國立交通大學

統計學研究所

碩 士 論 文

利用主成份分析監控剖面資料之研究

Profile Monitoring via Principal Components

研 究 生：王筱娟

指導教授：洪志真　博士

中 華 民 國 九 十 七 年 六 月

利用主成份分析監控剖面資料之研究

Profile Monitoring via Principal Component

研 究 生：王筱娟　　Student：Hshiao-Chuan Wang

指導教授：洪志真　　Advisor：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

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master

in

Statistics

June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

誌　　謝

兩年的碩士班生涯即將結束，馬上就要走入人生另一個旅程，回首剛回到校園進修時的期待，以及現在收成的喜悅，這一切都要感謝許多人對我的提攜與幫助。

這一篇論文能夠順利完成，首先要感謝我的指導老師—洪志真教授。感謝老師給予我細心的指導與叮嚀，並在數次的修改中，給予我細部的指導，論文方能夠如期的完成。除此之外，有感謝我的口試委員陳鄰安教授、黃榮臣教授及曾勝滄教授，在百忙之中為我口試，在口試時對於我的指導，更讓我發現本論文主題可以更深入探討的地方，這在我的工作領域上有很大的幫助，能夠真正達到學以致用的目標。

而我的同學們在這段時間給予了很多的協助，尤其是在我工作繁忙時，幫我解決了很多事務上的問題，讓我可以專注在論文的題目上。我要特別感謝怡玲和香菱，在課業及論文的撰寫給予我寶貴的建議及鼓勵。也感謝統計所上的同學，在碩二繁忙的日子裡帶給歡樂，減輕寫論文時的壓力。除此要特別感謝我們所上的助理郭碧芬小姐，她不辭勞苦幫我們處理很多行政事務上的問題。

最後要感謝我的家人給予我莫大的支持，與家人相處的時間很少，因為爸媽的體諒，才能專心完成學業。

另外還有很多曾經幫助過我的朋友，因為有大家的幫助，我才能有今天的成果。

2008 年 6 月　　王筱娟

利用主成份分析監控剖面資料之研究

研究生：王筱娟　　　　　　　　　　指導教授：洪志真博士

國立交通大學統計學研究所

摘要

在多數的品質監控應用上，我們使用一個或多個品質特性（一個變數或多個變數）來測量製程的品質。然而，在許多情況下，有興趣的反應變數不是一個變數，而是一個函數。這個函數稱為剖面（profile)。本研究中我們採用無母數的模型去監控剖面資料而不是用有母數的模型。

對於剖面資料間可允許的變化，我們利用主成份分析分析剖面資料變異的結構。我們利用剖面資料所包含的主成份分數(Score)的資訊監控剖面資料。傳統的T^2統計量視每一個主成份同等重要。然而，對於貢獻解釋剖面的變異較少的主成份而言，當他們的分數與製程在控制中的值有顯著不同時，在實務上我們未必希望因此而發出失控警訊。換句話說，這些主成份的影響在統計上是顯著的，而在實際上卻並不重要。因此我們提出一個新的監控方法，其只對於重要的主成份改變時是敏感的。我們用模擬來檢驗此方法的效能。

Profile Monitoring via Principal Components

Student: Hshiao-Chuan Wang Advisor: Jyh-Jen Horng Shiau

Institute of Statistics

National Chiao Tung University

Hsinchu, Taiwan

Abstract

In most of quality control applications, we use one or multiple quality characteristics (a single univariate or multivariate variable) to measure the process quality. However, in many situations, the response of interest is not a single variable but a function of one or more explanatory variables. This functional response is called a profile. We adopt nonparametric models to monitor profiles rather than parametric models.

For profiles with allowable profile-to-profile variation, we utilize the technique of principal components analysis to analyze the covariance structure in profiles. We consider monitoring profile by using the information contained in the principal component scores. Classical T^2 statistics treat each principal component equally important. However, for some principal components with little contribution in explaining the variability of profiles, it may not be desirable to signal out-of-control alarms when their scores are significantly different from the in-control values. In other words, effects of these components are statistically significant but not practically significant. Thus, we propose a new monitoring scheme that is only sensitive to shifts on "important" components. Simulation studies demonstrate the efficacy of the method.

Contents

1 Introduction **1**

2 Review on Background Methodologies **3**

2.1 Nonparametric Regression 3

2.2 Functional Data Analysis 4

2.3 Principal Components Analysis 5

2.4 Functional PCA vs. PCA 7

3 Monitoring Schemes **9**

3.1 Monitoring Statistics 9

3.2 Control Limits 11

4 Simulation Studies **13**

4.1 Data Generation 13

4.2 Number of Principal Components 14

4.3 Simulation Results 14

5 Conclusion **16**

List of Tables

1 Eigenvalues for smoothed VDP data (proportion of explaining variation) . . 20

2 ARL for PC1 21

3 ARL for PC2 21

4 ARL for PC23 22

5 ARL for PC1(m=200) 22

6 ARL for PC2(m=200) 23

7 ARL for PC23(m=200) 23

8 ARL for PC1(m=300) 24

9 ARL for PC2(m=300) 24

10 ARL for PC23(m=300) 25

11 ARL for PC1(m=400) 25

12 ARL for PC2(m=400) 26

13 ARL for PC23(m=400) 26

14 ARL for PC1(m=500) 27

15 ARL for PC2(m=500) 27

16 ARL for PC23(m=500) 28

17 ARL for PC1(m=600) 28

18 ARL for PC2(m=600) 29

19 ARL for PC23(m=600) 29

List of Figures

1 Original 24 VDP-profiles 30

2 Smoothed VDP 30

3 Density for T_3 31

4 Power of PC1 (true VDP) 32

5 Power of PC2 (true VDP) 32

6 Power of PC23 (true VDP) 32

7 Power for PC1 with m=200 33

8 Power for PC2 with m=200 33

9 Power for PC23 with m=200 33

10 Power for PC1 with m=300 34

11 Power for PC2 with m=300 34

12 Power for PC23 with m=300 34

13 Power for PC1 with m=400 35

14 Power for PC2 with m=400 35

15 Power for PC23 with m=400 35

16 Power for PC1 with m=500 36

17 Power for PC2 with m=500 36

18 Power for PC23 with m=500 36

19 Power for PC1 with m=600 37

20 Power for PC2 with m=600 37

21 Power for PC23 with m=600 37

1 Introduction

Statistical Process control (SPC) has been widely applied in many areas, especially in industries. Classical methods for SPC usually 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 (univariate or multivariate) quality variable but a function of one or more explanatory variables. For example, for a product item, we collect a set of measurements over time, space, or levels of parameter settings in production or experiments, etc. This functional response data is called a profile. An example of profiles given in Kang and Albin (2000) is the dissolving process of aspartame, an artificial sweetener, which is characterized by the amount of aspantane dissolved per liter of water at different levels of temperature. Mestek et al. (1994) used linear functions to monitor the performance of the process for the calibration of a mass flow controller. An interesting example of nonlinear profiles was presented in Walker and Wright (2002). The density of a particle board is measured with a profilometer that uses a laser devices to take measurements at fixed depths across the thickness of the board. These measurements on the sample form the vertical density profile (VDP) of the board. The VDP dataset contains 24 profiles of vertical density and each profile consists of 314 measurements taken 0.002 inches apart. The original 24 VDP profiles are shown in Figure 1. As to some other work in profile monitoring, Woodall, Williams, and Birch (2004) provided a good introduction to profile monitoring and its applications. For other related works, see in Jensen et al. (2007), Williams et al. (2007), and the reference cited therein.

Shiau and Weng (2004) extended the linear profile monitoring schemes of Kang and Albin (2000) and Kim et al. (2003) to profiles of more general forms. Fixed effects models are considered in these papers. Kang and Albin (2000) monitored linear profiles with a fixed effects model by monitoring slopes and intercept jointly with a multivariate T^2 charts. However, under the fixed effects model, the effects of some uncontrollable factors, such as variations in humidity or temperature, are all categorized as part of the error term. In many cases, the collective effects of these factors may affect the values of the slope and intercept of the linear profiles. If so, then these hard-to-control factors would be more appropriate to be considered as common causes of variation. Shiau, Lin, and Chen (2006) proposed an approach to monitoring linear profiles with a random effects model to account for these common causes of variation. In this study, we consider a parametric random effects model in order to incorporate more variability that we often observe in many profile data. To understand the profile-to-profile variation, we focus on the covariance structure of profiles. To analyze the covariance structure, we consider the technique of principal components analysis (PCA).

Classical Hotelling T^2 statistic treats all the principal components equally important. However, for principal components with little contribution in explaining the variability of the profiles, it may not be desirable to signal alarms when the corresponding scores are significantly deviated from the in-control process. This is an example of effects being statistically significant but not practically significant. We propose a new monitoring scheme that is only sensitive to shifts in "important" components. This method gives a weight to each principal component according to how much of the total variation it accounts for. By putting more weights on "important" components, we can detect shifts in these corresponding directions more effectively. Also by neglecting those principal components that contribute so little to the total variation, we can avoid unnecessary interrupts signaled by the shiftings in the corresponding directions.

We apply the principal components analysis to data to obtain the eigenvectors and the associated principal component scores. When the covariance matrix is not of full rank, some eigenvalues are zero. When this happens, neglect the eigenvectors associated with them. This will happen when the number of profiles is smaller than the number of the set points. This problem also can be solved by using the singular value decomposition (SVD) to data matrix. See Ramsary and Silvermen (2005). Then we utilize these eigenvalues/eigenvectors to construct monitoring statistics for Phase II profile monitoring.

In this study, we compare three statistics, the classical T^2 statistic, maximum score statistic, and a total-squared-scores T^2 statistic. A simulation study is conducted to evaluate the performances of the three statistics. It is found from the simulation study that the total-squared-scores T^2 statistic has better power than the classical Hotellong T^2 statistic and maximum score statistic for shifts in the most important principal component and is very insensitive to the "unimportant" principal components.

The rest of the thesis is organized as follows. Section 2 reviews the technique of PCA, which is employed to analyze the covariance structure of the in-control profiles, and the concept of functional data analysis. Section 3 describes the three monitoring schemes studied in this paper. Section 4 presents the simulation results of Phase II profile monitoring with the VDP example. Finally, Section 5 concludes the paper with a brief summary and some remarks.

2 Review on Background Methodologies

In an increasing number of cases, the quality of a product or a process cannot be represented by a single quality variable. Rather, a set of measurements are taken across some continuum, such as time or space. In this section, we review the concepts of profile monitoring, nonparametric regression, principal components analysis, and functional data analysis.

2.1 Nonparametric Regression

Using linear models to monitor nonlinear profiles definitely is not suitable. By treating the discretized profile as a vector, monitoring nonlinear profiles can be considered as a particular application of multivariate process control problems. However, many traditional statistical methods fail when dealt with functional data, e.g., the strong correlations between the variables will cause an ill-conditioned problem in multivariate models. Hence, nonparametric statistics have been developed accordingly. Recently, more and more research works have focused on monitoring nonlinear profiles. A common approach to nonlinear profile monitoring is to fit a parametric regression model to each profile and make inferences based on the estimated parameters. When a parametric form of the profiles is overly complicated or hard to determine from data, the nonparametric regression approach may be more appropriate. This approach fits profiles via some smoothing methods, like spline smoothing, local polynomial smoothing, and wavelets. The following nonparametric regression model is considered:

$$y_j = m(x_j) + \varepsilon_j, j = 1, ..., p, \tag{1}$$

where y_j is the j-th observation of the profile, $m(x)$ is a smooth regression curve to be estimated and ε_i are independent and identically distributed (i.i.d.) normal variates with mean zero and common variance. Each profile in the VDP data set is smoothed by the function "smooth" of statistical package R. Figure 2 shows the 24 smoothed VDP profiles.

