The Pennsylvania State University
The Graduate School

MEASURES OF AGREEMENT IN METHOD COMPARISON STUDIES FOR INTENSIVE LONGITUDINAL DATA

A Dissertation in
Statistics
by
Mina Yoo

© 2011 Mina Yoo

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2011

The dissertation of Mina Yoo was reviewed and approved* by the following:

Runze Li
Professor of Statistics
Dissertation Co-Adviser
Co-Chair of Committee

Mosuk Chow
Associate Professor of Statistics
Dissertation Co-Adviser
Co-Chair of Committee

Naomi Altman
Associate Professor of Statistics

Vernon M. Chinchilli
Distinguished Professor of Public Health Sciences
Professor of Statistics

Bruce G. Lindsay
Willaman Professor of Statistics
Head of the Department of Statistics

*Signatures are on file in the Graduate School.

Abstract

This dissertation is concerned with assessment of measurement agreement for intensive longitudinal data. Assessment of the measurement agreement encompasses a variety of applications. A number of indices for measuring agreement have been developed. However, these measures make a major assumption: that the mean and variation are stable over time. With recent developments in data collection methods and statistical models, intensive longitudinal studies and the analysis of intensive longitudinal data are gaining popularity across many areas. Intensive longitudinal data enable researchers to examine more detailed features of how processes change over time.

Due to its high intensity of assessments within subjects, it has different characteristics from traditional longitudinal data, which often involve a small number of repeated observations across many individuals. The overall mean of intensive longitudinal data is typically a smooth curve of time and variance of the error process may be time-varying over study duration. Moreover, heterogeneity of intra subject processes such as autocorrelation and instability exists. To overcome these challenges and provide accuracy estimates, we first propose a novel estimation procedure for functional mixed models and partially linear mixed models and study the asymptotic properties of the proposed estimation procedure. Then, we develop a new index of the agreement for intensive longitudinal data, the functional type of concordance correlation coefficient based on proposed models. The functional concordance correlation coefficient is robust with respect to model specification, compared with the popular index, the unified approach of concordance correlation coefficient. The proposed index improves the accuracy of measurement agreement by separating the time trend of measurements from the degree of agreement. All the proposed procedures are assessed by intensive finite sample simulation studies and most are illustrated with real data examples.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my sincere gratitude to my advisor, Dr. Runze Li, who has irreplaceable influence on my academic career. Without his invaluable advice and endless patient and inspiring guidance, it would be impossible for me to continue the research work and finish this dissertation. His broad knowledge, brilliant ideas and nice personality make him a great mentor. I learn from him not only how to do research in statistics, but also how to behave and work with people.

I am very grateful to my co-advisor, Dr. Mosuk Chow, for her time and patience for helping me and providing data sets for me to work on. I also would like to thank the committee members, Dr. Vernon M. Chinchilli and Dr. Naomi Altman, for their helpful comments on my dissertation.

Thanks to all faculty, staff and graduate students in the Department of Statistics, Penn State University, for making my stay here such an enriching experience.

Finally, I would like to acknowledge my husband, Mr. Hyeong Jae Lee. Without his true love, support and understanding, I could not finish this dissertation.

This research was supported by National Institute on Drug Abuse (NIDA) grants P50-DA10075, R21-DA024260 and National Science Foundation grant DMS 0348869.

Chapter 1

Introduction

1.1 A Brief Overview of the Literature

Assessment of measurement agreement encompasses a variety of applications. In clinical practice, measurements serve as a basis for diagnostic and prognostic evaluations. When new methods or devices for diagnostic and prognostic evaluations become available, assessing utility relies on the degree of closeness of their measurement to a standard. In quantitative psychology, rating agreements are often used to evaluate a new rating system or instrument.

Agreement indices have been studied over many decades. For nominal and ordinal scale of outcome variable, kappa (Cohen, 1960) and weighted kappa (Cohen, 1968; Fleiss et al., 1969) are typical methods for assessing agreement. The concordance correlation coefficient has gained popularity as a measure of agreement since its introduction by Lin (1989). The concordance correlation coefficient measures the variation of the linear relationship between two methods from the identity line including two meaningful components: the deviations of each measurement around the linear line (precision) and the distance between the identity line and the linear line (accuracy; Lin (1989)). The concordance correlation coefficient (CCC) originally aims for two methods, each taking a single observation on each subject.

Many extensions of the CCC have been suggested to involve repeated measures taken by multiple methods. King & Chinchilli (2001a) and Barnhart et al. (2002) proposed Generalized CCC and Overall CCC for multiple methods, each making a single repetition on a subject. The degree of agreement for repeated measures are evaluated by Chinchilli et al. (1996) and King et al. (2007) for paired data. The CCC is extended to the degree of agreement between multiple methods for repeated measurements by Carrasco & Jover (2003) and Barnhart et al. (2005). Carrasco & Jover (2003) estimated the CCC by variance components of the two-way mixed model and demonstrated that it yields more stable estimates of an agreement index then the moment method. Barnhart et al. (2005) factorized the source of disagreement into inter-, intra- and total- agreement so that the detailed information may be utilized to improve agreement. Lin et al. (2007) proposed the unified approach of concordance correlation coefficient (UCCC) to evaluate agreement for repeated measures taken by multiple methods. The UCCC integrated the approaches of Carrasco & Jover (2003) and Barnhart et al. (2005); three indices consist of variance components of the two-way mixed model. The UCCC performs well when observations have stable mean and variation of repeated measures over occasion. However, these underlying assumptions restrict the utility of the UCCC for densely measured longitudinal data, which have smooth curve and a time-varying variance. A more detailed review follows in the next chapter.

This dissertation is motivated by analysis of a real data set in Chapter 5. The real data set was collected in a physiological study, in which core body temperatures were collected every minute over an 81 minutes period for each of six trials from 14 subjects. Figure 1.1 illustrates the trend of core body temperature of trials by subjects over time. 14 black curves stand for the first trial and red, green, blue, light blue and purple represent trials 2 through 6, respectively. The goal of this research

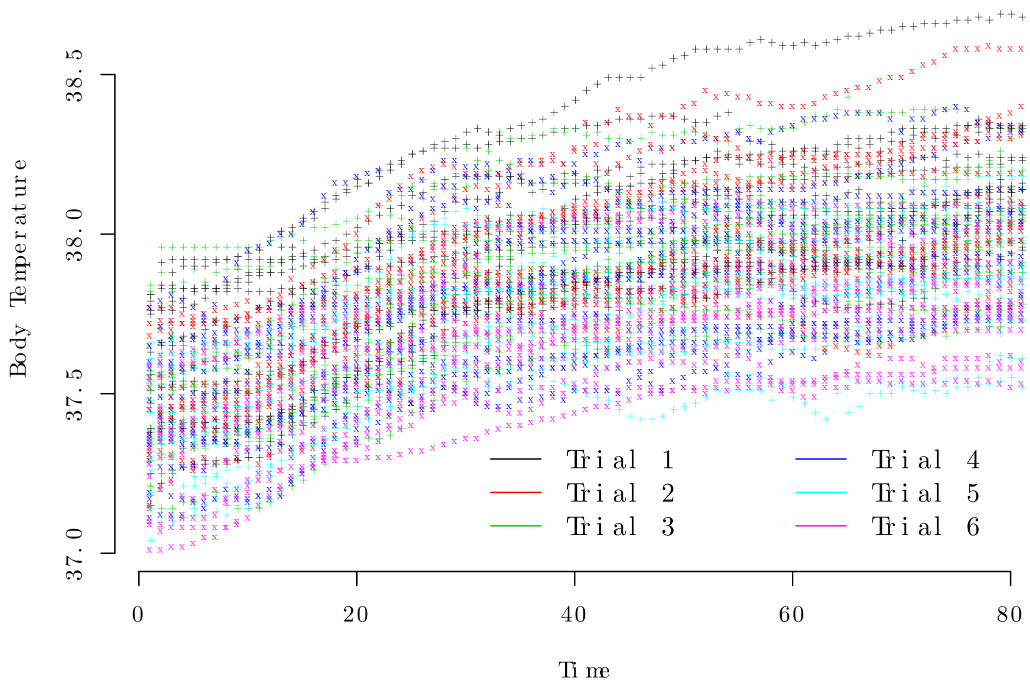

Figure 1.1. Plots for core body temperature of six trials by 14 subjects. 14 black curves stand for the first trial and red, green, blue, light blue, purple represents trial 2 through trial 6 respectively

is to assess the degree of agreement between trial curves of core body temperature. Because core body temperature smoothly increases over time, the aforementioned methods for assessment of measurement agreement cannot be directly applied for such functional data. This consideration motivates us to develop a novel tool for assessing agreement for intensive longitudinal data.

The recent developments in data collection methods enable researchers to collect data more frequently and flexibly with lower cost. For example, the development of automatic monitoring of pains in clinical research enables massive amounts of data collection faster and more efficiently. Web-based systems and smart phones are emerging tools in healthcare and social studies to gather data. Intensive longitudinal data have become more and more common in various research areas. Intensive longitudinal data (ILD) provide more detailed features on how processes change over time. Due to high intensity of assessments within subject, ILD often have different characteristics from traditional longitudinal data, which often involve a small number of repeated observations across many individuals. The overall mean of ILD is typically a smooth curve of time and the variance of the error process may be time-varying over study duration. In addition, heterogeneity of intra-subject processes such as autocorrelation and instability exists. To overcome these challenges in the estimation of ILD, various estimation and inference procedures for nonparametric and semiparametric models have been proposed and studied in the literature. However, most existing estimation procedures focus on the statistical accuracy of regression coefficients in ILD. This dissertation aims to develop a novel estimation procedure to achieve accuracy estimates of variance components as well as coefficients for the purpose of time-varying agreement index relying on the variance components in the base model.

1.2 The Contributions of This Thesis

With growing interest in ILD, this dissertation aims to develop a functional measure of the degree of agreement between intensive repeated measurements taken by multiple methods. Li & Chow (2005) introduced a measure of agreement for functional data and proposed functional type weighted CCC for overall agreement of paired repeated measures. Because the index accommodates the time trend on repeated measures, it is used to quantify agreement of curve data when assessing the amount of agreement between two distinct methods of measuring the same response variable without covariate.

Motivated by Li & Chow (2005) and Lin et al. (2007), we propose the functional concordance correlation coefficient (FCCC) for ILD to encompass time varying measurements taken by multiple methods considering possible confounding effects. The FCCC is a covariance-based index relying on the variance components in the base model. This index delineates the trend of concordance of measurements over time by summarizing the degree of agreement at each time point. Moreover, the FCCC is flexible for data structure. We show that, compared with the UCCC, the FCCC significantly improves detection of agreement by separating the time trend from the degree of agreement between densely repeated measurements of multiple methods regardless of the degree of serial correlation in repeated measurements.

Our second goal is to study an estimation procedure of nonparametric and semiparametric models with time-varying covariate for ILD considering serial correlation. Most existing estimation procedures are focused on the statistical accuracy of regression coefficients in ILD. Interestingly, we found that the existing estimation procedures perform poorly for estimating the variance components under serial correlation. We consider functional mixed models and partially linear mixed models, and develop a new estimating procedure for nonparametric and semiparametric models that are robust and computationally simple for ILD with serial correlation. In addition, we further investigate the performance of bootstrap inference of the FCCC and propose a modification of parametric bootstrap for functional mixed model and partially linear mixed effects model. Extensive Monte Carlo simulation studies are conducted to examine the finite sample performance of the proposed procedures. We show that the estimation procedure is robust against serial correlation, which poses challenges to accurate variance estimation.

1.3 The Organization of This Thesis

The dissertation is organized as follows. In Chapter 2, we provide the literature review for this dissertation research. Chapter 3 focuses on the FCCC in a nonparametric covariate setting. In Chapter 4, we study the FCCC based on the robust estimation for the partially linear mixed models. Simulation studies are conducted in Chapters 3 and 4 to assess the performance of the proposed estimation procedure and FCCC in various settings. A real data example is also illustrated in Chapter 5 to advocate the use of FCCC. Finally, concluding remarks and future research direction are given in Chapter 6.

Chapter 2

Literature Review

This chapter provides a brief literature review of the dissertation research. This dissertation uses research findings from two topics: measurement agreement indices and nonparametric smoothing techniques. Both areas have been studied for many decades and are still active topics in statistics.

2.1 Assessment of measurement agreement

2.1.1 Independent data - two methods

Measuring agreement between two variable is the objective of some comparison studies. For example, two raters observe the subjects in a study and the reliability between raters arises as to how well two raters agree. In clinical studies, it is desirable to measure the amount of agreement between two devices measuring the same response variable.

2.1.1.1 Kappa and weighted kappa

A popular index for categorical response is kappa by Cohen (1960) and weighted kappa by Cohen (1968) to measure agreement for binary or ordinal data. Kappa assesses agreements for two observers when they took single replications on a subject. Suppose two observers placed N subjects into m mutually exclusive and exhaustive categories independently. Let p_{ij} be the proportion of being placed into the cell (i, j), and $p_{i.} = \sum_{j=1}^{m} p_{ij}$ be the proportion of being placed in the i_{th} category by the first observer; $p_{.j}$ can be defined similarly for the second observer.

		observer 2				
		1	2	$\cdots$	m	
observer 1	1	p_{11}	p_{12}	$\cdots$	p_{1m}	$p_{1\cdot}$
	2	p_{21}	p_{22}	$\cdots$	p_{2m}	$p_{2\cdot}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	m	p_{m1}	p_{m2}	$\cdots$	p_{mm}	$p_{m\cdot}$
		$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot m}$	1

The kappa coefficient is defined as follows:

$$\kappa = \frac{p_o - p_c}{1 - p_c}, \tag{2.1}$$

where $p_o = \sum_{i=1}^{m} p_{ii}$, the observed proportion of agreement and $p_c = \sum_{i=1}^{m} p_{i.} p_{.i}$, the proportion of expected by chance. For two independent ratings of the n subjects, the kappa coefficient equals +1 when there is complete agreement of the observers. When the observed agreement (p_o) exceeds chance agreement (p_c), kappa is positive, with its magnitude reflecting the strength of agreement. Kappa is negative when the observed agreement is less than chance agreement. The minimum value of kappa is between -1 and 0, depending on the product-moment correlation coefficient. The sample estimate, $\hat{\kappa}$ is obtained by replacing the observed proportion of agreement. Cohen (1960) derived the variance, which was found to be incorrect later. The asymptotic variance of the kappa coefficient is estimated by the following, according to Fleiss et al. (1969).

$$\hat{\text{var}}(\hat{\kappa}) = \frac{\sum_j p_{jj}\left\{1-(p_{j\cdot}+p_{\cdot j})(1-\hat{\kappa})\right\}^2 + (1-\hat{\kappa})^2 \sum\sum_{i\neq j} p_{ij}(p_{\cdot j}+p_{j\cdot})^2 - \left\{\hat{\kappa}-p_c(1-\hat{\kappa})\right\}^2}{(1-p_c)^2 N}. \tag{2.2}$$

The confidence interval, $\hat{\kappa} \pm Z_{\alpha/2}\sqrt{\hat{\text{var}}(\hat{\kappa})}$, performs as a hypothesis test. The Kappa assesses non-chance (chance-corrected) agreement relative to the total non-chance agreement. All disagreement has the same weight, so it is limited where certain disagreement between two observers is more serious than others, or ordered. Weighted kappa proposes to overcome the limitations of kappa with cell weights. Let w_{ij}, assumed without loss of generality to lie between 0 and 1, be the weight assigned to the cell (i, j). Then the weighted kappa is given as follows:

$$\kappa_w = \frac{p_o^w - p_c^w}{1 - p_c^w}. \tag{2.3}$$

The weighted kappa has the same formula to the (2.1) except $p_o^w = \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}p_{ij}$ and $p_c^w = \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}p_{i\cdot}p_{\cdot i}$, where w_{ij} is the weight assigned to the cell (i, j) among between 0 and 1. The common used weight sets are

$$w_{ij} = 1 - \frac{(i-j)^2}{(m-1)^2}, \qquad i, j = 1, 2, \ldots, m$$

or

$$w_{ij} = 1 - \frac{|i-j|}{|m-1|}. \qquad i, j = 1, 2, \ldots, m$$

If w_{ij} is indicator function $I\{i = j\}$, the weighted kappa reduces to kappa. The asymptotic variance of the weighted kappa coefficient is estimated by Fleiss et al.

(1969) as follows:

$$\hat{\text{var}}(\hat{\kappa}_w) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\left\{w_{ij} - (\sum_{j=1}^{m} p_{\cdot j} w_{ij} + \sum_{i=1}^{m} p_{i\cdot} w_{ij})(1-\hat{\kappa}_w)\right\}^2 - \{\hat{\kappa}_w - p_c^w(1-\hat{\kappa}_w)\}^2}{(1-p_c^w)^2 N}. \quad (2.4)$$

Both kappa and weighted kappa are simple and easy to implement but they cannot make distinctions among various types and sources of disagreement; it is difficult to extract influence by observers' prevalence and base-rates, and it is impossible to compare when one is interested in measuring the consistency of ratings for observers that use different categories (e.g., one uses a scale of 1 to 5, another uses a scale of 1 to 7). Nevertheless, kappa and weighted kappa are the first approach as in measuring agreement for categorical data, and are the most popular being used for categorical data.

2.1.1.2 Concordance correlation coefficient

For continuous data with two methods, a popular index is concordance correlation coefficient(CCC). Lin (1989) reviewed and criticized the basic model on assessing agreement in continuous data, and introduced the CCC, which quantifies how similar measurements from paired two methods are under regularity conditions. Assume that (Y_{i1}, Y_{i2}), $i = 1, \cdots n$ are n paired samples randomly selected from a bivariate population with mean μ_1, μ_2, variance σ_1^2, σ_2^2 respectively and σ_{12} covariance. Then CCC is defined as

$$\begin{aligned} \rho_c &= 1 - \frac{E((Y_1 - Y_2)^2)}{(\mu_1 - \mu_2)^2 + \sigma_1^2 + \sigma_2^2} \\ &= \frac{2\sigma_{12}}{\sigma_1^2 + \sigma_2^2 + (\mu_1 - \mu_2)^2}. \end{aligned} \quad (2.5)$$

CCC decomposes into meaningful precision (ρ), and accuracy (χ_α) components, so it provides information on how far each observation deviates from the fitted line based on data and how far the line deviates from the identity line through the origin, which is the source of measurement error. CCC has similar properties as the Pearson correlation coefficient; the range is between -1 and 1; -1 stands for perfect negative agreement while 1 reflects perfect positive agreement. Lin (1989) proposed sample counterpart of marginal moments and covariance to compute $\hat{\rho}_c$.

$$\hat{\hat{\rho}}_c = \frac{2S_{12}}{S_1^2 + S_2^2 + (\bar{Y}_1 - \bar{Y}_2)^2}, \tag{2.6}$$

where $S_j = \sum_{i=1}^{n}(Y_{ij} - \bar{Y}_j)^2,\ j = 1, 2,\ \bar{Y}_j = \frac{1}{n}\sum_{i=1}^{n} Y_{ij}$. $\hat{\rho}_c$ is a consistent estimator of ρ_c and has an asymptotic normal distribution with mean ρ_c and variance

$$\sigma^2_{\hat{\rho}_c} = \frac{1}{n-2}\left[(1-\rho^2)\rho_c^2(1-\rho_c^2)/\rho^2 + 4\rho_c^3(1-\rho_c)u^2/\rho - 2\rho_c^4 u^4/\rho^2\right],$$

where $u = \frac{(\mu_{y_1} - \mu_{y_2})^2}{\sigma^2_{y_1}\sigma^2_{y_2}}$.

In order to obtain a more powerful test, Lin (1989) applied the inverse hyperbolic tangent transformation

$$\rho_Z = 0.5 \ln\left(\frac{1+\rho_c}{1-\rho_c}\right).$$

Then the variance becomes

$$\sigma_Z^2 = \frac{1}{n}\left[\frac{(1-\rho^2)\rho_c^2}{(1-\rho_c^2)\rho^2} + \frac{4\rho_c^2(1-\rho_c)u^2}{\rho(\rho_c^2)^2} - \frac{2\rho_c^2 u^4}{\rho^2(1-\rho_c^2)^2}\right].$$

Lin et al. (2002) proposed four agreement statistics besides CCC: precision(Pearson correlation coefficient), accuracy, mean squared deviation (MSD), total deviation index (TDI), and coverage probability (CP). TDI and CP offer better power for inference with intuitive interpretation.

2.1.1.3 Generalized concordance correlation coefficient for independent data

King & Chinchilli (2001b) argued that CCC may not be robust and may fail to assess the accurate agreement if there are extreme outliers. CCC by Lin (1989) is based on L_2 distance which is sensitive to outliers, so CCC is not robust with heavy-tail distributions. King & Chinchilli (2001b) proposed a robust version of the CCC by applying alternative distance functions and introduced robust estimators. Define $g(\cdot)$ as a distance function which satisfies three properties:

1. $g(0) = 0$
2. $g(z)$ is an even function, i.e., $g(-z) = g(z)$ for all z
3. $g(z)$ is an non-decreasing function of z for all $z \geq 0$

The robust version of the CCC is

$$\rho_g = \frac{E_{F_X F_Y}\left[g(X-Y) - g(X+Y)\right] - E_{F_{XY}}\left[g(X-Y) - g(X+Y)\right]}{E_{F_X F_Y}\left[g(X_i - Y_j) - g(X_i + Y_j)\right] + \frac{1}{2} E_{F_{XY}}\left[g(2X) + g(2Y)\right]},$$

where $g(X, Y)$ is an integrable function with respect to F_{XY}, which is the cumulative bivariate distribution for (X, Y), F_X and F_Y are the marginal cumulative distribution of X and Y respectively. Then the estimator of the CCC is expressed as

$$\begin{aligned}\hat{\rho}_g &= \frac{\frac{1}{n}\sum_i\sum_j \left[g(X_i - Y_j) - g(X_i + Y_j)\right] - \sum_i \left[g(X_i - Y_i) - g(X_i + Y_i)\right]}{\frac{1}{n}\sum_i\sum_j \left[g(X_i - Y_j) - g(X_i + Y_j)\right] + \frac{1}{2}\sum_i \left[g(2X_i) + g(2Y_i)\right]} \\ &= \frac{(n-1)(U_3 - U_1)}{U_1 + nU_2 + (n-1)U_3}, \end{aligned} \tag{2.7}$$

where

$$U_1 = \sum_{i \neq j} \frac{\frac{1}{2}\left[g(X_i - Y_i) + g(X_j - Y_j)\right] - \frac{1}{2}\left[g(X_i + Y_i) + g(X_j + Y_j)\right]}{n(n-1)},$$

$$U_2 = \sum_{i \neq j} \frac{\frac{1}{4}\left[g(2X_i) + g(2X_j) + g(2Y_i) + 2g(2Y_j)\right]}{n(n-1)},$$

$$U_3 = \sum_{i \neq j} \frac{\left[g(X_i - Y_j) - g(X_i + Y_j)\right] + \frac{1}{2}\left[g(X_j - Y_i) - g(X_j + Y_i)\right]}{n(n-1)}.$$

Here, U statistics U_1, U_2, and $U_1 + (n-1)U_3$ are unbiased estimators of $E_{F_X F_Y}$ $[g(X-Y) - g(X+Y)]$, $\frac{1}{2}E_{F_{XY}}\left[g(2X) + g(2Y)\right]$, and $E_{F_X F_Y}$ $[g(X_i - Y_j) - g(X_i + Y_j)]$ respectively. If $g(z)$ is a quadratic function, ρ_g reduces to CCC by Lin (1989). When X and Y each follow a multinomial distribution with ordinal levels, the robust version of the CCC can be specified to equal the kappa or weighted kappa coefficient.

The index is expressed as a function of U-statistics. Thus the estimator, $\hat{\rho}_g$ is normally distributed, and has consistent variance. King & Chinchilli (2001b) showed the asymptotic normality of (2.7) and consistent variance via U-statistics.

$$\text{var}(\hat{\rho}_g) = (\hat{\rho}_g^2)\left[\frac{\text{var}(\hat{H})}{\hat{H}^2} - \frac{2\text{cov}(\hat{H}, \hat{G})}{\hat{H}\hat{G}} + \frac{\text{var}(\hat{G})}{\hat{G}^2}\right], \tag{2.8}$$

where $H = (n-1)(U_3 - U_1)$ and $G = U_1 + nU_2 + (n-1)U_3$.

2.1.2 Independent data - multiple methods

Because the reliability and validity studies often involve more than two methods, there was a need to assess agreement among multiple methods.

2.1.2.1 Generalized concordance correlation coefficient

King & Chinchilli (2001a) proposed an generalized CCC for more than 2 responses $X_1, \cdots, X_p$. The basic idea is the same as (2.7) with all unique pair-wise responses $(X_i, X_j),\ \ 1 \leq i < j \leq p$.

$$\bar{\rho}_g = \frac{\sum\limits_{i<j} E_{F_{X_i}F_{X_j}}\left[g(X_i - X_j) - g(X_i + X_j)\right] - \sum\limits_{i<j} E_{F_{X_iX_j}}\left[g(X_i - X_j) - g(X_i + X_j)\right]}{\sum\limits_{i<j} E_{F_{X_i}F_{X_j}}\left[g(X_i - X_j) - g(X_i + X_j)\right] + \frac{1}{2}\sum\limits_{i<j} E_{F_{X_iX_j}}\left[g(2X_i) + g(2X_j)\right]}. \tag{2.9}$$

Similar to the case (2.7), the sample counterpart is

$$\widehat{\bar{\rho}}_g = \frac{(n-1)(U_{3s} - U_{1s})}{U_{1s} + nU_{2s} + (n-1)U_{3s}},$$

where U_{kij} is the U statistic U_k as defined in model (2.7), and $U_{ks} = \sum\limits_{i<j} U_{kij}$, the sum over all unique pairs. King & Chinchilli (2001a) showed $\widehat{\bar{\rho}}_g$ has a normal distribution asymptotically with mean ρ_g and a variance that can be consistently estimated with

$$\mathrm{var}(\widehat{\bar{\rho}}_g) = (\widehat{\bar{\rho}}_g)\left[\frac{\mathrm{var}(\hat{H}_s)}{\hat{H}_s^2} - \frac{2\mathrm{cov}(\hat{H}_s, \hat{G}_s)}{\hat{H}_s\hat{G}_s} + \frac{\mathrm{var}(\hat{G}_s)}{\hat{G}_s^2}\right], \tag{2.10}$$

where $H_s = (n-1)(U_{3s} - U_{1s})$ and $G_s = U_{1s} + nU_{2s} + (n-1)U_{3s}$.

The extended concordance correlation coefficient is applicable to both categorical data and continuous data, and statistical testing is possible based on its asymptotic normality and variance. The model is flexible for both categorical and continuous data, but it is difficult to take into account other meaningful covariate.

2.1.2.2 Overall concordance correlation coefficient

Barnhart et al. (2002) proposed the overall CCC (OCCC) for measure among k methods with each method measuring each of the subject once. Suppose that J methods measure N subjects with a continuous scale Y, a $J \times N$ matrix.

$$\begin{aligned} \rho_o^c &= 1 - \frac{E\left[\sum_{j=1}^{J-1}\sum_{k=j+1}^{J}(y_j - y_k)^2\right]}{E\left[\sum_{j=1}^{J-1}\sum_{k=j+1}^{J}(y_j - y_k)^2 | y_1, \cdots, y_J, \text{ are uncorrelated}\right]} \\ &= \frac{2\sum_{j=1}^{J-1}\sum_{k=j+1}^{J}\sigma_{jk}}{(J-1)\sum_{j=1}^{J}\sigma_j^2 + \sum_{j=1}^{J-1}\sum_{k=j+1}^{J}(\mu_j + \mu_k)^2}, \end{aligned} \tag{2.11}$$

where $\xi_{jk} = [E(Y_j - Y_k)^2 | y_1, \cdots, y_J, \text{ are uncorrelated}]$, $\rho_{ij} = \rho$ for all j and k, and $\chi_{ij} = \frac{2\sigma_j\sigma_k}{\{\sigma_j^2 + \sigma_k^2 + (\mu_j - \mu_k)^2\}}$, index of accuracy. ρ_o^c is interpreted as a weighted average of all pairwise CCC; a pair of measurements with higher variances and larger mean difference is assigned high weight.

The sample counterpart is defined as

$$\hat{\rho}_o^c = \frac{2\sum_{j=1}^{J-1}\sum_{k=j+1}^{J}S_{jk}}{(J-1)\sum_{j=1}^{J}S_j^2 + J\sum_{j=1}^{J}(Y_{\cdot j} - Y_{\cdot\cdot})^2},$$

where $Y_{\cdot j}$, S_j and S_{jk} are sample mean, variance and covariance, respectively. Barnhart et al. (2002) proposed a generalized estimating equation (GEE) approach to estimate OCCC and their variances. The GEE of the equations are

$$\sum_{i}^{n} D_i' V_i^{-1}(Y_i - \mu_i(\mu)) = 0,$$
$$\sum_{i}^{n} F_i' H_i^{-1}(Y_i^2 - \delta_i^2(\sigma^2, \mu)) = 0,$$
$$\sum_{i}^{n} C_i' W_i^{-1}(U_i^2 - \theta(\alpha, \sigma^2, \mu)) = 0,$$

where $D_i = \partial\mu_i/\partial\mu$, $F_i = \partial\delta_i^2/\partial\sigma^2$, $C_i = \partial\theta/\partial\alpha$, $U_i = y_{i1}y_{i2}$ and V_i, H_i, W_i are working covariance matrices. α is the z transformation of the OCCC. The use of GEE allows to assess the agreement without specifying the joint distribution of the data, and also makes the covariates adjustment.

2.1.2.3 Intra class coefficient

An intra-class coefficient (ICC) is defined under an ANOVA model. Because ICC is a covariance-based index, ICC relies on the ANOVA model. Fleiss (1986) showed several kinds of ICC according to experiment designs. Suppose that each of n subjects is measured m times by k methods. Fleiss (1986) considered the two-way ANOVA model:

$$y_{ijl} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \tag{2.12}$$

where $i = 1, \cdots, N$ and $j = 1, \cdots, J$ where α_i are independent and identically distributed with mean zero and variance σ_a^2, β_js are fixed with $\sum_{j=1}^{J}\beta_j = 0$, ε_{ij}s are independent and identically distributed with mean zero and variance σ_ε^2, and α_i and ε_j are mutually independent.

In terms of the analysis of variance, we have the total variation, which is the variation of all n subjects and k categories regardless whom they belong. Assume that all the effects in the model are either between effects or within effects. Then the ICC is defined as the ratio of between subject variance and total, within and between variance. Subject effects are considered as random. Method effects could be considered as either fixed or random. For fixed method effect ICC is expressed as

$$\rho_{ICC_1} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}, \tag{2.13}$$

and for random method effect, ICC is defined as

$$\rho_{ICC_2} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\varepsilon^2}. \tag{2.14}$$

Shrout & Fleiss (1979) argued that the index (2.13) is a measure of association and the index (2.14) is the degree of the agreement between methods. Carrasco & Jover (2003) also addressed the use of the between-method variation even if it is fixed effect. In the case that method is a fixed effect, σ_β^2 is defined as

$$\sigma_\beta^2 = \frac{1}{J(J-1)} \sum_{j=1}^{J-1} \sum_{j'=j+1}^{J} (\beta_j - \beta_{j'})^2.$$

While CCC and ICC are similar indices, the main difference is that ANOVA assumptions are required for ICC but not for CCC. However, if there are no replications, Carrasco & Jover (2003) demonstrated that ρ_{ICC_2} is the same as CCC without the ANOVA model assumption. In general, if the ANOVA model assumptions are correct, the CCC under the model reduces to the ICC defined by the ANOVA model.

