國 立 交 通 大 學

統計學研究所

碩 士 論 文

用無母數迴歸方法監控隨機效應曲線型品質特性之製程

Monitoring Random Effect Profiles by Nonparametric Regression

研 究 生 ： 蔡明曄
指導教授 ： 洪志真 博士

中 華 民 國 九 十 五 年 六 月

用無母數迴歸方法監控隨機效應曲線型品質特性之製程

Monitoring Random Effect Profiles by Nonparametric Regression

研 究 生：蔡明曄　　Student： Ming-Ye Tsai

指導教授：洪志真　　Advisor： Dr. Jyh-Jen Horng Shiau

國 立 交 通 大 學

統計學研究所

碩 士 論 文

A Thesis
Submitted to Institute of Statistics
College of Science
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master
in
Statistics
June 2006
Hsinchu, Taiwan, Republic of China

中華民國九十五年六月

用無母數迴歸方法監控隨機效應曲線型品質特性之製程

研究生： 蔡明曄　　指導教授： 洪志真 博士

國立交通大學統計學研究所

摘要

監控製程與產品之曲線型資料在統計品質管制中是一個非常熱門且有前景的研究領域。我們的研究將以對具有隨機效應之非線性曲線型資料的監控方法為主。

對隨機效應模式，我們利用主成分分析來分析曲線型資料的共變異結構，並利用每個曲線型資料所得之主成分分量來做監控。在第二階段中，因為個別的主成分分量很難區別開變動的趨勢，所以我們建議利用 combined chart，即合併主成分分量的方法來做監控。

在第一階段中，因為主成分分量的相依特性，我們採用 T^2 圖來檢測穩定性。我們利用大量模擬的方法來說明，當 outliers 出現的形式是短暫時，T_1^2 表現的較 T_2^2 佳。

而且，我們也發現用來建造控制圖的主成分個數會影響偵測力。我們採用交互驗證(cross-validation)的方法，來選擇主成分的個數。

Monitoring Random Effect Profiles by Nonparametric Regression

Student： Ming-Ye Tsai　Advisor： Dr. Jyh-Jen Horng Shiau

Institute of Statistic
National Chiao Tung University

ABSTRACT

The monitoring of process and product profiles is a very popular and promising area of research in statistical process control. This study is aimed at the monitoring scheme for nonlinear profiles with random effects.

For random effect models, we use the technique of principal component analysis to analyze the covariance structure of the profiles and use principal component scores of each profile to perform monitoring. In Phase II, since it is difficult for each principal component score to have identified direction of shifts, we recommend using a combined chart scheme that combines the principal component scores to perform monitoring.

In the historical analysis of Phase I data, due to the dependency of principal component scores, we adopt the T^2 chart to check for stability. We show by simulation that the sample-covariance-based T_1^2 performs better than the successive-difference-based T_2^2 for temporal shifts.

Also, the number of principal component scores used in constructing control charts has an effect on the detecting power. We adopt the cross-validation to choose the number of principal component scores.

誌　謝

在統研所短短的兩年裡，我得到了很多，除了書本上的知識，也交到了很多好朋友。這篇論文的完成，最要感謝我的指導老師 洪志真教授認真親切的指導跟辛勤的批閱，還有博士班碩慧學姊，無論是在學業上或生活上的指教及協助，使我能夠順利地完成論文。最後還要感謝我的研究所同學俊凱、冠達、泓毅、清峰、義家、小馬、瑩琪、火哥跟勢耀帶給我充滿歡笑的回憶，在學業上的切磋討論也使我獲益甚多。

此外還要感謝家人給我的支持與鼓勵，使我可以無後顧之憂的專心學習，最後僅將此論文獻給我最愛的家人、朋友們。

蔡 明 曄　　謹誌于

國立交通大學統計研究所

中華民國九十五年六月

Contents

List of Tables

List of Figures

1 Introduction

Statistical process control (SPC) has been widely applied in many areas, especially in industries. Classical methods for SPC assume that the quality of the product or process can be measured by one or multiple quality characteristics. However in many situations, the response of interest is not a single variable but rather a function of some independent variables. This functional response is called a profile. An example of profiles is the dissolving process of aspartame, an artificial sweetener, which is characterized by the amount of that dissolves per liter of water at different levels of temperature; see Kang and Albin (2000). For illustration, Figure 1 shows the plot of four hypothetical aspartame profiles. Another example mentioned in Kang and Albin (2000) is a semiconductor manufacturing problem that occurs during the etching step. In this example, the calibration of a mass flow controller in which the performance of the process is characterized by a linear function; see Mestek et al. (1994). There is also an interesting example introduced by Walker and Wright (2002), called vertical density profiles (VDP). The density is measured using a profilometer that uses a laser device to take measurements at fixed depths across the thickness of the engineered wood board. As an illustration, Figure 2 shows the VDP data from Walker and Wright (2002). The data set is available at http://bus.utk.edu/stat/walker/VDP/Allstack.TXT. The authors proposed a method using additive models to assess the sources of variation in the density profiles of particleboards.

Kang and Albin (2000) studied the problem of linear profile monitoring. Let the output of a process be a random variable Y which depends linearly on an independent variable X. They modeled the profiles by the simple linear regression model, $Y = A_0 + A_1 X + \epsilon,\ X_l \leq X \leq X_h$, where A_0 and A_1 are parameters and X_l and X_h define the range of X in which the process is characterized by a linear profile. The random variable ϵ are independent

and normally distributed with mean 0 and variance σ^2. They proposed two approaches to monitoring profile data: (i) monitoring the parameters, the intercept (A_0) and slope (A_1), simultaneously with the Hotelling T^2 chart and (ii) using the regular EWMA and R charts for profile monitoring by treating the residuals of the sample line to the in-control reference line as a rational subgroup. Kim et al. (2003) proposed another approach called $EWMA_3$, which they showed was comparable to the approach of Kang and Albin (2000). They centered the X-values to make the least squares estimators of the Y-intercept and slope independent of each other. So that the intercept and slope can be monitored separately. They constructed three EWMA charts to monitor intercept, slope, and the variation, separately.

Shiau and Weng (2004) extended the above profile monitoring scheme to a scheme for more general profiles. No assumption are made for the form of the profiles except the smoothness. The nonparametric regression model considered is $Y = g(x) + \epsilon$, where $g(x)$ is a smooth function and ϵ is the random error as before. Spline regression was adopted as the curve fitting/smoothing technique. They proposed an EWMA chart for detecting the mean shifts, an R chart for variation changes, and an exponentially weighted moving standard deviation (EWMSD) chart for the variation increase.

Note that the models described so far are a deterministic line/curve plus random noises. In this paper, based on Shiau and Weng (2004), we extend the fixed effect model to a random effect model in order to provide more variability that we often observe in many profile data, e.g., the aspartame example and VDP example. We use the function $Y = I + Me^{N(x-1)^2} + \epsilon$ given in Shiau and Weng (2004) as our illustrative example, where I, M, N are considered to be random variables. Like Shiau and Weng (2004), we use the B-spline regression technique to smooth profiles. With the random effect model, we put emphasis on the covariance structure. To analyze the

covariance matrix, it is natural to consider the technique of the principal component analysis (PCA).

The PCA is very useful in summarizing and interpreting a set of profile data with the same equally spaced values of the independent variable X for each profile. Castro et al. (1986) showed that the principal modes of variation consist of eigenfunctions of the process covariance function and gave procedures for estimating these eigenfunctions from a set of observed curves. Rice and Silverman (1991) proposed a method for estimating the mean function nonparametrically under the assumption that it is smooth. They also suggested a variant of the usual cross-validation for choosing the degree of smoothing to be employed. And in the estimation of covariance structure, they primarily concerned with the first few eigenfunctions that were smoothed and the eigenvalues that decayed rapidly. Jones and Rice (1992) suggested that a simple PCA be used to identify important modes of variation among the curves and that the principal component scores be used to identify particular curves which clearly demonstrate the form and extent of a particular mode of variation.

In this paper, we construct our monitoring schemes by utilizing the eigenvectors and principal component scores obtained from PCA. If these scores corresponds to separate modes of variation, then it is natural to monitor each principal component score for what it represents. If not, we suggest a combined-chart scheme that combines the charts of these principal component scores for profile monitoring. Finally, a simulation study is conducted to evaluate the performance of each principal-component-score chart and the combined-chart in terms of the average run length (ARL). The simulation study demonstrates that the combined chart scheme is comparable with the best principal-component-score chart.

In the study of Phase I monitoring, we compare two commonly used

T^2 statistics, T_1^2 and T_2^2, in which different covariance estimators are used. Also, it is found from the study that the number of principal component scores used in constructing T^2 statistics affects the power of detecting out-of-control profiles significantly. In this study, we use the cross-validation method, one of the most popular procedures for model selection, to choose the number of principal component scores.

Section 2 reviews the methods proposed by Shiau and Weng (2004) and Shiau and Lin (1999) in which a set of degradation paths was modeled by a stochastic process and PCA was employed to analyze the covariance structure. Section 3 describes some methodologies used in this paper, including the B-spline regression technique, principal component analysis, procedure of cross-validation, and the proposed control schemes in details. Section 4 presents the simulation results of Phase II monitoring with some discussions along with some recommendations for Phase I monitoring. Section 5 presents a case study using the VDP data from Walker and Wright (2002). Finally, Section 6 concludes the paper with a brief summary and some remarks.

2 Literature Review

Shiau and Weng (2004) proposed a profile monitoring scheme for profiles of flexible shape. Unlike the combined EWMA/R chart proposed by Kang and Albin (2000), they recommended using the EWMA to monitor the residual average for detecting mean shifts and the R chart to monitor the residual range for detecting standard deviation shifts of the residuals individually in order to have better detecting power. Also, they proposed an EWMSD control chart to detect shifts in process variation.

In Shiau and Weng (2004), an exponential profile of the form $Y = I + Me^{-N(x-1)^2} + \epsilon$ is used to mimic an aspartame profile. Shiau and Weng (2004)

showed by simulation studies that, for I, M, and N shifts in the exponential profile, the EWMA chart performed well in detecting shifts. And for standard deviation shifts, the EWMSD chart is a good choice while one can also use the R chart. Also, they showed that, for the linear profile example given in Kang and Albin (2000), their methods were slightly effective than the methods of Kang and Albin (2000) and Kim et al. (2003).

Shiau and Lin (1999) proposed a nonparametric regression accelerated life-stress (NPRALS) model to analyze the accelerated degradation tests (ADT) data. In order to model a set of degradation paths by a stochastic process, they made the following assumptions:

1. The sample degradation path of each experimental subject is a realization of an underlying stochastic process $\{X(t), t \in T\}$, where T is an interval.

2. The model for $X(t)$ is

$$X(t) = \mu(t) + \omega(t) + \epsilon(t), \tag{1}$$

   where $\mu(t) \equiv E[X(t)]$, $\omega(t) \equiv$ a stochastic process with mean 0 and covariance function $r(s,t) \equiv \text{Cov}[\omega(t), \omega(s)]$, and $\epsilon(t) \equiv$ uncorrelated error terms with $E[\epsilon(t)]=0$ and $Var[\epsilon(t)]=\sigma^2$.

3. The acceleration stress affects only the degradation rate, not the shape of the degradation curve.

With these assumptions, they used $\mu(t)$ to denote the mean curve under the normal condition (un-accelerated) and $\mu(a \cdot t)$ to denote the accelerated mean curve where a is the acceleration factor and usually $a > 1$.

In their model, there are m stress levels in the experiments. And for stress level i, there are n_i experimental units. Also, $t_1, \ldots, t_p$ are the measured

points of the product characteristic for each experimental unit. By using the functional principal component technique, they model the ADT data by

$$X_{i,j,k} = \mu(a_i \cdot t_k) + \sum_{q=1}^{L} \epsilon_{i,j,q} \cdot \rho_{i,q}(t_k), \tag{2}$$

where $1 \leq i \leq m, 1 \leq j \leq n_i, 1 \leq k \leq p, 1 \leq q \leq L$, and L is a positive integer no more then p. So the random component $\omega(t)+\epsilon(t)$ of equation (1) is then replaced by a random combination of functional principal components $\{\rho_{i,q}(\cdot), q = 1, \ldots, L\}$. And they provided an iterative algorithm to estimating $\mu(\cdot)$ and $a_1, \ldots, a_m$.

Also, they investigated the covariance structure of $X(\cdot)$. First, for $i = 1, \ldots, m$, compute

$$(V_i)_{s,r} = \frac{1}{n_i - 1} \cdot \sum_{j=1}^{n_i} [X_{i,j,s} - \hat{\mu}_i(t_s)] \cdot [X_{i,j,r} - \hat{\mu}_i(t_r)]$$

for $1 \leq s, r \leq p$, where $\hat{\mu}(\cdot)$ is the estimate of $\mu(\cdot)$. Let $\hat{\boldsymbol{\rho}}_{i,q} \equiv (\hat{\rho}_{i,q}(t_1), \ldots, \hat{\rho}_{i,q}(t_p))'$ be the eigenvector corresponding to the q-th largest eigenvalue $\lambda_{i,q}$ of V_i, where $i = 1, \ldots, m$ and $q = 1, \ldots, p$. Then estimate:

- $\sigma^2_{i,q}$ by $\lambda_{i,q}$.
- $\rho_{i,q}(\cdot)$ by smoothing $\{\hat{\rho}_{i,q}(t_k), k = 1, \ldots, p\}$ on $\{t_k, k = 1, \ldots, p\}$.

At the end of the paper, they presented a real data analysis with their NPRALS model. Also, in order to show how well their procedure does in recovering the truth, they conducted a simulation study. They simulated the degradation data curves by using equation (2). Treat $\hat{\mu}(\cdot)$, $\lambda_{i,q}$, and the smoothed $\{\hat{\rho}_{i,q}(t_k), k = 1, \ldots, p\}$ obtained from the previous real data analysis as the true values of $\mu(\cdot)$, $\sigma^2_{i,q}$, and $\rho_{i,q}(\cdot)$, respectively. The following are the steps of the data generation.

- Compute:

$$(V_i)_{s,r} = \frac{1}{n_i} \cdot \sum_{j=1}^{n_i} [X_{i,j,s} - \mu_i(t_s)] \cdot [X_{i,j,r} - \mu_i(t_r)]$$
$$\text{for } 1 \le i \le m, 1 \le j \le n_i, 1 \le k, s, r \le p.$$

- Generate $\epsilon_{i,j,q}$ from $N(0, \sigma_{i,q}^2), 1 \le j \le n_i, 1 \le q \le L$.

- Generate:

$$\tilde{X}_{i,j,k} = \mu_i(t_k) + \sum_{q=1}^{L} \epsilon_{i,j,q} \cdot \rho_{i,q}(t_k)$$
$$\text{for } 1 \le j \le n_i, 1 \le k \le p.$$

By simulation, they showed that their NPRALS model performed very well in recovering the features of the original degradation data curves.

3 Methodologies

3.1 Procedures

In this paper, as an illustrative example, we use the model

$$Y = I + Me^{N(x-1)^2} + \epsilon$$

where $I \sim N(1, 0.2^2)$, $M \sim N(15, 1)$, $N \sim N(-1.5, 0.3^2)$, $\epsilon \sim N(0, 0.3^2)$, and x=0.64, 0.8, ..., 3.52 to simulate the profile data of the aspartame example. We generate 200 profiles to serve as our historical data for Phase I analysis.

We apply the B-spline smoothing technique to filter out the noise. For the model $X(t) = \mu(t) + W(t) + \epsilon(t)$, this is to filter out $\epsilon(t)$ so that the actual signals can be better extracted from the data. These extracted signals will be smoother and can explain the variation among profiles better. As

the variance of $\epsilon(t)$ tends larger, the advantage of smoothing tends more profound.

In Phase I monitoring, first we select the number of principal component (PC) scores to be used by cross-validation and then use the T^2 statistics of the score vectors to detect the outliers. The details of these methods are given in Subsections 3.4 and 3.5. This procedure will continue until, all the remaining profiles are within the control limit. Then we can apply the technique of PCA to these "in-control" profiles and the principal components obtained will be used for the Phase II monitoring.

We summarize the whole simulation procedure as follows:

- Generate historical profile data at set points x=0.64, 0.8, ..., 3.52 by

$$y = I + Me^{N(x-1)^2} + \epsilon,$$

  where $I \sim N(1, 0.2^2)$, $M \sim N(15, 1)$, $N \sim N(-1.5, 0.3^2)$, and $\epsilon \sim N(0, 0.3^2)$.

- Smooth each of the discrete profile data with the B-spline smoothing technique.

- Select the number of effective principal components by the cross-validation method.

- Phase I: use T^2 statistics to remove outliers until the remaining profiles are all in control. Then, use PCA to obtain important features, by which the in-control process is characterized.

- Phase II: We compute PC scores for each of the incoming profiles. And use the independent property of the scores to check the stability of the incoming profile by monitoring each score separately or by the combined-chart scheme.

Details of our simulation results will be given in Section 4.

3.2 B-Splines

One basic objective of the data analysis is to identify the signals in a set of data. But this gets complicated with the existence of random noise in the data. That is, the goal is to filter out the random noise from the data in order to get the signal. In order to extend nonlinear profiles of a fixed form to smooth profiles of any shapes, a smoothing technique is needed for de-noising sample profiles. The idea of smoothing is to fit a flexible function whose final form is determined by the data and by the chosen level of smoothness of the curve. That is, let the data speak for themselves. One popular way is to fit noisy data by splines. Frequently, cubic splines−piecewise cubic polynomials with continuous second derivatives−are used for such approximation.

Consider the following nonparametric regression model:

$$y_i = g(x_i) + \epsilon_i, \quad i = 1, \ldots, p, \tag{3}$$

where $g(x)$ is a smooth regression curve and ϵ_i's are i.i.d. Gaussian variables with zero mean and common variance $\sigma^2 > 0$. In this paper, to estimate $g(x)$, we adopt the B-spline regression method for its popularity and simplicity. Simply put, the B-spline regression is just a multiple linear regression with B-spline bases.

Before we go on, we first give a brief description on B-spline bases. Details can be found in de Boor (1978). Let b denote the number of bases and k be the order of the B-splines. Note that the degree of the polynomials is $k-1$, hence each basis of order k is a $k-2$ continuously differentiable function. For example, $k = 4$ for cubic splines. Also let the interval $[u, v]$ be the domain of interest. For spline smoothing, knots are the points where the higher-order derivatives could be discontinuous. For b bases, we need $b+k$ knots, denoted by $t_1, \ldots, t_{b+k}$, such that $t_k = u$ and $t_{b+1} = v$. And let $B_{l,k}$ denotes the l-th B-spline basis of order k, where $l = 1, \ldots, b$. $B_{l,k}$ can be defined iteratively

by

$$B_{l,1}(t) = \begin{cases} 1 & \text{for } t_l \le t < t_{l+1}, \\ 0 & \text{for } t < t_l \text{ or } t \ge t_{l+1}. \end{cases}$$

$$\text{For } k \ge 2,\ B_{l,k}(t) = \frac{t - t_l}{t_{l+k-1} - t_l} B_{l,k-1}(t) + \frac{t_{l+k} - t}{t_{l+k} - t_{l+1}} B_{l+1,k-1}(t).$$

Note that $B_{l,k}$ is nonzero only on the interval (t_l, t_{l+k}). A B-spline of order k can then be constructed as

$$g(x) = \sum_{l=1}^{b} c_l B_{l,k}(x),\ \text{for } x \in [u, v],$$

where c_l's are the unknown B-spline coefficients to be estimated from data.

