faulty condition discrimination. SPC concepts are not directly applicable to MFD [2, 3], since it is focused on the analysis of multivariate scalar data, and no systematic approach exists to analyze the complexity of variations in MFD. Also, late process change detection may result in high quality-costs and large number of defectives, especially if more than faulty condition exists. Motivated by this fact, we developed an approach to interpret the variations of MFD for process change detection and process faulty condition discrimination. The approach works on both multivariate processes and univariate processes with multiple realizations.

We decomposed MFD into global and local components to reveal global and local variations in sensed signals. Global patterns were extracted by applying weighted spline smoothing on original data. Weights were determined using local moving average of the generalized residuals. This method provides weights which correspond to homogenous

variance among process cycle. Therefore, any variation in the process cycle can be captured. Local variation profiles were obtained by subtracting the global component from the original signals.

Controlling and monitoring processes that include MFD become a subsequent challenge. For process change detection purpose, we used global variation component to detect changes in process variation. In-control global variations used as training data sets in PCuR test; therefore, process faulty conditions can be detected.

For the purpose of process faulty condition discrimination, we used in control local variation as an input in the PCuR test. Process faulty condition that is due to local shifts appears out of control, whereas, process faulty condition that is due to global shifts remains in control since it is generated by shifting whole the cycle. We also conducted an exploratory study to extract the appropriate data sets needed to increase the performance of prediction in the PCuR model and to reduce the number of false alarms.

The approach proposed was applied on real life forging data. A simulation study was also conducted to verify the developed approach. We simulated bivariate normal in-control data sets with faulty conditions due to global shifts and faulty conditions that are due to multiple local problematic segments. Using the developed approach, no false alarms were observed, and results show the applicability of the developed approach on further applications.

Chapter 4. Global and Local Variations Decomposition

This chapter presents the methodology being developed to extract global and local variations in MFD. It provides a description about global and local components, the decomposition of global and local variations.

4.1 Global and Local Variations

This section provides a description about the global and local components. Before embarking to the approach, the assumptions for the MFD to be analyzed in this paper are: 1) the data are all cyclic; 2) the lengths of the cycle are the same; and 3) in each channel, the observations among cycles are identically independent distributed.

Since it is hard to interpret sensed MFD, unlike in vector characteristics where the mean and variance can describe process variability, we decomposed original signals into global and local patterns to extract process variation. The idea behind this decomposition is to separate global variations from local variations by which global signal shifts and local variations pattern can be captured. Global variation was extracted by developing smoothed curves using weighted spline smoothing in the sense that we increase local variations in problematic cycle intervals and decrease it elsewhere in the process cycle based on previous process and engineering knowledge. Local variation was extracted by

subtracting global variation from original signals. Using this signal decomposition we can separate defective samples that were caused by global shift in the process cycle from others which were caused by local variation in problematic process segments. Eq. 4.1 presents the decomposed components of the original signal.

$$Y(t) = G(t) + L(t) \tag{4.1}$$

where, $Y(t)$ is the original signal, $G(t)$ is the global variation pattern, and $L(t)$ is the local variation pattern.

Our approach was applied on real forging data. Three data sets were collected under different faulty conditions: normal condition, faulty condition one which corresponds to deformity in counterweight, and faulty condition two that occurs due to incorrect die setup. Each data set, under each process condition, includes 41 samples of four channels representing tonnage signals as a function of 90 design points of process time.

Such decomposition can be illustrated with faulty conditions in a real forging process. In Figure 4.1, three data sets were collected under different faulty conditions: normal condition, faulty condition one which corresponds to deformity in counterweight, and faulty condition two that occurs due to incorrect die setup. Each data set, under each process condition, includes 41 samples of four channels representing tonnage signals as a function of 90 design points of process time. We noticed that faulty condition two curves behave almost the same like normal condition curves. The faulty condition one occurs

most likely in the problematic interval [20, 30] of the cycle, whereas faulty condition two takes place over all the cycle as a shift in the tonnage signal. Noticing the dissimilarity in behavior for faulty condition one than the other two conditions in the problematic interval (see Figure 4.1) [1]. We can look to the problem in a way by which we can extract data from the curves that can distinguish faulty conditions. The following section presents the methodology of decomposing signals into global and local patterns based on previous knowledge.
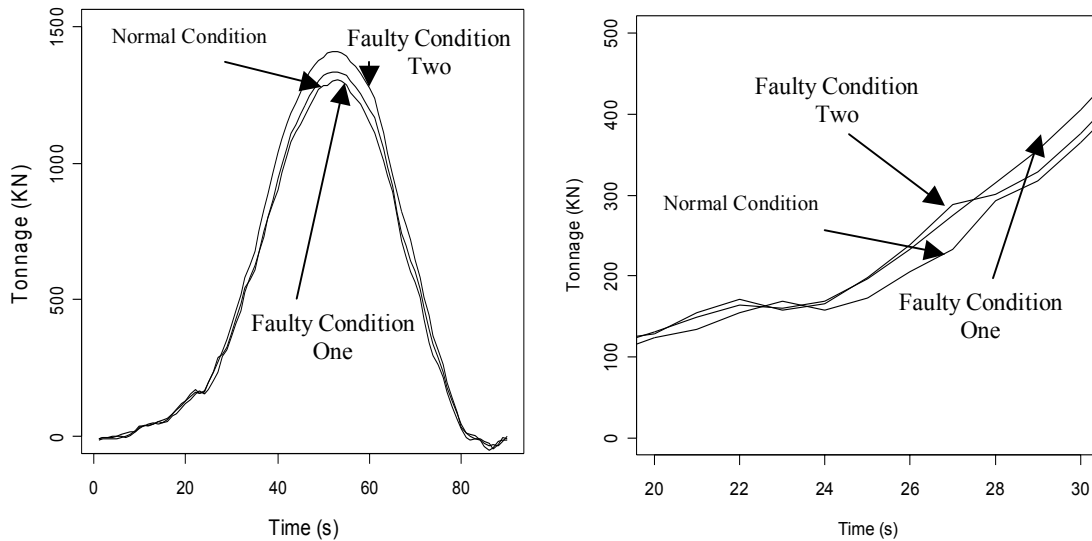

Figure 4.1 Different Process Conditions

Left: global pattern; Right: local pattern in the Problematic Time Interval

4.2   Global and Local Variations Extraction Using Weighted Spline Smoothing

4.2.1 Global and Local Variations Extraction

Using weighted spline smoothing for data sets under normal condition, we define the process weights vector (w) as a combination of different weights assigned based on local moving average of squared generalized residuals [38]. This weight vector is further used to produce global patterns for other process conditions to detect local and global variations for each process channel.

Using weights vector, the spline smoothing curves after then will be determined by three main factors: local ordinary residuals, roughness penalty, and the weights values. Eq. 4.2 represents the modified residual sum of squares demonstrating the trade off among the three factors.

$$S(g) = \sum_{i=1}^{n} w_i [y(t_i) - G(t_i)]^2 + \lambda \int [G''(t)]^2 dt \qquad (4.2)$$

The generalized residuals were used since it do not depend on the residuals variance and the smoother matrix corresponding value, besides it is useful in diagnostic procedures to detect outliers. We notice in (4.2) that generalized residuals are a scaled form of the ordinary residuals. The generalized residuals can be defined as follows [38]:

$$r_i = \frac{w^{1/2} \{Y(t_i) - G(t_i)\}}{\sigma \{1 - n^{-1} tr(S_\lambda)\}^{1/2}} \qquad (4.3)$$

where σ is an estimate of residuals standard deviation factor;

$$\sigma = \sqrt{\frac{\sum w_i \{Y(t_i) - G(t_i)\}^2}{n - tr(S_\lambda)}} \qquad (4.4)$$

28

Our approach starts with assigning a uniform vector of weights with $\{ w_i = 1 \}$. The initial values of the residuals standard deviation and generalized residuals are used in finding the first combination of weights based on local moving average of the generalized residuals. Eq. 4.5 shows the formula for finding the first weights vector for each design knots [38].

$$w_i = \frac{(d_i - m_i + 1)}{\sum\limits_{j=m_i}^{d_i} r_j^2} \tag{4.5}$$

where, $d_i = \min(n, i+k)$ and $m_i = \max(1, i-k)$

Since our data sets are of moderate size, k value was determined to be 5 as in [38].

The first weights combination produced is further can be fed back into the weight estimate model to produce new updated weight estimation $\hat{w}$ as follows:

$$\hat{w}_i = \frac{w_i (d_i - m_i + 1)}{\sum\limits_{j=m_i}^{d_i} r_j^2} \tag{4.6}$$

These steps can be applied till we reach iteration where the produced combination of weights generates a corresponding generalized residuals plot with homogeneous variance. i.e., random generalized residual plot or some sort of convergence occurs. See Figure 4.2.