2.2 Functional Data Analysis

There is an increasing number of situations coming from different fields of applied sciences in which the collected data are curves. However, the progress of the computing tools, both in terms of memory and computational capacities, allows us to deal with large sets of data. And we can observe a very large set of variables. In some situation, the variable can be observed over a period of time $\mathcal{T} = [T_0, T]$, then the observation can be expressed as a family of random variables, denoted by $\{X(t), t \in \mathcal{T}\}$. Since the functional data consist of noises that may affect the performance of analysis, smoothing techniques can filter out the noise and then the features revealed by PCA can explain the data more clearly.

For our model, we consider the sample mean function $\bar{X}(t)$ and the sample covariance function $Q(s,t)$ as below:

$$\bar{X}(t) = n^{-1} \sum_{i=1}^{n} X_i(t), \tag{2}$$

$$Q(s,t) = (n-1)^{-1} \sum_{i=1}^{n} \{X_i(s) - \bar{X}(s)\}\{X_i(t) - \bar{X}(t)\}. \tag{3}$$

Then, we can use the principal components analysis for the functional data to analyze the profile.

2.3 Principal Components Analysis

The principal components analysis (PCA) is very useful in summarizing and interpreting profiles with the same equally-spaced values for independent variable. The PCA is a technique for choosing more important combinations of these independent variables. Castro et al. (1986) showed that the principal modes of variation contain eigenfunctions of the process covariance function. The authors also gave methods for estimating these eigenfunctions from a set of observed curves. Rice and Silverman (1991) provided a nonparametric method to estimate the mean function from a set of curves under the assumption that it was smooth. They also suggested a variant of the usual cross-validation to choose the degree of smoothing for data to be smoothed. In the estimation of the covariance structure, they primarily concerned about the first few eigenfunctions that were smoothed and their eigenvalues decayed rapidly. Jones and Rice (2002) suggested to use a simple PCA to identify important modes of variation among these curves. And they also used the principal component scores to identify particular curves that demonstrate the form and extent of a particular mode of variation.

Below, we review the PCA with materials mainly taken from Anderson (2003). The PCA is a multivariate procedure that rotates the data set such that the maximum variabilities are projected onto the new axes. Essentially, a set of correlated variables are transformed into a set of uncorrelated variables. Then, these uncorrelated variables are called principal components (PC). These principal components are linear combinations of the original variables. And the main concept is to reduce the dimension of the data set while retaining as much information as possible. Principal components contain special properties or features of the covariance matrix. They turn out to be the eigenvectors of the covariance matrix. Let the random vector $\mathbf{Y}$ of p components have the covariance matrix Σ and assume that the mean vector of the random variable $\mathbf{Y}$ is zero without loss of generality.

Theorem 1 ***(Anderson, 2003)***
Let $\boldsymbol{Y}$ be a p-component random vector with $E(\boldsymbol{Y}) = \mathbf{0}$ and $E(\boldsymbol{Y}\boldsymbol{Y}') = \boldsymbol{\Sigma}$. Then there exists an orthogonal transformation $\boldsymbol{B}$ such that

$$\boldsymbol{U} = \boldsymbol{B}'\boldsymbol{Y} \tag{4}$$

has the covariance matrix

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \lambda_{p-1} & 0 \\ 0 & 0 & \dots & 0 & 0 & \lambda_p \end{pmatrix}, \tag{5}$$

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

Thus, $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_p$ are eigenvectors and $\lambda_1, \dots, \lambda_p$ are eigenvalues of the covariance matrix $\boldsymbol{\Sigma}$. In PCA, $\boldsymbol{\beta}_r$ is the r-th principal component of $\boldsymbol{Y}$ and U_r is called the score of the r-th principal component. These scores will catch the features of curves in the directions of the corresponding principal components. In our monitoring scheme, we aim at detecting changes based on these principal component scores.

When analyzing data, we apply PCA to the sample covariance matrix $\boldsymbol{S}$. Suppose we have n profiles and each profile contains p measurements. Compute the sample covariance matrix $\boldsymbol{S}$ by

$$\boldsymbol{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})', \tag{6}$$

where $\mathbf{y}_i$ the i-th profile and $\bar{\mathbf{y}}$ is the mean vector of these profiles, and $\bar{\mathbf{y}} = \sum_{i=1}^{n}\mathbf{y}_i/n$. Apply the eigenanalysis to the sample covariance $\boldsymbol{S}$. The eigenvector corresponding to the j-th largest eigenvalue is the j-th principal component, $j = 1, \ldots, p$. Then project each profile onto the eigenvectors to get the PC scores.

2.4 Functional PCA vs. PCA

In this study, we consider the nonparametric model of Karhunen-Loève expansion to model profile data. Let $\{X(t), t \in \mathcal{T}\}$ be a realization of a stochastic process with mean $\mu(t)$ and covariance function $G(s,t)$, where

$$Cov(X(s), X(t)) = G(s,t) = \sum_{k=1}^{\infty}\rho_k\phi_k(s)\phi_k(t)$$

and $\{\phi_1(t), \phi_2(t), \ldots\}$ is an orthognormal sequence of continuous eigenfunctions in L_2 and eigenvalues $\rho_1 \geq \rho_2 \geq \ldots \geq 0$ satisfy

$$\int_{\mathcal{T}} G(s,t)\phi_k(t)dt = \rho_k\phi_k(s), \quad k = 1, 2, \ldots \tag{7}$$

Equation (8) is the functional eigenequation. Then $X(t)$, $t \in \mathcal{T}$, has a (quadratic mean) representation:

$$X(t) = \mu(t) + \sum_{k=1}^{\infty} X_k\phi_k(t), \tag{8}$$

where $X_k = \int (X(t) - \mu(t))\phi_k(t)dt$. It is well known that X_k follows a normal distribution with mean zero and variance ρ_k and $\{X_k\}$ are linearly independent. In equation (9), the common modes of variation of a curve are represented by the Karhunen-Loève expansion in L_2.

In practice, we observe the curve data at some discrete points as in the following:

$$Y_j = X(t_j) + \varepsilon_j,\ j = 1, \ldots, p. \tag{9}$$

We apply Tukey's smooth smoothing techique code as a function in R to filter out the noise. After filtering out the noise ε_j, the actual signals of curves can be better extracted from the data and these signals can explain the variation of profiles more clearly.

In this study, we adopt the functional data analysis approach to model profiles data but computation is carried out for discretized profiles. Therefore, we need to understand the relationship between the eigenfunctions and eigenvectors.

Let $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_m$ be m discrete profiles data measured at the same p equally-spaced points $t_1, \ldots, t_p$. The standard multivariate principal components analysis produces eigenvalues λ and eigenvectors $\boldsymbol{u}$ satisfying

$$\boldsymbol{V}\boldsymbol{u} = \lambda \boldsymbol{u}, \tag{10}$$

where $\boldsymbol{V}$ is the sample covariance matrix of $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_m$. Note that the sample covariance matrix $\boldsymbol{V}$ has the elements $V(t_j, t_k)$ where $V(\cdot, \cdot)$ is the sample covariance function. Let ξ be any function and $\boldsymbol{\xi} = (\xi(t_1), \ldots, \xi(t_p))'$. Let $w = |\mathcal{T}|/n$, where $|\mathcal{T}|$ is the length of the interval $\mathcal{T}$. Since

$$\int V(t_j, t)\xi(t)dt \approx w \sum_{k=1}^{p} V(t_j, t_k)\xi(t_k),$$

the eigenfunction (8) can be approximated by

$$w\boldsymbol{V}\boldsymbol{\xi} = \rho \boldsymbol{\xi}. \tag{11}$$

The solution of (12) will correspond to those of the ordinary discrete eigenequation

$$\boldsymbol{V}\boldsymbol{u} = \lambda \boldsymbol{u}$$

with $\rho = w\lambda$ and $\boldsymbol{\xi} = w^{-1/2}\boldsymbol{u}$ if $w\|\boldsymbol{\xi}\| = \|\boldsymbol{u}\| = 1$.

Return to our case. Denote $\boldsymbol{\phi}_k = (\phi_k(t_1), \ldots, \phi_k(t_p))'$. Then, for $k = 1, \ldots, p$, we have $\rho_k = w\lambda_k$ and $\boldsymbol{\phi}_k = w^{-1/2}\boldsymbol{\beta}_k$, where λ_k and $\boldsymbol{\beta}_k = (\beta_{k1}, \beta_{k2}, \ldots, \beta_{kp})'$ are respectively the

k-th eigenvalues and eigenvectors of the sample covariance matrix. Let X_k and A_k be the scores of the functional PCA and PCA respectively. Then the score in functional PCA

$$\begin{aligned} X_k &= \int (X(t)-\mu(t))\phi_k(t)dt \approx w\sum_{j=1}^{p}(X(t_j)-\bar{X}(t_j))\phi_k(t_j) \\ &= w\sum_{j=1}^{p}(X(t_j)-\bar{X}(t_j))w^{-1/2}\beta_{kj} = w^{1/2}A_k \end{aligned}$$

Accordingly, we have $X_k^2/\rho_{kk}^2/\lambda_k$ and $\sum_{k=1}^{p} X_k\boldsymbol{\phi}_k \approx \sum_{k=1}^{p} A_k\boldsymbol{\beta}_k$. This means that we can approximate the functional PCA by multivariate PCA.

3 Monitoring Schemes

3.1 Monitoring Statistics

In this study, we focus on Phase II monitoring. In Phase II study, it is usually assumed that the in-control process is already characterized. Then, without loss of generality, assume the discrete in-control profiles are i.i.d. as $N_p(\mathbf{0},\boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}$ is known. It is well known that, when the process is in control, the score A_k is distributed as $N(0,\lambda_k)$. For example, see Anderson (2003). Consider the case of shifts in mean. Let δ be the shift size as a multiple of standard deviation. In this thesis, for Phase II monitoring, we study three types of control charts based on the PC scores as described below. Denote the number of nonzero eigenvalues by n'. The first chart is the usual T^2 chart, the the monitoring statistic:

$$T_1 = \sum_{k=1}^{n'} \lambda_k^{-1} A_k^2. \tag{12}$$

Since A_k is distributed as $N(0,\lambda_k)$, $A_k/\sqrt{\lambda_k}$ follows the standard normal distribution. Then T_1 is distributed as χ^2 with n' degrees of freedom. Thus, claim the process as out of control if

$$T_1 > \chi^2_{1-\alpha,n'}, \tag{13}$$

where $\chi^2_{1-\alpha,n'}$ denotes the $100(1-\alpha)$ percentile of the χ^2 distribution with n' degrees of freedom. By (13), it is clear that T_1 statistic treats each principal component (score) equally important.

Note that PC scores can be monitored individually since they are independent. But it would be too troublesome to monitor several charts at the same time. Thus we consider a combined chart scheme that combines all n' individual PC score charts. More specifically, a combined chart signals when any of the individual charts signals. This scheme is equivalent to monitor the maximum of the individual monitoring statistics. Then the monitoring statistic is as below:

$$T_2 = \max_{1 \leq k \leq n'} |\lambda_k^{-1/2} A_k|. \tag{14}$$

Thus, signal is out of control if

$$T_2 > Z_{\alpha'/2}, \tag{15}$$

where $\alpha' = 1 - (1-\alpha)^{1/n'}$. Then it can be easily shown that the false alarm rate ofthe combined chart is α.

