國 立 交 通 大 學

統 計 學 研 究 所

碩士論文

利用 Oja 資料縱深測度建構剖面資料之無母數監控方法

Profile Monitoring via Oja Data Depth

研 究 生：王新樺

指導教授：洪志真　博士

中華民國 九十八 年 六 月

利用 Oja 資料縱深測度建構剖面資料之無母數監控方法

Profile Monitoring via Oja Data Depth

研 究 生：王新樺　　Student: Shin-Hua Wang
指導教授：洪志真 博士　　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 2009
Hsinchu, Taiwan, Republic of China

中華民國 九十八 年 六 月

利用 Oja 資料縱深測度建構剖面資料之無母數監控方法

學生：王新樺　　指導教授：洪志真　博士

國立交通大學統計學研究所

摘　　要

對於品質特性可以藉由剖面資料加以描述的製程，基於資料縱深測度，我們提出一些無母數製程監控方法。我們首先回顧剖面資料監控、數據平滑化、主成分分析、階層式集群方法與資料縱深測度管制圖等文獻。在此我們聚焦於階段二的非線性剖面資料監控。對於所提出的製程監控方法，其最大優點在於不需要對剖面資料做有母數形式之分配假設。我們根據平均連串長度，透過模擬的方式與現存藉由無母數迴歸且在高斯分配假設下所發展出來的監控方法做比較。對於階段一的分析，我們提出一個具啟發性的診斷程序藉此去除潛在失控的剖面資料。在此透過一組生物鑑定實驗資料與一組垂直密度剖面資料(VDP)來展示我們所提出的診斷程序。

關鍵字：剖面、數據平滑化、主成分分析、階層式集群方法、資料縱深測度、平均連串長度。

Profile Monitoring via Oja Data Depth

Student: Shin-Hua Wang Advisor: Dr. Jyh-Jen Horng Shiau

Institute of Statistics

National Chiao Tung University

Abstract

We propose and study some nonparametric process monitoring schemes based on data depth for processes where the quality of a product could be characterized by a function or a profile. We start by reviewing some literatures on profile monitoring, data smoothing, principal component analysis, hierarchical clustering methods, and data-depth-based control charts. We focus on Phase II profile monitoring for non-linear profiles with random effects. The great advantage of the proposed monitoring schemes is that neither a parametric form nor distributional assumptions about the profiles are required. We compare the proposed schemes with some existing schemes developed via nonparametric regression but under Gaussian assumption in terms of the average run length by simulation. For Phase I analysis, we provide a heuristic diagnosis procedure for screening out potential out-of-control profiles. The bioassay data and vertical density profile (VDP) data are used to demonstrate the diagnostic scheme.

Key Words: profile, data smoothing, principal component analysis, hierarchical clustering, data depth, average run length.

誌　　謝

兩年的碩士班生活，十八年的求學生涯，即將在此刻暫時劃上句點。

在此由衷地感謝我的指導教授　洪志真老師，有了老師細心、耐心的指導，不厭其煩地為我解決疑惑，這篇論文才能夠順利完成。謝謝口試委員楊照崑老師、黃榮臣老師以及陳志榮老師，由於老師們對此篇論文的指正與建議，使整篇論文能夠更加完整、充實。

謝謝身邊的同學、朋友們，和你們一起成長、學習的感覺真的很棒，情緒低落時有人分享，遇到問題時一起討論，因為你們，我不是孤軍奮戰，有你們在真好。

最後謝謝一直陪伴著我的家人們，有你們的支持，讓我求學的一路上沒有後顧之憂，讓我知道有一個溫暖港口隨時讓我停靠休憩，謝謝你們，我最愛的家人們。

在此，將本論文獻給我的師長、家人和朋友們，致上我最誠摯的謝意，能和你們分享成果與喜悅是我最榮幸、快樂的事情。

新樺 於交通大學統計學研究所

中華民國九十八年六月

Contents

List of Tables

List of Figures

1 Introduction

Statistical process control (SPC) has been widely used for process improvement in many fields, especially in industries. The control chart is a primary and powerful tool in SPC for monitoring the quality of processes or products. Classical control charts are mostly developed for applications where the quality characteristic is a (univariate or multivariate) random *variable* measured from the processes or products of interest. However, in many applications, the quality characteristic is a *function* of one or more explanatory variables rather than a variable. This functional quality characteristic is often called a *profile*.

Classical control charting consists of two phases. In Phase I, the main objective is to quickly guide the process into an in-control state and collect some in-control data to characterize (or model) the in-control process, which would be used in Phase II to construct reliable control limits for effective on-line monitoring of future production.

In this study, we propose and study a new "nonparametric" control scheme based on the notion of data depth for profile monitoring. The term "nonparametric" here means we do not assume specific distributional properties on profile data. The only assumption required is that the underlying functions of profiles are smooth. Our focus is mainly on developing control schemes for Phase II profile monitoring. The control schemes are constructed by employing statistical methods to profile data, including data smoothing, principal component analysis, and data depth. For Phase I analysis, it is difficult to develop a control chart based on data depth. Nevertheless, we will describe a heuristic diagnosis procedure for screening out potential out-of-control profiles in the historical data set collected in Phase I. For this purpose, clustering techniques may be useful here.

The rest of this paper is organized as follows. Section 2 provides a brief literature review on profile monitoring including techniques for linear/nonlinear profiles with fixed-effects/random-effects as well as the techniques we use in constructing the proposed profile monitoring schemes. Section 3 presents the proposed Phase II profile monitoring schemes. Section 4 describes a screening procedure for finding out-of-control profiles in the Phase I sample. Section 5 gives the results of our simulation and comparative studies. Section 6 concludes this paper with a brief summary and some remarks.

2 Literature Review

In this section, we first review recent works on profile monitoring including linear and nonlinear profiles in the literature. Then we review the statistical techniques that we use in developing the proposed monitoring schemes in this study.

2.1 Profile Monitoring

2.1.1 Linear Profiles

Kang and Albin (2000) described the output quality characteristic of a process as a random response y that is a linear function (or profile) of an independent variable x. Assuming there are m profiles and each profile has p measurements, they modeled these linear profiles as

$$y_{ij} = A_0 + A_1 x_j + \epsilon_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, p, \tag{2.1.1}$$

where y_{ij} denotes the response of the ith profile at the jth measurement x_j, the intercept A_0 and the slope A_1 are fixed parameters, and ϵ_{ij}'s are independent and identically distributed (i.i.d.) as $N(0, \sigma_e^2)$. Kang and Albin (2000) used the least squares estimators

$$a_{i0} = \overline{y}_i - a_{i1}\overline{x} \tag{2.1.2}$$

and

$$a_{i1} = \frac{S_{xy(i)}}{S_{xx}} \tag{2.1.3}$$

to estimate A_0 and A_1 of the ith profile, respectively, where $\overline{y}_i = \sum_{j=1}^{p} y_{ij}/p$, $\overline{x} = \sum_{j=1}^{p} x_j/p$, $S_{xy(i)} = \sum_{j=1}^{p} y_{ij}(x_j - \overline{x})$, and $S_{xx} = \sum_{j=1}^{p}(x_j - \overline{x})^2$. By treating the residuals of each profile fitting as a subgroup, Kang and Albin (2000) combined the EWMA chart and the R chart for Phase II profile monitoring. Note that a_{i0} and a_{i1} are correlated.

Considering the same model (2.1.1), Kim et al. (2003) first coded the x-values by centering so that the average coded value is zero. With this decoding, the least squares estimators of the intercept and slope for each sample are independent, which allows them to combine three independent EWMA charts to monitor the intercept, slope, and σ_e^2 simultaneously. For more details, see the original papers. Since A_0 and A_1 are constants, we may say that (2.1.1) is a fixed-effect model.

For applications where slight random deviations in the intercept and slope are allowed, Shiau, Lin, and Chen (2006) considered the following random-effect model:

$$y_{ij} = A_{i0} + A_{i1}x_j + \epsilon_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, p, \tag{2.1.4}$$

where random coefficients A_{i0}'s and A_{i1}'s are i.i.d. as $N(\alpha_0, \sigma_0^2)$ and $N(\alpha_1, \sigma_1^2)$, respectively, and the error variables ϵ_{ij}'s are i.i.d. as $N(0, \sigma_e^2)$. Moreover, they assumed that A_{i0}, A_{i1}, and ϵ_{ij} are mutually independent. Following Kim et al. (2003), Shiau, Lin, and Chen (2006) first coded the x-values such that $\overline{x} = 0$ after centering. Then for $i = 1, \ldots, m$, the least squares estimators of A_{i0} and A_{i1} are

$$\widehat{\alpha}_{i0} = \overline{y}_i \tag{2.1.5}$$

and

$$\widehat{\alpha}_{i1} = \frac{S_{xy(i)}}{S_{xx}}, \tag{2.1.6}$$

respectively, where $\overline{y}_i = \sum_{j=1}^{p} y_{ij}/p$, $S_{xy(i)} = \sum_{j=1}^{p} x_j(y_{ij} - \overline{y}_i)/p = \sum_{j=1}^{p} x_j y_{ij}/p$, and $S_{xx} = \sum_{j=1}^{p} x_j^2/p$. Shiau, Lin, and Chen (2006) showed that $\widehat{\alpha}_{10}, \ldots, \widehat{\alpha}_{m0}$ are i.i.d. as $N(\alpha_0, \sigma_0^2 + \sigma_e^2/p)$, $\widehat{\alpha}_{11}, \ldots, \widehat{\alpha}_{m1}$ are i.i.d. as $N(\alpha_1, \sigma_1^2 + \sigma_e^2/S_{xx})$, and these statistics are mutually independent. With these distributional properties, they proposed a monitoring scheme accordingly. For more details, see Shiau, Lin, and Chen (2006).

2.1.2 Nonlinear Profiles

Profiles that can not be well described by a linear model are often regarded as nonlinear profiles. To model nonlinear profiles, one popular approach is via nonlinear parametric regression as follows:

$$y_{ij} = f(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_i) + \epsilon_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, p, \tag{2.1.7}$$

where $\boldsymbol{x}_{ij}$ is the $K \times 1$ vector of regressors for the jth observation of the ith profile, $\boldsymbol{\beta}_i$ is the $P \times 1$ vector of parameters for the ith profile, f is nonlinear in the parameters, and the random errors ϵ_{ij}'s are usually assumed to be i.i.d. as $N(0, \sigma_e^2)$. In many applications, there is only one regressor ($K = 1$) but having more than one parameter ($P > 1$) to monitor. In matrix form, the scalar model given in (2.1.7) is expressed by

$$\boldsymbol{y}_i = f(X_i, \boldsymbol{\beta}_i) + \boldsymbol{\epsilon}_i,\ i = 1, \ldots, m, \tag{2.1.8}$$

where $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{ip})'$, $X_i = (\boldsymbol{x}'_{i1}, \ldots, \boldsymbol{x}'_{ip})$, $f(X_i, \boldsymbol{\beta}_i) = (f(\boldsymbol{x}_{i1}, \boldsymbol{\beta}_i), \ldots, f(\boldsymbol{x}_{ip}, \boldsymbol{\beta}_i))'$, and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2}, \ldots, \epsilon_{ip})'$.

Having these estimates of $\boldsymbol{\beta}_i$, one can then monitor the process with standard multivariate control charting schemes. Similar to linear profiles, fixed-effect and random-effect models can be considered for different applications. For profiles with random effects, using the vertical density profiles (VDP) data given in Walker and Wright (2002) as an example, Williams et al. (2003) compared three T^2 control charts for monitoring VDP data by fitting a particular nonlinear regression model to data. For robustness, Shiau, Yen, and Feng (2006) proposed a Hotelling T^2 chart based on the Minimum Covariance Determinant (MCD) estimators. For getting more model flexibility, Shiau et al. (2009) employed the technique of principal component analysis to deal with the covariance structure of the nonlinear profiles with random effects via nonparametric regression. For more on monitoring nonlinear profiles, readers are referred to the afore mentioned articles and the references cited therein.

2.2 Data Smoothing

Data analysis is a process of obtaining insights from data by employing statistical techniques. But quite often, noise may be too large to reveal the interesting features in the data. For many applications, the underlying profiles are smooth and then smoothing is a useful pre-processing technique to filter out the random noises from raw sample profiles.

The notion of data smoothing is to fit a smooth function whose final form is flexible and depends on data and the level of smoothness chosen. In other words, the technique lets data speak for themselves in presenting their characteristics. One of the popular approaches of data smoothing is to fit noisy sample profiles by splines. Frequently, cubic splines (i.e., the piecewise cubic polynomials with continuous second derivatives) are used. The spline smoothing technique that we use in this paper is smoothing splines, which is available in statistical package R.

2.3 Principal Component Analysis

Principal component analysis (PCA) is a popular technique in multivariate data analysis. The following materials on PCA are from Johnson and Wichern (2007). The idea of PCA is to find a small number of directions (as linear combinations of the original variables)

that can explain most of the variability in data. Therefore, PCA is often employed for dimension reduction.

Suppose that $\boldsymbol{Y}$ is a p-variable random vector following a multivariate distribution. If we want to keep all the total variability (as in the variance-covariance structure) of the multivariate distribution, then p components are usually required. But often much of this variability could be accounted for with just a smaller number (denoted by k) of the principal components obtained from PCA. In other words, we could get almost as much information with these k principal components as there is in the original p variables. These k principal components then could be used as a surrogate of the initial p variables for subsequent analysis. In this sense, the general purpose of dimension reduction is achieved.

In fact, these principal components simply are eigenvectors of the variance-covariance matrix of the multivariate distribution. If this variance-covariance matrix needs to be estimated from data, then the sample version of PCA can be employed.

Since PCA is a dimension reduction technique, the important issue about how many principal components should be taken arises. One of popular criteria is to take the first k principal components such that the amount of the total variation explained by these k principal components reaches a satisfactory level.

2.4 Hierarchical Clustering Methods

Hierarchical clustering methods proceed by either a series of successive mergers or a series of successive divisions. Agglomerative hierarchical methods start with the individual objects. Initially, there are as many clusters as objects. These initial clusters are merged according to their similarities. The most similar objects are merged first. Repeat the merging process until all subgroups are merged into one single cluster.

Divisive hierarchical methods work in the opposite direction. The initial single cluster of all objects is divided into two subgroups such that the objects in one subgroup are most faraway from the objects in the other subgroup. Then these subgroups are further divided into dissimilar subgroups. The dividing procedure continues until there exist as many subgroups as objects, that is, until each object itself becomes one cluster. Results of agglomerative or divisive hierarchical methods can be displayed in a two-dimensional plot known as dendrogram. With this dendrogram, we can trace the mergers or divisions that have been made at successive levels.

Three popular linkage methods are often used in hierarchical clustering methods, single linkage (minimum distance or nearest neighbor), complete linkage (maximum distance or farthest neighbor), and average linkage (average distance). The single linkage occurs if groups are merged according to the distance between their nearest members. The complete linkage occurs if groups are merged according to the distance between their farthest members. For the average linkage, groups are merged according to the average distance between pairs of members in the respective groups.

In this paper, we adopt an agglomerative hierarchical procedure with single linkage. For the distance between objects, we use Euclidean distance in this paper. The following are the steps in the agglomerative hierarchical clustering algorithm for clustering N objects:

1. Start with N clusters, each cluster containing a single object and an $N \times N$ symmetric matrix of distances (or similarities) denoted by $D = \{d_{ij}\}$, where d_{ij} is the distance (similarity) between the ith and jth clusters.