For estimates of inference, Carrasco & Jover (2003) suggested using ML or REML to obtain the estimates of variance components and then using delta method to estimate the variance of ICC.

$$\text{var}(\widehat{\rho}_{ICC}) = \frac{(1-\rho_{ICC})^2 var(\sigma_\alpha^2) + \rho_{ICC}^2\{var(\sigma_\beta^2) + var(\sigma_\varepsilon^2) + 2cov(\sigma_\varepsilon^2, \sigma_\beta^2)\}}{(\sigma_\alpha^2 + \sigma)^2}$$
$$\frac{-2(1-\rho_{ICC})\rho_{ICC}\{cov(\sigma_\alpha^2, \sigma_\beta^2) + cov(\sigma_\alpha^2, \sigma_\varepsilon^2)\}}{(\sigma_\alpha^2 + \sigma)^2}. \tag{2.15}$$

Estimating the CCC with variance components via a mixed effects model allows the CCC to be easily extended for more than two methods and to be adjusted by potentially confounding effects, in which case other inter-subject sources of variability should be removed in order to improve detection of the estimated between-subject variance. Carrasco & Jover (2003) exploited variance component model to improve accuracy of CCC allowing more than two methods. Carrasco & Jover (2003) showed that OCCC yielded almost identical results as GCCC through simulation study, and indicated that OCCC can incorporate covariates easily but it is not obvious how one can incorporate covariates in the case of GCCC. This model can accommodate $k > 2$ methods, but the assumption of equal variance for all the methods is generally not reasonable.

2.1.3 Repeated measures data - two methods

Chinchilli et al. (1996) proposed weighted CCC, and King (2007) proposed repeated measures CCC, for data with repeated measures. The repeated measures are not considered replications because it is natural that the measurements from the same subject are more similar than the measurements across subjects, and the variability relies on the range of covariate and/or the time change; they may not be independently and identically distributed, given subject or sample. Suppose that for each method, there are q repeated measures for a subject, rather than K replications. There are then a total of q^2 pairs of measurements between the two methods. Chinchilli et al. (1996) constructed the CCC as a weighted average of q CCCs where the q CCCs are defined as the CCCs between two methods based on q fitted new variables obtained from a random coefficient model. For the i_{th} subject, $i = 1, \cdots, n$ as an average of q coefficients; that is,

$$\rho_{c,i} = \frac{1}{q}\sum_{j}^{q} \frac{2(\Delta_{XYi})_{jj}}{(\Delta_{XXi})_{jj} + (\Delta_{YYi})_{jj} + \{(\mu_{X_i} - \mu_{Y_i})(\mu_{X_i} - \mu_{Y_i})'\}_{jj}},$$

where the $q \times q$ matrices Δ_{XYi}, Δ_{YYi} are variance covariance matrices of transformed random vectors of two variables X and Y with repeated measurements. Variance covariance matrices are constructed from within-subject variances that are estimated uniquely for each subject, as well as between-subject variance covariance parameters that are constructed using the entire sample. For each subject, the CCC is estimated from these unbiased estimators, and a weighted average of the coefficients is formed

$$\rho_c = \left(\sum_{i}^{n} w_i\right)^{-1} \left(\sum_{i}^{n} w_i \rho_{c,i}\right),$$

using a weight function based on the amount of within-subject variability for each subject

$$w_i = 1/\left[(1 + \sigma_{XX_i})(1 + \sigma_{YY_i}) - \sigma^2_{XYi}\right],$$

where σ_{XX_i} and σ_{YY_i} are the variance of random errors subject to variables X_i and Y_i, respectively.

It performs well with adequate numbers of repeated measures on each subject. The model-based transformation of observations is most useful if the observers take different numbers of repeated measures over time or across subjects. Note, because the w_i is the inverse of within-subject variabilities, if the within-subject variabilities are the same for all subjects, then the weighted CCC equals the straight average of the q CCCs by Lin (1989) between two methods. When the number of readings within each subject is small, the estimates of within-subject variance may be unstable and skew the overall results because within-subject variance is estimated uniquely for respective subjects.

King et al. (2007) remedied WCCC's shortcoming by using an unstructured variance-covariance structure of the repeated measurements and estimates of population moments, rather than subject-specific moments. Let $Y_{i1} = (Y_{i11}, \cdots, Y_{i1p})'$ and $Y_{i2} = (Y_{i21}, \cdots, Y_{i2p})'$ be the observations made by two observers, respectively. King et al. (2007) defined repeated measures CCC by using a distance matrix D as

$$\begin{aligned}\rho_{c,rm} &= 1 - \frac{E\left[(Y_{i1} - Y_{i2})' D (Y_{i1} - Y_{i2})\right]}{E_I\left[(Y_{i1} - Y_{i2})' D (Y_{i1} - Y_{i2})\right]} \\ &= \frac{\sum_j^p \sum_k^p d_{jk}(\sigma_{X_j Y_k} + \sigma_{Y_j X_k})}{\sum_j^p \sum_k^p d_{jk}(\sigma_{X_j Y_k} + \sigma_{Y_j X_k}) + \sum_j^p \sum_k^p d_{jk}(\mu_{X_j} - \mu_{Y_j})(\mu_{X_k} - \mu_{Y_k})},\end{aligned}$$

where D is a $p \times p$ distance matrix, which can be thought of as a weight matrix. The repeated measure of CCC is the ratio of the weighted numerator and weighted denominator based on the pairwise CCC rather than the use of the weighted average of the pairwise CCC. The estimator of the above coefficient is based on sample estimates of population moments, as opposed to subject-specific moments.

The asymptotic distribution is derived based on U-statistic

$$\hat{\rho}_{c,rm} = 1 - \frac{nU}{U + (n-1)V} = \frac{(n-1)(VU)}{U + (n-1)V},$$

where nU and $U + (n-1)V$ are unbiased estimators of $E\left[(X_Y)'D(X_Y)\right]$ and $E_I\left[(X_Y)'D(X_Y)\right]$ respectively. The $\hat{\rho}_{c,rm}$ has a normal distribution asymptotically with mean $\rho_{c,rm}$ and variance $var(\hat{\rho}_{c,rm}) = d\Sigma d'$, where d is 2 size vector with a partial differential of g with respect to U and V, and $\Sigma = var(U)$ is the 2×2 variance covariance matrix of the U statistics.

2.1.4 Repeated measures data - multiple methods

Barnhart et al. (2005) extended model (2.11) to multiple methods with multiple readings. The remarkable achievement is to distinguish the source of disagreement into within-method, between-method, and total-method disagreement. Suppose that there are N randomly selected subjects and J fixed methods, where j_{th} method K_{ij} replicated measurements for the i_{th} subject. Let $E_k(Y_{ij}|ij) = \mu_{ij},\ V(Y_{ijk}|ij) = \sigma^2_{ij},\ E_i(\mu_{ij}|j) = \mu_{*j}, V_i(\mu_{ij}|j) = \delta^2_j,\ \text{Corr}_i(\mu_{ij}, \mu_{ij'}|jj') = \rho^{\mu}_{jj'}$. For the k_{th} replicated reading on subject i by method j, Barnhart et al. (2005) modeled

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk},$$

where ε_{ijk} is random error with $E_k(\varepsilon_{ij}|ij) = 0,\ V_k(\varepsilon_{ijk}|ij) = \sigma^2_{ij},\ E_i(\sigma_{ij}|j) = \sigma_{*j}$, and μ_{ij} are mutually independent from ε_{ijk}. The proposed intra-, inter-, and total-indexes are defined as

$$\begin{aligned}
\rho_{j,\text{intra}} &= \frac{\delta^2_j}{\sigma^2_j + \sigma^2_{*j}}, \qquad (2.16)\\
\rho_{\text{inter}} &= 1 - \frac{E\left(\sum_{j=1}^{J}(\mu_{ij} - \mu_{i.})^2/(j-1)\right)}{E\left(\sum_{j=1}^{J}(\mu_{ij} - \mu_{i.})^2/(j-1)|\mu_j\text{'s are independent}\right)}
\end{aligned}$$

$$= \frac{\sum_{j=1}^{J-1}\sum_{j'>j}\rho^{\mu}_{jj'}\delta_j\delta_{j'}}{(j-1)\sum_{j=1}^{J}\delta_j^2 + \sum_{j=1}^{J-1}\sum_{j'>j}(\mu_{*j}-\mu_{*j'})^2}, \tag{2.17}$$

$$\begin{aligned}\rho_{\text{total}} &= 1 - \frac{E\left(\sum_{j=1}^{J}(Y_{ijo}-Y_{i.o})^2/(j-1)\right)}{E\left(\sum_{j=1}^{J}(Y_{ijo}-Y_{i.o})^2/(j-1)|Y_{i1o},\cdots,Y_{iJo}\text{ are independent}\right)} \\ &= \frac{\sum_{j=1}^{J-1}\sum_{j'>j}\rho^{\mu}_{jj'}\delta_j\delta_{j'}}{(j-1)\sum_{j=1}^{J}\delta_j^2 + \sum_{j=1}^{J-1}\sum_{j'>j}(\mu_{*j}-\mu_{*j'})^2}, \end{aligned} \tag{2.18}$$

where Y_{ijo} is a randomly selected reading from the K_{ij} readings of i_{th} subject and j_{th} method. Total disagreement (2.18) is caused both by random error within a method and true differences attributed by the different methods. Intra-method agreement (2.16) measures consistency of readings taken by the same method and inter-method agreement (2.17) measures consistency of true readings attributed by J methods. Therefore, if there is disagreement between methods, then the discrepancy source is fractionated whether it originates from random error within a particular method, or from true differences attributed by the different methods. There are four vectors of parameters $\mu_* = (\mu_{*1},\cdots,\mu_{*J})^T$, $\delta^2 = (\delta_1^2,\cdots,\delta_J^2)^T$, $\sigma_*^2 = (\sigma_{*1}^2,\cdots,\sigma_{*J}^2)^T$, and $\rho^{\mu} = \left(\rho^{\mu}_{12},\cdots,\rho^{\mu}_{(J-1)J}\right)$, and these are estimated via four generalized estimation equations:

$$\sum_{i=1}^{N} D_{i1}^T V_{i1}^{-1}(Y_{i\cdot}-\mu_*) = 0, \tag{2.19}$$

$$\sum_{i=1}^{N} D_{i2}^{T} V_{i2}^{-1} \left(U_i - \sigma_*^2\right) = 0. \tag{2.20}$$

Estimate μ_* by GEE equation (2.19) where $D_{i1} = \partial\mu_*/\partial\mu_* = I_J$, V_{i1} is the working covariance matrix of $Y_{i\cdot}$. Through GEE equation (2.20), σ_* is estimated where $D_{i2} = \partial\sigma_*^2/\partial\sigma_*^2 = I_J$,, V_{i2} is the working covariance matrix of $U_i = \left(\frac{\sum_{k=1}^{K_{ij}}(Y_{i1k}-Y_{i1\cdot})^2}{K_{i1}-1}, \cdots, \frac{\sum_{k=1}^{K_{ij}}(Y_{iJk}-Y_{iJ\cdot})^2}{K_{iJ}-1}\right)^T$.

$$\sum_{i=1}^{N} D_{i3}^{T} V_{i3}^{-1} \left(W_i - g(\mu_*, \delta^2)\right) = 0, \tag{2.21}$$

$$\sum_{i=1}^{N} D_{i4}^{T} V_{i4}^{-1} \left(Z_i - h(\alpha, \mu_*, \sigma_*^2, \delta^2)\right) = 0. \tag{2.22}$$

Based on GEE equation (2.21), δ^2 is estimated where $D_{i3} = \partial g(\delta^2, \sigma_*^2, \mu_*)/\partial\delta_*^2 = I_J$, V_{i3} is the working covariance matrix of $W_i = \left(\frac{(Y_{i1\cdot}-U_{i1})^2}{K_{i1}}, \cdots, \frac{(Y_{iJ\cdot}-U_{iJ})^2}{K_{iJ}}\right)^T$, and $g(\mu_*, \delta^2) = E(W_i) = \mu_*^2 + \delta^2$. ρ^μ is estimated by the cross products $Y_{ij\cdot}Y_{ij'\cdot}I(j' > j)$ by $E(Y_{ij\cdot}Y_{ij'\cdot}) = \mu_{*j}\mu_{*j'} + \rho_{jj'}^{\mu}\delta_j\delta_{j'}$, where $D_{i4} = h(\alpha, \mu_*, \sigma_*^2, \delta^2)/\partial\alpha$, $h(\alpha, \mu_*, \sigma_*^2, \delta^2)$ $= E(Z_i)$ and V_{i4} is the working covariance matrix for Z_i in equation (2.22). The GEE method provides an asymptotic distribution and variance without specifying the joint distribution of the data and its implementation is easy. Furthermore, this index provides direction on how to improve agreement by identifying the source of disagreement, either intra-method variability or inter-method variability or both. This model only can evaluate agreement for continuous data.

Lin et al. (2007) integrated the approaches by Barnhart et al. (2005) and Carrasco & Jover (2003), and proposed unified CCC (UCCC) to quantify the degree of the agreement among k methods measuring multiple readings for every subject for continuous and categorical data. That is, the proposed UCCC is generalized CCC proposed by Lin (1989) in the mixed effects model point of view. Three indexes, inter-, intra- and total-, are proposed; these are denoted by the product of meaningful precision and accuracy. All proposed indices are expressed as functions of variance components in order to give stability of estimators, and the GEE method is used to obtain the estimates and perform inferences.

Suppose each of k methods measures m readings on every of n subjects. The model we use for measuring agreement is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}, \qquad i = 1, ..., n, j = 1, ..., k, l = 1, ..., m, \tag{2.23}$$

where μ is the overall mean. α_i and γ_{ij} are the random subject effects, which are independent and identically distributed with zero mean and variances σ_α^2 and σ_γ^2 respectively. β_j is the method effect, which is a fixed effect and assumed $\sum_{j=1}^{k} \beta_j = 0$. e_{ijk} is a random error, zero mean and σ_e^2 variance. Let $\sigma_\beta^2 = \frac{\sum \left(\beta_j - \bar{\beta}.\right)^2}{k-1}$, then consider it as the variance among all methods. Even the method effect is considered as a fixed effect, σ_β^2 has to be taken into consideration following the argument of Shrout & Fleiss (1979).

CCC$_{intra,j}$ measures for overall k methods, and it measures how well each method reproduces itself. Since $CCC_{intra,j}$ is not a function of j, we can simply the notion to CCC$_{intra}$. For a given method i,

$$\begin{aligned} CCC_{intra} &= 1 - \frac{E\left[\frac{\sum_{l=1}^{m}\left\{Y_{ijl} - \overline{Y}_{ij.}\right\}^2}{m-1}\right]}{E\left[\frac{\sum_{l=1}^{m}\left\{Y_{ijl} - \overline{Y}_{ij.}\right\}^2}{m-1} | Y_{ij1}, \cdots, Y_{ijm} \text{indep}\right]} \\ &= \frac{\sigma_\alpha^2 + \sigma_\gamma^2}{\sigma_\alpha^2 + \sigma_\gamma^2 + \sigma_e^2}. \end{aligned} \tag{2.24}$$

The expectation is that all readings from the same method are the same, so $CCC_{\text{intra},j}$ is the same as ρ and intra-method accuracy is 1.

Inter-method agreement is the index between any two methods say j and j'. It is considered as fixed after averaging m replications of each subject in a method. It quantifies the inter-method agreement based on the average of the m readings from each of the (i, j), and measures how close the means of respective methods are. Since $CCC_{inter,jj'}$ is not a function of j or j', we simplify to CCC_{inter}.

$$\begin{aligned} CCC_{inter} &= 1 - \frac{E\left[\frac{\sum_{l=1}^{k}\left\{\overline{Y}_{ij.}-\overline{Y}_{ij'.}\right\}^2}{k-1}\right]}{E\left[\frac{\sum_{l=1}^{k}\left\{\overline{Y}_{ij.}-\overline{Y}_{ij'.}\right\}^2}{k-1}|\overline{Y}_{i1.},\cdots,\overline{Y}_{ik.}\text{indep}\right]} \\ &= \frac{\sigma_\alpha^2}{\sigma_\alpha^2+\sigma_\gamma^2+\sigma_\beta^2+\sigma_e^2/m}. \end{aligned} \quad (2.25)$$

ρ_{inter} is $\frac{\sigma_\alpha^2}{\sigma_\alpha^2+\sigma_\gamma^2+\sigma_e^2/m}$, which is an precision index for overall k methods based on the average of the m readings. χ_{inter} is $\frac{\sigma_\alpha^2+\sigma_\gamma^2+\sigma_e^2/m}{\sigma_\alpha^2+\sigma_\gamma^2+\sigma_e^2/m+\frac{\sum_{j=1}^{k-1}\sum_{j'=j+1}^{k}(\mu_j-\mu_{j'})^2}{k(k-1)}}$, which is an index of accuracy depending on the difference among means because Lin et al. (2007) assumed that variances are the same for all methods.

Total agreement $CCC_{total,jj'}$ measures the inter-method agreement between methods j and j' based on any one of the m readings from each of the (i, j) combinations, and CCC_{total} shows how well all methods reproduce each other.

$$CCC_{total} = 1 - \frac{E\left[\frac{\sum_{l=1}^{k}\left\{Y_{ijo}-\overline{Y}_{i.o}\right\}^2}{k-1}\right]}{E\left[\frac{\sum_{l=1}^{k}\left\{Y_{ijo}-\overline{Y}_{i.o}\right\}^2}{K-1}|Y_{i1o},\cdots,Y_{iko}\text{indep}\right]}$$

$$= \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\gamma^2 + \sigma_\beta^2 + \sigma_e^2}. \tag{2.26}$$

Since $CCC_{total,jj'}$ is not a function of j and j', we simplify to CCC_{total}. This index does not depend on the number of replications because it is based on any one of the m readings of each method whereas CCC_{inter} has m in denominator.

The advantage of this approach is that the intra-subject correlation is (1) induced by a random (subject) effect without assuming correlation structure, (2) easily extended for more than two observers, and (3) can be adjusted by potentially confounding subject-covariate. All proposed indices are expressed as functions of variance components, and GEE is used to obtain the estimates and perform inferences. The model assumes, however, stable variations as well as time-invariant overall mean over the study duration. However, these underlying assumptions restrict the utility of the UCCC for ILD characterized by a smooth curve trajectory of response values and a time-varying variance.

2.1.5 Functional data - two methods

Li & Chow (2005) introduced the concept of functional data in reproducibility, and proposed functional-type CCC for an overall index of agreement of paired repeated measures through the duration of the study. Let $X(\cdot)$ and $Y(\cdot)$ be stochastic process as in $t \in \mathfrak{T},\ j = 1, \cdots, N$, where $\mathfrak{T}$ is a finite closed real interval. That is, given time t_0, $X(t_0)$ and $Y(t_0)$ are random variables in some probability space, and $X_i(\cdot)$ and $Y_i(\cdot),\ i = 1, \cdots, n$ are functional data about t. For the probability functional space $\mathfrak{T}$, an inner product is defined as

$$< X(\cdot), Y(\cdot) >= E \int_{\mathfrak{T}} X(t)Y(t)w(t)dt,$$

where $w(\cdot)$ is a weight function and takes non-negative values over $\mathfrak{T}$. The norm $||\cdot||$ is given as

$$||X|| = \sqrt{< X, X >}.$$

A functional concordance coefficient for $X(\cdot)$ and $Y(\cdot)$ is defined as follows:

$$\rho_c = \frac{2 < X - E(X), Y - E(Y) >}{||\,(E(X) - E(Y))\,||^2 + ||\,(X - E(X))\,||^2 + ||\,(Y - E(Y))\,||^2}.$$

The estimated ρ_f, then, is

$$\hat{\rho}_c(X,Y) = \frac{\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\{X_i(t_j) - \bar{X}(t_j)\}\{Y_i(t_j) - \bar{Y}(t_j)\}w(t_j)\triangle_j}{\sum_{j=1}^{N}\{\bar{X}(t_j) - \bar{Y}(t_j)\}^2 w(t_j)\triangle_j + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\left[\{X_i(t_j) - \bar{X}(t_j)\}^2 + \{Y_i(t_j) - \bar{Y}(t_j)\}^2\right]w(t_j)\triangle_j},$$

where $\triangle_j$ is the gap size between t_{j+1} and t_j, and $w(t)$ is a weighted function with non-negative values over $\mathfrak{T}$; $\bar{X}(t_j) = n^{-1}\sum_{i=1}^{n} X_i(t_j)$ and $\bar{Y}(t_j) = n^{-1}\sum_{i=1}^{n} Y_i(t_j)$ are the sample mean of $X(t_j)$ and $Y(t_j)$, respectively. The asymptotic normality of $\hat{\rho}_c$ is established. Li & Chow (2005) approached the measurement agreements problem from the stochastic perspective. Two allows the researcher to take functional data into consideration in proper way. This model is more adequate than a multi-dimensional vector if measurements are complicated curves over time.

2.2 Local Regression

Nonparametric regression provides a powerful tool to explore data with unknown structure for this functional type of data. There are many specific methods of nonparametric smoothing, such as kernel smoothing, local polynomial smoothing, spline smoothing, and wavelet-based methods. Most of them assume certain smoothness of the regression function. In this section, we will briefly review the kernel smoothing and local polynomial smoothing techniques, which are frequently used in our research.

Suppose the bivariate sample $\{(x_i, y_i), i = 1, \ldots, n\}$ is collected from the model:

$$y_i = m(x_i) + \varepsilon_i,$$

where $m(\cdot)$ is an unknown function ε is an error term, representing random errors with $E(\varepsilon|X) = 0$ and $var(\varepsilon|X = x) = \sigma^2(x)$. The mean function $m(\cdot)$ is the object to be estimated in a nonparametric regression problem. The shape of $m(\cdot)$ describes the underlying relationship between the response variable Y and the predictor variable X. Usually a point closer to x has more information about the value of $m(x)$, so a natural idea for estimating $m(x)$ is to use the running local average. Rather than global assumptions about the function, the function can be locally well approximated with a member of a simple class of parametric function, such as a constant or straight line.

2.2.1 Local constant

Kernel smoothing provides a simple way of finding structure in data without imposing a parametric model: use only those observations close to the target point x_0 to fit the simple model. Localization is achieved via a weighing function or kernel $K_h(x_0, x_i)$, which assigns a weight to x_i based on its distance from x_0, so the locally weighted average estimator defined as

$$\hat{m}(x_0) = \frac{\sum_i^n K_h(X_i - x_0)Y_i}{\sum_i^n K_h(X_i - x_0)},$$

where h is the bandwidth, and $k_h(t) = k(t/h)/h$. This estimator is often called the Nadaraya-Watson (NW) kernel regression estimator. By taking the kernel function to be the uniform kernel $K(u) = I(|u| < 1/2)$ the NW estimator becomes the running local average, which is similar to the K nearest neighbor estimator. The performance of the NW estimator can be summarized

$$E(\hat{m}(x)|X_1, \cdots, X_n) - m(x) = \mu(K)\left\{\frac{m''(x)}{2} + \frac{m'(x)f'(x)}{f(x)}\right\} h^2 + o(h^2),$$

and

$$var(\hat{m}(x)|X_1, \cdots, X_n) = R(K)\frac{\sigma^2}{f(x)nh} + o\left((nh)^{-1}\right),$$

where f is the marginal density of X, and x is an interior point of the support of f.

The kernel function K is usually chosen to be a unimodal probability density function that is symmetric about zero. Sometimes kernels that are not densities are utilized. It is interesting that the choice of the shape of the kernel function is not a particularly important issue (Fan & Gijbels (1996)). However, the choice of the smoothing parameter (bandwidth) is critical. It considerably influences the performance of the estimator. If h is chosen to be too small, then we will overfit the data and get an undersmoothed estimate. This estimate pays too much attention to the data in the local neighborhood. On the contrary, if h is chosen to be too large, then we will underfit the data and get an oversmoothed estimate. This estimate will miss some fine features of the data. Thus, bandwidth h effectively controls the complexity of the estimate. The bandwidth can be subjectively chosen by trial-error approach via visually inspecting estimates in graph, or chosen by data via minimizing a theoretical risk. One choice is the widely used mean squared error (MSE) criterion, the sum of variance and the squared bias. A good choice of bandwidth h should give the estimator an optimal balance between variance and bias. Applying the MSE criterion to the kernel regression estimator, MSE$\{\hat{m}(x)\}$ depends on x. The assessment of the performance of $m(x)$ is achieved at a fixed point x. Therefore, it will be more appropriate to analyze the integration of MSE$\{\hat{m}(x)\}$ over the entire real line. Such a criterion is called the mean integrated squared error (MISE): The optimal choice of bandwidth h is the one that minimizes the MISE. The MISE is

$$\text{MISE} = \frac{\mu(K)^2 h^4}{4} \int \left\{ m''(x) \right\}^2 w(x)dx + \frac{R(K)}{nh} \int \frac{\sigma^2(x)}{f(x)} w(x)dx, \tag{2.27}$$

where w is a nonnegative weight function. The optimal bandwidth of a local linear smoother via minimizing the MISE is

$$h_{\text{opt}} = \left[\frac{R(K) \int \sigma^2(x) f^{-1}(x) w(x) dx}{\mu^2(K) \int \{m(x)''(x)\}^2 w(x) dx} \right]^{1/5} n^{-1/5}. \tag{2.28}$$

Because the exact MSE or MISE depends on the bandwidth in a complicated way, we may calculate the asymptotic MSE or MISE instead and get the asymptotic MISE-optimal bandwidth h. Another classical approach is cross-validation. The basic idea is to prevent overfitting induced by data reuse on estimating coefficients and measuring residuals; it holds out part of the sample with which to evaluate the performance of a predictor. It is conceptually simple, but needs intensive computation. In small sample-size situation, an efficient method is leave-one-out: given i, use data except i to build a regression function and validate the model by examining difference between the response value of i and the interpolated value from the fitted model.

$$\mathrm{CV}(h) = n^{-1}\sum_{i=1}^{n}\left\{Y_i - \hat{m}_{h,(-i)}(X_i)\right\}^2. \tag{2.29}$$

The $\hat{h}_{\mathrm{CV}}$ is selected by minimizing the $\mathrm{CV}(h)$. It is common to plot the function $CV(h)$ against h to examine the chage of overall prediction errors. The visualization of the function $CV(h)$ helps in the process of the bandwidth selection.

2.2.2 Local polynomial

The kernel estimator is a local constant fit, which means that the mean function $m(\cdot)$ is locally approximated. This approach can be naturally extended to local polynomial fit, in which the mean function $m(\cdot)$ is locally approximated by a polynomial rather than a constant.

Suppose $m(x)$ is smooth function so that the $(p+1)^{\mathrm{th}}$ derivative at an arbitrary fixed point x_0 exists, and we approximate $m(\cdot)$ locally by a polynomial of order p via the Taylor expansion.

$$m(x) \approx m(x_0) + m^{(1)}(x_0)(x - x_0) + \frac{m^{(2)}(x_0)}{2}(x - x_0)^2 + \cdots + \frac{m^{(p)}(x_0)}{p!}(x - x_0)^p,$$

where $m^{(r)}(x_0)$ denotes the r_{th} derivative of $m(x)$ at x_o. At a point x_0, the local polynomial is obtained by fitting the p_{th}-degree polynomial model

$$E(y|x) = \beta_0 + \beta_1(x - x_0) + \beta_2(x - x_0)^2 + \cdots + \beta_p(x - x_0)^p.$$

Set $\beta_r = m^{(r)}(t_0)/r!,\ r = 0, \cdots, p$. After parametrization of $m(x)$, parametric approaches can be applied to estimate $m(x)$. Given error distribution, score function with kernel weight is one way to estimate regression coefficient. Without the distribution assumption, we can estimate the regression function via optimizing least squares. In the local polynomial regression, $\hat{\beta}_r,\ r = 0, \cdots, p$ is the estimator minimized locally via the weighted least squares function.

$$\min_{\beta} \sum_{i=1}^{n} \left(y_i - \sum_{j=0}^{p} \beta_j (X_i - x_0)^j \right)^2 K_h(X_i - x_0). \tag{2.30}$$

We minimize equation (2.30) with respect to $\beta = (\beta_0, \cdots, \beta_p)$ to obtain the local polynomial estimate $\hat{\beta}$

$$\hat{\beta} = (X^T W X)^{-1} X^T W y, \tag{2.31}$$

where W be the $n \times n$ diagonal matrix of weights:

$$\operatorname{diag}\left(K_h(t_1 - t_0), \cdots, K_h(t_n - t_0)\right),$$

and

$$X = \begin{pmatrix} 1 & (x_1 - x_0) & \cdots & (x_1 - x_0)^p \\ \cdots & \vdots & \ddots & \vdots \\ 1 & (x_n - x_0) & \cdots & (x_n - x_0)^p \end{pmatrix}, \quad y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_p \end{pmatrix}$$

Thus,

$$\hat{m}^{(r)}(t_0) = r! \beta_r, \quad r = 0, \cdots, p,$$

which is an estimator for $m^{(r)}(x_0)$, the r_{th} derivative of $m(x)$ evaluated at point x_0.

By using local polynomial regression, we can estimate not only the mean function $m(\cdot)$, but also the first p_{th} derivatives of $m(\cdot)$. This is an advantage of local polynomial regression compared to the NW estimators. Local polynomial estimators also have better asymptotic properties than NW.

Fan & Gijbels (1996) showed that polynomials with higher order reduce the bias by introducing more local parameters, while extra parameters produce variability by moving from an odd-order approximation to its consecutive even order approximation; the local linear for estimating the regression function outperforms the local constant (kernel), and the local cubic regression for estimating the regression function has better fit than the local quadratic fit. Thus, the odd p is superior to even in the case of estimating regression function. However, there is no direct answer to compare the local linear fit with the local cubic fit because of bias and variance trade off; a lower order polynomial results in a larger bias, but has a smaller variance. Local linear regression is easy to implement and requires weaker assumptions than local cubic regression. Furthermore, Fan & Gijbels (1992) showed that the local linear regression estimator automatically adjusts bias at the edge, and the leading term for the conditional bias does not involve the derivative of f. Fan (1992) calls an estimator with this property design-adaptive. The form in (2.32) is obviously linear in Y so that large sample normality for class of linear smoothers is immediate for a given bandwidth. The asymptotic minimax property of the local linear estimator is also demonstrated by Fan (1993). That is, the local linear estimator is superior to all linear smoothing estimators in a certain sense. Thus, we focus on the properties of local linear regression smoother from now on. The estimator is

$$\hat{m}(x_0) = \sum_{i=1}^{n} \frac{K_h(X_i - x)\{\mathcal{S}_{n,2} - (X_i - x)\mathcal{S}_{n,1}\}}{\mathcal{S}_{n,0}\mathcal{S}_{n,2} - \mathcal{S}_{n,1}^2} Y_i, \tag{2.32}$$

where $\mathcal{S}_{n,j}(x_0) = \sum_{i=1}^{n} K_h(X_i - x_0)(X_i - x)^j,\ j = 1, \cdots, p.$
The asymptotic bias and variance of local linear regression demonstrated by Fan (1992) are

$$\text{Bias}\{\hat{m}(x_0)|X\} = \mu(K)\frac{m''(x)}{2}h^2 + o(h^2). \tag{2.33}$$

$$\text{Var}\{\hat{m}(x_0)|X\} = R(K)\frac{\sigma^2(x)}{f(x)nh} + o\left(\frac{1}{nh}\right), \tag{2.34}$$

provided that the bandwidth h tends to zero in such a manner that $nh \to \infty$ where f is the marginal density of X and $\mu(K) = \int t^2 K(t)dt,\ \ R(K) = \int K^2(t)dt.$

Chapter 3

Functional measurement agreement assessment for intensive longitudinal data

3.1 Introduction

Linear mixed models have been widely used in longitudinal data analysis. However, they have limitations when characterizing a curve feature in ILD because they are parametric models that likely introduce model approximation error. Nonparametric mixed effects models are a natural extension to relax the assumptions on parametric forms and allow a flexible trend over time. Wu & Zhang (2002) utilized the combination of a local polynomial smoothing and a linear mixed model to estimate a population and individual curve in a nonparametric mixed effects model. Park & Wu (2006) proposed a local likelihood method along with a backfitting algorithm. Their procedure allows different bandwidths for a population curve and individual derivatives, and its estimates may improve the performance of estimates by Wu & Zhang (2002) with one bandwidth. The nonparametric mixed model is useful to investigate data structure for curve data allowing only one explanatory variable. In the same context of mixed-effects models, Wu & Liang (2004) proposed the random varying-coefficient (RVC) model, an extension of the varying coefficient model including random effects. Wu & Liang (2004) naturally imposed the within-subject correlation and subject-specific feature of longitudinal data into the estimation procedure by introducing a random varying-coefficient model. They estimate time-varying coefficients by a smoothing method after pooling the data from all subjects together, assuming working independence for within-subject in the varying coefficient regression. The model is transformed into parametric linear mixed-effect models in each step of proposed backfitting algorithm. Thus, it is easy to implement using the existing software, but it is restricted to the case in which random effects are nested in fixed effects. Guo (2002) proposed a general model, named a functional mixed model, which allows a different design matrix between population effects and individual derivatives. He applied a smoothing spline to estimate time-varying coefficient curves, and followed mixed model procedures to estimate parameters. However, a smoothing spline is computationally expensive, especially for the functional data measured densely. Compared with the spline method, the kernel method is less intensive because only a part of data is used.