As a result, the weights vectors are approximately similar for data variables extracted from the same sensor among different samples under normal process condition,

and there is no great need to have accurate weights values [38]. Therefore, an average

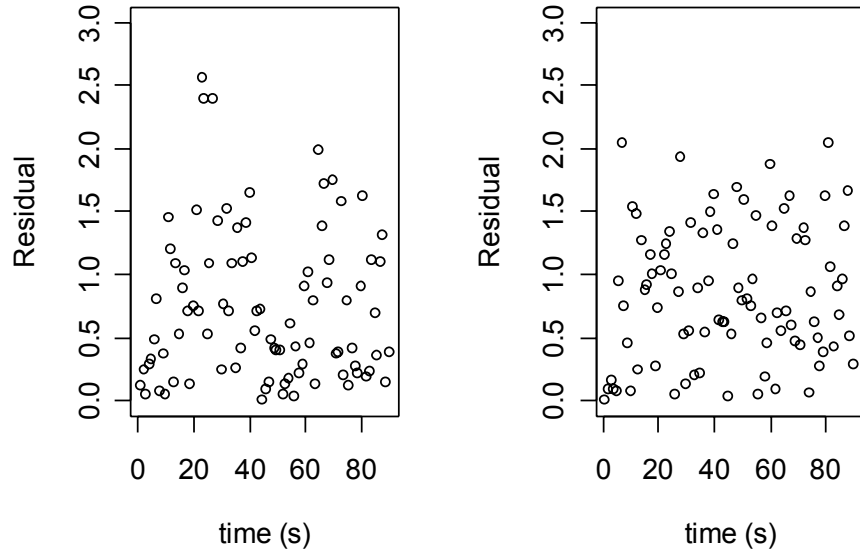weight value can be used to produce global pattern for future samples.

Figure 4.2 Absolute generalized residuals for reweighed Tonnage Signal
Left: First reweighing iteration; Right: second reweighing iteration

The basic idea of finding the cubic polynomial for a given interval can be shown

in the following equation [48]:

$$P(x_i) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \tag{4.7}$$

for i=1,2,…, n-1.

where, $a_i, b_i, c_i, and \ d_i$ are coefficients, and $(x - x_i)^3, (x - x_i)^2, and \ (x - x_i)$ are called

basis.

Then, the polynomial fit $G(t)$ over multi intervals will be represented as follows:

$$G(t_i) = \sum_{i=1}^{n} \theta_i N_i \tag{4.8}$$

where N_i is an n-dimensional set of basis functions for representing the natural splines, and theta (θ_i) stands for the basis coefficients.

Based on the spline features of continuity and first and second derivative existence, the polynomial coefficients can then be determined. Since the solution for the chosen global function is cubic spline, the function can be represented as follows:

$$G(t) = \sum_{j=1}^{n} w_i N_j(t_i)\theta_j \tag{4.9}$$

The new smoothing parameter (λ_G) will be then determined using the extended argument of the cross validation method developed by Silverman in [38] for the weighted case. Consequently, the effective degree of freedom generated using this approach (Df_G) can be determined using the following equation:

$$Df_G = trace(\lambda_G^{-1/4} w_i (\sum w_k)^{-3/4} 2^{-3/2} f(t_i)^{-3/4}) \tag{4.10}$$

where $\sum w_k$ is the sum of weights of the design knots at $t=i$, and $f(t_i)$ is the standard normal density function of the design points.

Using the weighted spline smoothing or the corresponding degree of freedom, the generated spline smoothing will have the required features for the purpose of process condition detection and discrimination. In Figure 4.3, we notice that this approach reveals that faulty condition one occurs due to high local variations in the problematic segment,

whereas faulty condition two occurs due to global shift in the tonnage values. This result

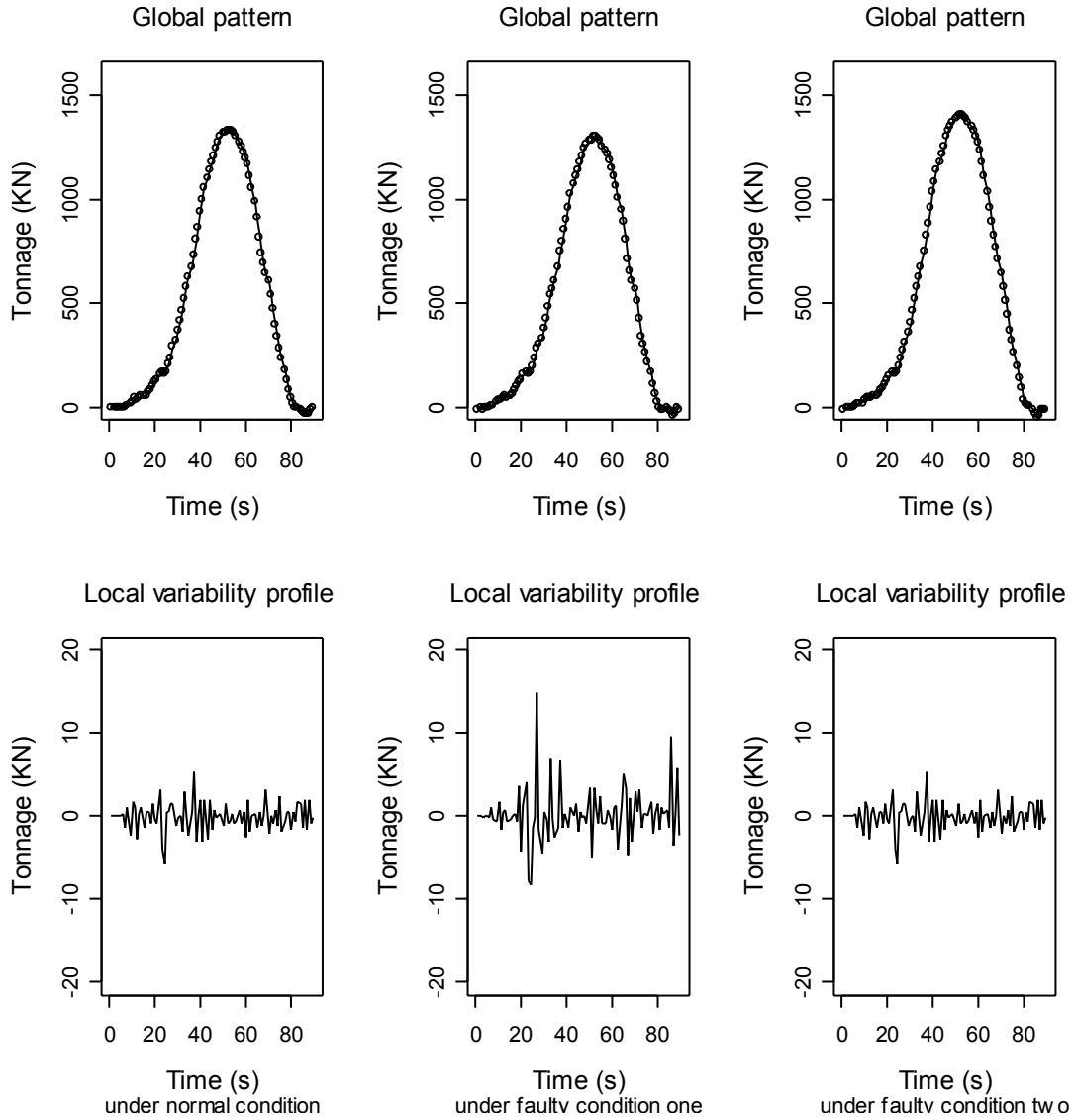complies with the study conclusions conducted in [1].

Figure 4.3 Sample of Global and Local Variation Extraction

4.3   Summary

In this chapter we discussed the spline smoothing technique used in developing global patterns to the original data in order to discriminate between different faulty conditions. The weighted spline smoothed curves were used to produce the global pattern of the original signals. Weights at each data knot were determined using the local moving average of the generalized residuals. Local variability profiles were extracted by subtracting original data curves from smoothed global patterns.

The next chapter presents the PCuR model and test used to detect faulty conditions and to distinguish between them. Besides, the selection of training data sets based on clustering technique is also discussed.

Chapter 5. Monitoring Global and Local Variations of MFD

Obtaining global and local variations from original signals form a key feature of signal variations. For the purpose of process conditions detection and discrimination, we used global and local patterns as an input to the PCuR model.

5.1 Review of Principal Curve Regression Model

Recalling PCuR in chapter 2, Huang [2] introduced the concept of Principal Curve Regression for tonnage signals as follows:

Denote by x_k the p-channel tonnage signal observed at the kth crank angle of a cycle, $k=1,2,\ldots,m$, where $x_k = (x_{k1},\ldots, x_{kp})^T$ and the signal is of the form $\{\mathbf{x}_k\}_1^m$. If tonnage signals of n parts are observed, denote by $\{\mathbf{x}_{k,i}\}_1^m$ the ith observation of $\{\mathbf{x}_k\}_1^m$, where $x_{k,i}=(x_{k1,i},\ldots, x_{kp,i})^T$ is the ith observation of x_k, $i=1, \ldots, n$ and $k=1, \ldots, m$. The n samples are of the form $\{\mathbf{X}_k\}_1^m$, where the $n \times p$ matrix $X_k=[x_{k,1}, x_{k,2}, ..., x_{k,n}]^T$.

The idea of PCuR is to extract principal curve from $\{\mathbf{x}_k\}_1^m$ and build regression model between the principal curve and $\{\mathbf{x}_k\}_1^m$.

Once the PCuR model is established from in-control tonnage signals, it can be used to determine whether process change occurs in future observations. The principal curve of new observation $\{\mathbf{z}_k\}_1^m$ is treated as new response $\{\mathbf{y}_k\}_1^m$.