2. In this distance matrix, search for the nearest (most similar) pair of clusters according to the distance elements. Denote the distance between the two most similar clusters, U and V, by d_{UV}.

3. First, merge the clusters U and V and label the newly formed cluster by (UV). Then update the elements in the distance matrix by (i) removing the rows and columns corresponding to clusters U and V and (ii) adding a row and a column giving the distances between the new cluster (UV) and the remaining clusters.

4. Repeat steps 2 and 3 until all objects are merged into one single cluster. While merging, record the identities of clusters that are merged and the corresponding dendrogram (i.e., clustering tree) levels at which the mergers take place.

For more detailed discussions on clustering methods, see, for example, Johnson and Wichern (2007) and the references cited therein.

2.5 Data Depth

The data depth is a measure of how central a given point is with respect to a multivariate distribution. There are several notions of data depth mentioned in Liu et al. (2004). For

convenience, the simplicial depth will be introduced as an illustrative example and the Oja depth will be used to construct our profile monitoring schemes in this study.

2.5.1 Simplicial Depth

First, we use the simplicial depth proposed by Liu (1990) as an example to illustrate the general notion of data depth.

When m random observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m$ come from a k-dimensional distribution with the cumulative distribution function (c.d.f.) F, we denote it by $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\} \sim F$. For simplicity, assume that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m$ are distinct and any $k+1$ of them will form a (nondegenerate) simplex. Let $s[\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_{k+1}}]$ denote the closed simplex with vertices $\{\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_{k+1}}\}$ and $I(\cdot)$ be the indicator function, i.e., $I(\Omega) = 1$ if Ω occurs and $I(\Omega) = 0$ otherwise, where Ω denotes a specific event. Suppose that $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\} \sim F$ are given. Each sample point is regarded as a $k \times 1$ vector and $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ is called a reference sample. Then the sample simplicial depth of any point $\boldsymbol{y} \in R^k$, $\boldsymbol{y} \notin \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$, is defined as

$$D_{F_m^*}(\boldsymbol{y}) = \binom{m+1}{k+1}^{-1} \sum_{(*)} I(\boldsymbol{y} \in s[\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_{k+1}}]), \tag{2.5.1}$$

where $(*)$ runs over all possible subsets of $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m, \boldsymbol{y}\}$ of size $(k+1)$ and the notation F_m^* stands for the empirical c.d.f. of the sample points in $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m, \boldsymbol{y}\}$. If $\boldsymbol{y} \in \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$, i.e. $\boldsymbol{y} = \boldsymbol{X}_i$ for some i, then

$$D_{F_m}(\boldsymbol{y}) = \binom{m}{k+1}^{-1} \sum_{(**)} I(\boldsymbol{y} \in s[\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_{k+1}}]), \tag{2.5.2}$$

where $(**)$ runs over all possible subsets of $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ of size $(k+1)$ and the notation F_m stands for the empirical c.d.f. of the sample points in $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$. A large value of $D_{F_m^*}(\boldsymbol{y})$ indicates that $\boldsymbol{y}$ is contained in many closed simplices constructed from the reference sample, and thus it is relatively central within the data cloud. On the contrary, a small value of $D_{F_m^*}(\boldsymbol{y})$ indicates a relatively outlying position of $\boldsymbol{y}$.

If the multivariate distribution F is known, then the simplicial depth of any point $\boldsymbol{y}$ with respect to the distribution F is defined as

$$D_F(\boldsymbol{y}) = P_F\{\boldsymbol{y} \in s[\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k+1}]\}, \tag{2.5.3}$$

where $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k+1}$ are $(k+1)$ random variables following the multivariate distribution F. Many basic properties of the simplicial depth $D_F(\cdot)$ can be found in Liu (1990). In

particular, the convergence of $D_{F_m}(\cdot)$ to $D_F(\cdot)$ allows us to approximate $D_F(\cdot)$ by $D_{F_m}(\cdot)$ when the multivariate distribution F is unknown.

2.5.2 Oja Depth

For any $k+1$ points $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k+1} \in R^k$, the simplex $s[\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k+1}]$ has a k-dimensional volume defined as

$$\begin{aligned} \triangle(s[\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k+1}]) &= \frac{1}{k!} \mid det((1, \boldsymbol{Z}_1')', \ldots, (1, \boldsymbol{Z}_{k+1}')') \mid \\ &= \frac{1}{k!} \mid det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_{11} & z_{21} & \cdots & z_{k+1,1} \\ z_{12} & z_{22} & \cdots & z_{k+1,2} \\ \vdots & \vdots & \ddots & \vdots \\ z_{1k} & z_{2k} & \cdots & z_{k+1,k} \end{pmatrix} \mid, \end{aligned} \tag{2.5.4}$$

where $\boldsymbol{Z}_1 = (z_{11}, \ldots, z_{1k})'$, $\ldots$, $\boldsymbol{Z}_{k+1} = (z_{k+1,1}, \ldots, z_{k+1,k})'$.

Oja (1983) proposed a location measure for the empirical distribution of the sample points in $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$, called the Oja median, as the location that minimizes the average volume of all the possible k-dimensional simplices as follows:

$$\min_{\boldsymbol{y} \in R^k} \frac{1}{\binom{m}{k}} \sum_{1 \le i_1 < \cdots < i_k \le m} \triangle(s[\boldsymbol{y}, \boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_k}]),$$

which is equivalent to

$$\min_{\boldsymbol{y} \in R^k} \frac{(m-k)!}{m!} \sum_{1 \le i_1 < \cdots < i_k \le m} \mid det((1, \boldsymbol{y}')', (1, \boldsymbol{X}_{i_1}')', \ldots, (1, \boldsymbol{X}_{i_k}')') \mid, \tag{2.5.5}$$

where $\boldsymbol{y}$ is any point and $\{\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_k}\}$ are k different points chosen from $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$.

The Oja depth function was proposed by Zuo and Serfling (2000) as

$$d_{Oja}(\boldsymbol{y} \mid T) = \big(1 + \frac{E(\triangle(s[\boldsymbol{y}, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k]))}{\sqrt{det(\Sigma)}}\big)^{-1}, \tag{2.5.6}$$

where $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k$ are k i.i.d. random vectors following a multivariate distribution T, E represents the expectation over $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k$, and Σ is the variance-covariance matrix of T. Accordingly, the sample Oja depth of any point $\boldsymbol{y} \in R^k$, $\boldsymbol{y} \notin \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$, is

$$\begin{aligned} & d_{Oja}(\boldsymbol{y} \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_m) \\ &= \big(1 + \frac{\frac{(m-k)!}{m!} \sum_{1 \le i_1 < \cdots < i_k \le m} \mid det((1, \boldsymbol{y}')', (1, \boldsymbol{X}_{i_1}')', \ldots, (1, \boldsymbol{X}_{i_k}')') \mid}{\sqrt{det(\Sigma_X)}}\big)^{-1}, \end{aligned} \tag{2.5.7}$$

where $\Sigma_X = \sum_{j=1}^{m}(\boldsymbol{X}_j - \overline{\boldsymbol{X}})(\boldsymbol{X}_j - \overline{\boldsymbol{X}})'/m$ with $\overline{\boldsymbol{X}} = \sum_{j=1}^{m} \boldsymbol{X}_j/m$. If $\boldsymbol{y} \in \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$, i.e. $\boldsymbol{y} = \boldsymbol{X}_i$ for some i, then the reference sample becomes $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{i-1}, \boldsymbol{X}_{i+1}, \ldots, \boldsymbol{X}_m\}$.

This Oja depth function is continuous in $\boldsymbol{x}$ and the maximum occurs at the Oja median while zero occurs at infinity. Also note that the Oja median may not be unique and the Oja depth function decreases monotonically on rays from all points in the Oja median set.

2.6 Relative Rank

For most of control charts, the fundamental judging criterion of in or out of control relies on a measure of how outlying an observation is. Data depth is such a measure but its distribution is too complicated to determine the control limit. To overcome this difficulty, Liu (1995) proposed an r-chart by representing a multivariate observation with its relative rank (induced from the data depth) as described below.

For any $\boldsymbol{y} \in R^k$, define the relative rank of $\boldsymbol{y}$ with respect to $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ as

$$r_{F_m}(\boldsymbol{y}) \equiv \frac{1}{m}\sum_{i=1}^{m} I(D_{F_m}(\boldsymbol{X}_i) < D_{F_m}(\boldsymbol{y})), \tag{2.6.1}$$

which is the proportion of the sample points in the reference sample that are more outlying than $\boldsymbol{y}$. For the population version, define

$$r_F(\boldsymbol{y}) \equiv P\{D_F(\boldsymbol{X}) < D_F(\boldsymbol{y}) | \boldsymbol{X} \sim F\} \tag{2.6.2}$$

as the relative rank of $\boldsymbol{y}$ with respect to the distribution F.

If the relative rank of $\boldsymbol{y}$ is extremely small, then the position of $\boldsymbol{y}$ is extremely outlying, indicating that $\boldsymbol{y}$ may not come from the same distribution that generates the reference sample $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$. This is the main idea behind data-depth control charts.

By reducing the dimension of data from k to 1, one can construct control charts based on these relative ranks following the same procedures as those for univariate control charts. Note that this approach is completely nonparametric, meaning that the resulting control charts are free of parametric distributional assumptions on process data. According to Liu and Singh (1993), these control charts can detect the location change and the scale increase in a process simultaneously.

For the Oja depth, we can define the relative rank as above by substituting the depth function (2.5.1) with (2.5.7). Liu et al. (2004) mentioned that the Oja depth can be used to construct multivariate control charts as well.

2.7 r-Chart

Denote the distribution of the k-dimensional in-control process by F. Assuming that an incoming random multivariate observation $\boldsymbol{Y}$ in Phase II on-line monitoring follows a k-dimensional distribution G. To monitor this incoming observation with a control chart is equivalent to test $H_0 : G = F$ versus H_a : there exists a location shift and/or a scale increase from F to G.

If H_0 is true, Liu and Singh (1993) showed that the relative rank follows a uniform distribution on the unit interval $[0, 1]$ as given in the following proposition.

Proposition 2.7.1 (Liu and Singh, 1993). *Assume that $\boldsymbol{Y} \sim G$ and $F = G$. Let $U[0, 1]$ denote the uniform distribution on the unit interval $[0, 1]$ and the notation "$\xrightarrow{L}$" stand for the convergence in law. If $D_F(\boldsymbol{Y})$ has a continuous distribution, then*

(1) $r_F(\boldsymbol{Y}) \sim U[0, 1]$, and
(2) as $m \to \infty$, $r_{F_m}(\boldsymbol{Y}) \xrightarrow{L} U[0, 1]$ along almost all $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ sequences, provided that $D_{F_m}(\cdot)$ converges to $D_F(\cdot)$ uniformly as $m \to \infty$.

The uniform convergence of $D_{F_m}(\cdot)$ holds for the simplicial depth if F is absolutely continuous. We shall assume F is absolutely continuous hereafter.

By Proposition 2.7.1, the r-chart can be constructed with the center line $CL = 0.5$ and the lower control limit $LCL = a$, where a is the desirable false-alarm rate. To monitor an incoming observation $\boldsymbol{y}$, the first step is to calculate the relative rank $r_F(\boldsymbol{y})$ of $\boldsymbol{y}$ with respect to the multivariate distribution F, or the relative rank $r_{F_m}(\boldsymbol{y})$ of $\boldsymbol{y}$ with respect to $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ when F is unknown and a reference sample $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ is available. If $r_F(\boldsymbol{y})$ or $r_{F_m}(\boldsymbol{y})$ is smaller than $LCL = a$, the process is considered to be out of control.

2.8 Q-Chart

The Q-chart proposed by Liu (1995) is the $\overline{X}$-chart version of the r-chart. Let $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots$ be the incoming observations to be monitored in time order. To construct the Q-chart, first we group $\boldsymbol{Y}_i$'s into subgroups of size q. Then we plot the averages of the subgroups of $r_F(\boldsymbol{Y}_i)$'s when the multivariate distribution F is known, or the averages of the subgroups of $r_{F_m}(\boldsymbol{Y}_i)$'s when only the reference sample $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ is available. For convenience, we denote the average of the jth subgroup of $r_F(\boldsymbol{Y}_i)$'s and $r_{F_m}(\boldsymbol{Y}_i)$'s by $Q(F, G_q^j)$ and

$Q(F_m, G_q^j)$, respectively, where the notation G_q^j denotes the empirical distribution of the $\boldsymbol{Y}_i$'s in the jth subgroup, $j = 1, 2, \ldots$. That is, when given G_q of a subgroup of q $\boldsymbol{Y}_i$'s,

$$Q(F, G_q) \equiv \frac{1}{q} \sum_{i=1}^{q} r_F(\boldsymbol{Y}_i) \tag{2.8.1}$$

and

$$Q(F_m, G_q) \equiv \frac{1}{q} \sum_{i=1}^{q} r_{F_m}(\boldsymbol{Y}_i). \tag{2.8.2}$$

In other words, the Q-chart plots exactly the statistics $\{Q(F, G_q^1), Q(F, G_q^2), \cdots\}$, or its empirical version. Liu (1995) gave asymptotic results for these statistics and provided the center line (CL) and lower control limit (LCL) of the Q-chart accordingly. The Q-chart also has $CL = 0.5$ and its LCL depends on the subgroup size q. If the subgroup size q is large, by Central Limit Theorem, LCL is approximately $0.5 - z_a(12q)^{-1/2}$ when plotting $\{Q(F, G_q^j)\}$'s, or $0.5 - z_a\sqrt{(1/m + 1/q)/12}$ when plotting $\{Q(F_m, G_q^j)\}$'s. If the subgroup size q is small and the false-alarm rate $a \leq 1/q!$, then LCL is exactly $(q!a)^{1/q}/q$. According to Liu (1995), this LCL is a reasonable approximation when the false-alarm rate a is slightly over $1/q!$.

2.9 DDMA-Chart

Liu et al. (2004) reported that the Q-chart could be inefficient in detecting small location shifts. So they proposed a moving-average control chart, called DDMA-chart, to enhance the efficiency.

If $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ is the reference sample following the multivariate distribution F and $\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots\}$ are new observations. Let q be the length of the moving window for computing moving averages. Then the moving averages are $\widetilde{\boldsymbol{Y}}_1 = (\boldsymbol{Y}_1 + \cdots + \boldsymbol{Y}_q)/q$, $\widetilde{\boldsymbol{Y}}_2 = (\boldsymbol{Y}_2 + \cdots + \boldsymbol{Y}_{q+1})/q$, …, $\widetilde{\boldsymbol{Y}}_{n-q+1} = (\boldsymbol{Y}_{n-q+1} + \cdots + \boldsymbol{Y}_n)/q$, and so on. For constructing the DDMA-chart to monitor the moving averages $\widetilde{\boldsymbol{Y}}_i$, $i = 1, 2, \ldots$, the corresponding reference sample is $\{\widetilde{\boldsymbol{X}}_1, \ldots, \widetilde{\boldsymbol{X}}_{m-q+1}\}$ where $\widetilde{\boldsymbol{X}}_1 = (\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_q)/q$, $\widetilde{\boldsymbol{X}}_2 = (\boldsymbol{X}_2 + \cdots + \boldsymbol{X}_{q+1})/q$, …, and $\widetilde{\boldsymbol{X}}_{m-q+1} = (\boldsymbol{X}_{m-q+1} + \cdots + \boldsymbol{X}_m)/q$.