We make several contributions in this chapter.

- We propose a functional concordance correlation coefficient. The proposed index separates the degree of measurement agreement from an overall time trend to enhance the ability to detect the agreement between methods for curve data.

- We utilize the semiparametric variance by Fan et al. (2007) for the variance-covariance of the random process. The semiparametric covariance for the random process increases the efficiency of the bootstrap variance of the random process by reducing the uncertainty while keeping the flexibility for the covariance function.

- To utilize the flexibility of nonparametric coefficients of potential covariates besides time and to take an advantage of the local likelihood approach, we construct a functional mixed model and estimate nonparametric functions with a local linear smoother.

The rest of this chapter is organized as follows. In section 3.2, we introduce the functional mixed model and propose an estimation procedure for functional mixed models in section 3.2.1. We develop FCCC based on the functional mixed model in section 3.3. We conduct a Monte Carlo simulation study to assess the finite sample performance of the proposed procedure in section 3.4. Conclusion is given in section 3.5. The technical conditions and proofs are given in the appendix.

3.2 Functional mixed models

Suppose that we have a sample of n subjects. For the ith subject, the response variable is $y_i(t_{ij})$, $x_{ij} = x_i(t_{ij})$ is the p-vector covariate and $z_{ij} = z_i(t_{ij})$ is the q-vector covariate at time points t_{ij}, $i = 1, \cdots, n$, $j = 1, \cdots, n_i$.
A functional mixed model is

$$y_i(t_{ij}) = x_{ij}^T \beta(t_{ij}) + z_{ij}^T \nu_i(t_{ij}) + \varepsilon_i(t_{ij}), \tag{3.1}$$

where $\beta(t_{ij})$ is a p-vector time-varying population characteristic at time t_{ij}. $\nu_i(t_{ij})$ is a q-vector of the ith subject contributions at time t_{ij}, considering realizations of a mean 0 process with a variance

$$G_i(t) = \mathcal{F}(t)^{1/2} G_i \mathcal{F}^{1/2}(t),$$

where G_i is a constant associated with between-subject and $\mathcal{F}(t)$ =diag$\big[f_1(t), \cdots, f_q(t)\big]$ is an unknown fixed function relying on time. $\varepsilon_i(t_{ij})$ is the error function that can not be explained by either the fixed effect or the random effect of an uncorrelated white noise process with mean zero and a covariance function $R_i(t_{ij}) = E\left(\varepsilon_i(t_{ij})\varepsilon_i^T(t_{ij})\right) = \text{diag}\big[\sigma_\varepsilon^2(t_{i1}), \cdots, \sigma_\varepsilon^2(t_{in_i})\big]$. That is, the errors are independent over subjects and over replications within subject. $\nu_i(t_{ij})$ and $\varepsilon_i(t_{ij})$ are independent. Thus the variance of $y_i(t_{ij})$ is $z_{ij}^T G_i z_{ij} f(t_{ij})+ \sigma_\varepsilon^2(t_{ij})$.

Fan et al. (2007) studied the modeling of the covariance function of the random error process $\varepsilon(t)$ for irregular longitudinal data in order to improve in the estimation of coefficients in the semiparametric regression model. They modeled a nonparametric variance and a semiparametric covariance function, which allows the random error $\varepsilon(t)$ to be nonstationary with time-varing variance $\sigma_\varepsilon^2(t)$. Because we exploit a random process to explain the time-varying heterogeneity, it is reasonable to impose the serial correlation into the random process. Thus, we utilize the semiparametric variance approach for the covariance of the random process by decomposing the source of variability: the variation subjected to subject differences and the variation along with time. For the estimation aspect, the semiparametric covariance function is the same as the fully nonparametric covariance. However, the two approaches yield different bootstrap variance estimates of the random process in section 3.3.1 and the use of semiparametric covariance provides more stable bootstrap variance estimates for ILD; the overall functional form of time of a time-varying variance is well estimated using densely measured observations and it leads to less fluctuation of the time-varying random process in bootstrap samples, especially when the original sample size is small. Thus, the semiparametric covariance for the random process increases the efficiency of the bootstrap variance of the random process by separating the source of variability while keeping the flexibility for the covariance function.

3.2.1 Estimation Procedure

Assume that $t_1, \cdots, t_L$ are distinct time points among the occasion set. For a given time t, time-varying functions are estimated based on a local log-likelihood. Suppose that $\beta(t_{ij})$ and $\nu_i(t_{ij})$ have second continuous derivatives. When t_{ij} lies in the local neighborhood of t for any given t, we locally approximate $\beta_k(t)$ and $\nu_{i,l}(t)$ by a linear function via Taylor expansions as follows:

$$\begin{aligned} \beta_k(t_{ij}) &\approx a_k + b_k(t_{ij} - t), \\ \nu_{i,l}(t_{ij}) &\approx c_{i,l} + d_{i,l}(t_{ij} - t), \end{aligned}$$

for $k = 1, \cdots, p$ and $l = 1, \cdots, q$. Then the model (3.1) is approximated by a local linear method at given time t,

$$y_{ij,t} = x_{ij}^T a_t + x_{ij}^T b_t(t_{ij} - t) + z_{ij}^T c_{i,t} + z_{ij}^T d_{i,t}(t_{ij} - t) + \varepsilon_{ij,t}, \tag{3.2}$$

where $a_t = (a_1, \cdots, a_p)^T$ and $b_t = (b_1, \cdots, b_p)^T$ are fixed effects, and $c_{i,t} = (c_{i,1}, \cdots, c_{i,q})^T$, $d_{i,t} = (d_{i,1}, \cdots, d_{i,q})^T$ are random effects. Here we add the subscript t to emphasize the local parameters depending on t.

Let $X_i = [x_{i1}, \cdots, x_{in_i}]^T$, $Z_i = [z_{i1}, \cdots, z_{in_i}]^T$, $T_{i,t} = \text{diag}(t_{i1} - t, \cdots, t_{in_i,t} - t)$, $Q_{i,t} = [X_i, T_{i,t}X_i]$, $U_{i,t} = [Z_i, T_{i,t}Z_i]$, $\alpha_t = [a_t^T, b_t^T]^T$ and $V_{i,t} = [c_{it}^T, d_{it}^T]^T$. Then $V_{i,t}$

are independent and identically distributed with $2 \times q$ dimensional normal distribution with mean 0 and covariance matrix $O_{i,t}$, where $O_{i,t} = E\left(V_{i,t}V_{i,t}^T\right)$ is $2q \times 2q$ matrix depending on t, $O_{1,t} = \cdots = O_{n,t} = O_t$.

Model (3.2) becomes

$$y_{i,t} = Q_{i,t}\alpha_t + U_{i,t}V_{i,t} + \varepsilon_{i,t}. \tag{3.3}$$

Model (3.3) can be approximately viewed as a linear mixed model for t_i in the neighborhood of any given t. The inference for a, b, c_i, and d_i is based on the likelihood method. To incorporate the local information in equation (3.3), under normality assumptions for c_i, d_i and ε_i, the following localized generalized log-likelihood is utilized to obtain estimates a, b, c_i and d_i at time t, $t = t_1, \cdots, t_L$.

$$\begin{aligned} l(\alpha, O_t; y_i) = \sum_{i=1}^{n} \{ & [y_i - Q_{i,t}\alpha_t - U_{i,t}V_{i,t}]^T K_{ih}^{1/2}(t) R_{i,t}^{-1} K_{ih}^{1/2}(t) \\ & [y_i - Q_{i,t}\alpha_t - U_{i,t}V_{i,t}] + V_{i,t}O_t^{-1}V_{i,t}^T\}, \end{aligned} \tag{3.4}$$

where $K_{ih}(t) = \mathrm{diag}\left[K\left(\frac{t_{i1}-t}{h}\right)/h, \cdots, K\left(\frac{t_{in_i}-t}{h}\right)/h\right]$ and $R_i = \mathrm{diag}\big(\sigma_\varepsilon^2(t_{i1}), \cdots, \sigma_\varepsilon^2(t_{in_i})\big)$. Given O_t and R_i, $\hat{\alpha}_t$ and $\hat{V}_i$ is obtained via minimizing (3.4) with respect to α and V

$$\begin{aligned} \hat{\alpha}_t &= \{\sum_{i=1}^{n} Q_{i,t}^T \Omega_{i,t} Q_{i,t}\}^{-1} \{\sum_{i=1}^{n} Q_{i,t}^T \Omega_{i,t} y_i\}, \\ \hat{V}_{i,t} &= O U_{i,t}^T \Omega_{i,t}(y_i - Q_{i,t}\hat{\alpha}_t), \end{aligned} \tag{3.5}$$

where $\Sigma_{i,t} = R_i + K_{ih}^{1/2} U_{i,t} O U_{i,t}^T K_{ih}^{1/2}$ and $\Omega_{i,t} = K_{ih}^{1/2}(t) \Sigma_{i,t}^{-1} K_{ih}^{1/2}(t)$.

The local linear estimator $\hat{\beta}(t)$ of $\beta(t)$ and $\hat{\nu}_i(t)$ of $\nu_i(t)$ is the part corresponding to $\hat{a}_t$ and $\hat{c}_{i,t}$,

$$\begin{aligned} \hat{a}_t &= e_{1,p;2p}\hat{\alpha}_t, \\ \hat{c}_{i,t} &= e_{1,q;2q}\hat{V}_{i,t}, \end{aligned} \tag{3.6}$$

where $e_{p,q;r}$ is a r vector with 1 at the pth through qth components and 0 otherwise.

Given $\hat{\beta}_t$ and $\hat{\nu}_{i,t}$, we estimate variance components using EM algorithm coupled with the restricted maximum likelihood (REML). The REML estimates of the variance components in the kernel weighted linear mixed model $K_{ih}^{1/2}(t_l - t)y_i = K_{ih}^{1/2}X_i\beta_t + K_{ih}^{1/2}Z_i\nu_{i,t} + \varepsilon_i$ are formed as follows.

$$\hat{\sigma}_{\varepsilon}^2(t) = \frac{1}{N}\sum^{n}\left\{\left[y_i - X_i\hat{\beta}_t - Z_i\hat{\nu}_{i,t}\right]^T K_{ih}\left[y_i - X_i\hat{\beta}_t - Z_i\hat{\nu}_{i,t}\right] + \sigma_{\varepsilon 0,t}^2(t)\left[1 - \sigma_0^2(t)\Lambda_i\right]\right\},$$

and

$$\hat{D}(t) = \frac{1}{n}\sum^{n}\left\{\hat{\nu}_{i,t}\hat{\nu}_{i,t}^T + D_{0,t}\left(I - Z_i^T K_{ih}^{1/2}\Lambda_i K_{ih}^{1/2} Z_i^T D_0\right)\right\},$$

where $D_{0,t}$ and $\sigma_{\varepsilon 0,t}^2$ indicate current estimates of D_t and $\sigma_{\varepsilon,t}^2$, respectively, $\Lambda_i = V^{*-1}\{I_{n_i} - K_{ih}^{1/2}X_i P^{-1} X_i^T K_{ih}^{1/2} V^{*-1}\}$, $P = \sum^n \left\{X_i^T K_{ih}^{1/2} V^{*-1} K_{ih}^{1/2} X_i\right\}$, $V^{*-1} = \sigma_0^2 I_{n_i} + K_{ih}^{1/2} Z_i D_0 Z_i^T K_{ih}^{1/2}$ and $N = \sum^n n_i$.

Replace $D_{0,t}$ and $\sigma_{\varepsilon 0,t}^2$ by $\hat{D}(t)$ and $\hat{\sigma}_{\varepsilon}^2(t)$, respectively, and repeat estimating $\hat{\beta}_t$ and $\hat{\nu}_{i,t}$ until converges.

Fixed and random coefficients a and c_i can be attained by well-developed software R or SAS.

$$K_{ih}^{1/2}(t_l - t)y_i(t_l) = K_{ih}^{1/2}(t_l - t)\left\{a + b(t - t_l) + c_i + d_i(t_l - t)\right\} + \varepsilon_i(t_l).$$

Fan & Zhang (2000) found that the usage of a different bandwidth for each covariate performs better than the usage of the single best bandwidth. In the same context, it may yield better fit to select two different bandwidths for fixed effects and random effects when they have different degrees of smoothness. Denote h_f and h_r to be bandwidths of fixed and random effects, respectively. The local kernel generalized log-likelihood at the neighbourhood of time t can be considered

$$l(\alpha, O_t; y_i) = \sum_{i=1}^{n} \{ [y_i - Q_{i,t}\alpha_t - \boldsymbol{K_{ih}} U_{i,t}V_{i,t}]^T K_{ih_f}^{1/2}(t) R_{i,t}^{-1} K_{ih_f}^{1/2}(t) \\ [y_i - Q_{i,t}\alpha_t - \boldsymbol{K_{ih}} U_{i,t}V_{i,t}] + V_{i,t}O_t^{-1}V_{i,t}^T\}, \tag{3.7}$$

where $K_{ih} = K_{ih_r}/K_{ih_f} = \text{diag}\,[h_{i,r}(t_{i1})/h_{i,f}(t_{i1}), \cdots, h_{i,r}(t_{in_i})/h_{i,f}(t_{in_i})]$. Then

$$\begin{aligned} \hat{\alpha}_t &= \{\sum_{i=1}^{n} Q_{i,t}^T \Omega_{i,t} Q_{i,t}\}^{-1} \{\sum_{i=1}^{n} Q_{i,t}^T \Omega_{i,t} y_i\}, \\ \hat{V}_{i,t} &= O U_{i,t}^T \Omega_{i,t}(y_i - Q_{i,t}\hat{\alpha}_t), \end{aligned} \tag{3.8}$$

where $\Sigma_{i,t} = R_i + K_{ih_r}^{1/2} U_{i,t} O U_{i,t}^T K_{ih_r}^{1/2}$ and $\Omega_{i,t} = K_{ih_f}^{1/2}(t) \Sigma_{i,t}^{-1} K_{ih_f}^{1/2}(t)$.

3.2.2 Asymptotic results

In this section, we study the asymptotic properties and inference of the local likelihood estimators of time-varying coefficients. The asymptotic bias and variance of estimators are investigated in the following four cases:

(i) there are an infinite number of repeated measurements within subjects, $n_i \to \infty$, and $h_f \neq h_r$;

(ii) there are an infinite number of repeated measurements within subjects, $n_i \to \infty$, and $h_f = h_r$;

(iii) there are finite but large enough repeated measurements $n_i < \infty$ and $h_f \neq h_r$;

(iv) there are finite but large enough repeated measurements $n_i < \infty$ and $h_f = h_r$;

For each case, it is assumed that $n \to \infty$ and $h \to 0$. The proof is deferred to the Appendix. Below are regularity conditions to facilitate the proofs.

(a) The design time points $t_{ij},\ j = 1, \ldots, n_i,\ i = 1, \ldots, n$ are i.i.d. with density $f(\cdot)$.

(b) $f(t) \neq 0$ for any time point t in the support of $f(t)$, and the continuous second derivative $f''(t)$ exits.

(c) $\beta(t)$ has a twice-continuous derivative at t_j, that is, $\beta''(t)$ exists and is continuous.

(d) $\gamma(s,t) =$cov$(\nu_i(s), \nu_i(t))$ is twice-continuous at time t.

(e) The variance function $\sigma^2(t)$ is continuous at time t.

(f) The kernel K is a bounded and symmetric density function with bounded support such that $\int K(u)du = 1$, $\int K(u)udu = 0$ and $\int K(u)^2du < \infty$.

(g) $n_i \to \infty,\ h \to 0,\ n_ih^2 \to \infty,\ h = O\left((nm)^{-1/5}\right)$ and $n_ih^3 \to 0$ where $m = n\left(\sum_{i=1}^{n} n_i^{-1}\right)^{-1}$

For notational simplicity, we only consider a model with a single random variable covariate ($q = 1$). Define $\eta_{it} = X_{i,t}\beta_t$, $\eta_{it}^{(2)} = X_{i,t}\beta_t^{(2)}$, $X_{i,t}^s = X_{i,t}^T X_{i,t}$, $\mu_a(b) = \int u^a K^b(u)$ and $\varphi_a(b,c) = \int u^a K^b(u) K^c\left(C_h u\right) du$ where n_{it} is a number of repeated observations of subject i at time t, $X_{i,t}$ is a $n_{it} \times p$ dimension covariance of subject i at time t and $C_h = h_r/h_f < \infty$.

Theorem 3.2.1. *(Case I: $n_i \to \infty$). Under Conditions a-g, the bias and variance of the local likelihood estimator $\hat{\beta}(t)$ are, respectively,* *(i)* $(h_f \neq h_r)$

$$
\begin{aligned}
Bias(\hat{\beta}) &= M_1 \frac{h_f^2}{2} \left(\left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{i,t}^{T} \eta_{it}^{(2)} \right) + O_p\left(\frac{1}{\sqrt{m h_f}} \right) \right), \\
Var(\hat{\beta}) &= \frac{1}{n} \gamma(t,t) \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} n_i^2 z_{it}^2 X_{it}^{s} \right) \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} + O_p\left(\frac{1}{n\sqrt{m h_f}} \right),
\end{aligned}
$$

where $M_1 = \frac{1 - C_h \varphi_0\left(\frac{1}{2},\frac{1}{2}\right) \varphi_2\left(\frac{1}{2},\frac{1}{2}\right)}{1 - C_h \varphi_0^2\left(\frac{1}{2},\frac{1}{2}\right)}$ *is a finite constant.*

(ii) $(h_f = h_r)$

$$
\begin{aligned}
Bias(\hat{\beta}) &= \frac{h^2}{2} \left(\frac{1}{n} \sum_{i=1}^{n} X_{it}^{-s} X_{it}^{T} \eta_{it}^{(2)} + O_p\left(\frac{1}{\sqrt{mh}} \right) \right), \\
Var(\hat{\beta}) &= \frac{\gamma(t,t)}{n} \left(\frac{1}{n} \sum_{i=1}^{n} z_{it}^2 X_{it}^{-s} + O_p\left(\frac{1}{\sqrt{mh}} \right) \right),
\end{aligned}
$$

Theorem 3.2.2. *(Case II: $n_i < \infty$). Under Conditions a-g, the bias and variance of the local likelihood estimator $\hat{\beta}(t)$ are, respectively,*

(iii) $(h_f \neq h_r)$

$$
\begin{aligned}
Bias(\hat{\beta}) &= M_1 \frac{h_f^2}{2} \left(\frac{1}{n} \left(\sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{i,t}^{T} \eta_{it}^{(2)} \right) + O_p\left(\frac{1}{\sqrt{N h_f}} \right) \right), \\
Var(\hat{\beta}) &= \frac{1}{n} \gamma(t,t) \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} n_i^2 z_{it}^2 X_{it}^{s} \right) \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} \\
&\quad + \frac{1}{n} \frac{M_2}{h f(x)} \tau^2(t) \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right) \left(\frac{1}{n} \sum_{i=1}^{n} n_i X_{it}^{s} \right)^{-1} \\
&\quad + O_p\left(\frac{1}{N\sqrt{N h_f}} \right),
\end{aligned}
$$

where $M_2 = \frac{1-2C_h\varphi_0\left(\frac{1}{2},\frac{1}{2}\right)\varphi_0\left(\frac{3}{2},\frac{1}{2}\right)+C_h^2\varphi_0\left(\frac{1}{2},\frac{1}{2}\right)\varphi_0(1,1)}{\left(1-C_h\varphi_0^2\left(\frac{1}{2},\frac{1}{2}\right)\right)^2}$ *is a finite constant.*

(iv) $(h_f = h_r)$

$$\begin{aligned} Bias(\hat{\beta}) &= \frac{h^2}{2}\left(\frac{1}{n}\sum_{i=1}^{n} X_{it}^{-s} X_{it}^{T} \eta_{it}^{(2)} + O_p\left(\frac{1}{\sqrt{mh}}\right)\right), \\ Var(\hat{\beta}) &= \frac{1}{n^2}\left(\gamma(t,t)\sum_{i=1}^{n} z_{it}^2 X_{it}^{-s} + \frac{\tau^2(t)}{nhf(x)}\sum_{i=1}^{n} X_{it}^{-s} + O_p\left(\frac{n}{\sqrt{mh}}\right)\right). \end{aligned}$$

With $q = 1$, O_t is assumed to be a diagonal matrix with nonzero elements $G_i(t) = \delta_1^2(t)$ and $H_i(t) = \delta_2^2(t)$, denoted by $O_t = \text{diag}\{\delta_1^2(t), \delta_2^2(t)\}$ for technical and notational convenience.

Note that if $X_{1\times 1}$ and $Z_{1\times 1}$ in model (3.1) are 1, then the model is a nonparametric mixed model. In this case, our estimator reduces to the local linear mixed estimator by Wu & Zhang (2002) . When $\gamma(t,t) = 0$ or the random effects $\nu_i(t)$ in model (3.1) are 0, then the serial correlation is 0. In this case, our estimator reduces to the corresponding estimator for independent data.

3.3 The Functional type of concordance correlation coefficient

Lin et al. (2007) proposed the UCCC to assess the agreement for k methods measuring multiple observations on a subject. The UCCC is subdivided into three indices-inter, intra and total CCC-each denoting a product of accuracy and precision as a function of variances. To extract the variance components, Lin et al. (2007) used two-way ANOVA including a fixed effect of method and random effects of a subject and interaction between the method and the subject. They proposed the asymptotic distribution of UCCCs using GEE. The index works well for comparison to multiple methods, each making repeated measures, as long as they have a stable mean and variation within each subject over time. Li & Chow (2005) introduced the concept of reproducibility for functional data, and proposed functional-type weighted CCC for an overall agreement of paired repeated measures, where the index accommodates the time trend on the response values of repeated measures. This index is limited to paired data, but it broadened the utilization of the measure of the agreement to curve data, which became popular by developing methods in data collection, computing technology and statistical modeling. In order to encompass functional data taken by multiple methods, we propose the FCCC based on the functional mixed model. Modeling the time-varying population and the time-varying variation may reduce modeling bias and enhance the accuracy of agreement in ILD.

The FCCC is a covariance-based index relying on the variance components in the functional mixed model (3.1). To improve the accuracy of the FCCC, the sources (i.e., gender, race, age) which may lead to an increase in the variance estimates in the FCCC that have to be taken into account (Carrasco & Jover, 2003). The functional mixed model will contain these covariates in order to attain this objective; then the FCCC may have a different form according to the potential covariates in model (3.1). We construct the two-way mixed model with one fixed effect, one random effect and their interaction effect for an analogy to the model (2.1.4) as an example and we will explore other cases in the simulation study.

The functional two-way mixed model is as following:

$$y_{ijk}(t_l) = \mu(t_l) + \alpha_i(t_l) + \beta_j(t_l) + \eta_{ij}(t_l) + \varepsilon_{ijk}(t_l), \qquad (3.9)$$

where $y_{ijk}(t_l)$ is the observed response variable of subject i by method j at time t_l, $i = 1, \cdots, n$, $j = 1, \cdots, J$, $k = 1, \cdots, K$, and $l = 1, \cdots, m$. $\mu(\cdot)$ is a baseline curve. $\beta_j(t)$ is the jth method effect at time t. $\alpha_i(\cdot)$ is a realization of the ith subject specific curve of a Gaussian process with mean zero and covariance $\gamma_\alpha(t, s)$, $\gamma_\alpha(t, t) = \sigma^2_\alpha(t)$. $\eta_{ij}(\cdot)$ is a method effect by subjects of a Gaussian process with mean zero and covariance $\gamma_\eta(t, s)$, $\gamma_\eta(t, t) = \sigma^2_\eta(t)$. $\varepsilon_{ijk}(t)$ is a measurement error of realization from a Gaussian process with mean zero and variance $\sigma^2_\varepsilon(t)$. Three random variables $\alpha_i(\cdot)$, $\eta_{ij}(\cdot)$ and $\varepsilon_{ijk}(\cdot)$ are mutually independent. Define $\sigma^2_\beta(t) = \frac{\sum\left(\beta_j(t)-\bar{\beta}_\cdot(t)\right)^2}{k-1}$ as the variance between methods at time t.

$\text{FCCC}_{\text{intra}}$ is used to measure the degree of the agreement among repeated measures taken by the same method. When a method took more than one observation at each time point on the same subject, or when there is another cross effect to the method, $\text{FCCC}_{\text{intra}}$ can be evaluated as a time-varying measure. For each method, assume there are K measurements on a subject at each time point. The intra FCCC between k and k' readings, $k = 1, \cdots, K$ is defined for a given method j:

$$\begin{aligned}\text{FCCC}_{\text{intra}}(t) &= \text{FCCC}_{\text{intra};j}(t) \\ &= 1 - \frac{E\left[W_{i;j}|X_i, Z_i\right]}{E\left[W_{i;j}|\text{all elements in } y_{ij1}(t), \cdots, y_{ijK}(t) \text{ are uncorrelated}, X_i, Z_i\right]} \\ &= \frac{\sigma^2_\alpha(t) + \sigma^2_\eta(t)}{\sigma^2_\alpha(t) + \sigma^2_\eta(t) + \sigma^2_\varepsilon(t)}, \qquad (3.10)\end{aligned}$$

where $W_{i;j} = \sum_{r=1}^{K} \left(y_{ijk}(t) - \bar{y}_{ij\cdot}(t)\right)^2$.

Because the $\text{FCCC}_{\text{intra};j}$ does not depend on j, it represents intra-method agreement for any method. Expected values of measurements taken by the same method are equal, which implies perfect accuracy. The $\text{FCCC}_{\text{intra}}$ represents the degree of the methods' reproducibility at time t, Pearson correlation $\rho_{kk';j}(t) = \rho(t)$,

$$\text{FCCC}_{\text{intra}}(t) = \rho(t) = \frac{\text{cov}\left(y_{ijk}(t), y_{ijk'}(t)\right)}{\sqrt{\text{var}y_{ijk}(t)}\sqrt{\text{var}y_{ijk'}(t)}} = \frac{\sigma_\alpha^2(t) + \sigma_\eta^2(t)}{\sigma_\alpha^2(t) + \sigma_\eta^2(t) + \sigma_\varepsilon^2(t)}.$$

The $\text{FCCC}_{\text{inter};j,j'}$ quantifies the agreement between methods j and j' based on the average of repeated measurements. Assume there are K readings on a subject at time t.

$$\begin{aligned}\text{FCCC}_{\text{inter}}(t) &= \text{FCCC}_{\text{inter};j,j'}(t) \\ &= 1 - \frac{E\left[W_{i;j,j'}|X_i, Z_i\right]}{E\left[W_{i;j,j'}|\bar{y}_{i1\cdot}(t), \cdots, \bar{y}_{iJ\cdot}(t) \text{ are uncorrelated}, X_i, Z_i\right]} \\ &= \frac{\sigma_\alpha^2(t)}{\sigma_\alpha^2(t) + \sigma_\eta^2 + \sigma_\beta^2(t) + \sigma_\varepsilon^2(t)/K}, \qquad (3.11)\end{aligned}$$

where $W_{i;j,j'} = \sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}\left(\bar{y}_{ij\cdot}(t) - \bar{y}_{ij'\cdot}(t)\right)^2$.

Since the $\text{FCCC}_{\text{inter};j,j'}$ does not depend on j and j', the $\text{FCCC}_{\text{inter}}$ represents inter-method agreement between methods. It quantifies the lack of agreement between methods using two disagreement sources;

$$\rho(t) = \frac{\sigma_\alpha^2(t)}{\sigma_\alpha^2(t) + \sigma_\eta^2(t) + \sigma_\varepsilon^2(t)/K},$$

represents to which the changes of methods are related at time t;

$$\chi(t) = \frac{\sigma_\alpha^2(t) + \sigma_\eta^2(t) + \sigma_\varepsilon^2(t)/K}{\sigma_\alpha^2(t) + \sigma_\eta^2(t) + \sigma_\beta^2(t) + \sigma_\varepsilon^2(t)/K},$$

is the degree of conformity between means of methods at time t. Thus, small $\text{FCCC}_{\text{inter}}$ occurs either when the method variance is large, (i.e., a respective method provides different measurements), or when measurements are scattered (i.e., the proportion of the variance attributed by random effects is small). Note that in the comparison study, the number of repeated measurements K may be the same between methods at each occasion whereas the number of repeated measurements can be different by occasion, $K = K(t)$ due to skipping the scheduled time set.

$\text{FCCC}_{\text{total};j,j'}$ is a measure of the agreement between methods j and j' based on individual measurement k given time t.

The $\text{FCCC}_{\text{total}}$ is defined as follows:

$$\begin{aligned} \text{FCCC}_{\text{total}}(t) &= \text{FCCC}_{\text{total};j,j'}(t) \\ &= 1 - \frac{E\left[W_{ik;j,j'}|X_i, Z_i\right]}{E\left[W_{ik;j,j'}|y_{i1k}(t), \cdots, y_{iJk}(t) \text{ are uncorrelated}|X_i, Z_i\right]} \\ &= \frac{\sigma_\alpha^2(t)}{\sigma_\alpha^2(t) + \sigma_\beta^2(t) + \sigma_\eta^2(t) + \sigma_\varepsilon^2(t)}, \end{aligned} \tag{3.12}$$

where $W_{ik;j,j'} = \sum_{j=1}^{J-1} \sum_{j'=j+1}^{J} \left(y_{ijk}(t) - y_{ij'k}(t)\right)^2$.

The $\text{FCCC}_{\text{total}}$ is not changed by j, and it represents an overall inter-method agreement profile based on individual measurements, implying how well individual observations agree with each other. $\text{FCCC}_{\text{total}}$ is also decomposed into precision and accuracy, which provide the same information about the the proportion of the variance attributed to subjects and the closeness between the means of methods. The only difference is that inter-method FCCC has σ_ε^2/K. Thus, the $\text{FCCC}_{\text{total}}$ is distinguished from $\text{FCCC}_{\text{inter}}$ when $K > 1$; methods measure multiple replications at each occasion; or there is a cross effect across a method.

The FCCCs separate the degree of measurement agreement from an overall time trend so that the FCCCs may enhance the ability to detect agreement between methods for curve data, while UCCCs may mix the between-method disagreement and the time trend by assuming stable mean and variance over time. Note that FCCCs require many observations to provide fine-grained information. If there are neither any effects crossed with the method effect nor replications on a subject at each time, $\eta_{ij}(t)$ is confounded with $\varepsilon_{ij}(t)$. In this case, only $\text{FCCC}^c_{\text{inter}}$ is evaluated based on two-way mixed model without interaction effect, which is an extension of Carrasco & Jover (2003). Based on the simulation results of Lin et al. (2007), the performance of the index of Carrasco & Jover (2003) is comparable to the UCCC for ordered and continuous data. We also check to verify that $\text{FCCC}^c_{\text{inter}}$ is comparable to $\text{FCCC}_{\text{inter}}$ in Chapter 4.