For the B-spline regression, the nonparametric regression model in equation (3) can then be replaced by the following linear model:

$$y_i = \sum_{l=1}^{b} c_l B_{l,k}(x_i) + \epsilon_i,\ i = 1, \ldots, p. \tag{4}$$

Given a set of data, $\{(x_i, y_i),\ i = 1, \ldots, p\}$, the spline regression method finds the best spline approximation via the following least squares problem:

$$\min_{\mathbf{c}} \sum_{i=1}^{p} \{y_i - \sum_{l=1}^{b} c_l B_{l,k}(x_i)\}^2, \tag{5}$$

where $\mathbf{c} = (c_1, \ldots, c_b)'$. Then the least squares estimator of $\mathbf{c}$ is

$$\hat{\mathbf{c}} = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{y}, \tag{6}$$

where $\mathbf{y} = (y_1, \ldots, y_p)'$, $\hat{\mathbf{c}} = (\hat{c}_1, \ldots, \hat{c}_b)'$, and $\mathbf{B}$ is the $p \times b$ design matrix with the (i, l)-th element $B_{l,k}(x_i)$, $l = 1, \ldots, b$, $i = 1, \ldots, p$. Under the model in equation (4), $\hat{\mathbf{c}}$ has a multivariate normal distribution with mean vector $\mathbf{c}$ and variance-covariance matrix $\Sigma = \sigma^2(B'B)^{-1}$.

Note that the number of bases b acts as the smoothing parameter in nonparametric regression and it is well known that the choice of the smoothing parameter is crucial in nonparametric regression estimation. Also, the

boundary effect inherited in most of the smoothing methods is another issue of concern.

3.3 Principal Component Analysis

The following materials on PCA are taken from Anderson (2003). PCA is a multivariate procedure that rotates the data such that maximum variabilities are projected onto the axes. Essentially, a set of correlated variables are transformed into a set of uncorrelated variables ordered by the amount of variation explained. Then these uncorrelated variables, called principal components, are linear combinations of the original variables, and the last of these variables can be removed with minimum loss of information contained in data. The main usage of PCA is to reduce the dimensionality of a data set while retaining as much information as possible.

Principal components have special properties or features in terms of variance. They turn out to be the characteristic vectors of the covariance matrix. Suppose the random vector $\mathbf{Y}$ of p-components has covariance matrix $\mathbf{\Sigma}$. Since we are only interested in the covariance matrix $\mathbf{\Sigma}$, without loss of generality, we may assume that the mean vector is $\mathbf{0}$. Let $\boldsymbol{\beta}$ be a p-component column vector such that $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$. The variance of $\boldsymbol{\beta}'\mathbf{Y}$ is

$$E(\boldsymbol{\beta}'\mathbf{Y})^2 = E(\boldsymbol{\beta}'\mathbf{Y}\mathbf{Y}'\boldsymbol{\beta}) = \boldsymbol{\beta}' E(\mathbf{Y}\mathbf{Y}')\boldsymbol{\beta} = \boldsymbol{\beta}'\mathbf{\Sigma}\boldsymbol{\beta} \,, \tag{7}$$

where E denotes the expectation. In order to determine $\boldsymbol{\beta}'\mathbf{Y}$ with the maximum variance subjected to $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$, let

$$\phi = \boldsymbol{\beta}'\mathbf{\Sigma}\boldsymbol{\beta} - \lambda(\boldsymbol{\beta}'\boldsymbol{\beta} - 1) \,, \tag{8}$$

where λ is a Lagrange multiplier. Then the partial derivatives of ϕ with respect to $\boldsymbol{\beta}$ is

$$\frac{\partial \phi}{\partial \boldsymbol{\beta}} = 2\mathbf{\Sigma}\boldsymbol{\beta} - 2\lambda\boldsymbol{\beta} \,, \tag{9}$$

since $\boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta}$ and $\boldsymbol{\beta}'\boldsymbol{\beta}$ have derivatives everywhere in the region that contains $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$. Thus, set

$$\frac{\partial \phi}{\partial \boldsymbol{\beta}} = 2\boldsymbol{\Sigma}\boldsymbol{\beta} - 2\lambda\boldsymbol{\beta} = 2(\boldsymbol{\Sigma} - \lambda I)\boldsymbol{\beta} = 0 \ . \tag{10}$$

In order to get the solution of equation (10) with $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$, we must have $\boldsymbol{\Sigma} - \lambda I$ singular. In other words,

$$|\boldsymbol{\Sigma} - \lambda I| = 0 \ . \tag{11}$$

Since $|\boldsymbol{\Sigma} - \lambda I|$ is a polynomial in λ of degree p, thus equation (11) has p roots. Without loss of generality, let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. Then

$$\boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta} = \lambda\boldsymbol{\beta}'\boldsymbol{\beta} = \lambda \ . \tag{12}$$

This shows us that if $\boldsymbol{\beta}$ satisfies equation (10) and $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$, then the variance of $\boldsymbol{\beta}'\mathbf{Y}$ is λ.

Let $\boldsymbol{\beta}^{(1)}$ be a normalized solution of $(\boldsymbol{\Sigma} - \lambda_1 I)\boldsymbol{\beta} = 0$. Then $\boldsymbol{\beta}^{(1)}$ is the corresponding first principal component vector and $U_1 = \boldsymbol{\beta}^{(1)'}\mathbf{Y}$ is a normalized linear combination of the original variables with maximum variance. Next, we want to find a normalized linear combination $\boldsymbol{\beta}'\mathbf{Y}$ that has maximum variance of all linear combination uncorrelated to U_1. That is,

$$0 = E(\boldsymbol{\beta}'\mathbf{Y}U_1') = E(\boldsymbol{\beta}'\mathbf{Y}\mathbf{Y}'\boldsymbol{\beta}^{(1)}) = \boldsymbol{\beta}'E(\mathbf{Y}\mathbf{Y}')\boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta}^{(1)} = \lambda_1\boldsymbol{\beta}'\boldsymbol{\beta}^{(1)} \ , \tag{13}$$

since $\boldsymbol{\Sigma}\boldsymbol{\beta}^{(1)} = \lambda_1\boldsymbol{\beta}^{(1)}$. So now, we want to maximize

$$\phi_2 = \boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta} - \lambda(\boldsymbol{\beta}'\boldsymbol{\beta} - 1) - 2\nu_1\boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta}^{(1)} \ , \tag{14}$$

where λ and ν_1 are both Lagrange multipliers. Then the partial derivatives of ϕ_2 is

$$\frac{\partial \phi_2}{\partial \boldsymbol{\beta}} = 2\boldsymbol{\Sigma}\boldsymbol{\beta} - 2\lambda\boldsymbol{\beta} - 2\nu_1\boldsymbol{\Sigma}\boldsymbol{\beta}^{(1)} \ . \tag{15}$$

Multiply $\boldsymbol{\beta}^{(1)'}$ on the left of the above equation, then set it to zero. We have

$$0 = 2\boldsymbol{\beta}^{(1)'}\boldsymbol{\Sigma}\boldsymbol{\beta} - 2\lambda\boldsymbol{\beta}^{(1)'}\boldsymbol{\beta} - 2\nu_1\boldsymbol{\beta}^{(1)'}\boldsymbol{\Sigma}\boldsymbol{\beta}^{(1)} = -2\nu_1\lambda_1 \ . \tag{16}$$

Thus, $\nu_1 = 0$ and $\boldsymbol{\beta}$ must satisfy equation (10), and therefore λ must satisfy equation (11). Let $\lambda_{(2)}$ be the maximum of $\lambda_1, \ldots, \lambda_p$ such that there is a vector $\boldsymbol{\beta}$ satisfying $(\boldsymbol{\Sigma} - \lambda_{(2)}I)\boldsymbol{\beta} = 0$, $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$, and equation (13). We denote this vector by $\boldsymbol{\beta}^{(2)}$ and $U_2 = \boldsymbol{\beta}^{(2)'}\mathbf{Y}$.

Continue the same procedure to the $(r+1)$-st step. All we need is to find a vector $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}'\mathbf{Y}$ has maximum variance of all linear combinations which are uncorrelated with $U_1, \ldots, U_r$. In other words, the above task is to find $\boldsymbol{\beta}$ such that

$$0 = E(\boldsymbol{\beta}'\mathbf{Y}U_i) = E(\boldsymbol{\beta}'\mathbf{Y}'\mathbf{Y}\boldsymbol{\beta}^{(i)}) = \boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta}^{(i)} = \lambda_{(i)}\boldsymbol{\beta}'\boldsymbol{\beta}^{(i)}$$
$$\text{for } i = 1, \ldots, r \ . \tag{17}$$

So we want to maximize

$$\phi_{r+1} = \boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta} - \lambda(\boldsymbol{\beta}'\boldsymbol{\beta} - 1) - 2\sum_{i=1}^{r}\nu_i\boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta}^{(i)} \ , \tag{18}$$

where $\lambda, \nu_1, \nu_2, \ldots, \nu_r$ are all Lagrange multipliers. First, take the partial derivatives of ϕ_{r+1}, that is,

$$\frac{\partial\phi_{r+1}}{\partial\boldsymbol{\beta}} = 2\boldsymbol{\Sigma}\boldsymbol{\beta} - 2\lambda\boldsymbol{\beta} - 2\sum_{i=1}^{r}\nu_i\boldsymbol{\Sigma}\boldsymbol{\beta}^{(i)} \ . \tag{19}$$

Second, multiply $\boldsymbol{\beta}^{(j)'}$ on the left of the above equation and then set this equation to zero. We then have

$$0 = 2\boldsymbol{\beta}^{(j)'}\boldsymbol{\Sigma}\boldsymbol{\beta} - 2\lambda\boldsymbol{\beta}^{(j)'}\boldsymbol{\beta} - 2\nu_j\boldsymbol{\beta}^{(j)'}\boldsymbol{\Sigma}\boldsymbol{\beta}^{(j)} = -2\nu_j\lambda_{(j)} \ . \tag{20}$$

If $\lambda_{(j)} \neq 0$, then $-2\nu_j\lambda_{(j)} = 0$ implies $\nu_j = 0$. On the other hand, if $\lambda_{(j)} = 0$, then $\boldsymbol{\Sigma}\boldsymbol{\beta}^{(j)} = \lambda_{(j)}\boldsymbol{\beta}^{(j)} = 0$ and the j-th term in the sum of equation (19)

vanishes. Hence, $\boldsymbol{\beta}$ must satisfy equation (10) and also λ must satisfy equation (11).

Let $\lambda_{(r+1)}$ be the maximum of λ_1 to λ_p such that there is a vector $\boldsymbol{\beta}$ satisfying $(\boldsymbol{\Sigma} - \lambda_{(r+1)}I)\boldsymbol{\beta} = 0$, $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$, and equation (17). Denote this vector by $\boldsymbol{\beta}^{(r+1)}$ and $U_{r+1} = \boldsymbol{\beta}^{(r+1)'}\mathbf{Y}$. If $\lambda_{(r+1)} = 0$ and $\lambda_{(j)} = 0$ for $j \neq r+1$, then $\boldsymbol{\beta}^{(j)'}\boldsymbol{\Sigma}\boldsymbol{\beta}^{(r+1)}$ does not imply that $\boldsymbol{\beta}^{(j)'}\boldsymbol{\beta}^{(r+1)} = 0$. But we may rewrite $\boldsymbol{\beta}^{(r+1)}$ as $\boldsymbol{\beta}^{(r+1)} + 0 \cdot \boldsymbol{\beta}^{(1)} + \cdots + 0 \cdot \boldsymbol{\beta}^{(r)}$, so that $\boldsymbol{\beta}^{(r+1)}$ is orthogonal to all $\boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)}, \cdots, \boldsymbol{\beta}^{(r)}$. This procedure is continued until at the $(e+1)$-st stage that one can not find a vector $\boldsymbol{\beta}$ satisfying $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$ and equation (17). Since $\boldsymbol{\beta}^{(1)}, \cdots, \boldsymbol{\beta}^{(e)}$ must be linearly independent, so either $e = p$ or $e < p$. But $e < p$ will leads to a contradiction, the details can be seen in Anderson (2003). So we must have $e = p$.

Now let $\mathbf{B} = (\boldsymbol{\beta}^{(1)}, \cdots, \boldsymbol{\beta}^{(p)})$ and

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_{(1)} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_{(2)} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_{(3)} & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_{(p-1)} & 0 \\ 0 & 0 & \cdots & 0 & 0 & \lambda_{(p)} \end{pmatrix} . \tag{21}$$

Then in matrix form, we can rewrite $\boldsymbol{\Sigma}\boldsymbol{\beta}^{(r)} = \lambda_{(r)}\boldsymbol{\beta}^{(r)}$ as

$$\boldsymbol{\Sigma}\mathbf{B} = \mathbf{B}\boldsymbol{\Lambda} \tag{22}$$

and $\boldsymbol{\beta}^{(r)'}\boldsymbol{\beta}^{(r)} = 1$, $\boldsymbol{\beta}^{(r)'}\boldsymbol{\beta}^{(s)} = 0$ for $r \neq s$ as

$$\mathbf{B}'\mathbf{B} = I \ . \tag{23}$$

From equation (22) and equation (23), we have

$$\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B} = \boldsymbol{\Lambda} \ . \tag{24}$$

From the fact that

$$\begin{aligned} |\mathbf{\Sigma} - \lambda I| &= |\mathbf{B}'| \cdot |\mathbf{\Sigma} - \lambda I| \cdot |\mathbf{B}| \\ &= |\mathbf{B}'\mathbf{\Sigma}\mathbf{B} - \lambda \mathbf{B}'\mathbf{B}| = |\mathbf{\Lambda} - \lambda I| \\ &= \prod_{i=1}^{p} (\lambda_{(i)} - \lambda) \ , \end{aligned} \tag{25}$$

it is clearly that the roots of this equation are the diagonal elements of $\mathbf{\Lambda}$. That is, $\lambda_{(1)} = \lambda_1, \lambda_{(2)} = \lambda_2, \cdots, \lambda_{(p)} = \lambda_p$. So that we can state the following theorem.

Theorem 1 ***(Anderson, 2003)***

Let $\mathbf{Y}$ *be a p-component random vector with* $E(\mathbf{Y})$*=0 and* $E(\mathbf{YY}') = \mathbf{\Sigma}$*. Then there exists an orthogonal linear transformation* $\mathbf{B}$ *such that*

$$\mathbf{U} = \mathbf{B}'\mathbf{X} \tag{26}$$

has the covariance matrix

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_3 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_{p-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 & \lambda_p \end{pmatrix} , \tag{27}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ *are roots of equation (11). Then the r-th column of* $\mathbf{B}$*,* $\boldsymbol{\beta}^{(r)}$*, satisfies* $(\mathbf{\Sigma} - \lambda_r I)\boldsymbol{\beta}^{(r)} = 0$*. And the r-th component of* $\mathbf{U}$*,* $U_r = \boldsymbol{\beta}^{(r)'}\mathbf{Y}$ *has maximum variance of all normalized linear combinations uncorrelated with* $U_1, \cdots, U_{r-1}$*.*

Thus $\boldsymbol{\beta}^{(1)}, \cdots, \boldsymbol{\beta}^{(p)}$ are eigenvectors and $\lambda_1, \cdots, \lambda_p$ are eigenvalues of Σ. In PCA, for $r = 1, \cdots, p$, $\boldsymbol{\beta}^{(p)}$ is the r-th principal components of $\mathbf{Y}$ and U_r

is called the score of the r-th principal components. These scores will catch the special properties or features of the curves corresponding to different principal components. Hence we can use these principal component scores to monitor the changes or shifts of the profiles.

Now for real data analysis, suppose we have n profiles, each with p set points. Compute the sample covariance matrix $\hat{\Sigma}$ by

$$\hat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})' \ ,$$

where $\boldsymbol{y}_i$ is the i-th profile and $\bar{\boldsymbol{y}} = \sum_{i=1}^{n} \boldsymbol{y}_i/n$. Apply the eigenanalysis to $\hat{\Sigma}$. The eigenvector corresponding to the j-th largest eigenvalue is the j-th principal component, $j = 1, \ldots, p$. Then simply project each profile onto these eigenvectors to have the corresponding principal component scores that we use in our monitoring scheme.

3.4 A Simple Cross-Validation Procedure

We adopt the principle of cross-validation (CV) as the criterion to decide how many PC scores to use. Here, we use an example to illustrate this procedure. Assume there are $n = 50$ profiles, each with $p = 19$ set points. We randomly divide the data set into $g = 5$ groups of 10 samples each. The CV procedure given in Jackson (1991) is as follows.
Repeat the following steps for $i = 1$ ro g:

1. Delete the i-th group from the original data set. Perform PCA on the remaining samples to obtain all nineteen eigenvectors. Denote them by $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{19}$.

2. Project each of the 10 deleted samples onto the eigenvectors obtained in step 1 to obtain its 19 PC scores.

3. Using, in turn, the first principal component, the first two principal components, and so on, to obtain the predicted values of the deleted sample $\boldsymbol{y} = (y_1, \cdots, y_{19})'$ by

$$\hat{\boldsymbol{y}} = \bar{y} + U\mathbf{Z} \tag{28}$$

where $\bar{y} = \sum_{i=1}^{19} y_i/19$ and $U = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{19})$ is the matrix of the eigenvectors obtained in step 1, and $\mathbf{Z}$ is the vector of the corresponding principal component scores obtained in step 2.

4. For each of the deleted samples, compute $Q = (\boldsymbol{y} - \hat{\boldsymbol{y}})'(\boldsymbol{y} - \hat{\boldsymbol{y}})$.

Then there will be 50 values of Q for the one-principal-component model, another 50 for the two-principal-component model, and so on. For each model, add up the 50 Q-statistics. These are called PRESS-statistics. Let PRESS(i) be the sum of the 50 values of Q for the i-principal-component model, $i = 1, \ldots, p$, and PRESS(0) be the sum of squares $\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \bar{y}_i)^2$. Finally, for k starting from 1, compute the statistic

$$W = \frac{[PRESS(k-1) - PRESS(k)]/D_M}{PRESS(k)/D_R}\ , \tag{29}$$

where $D_M = n + p - 2k$ and $D_R = p(n-1) - \sum_{i=1}^{k}(n + p - 2i)$. If $W > 1$, then retain the k-th principal component in the model and continue to test the $(k+1)$-st. The procedure stops when $W < 1$.

3.5 T^2 statistics

In this paper, for Phase I monitoring, two versions of T^2 statistic are studied. They are different in the estimation of the covariance matrix.

Method 1: Use the usual sample covariance matrix

$$T^2_{1,i} = (\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}})'\mathbf{S}_1^{-1}(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}})\ , \tag{30}$$

where $\hat{\boldsymbol{\beta}}_i = (\hat{\beta}_{i,1}, \hat{\beta}_{i,2}, \cdots, \hat{\beta}_{i,p})'$ is the score vector of the i-th sample profile, $\bar{\hat{\boldsymbol{\beta}}} = (\frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_{i,1}, \frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_{i,2}, \cdots, \frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_{i,p})'$, $i = 1, \ldots, n$, and $\mathbf{S}_1 = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}})(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}})'$. So according to William et al. (2003),

$$T_1^2 \frac{n}{(n-1)^2} \sim Beta\left[\frac{p}{2}, \frac{n-p-1}{2}\right]. \tag{31}$$

Thus, the control limit $UCL_1 = \frac{(n-1)^2}{n} Beta_{1-\alpha,p/2,(n-p-1)/2}$.

Method 2: Use the successive-difference-based sample covariance matrix

$$T_{2,i}^2 = (\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}})'\mathbf{S}_2^{-1}(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}), \tag{32}$$

where $\mathbf{S}_2 = \sum_{i=1}^{n-1}\mathbf{V}_i\mathbf{V}_i'/2(n-1)$ with $\mathbf{V}_i = \hat{\boldsymbol{\beta}}_{i+1} - \hat{\boldsymbol{\beta}}_i$, $i = 1, \ldots, n-1$. Then

$$T_2^2 \frac{n}{(n-1)^2} \sim Beta\left[\frac{p}{2}, \frac{f-p-1}{2}\right], \tag{33}$$

where $f = \frac{2(n-1)^2}{3n-4}$. Thus, the control limit $UCL_2 = \frac{(n-1)^2}{n} Beta_{1-\alpha,p/2,(f-p-1)/2}$. See Williams et al. (2003).