Similar to constructing the r-chart, for each $\widetilde{\boldsymbol{Y}}_i$, we calculate the corresponding relative rank, $r_{\widetilde{F}_{m-q+1}}(\widetilde{\boldsymbol{Y}}_i)$, with respect to the reference sample $\{\widetilde{\boldsymbol{X}}_1, \ldots, \widetilde{\boldsymbol{X}}_{m-q+1}\}$, i.e.,

$$r_{\widetilde{F}_{m-q+1}}(\widetilde{\boldsymbol{Y}}_i) = \frac{1}{m-q+1} \sum_{j=1}^{m-q+1} I(D_{\widetilde{F}_{m-q+1}}(\widetilde{\boldsymbol{X}}_j) < D_{\widetilde{F}_{m-q+1}}(\widetilde{\boldsymbol{Y}}_i)), \tag{2.9.1}$$

where $\widetilde{F}_{m-q+1}$ denotes the empirical distribution of the reference sample $\{\widetilde{\boldsymbol{X}}_1, \ldots, \widetilde{\boldsymbol{X}}_{m-q+1}\}$ and $D_{\widetilde{F}_{m-q+1}}(\cdot)$ denotes the sample simplicial depth of one point with respect to the reference sample $\{\widetilde{\boldsymbol{X}}_1, \ldots, \widetilde{\boldsymbol{X}}_{m-q+1}\}$.

Accordingly, the DDMA-chart is just the plot of the statistics $r_{\widetilde{F}_{m-q+1}}(\widetilde{\boldsymbol{Y}}_i)$'s, $i = 1, 2, \ldots$, versus their sequence orders i's. Following the same principle as in constructing the r-chart, this DDMA-chart has the center line $CL = 0.5$ and the lower control limit $LCL = a$. According to Liu et al. (2004), the advantage of the DDMA-chart compared with the Q-chart is that the DDMA-chart is more sensitive to small location shifts while retaining the same performance in detecting scale shifts.

3 Proposed Monitoring Schemes

In this study, we focus mainly on Phase II monitoring. Same as in most of Phase II studies, we assume that we have sufficient information about the in-control process, meaning that we have already obtained a large reference sample whose features are explicit enough to represent the realistic situation of the in-control process.

3.1 Pre-processing Steps

Before we proceed to the actual Phase II monitoring, we need to convert the sample profiles in the reference sample to multivariate data for data depth computation with the following two procedures:

1. Data Smoothing

   Although the profiles in the reference sample are from the in-control process, the existence of the random noise in the profile data may get in the way of the subsequent analysis. Therefore, a smoothing step is suggested to de-noise sample profiles. We use smoothing splines as our smoothing technique for de-noising sample profiles throughout this paper. Other popular smoothing techniques such as B-spline regression, local polynomial regression smoothing, and wavelets can be used as well.

2. Principal Component Analysis

   After data smoothing, we apply principal component analysis to the sample variance-covariance matrix of the sample profiles in the reference sample. By choosing an appropriate number k (often much smaller than p) of principal components, we reduce the dimension of the profile data from p to a much lower dimension k. For the illustrative examples shown in this paper, we choose 2 or 3 principal components according to the principle of achieving a satisfactory level of total variation explained.

Below, we summarize the above two pre-processing steps in more details:

1. Suppose that we have m in-control sample profiles $\{\boldsymbol{X}_1^*, \ldots, \boldsymbol{X}_m^*\}$, in which the ith profile is expressed by a p-dimensional vector $\boldsymbol{X}_i^*$ for $i = 1, \ldots, m$.

2. For each raw sample profile $\boldsymbol{X}_i^*$, we use smoothing splines to purge random noise from the profile. Denote these smoothed sample profiles by $\{\boldsymbol{X}_1^{**}, \ldots, \boldsymbol{X}_m^{**}\}$. Here each smoothed sample profile $\boldsymbol{X}_i^{**}, i = 1, \ldots, m$, is still a p-dimensional vector.

3. Perform PCA to these m smoothed sample profiles $\{\boldsymbol{X}_1^{**}, \ldots, \boldsymbol{X}_m^{**}\}$. That is, compute the eigenvalue-eigenvector decomposition of the $p \times p$ sample variance-covariance matrix of $\{\boldsymbol{X}_1^{**}, \ldots, \boldsymbol{X}_m^{**}\}$. The jth principal component is the eigenvector $\boldsymbol{v}_j$ corresponding to the jth largest eigenvalue λ_j, $j = 1, \cdots, p$.

4. Compute $S_{ij} = \boldsymbol{X}_i^{**'} \boldsymbol{v}_j$, the score of the jth principal component of the ith smoothed profile, $j = 1, \ldots, p,\ i = 1, \ldots, m$.

5. Choose an appropriate number of principal components that can explain the variation to a satisfactory level. Assume that we select the first k ($< p$) effective principal components $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. Consequently, for each smoothed profile $\boldsymbol{X}_i^{**}$, we can obtain a score vector

$$\boldsymbol{X}_i \equiv (S_{i1}, \ldots, S_{ik})', \ i = 1, \ldots, m. \tag{3.1.1}$$

Thus we have reduced the data from p-dimensional smoothed profile vectors to k-dimensional score vectors. These m score vectors form the reference sample to be used in the data depth computation.

3.2 Phase II Monitoring

For Phase II monitoring, we use the reference sample, $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$, described in the previous subsection to represent the in-control process. Below we describe how to perform the Phase II monitoring step by step:

1. Let $\boldsymbol{Y}^*$ denote an incoming p-dimensional profile vector to be monitored. Smooth it as described earlier. Denote the smoothed profile vector by $\boldsymbol{Y}^{**}$.

2. Project the smoothed profile $\boldsymbol{Y}^{**}$ onto the first k effective principal components $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ to obtain the score vector $\boldsymbol{Y} \equiv (\boldsymbol{Y}^{**'} \boldsymbol{v}_1, \ldots, \boldsymbol{Y}^{**'} \boldsymbol{v}_k)'$.

The above 2 steps focus on pre-processing the incoming profile data into a k-dimensional score vector suitable for further data depth computation.

The following steps focus on monitoring the score vectors of the incoming profiles, say, $\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_i, \ldots\}$. We first describe the control charting for the r-chart.

1. Initialization: Compute the Oja depths of the reference sample by equation (2.5.7). That is, $\{d_{Oja}(\boldsymbol{X}_1 \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_m), \ldots, d_{Oja}(\boldsymbol{X}_m \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_m)\}$.

2. For $i = 1, 2, \ldots,$

   (a) Compute the Oja depth of the ith incoming score data $\boldsymbol{Y}_i$ with respect to the reference sample by equation (2.5.7). That is, $d_{Oja}(\boldsymbol{Y}_i \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_m)$.

   (b) Compute the relative rank of $\boldsymbol{Y}_i$ with respect to $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ by

$$r_{F_m}(\boldsymbol{Y}_i) = \frac{1}{m} \sum_{j=1}^{m} I(d_{Oja}(\boldsymbol{X}_j \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_m) < d_{Oja}(\boldsymbol{Y}_i \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_m)). \tag{3.2.1}$$

   (c) Plot $r_{F_m}(\boldsymbol{Y}_i)$ against its sequence order i on the control chart. The center line and the lower control limit of the r-chart are $CL = 0.5$ and $LCL = a$, respectively. Here a is the desirable false-alarm rate. If $r_{F_m}(\boldsymbol{Y}_i)$ is below LCL, then the process is considered out of control.

The Q-chart and DDMA-chart can be constructed with the score vectors as described in Subsections 2.8 and 2.9, respectively.

4 Diagnosing Profiles in the Historical Data Set

It is desirable to have a control chart for Phase I analysis. Unfortunately, the relative ranks of the profiles in the reference sample cannot be used to discriminate out-of-control profiles from in-control profiles due to the nature of ranks. On the other hand, the Oja depths contain more information on how outlying a profile is. Although the distributional properties of the Oja depths for profiles are not available, the depth values might still be useful in screening out potential out-of-control profiles. In this section, we describe a heuristic screening procedure via clustering.

4.1 A Simple Diagnostic Procedure

Let $\{\boldsymbol{X}_1^*, \ldots, \boldsymbol{X}_l^*\}$ be a set of historical profile data. As described before, we first perform data smoothing and PCA to the profile data to convert them into k-dimensional score vectors $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_l\}$; then compute their Oja depths $\{d_{Oja}(\boldsymbol{X}_1 \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_l), \ldots, d_{Oja}(\boldsymbol{X}_l \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_l)\}$. We then apply the agglomerative hierarchical clustering procedure with single linkage to cluster these Oja depths into two groups. The profiles in the group with smaller Oja depths are considered to be more outlying than those in the other group. Pursue investigation on these potential out-of-control profiles. Remove doubtful profiles from the data set. The remaining data set could then be considered as the reference sample for Phase II profile monitoring.

We demonstrate this diagnostic procedure with two real data sets next.

4.2 Examples

4.2.1 Bioassay Data

- Data Description

  This data set is from DuPont Crop Protection (Williams et al., 2007). It consists of forty-four weeks ($l = 44$) of in vivo bioassay results run alongside experimental compounds over a one-year time period. Here, the commercial compound is diluted to eight doses, 0.003, 0.009, 0.028, 0.084, 0.25, 0.76, 2.27, 6.8 ($d = 8$) and replicates four times at each dose ($r = 4$) in 96-well microtiter plates for each sampling period. A spectrophotometer measures the optical density (OD) of the plant organism after the inoculation period. Both treated and untreated wells are measured for growth

inhibition. The percent control (PC) values are computed using the median OD (M_i) from 96 replications of untreated wells.

Following Williams et al. (2007), let y_{ijk} represent the kth response of the jth dose at sampling period i, where $i = 1, 2, \ldots, l$, $j = 1, 2, \ldots, d$, and $k = 1, 2, \ldots, r$. In this data set, $l = 44$, $d = 8$, and $r = 4$. Let M_i represent the median response of the untreated specimen at sampling period i. Then, the percent control of the chemical for the kth replication of the jth dose at sampling period i is calculated as

$$PC_{ijk} = \frac{M_i - y_{ijk}}{M_i}. \tag{4.2.1}$$

The plots of PC_{ijk} values for all $l = 44$ weeks for one of the commercial crop protection products from the DuPont are given in Figure 1.

- Diagnostic Results

  To proceed our proposed diagnostic scheme, we first note that there are four replications at each dose for each week in this bioassay data. The replication complicates the computation of smoothing splines. To simplify computation, we average the four replications and apply all the analyzing steps described earlier to the averaged bioassay data, i.e., smoothing splines, principal component analysis, Oja depths, and the agglomerative hierarchical clustering method. The smoothing spline fittings of 44 averaged profiles are shown in Figure 1.

  To understand the effects of the smoothing parameter as well as the number of principal components, various choices of these parameters are tried. The results are given in Tables 1-3. To read Tables 1-3,

  - "spar" is the smoothing parameter in smoothing splines;
  - the first diagnosis is the result that we perform the diagnostic procedure to the original 44 weeks/profiles;
  - the second diagnosis is the result that we repeat the diagnostic procedure to the data after removing the profiles diagnosed as out-of-control in the first diagnosis;
  - "cumulative" gives the number of principal components used in analysis and the proportion of the variation explained by the principal components used;

– "weeks" shows which weeks are diagnosed as out-of-control.

For example, the last row in Table 1 indicates that, under smoothing parameter spar=0.3, we choose 3 principal components, which explains 94.1% of variation, and the 34th week is suggested to be out of control in the first diagnosis. After removing the 34th week, repeat the diagnostic procedure. By choosing 4 principal components, which explains 97.7% of the variation, the 32th week is then diagnosed as potentially shifted in the second diagnosis.

From these results, we observe that both smoothing parameters and the number of effective principal components affect the proportion of variation explained and possibly further lead to different diagnostic results. It is also observed that, for those suggested outlying weeks in Tables 1-3, almost all of their original data profiles exhibit some unusual patterns.

The reason for performing second diagnosis is because the procedure only takes the last result of clustering into two groups. If the data set exhibits more than two clusters, say, one in-control and two out-of-control, then the sizes of the gaps between clusters determine the final clustering result. For example, it is possible that an extreme outlying profile will form a cluster by itself. A better solution for this is to look into the dendrogram (clustering tree) as well as the corresponding Oja depths and determine the cutoff point for separating in or out-of-control profiles. Also look into the profiles themselves to see if they have unusual patterns. In other words, the heuristic diagnostic procedure presented here is only providing some indications of potential outlying profiles in the historical data set. It should not be used as the sole rule to follow in Phase I analysis.

4.2.2 Vertical Density Profile Data

- Data Description

  This vertical density profile data set is from Walker and Wright (2002). In the manufacture of particleboards, the density properties of the produced boards are quality characteristics that are monitored through time. The density is measured using a profilometer which uses a laser device to scan fixed vertical depths across the thickness of the board and the density at each depth is recorded. A profilometer

takes multiple measurements on a sample to form the vertical density profile of the board. This data set contains 24 profiles of vertical density and each profile consists of 314 measurements. Figure 2 displays the VDP data.

- Diagnostic Results

  We apply the diagnostic procedure to these vertical density profiles. And the diagnostic results are given in Table 4. Similar to the bioassay data, different choices of smoothing parameters affect the proportion of variation explained. Although the diagnostic results are all the same for the four smoothing parameters we used in analyzing this VDP data, this does not mean that there are no other possibilities. Figures 3-5 show these potentially outlying profiles. They all exhibit some unusual patterns.

5 Simulation and Comparative Studies

5.1 Aspartame Example

Kang and Albin (2000) raised an interesting issue in process monitoring. In most applications, a single variable or multiple variables can characterize the state of the process. However, there exist a number of processes where the state is characterized by a profile or a function. Aspartame described in Kang and Albin (2000) is one of such examples. An important characteristic of the aspartame is the amount of aspartame that is dissolved per liter of water at different temperatures. For illustration, Figure 6 shows the plot of some hypothetical aspartame profiles.

Following Shiau et al. (2009), we adopt the profile form of $Y = I + Me^{N(x-1)^2} + \epsilon$ to simulate aspartame profiles. To simulate in-control profiles with random effects, their idea is to randomly disturb the parameters I, M, and N to generate natural variation among profiles. Accordingly, the following random-effect model can be considered to generate aspartame profiles:

$$Y_j = I + Me^{N(x_j-1)^2} + \epsilon_j,\ j = 1, \ldots, p, \tag{5.1.1}$$

where $I \sim N(\mu_I, \sigma_I^2)$, $M \sim N(\mu_M, \sigma_M^2)$, $N \sim N(\mu_N, \sigma_N^2)$, and $\epsilon \sim N(0, \sigma_\epsilon^2)$. All of these random parameters, i.e., $\{I, M, N, \epsilon\}$, are independent of each other.