A multivariate linear regression model is assumed to adequately model "response" $\mathbf{y}_k$ and "predictors" $\mathbf{x}_k$ at the kth crank angle, k=1, …, m, i.e.,

$$\mathbf{y}_k^{\mathrm{T}} = [1\ \mathbf{x}_k^{\mathrm{T}}]\,B_k + \varepsilon_k \qquad (5.1)$$

with

$$E(\varepsilon_k) = 0,\ \mathrm{Cov}(\varepsilon_k) = \Sigma_k, \qquad k = 1,\ \dots,\ m. \qquad (5.2)$$

By the results of Johnson, R.A, and Wichern, D.W. [50] in multivariate regression. The predicted ellipsoids for new response y_k at time k are

$$(Y_k - \hat{B}_k^T \begin{bmatrix} 1 \\ Z_k \end{bmatrix})^T \left(\frac{n}{n-p-1} \hat{\Sigma}_k \right)^{-1} (Y_k - \hat{B}_k^T \begin{bmatrix} 1 \\ Z_k \end{bmatrix}) / (1 + [1 z_k^T]([1X_k]^T[1X_k])^{-1} \begin{bmatrix} 1 \\ Z_k \end{bmatrix}) \leq \left(\frac{p(n-p-1)}{n-2p} \right) F_{p,n-2p}(\alpha) \quad (5.3)$$

We used Bonferroni correction procedure to decrease the α value. This way less false alarms can occur in PCuR test. The primary α value for each data knot is 10%, therefore, the actual testing α value for all the cycle is (α/M)=0.0011 for every time knot.

Huang [2] applied the PCuR test on the problematic interval and results in large number of out of control points appearing under faulty condition one compared to the samples of faulty condition two.

5.2 Monitoring Global and Local Variations Using PCuR

In this research we applied PCuR model to the forging samples without segmentation, i.e. a sample consists of 90 time knots and training data sets includes the first thirty samples of normal process condition. For process change detection purpose, we applied PCuR model to the global variation data sets, which were extracted by applying spline smoothing on the original signals. The results for normal condition data sets remain in control, whereas data points for faulty condition one and two exceeded threshold randomly, which indicates out of control status for data points.

For the purpose of faulty condition discrimination, we applied PCuR model on local variation profiles. As a result, all the samples under normal condition remain in-control. However, all samples of faulty condition one are out of control and 34 samples of faulty condition two remained in control.

The following figure presents a sample result of applying PCuR on extracted global variation using in-control data sets. Results of applying PCuR test on global and local data sets using other process conditions are shown at the end of this chapter.
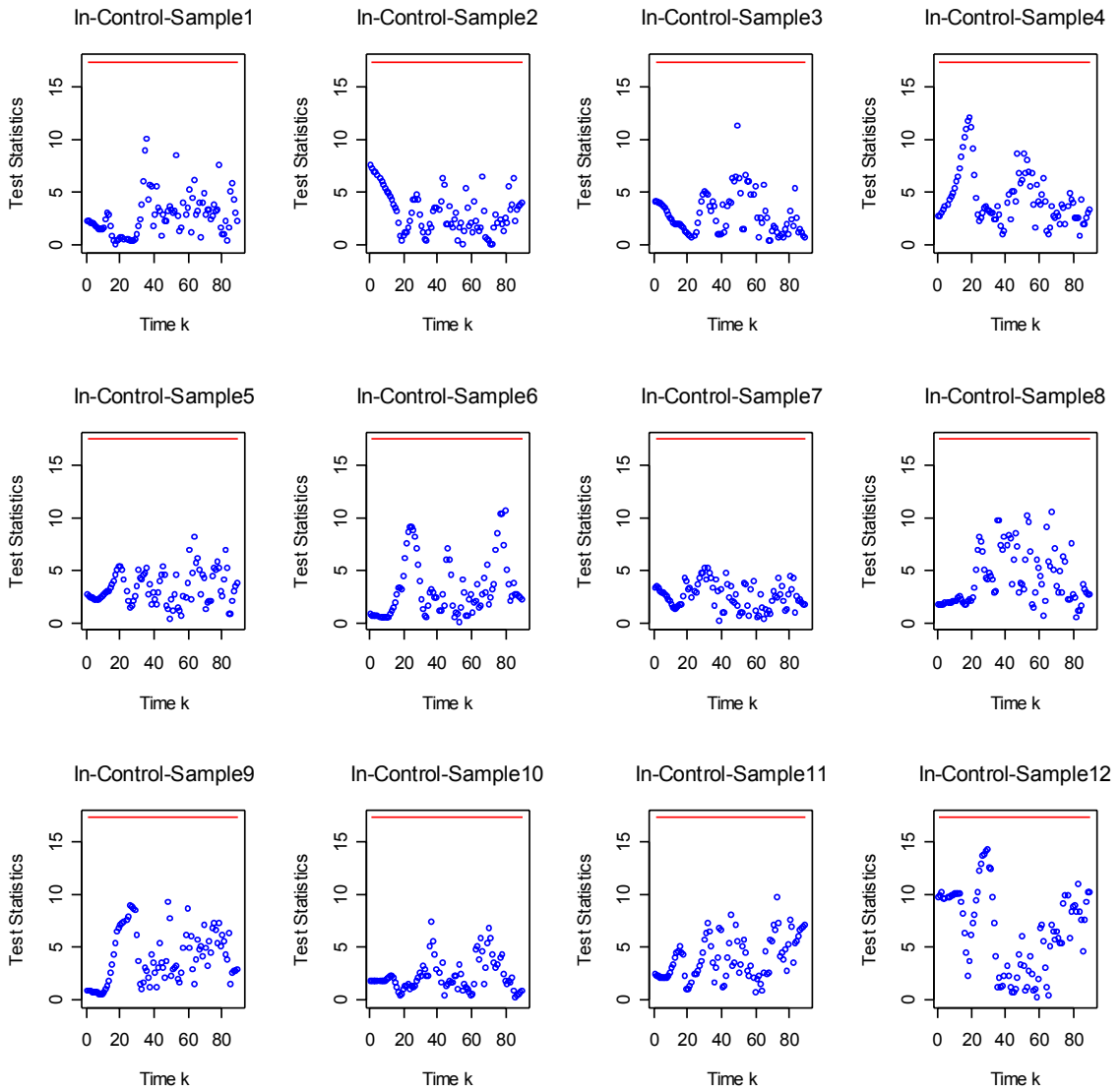
Figure 5.1 Result of PCuR Test on Global Pattern Under Normal Condition

Using global variation data sets for process change detection, global and local variations were detected. On the other hand, local variation data sets were used for faulty condition discrimination purposes. Since faulty condition two is due to a global shift in the cycle, the corresponding local profile bear a resemblance to the local profile of the

training data sets; therefore it appears in-control. However local profiles for faulty condition one appears of control due to local variations. The result indicates the existence of some false alarms in applying PCuR model on local variations for faulty condition discrimination purpose. To illustrate the main cause of false alarms, we conducted an exploratory study to capture the variability among samples.

5.3 Exploratory Study

According to Statistical Quality Control (SQC) concepts, an out of control status can be caused by variability and/or shift in the sample mean. This concept motivated the study of the a) RSS of the local variability data sets under normal condition for each sample composed of four channels p=4.

$$RSS_s = \sum_{c=1}^{p} \sum_{i=1}^{n} (Y_c(t_i) - G_c(t_i))^2 \tag{5.4}$$

and b) the mean of the residuals coming from the four channels for each sample under normal condition. See Figure 5.7-(a,b).

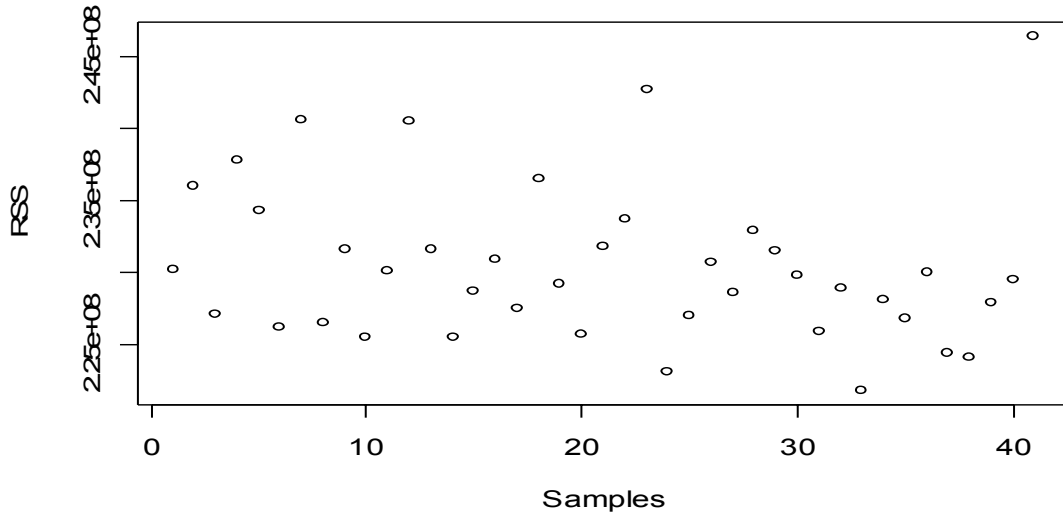$$\overline{R}_s = p^{-1} \sum_{c=1}^{p} \sum_{i=1}^{n} (Y_c(t_i) - G_c(t_i)) \tag{5.5}$$

where,

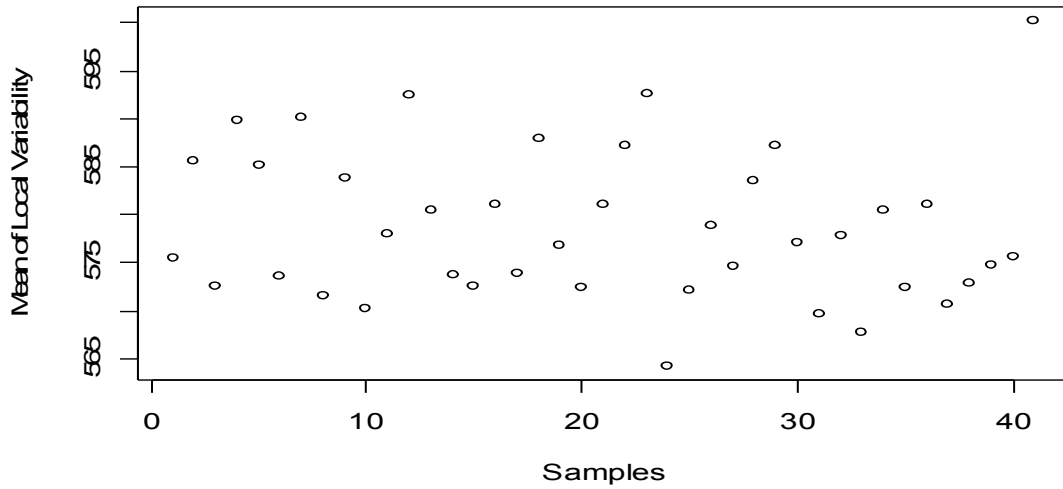RSS_s is the residual sum of squares for sample s, s=1,…,41

$\overline{R}_s$ is the residuals mean for sample s.