The first two approaches are often used in other contexts. The third approach is new and proposed practially for profile monitoring. The idea is simple. We wish to find a monitoring statistic that are sensitive to the shifts in “important” directions and insensitive to the shifts in “unimportant” directions. For simplicity, consider the following monitoring statistics:

$$T_3 = \sum_{k=1}^{n'} A_k^2. \tag{16}$$

Note that T_3 treats each of principal components differently with a weight according to how much it contributes to the variation of the profiles instead of standardizing each principal component score to standard normal. Satterwaite (1941) proposed an approximation to linear combinations of independent χ_1^2 distributions as a scaled χ^2 distribution, $c\chi_{df}^2$, where

$$c = \frac{\sum_{k=1}^{n'} \lambda_k^2}{\sum_{k=1}^{n'} \lambda_k} \text{ and } df = \frac{[\sum_{k=1}^{n'} \lambda_k]^2}{\sum_{k=1}^{n'} \lambda_k^2}. \tag{17}$$

Thus, the process is claimed out of control if

$$T_3 > c\chi_{df,1-\alpha}^2. \tag{18}$$

To see whether the control limit given in (19) is adequate or not, as an example, we generate 1,000,000 profiles from the eigenvectors/eigenvalues of the smoothed VDP data to approximate the distribution of T_3. The density of T_3 estimated by kernel density estimation is shown in Figure 3. Since the approximate distribution of T_3 is not close enough to $c\chi_{df}^2$, we proposed using empirical $(1-\alpha)$th quantitle of T_3 as the control limit of this chart. We refer to this chart as the "Total-Squared-Scores" chart hereafter.

3.2 Control Limits

In studying the performances of the monitoring schemes, researchers usually assume $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known in Phase II monitoring. In this study, we evaluate the performances of the proposed monitoring schemes in terms of the average run length (ARL). Several versions of control limits are considered as below.

1. T^2 Chart:

   For T^2 statistic, we consider another version and refer to the original T_1 statistic in (12) as

   $$T_{11} = \sum_{k=1}^{n'} \lambda_k^{-1} A_k^2. \tag{19}$$

   Since T_{11} has poor power when n' is large, we consider a T^2 statistic by adding only the tanderized scores of effective components as

   $$T_{12} = \sum_{k=1}^{K} \lambda_k^{-1} A_k^2, \tag{20}$$

where K is the number of the "effective" principal components. It is obvious that T_{11} and T_{12} have chi-squre distributions with degrees of freedom n' and K, respectively. Then the control limits for T_{11} and T_{12} are $\chi^2_{n',1-\alpha}$ and $\chi^2_{K,1-\alpha}$, respectively.

2. Combined Chart:

   The control limit of T_2 is $Z_{\alpha'/2}$, where $\alpha' = 1 - (1-\alpha)^{1/n'}$.

3. Total Squared Score (TSS) Chart:

   Since Figure 3 shows that the distribution of T_3 is not close to $c\chi^2_{df}$, we study two versions of control limits: one is the empirical $(1-\alpha)$ quantitle and the other is $c\chi^2_{df,1-\alpha}$. For the empirical quantitle, we take the eigenvalues $\{\lambda_k\}_{k=1}^{n'}$ of the smooth VDP data as the true eigenvalues/eigenvectors and generate 1,000,000 set of $\{A_k\}_{k=1}^{n'}$ with $A_k \sim N(0, \lambda_k)$, to obtain 1,000,000 T_3, and hence the empirical $(1-\alpha)$ quantitle of T_3. We refer to the T^2 control chart with this control limit as the T_{31} chart in simulation studies. For our simulation study, UCL_{31} is 8262.107. We refer to T^2 chart with the control limit $c\chi^2_{df,1-\alpha}$ as the "T_{32}", where c and df are estimated by eigenvalues and eigenvectors of the smoothed VDP.

On the other hand, if the parameters are unknown, we need to use some in-control historical data to estimate these parameters. Jensen et al. (2006b) mentioned that the effect of parameter estimation on control chart properties should not be ignored. Thus we consider the case that eigenvalues/eigenvectors are estimated from in-control profiles to investigate the effect of the estimation error. For this cases, the control limits are obtained as below. The T^2 and the Combined Chart have exact control limits as given above. Fore versions of control limits for T_3 are considered. The control limit for T_{31} and T_{32} are as the same as above. Assume that we have m in-control profiles available. Apply PCA to them to obtain eigenvalues and eigenvectors, generate 1,000,000 set of scores as before with these estimated eigenvalues to obtain 1,000,000 $T_3's$. Let the control limit be the empirical $(1-\alpha)$th quantitle of these values of T_3.

Since these estimated parameters depends on this particular set of m profiles, the performance of the control chart may not be the average case. Thus we will repeat the study, say b times to coverage sampling bias. The last version of the TSS Chart is referred to as "T_{34}" chart, in which the control limit is $c\chi^2_{df,1-\alpha}$ with c and f as in (17) but using the setimated eigenvalues for $\{\lambda_k\}_{k=1}^{n'}$.

4 Simulation Studies

4.1 Data Generation

In our simulation studies, we simulate data based on the vertical board density profiles. The vertical board density profiles from Walker and Wright (2002) consist of 24 profiles of vertical density, each profile has 314 measurements. These data set is available at http://bus.utk.edu/stat/walker/VDP/Allstack.TXT. First, we smooth each profile and then apply SVD to the sample covariance matrix of VDP data to get the "original" eigenvalues and eigenvectors, $\{(\lambda_k, \boldsymbol{\beta}_k)\}$, and treat them as the population version. Note that monitoring the functional data can be reduced to monitoring the discrete profile data as shown in Section 2.4. Thus, we only need to generate discrete profile data. Table 1 gives 23 eigenvalues of the smoothed VDP data. To generate a new profile as a data vector, we use:

$$\boldsymbol{Y} = \boldsymbol{\mu} + \sum_{k=1}^{23} A_k \boldsymbol{\beta}_k, \tag{21}$$

where $A_k \sim N(0, \lambda_k)$, $\boldsymbol{\mu}$ is the mean vector of the smoothed VDP data and is assumed known. A_k is the k-th PC score of a profile. When the process is in control, the score A_k is distributed as $N(0, \lambda_k)$. For an out-of-control process, we let the score A_k be distributed as $N(\delta\sqrt{\lambda_k}, \lambda_k)$. In our study, δ ranges from 1 to 12.

In our simulation study, two methods are considered in Phase II monitoring. One uses the "true" eigenvalues and eigenvectors of the smoothed VDP, the other uses the estimated eigenvalues and eigenvectors from m profiles. And each of the m profiles in our study is of the form (21).

For simulation study with "true" eigenvalues and eigenvectors of the smoothed VDP data, we only need to generate $\{A_k\}_{k=1}^{n'}$ 1,000,000 times for each shift in PC1, PC2, and PC23. For simulation study with estimated eigenvalues and eigenvectors, we first generate m profiles by equation (21). Apply PCA to these m profiles to obtain estimated eigenvalues $\hat{\lambda}_1, \ldots, \hat{\lambda}_{23}$ and eigenvectors $\hat{\boldsymbol{\beta}}_1, \ldots, \hat{\boldsymbol{\beta}}_{23}$. Generate 1,000,000 profiles of (21). Since the mean vector $\boldsymbol{\mu}$ of the process is known in Phase II monitoring, without loss of generality, we assume $\boldsymbol{\mu}=\mathbf{0}$. Project profiles onto these estimated eigenvectors to get their PC scores. We use these principal component scores to evaluate the performances of the three charts constructed with the statistics T_1, T_2, and T_3 in Phase II monitoring in terms of the detecting powers. Repeat the above procedure 1,000 times to average the detecting power so that the simulation results can represent an average case, not biased by the particular m profiles used for constructing the control limits. In our study, the false alarm rate α is set at 0.0027 and m=200, 300, 400, 500, 600.

4.2 Number of Principal Components

In this paper, we analyze the VDP data set and its eigenvalues are shown in Table 1. This table shows that the first four principal components account for 0.8451, 0.1076, 0.0192, and 0.0088 of the total variation in the profiles, respectively. The total is 0.9807. Thus, for T_{12} statistic, K is set at 4 for the monitoring scheme.

4.3 Simulation Results

In this studies, the eigenvalues/eigenvectors of the smoothed VDP data are treated as the true population parameters. The performance of the control charts are estimated in terms of ARLs. Table 2-4 presents the ARL values of the charts under study on PC1, PC2, and PC23,respectively. Figure 4-6 display the corresponding detecting power of charts. The results are summarized as below:

- Since the T^2 charts and the combined chart treats each principal component equally important, there is no difference in performance among all principal components for them.
- The T^2 chart using only the effective principal components, T_{12}, performs better than the T^2 chart using all principal components, T_{11}, for all PC scores.
- The combined chart performs in between the two T^2 charts, but fairly close to the T^2 chart with only four components for the first four PCs.
- The Satterwaite's approximation indeed is not good enough for process monitoring, which leads to an unsuitable in-control ARL. Hence, this control limit is not suggested.
- The TSS chart using the empirical $(1-\alpha)$ quantile as the control limit perform better than the T^2 charts and the combined chart for PC1.
- The detecting power of the TSS chart drops very quickly after the first component. This confirms the expectation that statistic T_3 is insensitive to "unimportant" PCs. But, the problem is that the TSS chart has very little detecting power for changes in other principal components.

Tables 5-19 show the ARL values when eigenvalues and eigenvectors need to be estimated from historical m profiles. The corresponding power curves are given in Figures 7-21. The results of this simulation study are summarized as below:

- For m not large enough, the ARL_0 of all the T^2 charts are disired This indicates the estimation error does play an important role in the performance of the charts. We would need a fairly large historical data set to get the estimation accurate enough.

- For the T^2 charts, the one using only effective components has ARL_0 values closer to 370.4, then the one using all components. The former also has a better detecting power than the latter.
- Again, the performance of the combined chart falls in between the two T^2 chart.
- As to the TSS chart, using sample eigenvalues to construct the control limit (T_{33}) has ARL_0 value than the one using the "true" control limit (T_{31}). In fact, they are not too far from the nominal value of 370.4. This indicates that this chart (T_{33}) may be useful in practice.

5 Conclusion

In recent years, the profile monitoring has become a popular area of research in statistical process control. In this study, we discuss profile monitoring schemes with nonparametric regression models. We use the principal components analysis to analyze the covariance structure of the profiles and use the principal component scores that capture special features of profiles for process monitoring.

In the thesis, we compare three statistics. The first statistic is the usual T^2. T^2 treats each principal component equally important but has poor power with an increasing the number of principal components. Thus, we only select the first K principal components according the proportion of the variation they explain. The second statistic is the maximum score, which corresponds to the combined chart that combines all the PC score charts. It monitors each component of a profile and also treats each principal component equally important. The third statistics is the TSS. It is sensitive to "important" principal components and insensitive to the "unimportant" principal components.

Using the VDP data as an illustrative example, our simulation shows that TSS performs better than T^2 and the combined chart for shifts in the most "important" component.

However, the effects of shifts on those components with little contribution to the overall profile may be statistically significant but in fact have no practical significance in quality. Wasting some power on “unimportant” components, classical T^2 statistics have less power in detecting changes in the “important” component. On the other hand, the statistic we propose giving more weights to more “important”principal components scores is sensitive to the change of the “important” principal components. Then, when shifts occur in the “important” principal component, the statistic we propose can detect more quickly than classical T^2 statistics. It is very insensitive to those “unimportant” components since almost no weights are given to them.