Shiau et al. (2009) mentioned that the distribution of the response profile $\boldsymbol{Y} = (Y_1, \ldots, Y_p)'$ generated by the above model (5.1.1) is fairly complicated. Specifically, this distribution has mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)'$ and variance-covariance matrix Σ as follows. For all $i,\ j = 1, \ldots, p$,

$$\mu_j = E(Y_j) = \mu_I + \mu_M e^{\mu_N (x_j-1)^2 + \frac{\sigma_N^2 (x_j-1)^4}{2}} \tag{5.1.2}$$

and

$$\begin{aligned} Cov(Y_i, Y_j) = \sigma_I^2 + (\mu_M^2 + \sigma_M^2)[e^{\mu_N[(x_i-1)^2+(x_j-1)^2] + \frac{\sigma_N^2[(x_i-1)^2+(x_j-1)^2]^2}{2}}] \\ - \mu_M^2 e^{\mu_N(x_i-1)^2 + \frac{\sigma_N^2(x_i-1)^4}{2} + \mu_N(x_j-1)^2 + \frac{\sigma_N^2(x_j-1)^4}{2}} + \sigma_\epsilon^2 \delta_{ij}, \end{aligned} \tag{5.1.3}$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. Unfortunately, this variance-covariance matrix Σ would be affected if the mean of the parameter M or N shifts and this model is too complicated to analyze the performance of the control charts under their study.

We remark that the control charts proposed in Shiau et al. (2009) were developed under the assumption that the random profile follows a Gaussian process. Thus, instead of using this original model (5.1.1), they modeled the aspartame profiles according to a Gaussian stochastic process with the mean function

$$\mu(x) = \mu_I + \mu_M e^{\mu_N (x-1)^2} \tag{5.1.4}$$

and a covariance function $G(s,t)$, here s and t are covered in the domain of x. In order to maintain the natural variation among profiles as it would be in the original random-effect model (5.1.1), they let the in-control aspartame profiles follow the distribution $MVN(\boldsymbol{\mu}_0, \Sigma)$, where $\boldsymbol{\mu}_0 = (\mu_{01}, \ldots, \mu_{0p})'$ with

$$\mu_{0j} = \mu_I + \mu_M e^{\mu_N (x_j - 1)^2}, \ j = 1, \ldots, p, \tag{5.1.5}$$

and Σ is the variance-covariance matrix given by equation (5.1.3).

If the mean function (5.1.4) is shifted, say, μ_I to $\mu_I + \alpha\sigma_I$, μ_M to $\mu_M + \beta\sigma_M$, and μ_N to $\mu_N + \gamma\sigma_N$, μ_{0j} is shifted from $\mu_I + \mu_M e^{\mu_N (x_j-1)^2}$ to

$$\widetilde{\mu}_j \equiv (\mu_I + \alpha\sigma_I) + (\mu_M + \beta\sigma_M) e^{(\mu_N + \gamma\sigma_N)(x_j - 1)^2}, \ j = 1, \ldots, p. \tag{5.1.6}$$

Let $\widetilde{\boldsymbol{\mu}} = (\widetilde{\mu}_1, \ldots, \widetilde{\mu}_p)'$. Then the shift on the mean of $\boldsymbol{Y}$ is $\boldsymbol{\delta} \equiv \widetilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0$.

Normally, when compared with a parametric method, a nonparametric method gains model flexibility but loses some efficiency (but *only* when the model is right for the parametric method). Although the "nonparametric" control charts developed in this paper are not restricted to distributional assumptions, we still use the setup that is suitable for the schemes of Shiau et al. (2009) in our comparison study. It is surprising to observe that even under a situation that suits the parametric methods, the new nonparametric charts perform quite well in terms of the detecting power.

5.2 Settings for Simulation

For the simulation studies, we generate aspartame profiles according to the distribution $MVN(\boldsymbol{\mu}_0, \Sigma)$, where $\boldsymbol{\mu}_0$ is the mean vector defined by (5.1.5) and Σ is the variance-covariance matrix given in (5.1.3). Here the parameters for simulation are $\mu_I = 1$, $\sigma_I = 0.2$, $\mu_M = 15$, $\sigma_M = 1$, $\mu_N = -1.5$, $\sigma_N = 0.3$, $\sigma_\epsilon = 0.3$, and $x = 0.64, 0.8, \ldots, 3.52$. Both of x and y values are scaled variables, not the real temperatures and the amount

of aspartame that is dissolved per liter of water. Following Liu et al. (2004), all control charts that we discuss here are designed to have the same in-control average run length (ARL), $ARL_0 = 20$, which corresponds to the false-alarm rate $a = 0.05$.

5.3 Comparative Studies

In order to assess the performance of our proposed Phase II monitoring scheme through ARL, we compute the ARL values for each control chart scheme under various shift sizes on parameters I, M, and N as follows.

Here we concentrate on the performances of r-chart, Q-chart (with three subgroup sizes 2, 4, 6), and DDMA-chart (with three subgroup sizes 2, 4, 6). Denote these control charts as Oja, Q2, Q4, Q6, D2, D4, and D6, respectively.

Consider shifting one parameter of I, M, and N at a time with shift size ranging from 0.00 to 3.00 in the increment of .25, i.e., $\{0.00, 0.25, 0.50, \ldots, 2.75, 3.00\}$.

For each scenario, we generate m (m=1008 in this study) in-control aspartame profiles according to the distribution $MVN(\boldsymbol{\mu}_0, \Sigma)$ given in Subsection 5.1. These m in-control aspartame profiles are used to construct the reference sample as described in Subsection 3.2. We choose $k = 2$ principal components for this example in creating the reference sample because the total variation explained by the first two principal components is over 90%.

We then generate n (n=1008 in this study) out-of-control aspartame profiles for each scenario according to the distribution $MVN(\widetilde{\boldsymbol{\mu}}, \Sigma)$, where $\widetilde{\boldsymbol{\mu}}$ is given in equation (5.1.6). Perform the Phase II monitoring schemes described in section 3 to these out-of-control profiles. Compute the proportion of the profiles that are considered out-of-control by each of the charts. The reciprocal of this proportion is an estimate of the ARL. Repeat this procedure 2000 times. Average the 2000 ARL estimates to obtain the final ARL estimate. The standard deviation of this ARL estimator can be estimated by dividing the sample standard deviation of 2000 ARL estimates by $\sqrt{2000}$. Results are given in Tables 5-7 for I, M, and N shifts, respectively.

In the following, we compare the ARL performance of the proposed monitoring scheme (r-chart) with those proposed in Shiau et al. (2009). In addition, we will compare the ARL performances among r-chart, Q-chart, and DDMA-chart.

5.3.1 ARL Comparisons with Shiau et al. (2009)

Before comparing our proposed scheme with the schemes proposed in Shiau et al. (2009) through ARL performance, we briefly introduce the Phase II monitoring schemes proposed in Shiau et al. (2009).

Shiau et al. (2009) characterized the in-control profiles after de-noising according to the distribution $N_p(\boldsymbol{\mu}_0, \Sigma_0)$, where the in-control variance-covariance matrix Σ_0 is in equation (5.1.3) without the $\sigma_\epsilon^2 \delta_{ij}$ term.

Their method goes like this. First, apply PCA to the in-control variance-covariance matrix Σ_0 to get eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$, and the corresponding eigenvectors, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$. Choose an appropriate number k of principal components that can explain high enough percentage of total variation.

Then for each incoming profile in Phase II monitoring, first smooth and then project this profile onto the k principal components (PCs) to obtain $S_1, \ldots, S_k$, the k scores. Because these k scores are independent and each S_r has distribution $N(\boldsymbol{v}_r'\boldsymbol{\mu}_0, \lambda_r)$ when the process is in control, hence, it is easy to construct a control chart for each of the k scores. With the false-alarm rate a, $\boldsymbol{v}_r'\boldsymbol{\mu}_0 \pm Z_{a/2}\sqrt{\lambda_r}$ are the control limits of the rth PC-score-chart for monitoring the statistic S_r, $r = 1, \ldots, k$.

The advantage of these PC-score-charts is that if a particular mode of process change can be caught by one of the k principal components, then one can use that particular PC-score-chart to monitor this type of process change. Unfortunately, a process change is often reflected on more than one principal component. Instead of considering only one particular PC-score-chart, Shiau et al. (2009) also considered a combined chart scheme by combining all k PC-score-charts as follows.

The combined chart scheme signals out-of-control if any of the k individual PC-score-charts signals. In other words, their combined chart scheme is equivalent to monitoring the statistic

$$\max_{1 \leq r \leq k} \mid \frac{S_r - \boldsymbol{v}_r'\boldsymbol{\mu}_0}{\sqrt{\lambda_r}} \mid . \tag{5.3.1}$$

This chart signals out-of-control if $\max_{1 \leq r \leq k} \mid (S_r - \boldsymbol{v}_r'\boldsymbol{\mu}_0)/\sqrt{\lambda_r} \mid > Z_{a'/2}$, where the individual false-alarm rate a' should be chosen at the level of $1 - (1-a)^{1/k}$ so that the overall false-alarm rate is at the desired level a.

Shiau et al. (2009) also considered a T^2 chart with the monitoring statistic

$$T^2 = \sum_{r=1}^{k} \frac{(S_r - \boldsymbol{v}_r'\boldsymbol{\mu}_0)^2}{\lambda_r}, \tag{5.3.2}$$

which follows the chi-square distribution with k degrees of freedom (denoted by χ_k^2) if the process is in control. Hence, the upper control limit of this T^2 chart is the $100(1-a)$ percentile of χ_k^2.

In the following, we will compare the Oja chart (i.e., r-chart) with the control charts proposed in Shiau et al. (2009) for Phase II monitoring. For profile monitoring, they chose $k = 3$ since the first three principal components account for a high enough level of variation. Figures 7-9 show the ARL curves of the control charts under comparison for the shifts in parameters I, M, and N, respectively. We observe the following:

- In Figure 7, we observe that the power of the Oja chart is only better than the PC2 chart and it is between the PC1 chart and PC2 chart. From Shiau et al. (2009), we know the mode of variation on I-shift is mostly caught by PC3 and some by PC1. Because we only consider PC1 and PC2 in our Oja chart, the Oja chart does not have as much power as the PC3 chart.

- In Figure 8, we observe that the power of the Oja chart is better than the other 5 control charts except for the small shift size $\beta < 0.50$, where these control charts are not too far from each other except for the PC3 chart. One reason for this could be that the Oja chart uses the power of PC1 and PC2 simultaneously.

- In Figure 9, we observe that the power of the Oja chart is better than the other 5 control charts when the shift size γ is between 0.75 and 2.00. For other magnitudes of shift, the T^2 chart has the power better than the other charts. The same situation in Figure 8 happens here: these control charts are not too far from each other except for the PC3 chart. The PC3 chart has a strange ARL curve in this figure. Shiau et al. (2009) explained that this may be caused by the fact that the shift in the mean vector when projected onto PC3, $\boldsymbol{v}_3'\boldsymbol{\delta}$, is not monotone in the shift multiple $\boldsymbol{\delta}$.

In summary, for this example, our proposed r-chart is comparable with the control charts proposed in Shiau et al. (2009) for Phase II monitoring. We remark that the example we choose here for comparison is to the advantage of the methods in Shiau et

al. (2009) in the sense that it satisfies the distributional assumptions under which the methods in Shiau et al. (2009) are based on. Nevertheless, our nonparametric chart performs almost equally well and sometimes slightly better.

5.3.2 ARL Comparisons for r-Chart, Q-Chart and DDMA-Chart

Figures 10-12 show the subgroups version of ARL curves of r-chart, Q-chart and DDMA-chart for shifts in parameters I, M, and N, respectively. According to these figures, we observe the following:

- From Figures 10-12, we note that the behavior of Q-charts are somewhat peculiar. When the process is in control, Q-charts with subgroup sizes 2, 4, and 6 have ARL_0 20.87, 21.87, 23.75, respectively, all larger than 20. Moreover, for small shift sizes (α=0.25,...,1.00, β=0.25, γ=0.25), the ARL_1 increases when the subgroup size (q) of the specific chart increases. This seems against our intuition. Fortunately, when shift size gets larger, the chart starts behaving as what we would expect. We have no good reasons for this but the asymptotic theory under which the control limit of the Q-chart was constructed as given in Liu (1995). The asymptotic theory requires the minimum between the reference sample size (m) and the subgroup size (q) to be infinite. But in our simulation, this q is only 2, 4, or 6, which is far from infinity.

- For all three types of shifts, the DDMA-charts perform the best among the three types of charts. Among DDMA-charts, D6 > D4 > D2, as expected. Here ">" means "performs better than".

- Despite the peculiar behavior of the Q-charts for small shift sizes, the ARL behaviors of all charts are pretty similar for all three types of shifts.

- When the shift size is large enough, the performance order of the charts is DDMA-charts > Q-charts > r-chart in general, with D2 sometimes intervening with Q4 and Q6.

- The performances of these control charts are very similar for M-shift and N-shift except that ARL values drop faster for N-shift.

To summarize, the control limit for the Q-chart needs to be modified to achieve an about-right false-alarm rate. The subgroup size is also an important issue when we adopt

the Q-chart or the DDMA-chart for Phase II monitoring. Here, their performances among different subgroup sizes, $q = 2, 4, 6$, behave as what we would expect for the DDMA-chart. As for the Q-chart, it starts behaving as what we would expect when shift size gets larger.

Since when using subgroups with different sizes, plotting one point on the control chart may require different number of profiles, it would be fairer to define run length as the number of profiles taken when the out-of-control alarm signals.

Figures 13-18 show the individual profiles version of ARL curves of Q-charts and DDMA-charts for shifts in parameters I, M, and N, respectively. They express the ARL in terms of the expected number of individual profiles sampled rather than the number of subgroups taken to detect a shift. Here, Figures 13-15 are for Q-charts and Figures 16-18 are for DDMA-charts. According to these figures, we observe the following:

- For Q-charts, we note that control chart with smaller subgroup size has smaller ARL values for all shift sizes from 0 to 3. Their performances do not behave as what we would expect when shift sizes are small.

- For DDMA-charts, we note that charts with different subgroup sizes ($q = 2, 4, 6$) perform differently (better or worse) at different shift sizes. Specifically, control chart with smaller subgroup size performs better (with smaller ARL) when the shift size becoming larger and larger. This is because the moving average has an overhead of size q at the beginning of the process monitoring.

- Figure 19 compares the average number of profiles needed to detect a shift for all the charts for I-, M-, N-shift, respectively. According to their performances, we recommend the DDMA-chart when shift size is small and the r-chart when shift size is large.

6 Conclusions

In this study, we propose and study some nonparametric schemes for Phase II profile monitoring based on the notion of data depth. We utilize data smoothing techniques to filter out random noises from raw sample profiles, principal component analysis to convert smoothed profiles to score-vectors of lower dimension, and then data-depth-based control charts to monitor incoming profiles.

According to the ARL performance comparison between our monitoring scheme and those proposed in Shiau et al. (2009) with the simulated aspartame example, we observe that our proposed monitoring scheme based on r-chart is comparable. Furthermore, the monitoring scheme based on DDMA-chart performs better than the monitoring scheme based on Q-chart and the monitoring scheme based on Q-chart performs better than the one based on r-chart in general.

The most important feature of the control schemes proposed in this paper is that they are nonparametric in the sense that no distributional assumptions are required for the profiles under monitoring. This broadens the scopes of applications of the proposed schemes. And interestingly, when compared with the profile monitoring schemes proposed in Shiau et al. (2009), the new nonparametric schemes perform equally well or even slightly better with an example satisfying the distributional assumptions of the existing schemes.