Because FCCCs consist of the variance component in model (3.9), with the sample estimates, $\hat{\beta}$, $\hat{\sigma}^2_u$, $\hat{\sigma}^2_\eta$, and $\hat{\sigma}^2_\varepsilon(t)$, the sample counterparts of FCCCs are given as

$$\begin{aligned}\widehat{\text{FCCC}}_{\text{intra}}(t) &= \frac{\hat{\sigma}^2_\alpha(t)+\hat{\sigma}^2_\eta(t)}{\hat{\sigma}^2_\alpha(t)+\hat{\sigma}^2_\eta(t)+\hat{\sigma}^2_\varepsilon(t)},\\ \widehat{\text{FCCC}}_{\text{inter}}(t) &= \frac{\hat{\sigma}^2_\alpha(t)}{\hat{\sigma}^2_\alpha(t)+\hat{\sigma}^2_\beta(t)+\hat{\sigma}^2_\eta(t)+\hat{\sigma}^2_\varepsilon(t)/K},\\ \widehat{\text{FCCC}}_{\text{total}}(t) &= \frac{\hat{\sigma}^2_\alpha(t)}{\hat{\sigma}^2_\alpha(t)+\hat{\sigma}^2_\beta(t)+\hat{\sigma}^2_\eta(t)+\hat{\sigma}^2_\varepsilon(t)},\end{aligned}$$

when there are K repeated measures for each method on a subject at each time point. We evaluate the performance of the index estimates by examining MSE in a simulation study.

All three time-varying indices, $\text{FCCC}_{\text{Intra}}(t)$, $\text{FCCC}_{\text{Inter}}(t)$ and $\text{FCCC}_{\text{Total}}(t)$ can be summarized into one number. Let $w(\cdot)$ be a weight function subjected to importance in different occasions, which takes non-negative values. If there is prior information on the importance of observing occasions, one may utilize this information as a weight function. The inverse of measurement error at each time can be a weight function if there is no information indicating that smaller variation within replicate measurements of the subjects receive larger weights.

$$\begin{aligned}\mathrm{FCCC}_{\mathrm{total}} &= \int \mathrm{FCCC}_{\mathrm{total}}(t)w(t)dt,\\ \mathrm{FCCC}_{\mathrm{inter}} &= \int \mathrm{FCCC}_{\mathrm{inter}}(t)w(t)dt,\\ \mathrm{FCCC}_{\mathrm{intra}} &= \int \mathrm{FCCC}_{\mathrm{intra}}(t)w(t)dt,\end{aligned} \tag{3.13}$$

where $\int w(t)dt = 1$.

3.3.1 The bootstrap variance estimation of indices

Statistical inference on the FCCC involves analytical derivation of the asymptotic distribution of the index. Because it is challenging to find out the relationship between variance components in the functional mixed model, we propose the use of bootstrap techniques in order to estimate the index distribution to overcome mathematical complexity. The bootstrap is a resampling method for statistical inference. It is used to obtain standard errors, confidence intervals and critical values of the sampling distribution of an estimator. There are three typical approaches to generate bootstrap samples: nonparametric bootstrap, semiparametric bootstrap and parametric bootstrap. In the linear model context, nonparametric bootstrap selects subjects and associated covariates randomly with replacement from a set of n subjects. It is flexible because no assumptions are made about the functional form of the population distribution as long as the sample adequately reflects the population with sufficient sample size. If not, unrepresentative sets of sample and population may result in small variance and large bias. The semiparameric bootstrap resamples residuals randomly after fitting a model. We found that the residual variance estimates through the semiparametric bootstrap are smaller than the actual variance and the coverage probability of indices was around 85% at the moderate sample size, $n = 50$. This finding is consistent with the result of Morris (2002), where the semiparametric bootstrap consistently underestimates the variation in the finite sample for the mixed effects model. He showed that coverage probability decreases as the number of repeated measures and sample size increase. Therefore, the residual bootstrap is inappropriate for ILD. The parametric bootstrap is less flexible but provides stable and accurate variance estimates of the FCCCs unless the underlying subject distribution is heavy-tailed. Thus, we generate subjects and residuals from a normal distribution, assuming subjects are sharply concentrated with no extreme values.

For the sake of simplicity, one random effect ($q = 1$) is assumed. The parametric bootstrap is as follows:

(1) Fit model (3.1) and obtain $\widehat{\beta}(\cdot)$, $\widehat{\sigma}_u^2(\cdot)$, $\widehat{f}(t)$ and $\widehat{\sigma}_e^2(\cdot)$.

(2) Draw n subjects $u_1^*, \ldots, u_n^*$ from $\mathrm{N}(0, \widehat{\sigma}_u^2)$ where

$$u_i^* = u_i(t)/\sqrt{\widehat{f}(t)}.$$

Draw N total observations $\varepsilon_i(t_1), \ldots, \varepsilon_i(t_{n_i})$ from $\mathrm{N}\Big(0, \hat{R}_i\Big)$, $\hat{R}_i = \mathrm{diag}(\hat{\sigma}_\varepsilon^2(t_{i1}), \cdots, \hat{\sigma}_\varepsilon^2(t_{in_i}))$ for $i = 1, \ldots, n$.

(3) Calculate bootstrap responses $\tilde{y}_i$ as

$$\tilde{y}_i(t) = x_i^T \widehat{\beta}(t) + u_i(t) + \varepsilon_i(t), \ i = 1, \cdots, n.$$

(4) Refit model (3.1) with $\tilde{y}_i(t)$ and calculate FCCCs on the bootstrap sample.

(5) Repeat 2- 4 B times.

B numbers of the FCCCs allows us to obtain standard deviation and quantiles, and to make inference about the FCCCs.

3.4 Simulation study

We access the finite sample properties of FCCCs based on the functional mixed model in Section 3.2 via comparison with the UCCCs based on the two-way mixed model for ILD. All simulation studies were coded in R.

We generated simulation data from the functional mixed model:

$$Y_{ijk}(t_l) = \beta_1(t_l) + u_i(t_l) + \beta_{2j}(t_l) + \beta_{3k}(t_l) + \varepsilon_{ijk}(t_l), \tag{3.14}$$

where $\varepsilon(t)$ is a realization of Gaussian process with mean 0 and variance $\sigma_\varepsilon^2(t)$. A time-varying intercept $\beta_1(t)$ is defined as $\beta_1(t) = 0.1t$. We let the fixed effects of β_2 and β_3 be three and two dimensional,

$$\begin{aligned} \beta_2(t) &= \left((\sin(-t\pi/L),\ \sin(-t\pi/L),\ 1.5 + \sin(-t\pi/L)\right)),\ \text{for}\ \ L = 50,\ j = 1,\ 2,\ 3 \\ \beta_3(t) &= \left(2t/L + 1,\ 1 - 2t/L\right). \end{aligned}$$

The random subject process $u_i(t)$ is normally distributed with mean 0 and covariance

$$f^{1/2}(s)\sigma_u^2 f^{1/2}(t),$$

between t and s where

$$f(t) = \exp\left(-t/L\right), \qquad i = 1 \cdots, n,$$

for $n = 50$ or 100. The observations from different subjects are independent, but the within-subject repeated measurements are correlated,

$$\rho(t,s) = \sigma_u^2 f(t) f(s) / \left(\sqrt{\sigma_e^2(t) + \sigma_u^2 f^2(t)} \sqrt{\sigma_e^2(s) + \sigma_u^2 f^2(s)} \right) \quad \text{for} \quad t \neq s.$$

To simulate the weakly, moderately and strongly correlated data, we select ($\sigma_\varepsilon^2(t)$, σ_u^2) as $(2,1)$, $(1,1)$, $(1,2)$, which correspond to $\rho = 0.2,\ 0.4,\ 0.7$ respectively. This reflects ILD with time-varying individual profiles. Each subject has the same set of 50 occasions $\{1, 2, \cdots, 50\}$ but same occasions may be randomly missed. In our simulation, we consider two schemes to generate the time.

Scheme I: We set $t = 1, \cdots, 50$.

Scheme II: We generate the set of time points in the following way: each subject has a set of time points, {1, 2, ..., 50}, and each observation time point, except the first time, has a 20% probability of being skipped.

We conduct 500 Monte Carlo simulations. For each simulation, we generate 500 bootstrap samples in order to construct the confidence intervals. Time-varying coefficients are estimated by three selected bandwidths (1, 3, 6). The Epanechnikov kernel is used whenever a kernel is needed. For each simulation study, mean and mean square error (MSE) of indices are calculated. The means are represented by the column labelled Mean in the tables. For each index, an estimate of the MSE

$$E||\hat{f}(t) - f(t)||^2,$$

are estimated by

$$\frac{1}{500} \sum_{t=1}^{500} \left\{ \widehat{f}(t) - f(t) \right\}^2,$$

and listed in the column labelled MSE in the tables. The standard deviation (SD) indicates SD of the 500 estimates of each index. Standard error (SE) implies the mean of 500 SEs of indices calculated by 500 bootstrap sample. In addition, for each index, the 95% coverage probability (CP) is calculated using a bootstrap percentile

$$\frac{1}{500}\sum_{t=1}^{500} I\left\{\delta \in \left[\hat{\delta}_{(0.025)},\ \hat{\delta}_{(0.975)}\right]\right\},$$

where δ is true index and $\hat{\delta}_{(a)}$ is the a% percentile among an index of estimates of bootstrap samples. In other words, the CP is the case that the bootstrap confidence interval contain the true index and it assess the overall performance of the proposed estimation procedure and parametric bootstrap. For comparison, we also estimate UCCC. Estimates of variance components in UCCCs are evaluated by the mixed effects model (Barnhart et al., 2005, 2007).

3.4.1 Time Scheme I

In this section, we present the finite-sample performance of the FCCC estimators for data generated by time scheme I. Figure 3.1 shows the performance of our estimating procedure for Model (3.14) under strong serial correlation, $\sigma_\varepsilon^2(t) = 1$ and $\sigma_u^2 = 2$ using Epanechnikov kernel with bandwidth 3. In Figure 3.1, four panels depict respective coefficient (black solid curve), its estimates (red dotted curve) and 95% confidence interval (gray curve). Note, the model (3.14) has two categorical variables, $\beta_{2j},\ j = 1, 2, 3$ and $\beta_{3k},\ k = 1, 2$. Since the first level is set to a base group in the regression notation, the intercept for the base group is the overall intercept for model, and the coefficient estimates for a particular group represent the estimated difference in intercepts between that group and the base group; $b_0 = \hat{\beta}_1 + \hat{\beta}_{21} + \hat{\beta}_{31}$, $b_1 = \hat{\beta}_{22} - \hat{\beta}_{21}$, $b_2 = \hat{\beta}_{23} - \hat{\beta}_{21}$, $b_3 = \hat{\beta}_{32} - \hat{\beta}_{31}$.

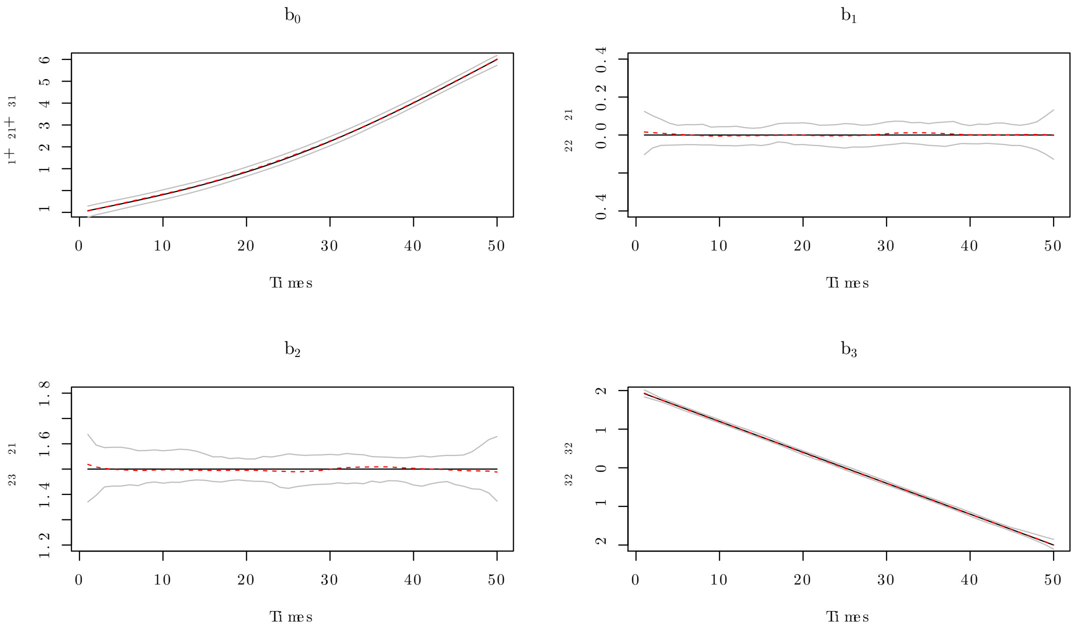

Figure 3.1. Plots for coefficients at $\sigma_e^2 = 1, \sigma_u^2 = 2 \times \exp(-t/50)$. The solid curve represents true values and red dotted curve indicates estimates. Gray curves stand for the 2.5th and 97.5th percentiles of 500 bootstrap estimates, respectively. Epanechnikov kernel with bandwidth 3 is used.

The estimates are on the true curves and 95% confidence intervals of estimates are narrow at the four plots in Figure 3.1. The mean estimates of constant variance σ_ε^2 is 1.0036 (near 1). Time-varying subject standard deviation (black solid curve), its estimates (red dotted curve) and its 95% confidence interval (gray curve) are displayed in Figure 3.2. The red dotted curve is overlapped to the solid curve over time periods, and the width of gray curves, the 95% interval, becomes narrower as n increases. As Figures 3.1 and 3.2 illustrate our estimation procedure provides consistent variance components as well as regression coefficients where both are the components of sample counterparts of FCCCs.

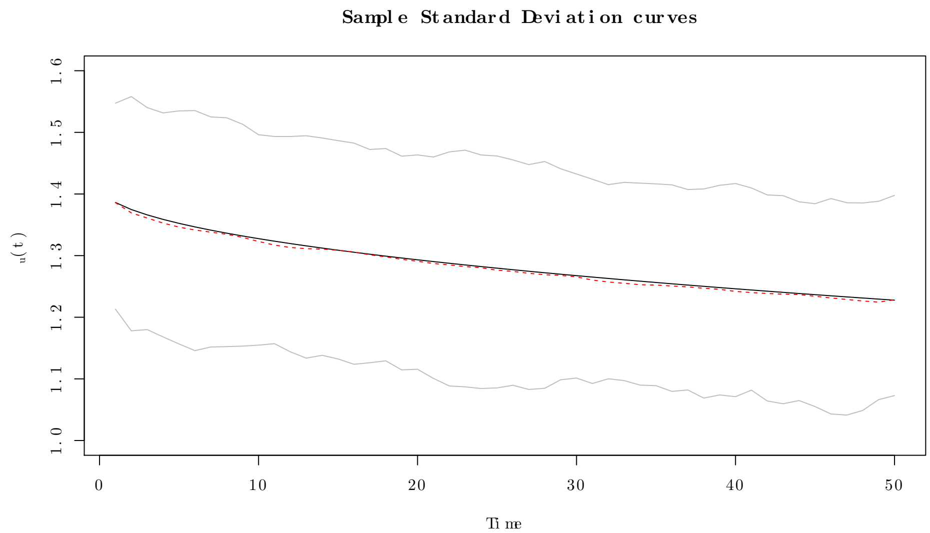

Figure 3.2. Plots for subject variance at $\sigma_e^2 = 1, \sigma_u^2 = 2 \times \exp(-t/50)$. Solid curve plots $\sigma_u^2(t)$ and red dotted curve is its estimates when model (3.14) is estimated with bandwidth 3. The gray curves stand for the 2.5th and 97.5th percentiles of 500 bootstrap estimates, respectively.

Based on model (3.14), the indices are defined as follows:

$$\begin{aligned} UCCC_{\text{intra}} &= \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2}, \\ UCCC_{\text{inter}} &= \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\beta^2 + \sigma_\varepsilon^2/100}, \\ UCCC_{\text{total}} &= \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\beta^2 + \sigma_\varepsilon^2}, \end{aligned} \tag{3.15}$$

where $\sigma_\beta^2 = \frac{1}{6}\sum_{k=1}^{2}\sum_{k'=2}^{3}(\beta_{2,k} - \beta_{2,k'})^2$.

$$\begin{aligned} FCCC_{\text{intra}} &= \frac{\sigma_a^2(t)}{\sigma_a^2(t) + \sigma_\varepsilon^2(t)}, \\ FCCC_{\text{inter}} &= \frac{\sigma_a^2(t)}{\sigma_a^2(t) + \sigma_\beta^2(t) + \sigma_\varepsilon^2(t)/2}, \\ FCCC_{\text{total}} &= \frac{\sigma_a^2(t)}{\sigma_a^2(t) + \sigma_\beta^2(t) + \sigma_\varepsilon^2(t)}. \end{aligned} \tag{3.16}$$

Table 3.1. Finite sample performance of Indices at time 10, 25 and 40 for strongly correlated data

$\sigma_\varepsilon^2(t) = 1, \quad \sigma_u^2(t) = 2 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		$n = 50$						$n = 100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	h	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.6380		.5850		.5018		.6380		.5850		.5018	
10	1	.6227	.0028	.5693	.0036	.4867	.0034	.6366	.0011	.5842	.0015	.5011	.0014
	3	.6350	.0019	.5823	.0023	.4997	.0022	.6380	.0010	.5851	.0013	.4935	.0013
	6	.6315	.0019	.5789	.0021	.4961	.0021	.6388	.0010	.5852	.0012	.5004	.0012
THEORETICAL		.6208		.5671		.4834		.6208		.5671		.4834	
25	1	.6187	.0032	.5517	.0039	.4698	.0037	.6190	.0011	.5653	.0014	.4839	.0013
	3	.6183	.0022	.5643	.0024	.4816	.0025	.6206	.0009	.5674	.0012	.4835	.0011
	6	.6136	.0021	.5580	.0023	.4769	.0023	.6211	.0010	.5712	.0012	.4835	.0012
THEORETICAL		.6083		.5540		.4702		.6083		.5540		.4702	
40	1	.5915	.0036	.5366	.0043	.4540	.0039	.6068	.0012	.5521	.0016	.4706	.0015
	3	.6047	.0023	.5526	.0026	.4690	.0026	.6090	.0009	.5539	.0012	.4705	.0011
	6	.6006	.0021	.5472	.0022	.4637	.0022	.6102	.0010	.5568	.0011	.4705	.0011
THEORETICAL		.6228		.6847		.4856		.6228		.6847		.4856	
UCCC	1	.3047	.1031	.6628	.0027	.2663	.0497	.3150	.0955	.6757	.0016	.2756	.0447
	3	.3132	.0972	.6752	.0033	.2742	.0482	.3154	.0951	.6705	.0082	.2760	.0445
	6	.3125	.0979	.6725	.0019	.2736	.0463	.3175	.0940	.6793	.0013	.2781	.0437

Table 3.2. SDs, SEs and 95% CPs of indices at time 10, 25 and 40 for strongly correlated data

$\sigma_\varepsilon^2(t) = 1, \quad \sigma_u^2(t) = 2 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC, $n = 50$

t	h	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0509	.0531 (.0038)	.9640	.0586	.0594 (.0031)	.9634	.0568	.0579 (.0020)	.9640
	3	.0432	.0451 (.0029)	.9681	.0471	.0483 (.0023)	.9461	.0474	.0490 (.0018)	.9513
10	6	.0428	.0430 (.0028)	.9508	.0454	.0456 (.0023)	.9512	.0459	.0462 (.0018)	.9512
	1	.0556	.0574 (.0036)	.9634	.0618	.0624 (.0027)	.9623	.0582	.0580 (.0021)	.9528
	3	.0473	.0490 (.0029)	.9681	.0494	.0501 (.0023)	.9641	.0498	.0490 (.0019)	.9424
25	6	.0455	.0460 (.0028)	.9459	.0474	.0486 (.0021)	.9512	.0479	.0487 (.0018)	.9512
	1	.0546	.0550 (.0033)	.9645	.0637	.0644 (.0024)	.9545	.0607	.0627 (.0023)	.9551
	3	.0482	.0490 (.0027)	.9642	.0515	.0519 (.0022)	.9521	.0510	.0517 (.0019)	.9543
40	6	.0449	.0464 (.0025)	.9512	.0468	.0479 (.0021)	.9578	.0464	.0479 (.0019)	.9444

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0332	.0363 (.0021)	.9788	.0386	.0413 (.0029)	.9708	.0377	.0407 (.0024)	.9728
10	3	.0370	.0352 (.0025)	.9438	.0363	.0360 (.0023)	.9326	.0358	.0369 (.0016)	.9526
	6	.0316	.0321 (.0018)	.9556	.0346	.0349 (.0016)	.9556	.0349	.0356 (.0010)	.9556
	1	.0337	.0373 (.0020)	.9707	.0377	.0401 (.0018)	.9708	.0366	.0400 (.0014)	.9746
25	3	.0329	.0341 (.0017)	.9538	.0345	.0345 (.0015)	.9413	.0338	.0350 (.0013)	.9438
	6	.0320	.0338 (.0015)	.9556	.0345	.0354 (.0012)	.9556	.0345	.0357 (.0010)	.9562
	1	.0347	.0379 (.0019)	.9702	.0404	.0424 (.0017)	.9729	.0386	.0401 (.0013)	.9648
40	3	.0332	.0356 (.0017)	.9638	.0352	.0378 (.0014)	.9551	.0342	.0379 (.0012)	.9538
	6	.0311	.0341 (.0015)	.9566	.0336	.0356 (.0012)	.9556	.0326	.0356 (.0010)	.9556

Table 3.1 summarizes the performance of the FCCC estimates under strong correlation, $\sigma_e^2 = 1, \sigma_u^2 = 2 \times \exp(-t/50)$, with three selected bandwidths 1, 3, 6 at time 10, 25, and 40; for comparison, the UCCC is also presented when $n = 50$ and 100. The average of estimates of FCCCs are close to the true values at $n = 100$, although we notice that the estimates are slightly smaller than the true values at $n = 50$. The decreasing bias and MSEs along with increasing sample size tells us that the estimates of the FCCCs are consistent with the corresponding true values. We test the accuracy of the bootstrap variance for indices in Table 3.2. The SD of the 500 estimated indices from 500 simulations can be considered the true SE. In Table 3.2, the average of the estimated SEs somewhat overestimates the SD, but the difference between the mean of the estimated SEs and the true value is less than half of one SD of the estimated SEs. The impact of bandwidth selection is not crucial but exists at $n = 50$. When the model (3.14) is fitted with bandwidth 1, the FCCCs are underestimated; the estimates with bandwidths 3 and 6 and MSEs are somewhat larger. The average of SD of FCCCs for each bootstrap sample overestimates SD, which is larger than SD of the estimates with bandwidths 3 and 6. The large SEs induce optimistic decision of the agreement via large CP, 94% $\sim$ 97%. With bandwidth 3 and 6, the average of the estimated SEs somewhat overestimates the SD, but the difference between the mean of the estimated SEs and the true value is less than half of one standard deviation of the estimated standard errors. In addition, 95% CPs are around desired probability 95%, which implies the inference based on the FCCCs is precise. The performance of the index is less affected by bandwidth selection at $n = 100$. Thus, simulation results indicate that our estimation works well even under strong correlation either with proper bandwidth or at sufficient sample size.

Tables 3.5 through 3.8 report the results under weak and mild correlation. For weak correlation, Table 3.5 indicates that bandwidths 3 and 6 provide unbiased estimates and the SEs in Table 3.6 are very close to SDs for $n = 50$. In Tables 3.7 and 3.8, there is subtle difference between true FCCCs and estimates, and SEs overestimate SDs at $n = 50$. Thus, the bias of indices grow and SDs and MSEs slowly increase as serial correlation becomes stronger when the sample is small ($n = 50$). For large sample ($n = 100$), the impact of correlation is negligible. Therefore, the FCCCs become robust against the degree of the serial correlation along with increasing sample size. Note that parametric bootstrap improves the variance estimation of the FCCCs. When semiparametric bootstrap is used for inference of the FCCCs, the CPs are around 85% $\sim$ 88% for any degree of serial correlation and it increases to 94% $\sim$ 97% with parametric bootstrap. Also note that the use of semiparametric covariance on a random (subject) process also improves the accuracy of the estimation of the time-varying variance and it reduces CP bias. As for UCCCs, the estimates are quite different from theoretical values. Because the curve feature in the response variable is not considered in the two-way mixed model, which uses overall mean and variance under the assumption of stable mean and variance, indices are dominated by residual variances at denominator in equation (4.15). The MSEs of UCCC are not changed by increasing sample size except Inter CCC, which is closer to the true value. inter CCC is relatively large because residual variance is divided by a large number of repetitions. In other words, although methods have large measurement errors, inter CCC will be close to 1 for densely measured data with flat curvature.

3.4.2 Time Scheme II

In this section, data are generated by scheme 2, 20% probability of each subject being skipped at each scheduled time except the first observation time point. Because the linear mixed model is flexible for unbalanced data, it provides unbiased estimates of β. In Figure 3.3, four panels depict four coefficients estimates at 100

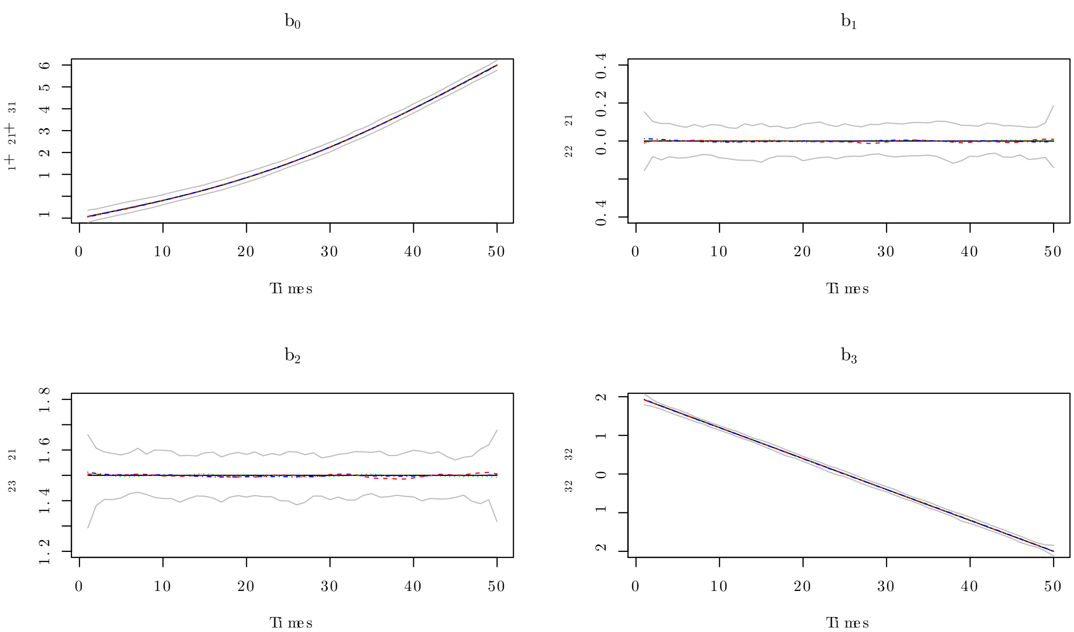

Figure 3.3. Plots for coefficients when $\sigma_e^2 = 1, \sigma_u^2 = 2 \times \exp(-t/50)$ after removing 20% of observations per occasion. The solid curve stands for true values and the dotted curves represent fitted values with bandwidth 1, 3 and 6 denoted by red, green and blue respectively. The gray curves represents the 2.5th and 97.5th percentiles of 500 bootstrap estimates with bandwidth 3.

sample size. The solid curve stands for true values and the doted curves indicate fitted values by Epanechnikov kernel with bandwidth 1, 3 and 6, denoted by red, green and blue, respectively. The gray curves stand for the 2.5th and 97.5th percentiles of 500 bootstrap estimates of Epanechnikov kernel with bandwidth 3 for comparison to time scheme I in Figure 3.1. Each dotted curve overlaps corresponding solid curves. The width of gray curves becomes narrower along with increasing sample size though the wider width of the interval compare to time scheme I indicates lower efficiency of estimates due to smaller sample size. Thus, the coefficients are well estimated at time scheme II regardless of bandwidth selection at $n = 100$. Figure 3.4 depicts time-varying subject variance (solid curve)

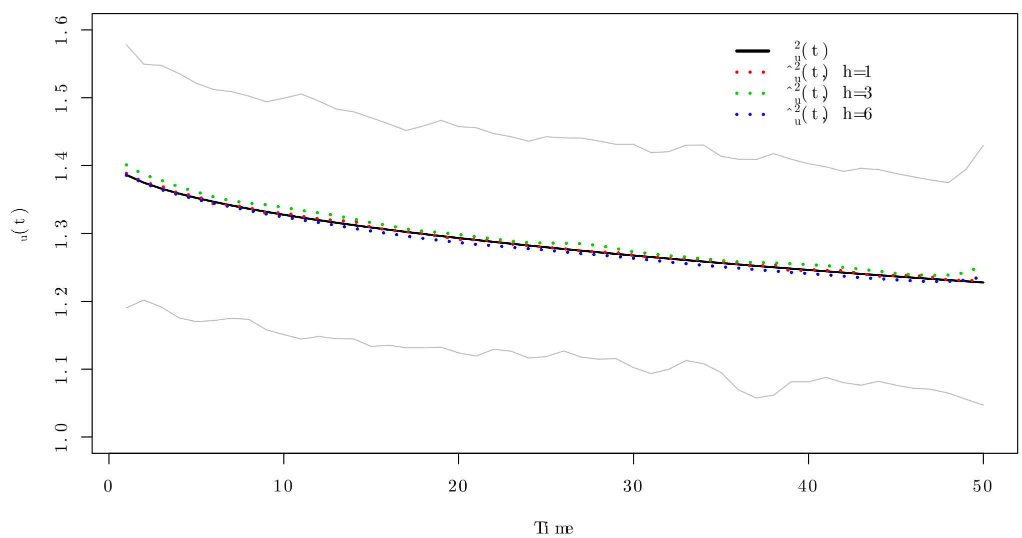

Figure 3.4. Plots for the subject random variance when $\sigma_e^2 = 1, \sigma_u^2 = 2 \times \exp(-t/50)$ after removing 20% of observations per occasion. The solid curve stands for $\sigma_u^2(t)$ and the doted curve for fitted values with bandwidth 1, 3 and 6, denoted by red, green and blue respectively. The gray curves stand for the 2.5th and 97.5th percentiles of 500 bootstrap estimates with bandwidth 3.

and its estimates (dotted curves) of specified bandwidth 1,3 and 6 denoted by red, green and blue, respectively, under strong serial correlation. All dotted curves are overlapped with true time-varying variance and $\hat{\sigma}_\varepsilon^2$s are 1.0005, 0.9995 and 1.0004 (near 1) for bandwidth 1, 3 and 6 respectively. Thus, the estimates of variance components are also accurate.