However, our simulation shows TSS only performs well in PC1, the power of TSS decreases very quickly after PC2. It has almost no power for PC3 and on. This is because the eigenvalues of the VDP data drop so quickly from. λ_1=0.8451 to λ_2=0.1076, then to λ_3=0.0192, and so on. It seems the weights need to be adjusted if the process changes are bound to be captured by some principal components other than PC1. This could be a potential future research topic.

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 Sylvester, E. A. (1986). "Principal Models of Variation for Processes With Continuius Sample Curves". *Technormetrics* 28, pp. 329-337.

[3] Jensen, W. A., Birch, J. B., and Woodall W. H. (2006a). "Profile Monitoring via Nonlinear Mixed Models", Techniquical Report. Virginia Polytechnic Institute and State University.

[4] Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., and Woodall W. H. (2006b). "Effects of Parameter Estimation on Control Chart Properties: A Literature Review",*Journal of Quality Technology* 38, pp. 349-366.

[5] 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.

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

[7] Mestek, O., Pavlik, J., and Suchanek, M. (1994). "Multivariate Control Charts: Control Charts for Calibration Curves". *Fresenius Journal of Analytical Chemistry* 350, pp. 344-351.

[8] 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.

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

[10] Ramsay, J. O. and Silverman, B. W. (2005) *Functinal Data Analysis,* 2nd. Springer, New York.

[11] Satterthwaite, F. W. (1941) "Synthesis of Variance", *Psychometrik*, 6, 309-316.

[12] Shiau, J.-J. H. and Weng, Z.-P (2004) "Profile Monitoring by Nonparametric Regression". Techniqual Report. Institute of Statistics, National Chiao Tung University.

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

[14] Williams, J. D., Woodall, W. H., and Birch, J. B. (2003), "Phase I Analysis of Nonlinear Product and Process Quality Profiles", Technical Report. Virginia Polytechnic Institute and State University.

[15] 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.

Table 1: Eigenvalues for smoothed VDP data (proportion of explaining variation)

k	λ_k	k	λ_k	k	λ_k
1	899.1265 (0.8451)	9	1.6551 (0.0015)	17	0.5106 (0.0005)
2	114.4609 (0.1075)	10	1.3654 (0.0012)	18	0.3880 (0.0003)
3	20.4119 (0.0191)	11	1.2072 (0.0011)	19	0.3788 (0.0003)
4	9.4104 (0.0088)	12	1.0337 (0.0009)	20	0.2879 (0.0002)
5	3.2126 (0.0030)	13	0.8993 (0.0008)	21	0.2800 (0.0002)
6	2.7177 (0.0025)	14	0.6758 (0.0006)	22	0.2529 (0.0002)
7	2.3202 (0.0021)	15	0.6040 (0.0005)	23	0.1787 (0.0001)
8	1.9736 (0.0018)	16	0.5144 (0.0004)		

Table 2: ARL for PC1

	T_{11}	T_{12}	T_2	T_{31}	T_{32}
0	370.3703	370.3703	370.3703	370.3703	287.3563
1	217.3913	103.7344	221.2389	45.2488	36.2318
2	60.9756	15.3280	28.4414	6.3259	5.6357
3	14.6842	3.5868	4.9677	2.0023	1.8912
4	4.3821	1.5755	1.7818	1.1791	1.1565
5	1.9051	1.1044	1.1418	1.0222	1.0187
6	1.2214	1.0111	1.0155	1.0014	1.0011
7	1.0384	1.0005	1.0000	1.0000	1.0000
8	1.0037	1.0000	1.0000	1.0000	1.0000
9	1.0002	1.0000	1.0000	1.0000	1.0000
10	1.0001	1.0000	1.0000	1.0000	1.0000
11	1.0000	1.0000	1.0000	1.0000	1.0000
12	1.0000	1.0000	1.0000	1.0000	1.0000

Table 3: ARL for PC2

	T_{11}	T_{12}	T_2	T_{31}	T_{32}
0	370.3703	370.3703	370.3703	370.3704	287.3563
1	217.3913	103.7344	221.2389	294.1176	242.7184
2	60.9756	15.3280	28.4414	303.0303	238.0952
3	14.6842	3.5860	4.9677	194.5525	143.6781
4	4.3821	1.5755	1.7818	108.4598	80.7754
5	1.9051	1.1044	1.1418	44.6827	33.9673
6	1.2214	1.0111	1.0155	15.8177	12.0569
7	1.0384	1.0005	1.0008	5.25762	4.12269
8	1.0037	1.0002	1.0000	2.0786	1.7925
9	1.0002	1.0000	1.0000	1.2426	1.1713
10	1.0001	1.0000	1.0000	1.0372	1.0227
11	1.0000	1.0000	1.0000	1.0028	1.0014
12	1.0000	1.0000	1.0000	1.0000	1.0000

Table 4: ARL for PC23

	T_{11}	T_2	T_{31}	T_{32}
0	370.3703	370.3703	333.3333	260.4166
1	217.3913	221.2389	384.6153	267.3796
2	60.9756	28.4414	381.6793	297.6190
3	14.6842	4.9677	335.5704	268.8172
4	4.3821	1.7818	406.5040	292.3976
5	1.9051	1.1418	406.5040	335.5704
6	1.2214	1.0155	387.5968	292.3976
7	1.0384	1.0008	359.7122	260.4166
8	1.0037	1.0000	359.7122	268.8172
9	1.0002	1.0000	373.1343	277.7777
10	1.0001	1.0000	387.5968	284.0909
11	1.0000	1.0000	359.7122	287.3563
12	1.0000	1.0000	347.3152	288.14523

Table 5: ARL for PC1(m=200)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	155.3097	236.9668	112.4590	214.5922	263.1578	367.6470	256.4102
1	34.4827	58.7544	74.9625	24.5700	36.6300	35.5114	28.1056
2	13.8927	9.96618	16.6444	4.3335	5.7730	5.4042	4.6576
3	5.2675	2.8088	3.7147	1.6500	1.9043	1.8492	1.7088
4	2.3243	1.4020	1.5549	1.1160	1.1671	1.1584	1.1281
5	1.3730	1.0685	1.0948	1.01268	1.0209	1.0189	1.0144
6	1.0842	1.0060	1.00888	1.0006	1.0011	1.0004	1.0006
7	1.0116	1.0002	1.0003	1.0000	1.0000	1.0002	1.0002
8	1.0008	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	1.0002	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 6: ARL for PC2(m=200)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	154.9451	235.8491	114.4165	201.6129	274.7253	337.8378	228.3105
1	40.2901	81.1688	88.0282	188.6792	265.9574	335.5705	226.2443
2	16.8237	13.5355	24.5459	153.8462	234.7418	264.5503	182.4818
3	6.0518	3.4729	5.1020	102.8807	149.2537	177.9359	125.6281
4	2.5431	1.5520	1.8230	58.6166	84.4595	104.6025	71.2251
5	1.4504	1.1000	1.1514	26.8528	37.3692	47.1698	32.4675
6	1.1048	1.0107	1.0163	9.5475	14.6071	17.0823	11.4784
7	1.0163	1.0006	1.0008	3.3697	4.8866	5.6218	3.9364
8	1.0014	1.0000	1.0000	1.6070	2.0075	2.1865	1.7538
9	1.0001	1.0000	1.0000	1.1237	1.2262	1.2752	1.1630
10	1.0000	1.0000	1.0000	1.0156	1.0345	1.0446	1.0224
11	1.0000	1.0000	1.0000	1.0009	1.0026	1.0036	1.0014
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0001	1.0001

Table 7: ARL for PC23(m=200)

	T_{11}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	154.1712	236.8224	116.0784	276.4167	364.8780	228.3105
1	39.2157	63.2111	200.0000	294.1176	352.1127	241.5459
2	18.7758	18.1159	196.8504	287.3563	320.5128	227.2727
3	7.3206	4.9796	204.9180	294.1176	352.1127	256.4103
4	3.1504	2.0962	204.0816	306.7485	316.4557	242.7184
5	1.6904	1.2916	206.6116	285.7143	375.9398	247.5248
6	1.1905	1.0657	210.0840	280.8989	384.6154	255.1020
7	1.0371	1.0094	200.0000	304.8780	328.9474	234.7418
8	1.0041	1.0006	195.3125	289.0173	344.8276	227.2727
9	1.0003	1.0000	217.3913	263.1579	333.3333	251.2563
10	1.0000	1.0000	207.2487	251.2884	324.1854	228.2159
11	1.0000	1.0000	219.1488	285.1587	334.7575	258.6783
12	1.0000	1.0000	211.2579	284.8423	315.8766	251.6453

Table 8: ARL for PC1(m=300)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	220.7729	403.2258	269.4915	359.7122	284.0909	375.9398	427.3504
1	81.6993	108.6957	132.2751	41.8410	39.9361	43.8982	47.8927
2	31.8066	17.1350	32.1750	6.2555	6.1013	6.4591	6.8213
3	10.0746	3.9448	5.5371	1.9874	1.9676	2.0188	2.0761
4	3.5865	1.6472	1.8963	1.1849	1.1807	1.1916	1.2044
5	1.7561	1.1258	1.1716	1.0232	1.0222	1.0242	1.0264
6	1.1930	1.0145	1.0195	1.0013	1.0011	1.0014	1.0015
7	1.0340	1.0007	1.0012	1.0000	1.0000	1.0000	1.0000
8	1.0035	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	1.0002	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 9: ARL for PC2(m=300)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	222.2494	416.6667	260.7717	362.3188	280.8989	400.0000	462.9630
1	82.3723	124.6883	149.2537	340.1361	242.7184	359.7122	406.5041
2	32.9598	20.2429	38.9408	248.7562	196.8504	270.2703	312.5000
3	10.7735	4.5106	6.7168	171.8213	141.6431	185.8736	205.7613
4	3.9711	1.7859	2.1251	92.5926	70.9220	99.6016	111.6071
5	1.8934	1.1570	1.2178	39.7772	31.0559	42.8816	48.9716
6	1.2369	1.0200	1.0279	14.1643	11.2587	15.3186	17.3792
7	1.0459	1.0009	1.0016	4.6155	3.8652	4.9417	5.5636
8	1.0052	1.0000	1.0000	1.9427	1.7212	2.0232	2.1824
9	1.0002	1.0000	1.0000	1.2076	1.1544	1.2294	1.2699
10	1.0000	1.0000	1.0000	1.0310	1.0200	1.0353	1.0438
11	1.0000	1.0000	1.0000	1.0018	1.0012	1.0023	1.0030
12	1.0000	1.0000	1.0000	1.0001	1.0000	1.0001	1.0001

Table 10: ARL for PC23(m=300)

	T_{11}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	225.3132	286.5671	381.2048	276.2430	402.5806	389.7122
1	82.3723	83.6120	312.5488	297.6190	328.9473	373.1343
2	24.5218	15.0966	324.6753	280.8984	337.8378	373.1343
3	8.2877	3.9491	342.4657	322.5806	359.7122	403.2258
4	2.9751	1.6767	349.6503	299.4011	373.1343	423.7288
5	1.5691	1.1436	328.9473	289.0173	354.6099	403.2255
6	1.1336	1.0203	364.9635	261.7801	375.9398	409.8360
7	1.0209	1.0015	375.9398	303.0303	393.7007	442.4778
8	1.0017	1.0001	316.4556	279.3296	340.1360	400.5156
9	1.0000	1.0000	328.4822	289.4891	348.5842	431.0344
10	1.0000	1.0000	357.8843	236.8997	316.1584	403.5941
11	1.0000	1.0000	318.2562	289.4879	369.4988	389.2971
12	1.0000	1.0000	316.158	325.5498	365.2594	408.4997