For Phase I analysis, we provide a heuristic diagnostic procedure with the objective of spotting unusual profiles in the Phase I data so that we can obtain a more reliable data set to create the reference sample for the subsequent construction of data depth control charts. The procedure is not a control chart because we cannot set a control limit that can control, say, the false-alarm rate.

Clustering depth values through different clustering methods in Phase I analysis and adopting other depth functions for Phase II profile monitoring could be interesting future studies. Another potential direction is to modify the lower control limit for the Q-chart to achieve an about-right false-alarm rate.

References

[1] Johnson, R. A. and Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis.* 6th ed. Prentice Hall, New Jersey.

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

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

[4] Liu, R. Y. (1990). "On a Notion of Data Depth Based on Random Simplices". *The Annals of Statistics.* 18, 405-414.

[5] Liu, R. Y. (1995). "Control Charts for Multivariate Processes". *Journal of the American Statistical Association.* 90, 1380-1387.

[6] Liu, R. Y. and Singh, K. (1993). "A Quality Index Based on Data Depth and Multivariate Rank Tests". *Journal of the American Statistical Association.* 88, 252-260.

[7] Liu, R. Y., Singh, K., and Teng, J. H. (2004). "DDMA-charts: Nonparametric Multivariate Moving Average Control Charts Based on Data Depth". *Allgemeines Statistisches Archiv.* 88, 235-258.

[8] Mosler, K. C. (2002). *Multivariate Dispersion, Central Regions and Depth: The Lift Zonoid Approach.* Springer, New York.

[9] Oja, H. (1983). "Descriptive Statistics for Multivariate Distributions". *Statistics and Probability Letters.* 1, 327-332.

[10] Shiau, J.-J. H., Huang, H.-L., Lin, S.-H., and Tsai, M.-Y. (2009). "Monitoring Nonlinear Profiles with Random Effects by Nonparametric Regression". *Communications in Statistics-Theory and Methods.* 38, 1664-1679.

[11] Shiau, J.-J. H., Lin, S.-H., and Chen, Y.-C. (2006). "Monitoring Linear Profiles Based on a Random-effect Model". Technical Report. Institute of statistics, National Chiao Tung University.

[12] Shiau, J.-J. H., Yen, C.-L., and Feng, Y.-W. (2006). "A New Robust Method for Phase I Monitoring of Nonlinear Profiles". Technical Report. Institute of statistics, National Chiao Tung University.

[13] Williams, J. D., Birch, J. B., Woodall, W. H., and Ferry, N. M. (2007). "Statistical Monitoring of Heteroscedastic Dose-Response Profiles from High-throughput Screening". *Journal of Agricultural, Biological and Environmental Statistics.* 2, 216-235.

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

[15] Williams, J. D., Woodall, W. H., and Birch, J. B. (2003). "Phase I Analysis of Nonlinear Product and Process Quality profiles". Technical Report No. 03-5. Department of Statistics, Virginia Polytechnic Institute and State University.

[16] Zuo, Y. and Serfling, R. (2000). "General Notions of Statistical Depth Function". *The Annals of Statistics.* 28, 461-482.

Table 1: Diagnostic results of Bioassay Data (smoothing parameter is 0.3).

	diagnosis								
	first				**second(removing first)**				
	cumulative			**weeks**	**cumulative**				**weeks**
spar	PC1	PC2	PC3		PC1	PC2	PC3	PC4	
0.3	0.550	0.849		20,22,24 34,45,46	0.616	0.811			32
0.3	0.550	0.849		20,22,24 34,45,46	0.616	0.811	0.910		26
0.3	0.550	0.849		20,22,24 34,45,46	0.616	0.811	0.910	0.980	26,32
0.3	0.550	0.849	0.941	34	0.598	0.868			20,22,24,45 46
0.3	0.550	0.849	0.941	34	0.598	0.868	0.942		20,22,24,26 32,45,46,48
0.3	0.550	0.849	0.941	34	0.598	0.868	0.942	0.977	32

Table 2: Diagnostic results of Bioassay Data (smoothing parameter is 0.4).

	diagnosis								
	first				**second(removing first)**				
	cumulative			**weeks**	**cumulative**				**weeks**
spar	PC1	PC2	PC3		PC1	PC2	PC3	PC4	
0.4	0.628	0.925		20,22,24 34,45,46	0.701	0.884			5,13,26,32 44
0.4	0.628	0.925		20,22,24 34,45,46	0.701	0.884	0.970		26
0.4	0.628	0.925		20,22,24 34,45,46	0.701	0.884	0.970	0.992	26,32
0.4	0.628	0.925	0.976	24	0.663	0.920			20,22,34,45 46
0.4	0.628	0.925	0.976	24	0.663	0.920	0.974		20,22,26,34 45,46,48
0.4	0.628	0.925	0.976	24	0.663	0.920	0.974	0.989	20,22,34,45

Table 3: Diagnostic results of Bioassay Data (smoothing parameter is 0.5).

	diagnosis								
	first				**second(removing first)**				
	cumulative			**weeks**	**cumulative**				**weeks**
spar	PC1	PC2	PC3		PC1	PC2	PC3	PC4	
0.5	0.690	0.969		20,22,24 34,45,46	0.763	0.941			26
0.5	0.690	0.969		20,22,24 34,45,46	0.763	0.941	0.995		26,48
0.5	0.690	0.969		20,22,24 34,45,46	0.763	0.941	0.995	0.999	16,26,32,48
0.5	0.690	0.969	0.995	24	0.721	0.966			20,22,34,45 46
0.5	0.690	0.969	0.995	24	0.721	0.966	0.995		20,22,26,34 45,46,48
0.5	0.690	0.969	0.995	24	0.721	0.966	0.995	0.999	20,22,45

Table 4: Diagnostic results of VDP Data with different smoothing parameters.

	diagnosis					
	first			**second(removing first)**		
	cumulative		**profiles**	**cumulative**		**profiles**
spar	PC1	PC2		PC1	PC2	
0.01	0.845	0.952	6,10	0.824	0.938	2,3,16,17,20
0.1	0.845	0.953	6,10	0.825	0.938	2,3,16,17,20
0.5	0.853	0.961	6,10	0.835	0.949	2,3,16,17,20
1	0.887	0.989	6,10	0.878	0.985	2,3,16,17,20

Table 5: ARL values and their standard errors (in parentheses) of r-chart, Q-charts, and DDMA-charts for various shift sizes of I-shift.

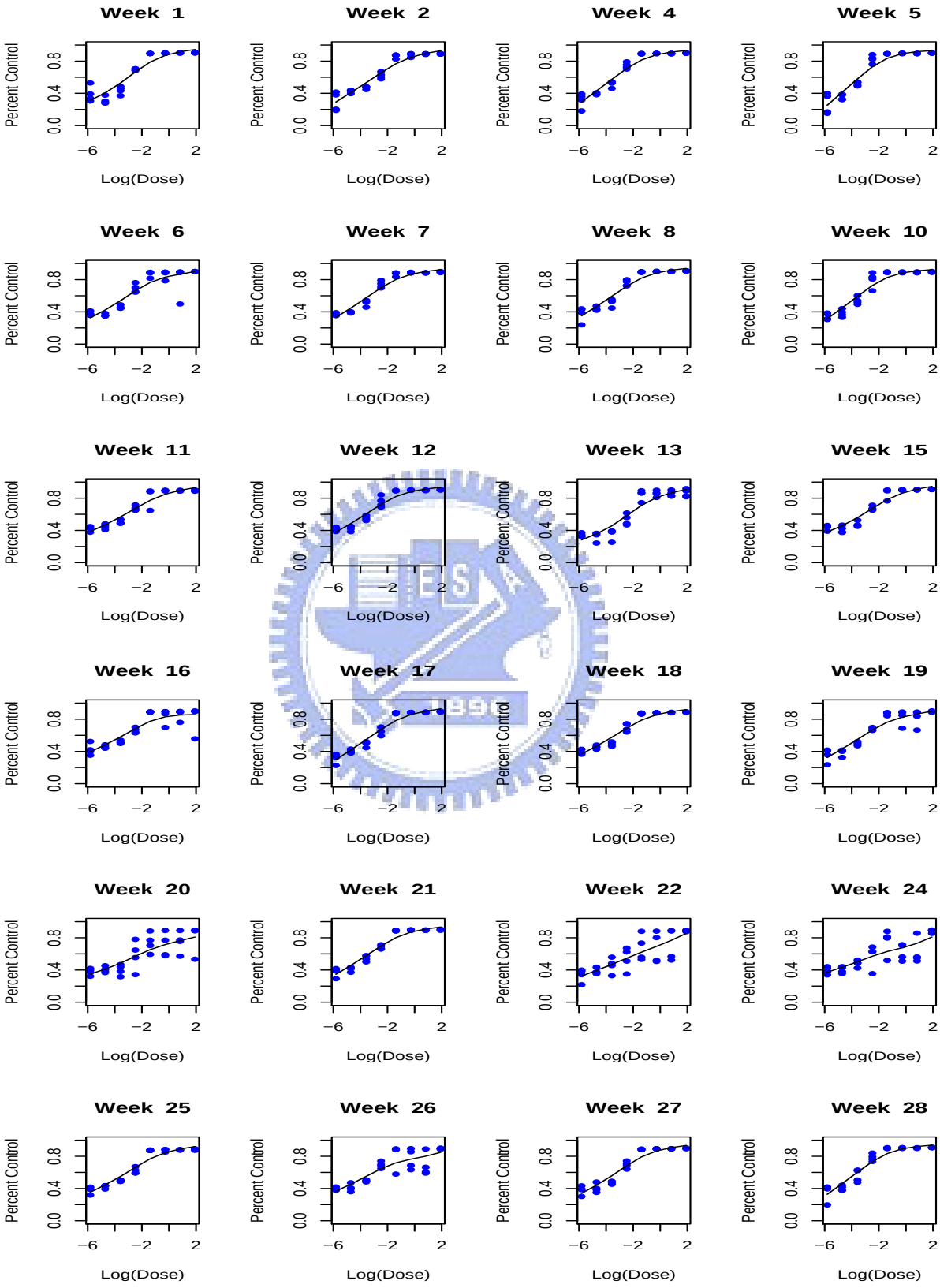
α	r-chart	Q2	Q4	Q6	D2	D4	D6
0.00	20.42	20.87	21.87	23.75	20.34	20.28	20.12
	(0.093)	(0.116)	(0.198)	(0.324)	(0.105)	(0.137)	(0.166)
0.25	20.15	20.76	21.88	23.59	20.01	19.66	19.40
	(0.091)	(0.119)	(0.185)	(0.304)	(0.101)	(0.128)	(0.159)
0.50	19.89	20.50	21.46	22.73	19.39	18.55	17.76
	(0.090)	(0.116)	(0.229)	(0.280)	(0.099)	(0.120)	(0.136)
0.75	19.41	19.85	20.34	21.39	18.41	16.75	15.35
	(0.089)	(0.110)	(0.157)	(0.241)	(0.091)	(0.108)	(0.118)
1.00	18.56	18.82	19.08	19.90	17.16	14.77	12.89
	(0.081)	(0.100)	(0.153)	(0.257)	(0.084)	(0.094)	(0.095)
1.25	17.77	17.95	17.80	18.06	15.69	12.79	10.63
	(0.076)	(0.091)	(0.136)	(0.178)	(0.076)	(0.076)	(0.073)
1.50	16.93	16.86	16.38	16.53	14.32	10.89	8.690
	(0.072)	(0.083)	(0.119)	(0.170)	(0.065)	(0.060)	(0.055)
1.75	15.86	15.68	15.01	14.63	12.78	9.160	7.050
	(0.064)	(0.076)	(0.110)	(0.132)	(0.056)	(0.049)	(0.042)
2.00	14.82	14.50	13.41	12.83	11.48	7.820	5.820
	(0.060)	(0.069)	(0.085)	(0.102)	(0.050)	(0.039)	(0.033)
2.25	13.76	13.16	11.95	11.27	10.12	6.530	4.750
	(0.055)	(0.057)	(0.071)	(0.087)	(0.040)	(0.029)	(0.023)
2.50	12.86	12.22	10.81	9.850	9.150	5.630	4.010
	(0.051)	(0.052)	(0.059)	(0.051)	(0.037)	(0.025)	(0.019)
2.75	11.75	11.02	9.580	8.510	8.070	4.820	3.380
	(0.042)	(0.045)	(0.051)	(0.054)	(0.031)	(0.021)	(0.015)
3.00	10.87	10.04	8.500	7.410	7.190	4.160	2.880
	(0.039)	(0.040)	(0.044)	(0.043)	(0.026)	(0.016)	(0.011)

Table 6: ARL values and their standard errors (in parentheses) of r-chart, Q-charts, and DDMA-charts for various shift sizes of M-shift.

β	r-chart	Q2	Q4	Q6	D2	D4	D6
0.00	20.42	20.87	21.87	23.75	20.34	20.28	20.12
	(0.093)	(0.116)	(0.198)	(0.324)	(0.105)	(0.137)	(0.166)
0.25	18.75	18.98	19.37	20.18	17.10	14.54	12.73
	(0.083)	(0.099)	(0.166)	(0.266)	(0.084)	(0.088)	(0.094)
0.50	14.72	14.51	13.49	12.95	11.50	7.780	5.800
	(0.059)	(0.068)	(0.093)	(0.108)	(0.049)	(0.038)	(0.030)
0.75	10.92	10.07	8.510	7.420	7.190	4.170	2.910
	(0.039)	(0.039)	(0.042)	(0.043)	(0.026)	(0.016)	(0.011)
1.00	7.840	6.790	5.240	4.250	4.620	2.500	1.770
	(0.025)	(0.022)	(0.021)	(0.018)	(0.015)	(0.007)	(0.004)
1.25	5.680	4.640	3.330	2.610	3.140	1.720	1.300
	(0.016)	(0.012)	(0.010)	(0.008)	(0.008)	(0.003)	(0.002)
1.50	4.180	3.260	2.250	1.770	2.260	1.330	1.100
	(0.010)	(0.007)	(0.005)	(0.004)	(0.004)	(0.002)	(0.001)
1.75	3.190	2.430	1.680	1.360	1.750	1.140	1.030
	(0.007)	(0.004)	(0.003)	(0.002)	(0.003)	(0.001)	(0.000)
2.00	2.510	1.890	1.350	1.150	1.430	1.050	1.010
	(0.004)	(0.003)	(0.002)	(0.001)	(0.002)	(0.000)	(0.000)
2.25	2.050	1.560	1.180	1.060	1.250	1.020	1.000
	(0.003)	(0.002)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
2.50	1.720	1.350	1.080	1.020	1.130	1.000	1.000
	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
2.75	1.490	1.210	1.030	1.000	1.070	1.000	1.000
	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
3.00	1.330	1.120	1.010	1.000	1.030	1.000	1.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 7: ARL values and their standard errors (in parentheses) of r-chart, Q-charts, and DDMA-charts for various shift sizes of N-shift.