4 Simulation Studies - Aspartame Example

4.1 Settings for Simulation

In our simulation study, we use the model

$$y = I + Me^{N(x-1)^2} + \epsilon, \tag{34}$$

where $I \sim N(1, 0.2^2)$, $M \sim N(15, 1)$, $N \sim N(-1.5, 0.3^2)$, $\epsilon \sim N(0, 0.3^2)$ and x=0.64, 0.8, ..., 3.52 to simulate profile data of the aspartame example. We assess the performance in terms of the average run length (ARL) through simulation studies. Let ARL_0 denote the in-control ARL. All charts are designed to have the same ARL_0=200, which corresponds to the false alarm rate α=0.005 .

To compute ARL, we simulate 200,000 profiles each time and count the number of profiles that are out of control. Then the proportion of the out-of-control profiles is an estimate of the out-of-control probability for each profile. Consequently, the reciprocal of this estimate is an ARL estimate. We repeat the above steps 1000 times to obtain 1000 ARL estimates. Compute the sample mean and sample standard deviation of these 1000 estimates. The final ARL estimate is taken as the sample mean and its standard error is the sample standard deviation divided by $1000^{1/2}$.

4.2 Number of Principal Components

An important issue arises in using PCA: how many principal components should be used for our purpose? More specifically, how does the choice affect the detecting power of our monitoring scheme? Here, the detecting power means the ability of our monitoring scheme detecting the real out-of-control profiles. For example, simulate fifty profiles with three out-of-control profiles in the data set. If our scheme catches two of these three, then we measure the detecting power by 2/3 for this data set. In order to study this issue, we choose some settings of I, M, N shifts and simulate fifty profiles for each setting. Then repeat each setting 20,000 times to get the average detecting power and the corresponding percentage of variation explained across different numbers of principal components.

Partial results are presented here for illustration. Figures 3 and 4 show the results for the situation of one out-of-control profile in fifty profiles with a mean shift of 5 sigma on I. Figures 5 and 6 are the plots that two out-of-control profiles in fifty profiles with a mean shift of 5 sigma on M. Figures 7 and 8 display the results for the situation that there are three out-of-control profiles in fifty profiles with a mean shift of 2 sigma on N. From these plots we can clearly see that more principal components does not lead to

more detecting power, although we do have more variation explained. So choosing number of principal components becomes an important task. We use cross-validation method to decide the number of principal components to use.

Here, we conduct another simulation study on this issue. We generate fifty profiles each time, apply the cross-validation method to select the number of principal components. Repeat the process 20,000 times and then plot the histogram of the results. Figures 9-11 display respectively the resulting histograms for the cases of I-shift, M-shift, and N-shift described above. We can see that for all these three settings, the cross-validation method seems in favor of choosing three principal components.

Also, we wonder if the result may be affected by the number of the dividing groups. In order to see this, we simulate fifty profiles and divide the sample profiles into two, ten, or fifty groups. Then apply the cross-validation method to select the number of principal components for each case. Repeat the procedure 20,000 times and plot the corresponding histograms. Figures 12-14 display the plots for two, ten, and fifty groups, respectively. We observe that as the number of groups increases, the selected number of principal components decreases. It is surprising to see that the delete-one cross-validation method chooses only one principal component.

4.3 Phase I monitoring

We now describe the performance of Phase I analysis by an example. In Phase I, we have 200 historical profiles each with 19 set points; see Figure 15. Fit these profiles with B-splines of effective degrees of freedom 5; see Figure 16. And we have a sample covariance matrix of dimension 19×19. Then we apply PCA to the covariance matrix and get the corresponding 19 eigenvalues. By ordering these eigenvalues, we have $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{19}$. For our sample

profiles, we divide the data set into ten groups of twenty observations each. The number of principal components selected by the cross-validation method is four. Then the first principal component is the eigenvector corresponding to λ_1, the second principal component is the eigenvector corresponding to λ_2, and so on. The explanation percentage corresponding to the i-th principal component is computed by $\lambda_i / \sum_{i=1}^{19} \lambda_i$. So the first four principal components account for 0.7644, 0.1742, 0.0309, and 0.0086 of variation in the profiles, respectively. The total is 0.9781.

Now for each profile, we project it onto the first four eigenvectors to get the scores. In our case, $\hat{\boldsymbol{\beta}}_i$ is a 4×1 vector of the scores of the first four principal components. Here the number of profiles n=200 and the number of principal components is 4. Set α at 0.005 in order to have $ARL_0 = 200$. Then compute both T_1^2 and T_2^2 given in Section 3.5. We then have the corresponding control charts displayed in Figures 17 and 18, respectively. We can see that in T_1^2 control chart, there is an out-of-control point. So, we remove the out-of-control profile and then take the remaining 199 profiles and recalculate the T^2 statistics again. Then plot the corresponding T_1^2 and T_2^2 control charts. If there are still some out-of-control profiles, then remove them, recalculate, and plot the control chart again until all of the remaining profiles are in control. In our case, the T_1^2 and T_2^2 control charts, displayed respectively in Figures 19 and 20, indicate that the remaining 199 profiles are in control.

Also, we notice some differences between the performance of T_1^2 and T_2^2. In general, T_2^2 performs better when out-of-control profiles occur successively, while T_1^2 performs better for temporal shifts. To see this, as before, we simulate sets of fifty profiles in which some are shifted. For one shifted profile only, we set it at the middle of the sequence of sample profiles. For two shifted profiles, we set them at the one-third and the two-third of the

sequenced sample profiles. And for three shifted profiles, we set them at the one-fourth, the middle, and the three-fourth of the sequenced sample profiles. Since I, M, N shifts contribute variation of the profiles in different aspects, we need to quantify the extra variation caused by the shift in terms of the response variable of the profile so that we can have a comparison basis for different types of shifts. The variance of $Y(x)$ can be approximated as follows.

Since

$$Y(x) = I + Me^{N(x-1)^2} + \epsilon \approx \mu_I + \mu_M \times e^{\mu_N(x-1)^2} + \epsilon + (I - \mu_I) \\ +(M - \mu_M) \times e^{\mu_N(x-1)^2} + (N - \mu_N)(x-1)^2\mu_M \times e^{\mu_N(x-1)^2} ,$$

we have

$$E(Y(x)) \approx \mu_I + \mu_M \times e^{\mu_N(x-1)^2},$$

$$Var(Y(x)) \approx \sigma_I^2 + \sigma_M^2(e^{\mu_N(x-1)^2})^2 + \sigma_N^2(\mu_M(x-1)^2 \times e^{\mu_N(x-1)^2})^2 + \sigma_\epsilon^2 .$$

Also, $Bias(Y(x)) = E(Y(x) - E(Y(x)))$. Note that, when the μ_I is shifted to $\mu_I + \delta\sigma_I$, $Bias(Y(x)) \approx \delta\sigma_I$; but when μ_N is shifted to $\mu_N + \delta\sigma_N$, then $Bias(Y(x)) \approx \delta\sigma_N(x-1)^2\mu_M e^{\mu_N(x-1)^2}$. Define

$$MSE(Y(x)) = E(Y(x) - E(Y(x)))^2 = Var(Y(x)) + Bias^2(Y(x)).$$

Then the integrated mean squared error ($IMSE$) of the profile is

$$IMSE = \int_{0.64}^{3.52} MSE(Y(x))dx = \int_{0.64}^{3.52} Var(Y(x))dx + \int_{0.64}^{3.52} Bias^2(Y(x))dx.$$

When the process is in control, denote the $IMSE$ by $IMSE_0$; when there are shifts in I, M, and N, denote the $IMSE$ by $IMSE_1$. We then define a measure called the variation expansion factor (vef) by

$$\text{vef} = \sqrt{\frac{IMSE_1}{IMSE_0}} .$$

Repeat 20,000 times to get the average of these measures for each setting. The performance is assessed by the detecting power and the false-alarm rate. Tables 1-3 display respectively the simulation results for I, M, N shifts. From these simulation results, we can clearly see that T_1^2 outperforms T_2^2 for temporal shifts with both criteria. Note that the detecting power decreases as the number of out-of-control profiles increases. This probability is caused by the fact that the parameter situation gets power as the data contamination gets worse.

4.4 Phase II monitoring

In Phase II monitoring, we use the true mean function $\mu(\cdot)$ and the sample covariance matrix Σ of the above 199 profiles as the in-control process parameters to perform our simulation. Let $\boldsymbol{\mu}$ be the 19×1 vector computed by $\mu(x) = 1 + 15e^{-1.5(x-1)^2}$ at $x = 0.64, 0.8, \ldots, 3.52$. We then generate in-control profiles from $MVN(\boldsymbol{\mu}, \Sigma)$. Consider I-shift from μ_I to $\mu_I + \alpha \times \sigma_I$, $\alpha = 0, 0.25, \ldots, 3$; M-shift from μ_M to $\mu_M + \beta \times \sigma_M$, $\beta = 0, 0.25, \ldots, 3$; N-shift from μ_N to $\mu_N + \gamma \times \sigma_N$, $\gamma = 0, 0.25, \ldots, 3$. Figures 21-23 display the curves of $f(x) = I + Me^{N(x-1)^2}$ with different value of I, M, N respectively. Figure 24 display the plot of the first four eigenvectors of Σ. And Figures 25-28 illustrate respectively the corresponding features captured by the first four principal components, by showing in each plot the mean profile and the profiles corresponding to the largest score and smallest score of the principal component. Now for each shift in I, M, N, we generate 200,000 profiles to computed an ARL estimate. Then repeat this 1000 times to get a more accurate estimate along with its standard error as before. Tables 4-6 display the ARL values for the shifts in I, M, N, respectively. We can see from Table 4 that PC3 and PC1 can capture the shift in I. PC3 has the best power for detecting I-shift. PC2 and PC4 hardly have any power on detecting I-shift. Also, we can see from Table 5 that PC2 and PC1 out-

perform the other two with PC2 slightly better than PC1 in capturing the shift in M. And for detecting N-shift, we can see from Table 6 that PC2 and PC1 are the best two. But as the shift tends to 3 units of σ_N, PC2 still has the best power and PC3 becomes the second best for detecting N-shift.

In fact, we can obtain the exact ARL values as follows. Let $P = (\mathbf{P}_1, \mathbf{P}_2, \cdots, \mathbf{P}_4)'$, where $\mathbf{P}_i$ is the i-th eigenvector of Σ. Then P projects profiles to the first four principal components. Let $\lambda_1, \cdots, \lambda_4$ be the corresponding eigenvalues of Σ. Since the in-control profile $\mathbf{Y} \sim MVN(\boldsymbol{\mu}, \Sigma)$ and $P\Sigma P' = diag(\lambda_1, \cdots, \lambda_4) \equiv \Lambda_4$, we have the score vector $P\mathbf{Y} \sim MVN(P\boldsymbol{\mu}, \Lambda_4)$. Denote the shifted profile by $\tilde{\mathbf{Y}}$ and the shifted mean profile by $\tilde{\boldsymbol{\mu}}$. Then $P\tilde{\mathbf{Y}} \sim MVN(P\tilde{\boldsymbol{\mu}}, \Lambda_4)$. Let $\boldsymbol{\delta} = \tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}$. The probability of detecting the shift of the i-th principal component chart is

$$p = P(\left|\frac{\mathbf{P}_i'(\mathbf{Y} - \boldsymbol{\mu})}{\sqrt{\lambda_i}}\right| \leq Z_{0.0025})$$

$$= P(\frac{-\mathbf{P}_i'\boldsymbol{\delta}}{\sqrt{\lambda_i}} - Z_{0.0025} \leq \frac{\mathbf{P}_i'(\mathbf{Y} - \tilde{\boldsymbol{\mu}})}{\sqrt{\lambda_i}} \leq \frac{-\mathbf{P}_i'\boldsymbol{\delta}}{\sqrt{\lambda_i}} + Z_{0.0025})$$

$$= P(\frac{-\mathbf{P}_i'\boldsymbol{\delta}}{\sqrt{\lambda_i}} - Z_{0.0025} \leq Z \leq \frac{-\mathbf{P}_i'\boldsymbol{\delta}}{\sqrt{\lambda_i}} + Z_{0.0025}),$$

where Z is the standard normal variate. The value $1/p$ is the actual ARL of the i-th principal-component chart. Tables 7-9 show respectively the actual ARL for each of the I, M, N shifts. By comparing Tables 4-9, we can see that our simulated ARLs are all very close to the actual ARLs, which verifies the correctness of the simulation.

Since each of the I, M, N shifts is captured by more than one principal-component chart, we recommend using the combined chart to monitor these shifts. A combined chart means we monitor the process by more than one chart and the combined chart signals out of control if any of the charts

signals. Set the overall false-alarm rate at $\alpha = 0.005$. Since these scores are independent, the individual false-alarm rate is $\alpha' = 1 - (0.995)^{1/4}$ for each of the four principal component scores. Then we can have the control limits of these four scores as shown below:

$$(\mathbf{P}_1'\boldsymbol{\mu} - Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_1},\ \mathbf{P}_1'\boldsymbol{\mu} + Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_1})$$
$$(\mathbf{P}_2'\boldsymbol{\mu} - Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_2},\ \mathbf{P}_2'\boldsymbol{\mu} + Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_2})$$
$$(\mathbf{P}_3'\boldsymbol{\mu} - Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_3},\ \mathbf{P}_3'\boldsymbol{\mu} + Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_3})$$
$$(\mathbf{P}_4'\boldsymbol{\mu} - Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_4},\ \mathbf{P}_4'\boldsymbol{\mu} + Z_{\frac{1-(0.995)^{1/4}}{2}}\sqrt{\lambda_4})$$

If one of the principal component scores is out of the control limits, then this profile is claimed as out of control. That is, only when all four principal component scores of that profile are within the control limits simultaneously, then it will be treated as an in-control profile. In order to do the comparison the above, we also simulate 200,000 profiles each time to get an ARL estimate. Then repeat 1000 times to get the final ARL estimate and its standard error. Tables 10-12 display the simulation results for ARL comparison for I, M, N shifts, respectively. Also, we can derive the exact control limits and compute the exact ARL. Tables 13-15 show the ARL computed from the exact control limits with shifts in I, M, N, respectively. By comparing the simulated ARL with the exact ARL for our combined chart, we can see that they are very closed to each other. Also, we plot the ARL comparison plots for PC1, PC2, PC3, PC4 and the combined chart with I, M, N shifts in Figures 29-31 respectively. As we can see for I-shifts, PC3 dominates primarily. And for the combined chart scheme, it was the second best in monitoring the shifts in I. As the shifts gets bigger, the difference in ARL between PC3 and the combined chart gets smaller. Also, for M-shifts, PC2 is the best scheme to monitor it. For the second best one, the PC1 and the combined chart are comparable. Note that two ARL curves intersect at about $1.25\sigma_M$ shift. That is, for shifts smaller than $1.25\sigma_M$, PC1 has better power. But the combined

chart scheme outperforms PC1 as the shift size tends larger than $1.25\sigma_M$. We can see the same phenomenon in N. From Figure 31, PC2 is still the best one to monitor shifts in N. PC1 and the combined chart scheme are comparable. Also the two ARL curves intersect at about $1.1\sigma_N$ shift.

Also, we show the ARL values for simultaneous shifts of I and M, I and N, and M and N, in Tables 16-18, respectively. We can see from Table 16 that PC1 takes the major job of monitoring simultaneous shifts in I and M while PC3 serves as the second best. And from Table 17, we can observe that PC2 dominates the simultaneous shifts in I and N while PC1 serves as the second best. Also from Table 18, we can see that PC1 plays the major role in monitoring simultaneous shifts in M and N while PC2 performs the second best. But notice that the power for PC2 here is not comparable with PC1.

Finally, we consider shifts in the variance of I, M, and N. Recall that the in-control situation is: $I \sim N(1, 0.2^2), M \sim N(15, 1), N \sim N(-1.5, 0.3^2)$. Now for variance shifts in I, we consider $\sigma_I = 0.3, 0.4, \ldots, 1$. For each shift, we simulate 200,000 profiles to get an ARL estimate and then repeat the procedure for 1000 times to get the mean and the corresponding standard deviation. The ARL results for each principal component chart and the combined chart are shown in Table 19. And Figure 32 shows the corresponding ARL curves. As we can see that PC3 dominates for the shifts in variance of I while the combined chart scheme also performs very well in monitoring this kind of shifts.

And for variance shifts in M, we consider $\sigma_M = 1.5, 2, \ldots, 5$. The simulation results are shown in Table 20 and Figure 33. Likewise, we can see that PC2 dominates for the shifts in variance of M. But here, the combined chart scheme is comparable with the PC2. As σ_M gets bigger, the combined chart scheme performs even better than PC2. At last, for shifts in variance of N,

we consider $\sigma_N = 0.4, 0.5, \ldots, 1.1$. Table 21 and Figure 34 shows the simulation results. From these simulation results, we can see that PC3 has the best detecting power for variance shifts in N. The combined chart scheme also has great detecting power for monitoring shifts in σ_N. Notice that as σ_N tends bigger than 0.7, the profile shape will be greatly affected. Therefore we can see that all these four principal components are sensitive to it and they all have great detecting power. While it is difficult to figure out which principal component dominates the shifts, we recommend the combined chart scheme for monitoring.

Since the model under study is a random effect model, we are curious of the power in monitoring variance shifts. From the above simulations, we have shown that the combined chart scheme is a very good choice for monitoring random effect model. So here, in order to compare the power of our combined chart in detecting mean shifts and variance shifts, we use the ARL of the combined chart scheme. Then, use our comparison basis vef to quantify the difference between the shifted profiles and the reference profile. Hence we can have the plot that compares the ARL of detecting mean shifts and variance shifts; see Figure 35.

We can see from Figure 35 that for both I, M, N shifts, our monitoring scheme performs better in detecting variance shifts than mean shifts. We can observe that there exist big differences for all I, M, N shifts at small values of comparison basis. And as the value of comparison basis tends bigger, the difference between mean shifts and variance shifts tends smaller. When the value of comparison basis getting large enough, there exist almost no differences between mean shifts and variance shifts. Also from Figure 35, we can see that our monitoring scheme does the best job in detecting variance shifts in N while the second best one is variance shifts in I.

5 A Case Study - VDP Example

5.1 Phase I monitoring

The VDP data set contains $n = 24$ profiles, each was measured at $p = 314$ set points. Figure 2 is the plot of the VDP data. We fit these 24 profiles by B-splines with 16 degrees of freedom. See Figure 36 for the plot of the fitted B-splines. Also Figures 37-40 respectively plot principal components 1 to 4 showing the modes of variation they capture. And the first four principal components account for 0.8535, 0.1084, 0.0190, 0.0084 of variation in the profiles, respectively. The total is 0.9893. Figure 41 is the plot of the corresponding eigenvectors. Now for Phase I monitoring, we use the same method as for the aspartame example. Figures 42 and 43 are the control charts of T_1^2 and T_2^2 respectively. We can see that there are no out-of-control profiles in the VDP data.

5.2 Phase II monitoring

Since, in the aspartame example, the first four principal components cannot distinguish the shifts in I, M, N very clearly, we suggest using the combined chart scheme to monitor these shifts. Now, we are going to demonstrate an example that its principal components can capture the shapes of shifts very clearly by simulation.

For demonstrating Phase I monitoring, we need to generate new VDP data. Applying PCA to the original VDP data, we then have the corresponding eigenvectors $\{\boldsymbol{\rho}_1, \cdots, \boldsymbol{\rho}_{314}\}$ and eigenvalues $\{\lambda_1, \cdots, \lambda_{314}\}$. Since the first four principal components can explain 97.57 percent of variation, we only use these four principal components to regenerate the new VDP data. And these first four principal components account for 0.8402, 0.1072, 0.0192, 0.0091 of variation in these profiles, respectively.

We use the following steps to regenerate a set of new VDP data.