Table 11: ARL for PC1(m=400)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	222.2494	324.2152	374.2160	340.4494	274.7253	352.1127	374.8252
1	82.3723	53.5906	92.4214	20.4666	23.3333	33.2882	41.6967
2	32.9598	8.9977	14.6413	3.8772	4.2676	5.9552	5.2827
3	10.7735	2.5788	3.2757	1.5708	1.8194	1.9418	1.6448
4	3.9711	1.3339	1.4532	1.0989	1.1498	1.1763	1.1137
5	1.8934	1.0537	1.0739	1.0102	1.0172	1.0218	1.0122
6	1.2369	1.0080	1.0111	1.0006	1.0009	1.0012	1.0007
7	1.0459	1.0002	1.0005	1.0000	1.0001	1.0000	1.0000
8	1.0052	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	1.0002	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	1.0002	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	1.0002	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	1.0002	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 12: ARL for PC2(m=400)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	250.1923	330.4147	390.8397	336.6120	290.6977	370.3704	368.3502
1	71.3267	71.7360	125.9446	119.9041	145.7143	285.4713	318.3488
2	27.2777	13.2240	25.3421	96.8992	190.1141	122.7391	271.2494
3	8.7199	3.4151	4.7985	65.3595	81.3333	133.1502	183.9672
4	3.1447	1.5249	1.7506	34.6741	43.5689	95.4198	72.9367
5	1.6138	1.1008	1.1387	15.3374	19.2110	31.4684	46.9442
6	1.2538	1.0184	1.0260	8.6790	11.1383	10.4745	16.5441
7	1.0505	1.0013	1.0019	3.1393	3.7222	3.2345	5.6668
8	1.0055	1.0000	1.0000	1.5305	1.6959	2.0870	1.6758
9	1.0002	1.0000	1.0000	1.1043	1.1453	1.2454	1.1384
10	1.0000	1.0000	1.0000	1.0116	1.0191	1.0395	1.0182
11	1.0000	1.0000	1.0000	1.0005	1.0015	1.0027	1.0009
12	1.0000	1.0000	1.0000	1.0000	1.0001	1.0001	1.0000

Table 13: ARL for PC23(m=400)

	T_{11}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	354.4165	386.5672	334.0483	282.4859	337.8378	367.2241
1	76.6871	83.6120	138.8889	271.7391	438.5965	372.4138
2	28.3607	15.0966	133.6898	260.4167	384.6154	360.7717
3	8.8778	3.9491	149.7006	247.5248	357.1429	284.5018
4	3.1702	1.6767	142.8571	299.4012	362.3188	280.5054
5	1.6202	1.1436	153.3742	271.7391	409.8361	289.3939
6	1.3185	1.0203	217.3913	322.5806	359.7122	255.1020
7	1.0639	1.0016	239.2344	310.5590	446.4286	271.7391
8	1.0081	1.0001	211.8644	301.2048	364.9635	251.2563
9	1.0005	1.0000	205.7613	271.7391	340.1361	250.4840
10	1.0000	1.0000	283.4983	284.4985	357.4894	275.1859
11	1.0000	1.0000	225.9714	274.4545	344.4891	269.1878
12	1.0000	1.0000	205.4187	268.4891	354.4894	288.8797

Table 14: ARL for PC1(m=500)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	287.2659	384.6154	338.0952	326.2443	292.3977	378.7879	367.3797
1	98.8142	80.6452	135.5014	27.6091	31.1865	35.6174	40.0559
2	33.1345	12.2070	20.7383	4.6777	5.3781	5.0411	5.9378
3	8.6957	3.1180	4.0806	1.7151	1.8186	1.9446	1.7826
4	3.0460	1.4529	1.6074	1.1272	1.1544	1.1765	1.1413
5	1.5534	1.0798	1.1037	1.0132	1.0183	1.0210	1.0154
6	1.1574	1.0103	1.0142	1.0010	1.0008	1.0015	1.0012
7	1.0239	1.0004	1.0008	1.0000	1.0000	1.0000	1.0000
8	1.0021	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 15: ARL for PC2(m=500)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	280.5054	423.7288	351.2563	309.2050	292.3977	354.6099	355.1020
1	126.5823	123.4568	188.6792	178.5714	213.7801	316.4557	261.6752
2	45.3309	19.3199	40.3551	154.3210	208.3333	292.1176	294.3077
3	13.0993	4.4045	6.4994	96.3391	149.2537	112.0104	173.8668
4	4.2878	1.7434	2.0464	55.9910	74.9625	68.4036	99.3060
5	1.9722	1.1476	1.2018	25.9875	31.0907	33.9484	46.6656
6	1.1535	1.0091	1.0139	8.7935	11.9646	11.4847	15.3973
7	1.0232	1.0003	1.0004	3.1620	4.0254	4.9039	3.8601
8	1.0020	1.0000	1.0000	1.5452	1.7687	2.0227	1.7422
9	1.0001	1.0000	1.0000	1.1108	1.1689	1.2353	1.1599
10	1.0000	1.0000	1.0000	1.0124	1.0228	1.0350	1.0204
11	1.0000	1.0000	1.0000	1.0005	1.0018	1.0029	1.0014
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0001	1.0000

Table 16: ARL for PC23(m=500)

	T_{11}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	297.6285	341.5459	392.3077	259.0674	326.7974	336.9668
1	129.5337	176.0563	210.9705	270.2703	390.6250	248.7562
2	49.9002	37.9939	207.4689	318.4713	326.7974	243.9024
3	14.4051	7.1388	219.2982	271.7391	393.7008	257.7320
4	4.9544	2.4116	199.2032	294.1176	326.7974	235.8491
5	2.1429	1.3365	206.6116	290.6977	318.4713	238.0952
6	1.2341	1.0275	250.0000	271.7391	393.7008	310.5590
7	1.0431	1.0017	210.9705	337.8378	342.4658	268.8172
8	1.0046	1.0000	229.3578	308.6420	431.0345	312.5000
9	1.0002	1.0000	235.4984	277.7778	364.9635	282.4859
10	1.0000	1.0000	208.4894	278.4874	372.4566	298.4894
11	1.0000	1.0000	248.8772	236.4898	369.4894	267.4897
12	1.0000	1.0000	225.5152	258.4898	356.4897	258.4897

Table 17: ARL for PC1(m=600)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	377.3585	386.1004	338.0952	327.2727	363.1579	367.6471	304.8780
1	221.2389	93.4579	156.7398	32.0102	33.6022	39.5236	44.0930
2	32.2165	13.5722	23.9006	5.2214	5.3214	5.4078	6.9538
3	8.9381	3.3174	4.5082	1.7981	1.8310	1.9998	1.9180
4	3.1534	1.5094	1.6862	1.1475	1.1553	1.1911	1.1749
5	1.5932	1.0946	1.1261	1.0168	1.0166	1.0239	1.0208
6	1.1384	1.0096	1.0128	1.0008	1.0009	1.0013	1.0011
7	1.0203	1.0005	1.0007	1.0000	1.0000	1.0000	1.0000
8	1.0019	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	1.0001	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 18: ARL for PC2(m=600)

	T_{11}	T_{12}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	380.2281	349.6503	325.2252	320.2643	348.7562	381.6794	322.5806
1	253.1646	120.1923	185.1852	207.4689	271.7391	270.6753	224.2703
2	42.1941	17.7873	33.9905	169.4915	199.2032	221.3797	267.2389
3	11.7288	4.0064	5.6664	111.1111	140.0560	151.5714	178.0574
4	3.9250	1.6501	1.8967	65.1890	79.1139	88.0664	107.4956
5	1.8258	1.1206	1.1654	28.0112	39.0320	39.2144	47.0625
6	1.2090	1.0152	1.0213	9.9483	13.2485	13.8634	16.7893
7	1.0372	1.0006	1.0010	3.4732	4.3948	4.3648	5.5372
8	1.0036	1.0000	1.0000	1.6342	1.8930	1.1329	2.9168
9	1.0001	1.0000	1.0000	1.1288	1.1972	1.2607	1.2025
10	1.0000	1.0000	1.0000	1.0160	1.0280	1.0417	1.0296
11	1.0000	1.0000	1.0000	1.0010	1.0017	1.0035	1.0022
12	1.0000	1.0000	1.0000	1.0001	1.0001	1.0001	1.0001

Table 19: ARL for PC23(m=600)

	T_{11}	T_2	T_{31}	T_{32}	T_{33}	T_{34}
0	375.9398	326.2443	325.2252	384.0909	340.1361	385.7143
1	210.9705	147.9290	232.5581	320.5128	354.6099	310.5590
2	40.6174	27.0856	221.2389	287.3563	373.1343	310.5590
3	12.3365	5.2598	231.4815	265.9574	387.5969	310.5590
4	4.0700	1.8858	233.6449	265.9574	347.2222	299.4012
5	1.8669	1.1814	208.3333	312.5000	314.4654	273.2240
6	1.2212	1.0236	227.2727	271.7391	367.6471	312.5000
7	1.0398	1.0018	226.2443	273.2240	370.3704	322.5806
8	1.0046	1.0000	209.2050	273.2240	359.7122	290.6977
9	1.0001	1.0000	236.9668	255.1020	384.6154	314.4654
10	1.0000	1.0000	286.4894	225.7587	357.5278	323.7755
11	1.0000	1.0000	254.4568	279.5630	324.7858	312.5278
12	1.0000	1.0000	244.5828	256.6786	328.5786	305.7885