γ	r-chart	Q2	Q4	Q6	D2	D4	D6
0.00	20.42	20.87	21.87	23.75	20.34	20.28	20.12
	(0.093)	(0.116)	(0.198)	(0.324)	(0.105)	(0.137)	(0.166)
0.25	18.88	19.19	19.35	20.19	17.65	15.25	13.41
	(0.086)	(0.108)	(0.156)	(0.243)	(0.091)	(0.096)	(0.097)
0.50	15.15	14.80	13.97	13.41	12.00	8.230	6.180
	(0.063)	(0.067)	(0.091)	(0.110)	(0.053)	(0.042)	(0.034)
0.75	10.81	10.01	8.390	7.350	7.150	4.110	2.870
	(0.039)	(0.039)	(0.042)	(0.041)	(0.026)	(0.016)	(0.011)
1.00	7.240	6.140	4.660	3.750	4.200	2.260	1.610
	(0.023)	(0.019)	(0.017)	(0.015)	(0.012)	(0.006)	(0.004)
1.25	4.790	3.800	2.660	2.080	2.600	1.470	1.170
	(0.013)	(0.009)	(0.007)	(0.005)	(0.006)	(0.003)	(0.001)
1.50	3.170	2.400	1.660	1.340	1.730	1.140	1.030
	(0.006)	(0.004)	(0.003)	(0.002)	(0.003)	(0.001)	(0.000)
1.75	2.180	1.660	1.220	1.080	1.300	1.030	1.000
	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
2.00	1.600	1.280	1.050	1.010	1.100	1.000	1.000
	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2.25	1.270	1.090	1.010	1.000	1.020	1.000	1.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2.50	1.100	1.020	1.000	1.000	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2.75	1.030	1.000	1.000	1.000	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
3.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)


Week 1
Week 2
Week 4
Week 5
Week 6
Week 7
Week 8
Week 10
Week 11
Week 12
Week 13
Week 15
Week 16
Week 17
Week 18
Week 19
Week 20
Week 21
Week 22
Week 24
Week 25
Week 26
Week 27
Week 28
Percent Control
0.0
0.4
0.8
−6
−2
2
Log(Dose)

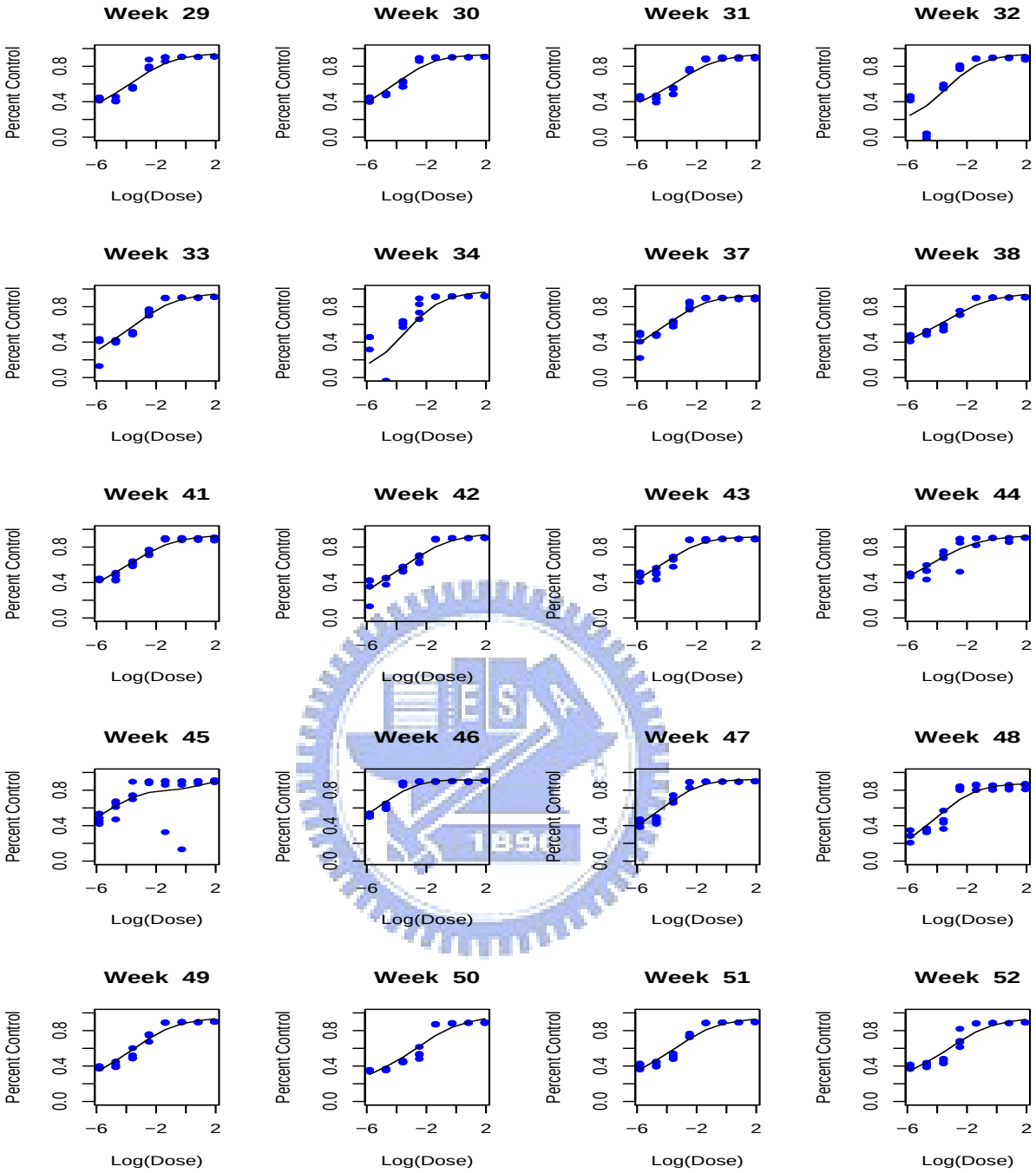


Figure 1: DuPont Dose-Response Data for 44 weeks with their smoothing spline fittings (smoothing parameter=0.5).

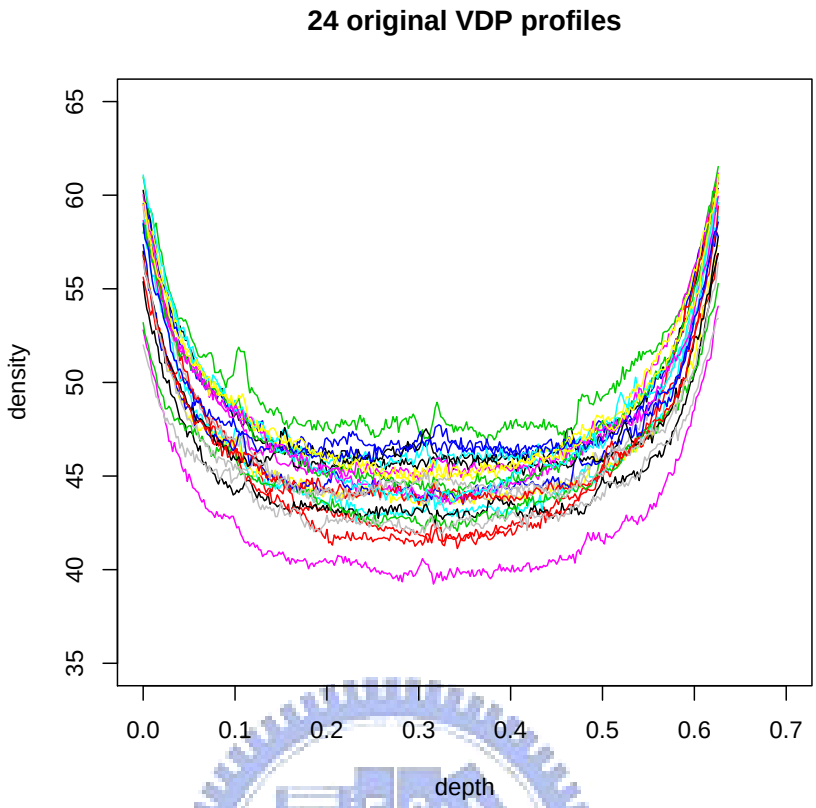


Figure 2: 24 VDP profiles.

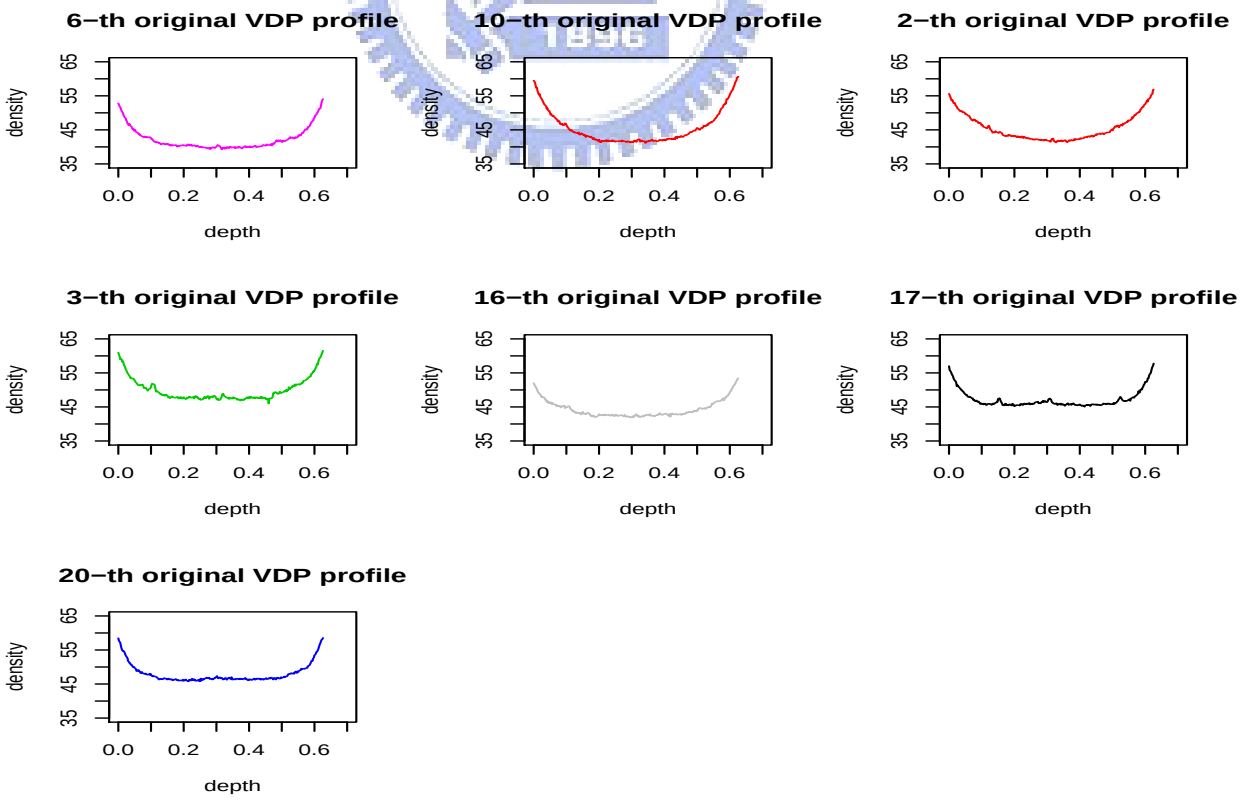


Figure 3: 7 potentially out-of-control profiles.

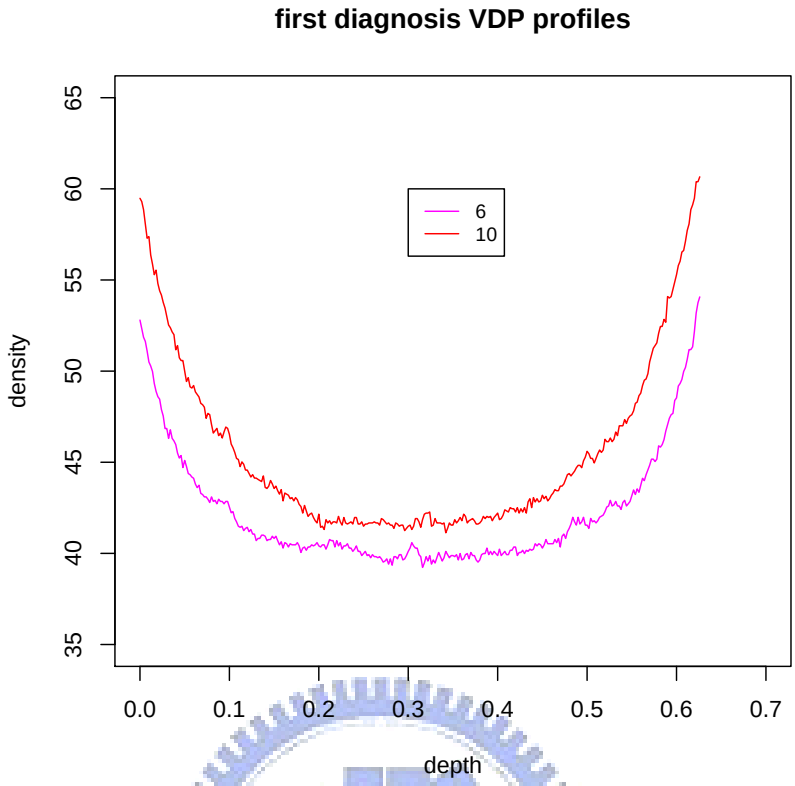


Figure 4: 2 potentially out-of-control profiles from the first diagnosis.

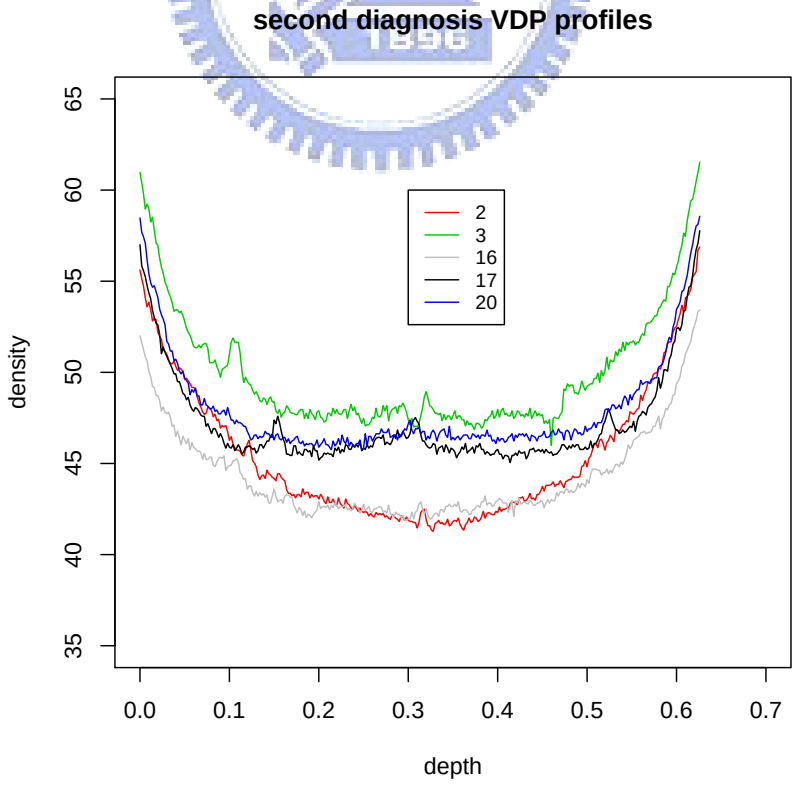


Figure 5: 5 potentially out-of-control profiles from the second diagnosis.