1. Generate i.i.d. $\epsilon_{j,q}$ from the normal distribution with zero mean and variance λ_q, $1 \le j \le 24$, $1 \le q \le 4$.

2. Generate $\tilde{X}_{j,k} = \mu(t_k) + \sum_{q=1}^{4} \epsilon_{j,q} \cdot \rho_q(t_k)$, $1 \le j \le 24$, $1 \le k \le 314$.

Note that $\lambda_1,\ \lambda_2,\ \lambda_3,\ \lambda_4$ are the four largest eigenvalues and $\rho_1,\ \rho_2,\ \rho_3,\ \rho_4$ are the corresponding eigenvectors with length 314. And $\mu(\cdot)$ is the mean profile of the 24 original VDP curves where t_k denotes the k-th set point. Figure 44 displays the simulated VDP data. We can see that the simulated data do capture the peaks and the shapes of the original VDP data.

Then, we apply our estimation method to the generated data to get the simulated mean profile, eigenvalues, and eigenvectors. Figure 45 displays the mean profile of the real data and the simulated data. We can see that these two curves are very close. Figures 46-49 show the first four eigenvectors of the real data and that of the simulated data, respectively. Likewise, from these four figures we can see that the simulated eigenvectors are all close to that of the real data. Also, Figure 50 shows that the simulated eigenvalues are fairly close to the eigenvalues of the original profile data. Therefore, we can say that the data generation method we adopt here can really capture the shapes and variations of the original data.

In Phase II monitoring, we treat the average profile vector $\boldsymbol{\mu}$ of the 24 smoothed VDP profiles and the sample covariance matrix Σ as the in-control process parameters to perform our simulation. Here, $\boldsymbol{\mu} = (\mu(t_1), \cdots, \mu(t_{314}))'$ is a 314×1 vector and the covariance matrix Σ is a 314×314 matrix. We can generate the in-control profiles by $\mathbf{Y} \sim MVN(\boldsymbol{\mu}, \Sigma)$. For out-of-control conditions, we shift the profiles in their first two principal components. That is, generate new profile data by

$$\tilde{\mathbf{Y}} \sim k\sqrt{\lambda_i} \times \boldsymbol{v}_i + MVN(\boldsymbol{\mu}, \Sigma)\ ,$$

where k=0, 0.25, ..., 2 and $\boldsymbol{v}_i$ is the i-th eigenvector of Σ, i=1, 2. We simulated 200,000 profiles to compute an ARL estimate for each out-of-control conditions considered. Then we repeat the procedure 1000 times to get our final ARL estimate along with its standard error. Tables 22 and 23 report the ARL results for shifts in principal component 1 and principal component 2 respectively.

As we can see from Table 22 that shifts in principal component 1 are solely captured by the first principal component score. Likewise, from Table 23, shifts in principal component 2 are captured by the second principal component score. The other three principal component scores make no contributions to the power of detection.

6 Conclusions

The monitoring of process or product profiles is a very popular and promising area of research in statistical process control in recent years. In this study, we discuss monitoring schemes for the random effect nonliner profiles. We use the principal component analysis to analyze the covariance matrix of the profile data and use the corresponding principal component scores that capture the main features of these profile data for process monitoring.

In the study of the aspartame example, our simulation shows that principal component 3 performs the best over the entire range of I shifts. And principal component 2 is the best for monitoring M shifts while principal component 1 also plays an important role in monitoring M shifts. And for shifts in N, principal component 2 performs the best over the entire range. Likewise, principal component 1 also plays an important role in monitoring N shifts. Note that as the shift tends bigger in N shift, all those four principal components are good at catching the out-of-control profiles. And for

each shift, we can also use the combined chart to perform monitoring. We have shown that although this monitoring scheme may not be the best for each shift, it still has comparably good power to monitor these shifts. Moreover, we have displayed in the tables the results when two of I, M, N shift simultaneously.

Using the example of vertical density profiles, we demonstrate that when the shift corresponds to a mode of variation that a particular principal component represents, then we can directly use the score of that principal component to perform monitoring.

So if a principal component can clearly identify the shift, a situation may be rare in real applications, then we recommend monitoring the score of that principal component. Otherwise we recommend the combined chart scheme because it still has comparably good power to monitor these shifts.

Moreover, we use a mean-squares-error-like (MSE-like) measure as a comparison basis to compare the ARL performance in mean shifts and variance shifts. Simulation results indicate that our monitoring scheme performs better in detecting variance shifts. And in the selection of number of principal components, we adopt the cross-validation method for its popularity and simplicity. Also, in the use of cross-validation, we have observed that the selection is affected by the number of groups that the data set is divided into. By our simulation result, we can see that as the number of groups increases, the selected number of principal components decreases. Further studies are needed on this issue.

Finally in the studies of Phase I monitoring, we compare T_1^2 and T_2^2. And we show by simulation that T_1^2 has a better overall performance than T_2^2 under temporal shifts.

References

[1] Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd edition. Wiley, New York.

[2] Castro, P. E., Lawton, W. H., and Sylvestre, E. A. (1986). "Principal Modes of Variation for Processes With Continuous Sample Curves". *Technometrics* 28, pp. 329–337.

[3] de Boor, C. (1978). *A Practical Guide to Splines*. Springer, Berlin.

[4] Jackson, J. E. (1991). *A User's Guide to Principal Components*. Wiley, New York.

[5] Jones, M. C. and Rice, J. A. (1992). "Displaying the Important Features of Large Collections of Similar Curves". *The American Statistician*, Vol. 46, No. 2, pp. 140–145.

[6] Kang, L. and Albin, S. L. (2000). "On-Line Monitoring When the Process Yields a Linear Profile". *Journal of Quality Technology* 32, pp. 418–426.

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

[8] Ramsay, J. O. and Silverman, B. W. (2002). *Applied Functional Data Analysis: Methods and Case Studies.* Springer, New York.

[9] Rice, J. A. and Silverman, B. W. (1991). "Estimating the Mean and Covariance Structure Nonparametrically When the Data Are Curves". *Journal of the Royal Statistical Society*, Ser. B, 53, pp. 233–243.

[10] Shiau, J.-J. H. and Lin, H.-H. (1999). "Analyzing Accelerated Degradation Data by Nonparametric Regression." *IEEE Transactions on Reliability*, Vol. 48, No.2, pp. 149–158.

[11] Shiau, J.-J. H. and Weng, Z.-P. (2004). "Profile Monitoring by Nonparametric Regression". Technical Report. NSC92-2118-M-009-012. To be submitted.

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

[13] Williams, J. D., Woodall, W. H., and Birch, J. B. (2003). "Phase I Monitoring of Nonlinear Profiles ". *Quality & Productivity Research Conference*, Yorktown Heights, NY.

[14] Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004). "Using Control Charts to Monitor Process and Product Quality Profiles". *Journal of Quality Technology* 36, pp. 309–320.

One shift set at the 100th curve among 200 curves

chart\ I shifts (vef)	$5 \times \sigma_I$ (1.543)	$7 \times \sigma_I$ (1.925)	$10 \times \sigma_I$ (2.554)
T_1^2 power	0.4506	0.8213	0.9842
false-alarm	0.0068	0.0062	0.0059
T_2^2 power	0.1467	0.4567	0.869
false-alarm	0.0028	0.0026	0.0025

Two shifts set at the 66th and 133th curve among 200 curves

chart\ I shifts (vef)	$5 \times \sigma_I$ (1.543)	$7 \times \sigma_I$ (1.925)	$10 \times \sigma_I$ (2.554)
T_1^2 power	0.2924	0.52	0.7339
false-alarm	0.0062	0.0059	0.0054
T_2^2 power	0.0726	0.1746	0.333
false-alarm	0.0026	0.0025	0.0022

Three shifts set at the 50th, 100th and 150th curve among 200 curves

chart\ I shifts (vef)	$5 \times \sigma_I$ (1.543)	$7 \times \sigma_I$ (1.925)	$10 \times \sigma_I$ (2.554)
T_1^2 power	0.1972	0.325	0.4515
false-alarm	0.0059	0.0054	0.005
T_2^2 power	0.0392	0.0745	0.1207
false-alarm	0.0025	0.0023	0.002

Table 1: The performance assessed by detecting power and false-alarm rate for T_1^2 and T_2^2 to detect I shifts with temporal shifts.
(i.e. Shift from μ_I=1 to 1+$k \times \sigma_I$ and vef (in parentheses), represents the impact of the shift on the whole function variation.)

One shift set at the 100th curve among 200 curves

chart\ M shifts (vef)	$5 \times \sigma_M$ (3.358)	$7 \times \sigma_M$ (4.580)	$10 \times \sigma_M$ (6.445)
T_1^2 power	0.716	0.9637	0.9989
false-alarm	0.007	0.0068	0.0067
T_2^2 power	0.3077	0.7331	0.9747
false-alarm	0.003	0.0029	0.003

Two shifts set at the 66th and 133th curve among 200 curves

chart\ M shifts (vef)	$5 \times \sigma_M$ (3.358)	$7 \times \sigma_M$ (4.580)	$10 \times \sigma_M$ (6.445)
T_1^2 power	0.4449	0.671	0.8207
false-alarm	0.0067	0.0066	0.0064
T_2^2 power	0.1209	0.2622	0.4298
false-alarm	0.0029	0.003	0.0028

Three shifts set at the 50th, 100th and 150th curve among 200 curves

chart\ M shifts (vef)	$5 \times \sigma_M$ (3.358)	$7 \times \sigma_M$ (4.580)	$10 \times \sigma_M$ (6.445)
T_1^2 power	0.2854	0.4197	0.52
false-alarm	0.0065	0.0064	0.0063
T_2^2 power	0.0583	0.1025	0.1502
false-alarm	0.0028	0.0028	0.0027

Table 2: The performance assessed by detecting power and the false-alarm rate for T_1^2 and T_2^2 to detect M shifts with temporal shifts.
(i.e. Shift from μ_M=15 to 15+$k \times \sigma_M$ and vef (in parentheses), represents the impact of the shift on the whole function variation.)

One shift set at the 100th curve among 200 curves

chart\ N shifts (vef)	$2 \times \sigma_N$ (2.284)	$2.5 \times \sigma_N$ (2.979)	$3 \times \sigma_N$ (3.923)
T_1^2 power	0.1587	0.4858	0.8961
false-alarm	0.006	0.0045	0.0028
T_2^2 power	0.0396	0.2023	0.6566
false-alarm	0.0024	0.0018	9.5612e-004

Two shifts set at the 66th and 133th curve among 200 curves

chart\ N shifts (vef)	$2 \times \sigma_N$ (2.284)	$2.5 \times \sigma_N$ (2.979)	$3 \times \sigma_N$ (3.923)
T_1^2 power	0.1152	0.3208	0.6172
false-alarm	0.005	0.0031	0.0016
T_2^2 power	0.0243	0.092	0.2507
false-alarm	0.002	0.0012	5.0625e-004

Three shifts set at the 50th, 100th and 150th curve among 200 curves

chart\ N shifts (vef)	$2 \times \sigma_N$ (2.284)	$2.5 \times \sigma_N$ (2.979)	$3 \times \sigma_N$ (3.923)
T_1^2 power	0.0906	0.2097	0.3781
false-alarm	0.0042	0.0025	0.0011
T_2^2 power	0.017	0.0451	0.0946
false-alarm	0.0016	8.8298e-004	3.1702e-004

Table 3: The performance assessed by detecting power and the false-alarm rate for T_1^2 and T_2^2 to detect N shifts with temporal shifts.

(i.e. Shift from μ_N= -1.5 to -1.5+$k \times \sigma_N$ and vef (in parentheses), represents the impact of the shift on the whole function variation.)

α							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.002)	(1.007)	(1.015)	(1.027)	(1.042)	(1.060)
PC1	200.38	196.64	187.86	174.44	157.88	140.51	123.51
	0.2004	0.1970	0.1795	0.1630	0.1394	0.1226	0.0940
PC2	200.12	200.14	200.16	200.21	199.77	199.26	198.53
	0.2003	0.1961	0.1981	0.1962	0.2006	0.1979	0.1987
PC3	200.13	172.04	119.20	76.094	48.049	30.992	20.497
	0.2027	0.1669	0.0933	0.0487	0.0239	0.0118	0.0066
PC4	200.40	200.34	200.27	199.69	199.23	198.58	197.87
	0.2031	0.1922	0.2027	0.2006	0.1964	0.1984	0.1957
chart	1.75	2	2.25	2.5	2.75	3	
	(1.081)	(1.105)	(1.131)	(1.160)	(1.190)	(1.223)	
PC1	107.43	93.12	80.410	69.369	59.919	51.721	
	0.0796	0.0643	0.0489	0.0415	0.0310	0.0268	
PC2	198.53	197.7	196.91	196.16	195.29	194.32	
	0.1965	0.1950	0.1936	0.1915	0.1858	0.1976	
PC3	13.999	9.859	7.1627	5.3581	4.1247	3.2674	
	0.0036	0.0020	0.0012	0.0008	0.0005	0.0003	
PC4	197.71	196.6	195.58	194.84	193.55	192.47	
	0.1951	0.1950	0.1900	0.1981	0.1839	0.1929	

Table 4: ARL comparison for I-shift ($\mu_I = 1 + 0.2 \times \alpha$).

β							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.019)	(1.062)	(1.126)	(1.207)	(1.302)	(1.408)
PC1	200.6	177.7	131.5	89.18	59.37	39.81	27.25
	0.2032	0.1735	0.1068	0.0601	0.0322	0.0116	0.0099
PC2	200.1	168.9	112.9	69.49	42.81	27.02	17.66
	0.1943	0.1549	0.0871	0.0411	0.0201	0.0098	0.0050
PC3	200.3	199.0	196.8	192.8	187.7	181.3	174.3
	0.1958	0.2064	0.1942	0.1884	0.1759	0.1689	0.1611
PC4	199.8	199.8	197.9	195.5	191.9	187.5	182.1
	0.1963	0.1986	0.2041	0.1906	0.1901	0.1841	0.1679
chart	1.75	2	2.25	2.5	2.75	3	
	(1.524)	(1.647)	(1.775)	(1.908)	(2.044)	(2.183)	
PC1	19.06	13.65	10.023	7.5328	5.7976	4.5616	
	0.0058	0.0034	0.0021	0.0014	0.0009	0.0006	
PC2	11.93	8.344	6.0404	4.5196	3.492	2.7828	
	0.0028	0.0016	0.0009	0.0006	0.0004	0.0003	
PC3	166.2	157.9	149.08	140.33	131.89	123.37	
	0.1489	0.1378	0.1276	0.1192	0.1080	0.0961	
PC4	176.4	169.6	163.41	156.41	149.01	142.01	
	0.1699	0.1562	0.1475	0.1354	0.1294	0.1229	

Table 5: ARL comparison for M-shift ($\mu_M = 15 + 1 \times \beta$).

γ							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.045)	(1.123)	(1.233)	(1.377)	(1.552)	(1.761)
PC1	200.6	175.46	123.33	77.235	46.416	27.765	16.783
	0.2045	0.1576	0.0990	0.0480	0.0221	0.0103	0.0047
PC2	200.07	170.31	109.92	61.319	32.571	17.17	9.2111
	0.1943	0.1512	0.0825	0.0339	0.0132	0.0051	0.0018
PC3	200.33	195.5	188.38	184.87	188.98	198.38	192.89
	0.1958	0.1895	0.1835	0.1750	0.1877	0.1956	0.1812
PC4	199.79	196.79	191.12	189.78	194.81	199.71	184.19
	0.1963	0.2030	0.1898	0.1835	0.1934	0.1935	0.1758
chart	1.75	2	2.25	2.5	2.75	3	
	(2.004)	(2.284)	(2.607)	(2.979)	(3.413)	(3.923)	
PC1	10.336	6.5317	4.2695	2.9118	2.0902	1.5929	
	0.0023	0.0011	0.0005	0.0003	0.0001	0.00009	
PC2	5.133	3.0404	1.9639	1.4142	1.1482	1.0383	
	0.0007	0.0003	0.0001	0.00006	0.00003	0.00001	
PC3	138.36	64.003	21.627	6.4314	2.1517	1.1355	
	0.1115	0.0351	0.0068	0.0011	0.0002	0.00003	
PC4	126.66	61.374	23.666	8.422	3.222	1.5783	
	0.1013	0.0342	0.0081	0.0017	0.0003	0.00009	

Table 6: ARL comparison for N-shift ($\mu_N = -1.5 + 0.3 \times \gamma$).

α							
chart	0 (1)	0.25 (1.002)	0.5 (1.007)	0.75 (1.015)	1 (1.027)	1.25 (1.042)	1.5 (1.060)
PC1	200.0	196.78	187.65	174.04	157.74	140.46	123.39
PC2	200.0	199.96	199.84	199.64	199.35	198.99	198.55
PC3	200.0	172.05	119.16	75.951	47.999	30.926	20.493
PC4	200.0	199.94	199.77	199.49	199.09	198.58	197.96
chart	1.75 (1.081)	2 (1.105)	2.25 (1.131)	2.5 (1.160)	2.75 (1.190)	3 (1.223)	
PC1	107.48	93.109	80.42	69.38	59.87	51.73	
PC2	198.03	197.44	196.7	196.0	195.2	194.3	
PC3	13.999	9.8623	7.162	5.358	4.126	3.267	
PC4	197.24	196.41	195.5	194.4	193.3	192.1	

Table 7: Real ARL comparison for I-shift ($\mu_I = 1 + 0.2 \times \alpha$).

β							
chart	0 (1)	0.25 (1.019)	0.5 (1.062)	0.75 (1.126)	1 (1.207)	1.25 (1.302)	1.5 (1.408)
PC1	200.0	177.55	131.31	89.198	59.356	39.824	27.235
PC2	200.0	168.89	112.77	69.516	42.801	27.032	17.646
PC3	200.0	199.19	196.78	192.89	187.67	181.32	174.07
PC4	200.0	199.45	197.83	195.17	191.56	187.09	181.87
chart	1.75 (1.524)	2 (1.647)	2.25 (1.775)	2.5 (1.908)	2.75 (2.044)	3 (2.183)	
PC1	19.056	13.655	10.02	7.5334	5.7958	4.5617	
PC2	11.926	8.3429	6.04	4.5196	3.4921	2.7826	
PC3	166.15	157.78	149.2	140.47	131.87	123.46	
PC4	176.04	169.71	163.0	156.08	149.01	141.88	

Table 8: Real ARL comparison for M-shift ($\mu_M = 15 + 1 \times \beta$).

γ							
chart	0 (1)	0.25 (1.045)	0.5 (1.123)	0.75 (1.233)	1 (1.377)	1.25 (1.552)	1.5 (1.761)
PC1	200.0	175.24	123.25	77.172	46.405	27.762	16.779
PC2	200.0	170.23	109.89	61.338	32.564	17.173	9.2121
PC3	200.0	195.66	188.26	184.73	188.84	198.28	192.84
PC4	200.0	196.55	191.12	189.61	194.36	199.99	184.32
chart	1.75 (2.004)	2 (2.284)	2.25 (2.607)	2.5 (2.979)	2.75 (3.413)	3 (3.923)	
PC1	10.333	6.5314	4.27	2.9119	2.0901	1.5928	
PC2	5.1331	3.0408	1.964	1.4141	1.1483	1.0383	
PC3	138.29	63.968	21.63	6.4329	2.1517	1.1355	
PC4	126.59	61.296	23.66	8.422	3.2224	1.578	

Table 9: Real ARL comparison for N-shift ($\mu_N = -1.5 + 0.3 \times \gamma$).

α							
chart	0 (1)	0.25 (1.002)	0.5 (1.007)	0.75 (1.015)	1 (1.027)	1.25 (1.042)	1.5 (1.060)
Combined	200.49 0.2053	189.43 0.1911	160.22 0.1427	124.07 0.0988	89.755 0.0622	62.165 0.0350	42.299 0.0191
chart	1.75 (1.081)	2 (1.105)	2.25 (1.131)	2.5 (1.160)	2.75 (1.190)	3 (1.223)	
Combined	28.733 0.0106	19.734 0.0061	13.784 0.0036	9.8526 0.0021	7.2125 0.0013	5.4224 0.0008	

Table 10: Combined chart ARL for I-shift ($\mu_I = 1 + 0.2 \times \alpha$).