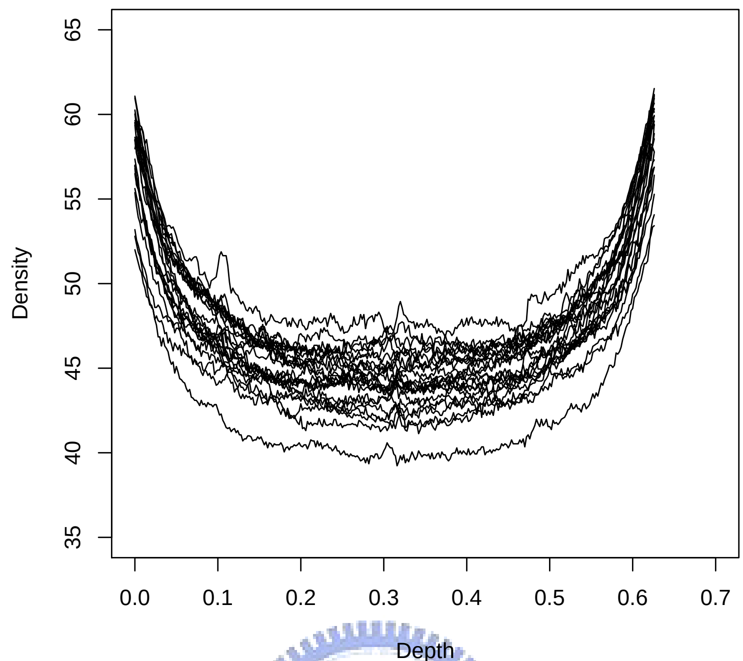

Figure 1: Original 24 VDP-profiles

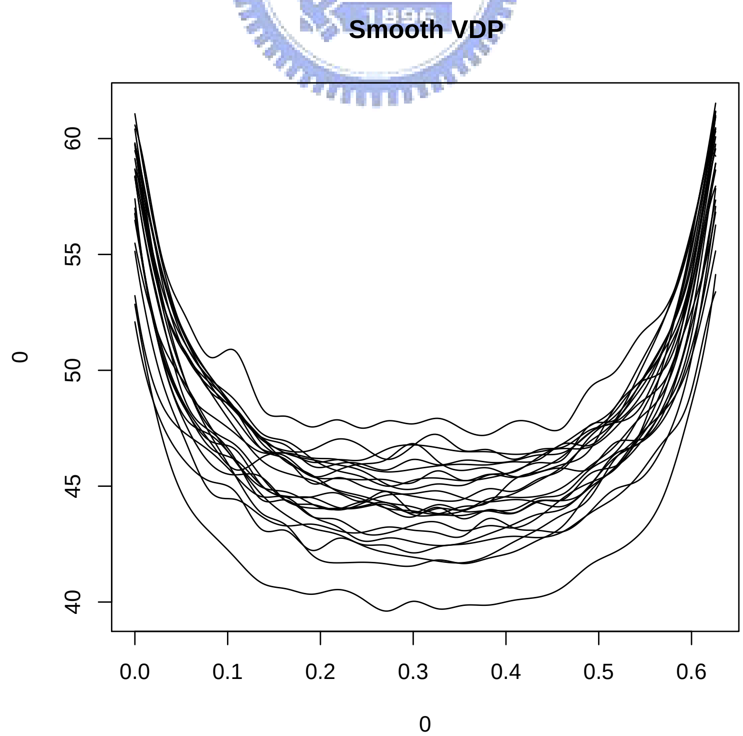

Figure 2: Smoothed VDP

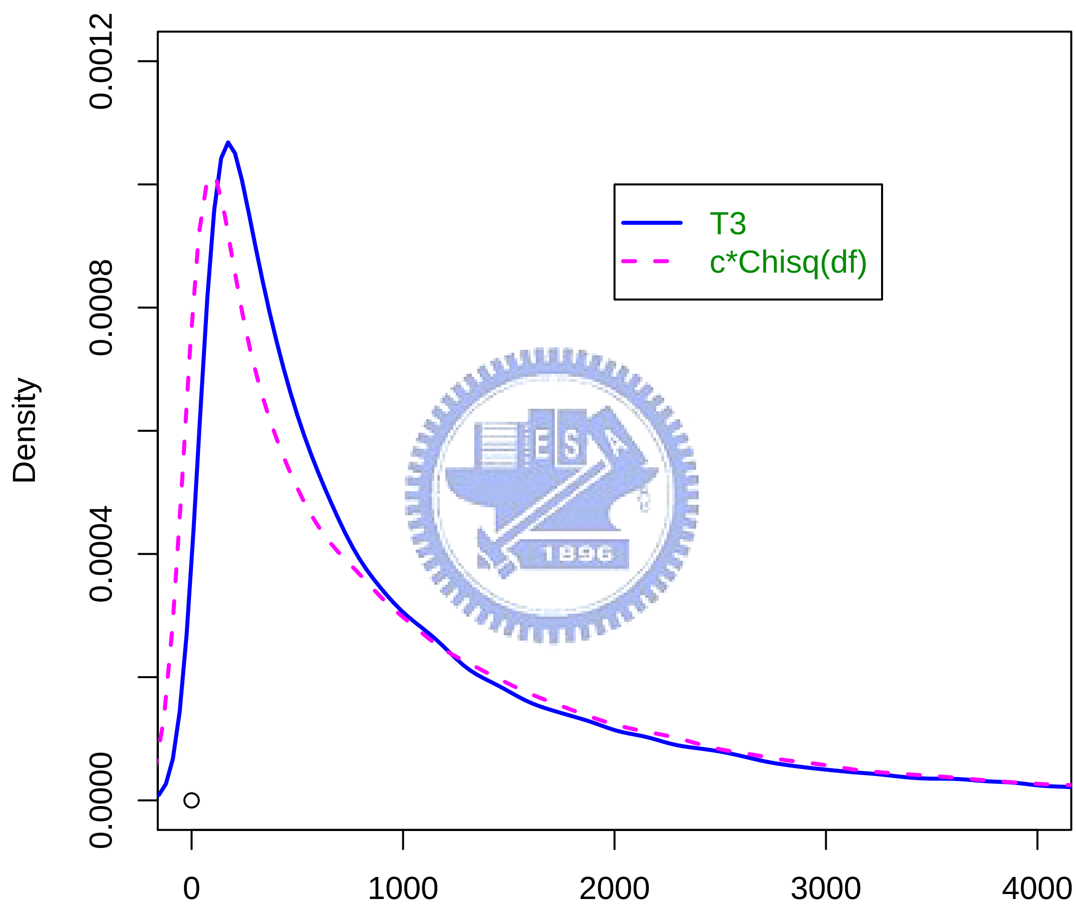

Figure 3: Density for T_3

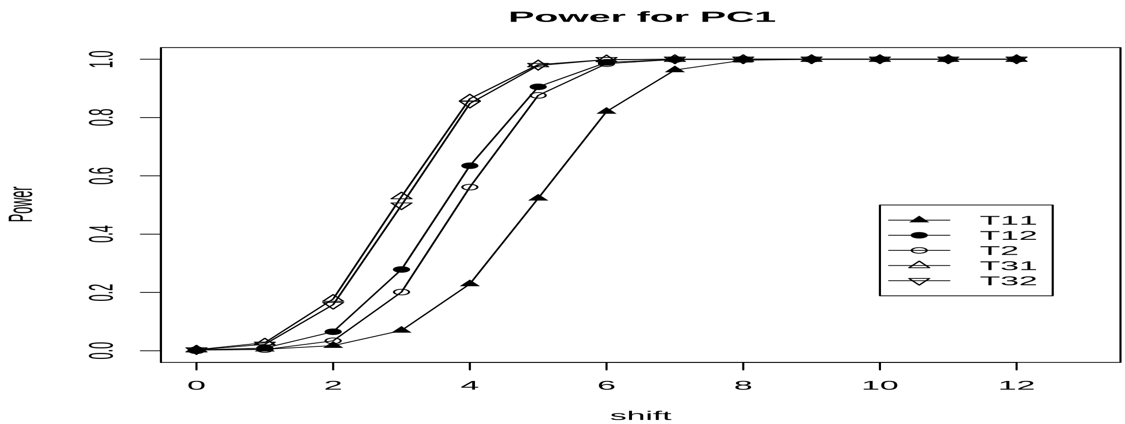

Figure 4: Power of PC1 (true VDP)

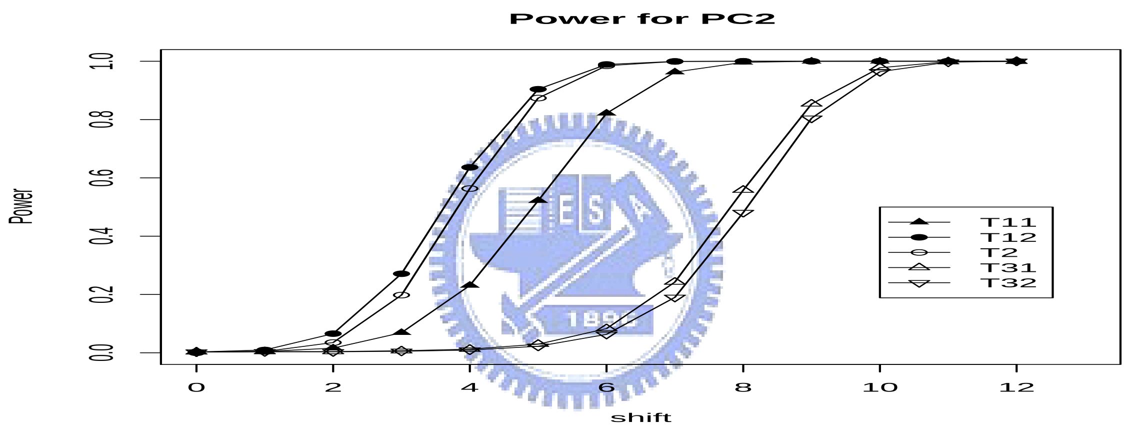

Figure 5: Power of PC2 (true VDP)

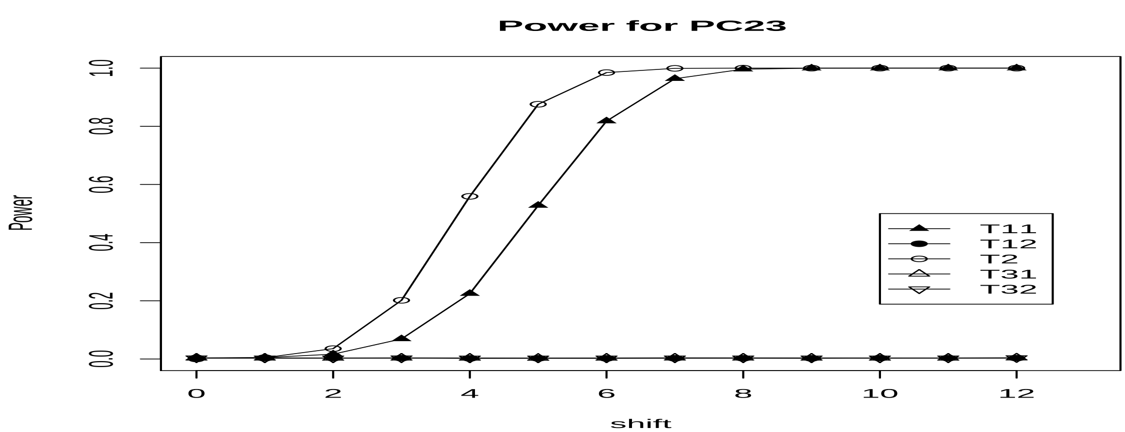

Figure 6: Power of PC23 (true VDP)