$Y_c(t_i)$ and $G_c(t_i)$ is the original and the spline smoothed data respectively at each knot i , i=1,…,n for channel c , c=1,…,p.

The following figure presents the RSS magnitudes and mean of residuals for all the samples collected under normal condition.

(a) RSS Under Normal Condition

(b) Mean of Residuals Under Normal Condition

Figure 5.2 RSS and Mean of Residuals

We chose twenty samples from each study as training data set that correspond to low variation for the PCuR model to predict faulty condition two as in control samples since it is caused by global signal shift. Eighteen samples were in common between the two studies groups N_{RSS} and $N_{\bar{R}}$ as follows:

$$N_{RSS} \cup N_{\bar{R}} = \begin{Bmatrix} 1 & 7 & 13 & 19 & 30 \\ 2 & 8 & 15 & 20 & 31 & 40 \\ 3 & 9 & 16 & 24 & 36 & 41 \\ 5 & 10 & 17 & 27 & 38 \end{Bmatrix}, \quad N_{RSS} \cap N_{\bar{R}} = \begin{Bmatrix} 1 & 13 & 19 & 30 \\ 5 & 15 & 20 & 31 & 40 \\ 9 & 16 & 24 & 36 & 41 \\ 10 & 17 & 27 & 38 \end{Bmatrix}$$

Using the obtained samples as training data set for PCuR test, the results improved by reducing the false alarms from 7 to 5 in detecting faulty condition two. The five false alarms samples of faulty condition two were out of control in few separate random knots i.e. not in the shape of cluster. Logically, this fact does not reflect defective condition occurrence on the working piece, since for a specimen to be defective, thermoplastic deformation may last for a time interval which will be reflected as cluster of out of control points in PCuR model. To interpret the result obtained by PCuR test for faulty condition two, we conducted the same exploratory study to capture the variability among samples collected under faulty condition two. The samples with high variation and shift in the mean $F_{RSS}, F_{\bar{R}}$ are represented below:

$$F_{RSS} = \begin{Bmatrix} 1 & 2 & 3 & 5 \\ 6 & 7 & 16 & 22\,39 \\ 26 & 27 & 32 & 33\,41 \\ 34 & 35 & 36 & 38 \end{Bmatrix} \qquad F_{\bar{R}} = \begin{Bmatrix} 1 & \\ 2 & 27 \\ 3 & 33 \\ 25 & 35 \end{Bmatrix}$$

The false alarm samples for faulty condition two (*F*) are:

$$(F) = \left\{ \begin{array}{l} 1 \\ 2 \\ 3 \end{array} \begin{array}{l} 33 \\ 35 \end{array} \right\}.$$

We notice that faulty condition two has two groups of variability in its data set. The existence of some samples with high local variation caused the existence of false alarms in discriminating faulty condition two.

5.4   Summary

In this chapter global and local variations under normal process condition were used as training data sets for the purpose of monitoring in the PCuR model and test. Global variation was used to detect process change, whereas local variation was used to discriminate among faulty conditions. However, some false alarms occurred, therefore, we conducted an exploratory study to capture the variability among the in control samples. The study was based on grouping and cluster analysis. The next chapter presents a simulation study performed to verify the proposed approach in decomposing MFD into global and local components.

Results of applying PCuR test on different process conditions using global and local variations are shown below:
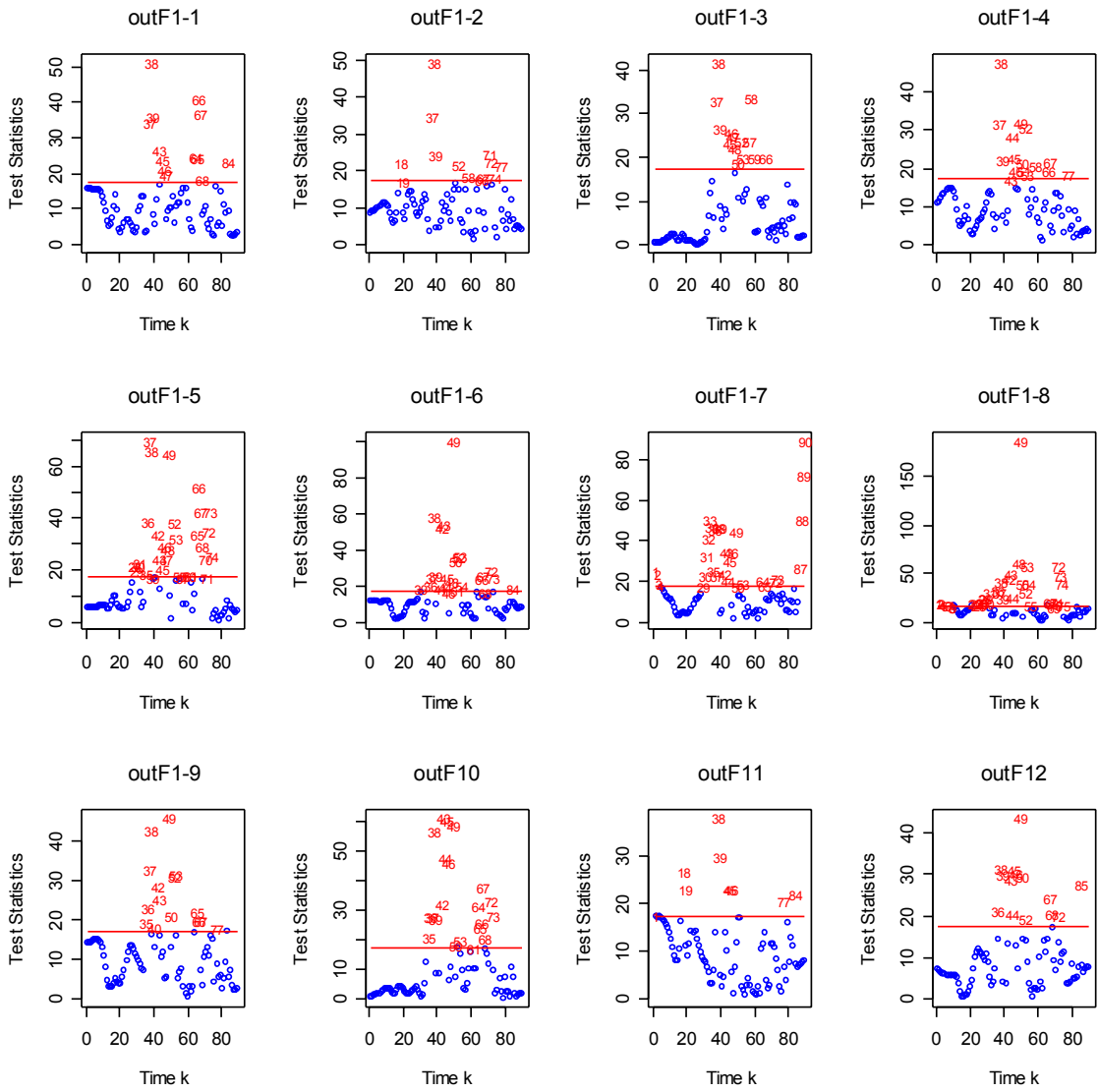
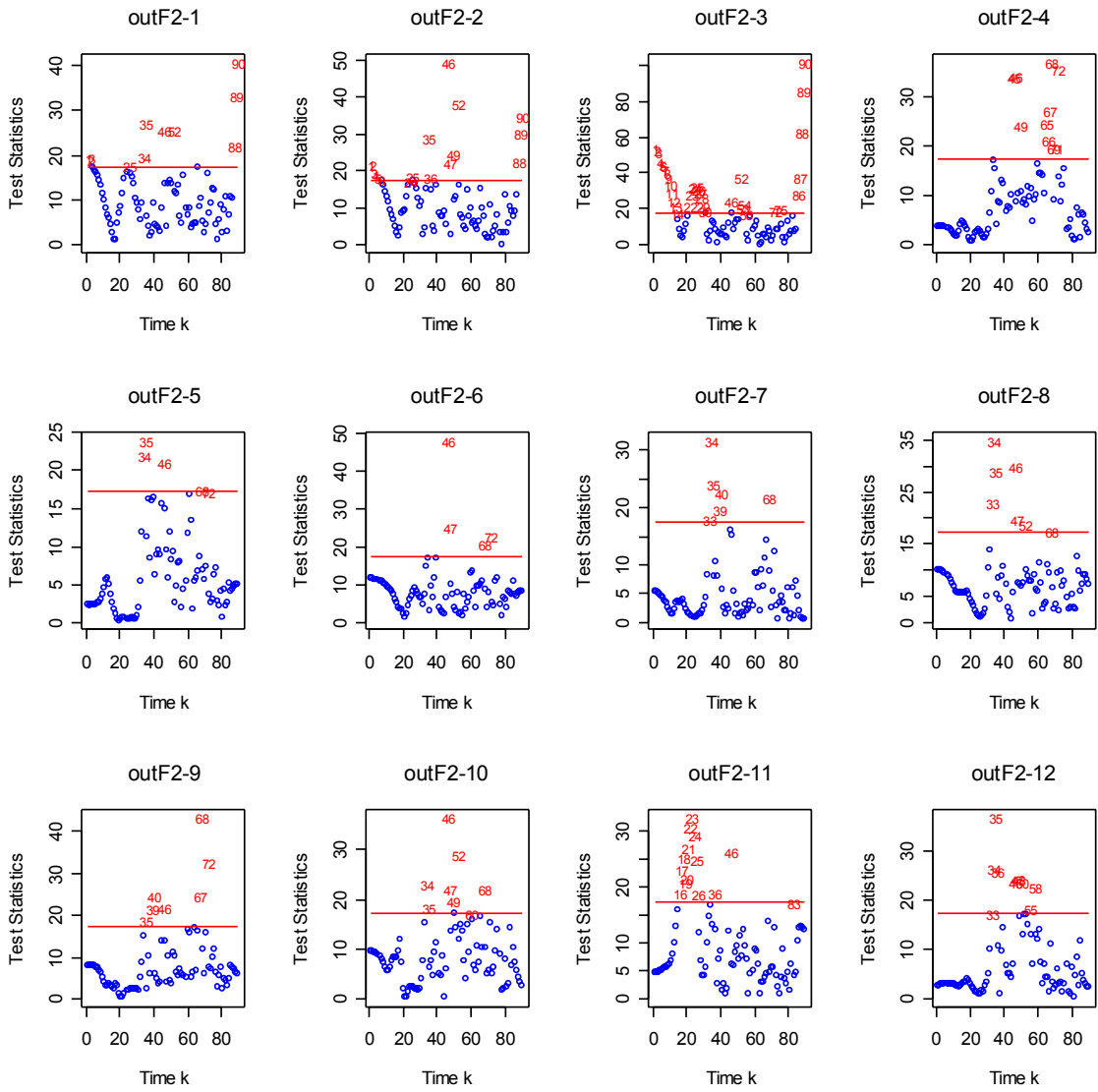Figure 5.3 Result of PCuR Test on Global Pattern Under Faulty Condition One