β							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.019)	(1.062)	(1.126)	(1.207)	(1.302)	(1.408)
Combined	200.49	181.31	139.05	95.667	62.436	40.059	25.88
	0.2053	0.1695	0.1162	0.0647	0.0353	0.0017	0.0091
chart	1.75	2	2.25	2.5	2.75	3	
	(1.524)	(1.647)	(1.775)	(1.908)	(2.044)	(2.183)	
Combined	17.02	11.46	7.9553	5.689	4.2035	3.2099	
	0.0048	0.0027	0.0015	0.0009	0.0005	0.0003	

Table 11: Combined chart ARL for M-shift ($\mu_M = 15 + 1 \times \beta$).

γ							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.045)	(1.123)	(1.233)	(1.377)	(1.552)	(1.761)
Combined	200.49	179.56	131.96	83.949	48.194	25.846	13.451
	0.2053	0.1616	0.1052	0.0530	0.0237	0.0089	0.0033
chart	1.75	2	2.25	2.5	2.75	3	
	(2.004)	(2.284)	(2.607)	(2.979)	(3.413)	(3.923)	
Combined	7.026	3.819	2.2347	1.4543	1.1021	1.0055	
	0.0012	0.0005	0.0002	0.00007	0.00002	0.00001	

Table 12: Combined chart ARL for N-shift ($\mu_N = -1.5 + 0.3 \times \gamma$).

α							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.002)	(1.007)	(1.015)	(1.027)	(1.042)	(1.060)
Combined	200.0	188.98	160.3	124.03	89.657	62.159	42.295
chart	1.75	2	2.25	2.5	2.75	3	
	(1.081)	(1.105)	(1.131)	(1.160)	(1.190)	(1.223)	
Combined	28.738	19.729	13.786	9.849	7.213	5.421	

Table 13: Real combined chart ARL for I-shift ($\mu_I = 1 + 0.2 \times \alpha$).

β							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.019)	(1.062)	(1.126)	(1.207)	(1.302)	(1.408)
Combined	200.0	181.28	138.9	95.578	62.398	40.086	25.892
chart	1.75	2	2.25	2.5	2.75	3	
	(1.524)	(1.647)	(1.775)	(1.908)	(2.044)	(2.183)	
Combined	17.022	11.469	7.953	5.689	4.203	3.21	

Table 14: Real combined chart ARL for M-shift ($\mu_M = 15 + 1 \times \beta$).

γ							
chart	0	0.25	0.5	0.75	1	1.25	1.5
	(1)	(1.045)	(1.123)	(1.233)	(1.377)	(1.552)	(1.761)
Combined	200.0	179.26	131.93	83.893	48.162	25.855	13.446
chart	1.75	2	2.25	2.5	2.75	3	
	(2.004)	(2.284)	(2.607)	(2.979)	(3.413)	(3.923)	
Combined	7.025	3.8194	2.235	1.454	1.102	1.006	

Table 15: Real combined chart ARL for N-shift ($\mu_N = -1.5 + 0.3 \times \gamma$).

$\alpha\backslash\beta$		1.5	1.75	2	2.25	2.5
1.5	PC1	12.9351	11.5453	10.3478	9.2981	8.3825
		0.0031	0.0027	0.0022	0.0019	0.0015
	PC2	19.1699	19.4443	19.6996	19.9855	20.2791
		0.0058	0.0058	0.0061	0.0063	0.0065
	PC3	30.4043	20.1576	13.7867	9.7243	7.0723
		0.0111	0.0063	0.0035	0.0020	0.0013
	PC4	192.091	192.6035	193.8697	195.1105	195.8213
		0.1893	0.1842	0.1964	0.1832	0.1872
1.75	PC1	9.5331	8.587	7.7562	7.0278	6.3915
		0.0020	0.0017	0.0014	0.0013	0.0010
	PC2	12.8624	13.0283	13.1982	13.3635	13.5408
		0.0031	0.0032	0.0031	0.0033	0.0033
	PC3	32.538	21.4606	14.6218	10.259	7.4253
		0.0125	0.0066	0.0039	0.0023	0.0013
	PC4	186.8353	188.369	189.5544	191.5863	192.4798
		0.1808	0.1826	0.1786	0.1837	0.1966
2	PC1	7.1927	6.5333	5.9532	5.4399	4.9876
		0.0012	0.0011	0.0010	0.0008	0.0007
	PC2	8.938	9.0445	9.1467	9.2587	9.3672
		0.0018	0.0018	0.0018	0.0019	0.0018
	PC3	34.8378	22.924	15.5161	10.8357	7.802
		0.0146	0.0076	0.0042	0.0024	0.0015
	PC4	181.7749	183.6495	185.2074	186.4019	188.2948
		0.1767	0.1735	0.1716	0.1767	0.1765

Table 16: ARL comparison for simultaneous I(row) and M(column) shifts.
($\mu_I = 1 + 0.2 \times \alpha$, $\mu_M = 15 + 1 \times \beta$)

$\alpha\backslash\gamma$		0.25	0.5	0.75	1	1.25
1.5	PC1	14.8805	13.2395	11.8106	10.5787	9.4991
		0.0038	0.0033	0.0029	0.0023	0.0020
	PC2	9.108	9.0018	8.8951	8.7957	8.6955
		0.0019	0.0018	0.0018	0.0018	0.0017
	PC3	146.9559	96.3573	60.8213	38.7273	25.3012
		0.1283	0.0654	0.0335	0.0165	0.0089
	PC4	186.3605	187.5394	189.4829	190.5522	192.2426
		0.1789	0.1850	0.1811	0.1892	0.1908
1.75	PC1	9.2862	8.3691	7.5679	6.8644	6.2446
		0.0018	0.0016	0.0014	0.0012	0.0010
	PC2	5.0842	5.036	4.9869	4.941	4.894
		0.0009	0.0007	0.0007	0.0007	0.0007
	PC3	89.8152	56.6432	36.1729	23.7163	16.0212
		0.0630	0.0319	0.0148	0.0081	0.0044
	PC4	128.9885	131.0008	133.2422	135.5495	138.0885
		0.1019	0.1035	0.1079	0.1089	0.1143
2	PC1	5.9526	5.4386	4.9852	4.5818	4.2257
		0.0009	0.0008	0.0007	0.0006	0.0005
	PC2	3.0185	2.9961	2.9734	2.9521	2.9305
		0.0003	0.0003	0.0003	0.0003	0.0003
	PC3	40.6474	26.4623	17.7185	12.2469	8.7252
		0.0180	0.0094	0.0050	0.0029	0.0017
	PC4	62.5511	63.7505	64.95	66.2724	67.6088
		0.0341	0.0365	0.0357	0.0387	0.0387

Table 17: ARL comparison for simultaneous I(row) and N(column) shifts.
$(\mu_I = 1 + 0.2 \times \alpha \ , \ \mu_N = -1.5 + 0.3 \times \gamma)$

$\beta\backslash\gamma$		0.25	0.5	0.75	1	1.25
1	PC1	30.5638	20.689	14.3699	10.2671	7.5407
		0.0114	0.0065	0.0038	0.0022	0.0014
	PC2	50.3047	79.0151	122.9198	174.99932	200.5319
		0.0255	0.0497	0.0923	0.1620	0.2037
	PC3	182.2847	174.6372	166.2471	157.5547	148.3484
		0.1778	0.1639	0.1544	0.1457	0.1243
	PC4	197.3178	199.2802	200.3078	200.1278	198.8937
		0.2038	0.1917	0.2043	0.2017	0.1992
1.25	PC1	18.7555	13.0269	9.328	6.8656	5.1969
		0.0058	0.0031	0.0019	0.0011	0.0007
	PC2	25.1209	37.7346	57.9919	90.1817	136.5484
		0.0087	0.0161	0.0317	0.0575	0.1122
	PC3	195.3024	190.4205	184.3617	177.4459	169.7167
		0.1861	0.1883	0.1771	0.1725	0.1614
	PC4	199.6297	198.3851	195.6613	192.3238	187.6948
		0.1980	0.1996	0.1893	0.1898	0.1784
1.5	PC1	11.6777	8.3743	6.1883	4.704	3.6792
		0.0027	0.0016	0.0010	0.0006	0.0004
	PC2	12.6638	17.8661	25.8734	38.4959	58.4024
		0.0032	0.0053	0.0090	0.0176	0.0296
	PC3	196.6974	198.9704	199.9054	199.5757	197.7061
		0.1938	0.2075	0.2060	0.2060	0.1948
	PC4	178.1878	171.2192	164.4912	156.8125	148.776
		0.1727	0.1536	0.1485	0.1397	0.1289

Table 18: ARL comparison for simultaneous M(row) and N(column) shifts.
$(\mu_M = 15 + 1 \times \beta \ , \ \mu_N = -1.5 + 0.3 \times \gamma)$

σ_I				
chart	0.3	0.4	0.5	0.6
	(1.034)	(1.080)	(1.136)	(1.201)
PC1	92.8076	48.6820	79.2516	49.4240
	0.0629	0.0233	0.0486	0.0246
PC2	46.3985	30.8190	53.2247	70.6521
	0.0218	0.0121	0.0276	0.0405
PC3	21.8815	9.8422	7.4607	3.8953
	0.0073	0.0022	0.0014	0.0005
PC4	172.7439	82.5331	158.2309	279.7896
	0.1601	0.0555	0.1467	0.3339
combined	29.3373	12.0996	10.4196	5.0304
	0.0115	0.0029	0.0022	0.0007
chart	0.7	0.8	0.9	1
	(1.273)	(1.352)	(1.437)	(1.525)
PC1	58.3326	16.5762	18.5192	16.2502
	0.0318	0.0046	0.0053	0.0045
PC2	61.6944	57.2272	51.1069	41.3881
	0.0342	0.0292	0.0249	0.0187
PC3	3.5062	2.5973	2.4885	2.0873
	0.0004	0.0002	0.0002	0.0002
PC4	167.1431	137.9471	170.3432	139.4717
	0.1464	0.1140	0.1561	0.1174
combined	4.4020	3.0447	2.8854	2.3577
	0.0006	0.0003	0.0003	0.0002

Table 19: ARL comparison for I-shift in variance.

σ_M				
chart	1.5	2	2.5	3
	(1.224)	(1.481)	(1.757)	(2.045)
PC1	33.3123	10.6511	7.7653	4.4383
	0.0132	0.0024	0.0014	0.0006
PC2	16.4597	6.7316	5.1548	3.3771
	0.0047	0.0011	0.0007	0.0004
PC3	86.1611	139.2719	69.8514	115.3932
	0.0576	0.1160	0.0417	0.0886
PC4	155.9135	73.3568	119.8870	61.2410
	0.1387	0.0470	0.0912	0.0357
combined	20.7851	7.0708	5.2722	3.2934
	0.0065	0.0012	0.0007	0.0003
chart	3.5	4	4.5	5
	(2.340)	(2.640)	(2.942)	(3.248)
PC1	3.9680	2.8805	2.5122	2.8277
	0.0005	0.0003	0.0002	0.0003
PC2	3.0209	2.3876	2.1357	2.2428
	0.0003	0.0002	0.0002	0.0002
PC3	54.7949	89.3402	95.7994	136.0448
	0.0286	0.0604	0.0653	0.1155
PC4	87.2543	49.0418	64.6406	142.1053
	0.0576	0.0239	0.0371	0.1195
combined	2.9894	2.8168	2.1886	2.0442
	0.0003	0.0003	0.0002	0.0002

Table 20: ARL comparison for M-shift in variance.

σ_N				
chart	0.4	0.5	0.6	0.7
	(1.153)	(1.323)	(1.506)	(1.670)
PC1	28.7742	9.4391	8.5725	3.0988
	0.0107	0.0019	0.0017	0.0003
PC2	11.5527	4.3324	4.3710	2.1099
	0.0027	0.0006	0.0006	0.0002
PC3	4.0039	2.2513	2.5591	1.0304
	0.0005	0.0002	0.0002	0.00001
PC4	42.1725	6.7205	10.8301	1.0207
	0.0190	0.0012	0.0024	0.00001
combined	4.4747	2.3922	2.1989	1.0057
	0.0006	0.0002	0.0002	0.00001
chart	0.8	0.9	1	1.1
	(1.893)	(2.093)	(2.270)	(2.502)
PC1	1.5877	1.1751	1.3701	1.0018
	0.0001	0.00003	0.00006	0.00001
PC2	1.7362	1.5663	1.3526	1.0008
	0.0001	0.0001	0.00006	0.00001
PC3	1.0072	1.0021	1.0057	1.0000
	0.00001	0.00001	0.00001	0.00001
PC4	1.0037	1.0010	1.0030	1.0000
	0.00001	0.00001	0.00001	0.00001
combined	1.0004	1.0000	1.0001	1.0000
	0.00001	0.00001	0.00001	0.00001

Table 21: ARL comparison for N-shift in variance.

k					
chart	0	0.25	0.5	0.75	1
PC1	201.141	157.7244	91.6269	50.1074	28.1935
	0.6537	0.4263	0.1869	0.0773	0.0339
PC2	202.5093	201.1011	202.7793	201.3586	202.3857
	0.6897	0.6687	0.6295	0.6704	0.6432
PC3	202.2185	202.1651	201.4955	203.475	202.0911
	0.6937	0.6648	0.6495	0.6207	0.6461
PC4	202.4143	200.7331	201.7661	202.0241	202.1318
	0.6594	0.6526	0.6303	0.6551	0.6248
chart	1.25	1.5	1.75	2	
PC1	16.7264	10.4605	6.8874	4.7689	
	0.0153	0.0071	0.0037	0.0020	
PC2	202.5093	201.1011	202.7793	201.3586	
	0.6897	0.6687	0.6295	0.6704	
PC3	202.2185	202.1651	201.4955	203.475	
	0.6937	0.6648	0.6495	0.6207	
PC4	202.4143	200.7331	201.7661	202.0241	
	0.6594	0.6526	0.6303	0.6551	

Table 22: ARL comparison for shifts in principal component 1.

k					
chart	0	0.25	0.5	0.75	1
PC1	203.1434	202.2938	202.2304	201.9739	202.0307
	0.6599	0.6496	0.6643	0.6826	0.6423
PC2	202.3857	157.2349	91.368	50.138	28.2533
	0.6432	0.4475	0.1938	0.0783	0.0318
PC3	202.0911	201.3132	202.4239	202.4192	201.6323
	0.6461	0.6695	0.6455	0.6580	0.6622
PC4	202.1318	201.4323	201.5589	203.1134	201.6549
	0.6248	0.6797	0.6407	0.6532	0.6349
chart	1.25	1.5	1.75	2	
PC1	200.6637	202.6417	202.2779	201.5532	
	0.6327	0.6492	0.6344	0.6624	
PC2	16.75	10.455	6.8888	4.771	
	0.0141	0.0073	0.0038	0.0020	
PC3	201.7072	202.043	202.6156	202.3369	
	0.6355	0.6418	0.6803	0.6310	
PC4	203.0749	202.5458	202.4285	201.5443	
	0.6780	0.6515	0.6505	0.6427	

Table 23: ARL comparison for shifts in principal component 2.

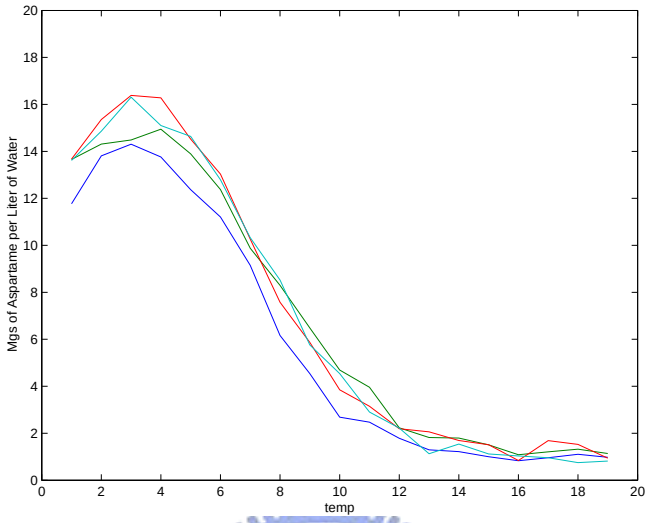

Figure 1: Four hypothetical aspartame profiles.

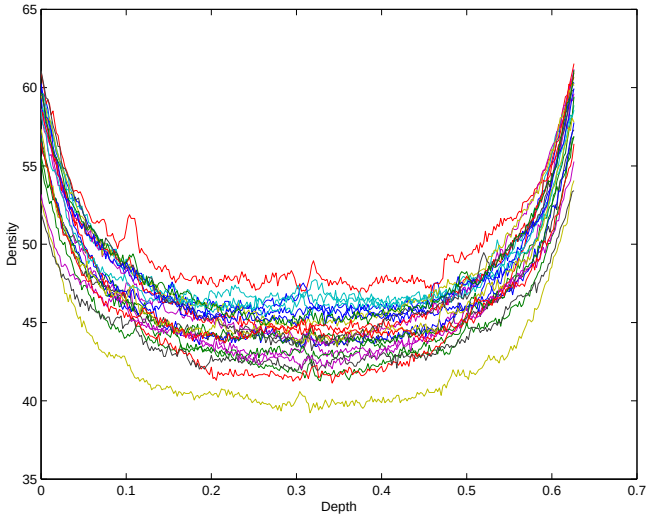

Figure 2: Original 24 VDP–profiles.