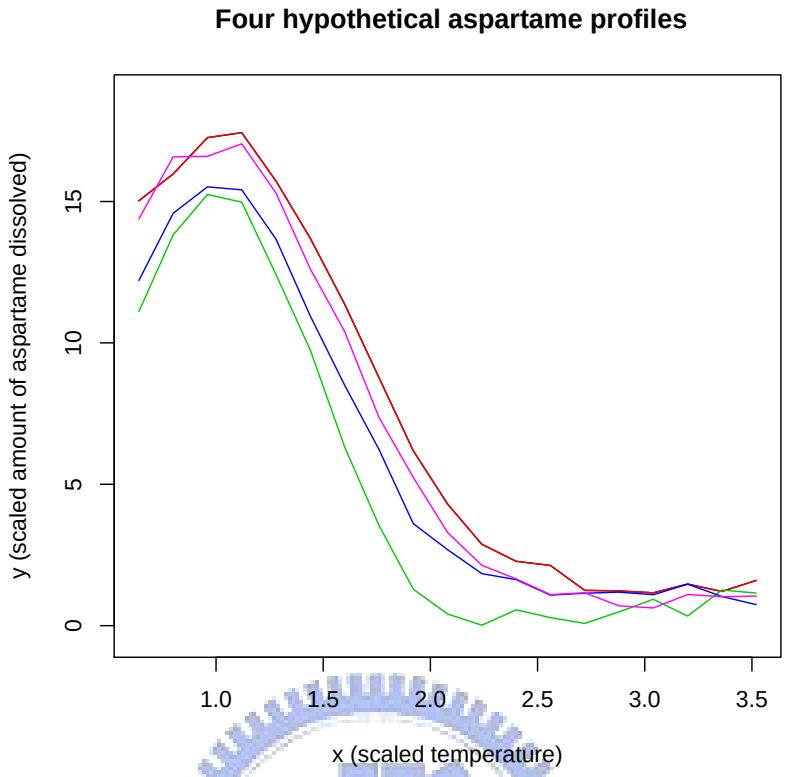


Figure 6: 4 hypothetical aspartame profiles.

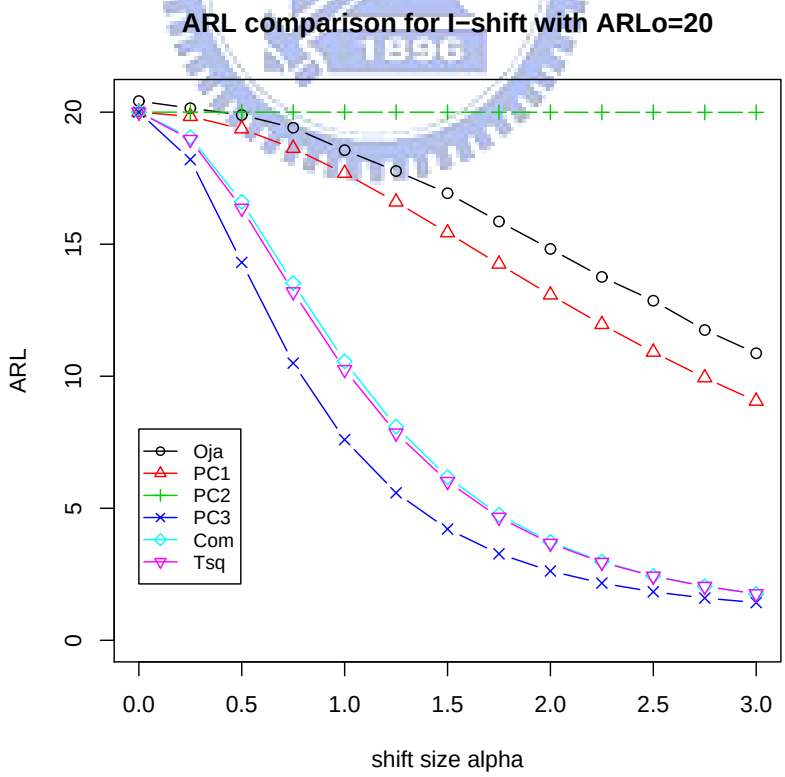


Figure 7: ARL comparisons for the I-shift.

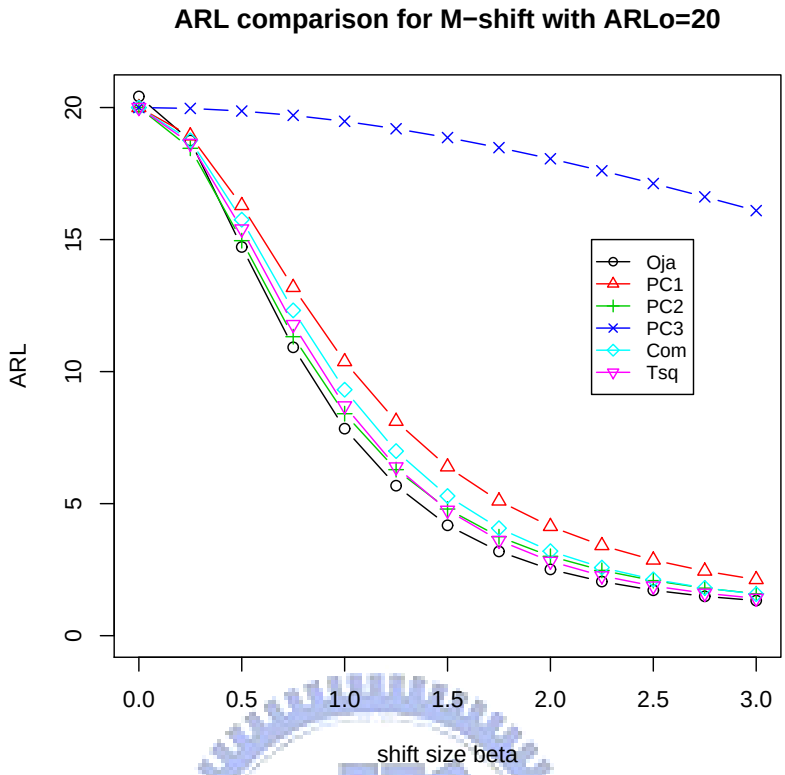


Figure 8: ARL comparisons for the M-shift.

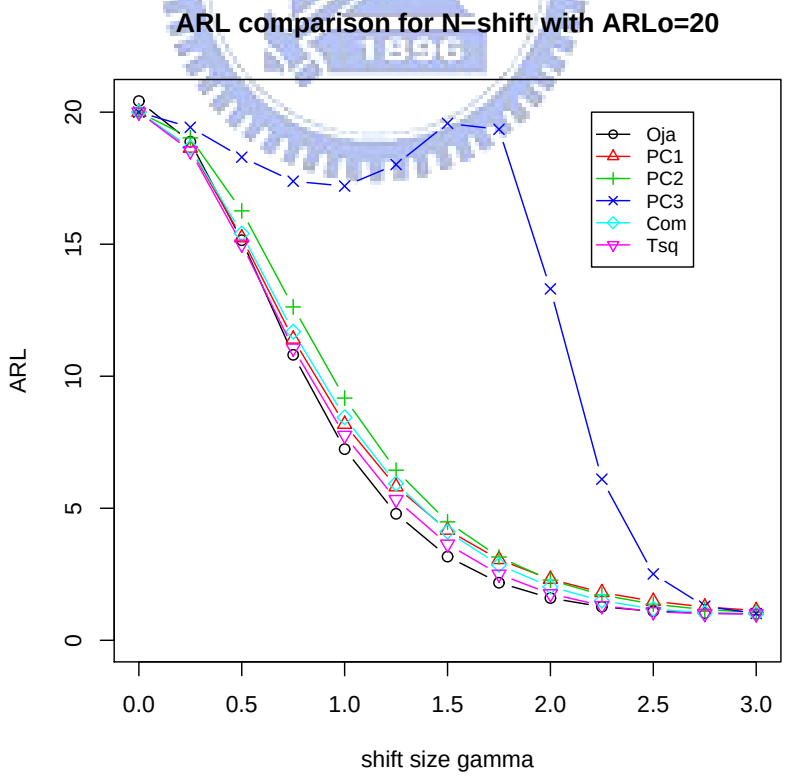


Figure 9: ARL comparisons for the N-shift.

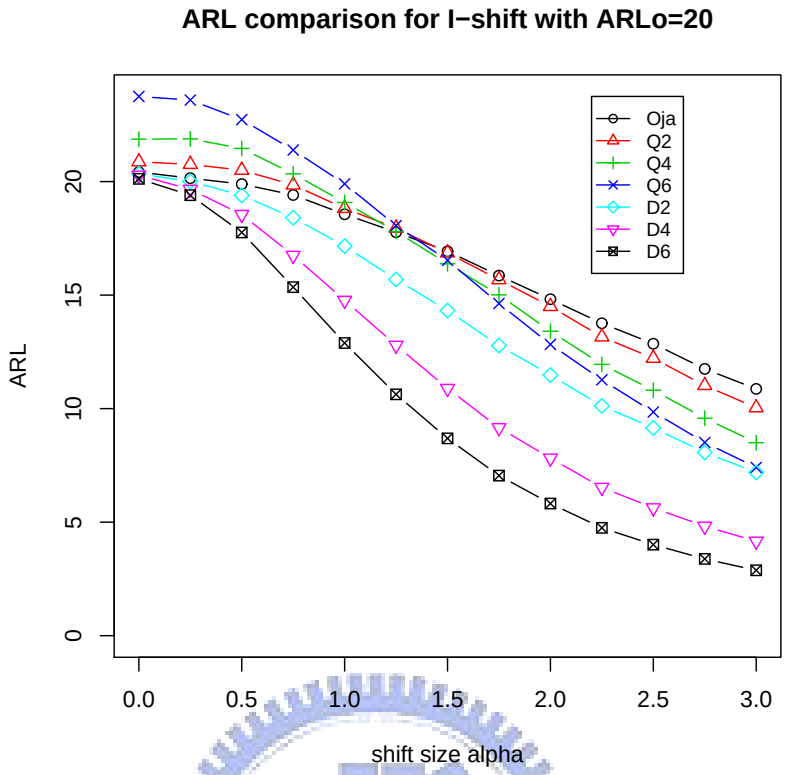


Figure 10: ARL comparisons of r, Q, and DDMA charts for the I-shift.

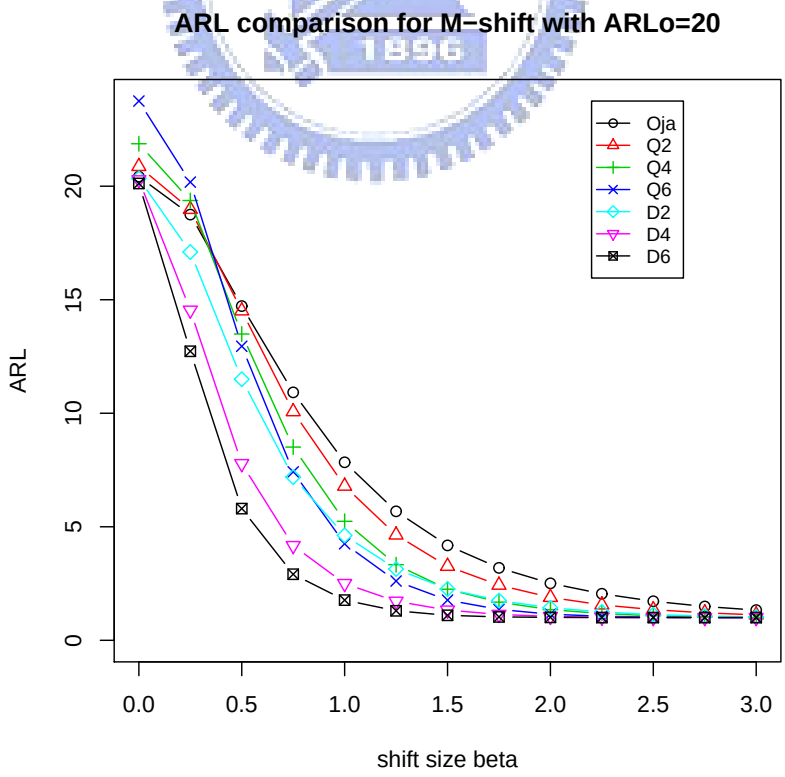


Figure 11: ARL comparisons of r, Q, and DDMA charts for the M-shift.

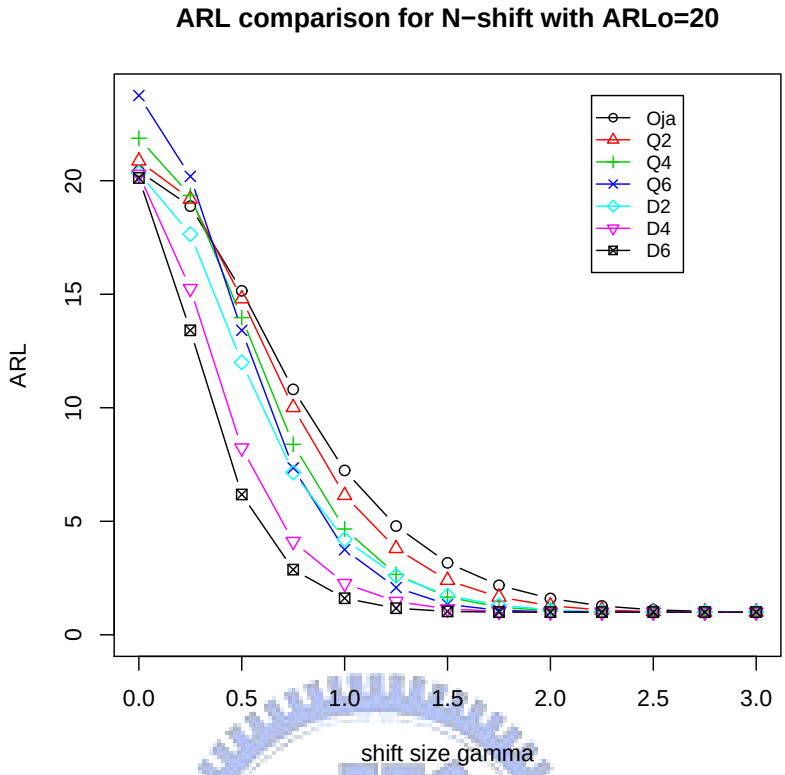


Figure 12: ARL comparisons of r, Q, and DDMA charts for the N-shift.