Table 3.3 reports the mean estimates of FCCCs and MSEs by three bandwidths 1, 3 and 6 at $t = 10, 25$ and 40; each scheduled time has 20% probability of being skipped except the first time. The FCCC is underestimated and MSE itself is

Table 3.3. Finite sample performance of Indices with 20% removed observations for each occasion under strong correlation

$\sigma^2_\varepsilon(t) = 1, \quad \sigma^2_u(t) = 2 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		$n = 50$						$n = 100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	h	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.6380		.5850		.5018		.6380		.5850		.5018	
10	1	.6280	.0031	.5735	.0040	.4914	.0038	.6371	.0011	.5846	.0015	.5015	.0014
	3	.6350	.0019	.5825	.0023	.5000	.0023	.6429	.0010	.5902	.0013	.5026	.0013
	6	.6325	.0019	.5803	.0021	.4974	.0022	.6353	.0010	.5885	.0012	.5007	.0010
THEORETICAL		.6208		.5671		.4834		.6208		.5671		.4834	
25	1	.6115	.0032	.5579	.0040	.4751	.0037	.6188	.0012	.5647	.0015	.4881	.0014
	3	.6175	.0024	.5640	.0025	.4812	.0026	.6245	.0009	.5714	.0012	.4881	.0013
	6	.6144	.0022	.5606	.0024	.4777	.0024	.6207	.0010	.5659	.0012	.4823	.0012
THEORETICAL		.6083		.5540		.4702		.6083		.5540		.4702	
40	1	.5983	.0034	.5437	.0042	.4611	.0038	.6068	.0012	.5517	.0016	.4749	.0015
	3	.6052	.0023	.5531	.0027	.4696	.0027	.6127	.0009	.5580	.0012	.4675	.0011
	6	.6034	.0021	.5505	.0022	.4671	.0022	.6047	.0010	.5521	.0011	.4674	.0011
THEORETICAL		.6228		.6847		.4856		.6228		.6847		.4856	
UCCC	1	.3105	.0995	.6689	.0027	.2717	.0474	.3149	.0956	.6756	.0010	.2755	.0448
	3	.3141	.0971	.6738	.0020	.2751	.0458	.3193	.0931	.6800	.0010	.2796	.0432
	6	.3111	.0988	.6710	.0020	.2723	.0469	.3125	.0971	.6736	.0010	.2734	.0457

Table 3.4. SD, SE and 95% CP of FCCC with 20% removed observations for each occasion under strong correlation

$\sigma_\varepsilon^2(t) = 1, \quad \sigma_u^2(t) = 2 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		n = 50								
t	h	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0545	.0524 (.0040)	.9489	.0621	.0593 (.0033)	.9349	.0606	.0589 (.0021)	.9357
	3	.0441	.0471 (.0029)	.9726	.0476	.0511 (.0023)	.9622	.0482	.0507 (.0018)	.9652
10	6	.0434	.0465 (.0027)	.9657	.0467	.0501 (.0021)	.9657	.0465	.0502 (.0018)	.9651
	1	.0557	.0536 (.0038)	.9560	.0618	.0607 (.0030)	.9572	.0602	.0577 (.0021)	.9571
	3	.0491	.0488 (.0029)	.9487	.0516	.0516 (.0023)	.9565	.0519	.0516 (.0020)	.9478
25	6	.0465	.0484 (.0027)	.9488	.0489	.0507 (.0021)	.9518	.0490	.0508 (.0020)	.9518
	1	.0573	.0555 (.0036)	.9518	.0638	.0603 (.0028)	.9429	.0611	.0598 (.0023)	.9500
	3	.0487	.0494 (.0025)	.9639	.0527	.0519 (.0022)	.9521	.0519	.0515 (.0020)	.9522
40	6	.0460	.0489 (.0025)	.9619	.0478	.0509 (.0021)	.9672	.0477	.0507 (.0020)	.9574

FCCC		n = 100								
t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0335	.0363 (.0020)	.9788	.0391	.0413 (.0029)	.9658	.0382	.0406 (.0024)	.9688
10	3	.0322	.0334 (.0018)	.9638	.0362	.0358 (.0020)	.9626	.0362	.0365 (.0016)	.9526
	6	.0315	.0321 (.0018)	.9565	.0346	.0349 (.0018)	.9565	.0348	.0356 (.0013)	.9610
	1	.0347	.0373 (.0020)	.9701	.0381	.0411 (.0018)	.9718	.0372	.0409 (.0014)	.9746
25	3	.0340	.0342 (.0017)	.9633	.0366	.0343 (.0015)	.9444	.0365	.0365 (.0013)	.9551
	6	.0320	.0345 (.0018)	.9565	.0345	.0353 (.0017)	.9565	.0345	.0357 (.0014)	.9630
	1	.0349	.0379 (.0019)	.9701	.0404	.0424 (.0017)	.9739	.0386	.0403 (.0013)	.9659
40	3	.0348	.0346 (.0017)	.9633	.0381	.0366 (.0014)	.9491	.0375	.0365 (.0012)	.9436
	6	.0319	.0349 (.0016)	.9555	.0336	.0356 (.0012)	.9556	.0326	.0356 (.0010)	.9556

greater than MSE of the data generated by time scheme I at $n = 50$. In contrast to complete data, there still exists some bias of FCCCs at $n = 100$ but the bias and MSE decrease as sample size increases. In Table 3.4, bootstrap standard errors of the FCCCs overestimate SDs at $n = 50$ and this leads to large CPs, similar to the data generated by time scheme I. Due to the loss of efficiency induced by a smaller number of observations, SEs are larger than SDs, but the difference is decreased at $n = 100$. For data generated by time scheme II, a large sample is required for reliable index because the convergence rate is slower than complete data, although indices is consistent. Tables 3.9 through 3.12 correspond to Tables 3.5 through 3.8. Table 3.9 shows the bias and MSE of indices under weak correlation; in contrast to the complete data case, there exists a bias which disappears at $n = 100$. In Table 3.10, SEs are larger than SD within half of one standard deviation of the estimated SEs at $n = 50$ and the difference decreases at $n = 100$. Just as data with occasion set of time scheme I, bias, SD and MSE slowly increase as the degree of serial correlation becomes stronger at $n = 50$, whereas SD and MSEs are stable over the amount of correlation for large sample, $n = 100$. Thus, overall patterns of the FCCC estimates are the same as the estimates for compete data, except for the convergence rate.

In summary, our simulation studies demonstrate that FCCCs improve the detection of agreement by separating the sources of disconcordance and efficiently estimating the functional mixed model. The estimates of FCCCs based on functional mixed models are accurate regardless of the time scheme and serial correlation when the model is fitted with a proper bandwidth. For highly correlated data having missing subjects at each time, FCCCs may be underestimated, and the variance of FCCC may be overestimated, which induces optimistic decision for a small sample.

3.5 Conclusion

Intensive longitudinal study is an emerging area of research in psychology, clinical study, and epidemiology. To develop a reliable agreement index for ILD, we use functional mixed models which describe the ILD efficiently, and provide accurate estimates of variance components and functional coefficients at every time point. The proposed FCCC based on functional mixed models is a time-varying index and illustrates the trend of measurement agreement over time. The FCCC is also robust against a functional form and serial correlation in data because functional mixed models are flexible enough to accommodate a smooth curve regardless of the degree of serial correlation for sufficient sample size. The bandwidth selection technique for functional mixed model under serial correlation is a potential area for future research.

3.6 Appendix

Proof of Theorem 3.2.1.

Use an estimating equation of α. Based on the fact $L(\hat{\alpha}) = 0$, we have

$$\hat{\alpha}_t - \alpha(t) = -\dot{L}(\alpha(t))^{-1} L(\alpha(t)),$$

where $L(\alpha_t) = -\sum_{i=1}^{n} Q_{i,t}^T \Omega^{-1}(y - Q_{i,t}\alpha_t)$ and $\dot{L}(\alpha_t) = \sum_{i=1}^{n} Q_{i,t}^T \Omega_{i,t}^{-1} Q_{i,t}$.
Let $\mathcal{T}$ be the collection of all the design time points, then

$$\begin{aligned} E(L(\alpha)|\mathcal{T}) &= -\sum_{i=1}^{n} Q_i^T \Omega^{-1} \eta_i \\ &= -\sum_{i=1}^{n} \Big[Q_i^T K_{ih_f}^{1/2} \{ R_i^{-1} - R_i^{-1} K_{ih_2}^{1/2} U_i (O_i^{-1} + U_i^T K_{ih_r}^{1/2} R_i^{-1} \\ &\qquad K_{ih_r}^{1/2} U_i) U_i^T K_{ih_2}^{1/2} R_i^{-1} \}^{-1} K_{ih_f}^{1/2} \eta_i \Big], \end{aligned}$$

where $\eta_{ij} = y_{ij} - Q_{ij,t}^T \alpha_t = x_{ij}^T \beta^{(2)}(t) \frac{(t_{ij}-t)^2}{2!} + x_{ij}^T \beta^{(3)}(t) \frac{(t_{ij}-t)^3}{3!} + \cdots = x_{ij}^T \widetilde{\beta}_{ij,t}$ and $\eta_i = (\eta_{i1}, \cdots, \eta_{in_i})$ is a n_i vector for given time t.

$$\begin{aligned} V(L(\alpha)|\mathcal{T}) &= \sum_{i=1}^{n} \{ Q_i^T \Omega_i^{-1} \Lambda_i \Omega_i^{-1} Q_i \} \\ &= \sum_{i=1}^{n} (Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \Lambda_i K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} Q_i \\ &\quad - Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \Lambda_i K_{ih_f}^{1/2} R_i^{-1} K_{ih_r}^{1/2} U_i \Gamma_i^{-1} U_i^T K_{ih_r}^{1/2} R_i^{-1} K_{ih_f}^{1/2} Q_i \\ &\quad - Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_r}^{1/2} U_i \Gamma_i^{-1} U_i^T K_{ih_r}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \Lambda_i K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} Q_i \\ &\quad + Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_r}^{1/2} U_i \Gamma_i^{-1} U_i^T K_{ih_r}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \Lambda_i K_{ih_f}^{1/2} R_i^{-1} \\ &\qquad K_{ih_r}^{1/2} U_i \Gamma_i^{-1} U_i K_{ih_r}^{1/2} R_i^{-1} K_{ih_f}^{1/2} Q_i). \end{aligned}$$

Here, consider the simple and popular case in which a model contains one random effect ($q = 1$); O_t is assumed to be a diagonal matrix with nonzero elements $G_i(t) = \delta_1^2(t)$ and $H_i(t) = \delta_2^2(t)$, denoted by $O_t = \text{diag}\{\delta_1^2(t), \delta_2^2(t)\}$ for technical and notational convenience.

Let $A_i = Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} X_i \widetilde{\beta}_i, \quad B_i = Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_r}^{1/2} U_i,$

$D_i = U_i^T K_{ih_r}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \eta_i, \quad \Gamma_i = O_t^{-1} + U_i^T K_{ih_r}^{1/2} R_i^{-1} K_{ih_r}^{1/2} U_i,$

$F_i = Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \Lambda_i K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} Q_i,$

$J_i = Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \Lambda_i K_{ih_f}^{1/2} R_i^{-1} K_{ih_r}^{1/2} U_i,$

$P_i = U_i^T K_{ih_r}^{1/2} R_i^{-1} K_{ih_f}^{1/2} \Lambda_i K_{ih_f}^{1/2} R_i^{-1} K_{ih_r}^{1/2} U_i,$

$E_i = Q_i^T K_{ih_f}^{1/2} R_i^{-1} K_{ih_f}^{1/2} Q_i.$

Then

$$
\begin{aligned}
E\left(L(\alpha)|\mathcal{T}\right) &= -\sum_{i=1}^{n} \{A_i - B_i \Gamma_i^{-1} D_i\}, \\
V\left(L(\alpha)|\mathcal{T}\right) &= \sum_{i=1}^{n} \left(F_i - 2B_i \Gamma_i^{-1} J_i^T + B_i \Gamma_i^{-1} P_i \Gamma_i^{-1} B_i^T\right), \\
\dot{L}(\alpha) &= \sum_{i=1}^{n} \left\{E_i - B_i \Gamma^{-1} B^T\right\},
\end{aligned}
$$

where $\Lambda_i(t) = Z_i^T \gamma_i(t,t) Z_i + R_i$.

Note that for any random variable A with finite first two moments, A denotes

$E(A) + O_p\left(\sqrt{\mathrm{Var}(A)}\right)$, which approximation leads to the following results.

$$\sum_{j=1}^{n_i} \frac{(t_{ij}-t)^a K^b_{ih_f}(t_{ij}-t)}{\sigma^2(t_{ij})}$$

$$= \frac{n_i h_f^{a-b+1}}{\sigma^2(t)} \left[f(t)\mu_a(b) + f(t)^{(1)}\mu_{a+1}(b)h_f + O_p\left(h_f^2 + 1/\sqrt{nh_f}\right)\right],$$

$$\sum_{j=1}^{n_i} \frac{(t_{ij}-t)^a K_{ih_r}(t_{ij}-t)}{\sigma^2(t_{ij})}$$

$$= \frac{n_i h_f^a C_h^a}{\sigma^2(t)} \left[f(t)\mu_a(1) + f(t)^{(1)}\mu_{a+1}(1)h_f C_h + O_p\left(h_f^2 + 1/\sqrt{nh_f}\right)\right],$$

$$\sum_{j=1}^{n_i} \frac{x_{ij,p}(t_{ij}-t)^a K_{ih_f}(t_{ij}-t)\eta^t_{ij}}{\sigma^2(t_{ij})}$$

$$= \frac{n_i h_f^{a+2} x_{it,p}}{\sigma^2(t)} \left[\frac{x_{it}^T\beta^{(2)}(t)f(t)}{2!}\mu_{a+2}(1)\right.$$

$$\left. + \left(\frac{x_{it}^T\beta^{(2)}(t)f^{(1)}(t)}{2!} + \frac{x_{it}^T\beta^{(3)}(t)f(t)}{3!}\right) h_f\mu_{a+3}(1) + O_p\left(h_f^2 + 1/\sqrt{nh_f}\right)\right],$$

$$\sum_{j=1}^{n_i} \frac{z_{ij}(t_{ij}-t)^a K^{1/2}_{ih_f}(t_{ij}-t)K^{1/2}_{ih_r}(t_{ij}-t)\eta^t_{ij}}{\sigma^2(t_{ij})}$$

$$= \frac{n_i z_{it} h_f^{a+1} C_h^{1/2}}{\sigma^2(t)} \left[x_{it}^T\beta^{(2)}(t)\frac{f(t)}{2!}\varphi_{a+2}(\tfrac{1}{2}, \tfrac{1}{2})\right.$$

$$\left. + \left(\frac{x_{it}^T\beta^{(2)}f^{(1)}(t)}{2!} + \frac{x_{it}^T\beta^{(3)f(t)}}{3!}\right) h_f\varphi_{a+3}(\tfrac{1}{2}, \tfrac{1}{2}) + O_p\left(h_f^2 + 1/\sqrt{nh_f}\right)\right],$$

$$\sum_{k\neq j}^{n_i}\ \sum_{j=1}^{n_i}\left(\frac{(t_{ij}-t)^a K^r_{ih_f}(t_{ij}-t)K^s_{ih_r}(t_{ij}-t)}{\sigma^2(t_{ij})}\right) z_{ij}\gamma(t_{ij},t_{ik})z_{ik}\left(\frac{(t_{ik}-t)^b K^u_{ih_f}(t_{ik}-t)K^v_{ih_r}(t_{ik}-t)}{\sigma^2(t_{ik})}\right)$$

$$= \frac{n_i(n_i-1)h_f^{2+a+b-r-u-s-v} z_{it}^2 C_h^{s+v}}{\sigma^4(t)}\left[f(t)^2\gamma(t,t)\varphi_a(r,s)\varphi_b(u,v)\right.$$

$$+ f(t)f^{(1)}(t)\gamma(t,t)h_f\left\{\varphi_{a+1}(r,s)\varphi_b(u,v)+\varphi_a(r,s)\varphi_{b+1}(u,v)\right\}$$

$$+ f^2(t)h_f\left\{\gamma_1^{(1)}(t,t)\varphi_{a+1}(r,s)\varphi_b(u,v)+\gamma_2^{(1)}(t,t)\varphi_a(r,s)\varphi_{b+1}(u,v)\right\}$$

$$\left.+O_p\left(h_f^2+1/\sqrt{nh_f}\right)\right],$$

where $\gamma_1^{(1)}(t,t) = \partial\gamma(x,y)/\partial x|_{x=t,y=t}$, $\gamma_2^{(1)}(t,t) = \partial\gamma(x,y)/\partial y|_{x=t,y=t}$, $\mu_a(b) = \int y^a K^b(y)dy$, $\varphi_a(p,q) = \int u^a K^p(u)K^q(cu)du$, $\varphi_a(b,0) = \mu_a(b)$, and $\tau_i(t) = Z_i^T(t)$ $\gamma(t,t)Z_i(t) + \sigma^2(t)$.

Thus, $\dot{L}(\alpha_t)$, $E\left(L(\alpha)|\mathcal{T}\right)$ and $V\left(L(\alpha)|\mathcal{T}\right)$ are as following:

$$\dot{L}\left(\alpha_t\right) = \sum_{i=1}^{n}\frac{n_i f(t)}{\sigma^2(t)}\left(\left(1-C_h\varphi_0^2\left(\frac{1}{2},\frac{1}{2}\right)\right)x_{it}x_{it}^T\begin{bmatrix} 1+O_p\left(\frac{1}{\sqrt{n_i h_f}}\right) & O_p\left(\sqrt{\frac{h_f}{n_i}}\right) \\ O_p\left(\sqrt{\frac{h_f}{n_i}}\right) & O_p\left(h_f^2\right)\end{bmatrix}\right),$$

$$E\left(L(\alpha)|\mathcal{T}\right) = \sum_{i=1}^{n}\frac{n_i^2 f^2(t)h^2}{\sigma^2(t)}\left(x_{it}^T\eta(t)^{(2)}\begin{bmatrix} 1+O_p\left(\frac{1}{\sqrt{n_i h_f}}\right) \\ O_p\left(h_f\right)\end{bmatrix}\right),$$

$$
\begin{aligned}
V\left(L(\alpha)|\mathcal{T}\right) &= \sum_{i=1}^{n} \frac{n_i^2 f^2(t)}{\sigma^4(t)} \Bigg(\Bigg\{ z_{it}^2 \left(1 - C_h \varphi_0^2 \left(\frac{1}{2}, \frac{1}{2}\right)\right)^2 x_{it} x_{it}^T \gamma(t,t) \\
&\quad + \frac{\mu_0(2) - 2C_h \varphi_0\left(\frac{1}{2}, \frac{1}{2}\right)\varphi_0\left(\frac{3}{2}, \frac{1}{2}\right) + C_h^2 \varphi_0^2\left(\frac{1}{2}, \frac{1}{2}\right)}{n_i h_f f(x)} x_{it} x_{it}^T \tau^2(t) \Bigg\} \\
&\quad \begin{bmatrix} 1 + O_p\left((n_i h_f)^{-1/2}\right) & O_p\left(h_f\right) \\ & \\ O_p\left(h_f^{1/2} n_i^{-1/2}\right) & O_p\left(h_f n_i^{-1}\right) \end{bmatrix} \Bigg),
\end{aligned}
$$

where x_{it} is a p-vector of coefficient at time t.

It follows that

$$
\begin{aligned}
\text{Bias}(\hat{\beta}) &= -e_{p,2p}^T \cdot L(\alpha)^{-1} E(\alpha) \\
&= M_1 \frac{h^2}{2} \left(\left(\sum_{i=1}^{n} n_i X_{it}^s \right)^{-1} \left(\sum_{i=1}^{n} n_i X_{i,t}^T \eta_{it}^{(2)} \right) O_p\left(1 + \sqrt{\frac{1}{mh}}\right) \right), \\
\text{Var}(\hat{\beta}) &= e_{p,2p}^T \cdot L(\alpha)^{-1} V(\alpha) \cdot L(\alpha)^{-1} e_{p,2p} \\
&\overset{(n_i \to \infty)}{=} \frac{\gamma(t,t)}{n} \left(\frac{1}{n} \sum_{i=1}^{n} z_{it}^2 X_{it}^{-s} + O_p\left(\frac{1}{\sqrt{mh}}\right) \right), \\
&\overset{(n_i < \infty)}{=} \gamma(t,t) \left(\sum_{i=1}^{n} n_i X_{it}^s \right)^{-1} \left(\sum_{i=1}^{n} n_i^2 z_{it}^2 X_{it}^s \right) \left(\sum_{i=1}^{n} n_i X_{it}^s \right)^{-1} \\
&\quad + \frac{M_2}{hf(x)} \tau^2(t) \left(\sum_{i=1}^{n} n_i X_{it}^s \right)^{-1} \left(\sum_{i=1}^{n} n_i X_{it}^s \right) \left(\sum_{i=1}^{n} n_i X_{it}^s \right)^{-1} \\
&\quad + O_p\left(\frac{1}{N\sqrt{Nh_f}}\right),
\end{aligned}
$$

where $e_{p,2p}$ is a $2p$ vector with 1 though p elements being 1, $M_1 = \frac{1 - C_h \varphi_0\left(\frac{1}{2},\frac{1}{2}\right)\varphi_2\left(\frac{1}{2},\frac{1}{2}\right)}{1 - C_h \varphi_0^2\left(\frac{1}{2},\frac{1}{2}\right)}$ and $M_2 = \frac{1 - 2C_h \varphi_0\left(\frac{1}{2},\frac{1}{2}\right)\varphi_0\left(\frac{3}{2},\frac{1}{2}\right) + C_h^2 \varphi_0\left(\frac{1}{2},\frac{1}{2}\right)\varphi_0(1,1)}{\left(1 - C_h \varphi_0^2\left(\frac{1}{2},\frac{1}{2}\right)\right)^2}$ are fixed and bounded for n_i and h_f.

Table 3.5. Finite sample performance of Indices at time 10, 25 and 40 for weakly correlated data

$\sigma_\varepsilon^2(t) = 2, \quad \sigma_u^2(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		n = 50						n = 100					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	h	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.3058		.3349		.2427		.3058		.2294		.2427	
10	1	.2989	.0036	.3258	.0044	.2364	.0027	.3037	.0017	.3323	.0022	.2411	.0014
	3	.3056	.0013	.3344	.0016	.2428	.0011	.3072	.0010	.3365	.0013	.2442	.0008
	6	.3058	.0014	.3349	.0016	.2431	.0011	.3061	.0010	.3354	.0012	.2434	.0008
THEORETICAL		.2905		.3187		.2294		.2905		.3187		.2294	
25	1	.2854	.0038	.3122	.0048	.2255	.0029	.2897	.0016	.3168	.0018	.2286	.0013
	3	.2910	.0020	.3184	.0022	.2300	.0015	.2889	.0011	.3170	.0012	.2283	.0008
	6	.2903	.0016	.3181	.0018	.2296	.0012	.2908	.0009	.3198	.0012	.2304	.0007
THEORETICAL		.2797		.3073		.2202		.2797		.3073		.2202	
40	1	.2735	.0036	.2993	.0046	.2152	.0027	.2777	.0014	.3042	.0018	.2183	.0011
	3	.2772	.0017	.3058	.0020	.2192	.0013	.2807	.0010	.3080	.0011	.2211	.0007
	6	.2799	.0015	.3082	.0017	.2212	.0011	.2799	.0008	.3071	.0009	.2204	.0006
THEORETICAL		.2925		.5175		.2311		.2925		.5175		.2311	
UCCC	1	.1500	.0209	.5006	.0028	.1318	.0104	.1515	.0210	.5054	.0016	.1331	.0100
	3	.1516	.0204	.5043	.0024	.1332	.0101	.1507	.0204	.5047	.0016	.1324	.0100
	6	.1515	.0204	.5046	.0022	.1331	.0100	.1509	.0204	.5055	.0016	.1327	.0100

Table 3.6. SDs, SEs and 95% CPs of indices at time 10, 25 and 40 for weakly correlated data

$\sigma^2_\varepsilon(t) = 2, \quad \sigma^2_u(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		$n = 50$								
		Intra-			Inter-			Total-		
t	h	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	1	.0590	.0591 (.0030)	.9818	.0660	.0652 (.0031)	.9734	.0521	.0519 (.0036)	.9810
	3	.0375	.0457 (.0031)	.9775	.0412	.0488 (.0029)	.9711	.0332	.0399 (.0034)	.9725
10	6	.0376	.0441 (.0030)	.9663	.0403	.0468 (.0029)	.9676	.0329	.0345 (.0034)	.9561
	1	.0618	.0586 (.0036)	.9820	.0618	.0648 (.0033)	.9813	.0541	.0520 (.0038)	.9812
	3	.0424	.0444 (.0034)	.9625	.0473	.0477 (.0034)	.9500	.0383	.0384 (.0034)	.9500
25	6	.0405	.0429 (.0032)	.9559	.0431	.0456 (.0031)	.9589	.0348	.0350 (.0035)	.9519
	1	.0599	.0584 (.0033)	.9745	.0644	.0646 (.0032)	.9760	.0521	.0503 (.0034)	.9768
	3	.0418	.0427 (.0032)	.9658	.0455	.0438 (.0032)	.9550	.0361	.0377 (.0034)	.9650
40	6	.0384	.0401 (.0034)	.9526	.0409	.0428 (.0033)	.9526	.0348	.0361 (.0035)	.9623

FCCC		$n = 100$								
		Intra-			Inter-			Total-		
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	1	.0412	.0417 (.0015)	.9757	.0466	.0462 (.0019)	.9718	.0368	.0368 (.0024)	.9721
10	3	.0317	.0319 (.0018)	.9535	.0345	.0345 (.0018)	.9504	.0279	.0281 (.0020)	.9652
	6	.0321	.0309 (.0018)	.9403	.0347	.0327 (.0018)	.9403	.0215	.0269 (.0021)	.9453
	1	.0394	.0404 (.0015)	.9704	.0427	.0460 (.0019)	.9724	.0338	.0362 (.0019)	.9747
25	3	.0327	.0314 (.0020)	.9419	.0341	.0337 (.0020)	.9419	.0283	.0271 (.0022)	.9419
	6	.0317	.0311 (.0021)	.9403	.0341	.0321 (.0020)	.9403	.0274	.0269 (.0022)	.9492
	1	.0396	.0411 (.0018)	.9747	.0419	.0454 (.0020)	.9704	.0325	.0346 (.0019)	.9712
40	3	.0309	.0309 (.0018)	.9501	.0333	.0332 (.0017)	.9459	.0266	.0265 (.0019)	.9507
	6	.0290	.0296 (.0019)	.9552	.0307	.0311 (.0018)	.9601	.0247	.0253 (.0020)	.9473

Table 3.7. Finite sample performance of Indices at time 10, 25 and 40 for moderately correlated data

$\sigma_\varepsilon^2(t) = 1, \quad \sigma_u^2(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		$n = 50$						$n = 100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	h	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.4684		.4135		.3349		.4684		.4135		.3349	
10	1	.4594	.0040	.4043	.0044	.3277	.0034	.4668	.0016	.4129	.0018	.3345	.0014
	3	.4637	.0021	.4086	.0022	.3312	.0018	.4698	.0012	.4157	.0013	.3371	.0011
	6	.4630	.0019	.4088	.0018	.3311	.0015	.4676	.0012	.4128	.0012	.3347	.0010
THEORETICAL		.4502		.3958		.3187		.4502		.3958		.3187	
25	1	.4423	.0038	.3885	.0039	.3131	.0030	.4476	.0015	.3934	.0016	.3169	.0013
	3	.4462	.0024	.3923	.0024	.3163	.0018	.4482	.0013	.3944	.0013	.3178	.0011
	6	.4451	.0021	.3912	.0020	.3151	.0016	.4482	.0013	.3947	.0012	.3180	.0010
THEORETICAL		.4371		.3832		.3073		.4371		.3832		.3073	
40	1	.4293	.0040	.3765	.0041	.3022	.0031	.4349	.0016	.3805	.0017	.3055	.0013
	3	.4312	.0024	.3795	.0023	.3043	.0018	.4382	.0013	.3839	.0012	.3085	.0010
	6	.4308	.0023	.3788	.0019	.3036	.0015	.4359	.0012	.3822	.0011	.3068	.0009
THEORETICAL		.4542		.5208		.3208		.4524		.5208		.3208	
UCCC	1	.1842	.0729	.5037	.0030	.1574	.0274	.1859	.0730	.5087	.0024	.1589	.0274
	3	.1840	.0728	.5041	.0025	.1572	.0274	.1837	.0729	.5043	.0024	.1570	.0274
	6	.1834	.0731	.5035	.0023	.1567	.0275	.1831	.0731	.5041	.0020	.1567	.0274

Table 3.8. SDs, SEs and 95% CPs of indices at time 10, 25 and 40 fo moderately correlated data

$\sigma^2_\varepsilon(t) = 1, \quad \sigma^2_u(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		$n = 50$								
		Intra-			Inter-			Total-		
t	h	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	1	.0625	.0601 (.0021)	.9617	.0656	.0619 (.0027)	.9410	.0577	.0547 (.0032)	.9440
	3	.0457	.0510 (.0018)	.9639	.0466	.0504 (.0023)	.9602	.0419	.0454 (.0030)	.9639
10	6	.0438	.0467 (.0019)	.9597	.0438	.0483 (.0021)	.9562	.0389	.0421 (.0026)	.9577
	1	.0611	.0603 (.0021)	.9683	.0618	.0604 (.0027)	.9619	.0555	.0533 (.0034)	.9670
	3	.0489	.0516 (.0020)	.9565	.0474	.0507 (.0026)	.9565	.0429	.0449 (.0032)	.9565
25	6	.0469	.0506 (.0019)	.9577	.0442	.0451 (.0022)	.9517	.0388	.0415 (.0029)	.9517
	1	.0624	.0605 (.0025)	.9645	.0631	.0617 (.0029)	.9651	.0555	.0533 (.0034)	.9651
	3	.0489	.0504 (.0021)	.9562	.0484	.0502 (.0027)	.9565	.0429	.0446 (.0034)	.9565
40	6	.0449	.0483 (.0020)	.9507	.0435	.0440 (.0025)	.9507	.0388	.0404 (.0029)	.9507

FCCC		$n = 100$								
		Intra-			Inter-			Total-		
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	1	.0406	.0415 (.0015)	.9721	.0429	.0441 (.0015)	.9701	.0370	.0389 (.0017)	.9702
10	3	.0353	.0364 (.0013)	.9485	.0364	.0362 (.0014)	.9485	.0326	.0329 (.0017)	.9485
	6	.0349	.0358 (.0013)	.9433	.0351	.0357 (.0016)	.9433	.0317	.0321 (.0017)	.9501
	1	.0392	.0416 (.0015)	.9701	.0405	.0421 (.0018)	.9711	.0361	.0385 (.0019)	.9661
25	3	.0361	.0363 (.0014)	.9485	.0361	.0357 (.0017)	.9538	.0322	.0320 (.0018)	.9485
	6	.0358	.0356 (.0014)	.9461	.0355	.0346 (.0018)	.9461	.0318	.0312 (.0018)	.9433
	1	.0395	.0407 (.0015)	.9661	.0416	.0427 (.0017)	.9702	.0359	.0370 (.0018)	.9621
40	3	.0358	.0362 (.0013)	.9538	.0338	.0354 (.0015)	.9538	.0310	.0315 (.0018)	.9538
	6	.0339	.0354 (.0014)	.9433	.0330	.0343 (.0011)	.9537	.0300	.0306 (.0015)	.9432

Table 3.9. Finite sample performance of Indices at time 10, 25 and 40 with 20% removed observations for each occasion under weak correlation

$\sigma_\varepsilon^2(t) = 2, \quad \sigma_u^2(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC		$n = 50$						$n = 100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	h	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.3058		.3349		.2427		.3058		.2294		.2478	
10	1	.2969	.0039	.3229	.0048	.2346	.0030	.3078	.0017	.3364	.0020	.2145	.0013
	3	.3043	.0016	.3327	.0019	.2416	.0013	.3088	.0012	.3383	.0012	.2457	.0008
	6	.3041	.0015	.3333	.0017	.2418	.0011	.3068	.0009	.3348	.0010	.2430	.0007
THEORETICAL		.2905		.3187		.2294		.2905		.3187		.2294	
25	1	.2838	.0037	.3097	.0044	.2236	.0027	.2899	.0015	.3163	.0017	.2283	.0011
	3	.2914	.0020	.3188	.0023	.2304	.0015	.2913	.0010	.3190	.0011	.2301	.0007
	6	.2879	.0017	.3159	.0019	.2277	.0012	.2893	.0008	.3180	.0010	.2290	.0006
THEORETICAL		.2797		.3073		.2202		.2797		.3073		.2202	
40	1	.2721	.0036	.2982	.0044	.2141	.0027	.2810	.0014	.3080	.0019	.2213	.0011
	3	.2766	.0020	.3053	.0023	.2188	.0018	.2817	.0010	.3087	.0011	.2218	.0007
	6	.2769	.0015	.3049	.0018	.2185	.0011	.2788	.0007	.3062	.0008	.2196	.0005
THEORETICAL		.2925		.5175		.2311		.2925		.5175		.2311	
UCCC	1	.1496	.0211	.4998	.0031	.1316	.0105	.1529	.0198	.5085	.0012	.1343	.0100
	3	.1513	.0205	.5035	.0025	.1329	.0101	.1518	.0201	.5063	.0015	.1334	.0100
	6	.1499	.0209	.5011	.0025	.1317	.0103	.1505	.0205	.5041	.0015	.1323	.0100