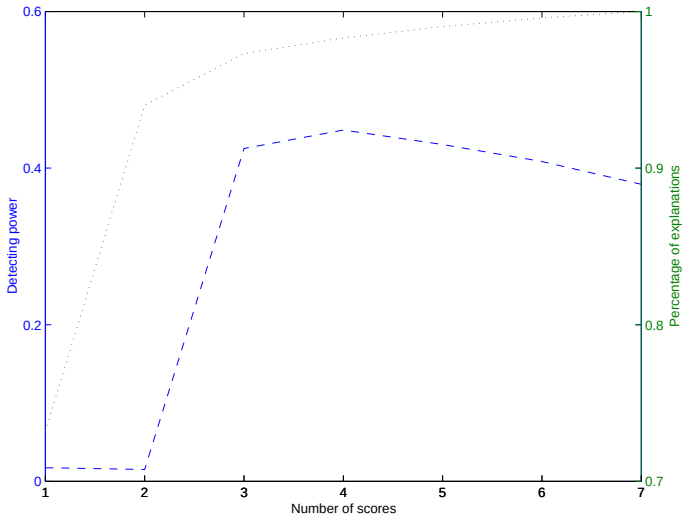

Figure 3: The relation between the detecting power and the percentage of explanation in T_1^2 chart. In 50 curves, there is one outlier with I shift from 1 to 1+5×0.2.

detecting power: '- -', percentage of explanation: '—'

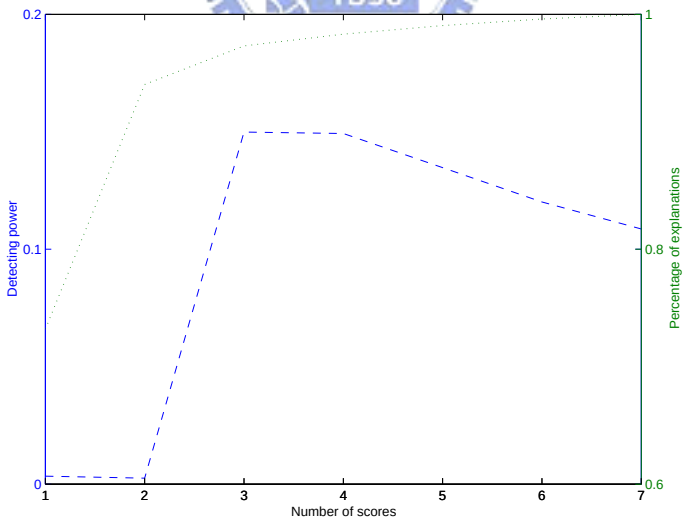

Figure 4: The relation between the detecting power and the percentage of explanation in T_2^2 chart. In 50 curves, there is one outlier with I shift from 1 to 1+5×0.2.

detecting power: '- -', percentage of explanation: '—'

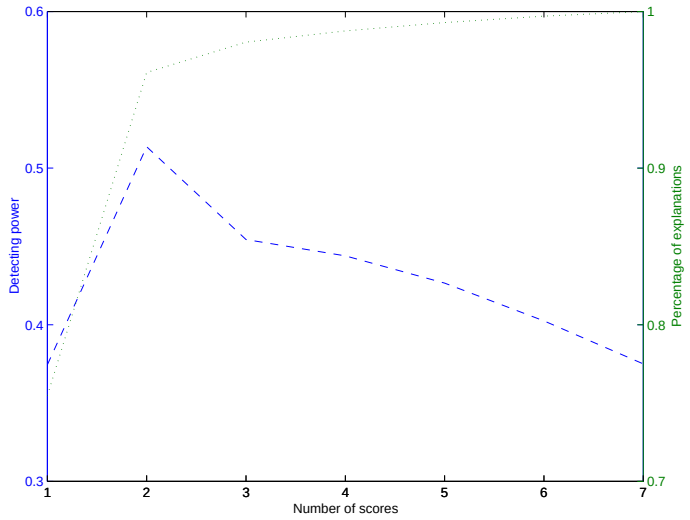

Figure 5: The relation between the detecting power and the percentage of explanation in T_1^2 chart. In 50 curves, there are two outliers with M shift from 15 to 15+5×1.

detecting power: '- -', percentage of explanation: '—'

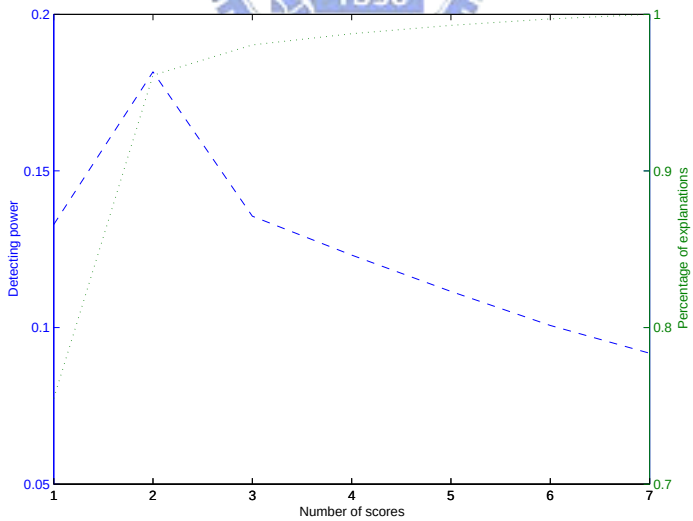

Figure 6: The relation between the detecting power and the percentage of explanation in T_2^2 chart. In 50 curves, there are two outliers with M shift from 15 to 15+5×1.

detecting power: '- -', percentage of explanation: '—'

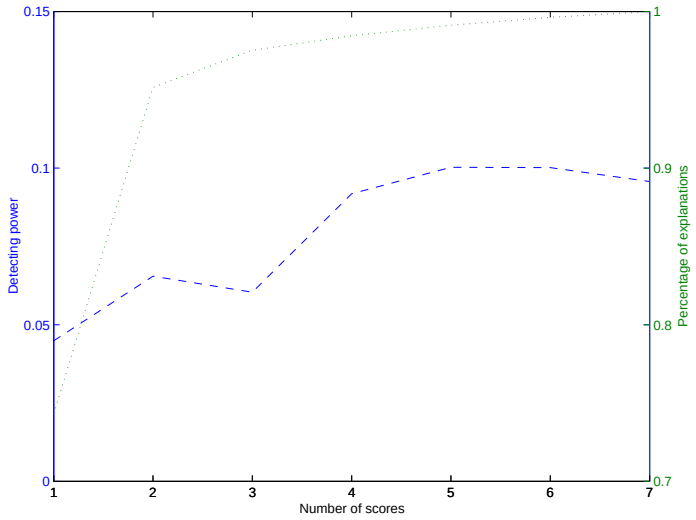

Figure 7: The relation between the detecting power and the percentage of explanation in T_1^2 chart. In 50 curves, there are three outliers with N shift from -1.5 to -1.5+2×0.3.

detecting power: '- -', percentage of explanation: '—'

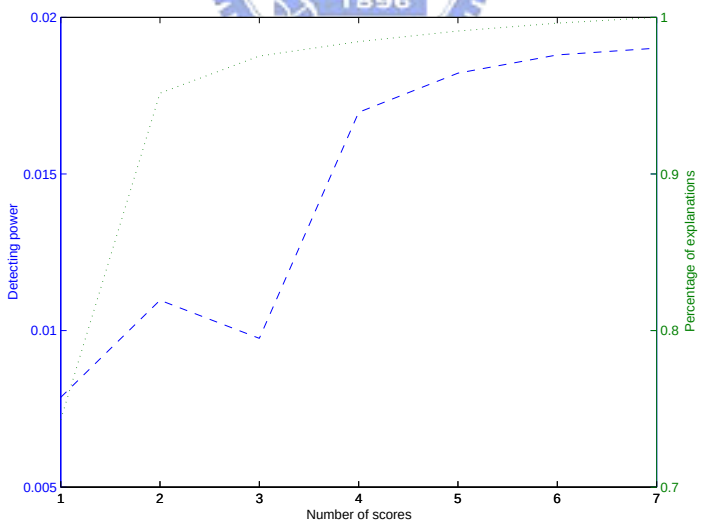

Figure 8: The relation between the detecting power and the percentage of explanation in T_2^2 chart. In 50 curves, there are three outliers with N shift from -1.5 to -1.5+2×0.3.

detecting power: '- -', percentage of explanation: '—'

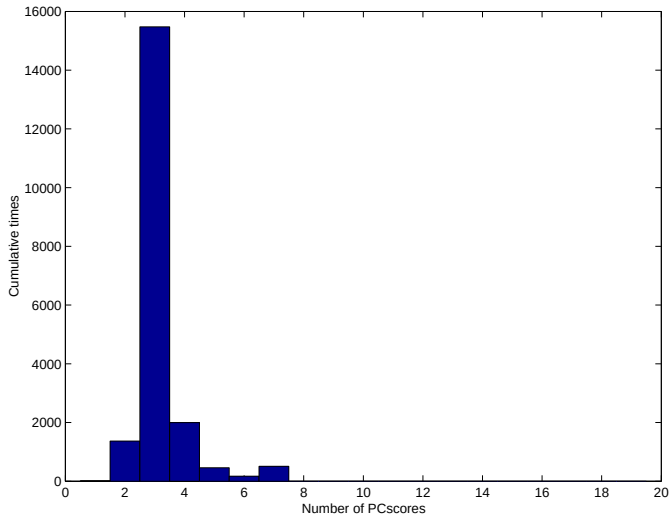

Figure 9: The cross-validation results for one outlier with a shift of five sigma in I among fifty profiles.

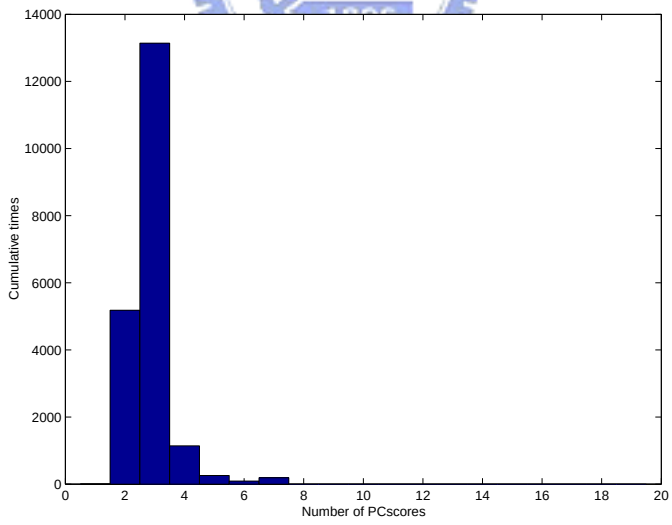

Figure 10: The cross-validation results for two outliers with a shift of five sigma in M among fifty profiles.

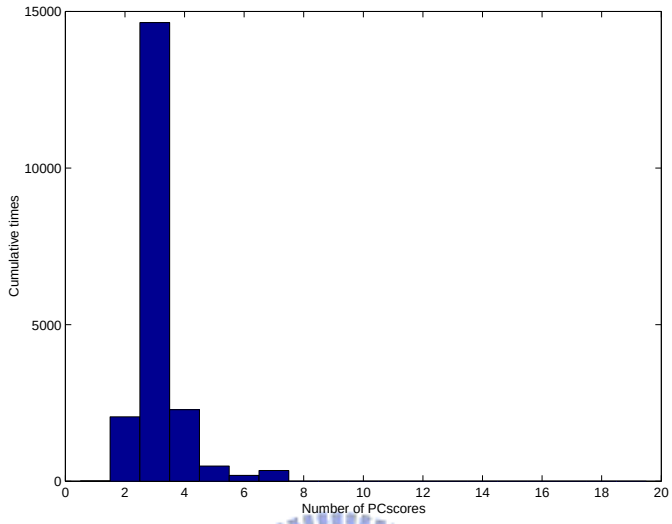

Figure 11: The cross-validation results for three outliers with a shift of two sigma in N among fifty profiles.

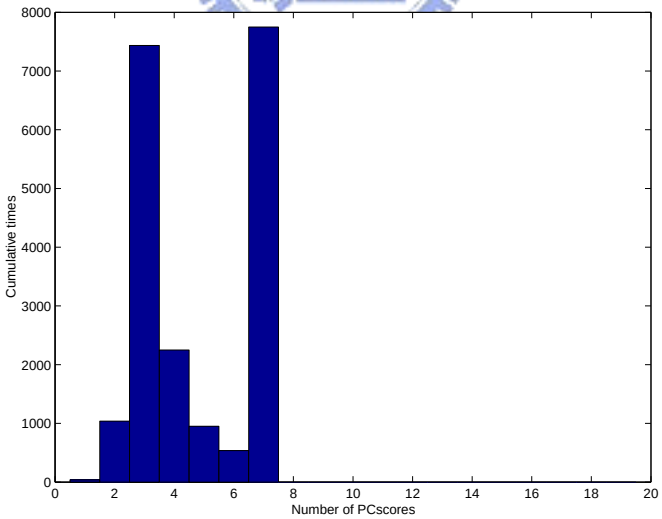

Figure 12: The resulting histogram for two-fold cross-validation.

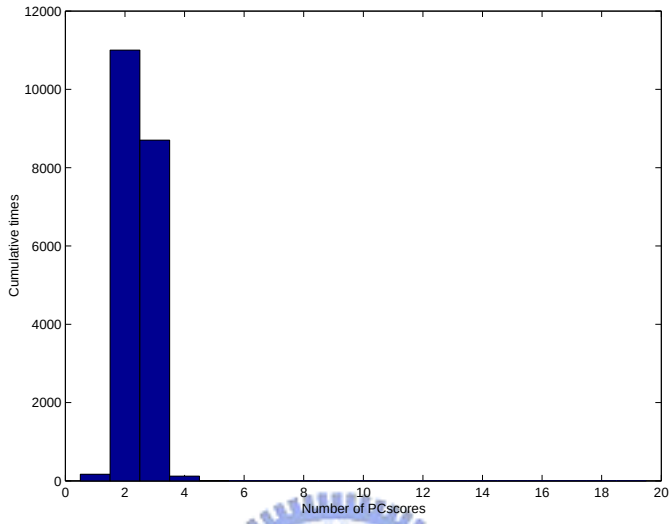

Figure 13: The resulting histogram for ten-fold cross-validation.

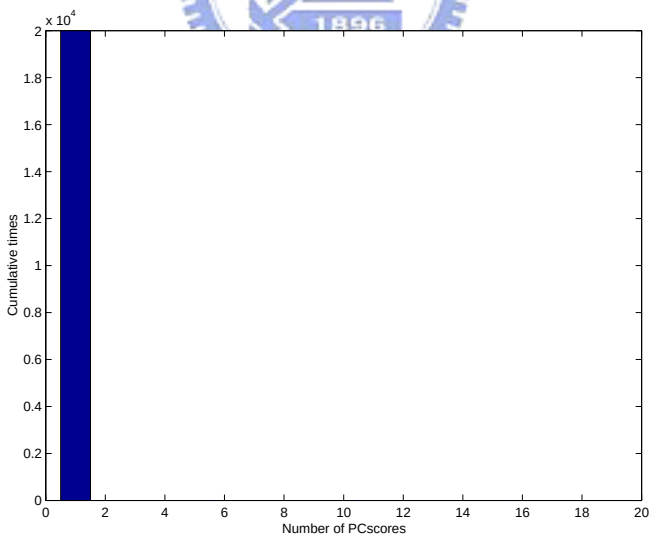

Figure 14: The resulting histogram for delete-one cross-validation.

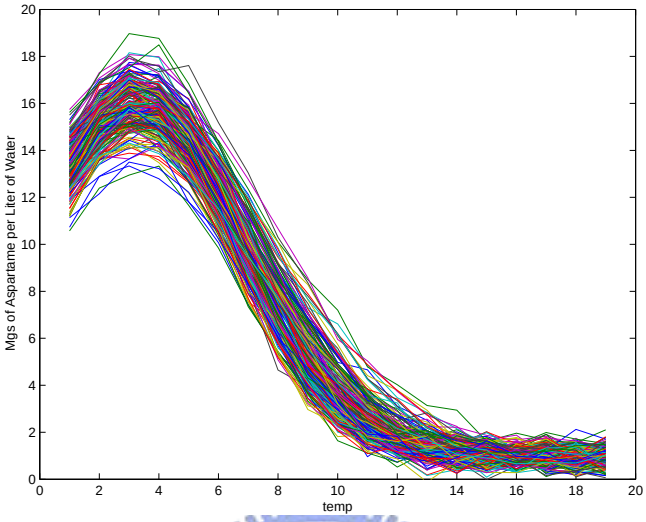

Figure 15: Original 200 aspartame profiles.

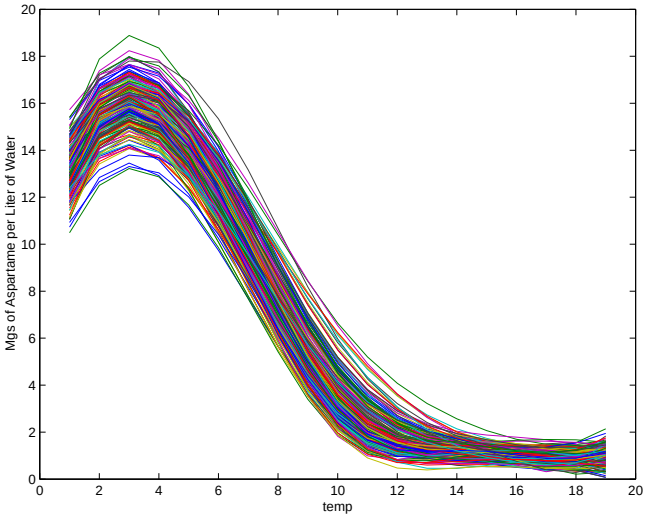

Figure 16: Fitted B-splines with d.f.=5 .

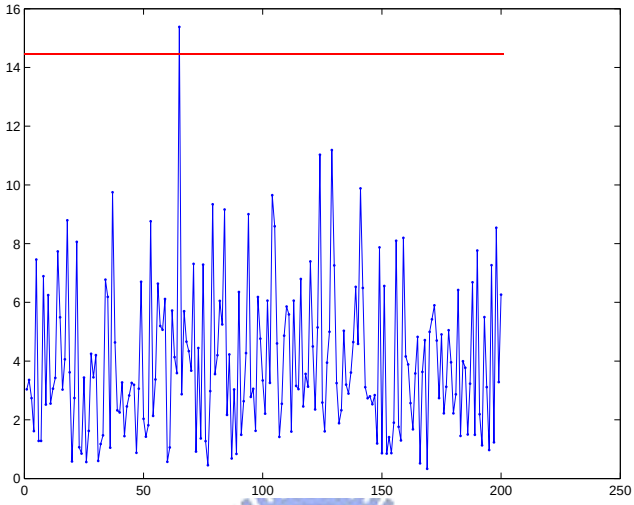

Figure 17: T_1^2 control chart.

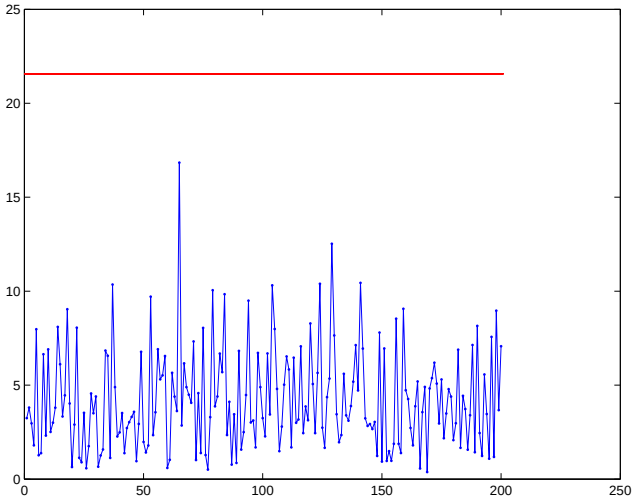

Figure 18: T_2^2 control chart.

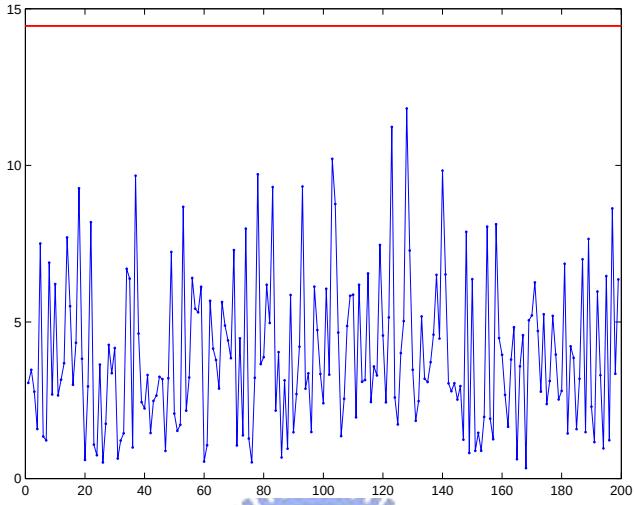

Figure 19: T_1^2 control chart after the out-of-control profile was removed.

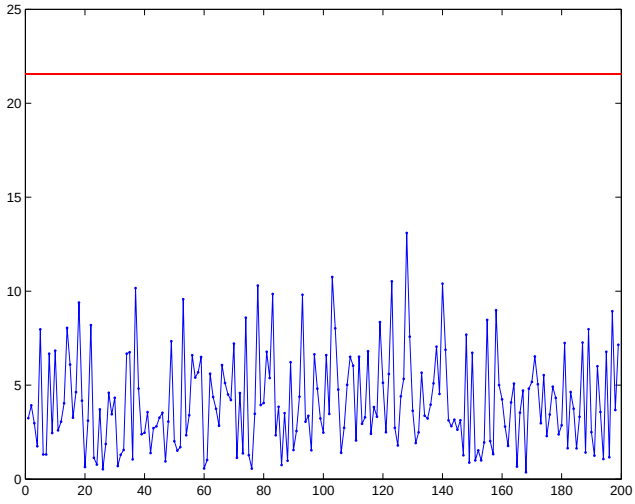

Figure 20: T_2^2 control chart after the out-of-control profile was removed.