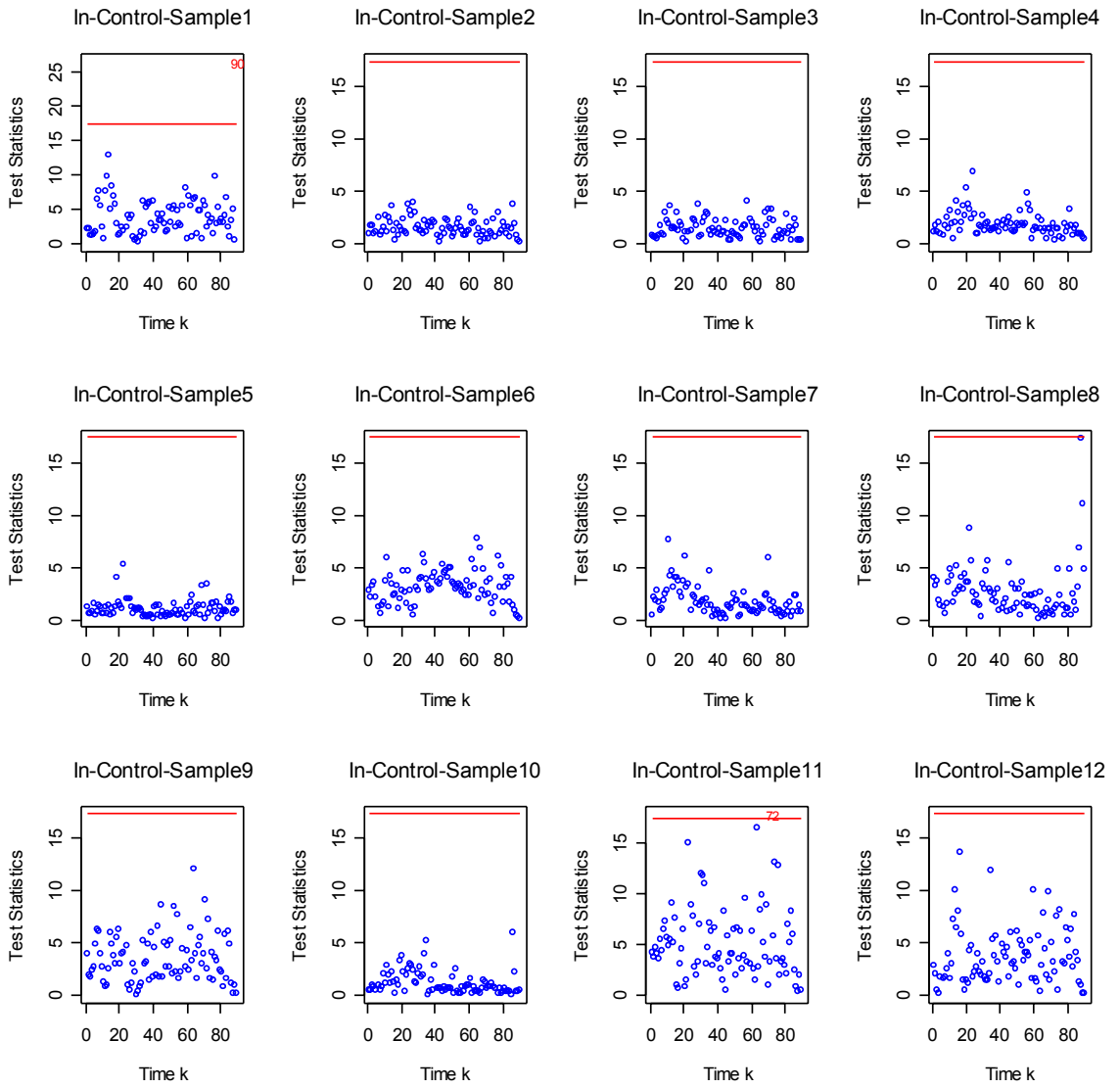Figure 5.4 Result of PCuR Test on Global Pattern Under Faulty Condition Two

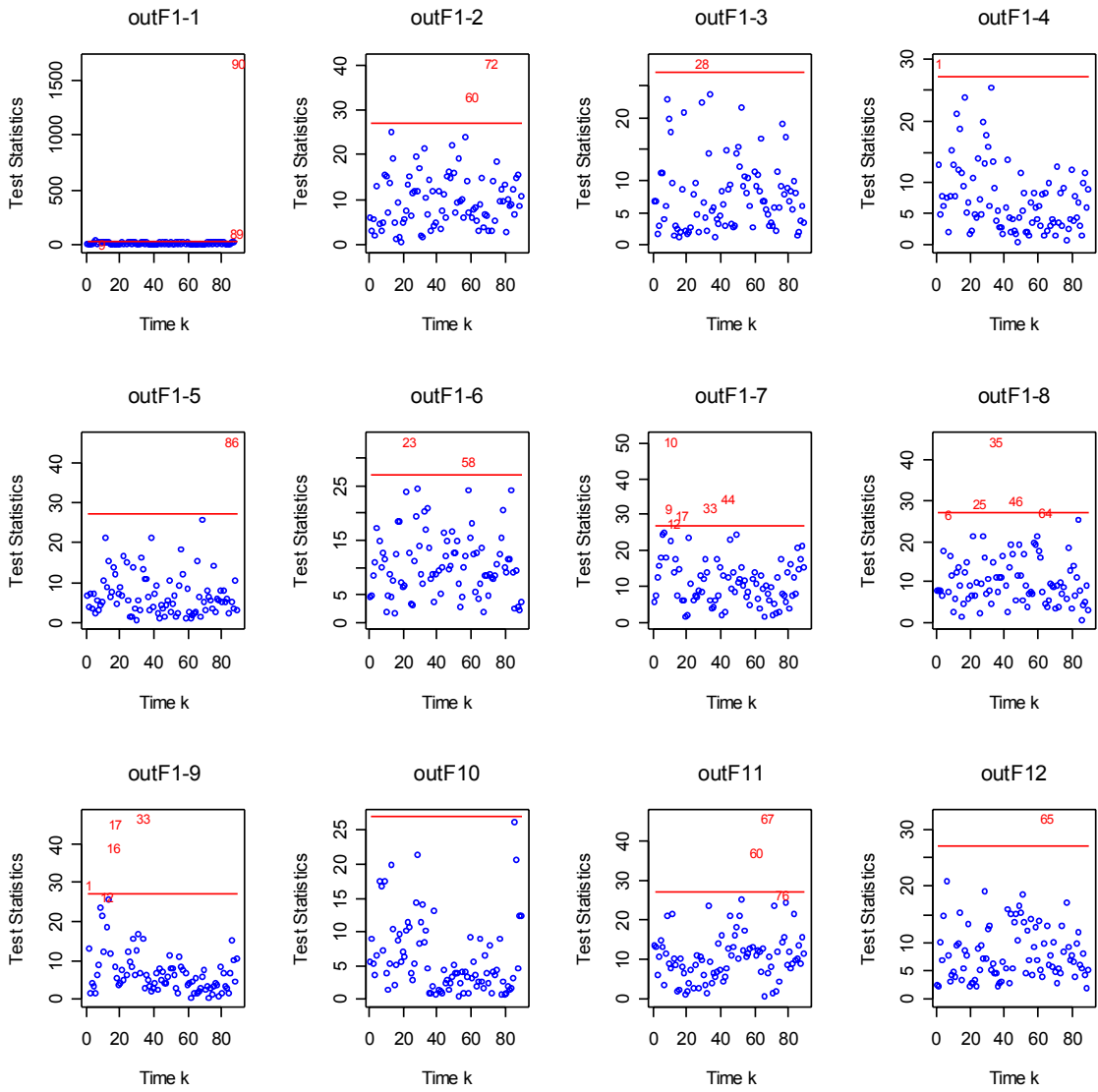Figure 5.5 Result of PCuR Test on Local Pattern Under Normal Condition