Table 3.10. SDs, SEs and 95% CPs of indices at time 10, 25 and 40 with 20% removed observations for each occasion under weak correlation

$\sigma_\varepsilon^2(t) = 2, \quad \sigma_u^2(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC, $n = 50$

t	h	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0616	.0591 (.0023)	.9681	.0685	.0650 (.0034)	.9713	.0510	.0517 (.0036)	.9658
	3	.0403	.0457 (.0029)	.9689	.0460	.0489 (.0028)	.9668	.0355	.0399 (.0033)	.9638
10	6	.0390	.0441 (.0030)	.9601	.0445	.0467 (.0028)	.9535	.0339	.0364 (.0033)	.9535
	1	.0601	.0584 (.0034)	.9644	.0660	.0646 (.0033)	.9634	.0519	.0507 (.0034)	.9634
	3	.0453	.0446 (.0033)	.9551	.0477	.0578 (.0033)	.9541	.0388	.0386 (.0031)	.9545
25	6	.0412	.0418 (.0030)	.9535	.0438	.0455 (.0030)	.9535	.0354	.0369 (.0032)	.9535
	1	.0591	.0582 (.0032)	.9694	.0660	.0607 (.0034)	.9634	.0513	.0502 (.0034)	.9649
	3	.0444	.0438 (.0031)	.9613	.0482	.0473 (.0031)	.9583	.0383	.0377 (.0031)	.9573
40	6	.0389	.0401 (.0030)	.9535	.0420	.0448 (.0032)	.9535	.0334	.0360 (.0030)	.9551

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0409	.0411 (.0016)	.9688	.0456	.0462 (.0018)	.9698	.0363	.0369 (.0019)	.9728
10	3	.0318	.0324 (.0017)	.9599	.0348	.0346 (.0018)	.9596	.0281	.0282 (.0021)	.9526
	6	.0303	.0313 (.0017)	.9462	.0331	.0331 (.0017)	.9534	.0266	.0271 (.0021)	.9534
	1	.0384	.0406 (.0020)	.9629	.0411	.0460 (.0021)	.9715	.0328	.0353 (.0018)	.9742
25	3	.0316	.0316 (.0020)	.9500	.0345	.0338 (.0021)	.9500	.0273	.0273 (.0020)	.9500
	6	.0295	.0303 (.0020)	.9532	.0317	.0322 (.0019)	.9534	.0255	.0261 (.0019)	.9534
	1	.0376	.0401 (.0019)	.9752	.0435	.0456 (.0019)	.9715	.0337	.0357 (.0018)	.9615
40	3	.0311	.0310 (.0019)	.9509	.0339	.0334 (.0018)	.9572	.0269	.0267 (.0012)	.9508
	6	.0274	.0394 (.0020)	.9523	.0291	.0301 (.0019)	.9531	.0234	.0344 (.0011)	.9504

Table 3.11. Finite sample performance of Indices at time 10, 25 and 40 with 20% removed observations for each occasion under moderate correlation

		$\sigma_\varepsilon^2(t) = 1, \quad \sigma_u^2(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$											
FCCC		$n = 50$						$n = 100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	h	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.4684		.4135		.3349		.4684		.4135		.3349	
10	1	.4573	.0040	.4020	.0044	.3257	.0033	.4670	.0017	.4129	.0019	.3345	.0015
	3	.4632	.0022	.4086	.0022	.3311	.0018	.4711	.0012	.4174	.0013	.3387	.0010
	6	.4579	.0024	.4040	.0022	.3268	.0019	.4637	.0012	.4093	.0012	.3313	.0010
THEORETICAL		.4502		.3958		.3187		.4502		.3958		.3187	
25	1	.4421	.0040	.3880	.0039	.3127	.0031	.4477	.0016	.3934	.0017	.3170	.0013
	3	.4467	.0024	.3930	.0024	.3168	.0018	.4509	.0014	.3968	.0013	.3200	.0011
	6	.4424	.0027	.3862	.0020	.3107	.0020	.4445	.0013	.3912	.0012	.3147	.0010
THEORETICAL		.4371		.3832		.3073		.4371		.3832		.3073	
40	1	.4279	.0040	.3748	.0041	.3007	.0031	.4352	.0016	.3805	.0017	.3056	.0013
	3	.4307	.0024	.3792	.0023	.3039	.0018	.4397	.0013	.3856	.0013	.3099	.0010
	6	.4258	.0025	.3741	.0024	.3009	.0019	.4315	.0012	.3778	.0011	.3049	.0009
THEORETICAL		.4542		.5208		.3208		.4524		.5208		.3208	
UCCC	1	.1837	.0739	.5031	.0031	.1570	.0276	.1861	.0713	.5090	.0015	.1591	.0265
	3	.1841	.0728	.5044	.0024	.1574	.0278	.1875	.0706	.5116	.0014	.1604	.0261
	6	.1813	.0741	.5005	.0029	.1551	.0285	.1837	.0731	.5040	.0015	.1562	.0274

Table 3.12. SDs, SEs and 95% CPs of indices at time 10, 25 and 40 with 20% removed observations for each occasion under moderate correlation

$\sigma_\varepsilon^2(t) = 1, \quad \sigma_u^2(t) = 1 \times f(t), \quad f(t) = \exp(-t/n_{ij})$

FCCC, $n = 50$

t	h	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0624	.0602 (.0022)	.9622	.0651	.0620 (.0027)	.9505	.0571	.0547 (.003)2	.9520
	3	.0459	.0508 (.0019)	.9658	.0464	.0503 (.0024)	.9573	.0418	.0434 (.0030)	.9608
10	6	.0484	.0511 (.0019)	.9377	.0472	.0500 (.0023)	.9331	.0428	.0454 (.0030)	.9435
	1	.0622	.0603 (.0021)	.9604	.0630	.0616 (.0027)	.9614	.0552	.0534 (.0032)	.9623
	3	.0494	.0506 (.0020)	.9573	.0479	.0502 (.0026)	.9573	.0432	.0453 (.0032)	.9487
25	6	.0514	.0507 (.0020)	.9336	.0497	.0491 (.0024)	.9336	.0445	.0441 (.0033)	.9336
	1	.0622	.0609 (.0029)	.9609	.0635	.0614 (.0029)	.9623	.0550	.0533 (.0033)	.9623
	3	.0484	.0511 (.0021)	.9487	.0484	.0500 (.0027)	.9513	.0428	.0445 (.0033)	.9487
40	6	.0494	.0504 (.0020)	.9336	.0481	.0487 (.0027)	.9336	.0426	.0435 (.0034)	.9336

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	1	.0413	.0422 (.0016)	.9671	.0436	.0440 (.0016)	.9675	.0384	.0390 (.0017)	.9713
10	3	.0350	.0367 (.0012)	.9516	.0356	.0364 (.0014)	.9561	.0321	.0332 (.0017)	.9516
	6	.0343	.0346 (.0015)	.9412	.0343	.0334 (.0015)	.9412	.0310	.0302 (.0014)	.9412
	1	.0398	.0416 (.0015)	.9703	.0407	.0419 (.0017)	.9658	.0356	.0385 (.0017)	.9769
25	3	.0373	.0366 (.0013)	.9516	.0367	.0360 (.0016)	.9516	.0329	.0324 (.0016)	.9516
	6	.0347	.0334 (.0014)	.9412	.0346	.0338 (.0019)	.9412	.0308	.0301 (.0016)	.9412
	1	.0396	.0417 (.0014)	.9712	.0416	.0437 (.0016)	.9725	.0360	.0370 (.0017)	.9664
40	3	.0364	.0364 (.0013)	.9507	.0359	.0356 (.0015)	.9516	.0319	.0318 (.0018)	.9355
	6	.0342	.0331 (.0019)	.9412	.0334	.0329 (.0017)	.9412	.0297	.0321 (.0019)	.9412

Chapter 4

Measurement agreement assessment for intensive longitudinal data based on partially mixed effects models

4.1 Introduction

Linear mixed models have been widely used in longitudinal studies due to their flexibility in terms of a subject-specific occasion and missingness. However, parametric models are likely to introduce approximation error because the parametric approach is not flexible enough to characterize curve features in ILD. Various nonparametric models have been proposed for longitudinal data analysis in order to relax the assumptions on parametric forms and allow flexible functional features in observations. On the other hand, fully nonparametric models may be too flexible to make statistical inference efficiently. The compromise between parametric models and nonparametric models is a semiparametric model. The most commonly used semiparametric regression model is a partially linear model (Härdle et al. (2000)). It avoids the curse of dimensionality by assuming some parametric structure and reduces model bias by introducing a smooth intercept function. For the jth observation of the ith subject, let Y_{ij} be the response variable, X_{ij} be a p-vector covariate, and t_{ij} be a scalar covariate, $j = 1, \cdots, n_i$ and $i = 1, \cdots, n$. The partially linear regression model is defined as follows:

$$Y_{ij} = \alpha(t_{ij}) + X_{ij}^T \beta + \varepsilon_{ij}, \qquad i = 1, \cdots, n, \; j = 1, \cdots, n_i \tag{4.1}$$

where $\alpha(\cdot)$ is an arbitrary smooth function with infinite dimension, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ are random errors with mean zero and finite variances, and $\beta = (\beta_1, \beta_2, \cdots, \beta_p)^T$ is an unknown regression coefficient vector.

The partially linear model has been well studied for independent data ($n_i = 1$) in the literature (Heckman, 1986; Speckman, 1988; Härdle et al., 2000). For repeated measurements ($n_i > 1$), Lin & Carroll (2001a,b) considered marginal models for categorical response through a kernel generalized estimation equation (GEE) under mean and variance specification. Fan & Li (2004) proposed two estimators; a different-based estimator (DBE) and profile least squares estimator for β. With an estimated coefficient $\hat{\beta}$, Fan & Li (2004) employed the local polynomial regression technique to estimate the nonparametric smooth intercept function.

As a natural extension of the partially linear model, a partially linear mixed effects model has been proposed in Zeger & Diggle (1994) to model individual random effects. The partially linear mixed model is defined as follows:

$$y_i(t_{ij}) = \alpha(t_{ij}) + x_{ij}^T \beta + u_i(t_{ij}) + \varepsilon_{ij}, \tag{4.2}$$

where $\alpha(\cdot)$ is a twice differentiable smooth function of time; $\beta = (\beta_1, \cdots, \beta_p)^T$ is a vector of unknown parameters; $u_i(t)$ are stationary random processes with mean 0 and covariance function $\gamma(u) = \sigma^2\rho(u : \theta)$; and ε_i are mutually independent measurement errors, each distributed as $N(0, \sigma_\varepsilon^2)$, uncorrelated with u_i.

The inclusion of a random effect inducing heterogeneity improves the efficiency of fixed-effect estimation and utilizes the random effects for prediction; thus, an alternative approach to the partially linear model for longitudinal data is to employ random individual effects separating from random error. The random component varies subject by subject, thereby accounting for a part of sources of heterogeneity in the population and estimating the variation by incorporating random effects on the scale of the linear predictor. Moyeed & Diggle (1994) proposed a partial residual approach to estimate $\alpha(\cdot)$ and β. They further investigated the rate of convergence for such estimators. Zhang et al. (1998) considered various stationary and nonstationary stochastic process to time-series within-subject correlation and adding time-invariant subject-specific covariates. Zeger & Diggle (1994) and Moyeed & Diggle (1994) set a function of serial correlation in a random process and utilized the correlation for the coefficient estimation by backfitting iteration. Fan & Li (2004) assumed the working independence between time points since the the serial correlation does not affect the consistency of coefficients. Several variations of the partially linear model and the partially linear mixed effects model and their estimation procedures can be found in a review paper by Fan & Li (2006), reference book by Wu & Zhang (2006) and references therein. In the recent literature, interest in time-varying covariance functions has surged. Fan et al. (2007) pointed out both efficiency for regression coefficients and predicting trajectories of individuals in longitudinal data are improved by involving the covariance function on estimating procedure. They proposed a semiparametric model for the covariance function of longitudinal data and developed an estimation procedure for a time-varying variance using kernel regression and a correlation matrix between time periods using quasi-likelihood. The primary interest of the aforementioned works is the estimation of $\alpha(\cdot)$ and β rather than the random effects. Interestingly, we found that the existing procedures perform badly for estimating the FCCC because the estimation of the FCCC requires accurate estimator of the variance components. Thus, we propose an estimation procedure for the purpose of estimating the FCCC.

We make the following contributions in this chapter.

- We propose a FCCC of measuring agreement among densely repeated measures taken by more than two methods based on partially linear mixed model.
- We propose a new estimating procedure for the partially linear mixed model for serially correlated data. It provides accurate estimates of variance components as well as coefficient estimates even under strong serial correlation.

This chapter is organized as follows. Section 4.2 introduces a partially linear mixed model for intensive longitudinal data along with its estimation procedure. We define the FCCC for ILD and propose an estimator for the FCCC in section 4.3. We assess the performance of the proposed estimator by Monte Carlo simulation in section 4.4. Discussion is given in section 4.5.

4.2 Partially linear mixed models

Suppose that we have a sample of n subjects. For the ith subject, the response variable is $y_i(t)$; the $n_{ij} \times p$ sized fixed covariate is X_{ij}; $n_{ij} \times q$ sized random covariate matrix is Z_{ij}, $i = 1, \cdots, n$, $j = 1, \cdots, n_i$, and $\sum_{i=1}^{n}\sum_{j=1}^{n_i} n_{ij} = N$.

The partially linear mixed model is defined as follows:

$$y_i(t_j) = \alpha(t_j) + X_{ij}^T \beta + Z_{ij}^T u_i + \varepsilon_i(t_j), \tag{4.3}$$

where $\alpha(\cdot)$ is a smooth function with a second derivative and β is a p-vector of unknown parameters. u_i is a q-vector of random effects with mean 0 and variance G, ε_i is a measurement error with mean zero, finite time varying variance $R_i(t) = E\left(\varepsilon_i(t)\varepsilon_i^T(t)\right)$, the serial correlation $\rho(t,s)$ only depends on time and is uncorrelated with u_i. t indexes time in the longitudinal data. The true variance of $y_i(t)$ is $\Sigma_i(t) = Z_i^T G Z_i + R_i(t)$.

In the semiparametric literature, two approaches have been applied to implement nonparametric function and parametric estimates simultaneously: backfitting and profiling. Both involve some smoothing techniques differing in the way they treat the nonparametric function. Zeger & Diggle (1994) and Moyeed & Diggle (1994) applied the iterative backfitting approach. Fan & Li (2004) employed profile-least squares method. These popular approaches provide comparable estimates for unknown function and regression coefficients. At first, we utilized a backfitting algorithm to estimate both time-varying and time-invariant components in model (4.3) simultaneously; $\alpha(\cdot)$ and $\sigma_\varepsilon^2(\cdot)$ were estimated by a local linear regression; the estimators of a linear mixed model were adopted as a sample counterpart of β and G. The algorithm provides unbiased estimates of β for any degree of the serial correlation but biased estimates for $\sigma_\varepsilon^2(\cdot)$, G and $\alpha(\cdot)$. If there is a positive correlation in repeated measures over time, the measurements will look more similar other than they really are. The negative correlations make it seem like there is more variation in the repeated measures than there really is. For example, in a discrete time-stationary process X_t with mean zero and covariance

$\rho_{t-s}\sigma^2$ between X_t and X_s for $s, t = 1, \cdots, N$, the variance of mean of X_t over time is

$$V(\bar{X}) = \frac{1}{N}\left(\sigma^2 + 2\sum_{\tau=1}^{N-1}\rho_\tau\left(1 - \frac{\tau}{N}\right)\right), \tag{4.4}$$

where $\tau = i - j$. Ignoring the serial correlation for estimation σ^2 may result in an underestimating when $\rho > 0$. Similarly, the ignorance of the negative serial correlation may lead to an overestimate of the true variance. Therefore, a local linear estimator, locally weighted average, of residuals can underestimate (overestimate) $\sigma^2_\varepsilon(t)$ when $\varepsilon_i(t_1), \ldots, \varepsilon_i(t_L)$ are positively (negatively) correlated and its behaviour affects the estimate of G. Thus, we propose an estimating procedure to relieve the influence of the serial correlation on the estimation of $\alpha(\cdot)$, β and G and yield accurate $\sigma^2_\varepsilon(\cdot)$ even under strong correlation.

Define $\bar{y}$, $\bar{X}$ and $\bar{Z}$ to be the average of the time-sequence of response variables and fixed and random covariates over time, respectively. Hence, $\bar{y}_i$ can be a vector if there is a categorical covariate with more than one level and each level is observed repeatedly over time. In the cross-over experiment for clinical trials, for example, subjects receive a sequence of different treatments after set number of days of wash-out period. Then $\bar{y}_i$ is the mean of response variables for each treatment. $\bar{y}_i$ is normally distributed with mean $\bar{\alpha} + \bar{X}_i\beta$ and variance $\bar{Z}_i G_i \bar{Z}_i + \tilde{R}_i$ where $\tilde{R}_i = E\left(\bar{\varepsilon}_i(\cdot)\bar{\varepsilon}_i^T(\cdot)\right)$. For densely repeated data, the serial correlation may be relieved by taking the averages of a large number of replicates across time points while keeping the impact of time-invariant parameters. Thus, it is reliable to use the estimates $\hat{\beta}$ and $\hat{G}$ of the parametric linear mixed model using $\bar{y}$, $\bar{X}$ and $\bar{Z}$. Denote the distinct event times by $t_{(1)} < \cdots < t_{(L)}$ and the number of events at time $t_{(l)}$ by d_l, $d_l = \sum_i^n\sum_j^{n_i} I(t_{ij} = t_{(l)})$ for l, $l = 1, \cdots, L$. Given the estimates of $\widehat{\beta}$ and $\widehat{u}_i$, define $\tilde{y}_i = y_i - x_i\widehat{\beta} - z_i\widehat{u}_i$ at each distinct time point l, $l = 1, \cdots, L$. The model (4.3) with $\tilde{y}_i$ is written as

$$\tilde{y}_i(t_l) = \alpha(t_l) + \varepsilon_i(t_l). \tag{4.5}$$

Equation (4.5) can be treated as a typical nonparametric regression problem if random errors have negligible serial correlation. However, kernel estimates of $\alpha(\cdot)$ are undersmoothed for the mild positive serial correlation as shown by Altman (1990), and the kernel estimates of $\sigma^2_\varepsilon(\cdot)$ proposed by Fan & Li (2004) are considerably underestimated as mentioned above. Altman (1990) overcame the biased estimation of the nonparametric mean function in correlated data by new optimal bandwidth selection of the kernel smoother and showed that kernel smoother with adjusted bandwidth performs well for correlated data unless the correlations are over a long term. Since densely repeated measurements may be highly correlated over many successive measurements, we avoid a smoothing method and utilize point-wise estimation for $\alpha(\cdot)$. The estimates are shown below:

$$\hat{\alpha}(t_l) = \sum_i \sum_k \tilde{y}_i(t_{ij}) I\left(t_{ij} = t_l\right)/d_l. \tag{4.6}$$

This implies that all subjects were observed at the same t_l. In practice, different subjects may be observed at different time points. In such a situation, one may use interpolation to compute the sample means.

As for $\sigma^2_\varepsilon(t)$ estimation, the incorporation of the serial correlation structure may provide accurate residual variance estimates. However, diverse serial covariances are possible in practice and it is difficult to find the exact covariance structure between observations taken from different time points. Our strategy is block-wise estimation instead of global estimation; we fit the linear mixed model with consecutive observations in a proper time window at time t, $y_i(t-h), y_i(t-(h+1)), \ldots, y_i(t+h)$, by assuming auto-correlated errors. Because the serial correlation typically relies on the time separation between observations, if the block size is not large but enough to yield correlation structure, autoregressive (AR) correlation assumption may be close to the true correlation coefficient even though overall serial correlation structure is complicated. Assume $\varepsilon_{i;t} \sim \left(0,\ \lambda_{i;t}^{1/2} R_{i;t} \lambda_{i;t}^{1/2}\right)$ where $\lambda_{i;t}$ is the AR serial correlation coefficient at times t, $R_{i;t} = R_i I(y_i \in [t-h,\ t+h])$ and $2h$ is a block size. Because AR correlation function $\lambda_{i;t}$ is positive-definite, we transform data with an invertible square-root correlation function with inverse then fit linear mixed model with

$$\tilde{y}_{i;t} = \lambda_{i;t}^{-1/2} y_i I_{2h}, \qquad \tilde{X}_i = \lambda_{i;t}^{-1/2} X_i I_{2h}, \qquad \tilde{Z}_i = \lambda_{i;t}^{-1/2} Z_i I_{2h},$$

where $I_{2h} = I(y_i \in [t-h,\ t+h])$.

The summary of the estimation procedure follows below:

Step 1 Average the consecutive observed values of response variables, fixed and random covariates over time, $\bar{y}_i$, $\bar{X}_i$ and $\bar{Z}_i$.

Step 2 Estimate time-invariant parameters: coefficients and variances of random effect via linear mixed models with average variables in Step 1.

Step 3 Calculate $\hat{\alpha}(t)$ by point-wise estimation of $\tilde{y}_i(t_l)$,

$$\tilde{y}_i(t_l) = y_i(t_l) - x_i^T \widehat{\beta} - z_i^T \hat{u}_i = \alpha(t_l) + \varepsilon_i(t_l),$$

for the distinct time points $l = 1, \cdots, L$. Suppose that $d_l = \sum_i^n \sum_j^{n_i} I(t_{ij} = t_l)$,

$$\hat{\alpha}(t_l) = \sum_i \sum_j \tilde{y}_i(t_{ij}) I\left(t_{ij} = t_l\right)/d_l \qquad \forall l,\ l = 1, \cdots, L.$$

Step 4 Obtain $\hat{\sigma}_\varepsilon^2(t)$ through the estimation of the linear mixed model using the observations within the local neighborhood of t taking AR serial correlation.

Existing software such as R function LME or SAS PROC MIXED may be used to fit linear mixed model with AR correlation (Pinheiro & Bates, 2009). The estimating procedure is simple and easy to evaluate with statistical software and provides accurate estimates in terms of small MSE and 95% coverage probability. The estimators of time-constant coefficients in model (4.3) have the same asymptotic behavior to the estimators of parametric linear mixed models.

4.3 Functional type of Concordance Correlation Coefficient

The Concordance Correlation Coefficient (CCC) is a popular index of agreement between observations made on a continuous scale of two methods. The pairwise CCC assesses the strength of agreement by the expected value of the squared difference and yield two measures: how far the linear relationship of two methods deviated from the identity line (45°), and how far each observation deviates from the fitted line to the data (Lin, 1989). Lin et al. (2007) presented the unified approach of concordance correlation coefficient (UCCC) for repeated measures data. The indices are used to measure inter, intra, total agreement expressed as a function of variance components extracted from a two-way mixed model.

The advantage of UCCC is that the intra-subject correlation is induced by a random (subject) effect without assuming correlation structure, and indices are easily adjusted by potentially confounding effects. The model assumes, however, stable variations as well as time-invariant overall mean over study duration. In longitudinal studies, it is common that variance and overall mean change over time. Thus, allowing time-varying overall mean (i.e., time-varying intercept function) and time-varying variance of the random errors may significantly reduce model bias and yield appropriate measurement agreement indices for longitudinal data.

We shall extend UCCC and propose FCCC for ILD using the partially linear mixed model, which provides variance components for constructing the index of agreement. FCCC is motivated by Li & Chow (2005), which introduced the concept of reproducibility for functional data and proposed functional type weighted CCC for an overall agreement of paired repeated measures. Li & Chow (2005) the approached measurement agreements problem in the stochastic perspective rather than a multi-dimensional vector perspective. This allows us to account for the structure of functional data in a proper way, although it is limited to paired data without covariates. Here we extend their ideas and propose time-varying indices.

FCCC is a covariance-based index relying on the variance components in the partially linear mixed model. To improve the accuracy of FCCC, the potential variance (i.e., gender, race, age), which may induce an increase of the variance estimates in FCCC, have to be taken into account (Carrasco & Jover, 2003). To achieve this objective, the partially linear mixed model includes these covariates; then, FCCC may have a different form according to the potential covariates in the partially linear mixed model. As an example, we present FCCC based on the two-way mixed model with one fixed effect, one random effect and their interaction effect for an analogy to the two-way mixed model proposed by Lin et al. (2007).

We will explore other cases in the simulation study.

The partially two-way mixed model is defined as follows:

$$y_{ij}(t_l) = \mu(t_l) + \alpha_i + \beta_j + \eta_{ij} + \varepsilon_{ij}(t_l), \tag{4.7}$$

where $y_{ij}(t_l)$ is the observed response variable of subject i by method j at time t_l, $i = 1, \cdots, n$, $j = 1, \cdots, J$ and $l = 1, \cdots, m$. $\mu(\cdot)$ is a baseline curve; β_j is a jth method effect; α_i is an i.i.d. subject specific deviation for the ith subject with mean zero and variance σ_α^2; η_{ij} is the method deviation by subjects with mean zero and variance σ_η^2; $\varepsilon_{ij}(t_l)$ is a measurement error of realization from a Gaussian process with mean zero and finite time-varying variance $R_i(t) = E\left(\varepsilon_i(t)\varepsilon_i^T(t)\right)$. Three random variables α_i, η_{ij} and $\varepsilon_{ij}(t)$ are mutually independent. Define $\sigma_\beta^2 = \frac{\sum\left(\beta_j - \bar{\beta}_.\right)^2}{k-1}$, a between-method variance.

For each method, assume that there are K repeated measures on a subject at each time point. The intra FCCC between k and k', $k = 1, \cdots, K$ readings is defined for the method j:

$$\begin{aligned} \text{FCCC}_{\text{intra}}(t) &= \text{FCCC}_{\text{intra};j}(t) \\ &= 1 - \frac{E\left[W_{i;j} | X_i, Z_i\right]}{E\left[W_{i;j} | \text{all elements in } y_{ijk}(t), \cdots, y_{ijK}(t) \text{ are uncorrelated}, X_i, Z_i\right]} \\ &= \frac{\sigma_\alpha^2 + \sigma_\eta^2}{\sigma_\alpha^2 + \sigma_\eta^2 + \sigma_\varepsilon^2(t)}, \end{aligned} \tag{4.8}$$

where $W_{i;j} = \sum_{k=1}^{K} \left(y_{ijk}(t) - \bar{y}_{ij\cdot}(t)\right)^2$.

Because the $\text{FCCC}_{\text{intra};j}$ does not depend on j, it represents intra-method agreement for any method. The expected value of the observations taken by a method is the same; thus $\text{FCCC}_{\text{intra}}$ has perfect accuracy and measures how strongly repeated measurements resemble each other, which indicates the reproducibility of a method. $\text{FCCC}_{\text{intra}}$ is a time-varying index as long as a method takes more than one observation at time t on the same subject, or there are other cross-effects to the method.

The inter-method index is used to measure the agreement between methods eliminating any internal error of a method by the average of multiple replications for any two method j and j'.

$$\begin{aligned}
\text{FCCC}_{\text{inter}}(t) &= \text{FCCC}_{\text{inter};j,j'}(t) \\
&= 1 - \frac{E\left[W_{i;j,j'}|X_i, Z_i\right]}{E\left[W_{i;j,j'}|\bar{y}_{i1\cdot}(t), \cdots, \bar{y}_{iJ\cdot}(t) \text{ are uncorrelated}, X_i, Z_i\right]} \\
&= \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2(t)/K},
\end{aligned} \tag{4.9}$$

where $W_{i;j,j'} = \sum_{j=1}^{J-1} \sum_{j'=j+1}^{J} \left(\bar{y}_{ij\cdot}(t) - \bar{y}_{ij'\cdot}(t)\right)^2$.
Since the $\text{FCCC}_{\text{inter};j,j'}$ does not depend on j and j', the $\text{FCCC}_{\text{inter}}$ represents inter-method agreement, and helps to track the degree of the agreement between methods by indicating how close means of different methods are over time. The $\text{FCCC}_{\text{inter}}$ measures the lack of agreement using two meaningful statistics: precision

$$\rho(t) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2 + \sigma_\varepsilon^2(t)/K},$$

represents the consistency between the average measurements of methods at time t; accuracy

$$\chi(t) = \frac{\sigma_\alpha^2 + \sigma_\eta^2 + \sigma_\varepsilon^2(t)/K}{\sigma_\alpha^2 + \sigma_\eta^2 + \sigma_\beta^2 + \sigma_\varepsilon^2(t)/K},$$

is the measurement of how close the average measurements of methods are at time t.

Note that in the comparison study, the number of repeated measurements K may be the same between methods at the same occasion whereas the number of repeated measurements can be different by occasion, $K = K(t)$ due to skipping the scheduled time set.

The total-method index for the methods j and j' is used to quantify the agreement based on individual reading k of (i, j) combinations defined as

$$\begin{aligned}\text{FCCC}_{\text{total}}(t) &= \text{FCCC}_{\text{total};j,j'}(t) \\ &= 1 - \frac{E\left[W_{ik;j,j'}|X_i, Z_i\right]}{E\left[W_{ik;j,j'}|y_{i1k}(t), \cdots, y_{iJk}(t) \text{ are uncorrelated}|X_i, Z_i\right]} \\ &= \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2(t)}, \end{aligned} \tag{4.10}$$

where $W_{ik;j,j'} = \sum_{j=1}^{J-1} \sum_{j'=j+1}^{J} \left(y_{ijk}(t) - y_{ij'k}(t)\right)^2$.

The $\text{FCCC}_{\text{total}}$ is not changed by j, and it represents an overall inter-method agreement profile based on individual measurements. It is also composed of precision and accuracy. The $\text{FCCC}_{\text{total}}$ equals to the $\text{FCCC}_{\text{inter}}$ when there is neither replication at each occasion nor other cross effect to the method effect. Note that between-subject variability is assumed to be similar across the range of covariate values and across time. Because FCCC relies on between-subject variability, if it changes over the range of covariate variables and over time, it has to be considered in model (4.7).

FCCCs separate the degree of agreement between measures from the time effect so that FCCC may enhance detection of agreement between methods especially for curve data. UCCCs mixes the source of disagreement, between-method difference and time trend in observations. It is useful to calculate pairwise FCCC of multiple methods to investigate the cause of disagreement when the agreement of multiple methods is unsatisfactory based on inter FCCC. These FCCCs are an extension of UCCCs proposed by Lin et al. (2007), and reduce to UCCCs if the time effect is not significant and repeated measures are taken in quick succession. Note that FCCC provides only inter-method agreement if there is no effect crossed with method effect. This is the cost to extract the time effect and researchers may choose between the detail in source of disagreement and time trend of index on the basis of data features.

It is natural to use the sample counterparts of parameters $\hat{\beta}$, $\hat{\sigma}_u^2$, $\hat{\sigma}_\eta^2$ and $\hat{\sigma}_\varepsilon^2(t)$ proposed in section 4.2 to estimate the FCCCs in (4.8), (4.9) and (4.10). For n independent subjects, the estimators of the FCCCs are given as

$$\begin{aligned}
\widehat{\text{FCCC}}_{\text{intra}}(t) &= \frac{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\eta^2}{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\eta^2 + \hat{\sigma}_\varepsilon^2(t)}, \\
\widehat{\text{FCCC}}_{\text{inter}}(t) &= \frac{\hat{\sigma}_\alpha^2}{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\beta^2 + \hat{\sigma}_\eta^2 + \hat{\sigma}_\varepsilon^2(t)/K}, \\
\widehat{\text{FCCC}}_{\text{total}}(t) &= \frac{\hat{\sigma}_\alpha^2}{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\beta^2 + \hat{\sigma}_\eta^2 + \hat{\sigma}_\varepsilon^2(t)},
\end{aligned}$$

when there are K repeated measures for each method on a subject at each time point. We evaluate performance of the index estimates by examining MSE in Chapter 5.