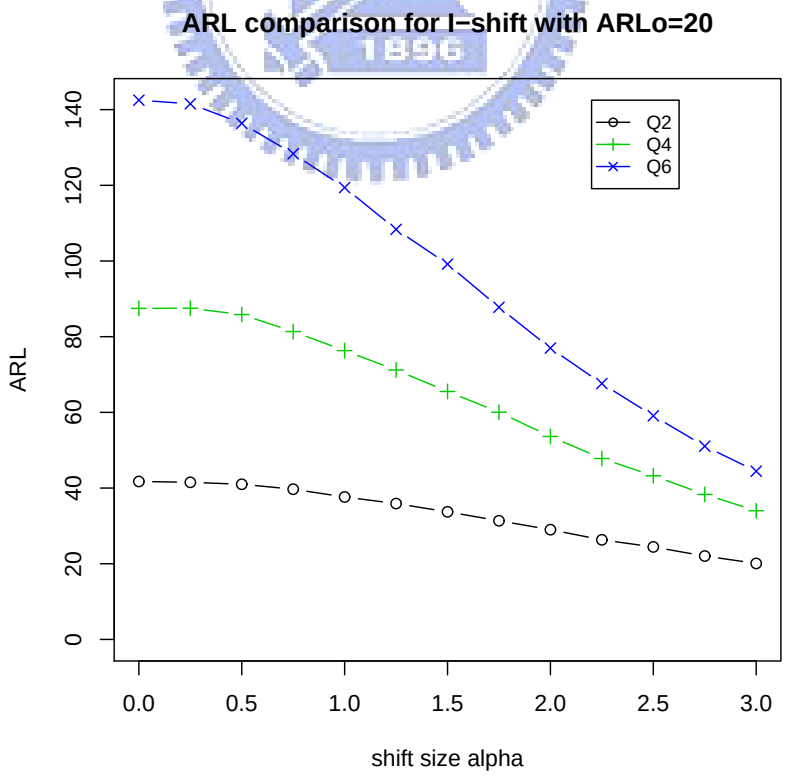


Figure 13: ARL (individual profiles version) comparisons of Q charts for the I-shift.

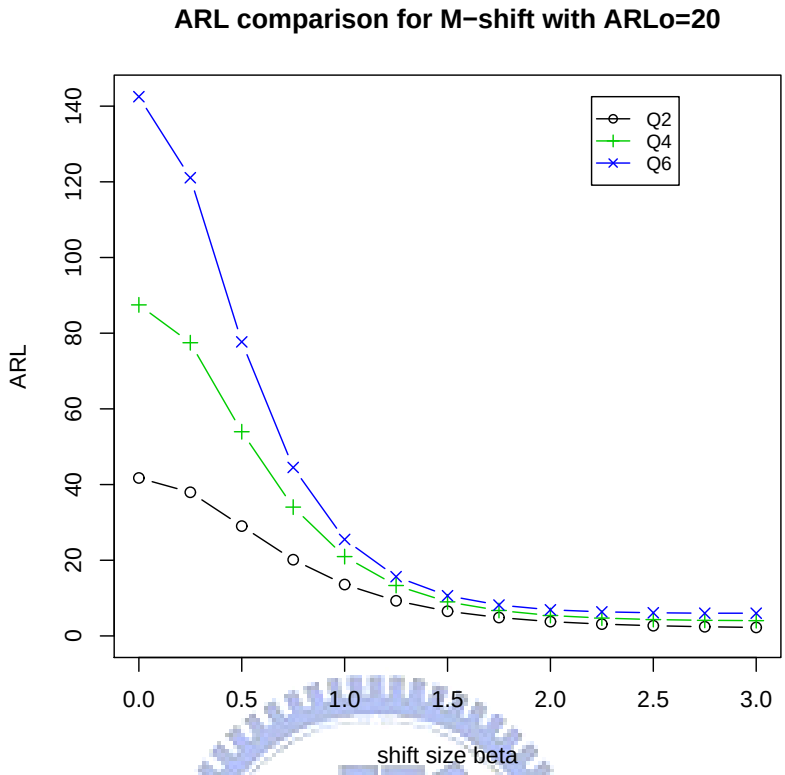


Figure 14: ARL (individual profiles version) comparisons of Q charts for the M-shift.

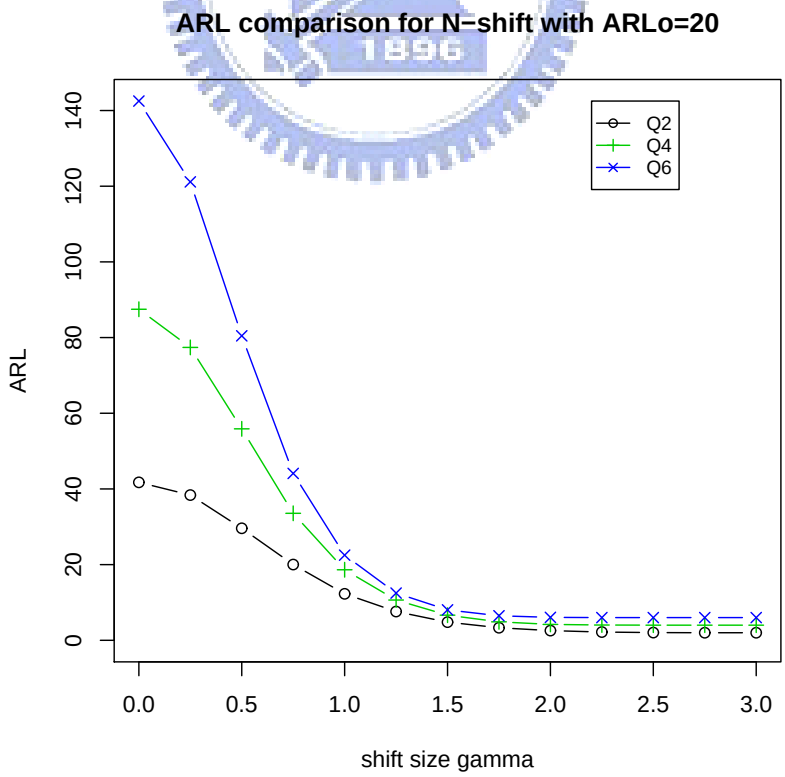


Figure 15: ARL (individual profiles version) comparisons of Q charts for the N-shift.

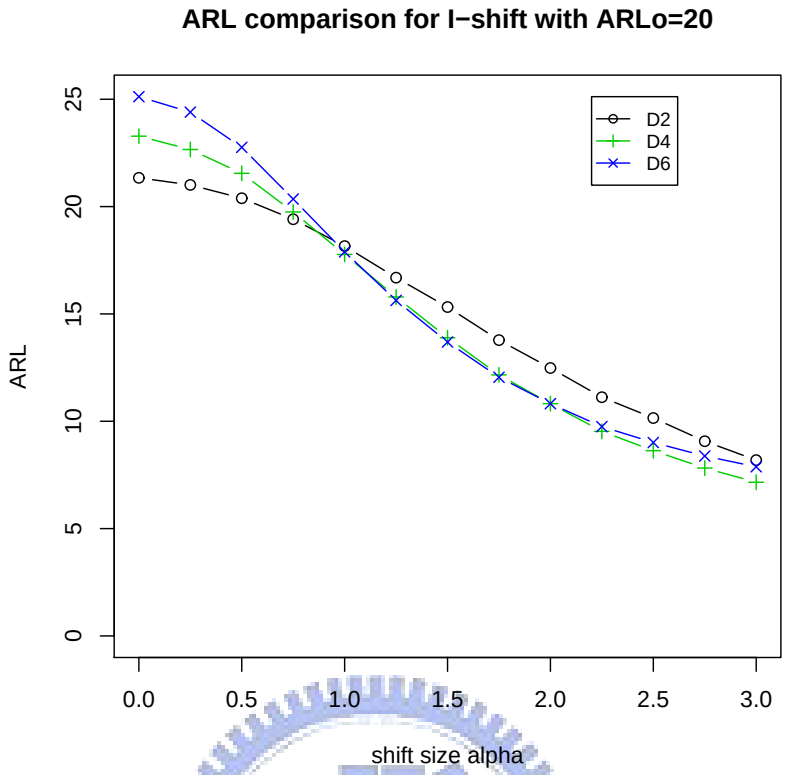


Figure 16: ARL (individual profiles version) comparisons of DDMA charts for the I-shift.

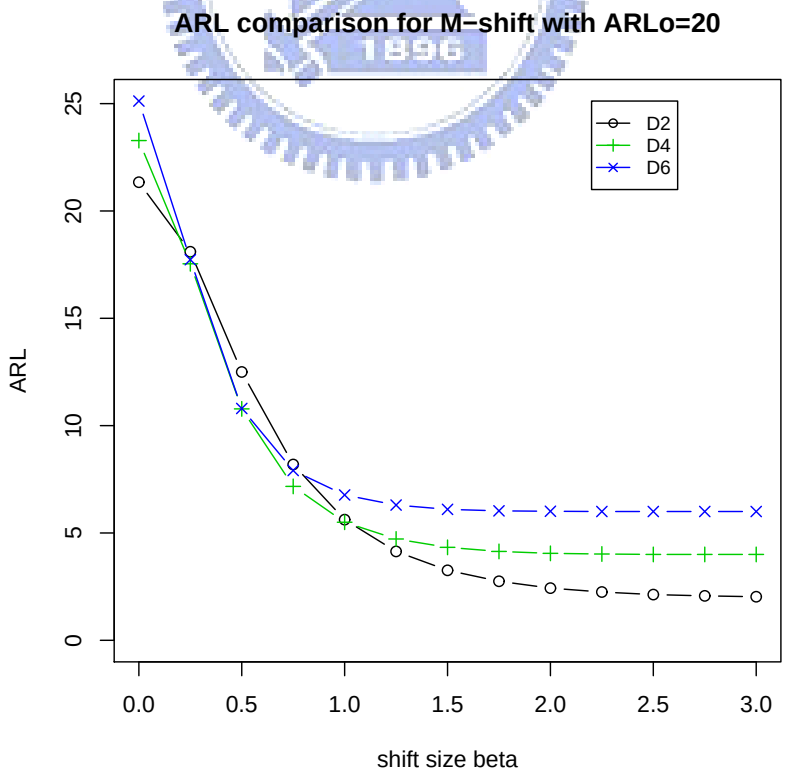


Figure 17: ARL (individual profiles version) comparisons of DDMA charts for the M-shift.

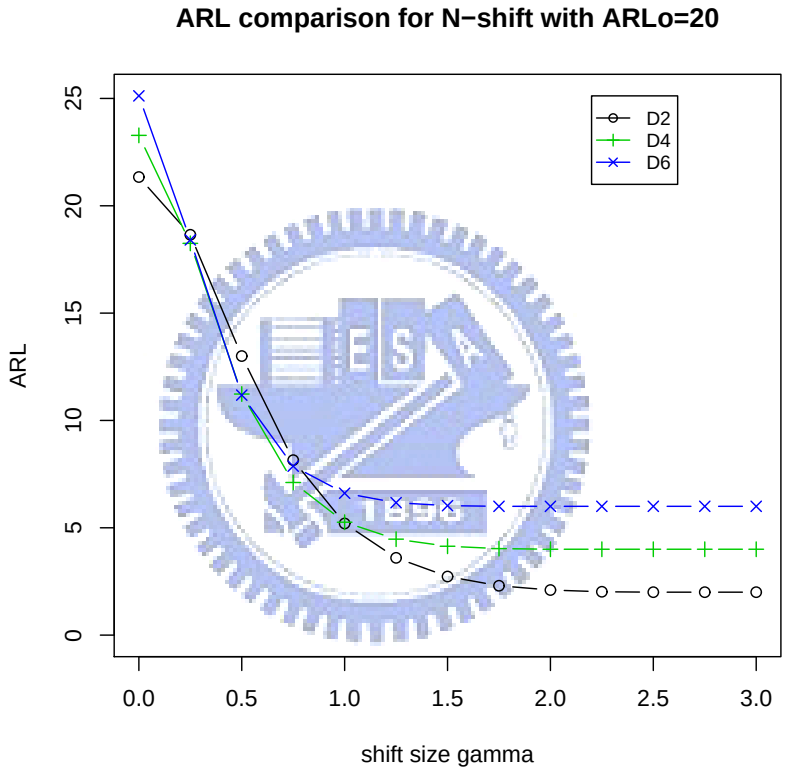


Figure 18: ARL (individual profiles version) comparisons of DDMA charts for the N-shift.

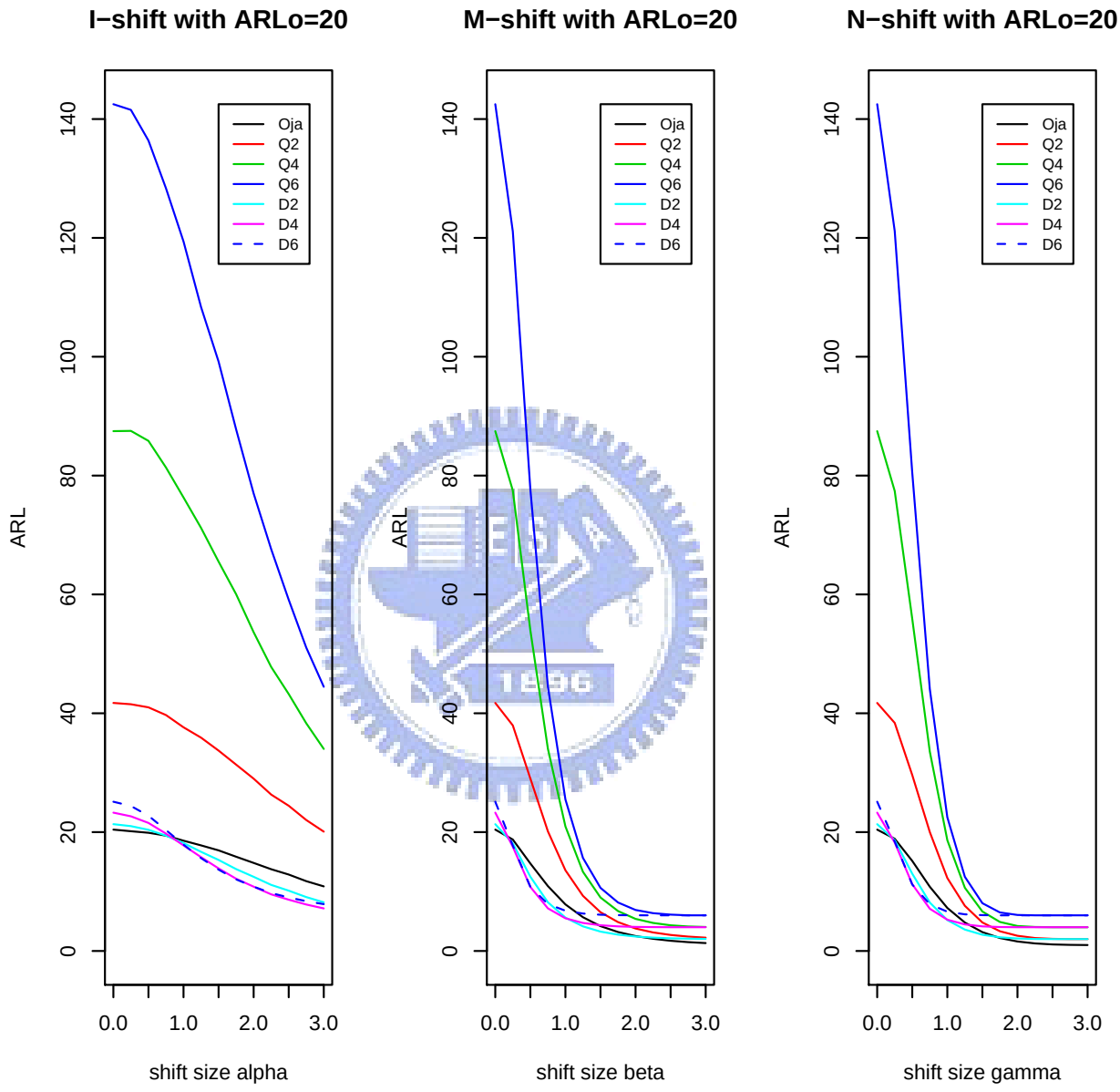


Figure 19: ARL (individual profiles version) comparisons of all charts for all three types of shifts.