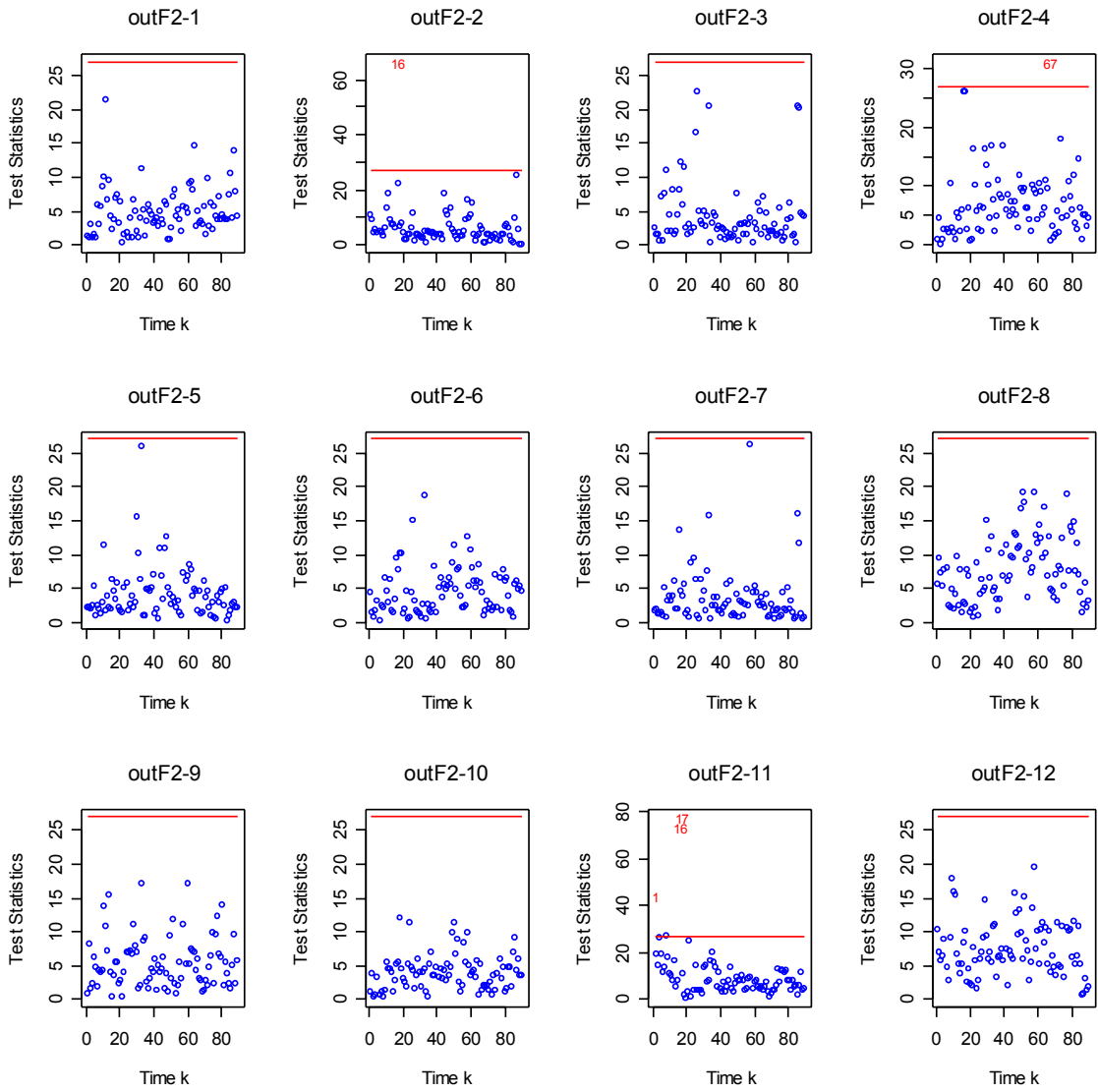Figure 5.6 Result of PCuR Test on Local Pattern Under Faulty Condition One

Figure 5.7 Result of PCuR Test on Local Pattern Under Faulty Condition Two

Chapter 6. Simulation Study

6.1 Simulation Data

We conducted the simulation study to verify our approach applicability on further applications and different types of process variables. We simulated a bivariate process with three process conditions: normal condition, a faulty condition due to two problematic segments in the cycle (faulty condition one), and a faulty condition due to global shift in the signal (faulty condition two). Thirty samples were simulated for every process variable with a cycle length equals to 90 knots. The signal functions under normal condition are given in the following equations:

$$Y_K = [1 + \cos(x), 1 + \sin(x)]^T + \varepsilon_K \tag{6.1}$$

where, Y_K is channel denotation at time k, k=1,… ,90. Noise ε_K is generated from normal distribution. The covariance structure of the noise is $\sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Different values of variance σ^2 were used to verify the approach under different variability structures in signals. The approach showed capability of faulty condition detection and discrimination using different combinations of covariance structures.

Faulty condition one was generated by adding two deviations to shift every in-control sample for both variables in two problematic time intervals. The length of the problematic segment is 10 knots.

The whole simulated signal was shifted to generate channels for faulty condition two. Figure 6.1 presents a sample of the simulated channels with their principal curve under normal condition.
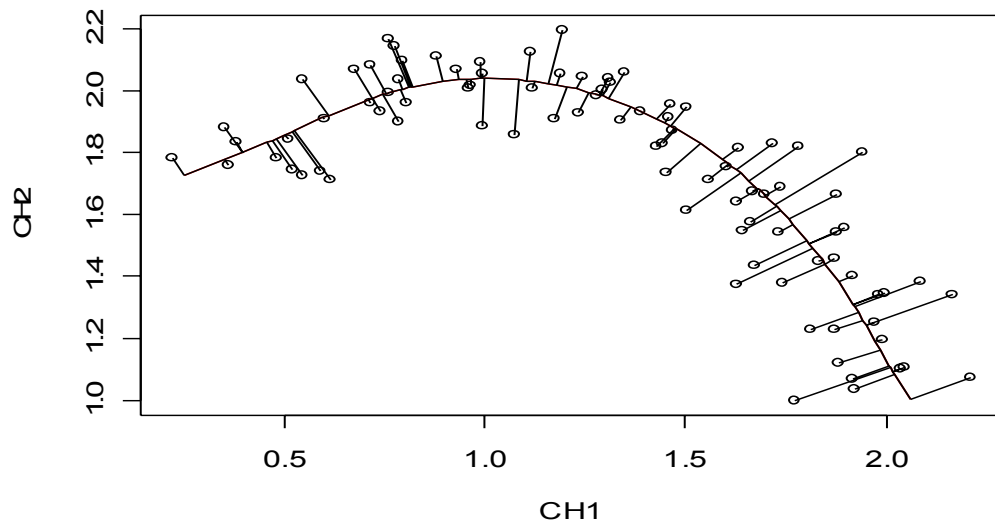