All three time-varying indices, $\text{FCCC}_{\text{Intra}}(t)$, $\text{FCCC}_{\text{Inter}}(t)$ and $\text{FCCC}_{\text{Total}}(t)$ can be aggregated over time into one number. Let $w(\cdot)$ be a nonnegative weight function with $\int w(t)dt = 1$. Define

$$\begin{aligned}
\text{FCCC}_{\text{total}} &= \int \text{FCCC}_{\text{total}}(t)w(t)dt, \\
\text{FCCC}_{\text{inter}} &= \int \text{FCCC}_{\text{inter}}(t)w(t)dt, \\
\text{FCCC}_{\text{intra}} &= \int \text{FCCC}_{\text{intra}}(t)w(t)dt.
\end{aligned} \tag{4.11}$$

If there is prior information on the importance of observing occasions, one may utilize the information as a weight function. Otherwise, we may set $w(t)$ to be proportional to 1.

4.3.1 Parametric Bootstrap for Indices

It is challenging to derive the asymptotic distribution of the proposed index estimators. In this paper, we propose using the bootstrap to estimate the finite sample distribution of the proposed estimators. This enables us to make statistical inference on the proposed indices.

The bootstrap method is a popular and powerful alternative technique to construct confidence intervals of parameters of interest when the sampling distribution of the corresponding estimator is impossible to derive or too complicated to be used. It is also an appropriate way to check the reliability of estimates. A number of bootstrap methods can be applied to generate the data: sampling from actual data (nonparametric bootstrap), sampling from the fitted model distribution (fully parametric bootstrap) or resampling with replacement residuals after fitting expected part in the model (residual/semi-parametric bootstrap). In a linear model for longitudinal data, nonparametric bootstrap randomly and independently picks subjects, the vector of the response variables, and the associated covariates, with equal mass. No assumptions are made about the parameters of the population; this allows flexibility if the sample adequately reflects the population with sufficient sample size. If not, unrepresentative sets of samples may result in biased estimates. The residual bootstrap fits the model first and bootstraps the residuals with associated covariates. It is flexible in terms of sampling without distribution assumption. However, we found through our simulation study that bootstrapping variances are much smaller than the true variances, even for the data with no serial correlation. This is consistent with findings in Morris (2002), in which the author demonstrated that semiparametric bootstraps consistently underestimate the variation in finite samples for mixed effects models. Morris (2002) also found that the coverage probability decreases as the number of repeated measures and sample size increase. Therefore, residual bootstrap may be inappropriate for ILD. In this paper, we consider the parametric bootstrap to generate random samples. Specifically, we fit model (4.3) to the observed data, draw bootstrap samples from the specified model with fitted parameters, and refit the model with the simulated sample. Because the serial correlation considerably affects coverage probability of indices, coverage probability decrease with increasing the degree of correlation. We employ AR correlation in the resampling procedure in order to enhance the representativeness of the bootstrap sample about collected data. The algorithm for the parametric bootstrap for correlated data is as follows:

1. Fit (4.3) following the four steps described above and obtain $\widehat{\alpha}(\cdot)$, $\widehat{\beta}$, $\widehat{\sigma}_u^2$, $\widehat{\sigma}_e^2(\cdot)$ and $\hat{\lambda}$.

2. Draw n subjects $u_1, \ldots, u_n$ from $\mathrm{N}(0, \widehat{\sigma}_u^2)$ and draw N total observations $\varepsilon_i(t_1), \ldots, \varepsilon_i(t_{n_{ij}})$ from $\mathrm{N}\Big(0, \hat{\lambda}_i^{1/2} \hat{R}_i \hat{\lambda}_i^{1/2}\Big)$, $\hat{R}_i = \mathrm{diag}\big(\hat{\sigma}_\varepsilon^2(t_{i1}), \cdots, \hat{\sigma}_\varepsilon^2(t_{in_{ij}})\big)$ for $i = 1, \ldots, n$, $N = \sum_i n_i$.

3. Calculate bootstrap responses $\tilde{y}_i$ as $\tilde{y}_i(t) = \widehat{\alpha}(t) + x_i^T \widehat{\beta} + z_i^T u_i + \varepsilon_i(t)$, $i = 1, \cdots, n$.

4. Refit (4.3) with $\tilde{y}_i(t)$ and calculate FCCCs via bootstrap sample coefficients and variance components.

5. Repeat 2-4 B times.

The bootstrap indices can be used to make inferences about FCCCs.

4.4 Simulation study

We conducted a simulation study to investigate the finite sample performance of the estimation procedure and proposed agreement indices. All simulation studies were conducted in R. We generated simulation data from the partially linear mixed model with two categorical fixed effects and a subject random effect.
The simulation model is designed as follows: $i = 1, \cdots, n,\ j = 1, 2,\ k = 1, 2, 3,\ l = 1, \cdots, 50$

$$Y_{ijk}(t) = \alpha(t_l) + u_i + \beta_{1j} + \beta_{2k} + \gamma_{ik} + \varepsilon_{ijk}(t_l). \tag{4.12}$$

The random process $\varepsilon(t)$ in model (4.12) is taken to be a Gaussian process with mean 0, variance function

$$\sigma_\varepsilon^2(t) = a \times \left(.9 + .1 \sin(2\pi t/50)^2)\right),$$

and covariance with AR(1) serial correlation

$$E\varepsilon(s)\varepsilon^T(t) = \rho^{|t-s|}\sigma_\varepsilon(t)\sigma_\varepsilon(s),$$

for $s \neq t$. We consider three values ρ, 0, .5, .9 which correspond to non, mild and strong serial correlation. The overall mean function $\alpha(t)$ is defined as

$$\alpha(t) = 20 \times (\cos(2t/100 + 2) + 4).$$

We let the fixed effects β_1 and β_2 be two- and three-dimensional,

$$\beta_1 = (5, 20)^T, \qquad \beta_2 = (2.2,\ 2.3,\ 2.5)^T.$$

We generate the u_i and η_{ik} from Gaussian distributions with means 0 and variances σ_u^2 and σ_g^2, respectively, and consider three pairs of $\left(\sigma_u^2,\ \sigma_g^2,\ a\right)$ as (1, 1, 1), (1, 1, 2), and (.5, .5, 1) where the second set has larger time-varying variation than the first set and the third set has half the noise than the first set. Three sets enable to investigate the impact of the degree of variability in the data on the performance of FCCCs.

In our simulation, we consider two schemes to generate the times.

Scheme I: We set $t = 1, \cdots, 50$.

Scheme II: We generate the set of time points in the following way: each subject has a set of time points, $\{1, 2, ..., 50\}$, and each observation time point except the first time has a 20% probability of being skipped.

We conduct 1000 Monte Carlo simulations. For each simulation, we generate 500 bootstrap samples in order to construct the confidence intervals. For each simulation study, mean and MSE of indices are calculated. The means are represented by the column labeled Mean in the tables. For each index, an estimate of the mean square error (MSE)

$$E||\hat{f}(t) - f(t)||^2,$$

is estimated by

$$\frac{1}{1000}\sum_{t=1}^{1000}\left\{\widehat{f}(t) - f(t)\right\}^2,$$

and listed as the column labeled MSE in the tables. The standard deviation (SD) indicates the standard deviation of the 1000 estimates of an index. Standard error (SE) indicates the mean of 1000 standard errors of indices calculated by a 500 bootstrap sample. In addition, for each index the 95% coverage probability (CP) is calculated using a bootstrap percentile,

$$\frac{1}{1000}\sum_{t=1}^{1000} I\left\{\delta \in \left[\hat{\delta}_{(0.025)},\ \hat{\delta}_{(0.975)}\right]\right\},$$

where δ is true index and $\hat{\delta}_{(a)}$ is the a% percentile among an index estimates of bootstrap samples. For comparison, the UCCC is provided where estimates of variance components in UCCCs are evaluated by the mixed effects model (Barnhart et al., 2005, 2007). We explore the sensitivity of indices with respect to the magnitude of variability and the amount of serial correlation for both balanced longitudinal data and incomplete longitudinal data. We also investigate the overall performance of estimation procedures and the bootstrap variance of indices by CP.

The UCCCs and FCCCs of β_2 in model (4.12) defined as

$$\begin{aligned}
UCCC_{\text{intra}} &= \frac{\sigma_u^2 + \sigma_g^2}{\sigma_u^2 + \sigma_g^2 + \sigma_\varepsilon^2}, \\
UCCC_{\text{inter}} &= \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\gamma^2 + \sigma_g^2 + \sigma_\varepsilon^2/100}, \\
UCCC_{\text{total}} &= \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\gamma^2 + \sigma_g^2 + \sigma_\varepsilon^2},
\end{aligned} \tag{4.13}$$

where $\sigma_\gamma^2 = \frac{1}{6}\sum_{k=1}^{2}\sum_{k'=2}^{3}(\beta_{2,k} - \beta_{2,k'})^2$.

$$\begin{aligned} FCCC_{\text{intra}} &= \frac{\sigma_a^2 + \sigma_g^2}{\sigma_a^2 + \sigma_g^2 + \sigma_\varepsilon^2(t)}, \\ FCCC_{\text{inter}} &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\gamma^2 + \sigma_g^2 + \sigma_\varepsilon^2(t)/2}, \\ FCCC_{\text{total}} &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\gamma^2 + \sigma_g^2 + \sigma_\varepsilon^2(t)}. \end{aligned} \tag{4.14}$$

4.4.1 Time Scheme I

In this section, we present the finite-sample performance with observed time following scheme I, $t = 1, 2, \cdots, 50$. Figure 4.1 and 4.2 display the performance of model (4.12) and the linear mixed model when $\sigma_u^2 = 1, \sigma_g^2 = 1$, $\sigma_\varepsilon^2(t) = 2 \times (.9 + .1\sin(2\pi t/50)^2)$ and $\rho = 0.5$.

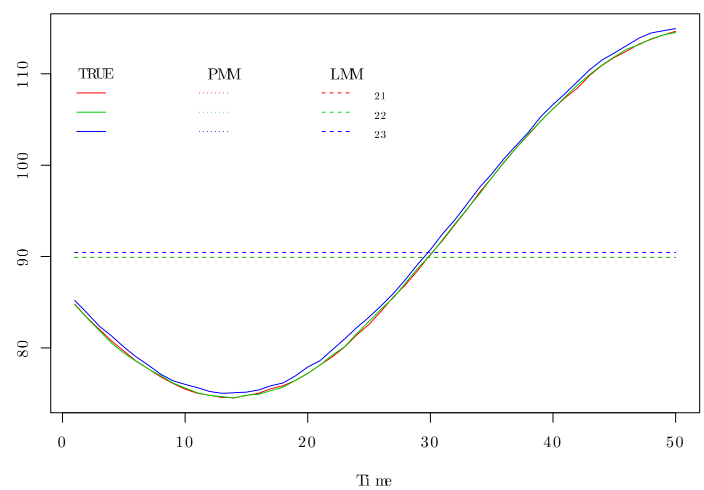

Figure 4.1. Plots for estimated curves at $\sigma_u^2 = 1, \sigma_g^2 = 1$, $\sigma_\varepsilon^2(t) = 2 \times (.9 + .1\sin(2\pi t/50)^2)$ and $\rho = 0.5$. For $k = 1, 2, 3$, $\mu_k(t) = \alpha(t) + u. + \beta_{1\cdot} + \beta_{2k} + \gamma_{\cdot k}$ (solid curves), $\hat{\mu}_k(t)$ (dotted curves) and the estimates (dashed lines) of the corresponding linear mixed model are depicted.

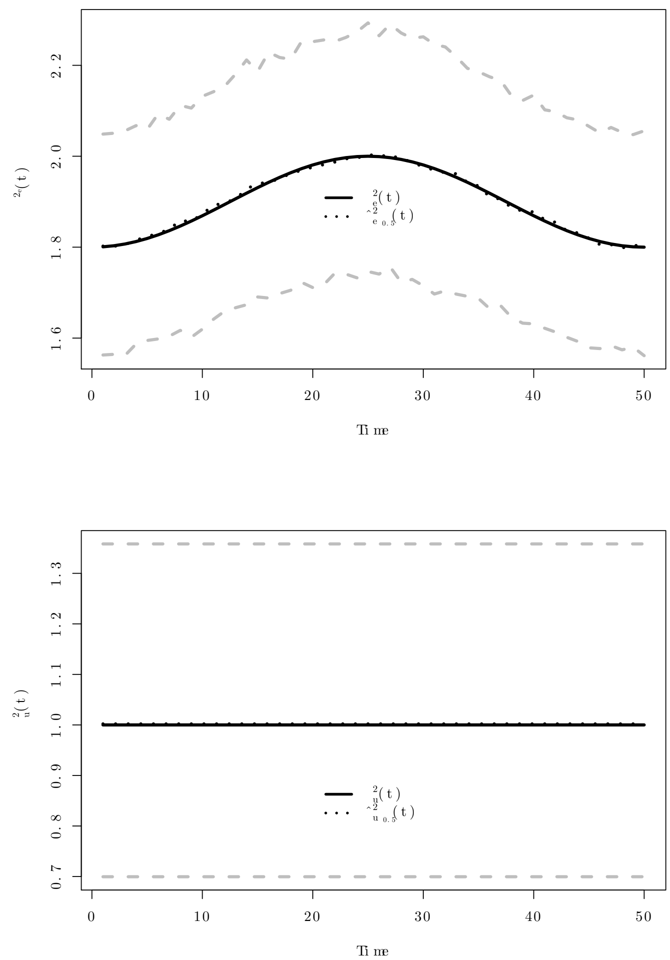

Figure 4.2. Plots for variance components at $\sigma_u^2 = 1, \sigma_g^2 = 1$, $\sigma_\varepsilon^2(t) = 2 \times (.9 + .1\sin(2\pi t/50)^2)$, and $\rho = 0.5$. In the upper panel, σ_ε^2 (solid) and $\hat{\sigma}_\varepsilon^2$ (dotted curves) with its 95% confidence interval (dashed curves) are presented. The bottom panel shows σ_u^2 (solid) and $\hat{\sigma}_u^2$ (dotted curves) and its 95% confidence interval (dashed lines).

True values $\mu_k(t) = \alpha(t) + u. + \beta_{1\cdot} + \beta_{2k} + \gamma_{\cdot k}$, $k = 1, 2, 3$ are presented by solid curves, dotted curves stand for the estimates by the proposed estimation procedure, and the estimates of the corresponding linear mixed model are depicted with dashed lines. Red, green, and blue stand for the level 1, 2, and 3 of β_2, respectively. The three fitted mean curves are on the true values. In Figure 4.2, the top panel displays the variance estimates $\hat{\sigma}_u^2$, which are plotted on the true values and $\hat{\sigma}_g^2$ is 0.99658 (near 1) in the bottom panel. On the other hand, for the estimates by the linear mixed model, the difference between the three flat lines $\hat{\beta}_{2k}$ in Figure 4.1 are well estimated but the time-constant overall mean function triggers large residual variance $\hat{\sigma}_\varepsilon^2 = 196.4$ and biased variance of a random effect, $\hat{\sigma}_u^2$=0.86.

Table 4.1 reports the average of FCCCs estimate for independent and correlated data at time 10, 25, and 40. UCCCs are presented for comparison when $n =$ 50 and 100. Estimated FCCCs are very close to the true values, although they slightly underestimate the true agreement. Estimates of FCCCs are consistent and robust against correlation so that MSE is small and decreases as n increases, regardless of the amount of serial correlation. The accuracy of the bootstrap variance estimates of indices denoted by SE is presented in Table 4.2. The SD of the 1000 estimated indices from the 1000 Monte Carlo datasets can be considered the true SE. The SE and SD of the 500 variance estimates of the FCCC gauge the overall performance of the parametric bootstrap. The CP indicates overall performance of our estimation procedure and bootstrap variance. In Table 4.2, the average of estimated SEs is slightly greater than the SD, but the difference between the average of the estimated standard errors and the true value is less than half of one standard deviation of the estimated standard errors. This implies that the parametric bootstrap with AR serial correlation performs fairly well in moderate sample size under mild correlation, but CP is slightly conservative for highly correlated data. Note that before adopting ρ structure in step 4 for $\sigma_\varepsilon^2(t)$ estimation based on the semiparametric bootstrap, the observed 95% CP of indices were 10% $\sim$ 30% for $\rho = 0.9$, 50% $\sim$ 70% for $\rho = .5$ and around 85% $\sim$ 88% for $\rho = 0$, as in the simulation results of Morris (2002). After sampling data using parametric bootstrap, the observed 95% CP of indices are increasing but CP is

Table 4.1. Finite sample performance of Indices at time 10, 25 and 40

$\sigma_\varepsilon^2(t) = 0.9 + 0.1\left(\sin(2\pi t/100)\right)^2, \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$

FCCC		n = 50						n = 100					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	ρ	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.681		.402		.338		.681		.402		.338	
	0	.680	.0017	.393	.0056	.332	.0030	.681	.0008	.398	.0028	.336	.0023
10	.5	.680	.0016	.393	.0050	.332	.0039	.681	.0008	.399	.0026	.337	.0022
	.9	.681	.0017	.393	.0054	.332	.0032	.681	.0008	.398	.0031	.336	.0025
THEORETICAL		.667		.396		.331		.667		.396		.331	
	0	.666	.0017	.388	.0045	.325	.0038	.666	.0008	.393	.0027	.329	.0022
25	.5	.665	.0017	.388	.0050	.324	.0038	.666	.0008	.394	.0026	.329	.0021
	.9	.666	.0017	.388	.0053	.325	.0039	.665	.0009	.393	.0030	.328	.0024
THEORETICAL		.682		.402		.338		.682		.402		.338	
	0	.680	.0016	.393	.0046	.331	.0030	.681	.0008	.398	.0028	.336	.0023
40	.5	.681	.0015	.393	.0051	.331	.0039	.680	.0009	.399	.0026	.336	.0022
	.9	.681	.0017	.393	.0054	.332	.0032	.681	.0009	.398	.0031	.336	.0025
THEORETICAL		.678		.492		.336		.678		.492		.336	
UCCC	0	.0035	.4551	.2472	.0655	.0035	.1108	.0035	.4551	.2512	.0607	.0035	.1108
	.5	.0035	.4551	.2485	.0648	.0035	.1108	.0035	.4551	.2539	.0593	.0035	.1108
	.9	.0037	.4548	.2595	.0596	.0037	.1107	.0037	.4548	.2638	.0548	.0037	.1107

Table 4.2. Standard deviations, standard errors and 95% coverage probability of indices at time 10, 25 and 40

$$\sigma_\varepsilon^2(t) = 0.9 + 0.1\left(\sin(2\pi t/100)\right)^2, \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$$

FCCC, $n = 50$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0407	.0398(.0040)	.9423	.0871	.0833(.0087)	.9304	.0775	.0742(.0078)	.9245
10	.5	.0393	.0397(.0046)	.9591	.0772	.0845(.0077)	.9373	.0694	.0752(.0069)	.9345
	.9	.0415	.0397(.0042)	.9382	.0798	.0847(.0079)	.9145	.0716	.0754(.0072)	.9236
	0	.0408	.0406(.0040)	.9493	.0859	.0827(.0086)	.9344	.0759	.0732(.0059)	.9344
25	.5	.0411	.0405(.0041)	.9418	.0771	.0838(.0077)	.9318	.0693	.0741(.0070)	.9282
	.9	.0417	.0406(.0042)	.9432	.0787	.0840(.0079)	.9162	.0700	.0743(.0070)	.9201
	0	.0396	.0398(.0040)	.9553	.0868	.0833(.0087)	.9344	.0771	.0742(.0077)	.9344
40	.5	.0393	.0396(.0039)	.9482	.0775	.0845(.0078)	.9236	.0699	.0752(.0070)	.9291
	.9	.0407	.0396(.0041)	.9388	.0799	.0847(.0080)	.9135	.0717	.0754(.0071)	.9168

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0282	.0281(.0014)	.9583	.0525	.0561(.0083)	.9322	.0477	.0504(.0061)	.9364
10	.5	.0280	.0281(.0015)	.9393	.0512	.0560(.0079)	.9436	.0465	.0503(.0057)	.9393
	.9	.0287	.0280(.0015)	.9373	.0557	.0564(.0086)	.9273	.0503	.0506(.0062)	.9291
	0	.0290	.0288(.0014)	.9458	.0523	.0557(.0081)	.9312	.0473	.0497(.0058)	.9312
25	.5	.0284	.0288(.0014)	.9479	.0509	.0556(.0077)	.9500	.0460	.0496(.0055)	.9435
	.9	.0301	.0287(.0015)	.9391	.0550	.0558(.0084)	.9282	.0493	.0498(.0061)	.9291
	0	.0286	.0281(.0015)	.9395	.0526	.0561(.0083)	.9301	.0478	.0504(.0060)	.9249
40	.5	.0288	.0281(.0015)	.9443	.0513	.0560(.0079)	.9443	.0466	.0503(.0058)	.9443
	.9	.0294	.0281(.0015)	.9436	.0555	.0563(.0086)	.9255	.0501	.0506(.0063)	.9264

still considerably underestimated for highly correlated data: 30% $\sim$ 50% CP at $\rho = .9$ and 80% $\sim$ 85% at $\rho = .5$. The use of correlation structure in the bootstrap stage dramatically increases CP to 93% $\sim$ 95% for non and mildly correlated data. Under strong correlation there is slightly smaller coverage probability, 91% $\sim$ 93% with underestimated SE but it improves as sample size increases. Thus, we gain efficiency of FCCCs estimate by incorporating the correlation structure in the estimation and bootstrap procedure. On the other hand, estimates of UCCCs provide quite different values from theoretical values. Because curve data is not considered in the UCCC model where it uses overall mean and variance under the assumption that repeated measures are stable, indices are dominated by residual variances in the denominator in equation (4.13). Residual variance is divided by large number of repetitions over time and it leads to large Inter UCCC. In other words, although methods have large measurement errors, Inter UCCC will be close to 1 for densely measured data with flat curvature.

Tables 4.5 through 4.8 assess the performance with respect to the degree of noise in the data; Tables 4.5 and 4.6 report the results when there is twice the time-varying variability; Tables 4.7 and 4.8 show when there is half the variability of the results described above. The means and MSEs in Tables 4.5 and 4.7 are comparable. Tables 4.6 and 4.8 demonstrate that the estimated SEs of FCCCs are stable regardless of the amount of variability. The amount of noise in data does not have a large impact on the accuracy of indices.

In order to investigate the impact of the model selection, we consider model (4.12) without interaction effect between a random effect and β_2; then, the indices are defined as

$$UCCC_{\text{intra}} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2},$$

$$\begin{aligned} UCCC_{\text{inter}} &= \frac{\sigma_u^2}{\sigma_u^2+\sigma_\gamma^2+\sigma_\varepsilon^2/200}, \\ UCCC_{\text{total}} &= \frac{\sigma_u^2}{\sigma_u^2+\sigma_\gamma^2+\sigma_\varepsilon^2}, \end{aligned} \tag{4.15}$$

where $\sigma_\gamma^2 = \frac{1}{6}\sum_{k=1}^{2}\sum_{k'=2}^{3}(\gamma_k-\gamma_{k'})^2$.

$$\begin{aligned} FCCC_{\text{intra}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma_\varepsilon^2(t)}, \\ FCCC_{\text{inter}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma_\gamma^2+\sigma_\varepsilon^2(t)/2}, \\ FCCC_{\text{total}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma_\gamma^2+\sigma_\varepsilon^2(t)}. \end{aligned} \tag{4.16}$$

Tables 4.9 and 4.10 report the performance, which is similar to the index based on the model with interaction effects. This result is consistent with the findings of Lin et al. (2007). They showed that the index without interaction proposed by Carrasco & Jover (2003) provides comparable results with the UCCC for ordinal and continuous data. Based on our approach, we point out that the base model can be generalized to include various covariates and FCCCs are easily modified with good performance.

We investigated the minimum sample size for reliable inference on FCCC. When $n = 15$, the subject variance fluctuated sample by sample and it affected the performance of indices. When sample size is larger, e.g. $n = 30$, the subject variance is stabilized and the 95% CPs of indices are around to 95% regardless of the degree of serial correlation in Tables 4.11 and 4.12,.

We tried to understand the model misspecification; under the AR(1) assumption, our estimating procedure was applied to datasets with errors having serial correlation of the stationary AR(2) model:

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \eta_t,$$

where ϕ_1 and ϕ_2 are the parameters of the model and η_t is white noise.
The 95% CP of intra- inter-, and total FCCC are around .91, .93, and .93 respectively at $n = 50$ and .93, .95, and .97 respectively at $n = 100$ for AR(2) parametrization with coefficients (ϕ_1, ϕ_2), (.2, .1); a similar the autocorrelation function as estimated serial correlation of the local approximation of AR(1) assumption. For $n = 100$, 95% CP of intra- inter-, and total FCCC are around .2, .94, and .91 respectively for AR(2) serial correlation (.5,.1), .1, .9, and .8 respectively for (.5,.2), .08, .8, and .6 respectively for (.5,.3), and 0, .4, and .2 respectively for (.5,.4). As the range of dependence gets higher and there is a higher number of related measurements, the misspecification of AR(1) serial correlation induces large bias. Thus, misspecified serial correlation structures affect the performance of indices, and the choice of the serial correlation structure is very important for accurate local approximation of serial correlation and inference on indices.

4.4.2 Time scheme II

In this section, data are generated by scheme II, 20% probability of subjects being skipped at each occasion. Because the linear mixed model is flexible for imbalanced data, it provides unbiased estimates of β. Figure 4.3 depicts the variance components, $\sigma_\varepsilon^2(\cdot)$, σ_u^2 and its estimates at $\rho = 0,\ 0.5,\ 0.9$. The 95% confidence interval at $\rho = .5$ is presented for comparison to the results in Figure 4.3. The variance estimates are placed near true values but the 95% confidence interval is wider than in time scheme I; thus the variance estimates are unbiased regardless

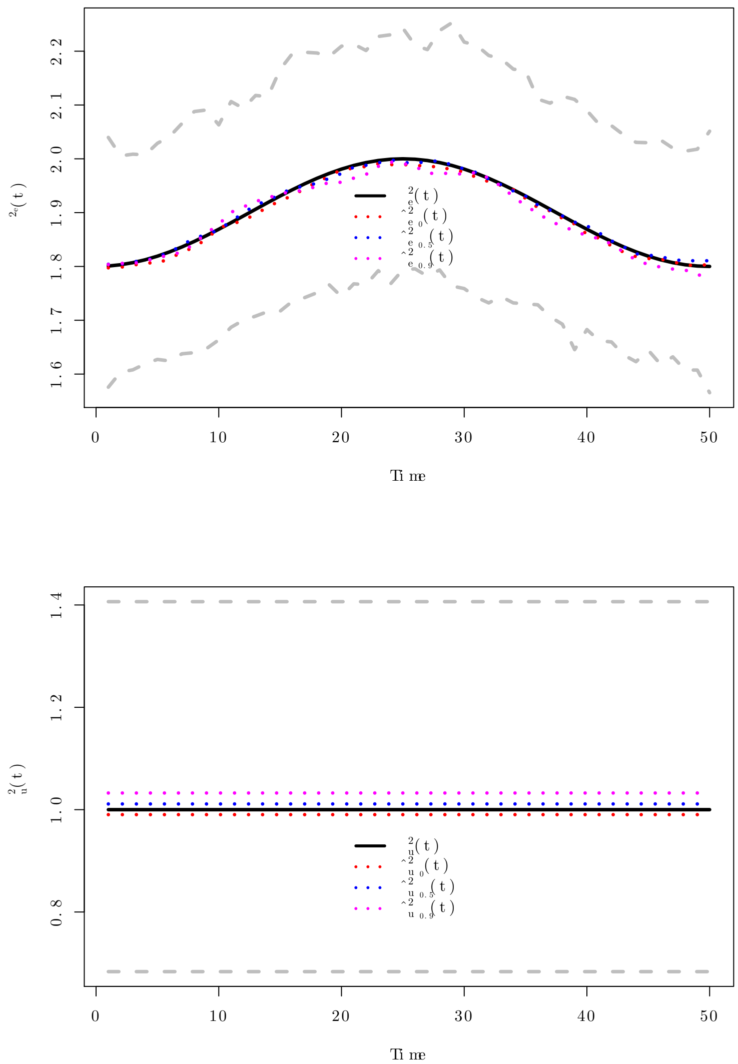

Figure 4.3. Plots for variance components at $\sigma_u^2 = 1, \sigma_g^2 = 1,\ \sigma_\varepsilon^2(t) = 2 \times (.9 + .1\sin(2\pi t/50)^2)$ after removing 20% of observations per occasion. The solid curve represents the true variance and 95% confidence interval of $\rho = 0.5$ is presented with dashed curves. Dotted curves stand for estimates; red, blue and purple stands for the estimates at $\rho = 0,\ 0.5,\ 0.9$ respectively. $\sigma_\varepsilon^2(\cdot)$ is shown in the upper panel. σ_u^2 are displayed in the bottom panel.

of the degree of ρ but less efficient due to the small sample size. The average of 1000 estimates $\hat{\sigma}_g^2$ is 1.0003; close to true value 1. Thus, variance components are well estimated for unbalanced data by our estimation procedure.