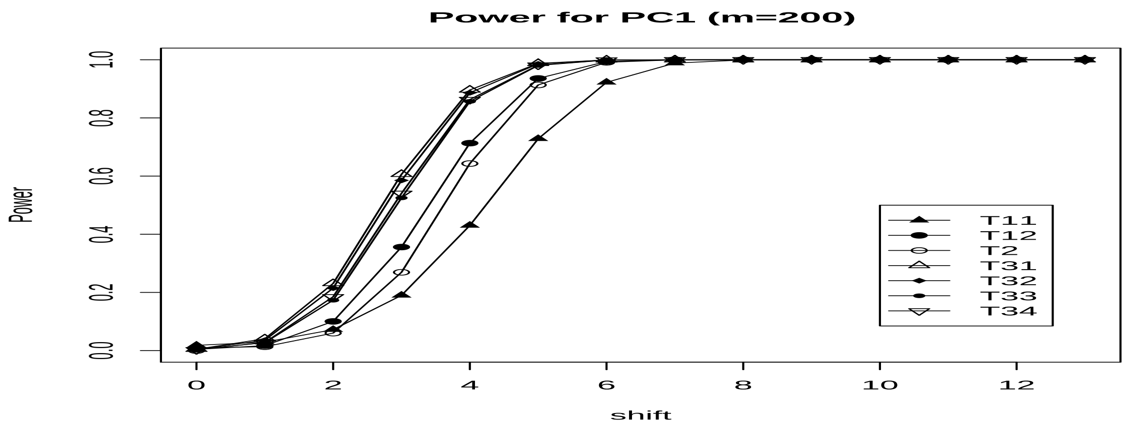

Figure 7: Power for PC1 with m=200

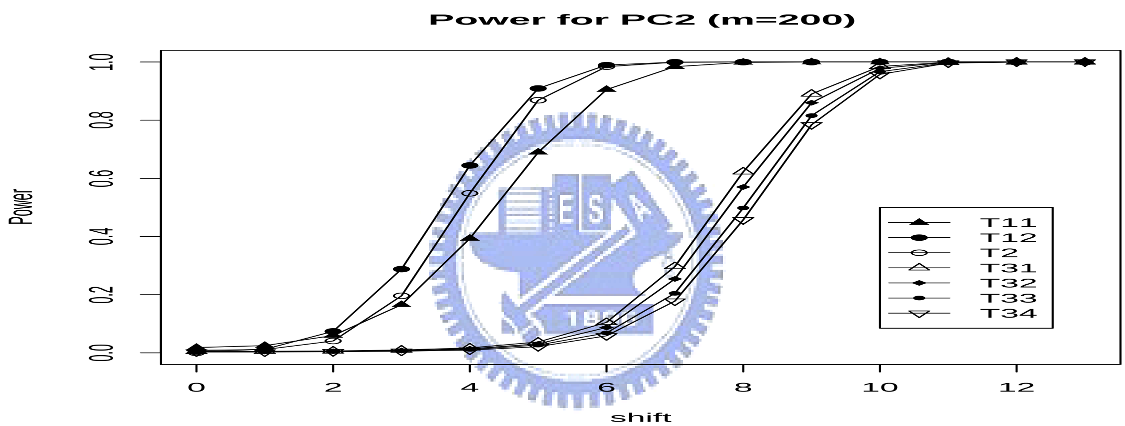

Figure 8: Power for PC2 with m=200

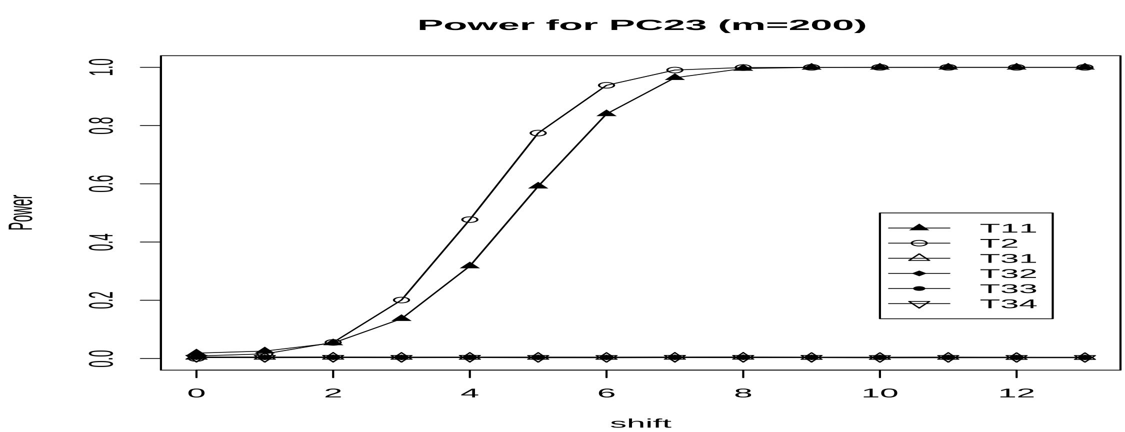

Figure 9: Power for PC23 with m=200

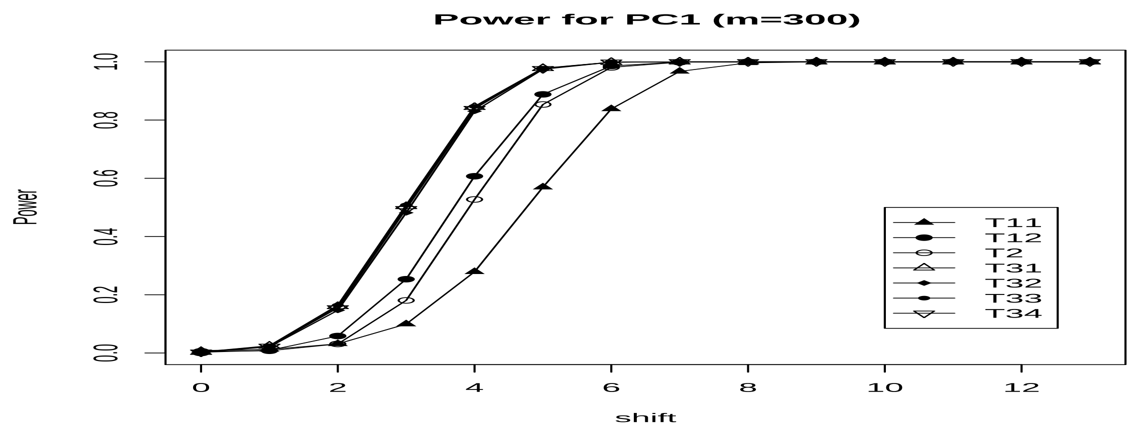

Figure 10: Power for PC1 with m=300

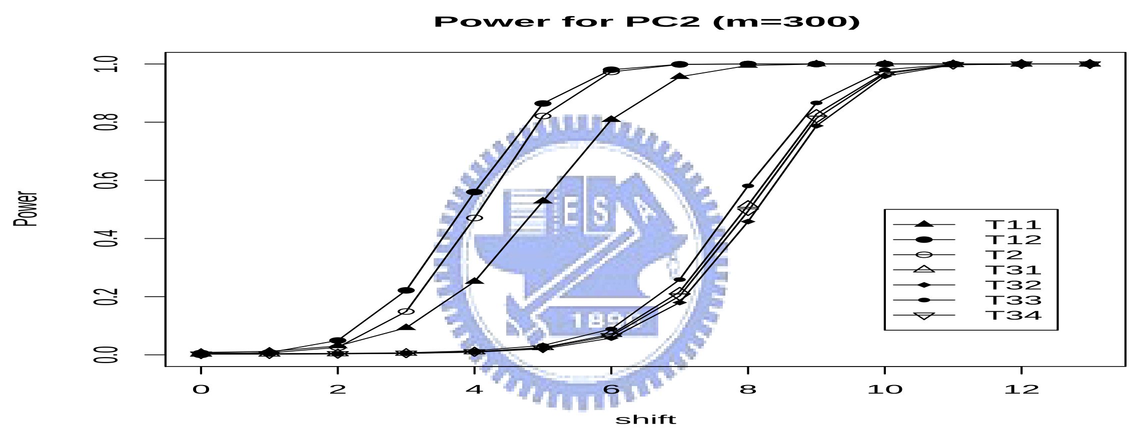

Figure 11: Power for PC2 with m=300

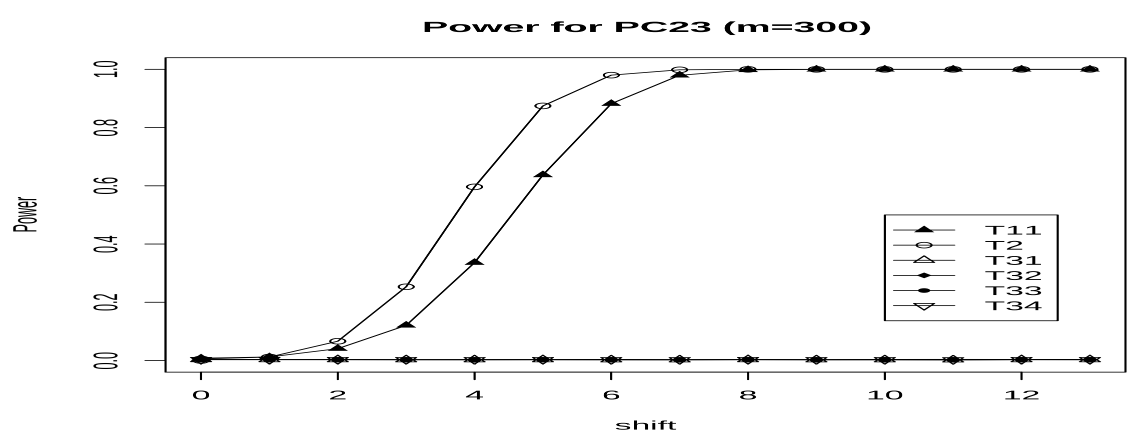

Figure 12: Power for PC23 with m=300

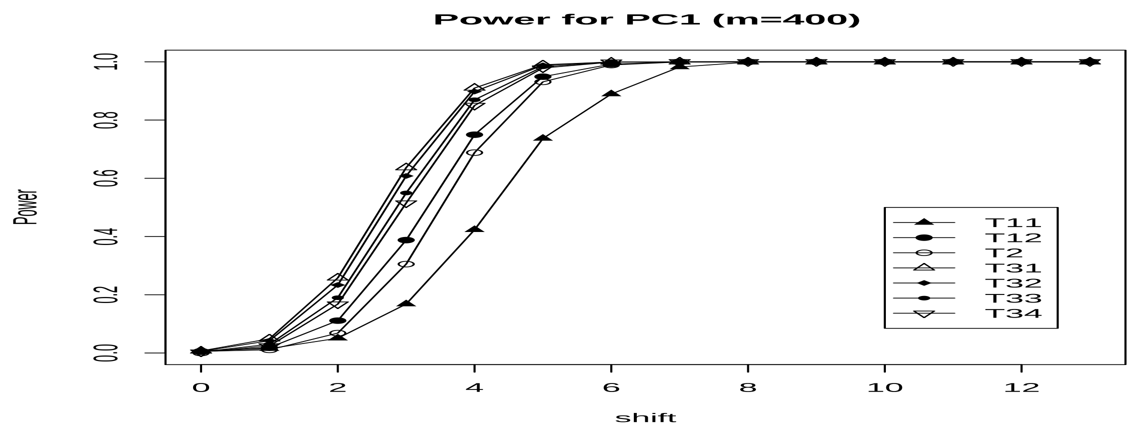

Figure 13: Power for PC1 with m=400

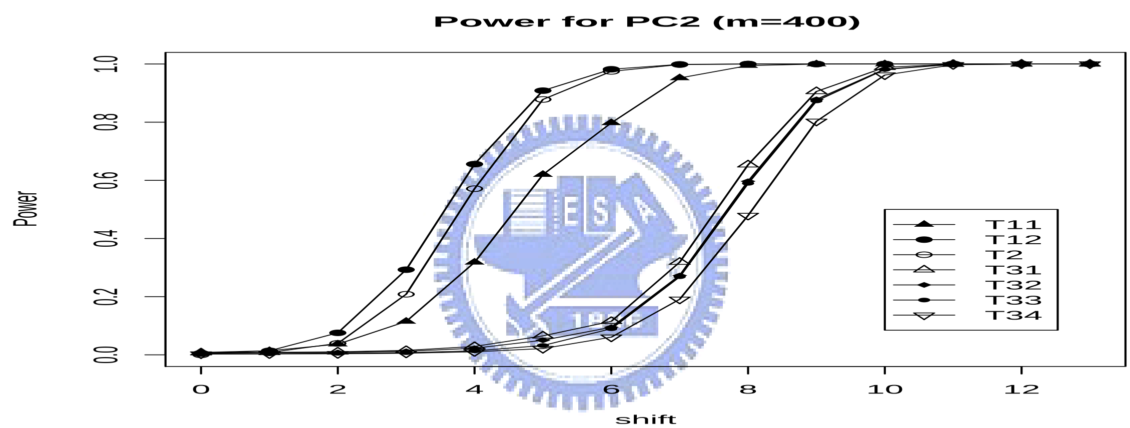