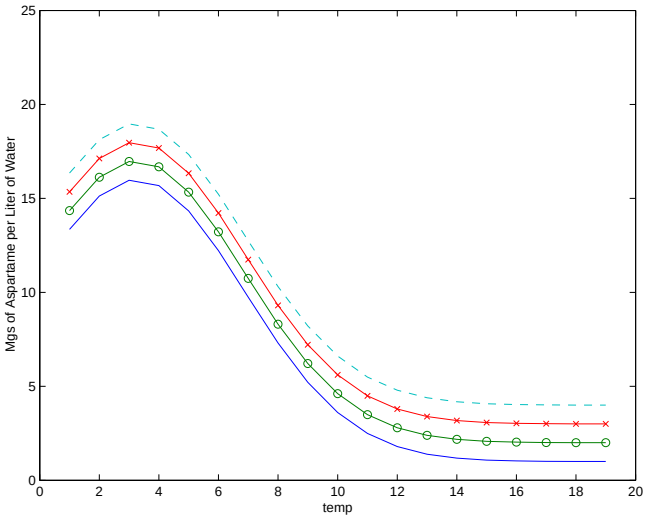

Figure 21: Shifts in I.

I=1: '—', I=2: 'o-', I=3: 'x-', I=4: '- -'

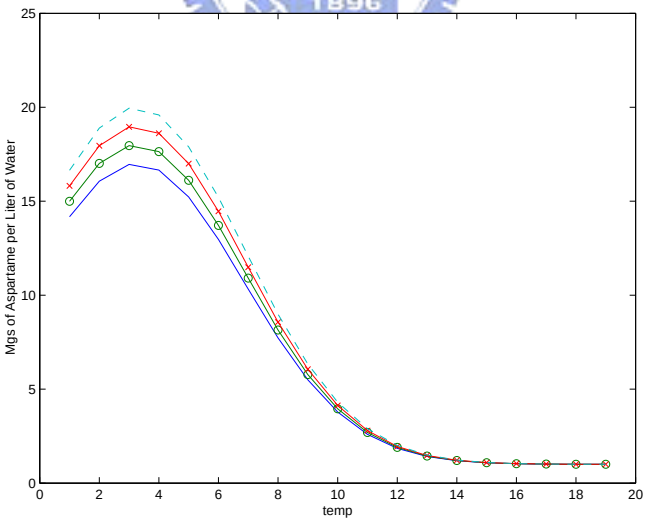

Figure 22: Shifts in M.

M=16: '—', M=17: 'o-', M=18: 'x-', M=19: '- -'

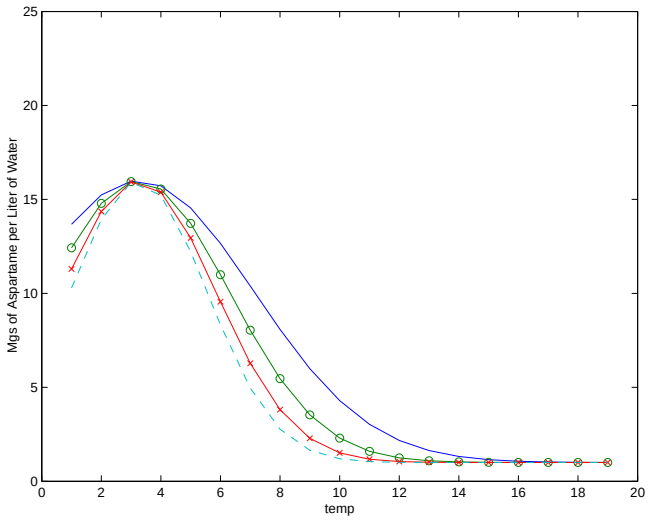

Figure 23: Shifts in N.

N=-1.3: '—', N=-2.1: 'o-', N=-2.9: 'x-', N=-3.7: '- -'

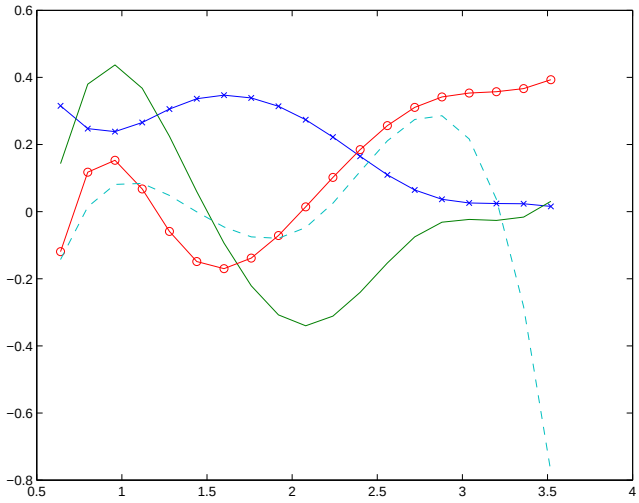

Figure 24: First four eigenvectors.

1st eigenvector: 'x-', 2nd: '—', 3rd: 'o-', 4th: '- -'

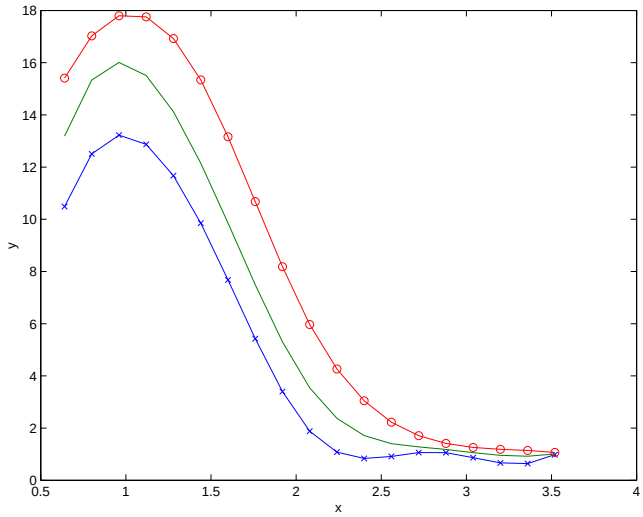

Figure 25: Principal component 1.
smallestscore: 'x-', meanprofile: '—', largestscore: 'o-'

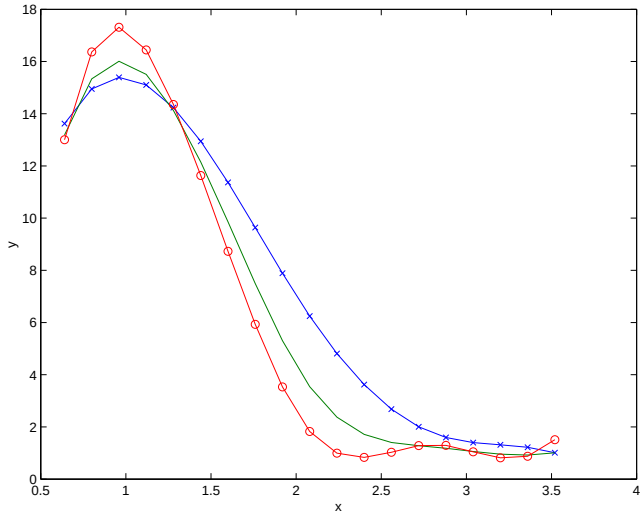

Figure 26: Principal component 2.
smallestscore: 'x-', meanprofile: '—', largestscore: 'o-'

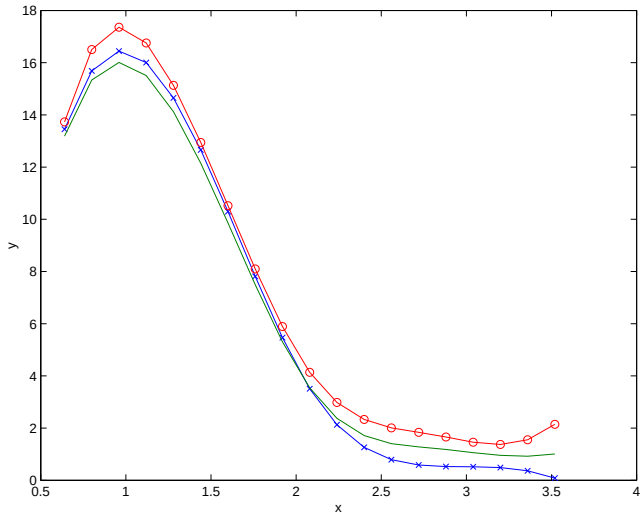

Figure 27: Principal component 3.
smallestscore: 'x-', meanprofile: '—', largestscore: 'o-'

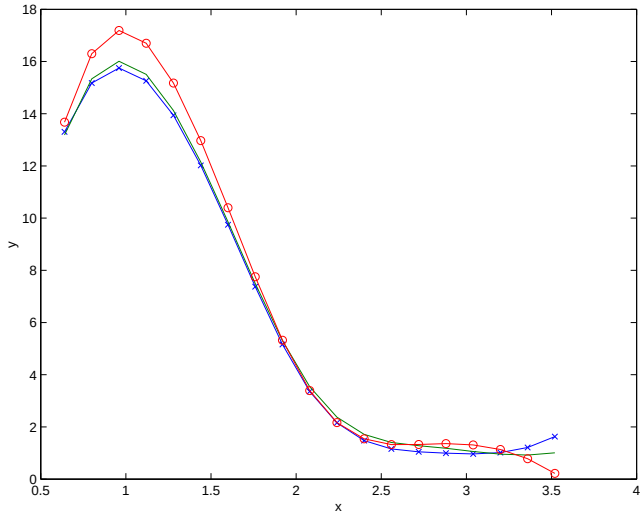

Figure 28: Principal component 4.
smallestscore: 'x-', meanprofile: '—', largestscore: 'o-'

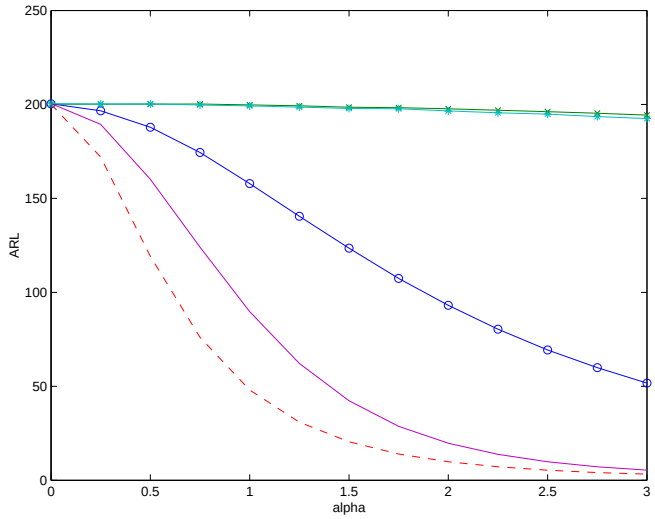

Figure 29: ARL comparison for I shifts in mean.
PC1: 'o-', PC2: 'x-', PC3: '- -', PC4: '*-', Combined: '—'

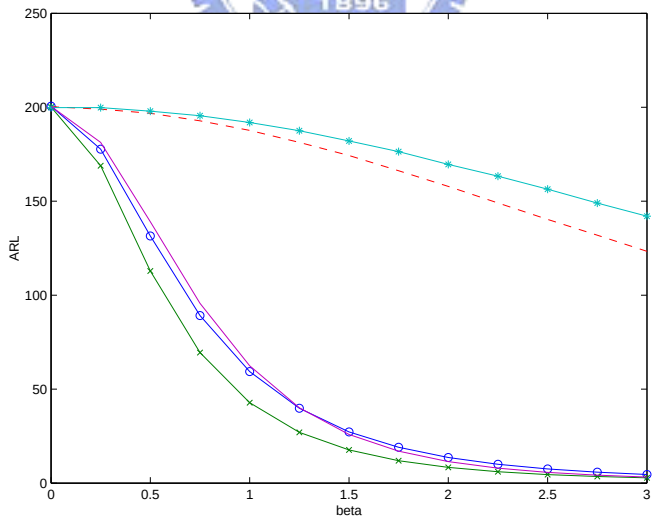

Figure 30: ARL comparison for M shifts in mean.
PC1: 'o-', PC2: 'x-', PC3: '- -', PC4: '*-', Combined: '—'

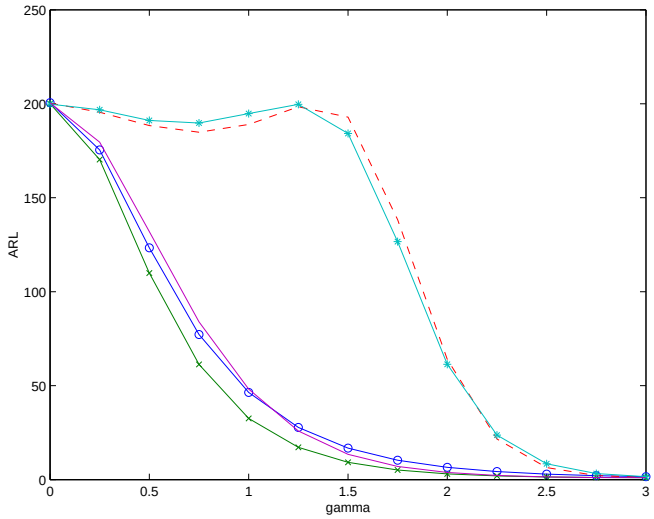

Figure 31: ARL comparison for N shifts in mean.
PC1: 'o-', PC2: 'x-', PC3: '- -', PC4: '*-', Combined: '—'

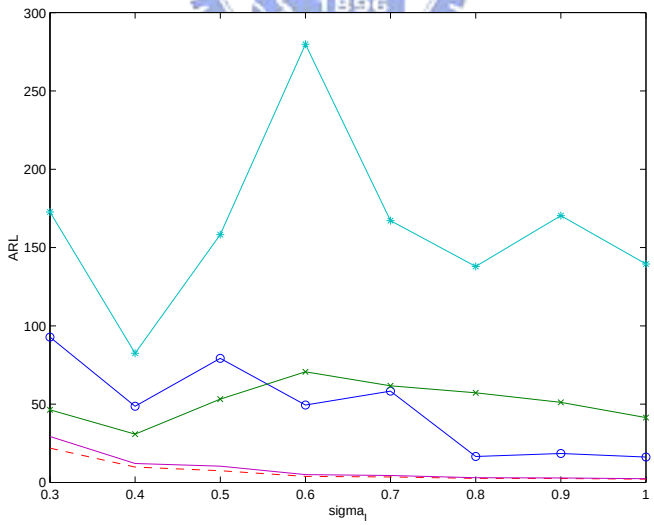

Figure 32: ARL comparison for I shifts in variance.
PC1: 'o-', PC2: 'x-', PC3: '- -', PC4: '*-', Combined: '—'

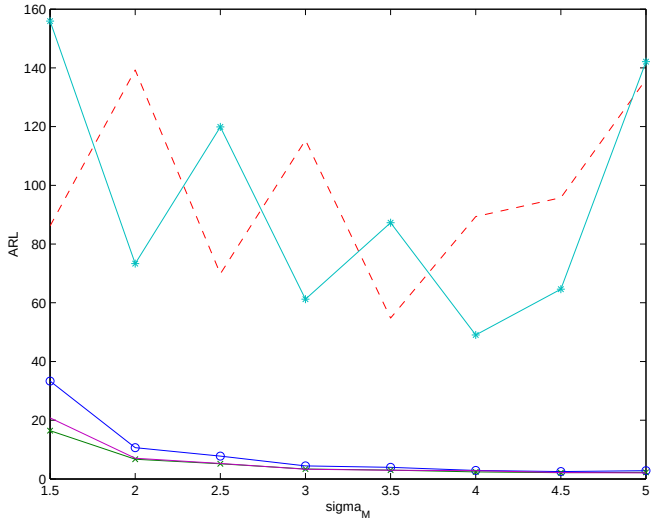

Figure 33: ARL comparison for M shifts in variance.
PC1: 'o-', PC2: 'x-', PC3: '- -', PC4: '*-', Combined: '—'

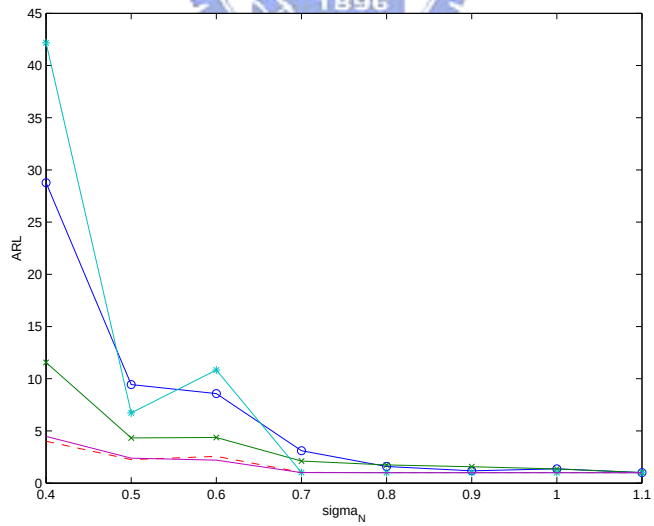

Figure 34: ARL comparison for N shifts in variance.
PC1: 'o-', PC2: 'x-', PC3: '- -', PC4: '*-', Combined: '—'

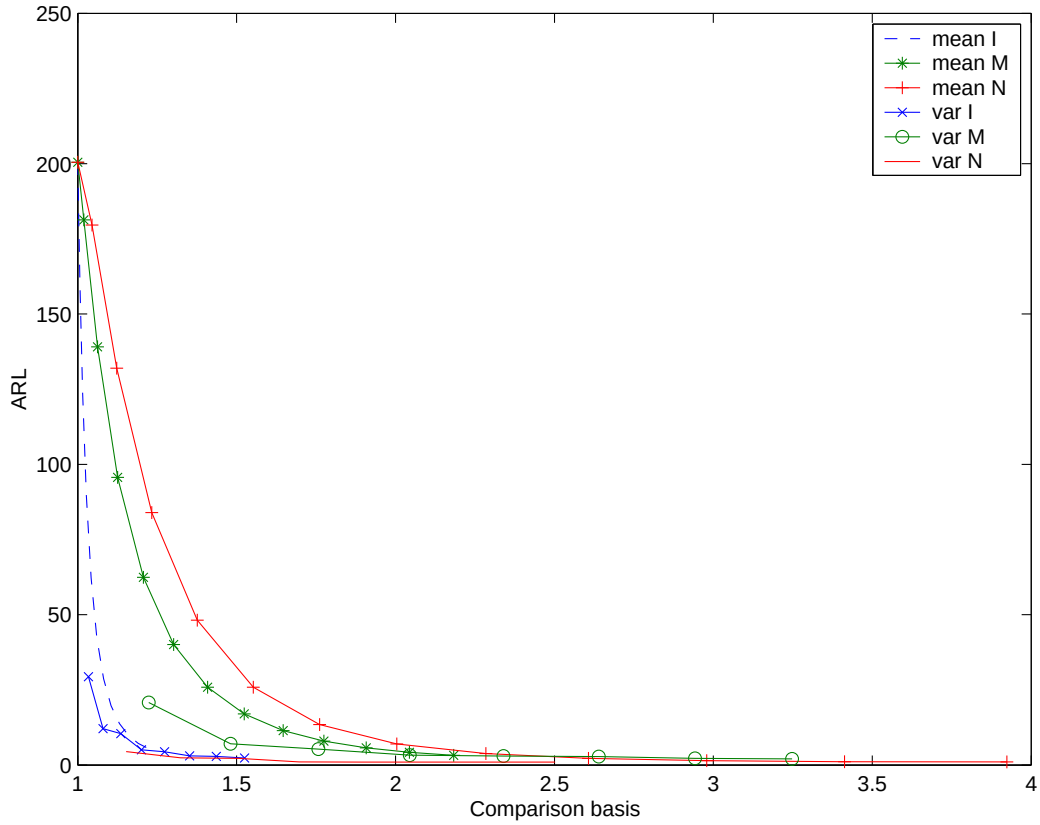

Figure 35: ARL comparison for mean shifts and variance shifts for the combined chart.