Figure 6.1 Simulated Channels and their Corresponding Principal Curve

6.2 Simulation Result

Global and local patterns were extracted by applying weighted spline smoothing on in-control data sets. Weights were determined using local moving average of generalized residuals. Two iterations were generated till convergence occurred. Since the two channels present two different variables, two average weight vectors were produced. The weight vectors were used in fitting the two channels of future coming samples under different process conditions. Figure 6.2 shows global and local variations extractions for one channel of the simulated data.
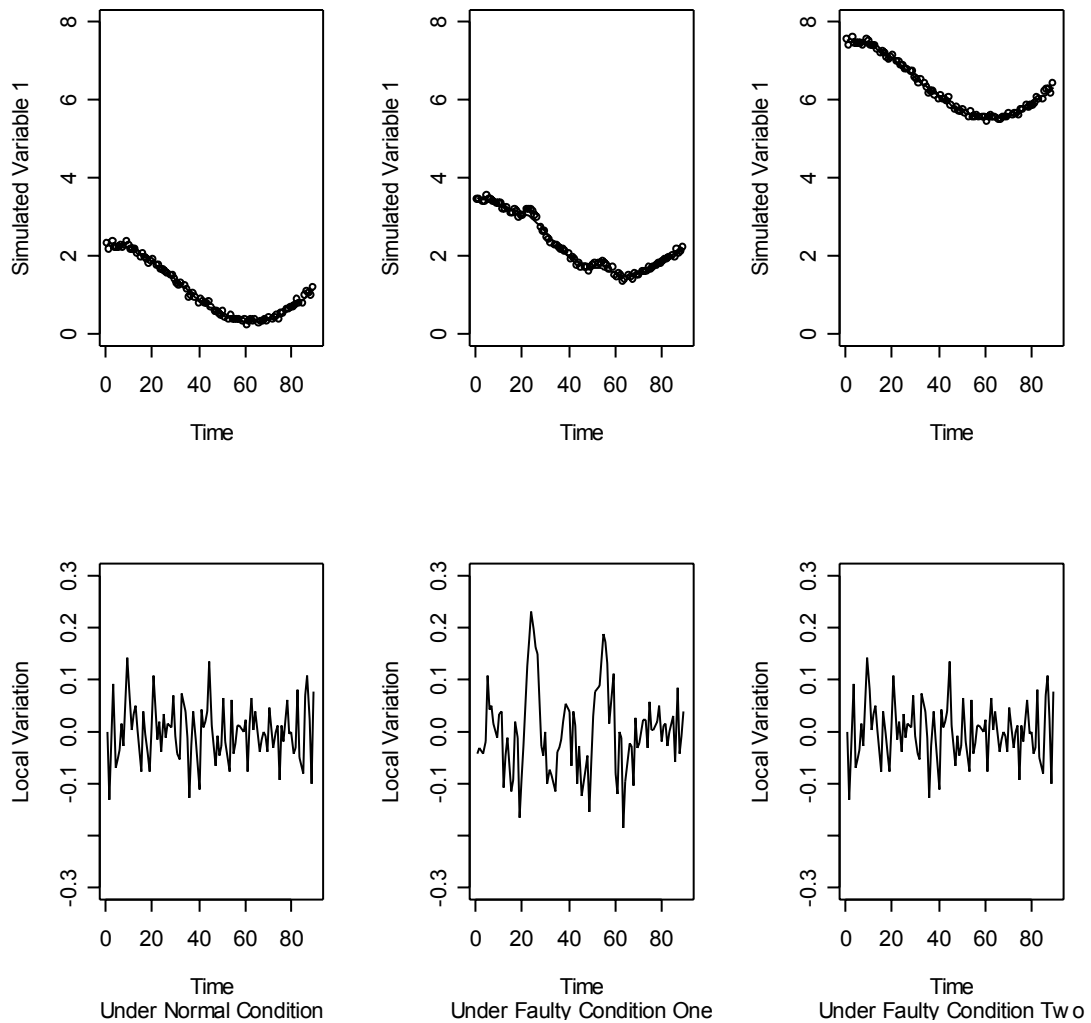
Figure 6.2 Example of Global and Local Variation Extraction-Simulated Data

The local variation under normal condition represents white noise. However, in the problematic segments, under faulty condition one, the variation magnitude is relatively higher than elsewhere in the simulated cycle. This indicates that faulty condition one is due to two problematic segments in the time intervals [20:30], and [50:60]. On the other hand, local variation under faulty condition two is similar to the

49

variation under normal condition, which shows that the cause of faulty condition two is a

global shift in the signal.
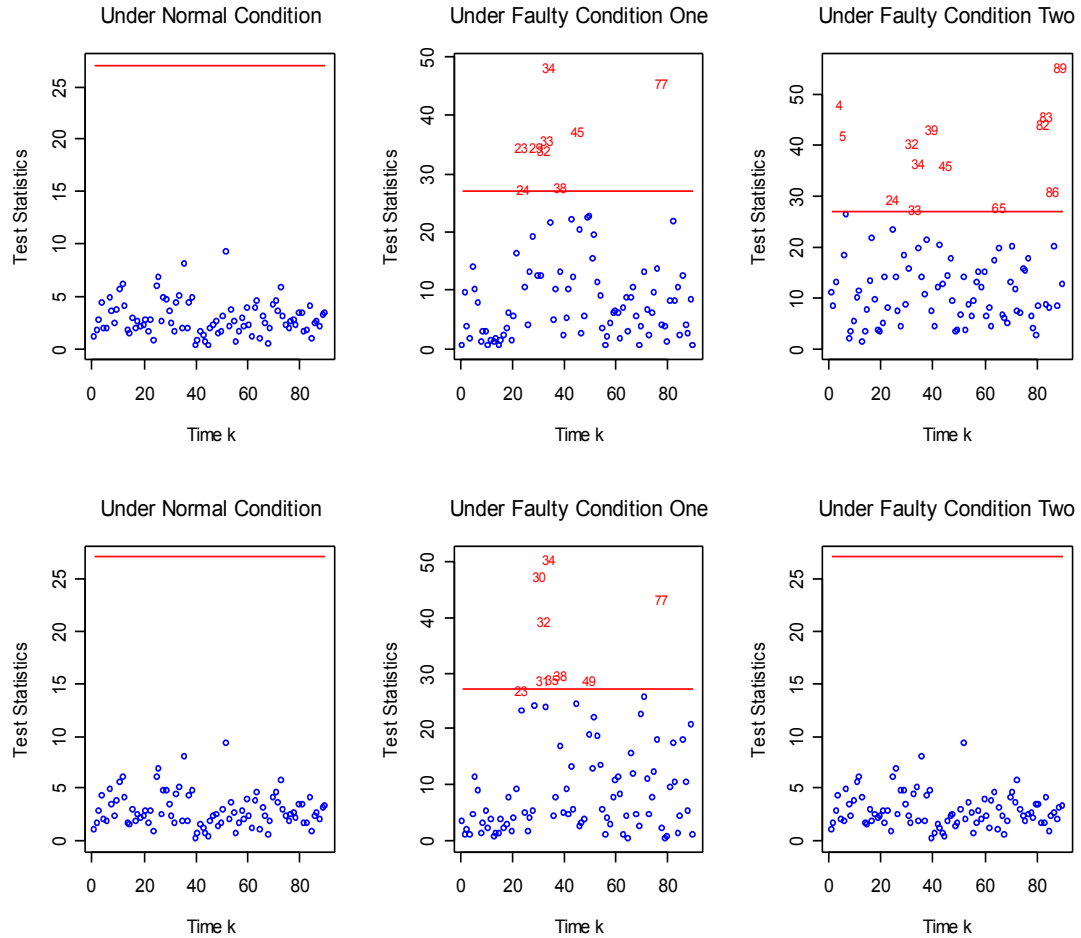
Figure 6.3 Sample Results of Applying PCuR on Simulated Data
First row is global data results; Second row presents results of local variations

For the purpose of process change detection, we used global variation as input in

the PCuR test. As a result, all samples under normal condition remains in-control,

however, all samples for faulty condition one and two are out of control. On the other

hand, using local variation as an input in the PCuR test, we could discriminate among

process faulty conditions. All training data sets under normal condition remains in

50

control, and all samples of faulty condition one were out of control, whereas, all the samples of faulty condition two remained in control. See Figure 6.1.

6.3 Summary

In this chapter, a simulation study was conducted to verify the developed approach. Three data sets were simulated representing bivariate data under three different process conditions: normal condition; faulty condition one, which is due to two local variations in the simulated cycles and; faulty condition two, which appears as global shift in the whole simulated cycles. Results were satisfactory. Global data appeared out of control for the faulty conditions and in control for the training data using the PCuR test. However, local variation data showed out of control status for faulty condition one, and remained in control for the in-control data sets and faulty condition two.

Chapter 7. Conclusion and Future Research Suggestions

This chapter presents the conclusion of the work that has been done in this research. A summary of the main contributions is also provided. Finally, directions for future research are also suggested in this chapter.

7.1 Conclusion

This research presents a monitoring scheme that is able to handle and analyze multichannel functional data by decomposing data into global and local variations. In our methodology, we addressed weighted spline smoothing technique to generate global patterns for the purpose of detecting global shifts and variations that causes faulty conditions. Weights were determined using local moving average of the generalized residuals for diagnostic procedures. Local variations were obtained by subtracting global variations from original data. Using local variations, as a key feature to predict samples with local variations, we could discriminate among different faulty conditions. The data extracted from both global and local approaches were tested using PCuR test. Our example and results were illustrated using forging data sets from the industry. Furthermore, we simulated data to verify the proposed approach and results were satisfactory since no false alarms were observed. Therefore, our main contributions and achievements in this research can be summarized as follows:

- MFD interpretation by decomposition MFD into global and local components.

- Process change detection using the extracted global variations from MFD by analyzing the whole cycle in a given industrial process.

- Discriminating among different process conditions using local variation as input in the PCuR test.

7.2 Future Research Suggestions

A further extension of this study can be focused on determining which process variable is significantly contributing in causing faulty conditions to occur. This work can start by decomposing the Hotelling T^2 statistic into independent components for each variable as in [51]. However, correlation among time should be considered in MFD.

Another desirable research direction is to estimate a confidence region for MFD or curve data using multiple samples. This study was accomplished for one sample depending on the finite dimensional Bayesian formulation for the curve estimation [38]. Finding confidence region for multi sample data provides solid structure for detecting faulty process conditions with MFD.

Last but not least research suggestion is to formulate an automatic approach to determine the amount of smoothness for the purpose of process change detection. This can be achieved by determining an optimal value of the effective degree of freedom by which local variations can be revealed. The optimal value can be controlled by assigning an objective function that maximizes the ratio of local variation in the problematic

segment to the variation at the rest of a process cycle. Other spline smoothing parameters can be determined using the variation and the smoother matrix as illustrated in [38].

References

[1]   J. Kim, Q. Huang, J. Shi, 2005. Online Multi-Channel Forging Tonnage Monitoring and Fault Pattern Classification using Principal Curve, ASME Transactions, Journal of Manufacturing Science and Engineering, in press.