Figure 14: Power for PC2 with m=400

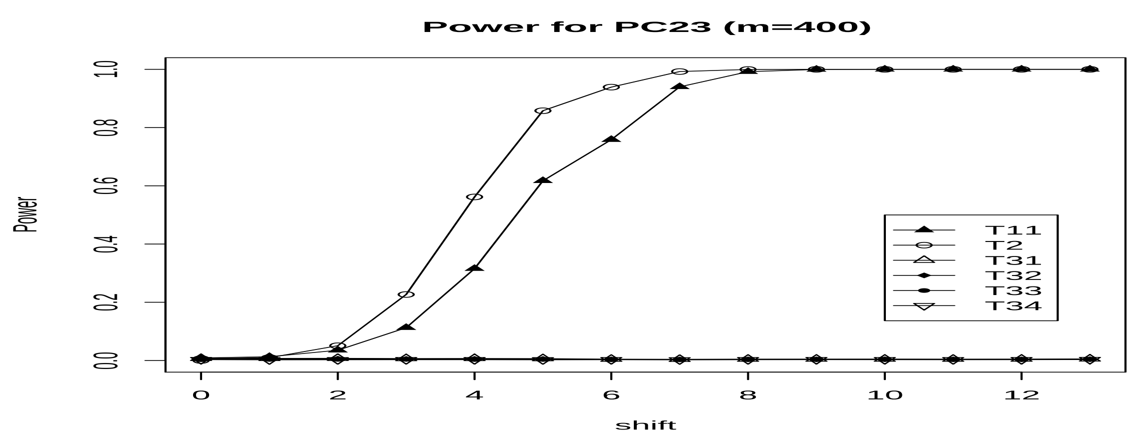

Figure 15: Power for PC23 with m=400

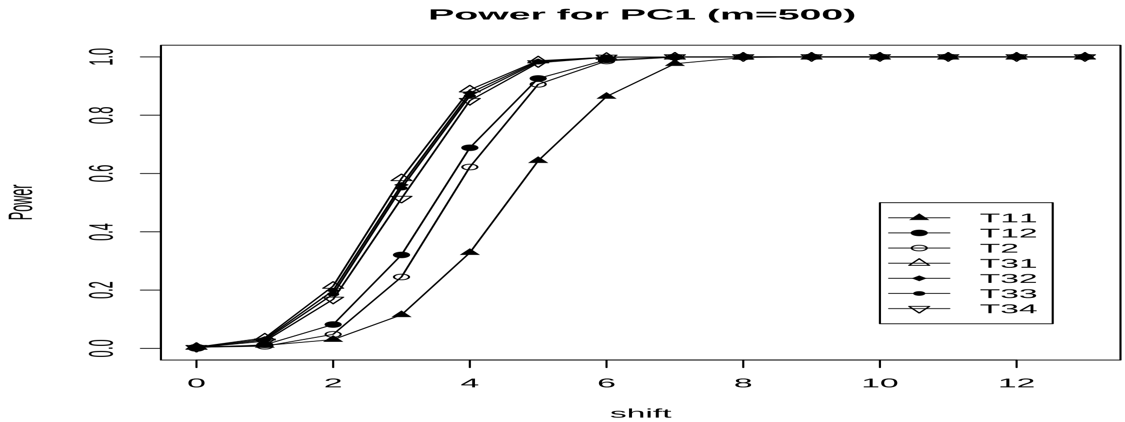

Figure 16: Power for PC1 with m=500

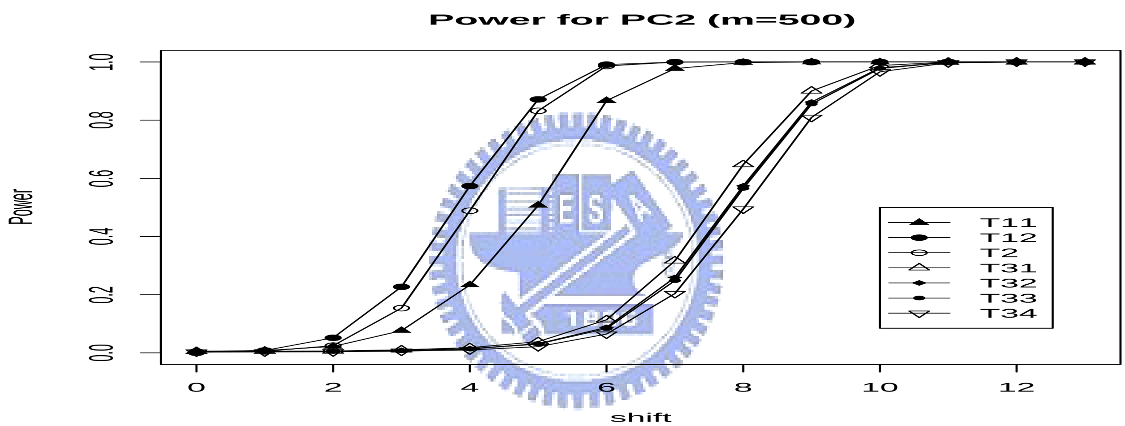

Figure 17: Power for PC2 with m=500

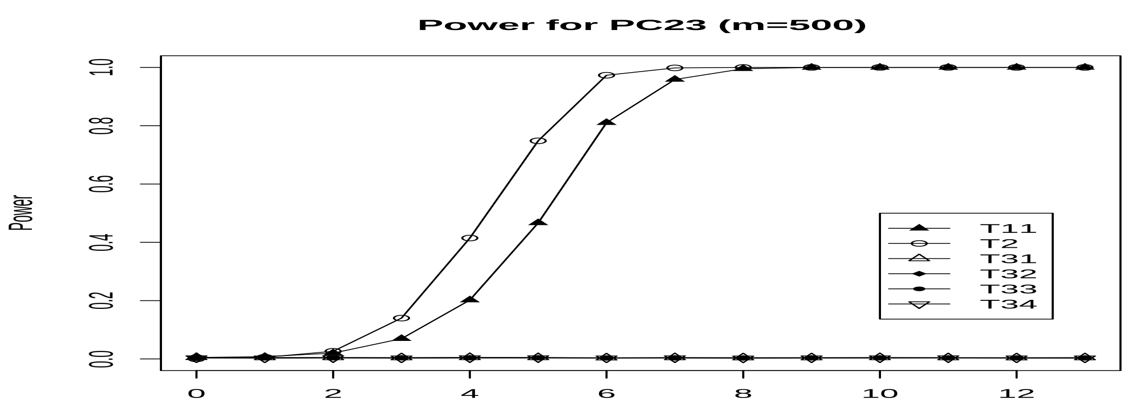

Figure 18: Power for PC23 with m=500

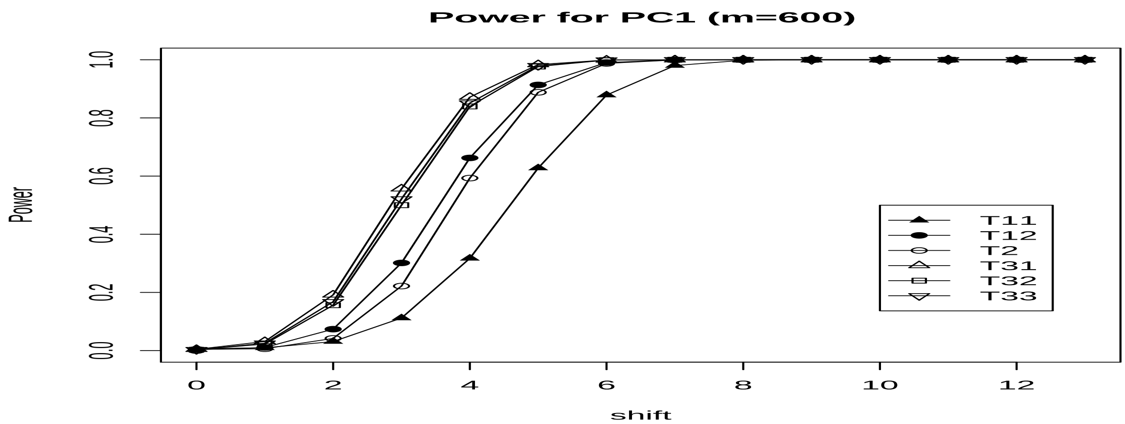

Figure 19: Power for PC1 with m=600

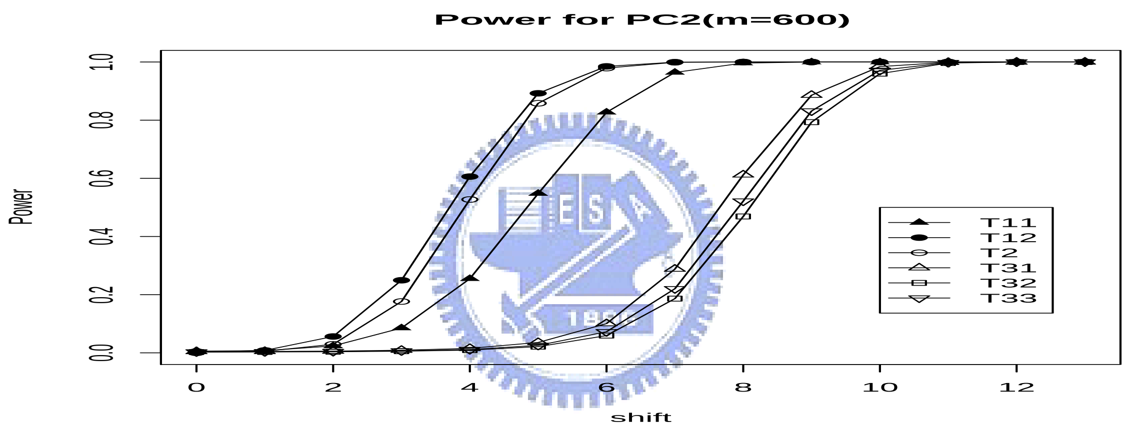

Figure 20: Power for PC2 with m=600

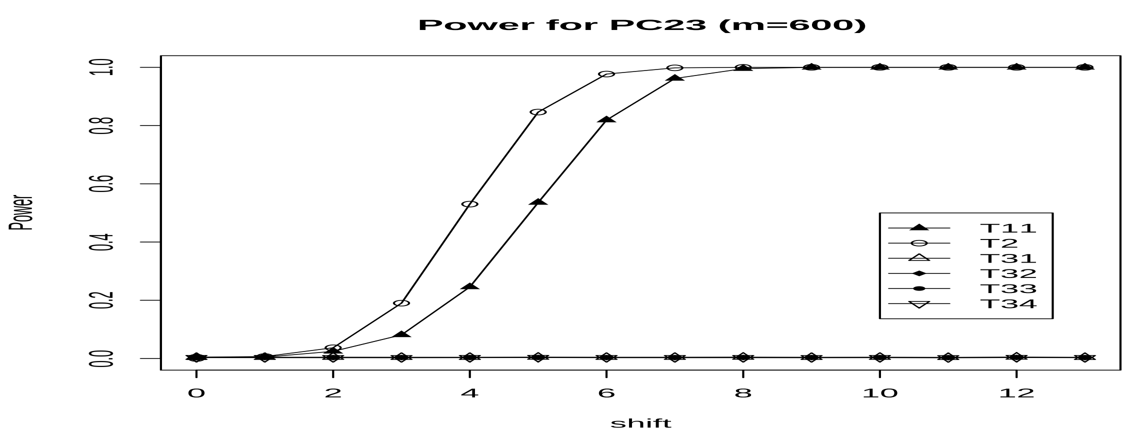

Figure 21: Power for PC23 with m=600