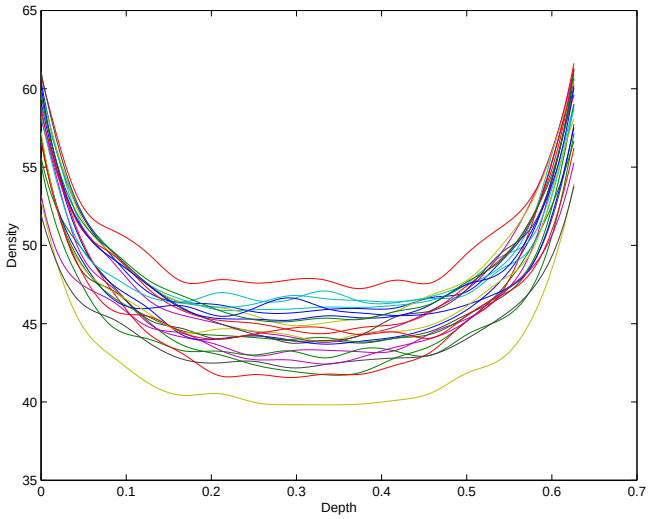

Figure 36: Fitted B-splines with d.f.=16.

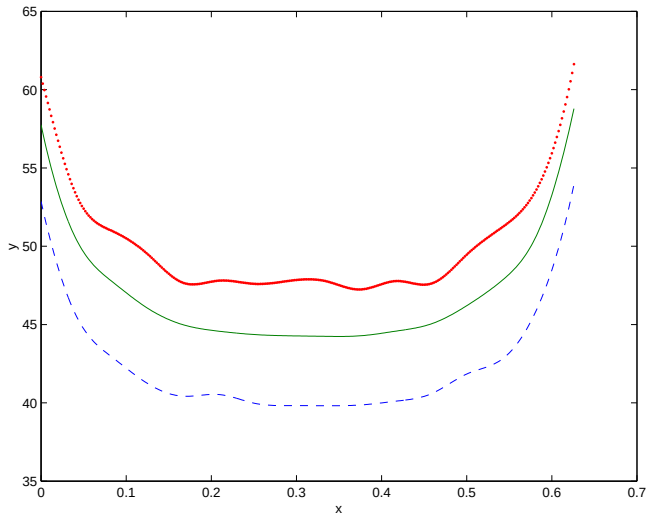

Figure 37: Principal component 1.
smallestscore: '- -', meanprofile: '—', largestscore: '.'

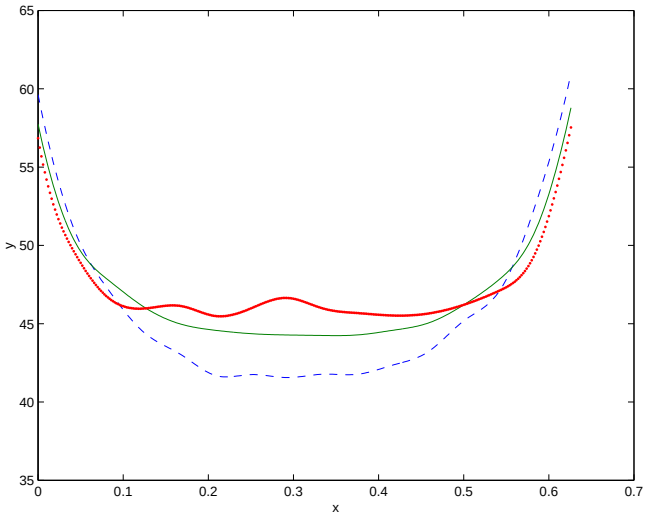

Figure 38: Principal component 2.
smallestscore: '- -', meanprofile: '—', largestscore: '.'

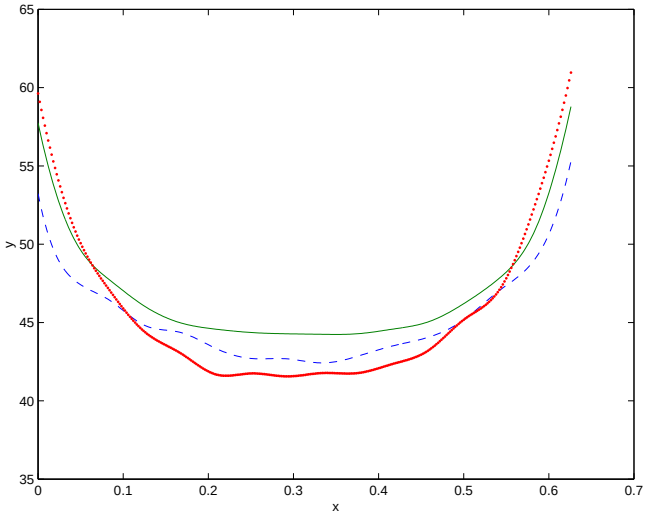

Figure 39: Principal component 3.
smallestscore: '- -', meanprofile: '—', largestscore: '.'

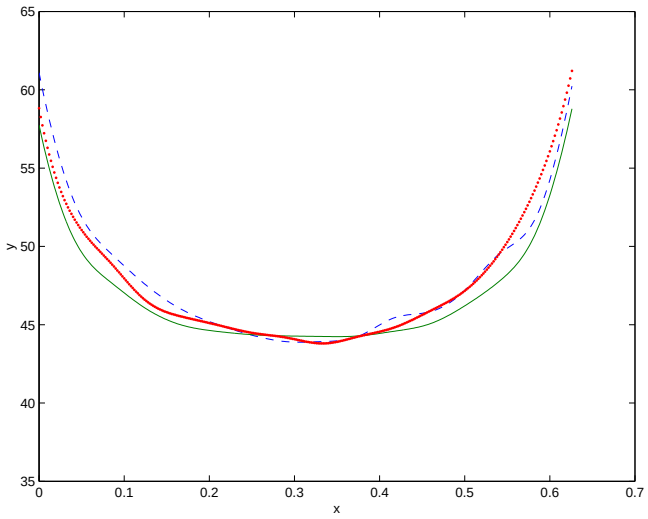

Figure 40: Principal component 4.
smallestscore: '- -', meanprofile: '—', largestscore: '.'

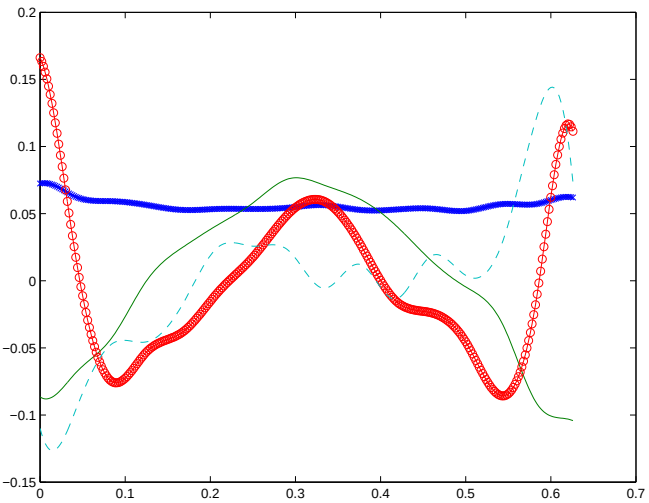

Figure 41: First four eigenvectors.
1st eigenvector: 'x-', 2nd: '—', 3rd: 'o-', 4th: '- -'

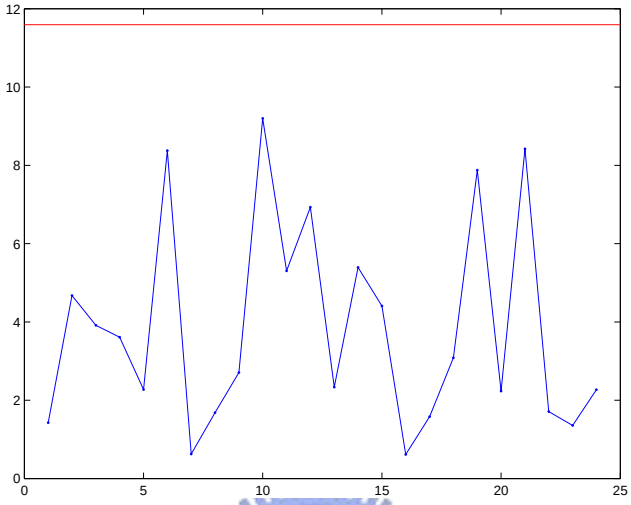

Figure 42: T_1^2 control chart.

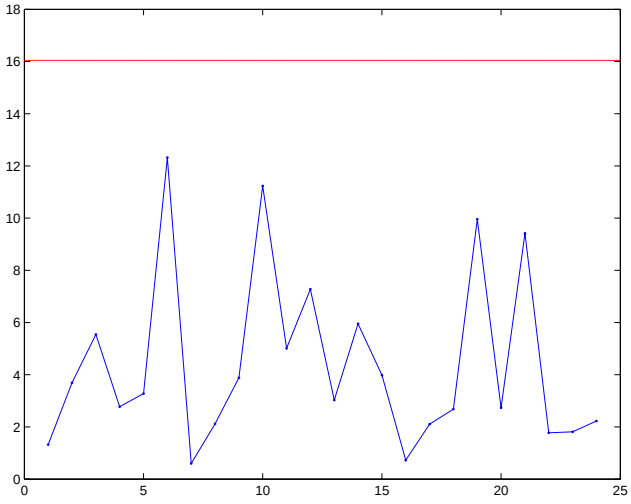

Figure 43: T_2^2 control chart.

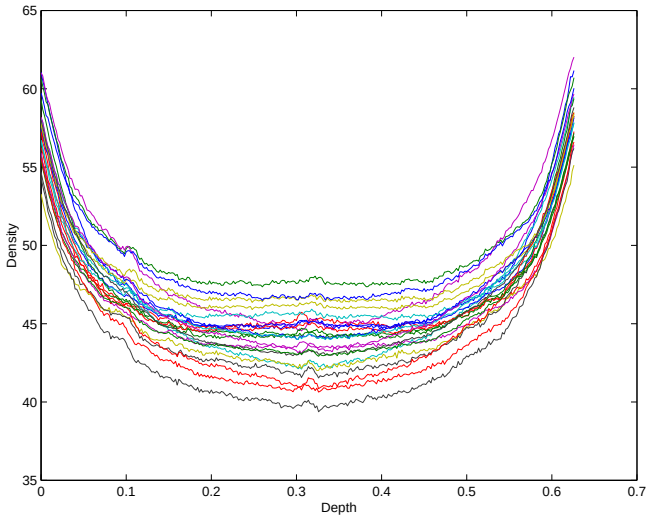

Figure 44: Simulated 24 VDP–profiles.

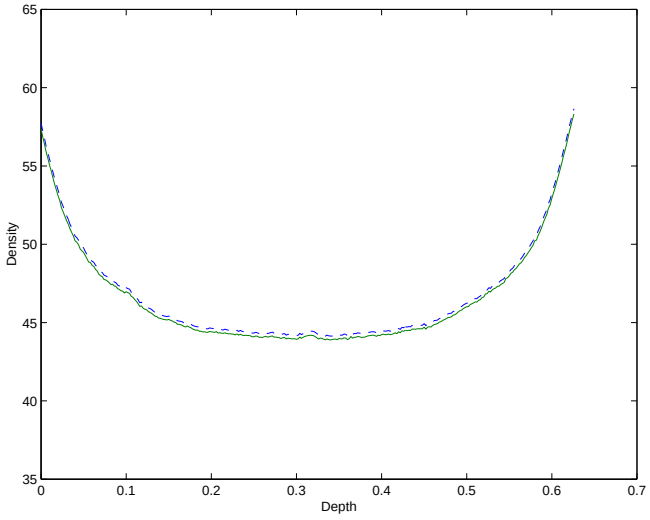

Figure 45: Comparison of the real mean profile and the mean profile estimate of the simulated data.

real profile: '- -', simulated profile: '—'

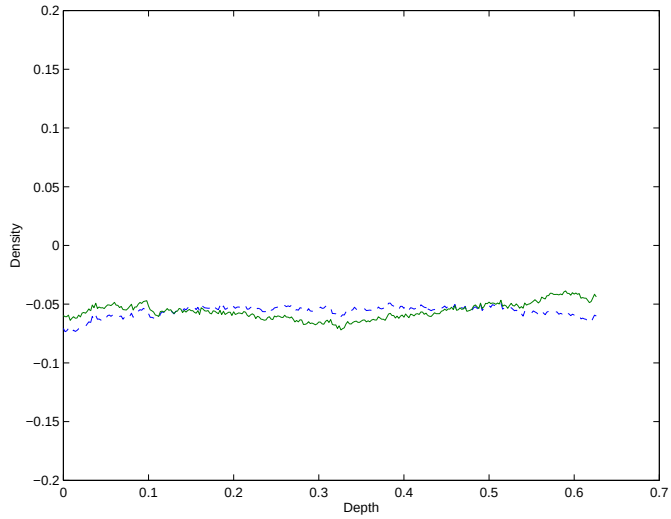

Figure 46: Comparison of the real profiles and the simulated profiles for the first eigenvector.

real profile: '- -', simulated profile: '—'

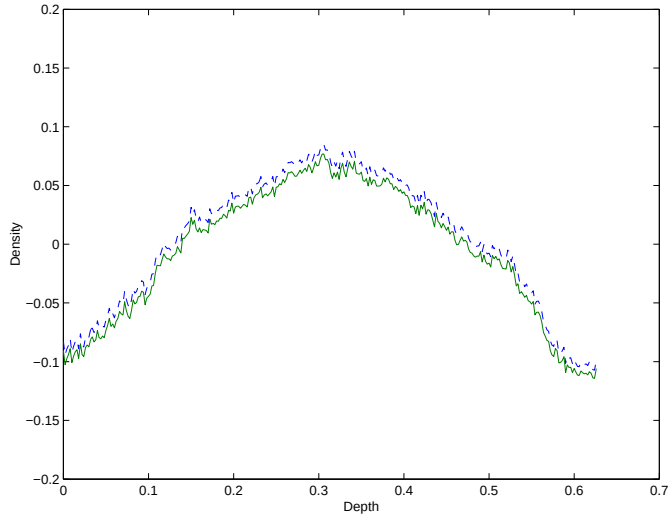

Figure 47: Comparison of the real profiles and the simulated profiles for the second eigenvector.

real profile: '- -', simulated profile: '—'

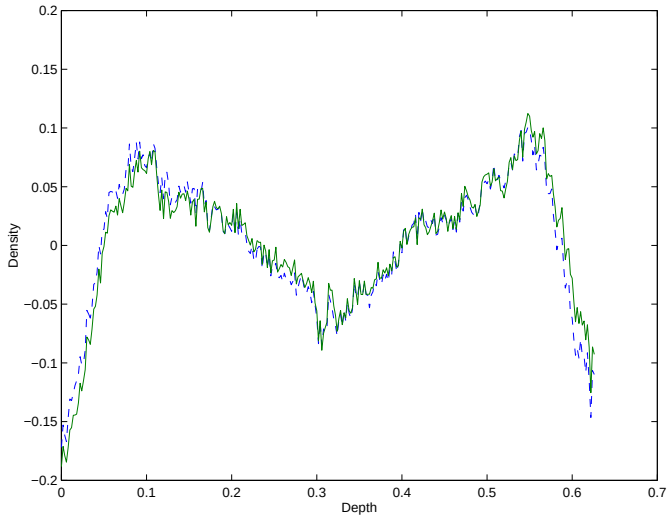

Figure 48: Comparison of the real profiles and the simulated profiles for the third eigenvector.

real profile: '- -', simulated profile: '—'

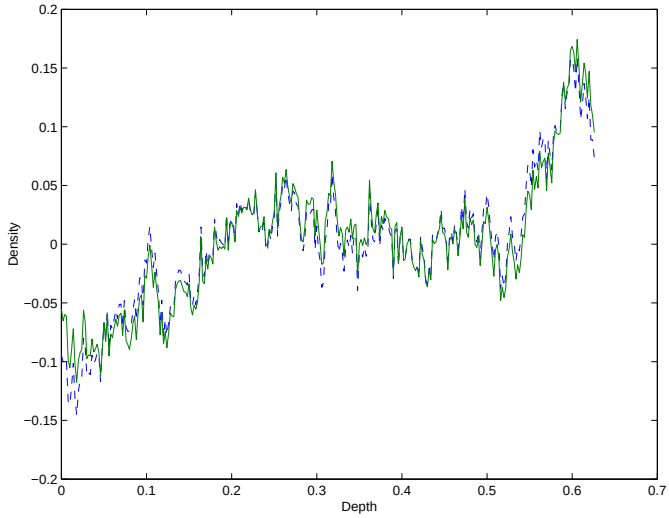

Figure 49: Comparison of the real profiles and the simulated profiles for the fourth eigenvector.

real profile: '- -', simulated profile: '—'

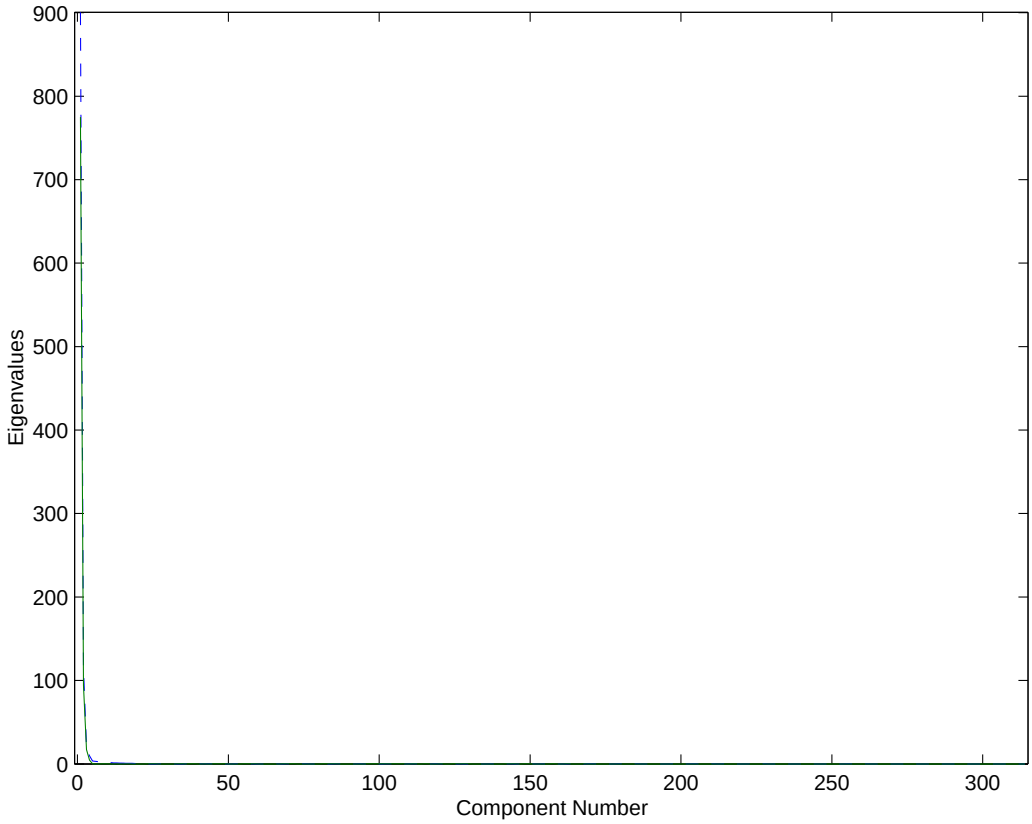

Figure 50: Comparison of the real eigenvalues and the simulated ones. real profile: '- -', simulated profile: '—'