[2]   Q. Huang, 2005, Principal Curve Regression and Analysis of Multiple Curve Data. Submitted to IIE Transactions.

[3]   M. Jeong, J. Lu, 2003, Wavelet-Based SPC Procedure for Complicated Functional. Conference paper, INFORMS Annual Meeting.

[4]   C. A. Lowry, and D. C. Montgomery, 1995, A Review of Multivariate Control Charts, IEE Transactions, 27(6), pp. 800-811.

[5]   F. B. Alt, 1985, Multivariate Quality Control, S. Kontz & N. L. Johnson (Eds), Encyclopedia of Statistical Sciences, New York: John Wiley & Sons,6, pp. 110-122.

[6]   J. A. Jr. Alloway, M. Raghavachari, 1990, Multivariate Control Charts Based on trimmed Means, ASQC Quality Congress Transactions, San Francisco.

[7]   J. William, W. Woodall, J. Birch, 2003, Phase I Analysis of Nonlinear Product and Process Quality Profiles, submitted to journal of quality and technology.

[8]   D. M. Hawkiins, 1991, Multivariate Quality Control Based on Regression-Adjusted Variables, Technometrics, 33(1), pp. 61-75.

[9]   D. M. Hawkiins, 1993, Regression Adjustment for Variables in Multivariate Quality Control , Journal of Quality Technology, 25(3), pp170-181.

[10]   J. J. Jr. Pignatiello, 1985, Development of Multivariate CUSUM chart, Computers in Engineering, Proceedings of the international computers in engineering conference, 2, pp. 427-432.

[11]   C. Fuchs, and Y. Benjamini, 1994, Multivariate Profile Charts for Statistical Process Control, Technometrics,36(2), pp. 182-195.

[12]   J. E. Jackson, R. H. Morris, 1957, An Application of Multivariate Quality Control to Photographic Processing, American Statistical Association Journal (June), 2, pp. 186-199.

[13]   R. Sparks, A Adolphson, and A. Phatak, 1997, Multivariate Process Monitoring Using The Dynamic Biplot, International Statistical Review / Revue Internationale de Statistique ,65(3), pp. 325-349.

[14]   D. W. Apley, J. Shi, 2001, A Factor-Analysis Method for Diagnosing Variability in Multivariate Manufacturing Processes, Technometrics, 43(1), pp. 84-95.

[15]   Y. Ding, P. Kim, D. Ceglarek, J. Jin, 2003, Optimal Sensor Distribution for Variation Diagnosis in Multistation Assembly Processes, IEEE Transaction on Robotics and Automation,19(4), pp. 543-556.

[16]   M. Rodríguez-Méndez, A. Arrieta, V. Parra, A.Bernal, A. Vegas, S. Villanueva, R. Gutiérrez-Osuna, J. de Saja, 2004, Fusion of Three Sensory Modalities for the Multimodal Characterization of Red Wines, IEEE, Sensors Journal, 4(3), pp 348-355.

[17]   R. Luo, C. Yih, K. Su, 2002, Multisensor Fusion and Integration: Approaches, Applications, and Future Research Directions, IEEE, Sensors Journal, 2(2), pp. 107-119.

[18]   S. Suranthiran and S. Jayasuriya, 2004, Optimal Fusion of Multiple Nonlinear Sensor Data, IEEE, Sensors Journal, 4(5), pp. 651-663.

[19]   M. Mahmoud, W. Woodall ,P. Paker, D. Hawkins, 2004. A Change Point Method for linear Profile Data, Submitted to Journal of Quality and Technology.

[20]   J. Jin,and J. Shi, 2001, Automatic Feature Extraction of Waveform Signals for In-process Diagnostic Performance Improvement, Journal of Intelligent Manufacturing, 12, pp267-268.

[21]   J. Jin, and J. Shi 2000, Diagnostic Feature Extraction from Stamping Tonnage Signals Based on Design of Experiment, ASME Transactions, Journal of Manufacturing Science and Engineering, 122(2), pp. 360-369

[22]   J. Jin, and J. Shi, 1999, Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets", Technometrics, 41, pp. 327-339.

[23]   C. K. H. Koh, J. Shi, and J. Black, 1996, Tonnage Signature Attribute Analysis for Stamping Process, NAMRI/SME Transactions, 23, pp193-198.

[24]   C. K. H. Koh, J. Shi, W. Williams, J. Ni, 1999, Multiple Fault Detection and Isolation Using the Haar Transform - Part 1: Theory, ASME. Transactions, Journal of Manufacturing Science and Engineering, 121(2), pp290-294.

[25]   J. Jin, 2004, Individual Station Monitoring Using Press Tonnage Sensors for Multiple Operation Stamping Processes, ASME Transactions, Journal of Manufacturing Science and Engineering, 126(1), pp. 83-90.

[26]   S. Zhou, N. Jin, and J. Jin, 2004, Cycle-based Signal Monitoring Using a Directionally Variant Control Chart System, accepted by IIE Transactions on Quality and Reliability.

[27]   T. Hastie, and, W. Stuetzle (1989). Principal Curves. Journal of the American Statistical Association, 84, pp.502-516.

[28]   T. Duchamp, W Stuezle, 1996, Extremal Properties of Principal Curves in the Plane, Annals of Statistics, 24, pp. 1511-1520.

[29]   B. Kégl, A. Krzyzak, T. Linder, and K. Zeger, 2000, Learning and Design of Principal Curves, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, pp. 281-297.

[30]   J.D. Banfield, A.E. Raftery, 1992, Ice Floe Identification in Satellite Images Using Mathematical Morphology and Clustering about Principal Curves, Journal of the American Statistical Association, 87, pp. 7-16.

[31]   M. LeBlance, R.J. Tibshirani, 1994, Adaptive Principal Surfaces, Journal of American Statistical Association, 89, pp. 53-64.

[32]   D. Standford, A.E. Raftery, 1997, Principal Curve Clustering with Noise, Technical Report 317, Department of Statistics, University of Washington.

[33]   K. Chang J. Ghosh, 1998, Principal curves for nonlinear feature extraction and classification, Applications of Artificial Neural Networks in Image Processing III, SPIE Photonics West '98 Electronic Image Conference, SPIE ,3307 pp. 120-129.

[34] O. l. Mangasarian, L. L. Shumakher, 1969, Splines Via Optimal Control, Approximations With Special Emphasis On Spline Functions, (Ed. I. J, Schoenberg , New York, Academic Press, pp. 119-156.

[35] E. Wegman, I Wright, 1982, Splines in Statistics, Journal of American Statistical Association, 78(382), pp.351-365.

[36] P. Eilers and B. Marx, 1996, Flexible Smoothing with B-Spline and Penalties, Statistical Science., 11(2), pp. 112-114.

[37] T. Hastie, R. Tisbshirani, J. Friedman, The Elements of Statistical Learning, Data Mining, Inference, and Prediction.   New York: Springer-Verlag, 2001, ch. 5.

[38] B. W. Silverman., 1985. Some Aspects of the Spline Smoothing Approach to Non-Parametric Regression Curve Fitting, Journal of Royal Statistical Society, 47(1), pp. 1-52.

[39] G. Wahba, S. Wold, 1975a, A Completely Automatic French Curve: Fitting Spline Functions by Cross-Validation", Communications in Statistics", 4, pp. 1-17.

[40] G. Wahba, S. Wold, 1975b, Periodic Splines for Spectral Density Estimation: The Use of Cross-Validation for Determining the Degree of Smoothing,, Communications in Statistics, 4, pp. 125-141.

[41] G. Wahba, 1976, A survey of Some Smoothing Problems and the Methods of Generalized Cross-Validation for Solving Them, Applications of Statistics (ed. P.R. Krishnaiah), 1, pp. 507-534.

[42] G. Wahba, 1979, Automatic Smoothing of the Log Periodgram, Journal of the American Statisticl Association, 75, pp. 122-132.

[43] G. Wahba, 1980, Spline Bases, Regularization, and Gerealized Cross-Validation for Solving Approximation Problems with Large Quantities of Noisy Data, Proceedings of the International Conference on Approximation Theory in Honor of George Lornez, (ed. Ward Cheney, Academic Press.

[44] R.D. Cook, and S. Weisberg ,1982, Residuals and Influence in Regression, London: Chapman and Hall.

[45] P. Craven, and G. Wahba, 1979, Smoothing Noisy Data with Spline Functions, Numer. Math., 31, pp. 377-403.

[46] J, Fleisher, 1979, Spline Smoothing Routines, Reference Manual for the 1110, Academic Computer Center, The University of Wisconsin, Madison.

[47] P. Merz, 1978, Spline Smoothing by Generalized Cross-Validation: a Technique for Data Smoothing, Chevron Research Company, Richmond, CA.

[48] L. Paihua-Montes, 1979, Quelques Methodes Numeriques Pour le Calcul de Fanctions Splines a Une et Plusier Variable, Thesis, Universite Scientifique et Medicale de Grenoble.

[49] B. W. Silverman, 1984, A Fast and Efficient Cross Validation Method for Smoothing Parameter Choice in Spline Regression, Journal of American Statistics Association, 79, pp. 584-589.

[50] R. Johnson, and D. Wichern, 2002, Applied Multivariate Statistical Analysis, (5th ed.), Prentice Hall, pp. 383-398.

[51] R. Mason, N. Tracy, and J. Young, 1995, Decomposition of T^2 for Multivariate control chart interpretation. Journal of Quality Technology, Vol. 27(2), pp. 99-108.