Table 4.3. Finite sample performance of Indices with 20% removed observations for each occasion

		$\sigma^2_\varepsilon(t) = 0.9 + 0.1\,(\sin(2\pi t/100))^2, \quad \sigma^2_u = 1, \quad \sigma^2_g = 1$											
FCCC		$n = 50$						$n = 100$					
t	ρ	Intra-		Inter-		Total-		Intra-		Inter-		Total-	
		Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.681		.402		.338		.681		.402		.338	
	0	.679	.0017	.400	.0077	.337	.0054	.680	.0007	.395	.0038	.333	.0029
10	.5	.680	.0017	.395	.0074	.333	.0058	.680	.0008	.396	.0039	.333	.0031
	.9	.679	.0023	.394	.0083	.333	.0066	.683	.0010	.404	.0042	.341	.0033
THEORETICAL		.667		.396		.331		.667		.396		.331	
	0	.663	.0017	.394	.0070	.329	.0053	.665	.0008	.390	.0039	.326	.0028
25	.5	.666	.0017	.390	.0071	.327	.0055	.664	.0009	.390	.0039	.325	.0021
	.9	.661	.0024	.389	.0081	.325	.0063	.669	.0011	.398	.0040	.334	.0031
THEORETICAL		.682		.402		.338		.682		.402		.338	
	0	.678	.0015	.393	.0073	.337	.0057	.680	.0006	.395	.0038	.333	.0029
40	.5	.681	.0015	.395	.0073	.334	.0057	.679	.0008	.395	.0040	.333	.0031
	.9	.677	.0023	.394	.0084	.333	.0066	.683	.0010	.404	.0042	.341	.0032
THEORETICAL		.678		.492		.336		.678		.492		.336	
	0	.0036	.4549	.2521	.0633	.0036	.1107	.0035	.4550	.2518	.0606	.0035	.1108
UCCC	.5	.0036	.4549	.2532	.0634	.0036	.1107	.0035	.4550	.2520	.0608	.0035	.1108
	.9	.0038	.4547	.2609	.0612	.0038	.1106	.0039	.4545	.2721	.0514	.0039	.1105

Table 4.4. SDs, SEs and 95% CPs of indices with 20% removed observations for each occasion

$\sigma_\varepsilon^2(t) = 0.9 + 0.1\left(\sin(2\pi t/100)\right)^2, \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$

FCCC, $n = 50$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0360	.0335(.005)	.9311	.0842	.0829(.0136)	.9224	.0738	.0736(.0098)	.9235
10	.5	.0407	.0399(.0042)	.9467	.0856	.0782(.0097)	.9212	.0760	.0721(.0074)	.9288
	.9	.0484	.0432(.0036)	.9317	.0912	.0806(.0045)	.9215	.0811	.0738(.0037)	.9181
	0	.0380	.0347(.0025)	.9337	.0839	.0821(.0132)	.9284	.0734	.0724(.0093)	.9286
25	.5	.0403	.0371(.0041)	.9448	.0843	.0808(.0094)	.9418	.0740	.0711(.0074)	.9214
	.9	.0491	.0432(.0033)	.9352	.0901	.0830(.0044)	.9185	.0794	.0677(.0037)	.9113
	0	.0379	.0341(.0025)	.9388	.0853	.0828(.0014)	.9250	.0754	.0735(.0096)	.9207
40	.5	.0393	.0381(.0026)	.9467	.0853	.0782(.0095)	.9218	.0755	.0701(.0074)	.9216
	.9	.0484	.0425(.0036)	.9488	.0913	.0846(.0065)	.9215	.0717	.0688(.0035)	.9248

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0271	.0233(.0014)	.9359	.0612	.0590(.0104)	.9372	.0536	.0507(.0079)	.9453
10	.5	.0285	.0282(.0014)	.9592	.0629	.0591(.0062)	.9218	.0553	.0521(.0049)	.9286
	.9	.0318	.0301(.0018)	.9533	.0648	.0598(.0026)	.9266	.0513	.0511(.0022)	.9418
	0	.0268	.0248(.0014)	.9395	.0606	.0649(.0102)	.9412	.0528	.0500(.0020)	.9453
25	.5	.0296	.0289(.0014)	.9395	.622	.0617(.0062)	.9420	.0543	.0510(.0055)	.9350
	.9	.0301	.0287(.0015)	.9391	.0550	.0558(.0084)	.9282	.0493	.0498(.0061)	.9291
	0	.0254	.0243(.0014)	.9424	.0613	.0580(.0104)	.9310	.0536	.0507(.0079)	.9359
40	.5	.0285	.0259(.0014)	.9490	.0628	.0591(.0062)	.9286	.0551	.0513(.0058)	.9320
	.9	.0314	.0301(.0019)	.9533	.0646	.0611(.0026)	.9297	.0571	.0552(.0020)	.9419

Table 4.3 reports the average of FCCC estimates for independent and correlated data at t=10, 25, 40. The index estimates are almost the same as the true values for any degree of correlation at $n = 100$, similar to the estimates for data, generated by time scheme I, although there is subtle bias with small samples. The MSE itself is greater than the MSE in Table 4.2 but it decays as sample size increases. In Table 4.4, although the average of estimated SEs is slightly smaller than SD, the difference between the average of the estimated standard errors and the true value is less than half of one standard deviation of the estimated standard errors. The SE and SD are greater than those of complete data and this leads to the higher CP in the presence of 20 % probability of each subject being skipped at time points. Tables 4.13 through 4.18 correspond to Tables 4.5 through 4.10. Tables 4.13 and 4.14 shows bias, MSE, SE, and SD of indices when the data have twice the time-varying variability over time, $\sigma_\varepsilon^2(t) = 2 \times \left(0.9 + 0.1\,(\sin(2\pi t/100))^2\right)$, $\sigma_u^2 = 1$, and $\sigma_g^2 = 1$. Tables 4.15 and 4.16 report four values of indices when the data have half the variability over time, $\sigma_\varepsilon^2(t) = 0.9 + 0.1\,(\sin(2\pi t/100))^2$, $\sigma_u^2 = .5$, and $\sigma_g^2 = .5$. FCCC indices are subtly underestimated and MSEs are subtly larger than those for data of time scheme I in Tables 4.13 and 4.15. The small bias vanishes and MSE decreases with increasing sample size just as the results in Table 4.3. In Tables 4.14 and 4.16, SEs underestimate SDs, a larger value than SDs of time scheme I at small sample size, and it is improved by increasing sample size. Based on the results that the pattern on the performance of estimates are comparable to the case of time scheme I, estimates of FCCC and bootstrap variances may be robust against the amount of variability in data, although they are slightly influenced by skipping individual time points in small sample. Tables 4.17 and 4.18 show consistent results; thus they are evidence that the extension of Carrasco & Jover (2003) is also flexible for skipping individual time points in large sample. Thus, FCCCs are consistent to the true values with slight loss of precision for imbalance by 20% probability of subjects being skipped at each occasion.

In summary, our simulation studies demonstrate that the FCCCs improve the detection of agreement for ILD with curve features by isolating the time impact caused by smooth mean curve from the components contributed to the concordance between methods.

4.5 Discussion

When new device, assay, methodology and theory are developed, it is necessary to evaluate agreement for comparison studies as theories and technologies continue to evolve. Moreover, advanced measuring devices allow investigators to collect data more rapidly, frequently and flexibly with lower cost. There is the need for assessment of measurement agreement for densely repeated measures. In this paper we have proposed the functional type of concordance correlation coefficient for measuring agreement among densely repeated measures taken by more than two methods. FCCCs are time varying indices and are informative about the trend of measurement agreement over time periods. Moreover, FCCCs are robust against the functional form of the data because partially linear mixed models are flexible enough to accommodate smooth curves. The variance estimates and coefficient estimates are accurate even under strong correlation. The parametric bootstrap for data with AR correlation provides reliable variance of indices when true correlation structure is AR. However, our bootstrap method slightly underestimates residual variance under strong correlation. Thus, more study is necessary on the adjustment of bootstrap methods to account for highly correlated data with general correlation structure.

Table 4.5. Finite sample performance of Indices at time 10, 25 and 40

$\sigma_\varepsilon^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right), \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$

FCCC		n = 50						n = 100					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	ρ	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.517		.338		.257		.517		.338		.257	
	0	.517	.0020	.331	.0056	.253	.0038	.517	.0011	.338	.0023	.257	.0016
10	.5	.516	.0020	.335	.0047	.254	.0033	.516	.0010	.336	.0023	.256	.0016
	.9	.517	.0025	.331	.0057	.253	.0039	.517	.0011	.335	.0028	.256	.0019
THEORETICAL		.500		.331		.249		.500		.331		.249	
	0	.500	.0020	.324	.0053	.245	.0035	.499	.0011	.328	.0023	.247	.0016
25	.5	.499	.0022	.325	.0047	.245	.0032	.499	.0011	.329	.0023	.248	.0016
	.9	.500	.0023	.324	.0055	.245	.0036	.499	.0012	.328	.0027	.247	.0018
THEORETICAL		.516		.338		.257		.516		.338		.257	
	0	.516	.0020	.331	.0056	.253	.0037	.517	.0011	.336	.0023	.256	.0016
40	.5	.517	.0020	.332	.0048	.254	.0033	.517	.0011	.336	.0023	.256	.0016
	.9	.517	.0023	.332	.0057	.254	.0039	.517	.0012	.335	.0028	.256	.0019
THEORETICAL		.513		.490		.255		.513		.490		.255	
	0	.0034	.2597	.2455	.0650	.0034	.0633	.0035	.2597	.2502	.0602	.0035	.0633
UCCC	.5	.0035	.2594	.2487	.0636	.0035	.0653	.0035	.2596	.2543	.0582	.0035	.0624
	.9	.0039	.2592	.2702	.0537	.0039	.0630	.0039	.2592	.2747	.0489	.0039	.0631

Table 4.6. SDs, SEs and 95% CPs of indices at time 10, 25 and 40

$\sigma_\varepsilon^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right), \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$

FCCC		$n = 50$								
		Intra-			Inter-			Total-		
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0451	.0450(.0031)	.9418	.0746	.0743(.0113)	.9339	.0613	.0611(.0084)	.9310
10	.5	.0452	.0454(.0020)	.9569	.0688	.0759(.0098)	.9349	.0573	.0625(.0068)	.9359
	.9	.0496	.0453(.0022)	.9264	.0752	.0759(.0107)	.9073	.0622	.0625(.0077)	.9127
	0	.0445	.0451(.0031)	.9546	.0727	.0733(.0112)	.9418	.0591	.0598(.0083)	.9369
25	.5	.0465	.0454(.0021)	.9445	.0686	.0747(.0094)	.9340	.0570	.0609(.0065)	.9301
	.9	.0487	.0454(.0022)	.9227	.0736	.0748(.0104)	.9045	.0601	.0610(.0075)	.9091
	0	.0445	.0451(.0031)	.9497	.0743	.0743(.0011)	.9369	.0610	.0612(.0084)	.9359
40	.5	.0453	.0454(.0020)	.9483	.0691	.0759(.0097)	.9273	.0578	.0624(.0066)	.9292
	.9	.0482	.0453(.0022)	.9218	.0753	.0758(.0107)	.9173	.0623	.0625(.0076)	.9164

FCCC		$n = 100$								
		Intra-			Inter-			Total-		
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0322	.0322(.0012)	.9508	.0489	.0513(.0073)	.9345	.0417	.0426(.0048)	.9331
	.5	.0323	.0322(.0012)	.9418	.0484	.0514(.0073)	.9308	.0406	.0426(.0047)	.9327
10	.9	.0342	.0322(.0013)	.9327	.0531	.0520(.0082)	.9118	.0441	.0430(.0054)	.9191
	0	.0329	.0322(.0012)	.9419	.0484	.0512(.0079)	.9341	.0401	.0415(.0046)	.9281
	.5	.0325	.0323(.0012)	.9464	.0477	.0507(.0070)	.9345	.0397	.0416(.0046)	.9318
25	.9	.0351	.0322(.0014)	.9291	.0519	.0512(.0080)	.9182	.0426	.0420(.0053)	.9182
	0	.0329	.0322(.0012)	.9409	.0491	.0514(.0072)	.9301	.0411	.0426(.0047)	.9311
	.5	.0329	.0322(.0012)	.9436	.0482	.0514(.0073)	.9355	.0404	.0425(.0048)	.9291
40	.9	.0349	.0322(.0013)	.9309	.0529	.0520(.0083)	.9145	.0244	.0430(.0056)	.9191

Table 4.7. Finite sample performance of Indices at time 10, 25 and 40

$\sigma_\varepsilon^2(t) = 0.9 + 0.1\left(\sin(2\pi t/100)\right)^2, \quad \sigma_u^2 = 0.5, \quad \sigma_g^2 = 0.5$

FCCC		$n=50$						$n=100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	ρ	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.517		.0.335		.255		.517		.0.335		.255	
	0	.515	.0023	.332	.0052	.254	.0036	.517	.0011	.335	.0023	.255	.0016
10	.5	.516	.0021	.329	.0047	.251	.0033	.516	.0010	.334	.0023	.255	.0016
	.9	.517	.0025	.329	.0056	.252	.0038	.517	.0012	.333	.0028	.254	.0019
THEORETICAL		.500		.328		.247		.500		.328		.247	
	0	.499	.0022	.325	.0049	.246	.0034	.5000	.0011	.326	.0023	.246	.0016
25	.5	.500	.0022	.322	.0047	.244	.0032	.499	.0011	.327	.0023	.247	.0016
	.9	.500	.0024	.326	.0054	.246	.0036	.500	.0012	.322	.0027	.244	.0018
THEORETICAL		.517		.335		.255		.517		.335		.255	
	0	.515	.0021	.332	.0051	.254	.0035	.517	.0011	.333	.0023	.254	.0016
40	.5	.518	.0021	.329	.0048	.252	.0033	.516	.0011	.334	.0023	.254	.0016
	.9	.517	.0023	.329	.0057	.252	.0039	.517	.0012	.333	.0028	.254	.0019
THEORETICAL		.5129		.4841		.2535		.5129		.4841		.2535	
	0	.0003	.2629	.0288	.2090	.0003	.0641	.0002	.2630	.0221	.2144	.0002	.0641
UCCC	.5	.0005	.2629	.0446	.2081	.0005	.0641	.0001	.2629	.0241	.2126	.0003	.0641
	.9	.0005	.2628	.0447	.2080	.0005	.0602	.0004	.2628	.0366	.2016	.0004	.0641

Table 4.8. SDs, SEs and 95% CPs of indices at time 10, 25 and 40

$\sigma_\varepsilon^2(t) = 0.9 + 0.1\,(\sin(2\pi t/100))^2, \quad \sigma_u^2 = 0.5, \quad \sigma_g^2 = 0.5$

FCCC		$n = 50$								
			Intra-			Inter-			Total-	
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0475	.0455(.0021)	.9460	.0718	.0738(.0088)	.9290	.0599	.0609(.0064)	.9276
10	.5	.0455	.0454(.0021)	.9530	.0683	.0748(.0090)	.9340	.0570	.0616(.0063)	.9322
	.9	.0494	.0453(.0022)	.9264	.0749	.0746(.0096)	.9075	.0620	.0616(.0069)	.9145
	0	.0467	.0456(.0020)	.9441	.0702	.0728(.0088)	.9293	.0579	.0596(.0064)	.9270
25	.5	.0465	.0454(.0021)	.9464	.0712	.0734(.0085)	.9345	.0596	.0602(.0060)	.9318
	.9	.0487	.0454(.0022)	.9286	.0733	.0735(.0094)	.9212	.0599	.0601(.0068)	.9212
	0	.0455	.0456(.0021)	.9361	.0713	.0738(.0089)	.9293	.0593	.0609(.0064)	.9270
40	.5	.0454	.0454(.0020)	.9436	.0719	.0748(.0088)	.9345	.0587	.0616(.0061)	.9318
	.9	.0482	.0453(.0022)	.9295	.0750	.0745(.0096)	.9128	.0620	.0615(.0068)	.9193

FCCC		$n = 100$								
			Intra-			Inter-			Total-	
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0324	.0322(.0012)	.9555	.0477	.0452(.0071)	.9361	.0400	.0424(.0033)	.9339
10	.5	.0323	.0322(.0012)	.9400	.0482	.0512(.0070)	.9282	.0404	.0423(.0046)	.9327
	.9	.0342	.0322(.0013)	.9323	.0529	.0518(.0081)	.9137	.0439	.0429(.0054)	.9193
	0	.0326	.0323(.0012)	.9441	.0474	.0500(.0067)	.9327	.0394	.0414(.0044)	.9340
25	.5	.0325	.0323(.0012)	.9464	.0476	.0504(.0067)	.9345	.0395	.0413(.0044)	.9318
	.9	.0351	.0326(.0014)	.9286	.0517	.0511(.0070)	.9212	.0424	.0418(.0053)	.9221
	0	.0331	.0322(.0012)	.9361	.0478	.0512(.0071)	.9293	.0402	.0424(.0046)	.9270
40	.5	.0329	.0322(.0012)	.9436	.0481	.0519(.0070)	.9345	.0403	.0423(.0046)	.9318
	.9	.0351	.0322(.0013)	.9295	.0527	.0518(.0082)	.9128	.0437	.0428(.0055)	.9193

Table 4.9. Finite sample performance of Indices at time 10, 25 and 40

$\sigma_\varepsilon^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right), \quad \sigma_u^2 = 1$

FCCC		n = 50						n = 100					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	ρ	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.349		.511		.346		.349		.511		.346	
	0	.345	.0025	.505	.0031	.343	.0025	.347	.0013	.508	.0016	.345	.0013
10	.5	.348	.0026	.508	.0031	.345	.0026	.347	.0014	.507	.0017	.344	.0014
	.9	.348	.0031	.506	.0038	.344	.0031	.350	.0012	.511	.0014	.347	.0011
THEORETICAL		.333		.494		.331		.333		.494		.331	
	0	.331	.0025	.490	.0032	.329	.0025	.333	.0013	.492	.0017	.330	.0013
25	.5	.333	.0026	.491	.0032	.330	.0025	.332	.0013	.491	.0016	.329	.0013
	.9	.335	.0029	.494	.0037	.332	.0029	.333	.0015	.490	.0019	.330	.0015
THEORETICAL		.349		.511		.346		.349		.511		.346	
	0	.346	.0025	.506	.0031	.343	.0025	.347	.0014	.508	.0017	.344	.0014
40	.5	.348	.0026	.508	.0031	.345	.0025	.347	.0014	.508	.0017	.344	.0013
	.9	.348	.0031	.507	.0038	.345	.0031	.350	.0016	.512	.0019	.348	.0016
THEORETICAL		.3450		.9594		.3422		.3450		.9594		.3422	
UCCC	0	.0018	.1179	.1433	.6713	.0019	.1160	.0018	.1179	.1461	.6642	.0018	.1160
	.5	.0019	.1178	.1498	.6609	.0019	.1159	.0018	.1179	.1504	.6572	.0018	.1160
	.9	.0022	.1175	.1774	.6713	.0025	.1157	.0023	.1175	.1823	.6062	.0023	.1157

Table 4.10. SDs, SEs and 95% CPs of indices at time 10, 25 and 40

$$\sigma_\varepsilon^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right), \quad \sigma_u^2 = 1$$

FCCC, $n = 50$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0503	.0494(.0037)	.9341	.0555	.0546(.0020)	.9334	.0500	.0491(.0037)	.9341
10	.5	.0510	.0495(.0037)	.9436	.0560	.0546(.0019)	.9445	.0507	.0492(.0037)	.9427
	.9	.0557	.0494(.0040)	.9100	.0616	.0544(.0020)	.9055	.0553	.0490(.0041)	.9082
	0	.0498	.0483(.0039)	.9459	.0560	.0545(.0019)	.9466	.0495	.0481(.0039)	.9459
25	.5	.0507	.0485(.0039)	.9382	.0567	.0546(.0019)	.9373	.0503	.0482(.0039)	.9400
	.9	.0542	.0484(.0042)	.9182	.0610	.0545(.0020)	.9136	.0539	.0481(.0042)	.9164
	0	.0503	.0494(.0037)	.9438	.0555	.0546(.0019)	.9438	.0500	.0491(.0037)	.9438
40	.5	.0506	.0496(.0037)	.9509	.0556	.0547(.0019)	.9518	.0503	.0493(.0037)	.9509
	.9	.0558	.0494(.0040)	.9118	.0616	.0544(.0021)	.9082	.0554	.0491(.0041)	.9109

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0367	.0351(.0020)	.9336	.0403	.0386(.0013)	.9345	.0365	.0349(.0020)	.9318
	.5	.0371	.0350(.0020)	.9273	.0401	.0386(.0012)	.9273	.0369	.0348(.0020)	.9273
10	.9	.0340	.0352(.0020)	.9637	.0376	.0387(.0013)	.9637	.0339	.0350(.0020)	.9637
	0	.0362	.0343(.0021)	.9318	.0406	.0386(.0013)	.9327	.0360	.0341(.0021)	.9327
	.5	.0360	.0343(.0021)	.9355	.0405	.0386(.0013)	.9355	.0368	.0341(.0021)	.9355
25	.9	.0392	.0346(.0024)	.9345	.0435	.0387(.0015)	.9282	.0389	.0344(.0024)	.9355
	0	.0370	.0351(.0021)	.9364	.0406	.0386(.0013)	.9364	.0368	.0341(.0021)	.9364
	.5	.0375	.0343(.0021)	.9364	.0413	.0386(.0013)	.9327	.0372	.0341(.0021)	.9345
40	.9	.0395	.0346(.0024)	.9182	.0436	.0387(.0015)	.9182	.0393	.0344(.0024)	.9191

Table 4.11. Finite sample performance of Indices at time 10, 25 and 40

$$\sigma_\varepsilon^2(t) = 0.9 + 0.1\left(\sin(2\pi t/100)\right)^2, \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$$

FCCC		$n = 15$						$n = 30$					
t	ρ	Intra-		Inter-		Total-		Intra-		Inter-		Total-	
		Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.681		.402		.338		.681		.402		.338	
	0	.672	.0053	.354	.0214	.300	.0163	.668	.0032	.388	.0085	.327	.0069
10	.5	.680	.0041	.354	.0310	.294	.0233	.682	.0022	.384	.0079	.325	.0061
	.9	.672	.0040	.363	.0224	.306	.0168	.671	.0035	.391	.0064	.329	.0052
THEORETICAL		.667		.396		.331		.667		.396		.331	
	0	.652	.0061	.349	.0214	.292	.0162	.659	.0036	.385	.0084	.322	.0067
25	.5	.663	.0059	.340	.0311	.286	.0235	.666	.0028	.378	.0075	.317	.0057
	.9	.655	.0061	.358	.0288	.300	.0174	.665	.0036	.388	.0064	.326	.0051
THEORETICAL		.682		.402		.338		.682		.402		.338	
	0	.668	.0049	.353	.0213	.298	.0161	.671	.0028	.389	.0083	.328	.0066
40	.5	.684	.0041	.345	.0306	.294	.0226	.676	.0028	.382	.0080	.323	.0062
	.9	.675	.0058	.365	.0234	.309	.0184	.678	.0032	.393	.0065	.332	.0052
THEORETICAL		.678		.492		.336		.678		.492		.336	
	0	.0035	.4554	.2319	.0929	.0035	.1110	.0034	.4552	.2379	.0737	.0034	.1109
UCCC	.5	.0035	.4551	.2328	.0840	.0035	.1109	.0034	.4552	.2393	.0724	.0034	.1109
	.9	.0036	.4550	.2366	.0838	.0036	.1108	.0038	.4537	.2598	.0638	.0038	.1107

Table 4.12. Standard deviations, standard errors and 95% coverage probability of indices at time 10, 25 and 40

$$\sigma_\varepsilon^2(t) = 0.9 + 0.1\left(\sin(2\pi t/100)\right)^2, \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$$

FCCC		$n = 15$								
			Intra-			Inter-			Total-	
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0718	.0727(.0077)	.9622	.1386	.1433(.0225)	.9622	.1219	.1269(.0227)	.9800
10	.5	.0639	.0727(.0073)	.9168	.1672	.1389(.0303)	.9202	.1462	.1229(.0280)	.8553
	.9	.0626	.0738(.0077)	.9157	.1448	.1415(.0242)	.9099	.1260	.1258(.0228)	.9270
	0	.0770	.0742(.0068)	.9622	.1384	.1441(.0230)	.9500	.1215	.1238(.0228)	.9622
25	.5	.0769	.0735(.0080)	.9168	.1674	.1361(.0299)	.8514	.1471	.1200(.0284)	.8333
	.9	.0776	.0748(.0077)	.9359	.1461	.1398(.0244)	.8970	.1284	.1234(.0237)	.8670
	0	.0691	.0738(.0072)	.9622	.1327	.1431(.0226)	.9600	.1206	.1265(.0220)	.9600
40	.5	.0691	.0720(.0075)	.9025	.1654	.1384(.0380)	.8320	.1440	.1233(.0282)	.8441
	.9	.0759	.0725(.0081)	.9571	.1492	.1414(.0247)	.9099	.1329	.1258(.0239)	.9142

FCCC		$n = 30$								
			Intra-			Inter-			Total-	
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0550	.0525(.0042)	.9250	.0914	.1074(.0150)	.9550	.0822	.0952(.0119)	.9483
10	.5	.0469	.0508(.0041)	.9651	.0871	.1005(.0156)	.9652	.0772	.0792(.0127)	.9552
	.9	.0580	.0520(.0043)	.9447	.0795	.0985(.0142)	.9672	.0713	.0860(.0119)	.9677
	0	.0597	.0529(.0043)	.9367	.0912	.1062(.0151)	.9467	.0816	.0941(.0122)	.9383
25	.5	.0528	.0521(.0039)	.9601	.0852	.0986(.0159)	.9652	.0746	.0858(.0134)	.9652
	.9	.0583	.0524(.0047)	.9217	.0792	.0980(.0139)	.9624	.0712	.0855(.0111)	.9670
	0	.0516	.0522(.0043)	.9350	.0903	.1077(.0153)	.9367	.0804	.0955(.0122)	.9367
40	.5	.0527	.0515(.0044)	.9601	.0834	.0991(.0101)	.9652	.0775	.0967(.0137)	.9601
	.9	.0562	.0513(.0046)	.9447	.0800	.0988(.0140)	.9624	.0722	.0867(.0115)	.9624

Table 4.13. Finite sample performance of Indices with 20% removed observations for each occasion

		$\sigma_\varepsilon^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right)$, $\sigma_u^2 = 1$, $\sigma_g^2 = 1$											
FCCC		$n = 50$						$n = 100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	ρ	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.517		.338		.257		.517		.338		.257	
	0	.518	.0018	.339	.0056	.259	.0037	.517	.0010	.338	.0028	.257	.0018
10	.5	.516	.0025	.333	.0061	.252	.0040	.517	.0011	.337	.0031	.257	.0021
	.9	.515	.0032	.335	.0073	.256	.0049	.519	.0016	.334	.0031	.260	.0021
THEORETICAL		.500		.331		.249		.500		.331		.249	
	0	.500	.0019	.331	.0054	.240	.0036	.498	.0010	.329	.0027	.247	.0011
25	.5	.497	.0023	.325	.0058	.244	.0038	.501	.0012	.330	.0029	.249	.0012
	.9	.496	.0033	.326	.0071	.246	.0046	.505	.0016	.334	.0031	.246	.0021
THEORETICAL		.516		.338		.257		.516		.338		.257	
	0	.517	.0020	.339	.0057	.259	.0039	.515	.0010	.336	.0028	.256	.0018
40	.5	.512	.0024	.331	.0061	.252	.0040	.516	.0011	.337	.0030	.257	.0011
	.9	.513	.0033	.333	.0074	.255	.0050	.520	.0014	.340	.0031	.261	.0021
THEORETICAL		.513		.490		.255		.513		.490		.255	
UCCC	0	.0037	.2594	.2546	.0614	.0037	.0632	.0033	.2601	.2458	.0640	.0034	.0630
	.5	.0036	.2594	.2517	.0631	.0036	.0632	.0036	.2593	.2578	.0579	.0036	.0632
	.9	.0040	.2591	.2711	.0550	.0040	.0630	.0041	.2589	.2842	.0409	.0041	.0629

Table 4.14. SDs, SEs and 95% CPs of indices with 20% removed observations for each occasion

$\sigma_{\varepsilon}^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right), \quad \sigma_u^2 = 1, \quad \sigma_g^2 = 1$

FCCC		$n = 50$								
			Intra-			Inter-			Total-	
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0420	.0382(.0024)	.9398	.0746	.0739(.0111)	.9357	.0607	.0605(.0075)	.9357
	.5	.0498	.0399(.0026)	.9358	.0784	.0674(.0068)	.9469	.0637	.0555(.0061)	.9469
10	.9	.0573	.04780(.0027)	.9567	.0860	.0723(.0040)	.9565	.0701	.0607(.0049)	.9348
	0	.0433	.0382(.0025)	.9418	.0761	.0738(.0106)	.9237	.0624	.0604(.0073)	.9237
	.5	.0475	.0400(.0026)	.9491	.0765	.0662(.0065)	.9425	.0612	.0541(.0057)	.9403
25	.9	.0579	.0480(.0026)	.9631	.0843	.0709(.0039)	.9520	.0680	.0581(.0048)	.9348
	0	.0329	.0322(.0012)	.9409	.0491	.0514(.0072)	.9301	.0411	.0426(.0047)	.9311
	.5	.0487	.0400(.0027)	.9381	.0781	.0672(.0067)	.9381	.0633	.0555(.0059)	.9447
40	.9	.0575	.0478(.0026)	.9510	.0855	.0721(.0037)	.9529	.0708	.0600(.0046)	.9384

FCCC		$n = 100$								
			Intra-			Inter-			Total-	
t	ρ	SD	SE(std)	CP	SD	SE(std)	CP	SD	SE(std)	CP
	0	.0314	.0268(.0013)	.9399	.0537	.0519(.0098)	.9370	.0433	.0424(.0067)	.9370
10	.5	.0335	.0284(.0016)	.9564	.0553	.0547(.0023)	.9419	.0453	.0391(.0024)	.9502
	.9	.0400	.0339(.0022)	.9514	.0559	.0519(.0040)	.9413	.0460	.0429(.0033)	.9570
	0	.0305	.0267(.0013)	.9484	.0525	.0511(.0095)	.9398	.0420	.0423(.0065)	.9427
25	.5	.0342	.0284(.0016)	.9627	.0544	.0464(.0023)	.9508	.0442	.0386(.0027)	.9302
	.9	.0397	.0335(.0023)	.9516	.0555	.0511(.0039)	.9577	.0454	.0421(.0032)	.9531
	0	.0445	.0451(.0031)	.9497	.0743	.0743(.0011)	.9369	.0610	.0612(.0084)	.9359
40	.5	.0335	.0284(.0015)	.9506	.0550	.0490(.0023)	.9481	.0449	.0391(.0025)	.9481
	.9	.0381	.0332(.0023)	.9570	.0558	.0519(.0040)	.9631	.0458	.0431(.0032)	.9421

Table 4.15. Finite sample performance of Indices with 20% removed observations for each occasion

$\sigma_\varepsilon^2(t) = 0.9 + 0.1\left(\sin(2\pi t/100)\right)^2, \quad \sigma_u^2 = 0.5, \quad \sigma_g^2 = 0.5$

FCCC		$n = 50$						$n = 100$					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	ρ	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.517		.0.335		.255		.517		.0.335		.255	
	0	.514	.0024	.337	.0067	.257	.0042	.516	.0012	.337	.0037	.257	.0023
10	.5	.512	.0028	.329	.0077	.252	.0049	.516	.0013	.332	.0042	.253	.0026
	.9	.516	.0035	.334	.0094	.256	.0060	.516	.0016	.337	.0045	.257	.0028
THEORETICAL		.500		.328		.247		.500		.328		.247	
	0	.497	.0025	.330	.0066	.250	.0041	.500	.0013	.330	.0036	.249	.0022
25	.5	.496	.0027	.323	.0073	.249	.0045	.499	.0014	.325	.0041	.245	.0025
	.9	.498	.0037	.327	.0089	.248	.0057	.501	.0019	.331	.0044	.250	.0027
THEORETICAL		.517		.335		.255		.517		.335		.255	
	0	.514	.0026	.337	.0069	.258	.0044	.515	.0012	.336	.0027	.256	.0023
40	.5	.513	.0027	.330	.0076	.252	.0048	.515	.0014	.332	.0042	.253	.0026
	.9	.516	.0036	.334	.0073	.256	.0060	.517	.0019	.337	.0045	.257	.0028
THEORETICAL		.513		.484		.253		.513		.484		.253	
	0	.0003	.2628	.0272	.2040	.0003	.0641	.0003	.2629	.0271	.2099	.0003	.0641
UCCC	.5	.0003	.2628	.0272	.2040.	.0003	.0641	.0003	.2629	.0271	.2100	.0003	.0641
	.9	.0005	.2626	.0423	.1918	.0005	.0640	.0005	.2627	.0273	.1971	.0005	.0640

Table 4.16. SDs, SEs and 95% CPs of indices with 20% removed observations for each occasion

$\sigma_\varepsilon^2(t) = 0.9 + 0.1\,(\sin(2\pi t/100))^2, \quad \sigma_u^2 = 0.5, \quad \sigma_g^2 = 0.5$

FCCC, $n = 50$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0491	.0484(.0028)	.9265	.0819	.0717(.0108)	.9161	.0652	.0696(.0077)	.9323
10	.5	.0531	.0504(.0029)	.9335	.0878	.0659(.0041)	.9041	.0702	.0546(.0048)	.9159
	.9	.0597	.0547(.0031)	.9526	.0971	.0715(.0042)	.9245	.0779	.0796(.0049)	.9447
	0	.0499	.0384(.0028)	.9129	.0812	.0714(.0104)	.9258	.0641	.0681(.0074)	.9387
25	.5	.0518	.0404(.0028)	.9354	.0854	.0642(.0041)	.9178	.0674	.0598(.0049)	.9279
	.9	.0613	.0489(.0031)	.9529	.0744	.0704(.0042)	.9217	.0748	.0600(.0048)	.9130
	0	.0505	.0384(.0029)	.9232	.0833	.0743(.0175)	.9392	.0666	.0596(.0077)	.9290
40	.5	.0521	.0434(.0029)	.9335	.0870	.0652(.0041)	.9139	.0693	.0542(.0048)	.9256
	.9	.0602	.0479(.0030)	.9447	.0968	.0714(.0048)	.9125	.0779	.0594(.0048)	.9249

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0353	.0269(.0015)	.9154	.0606	.0515(.0106)	.9248	.0479	.0423(.0078)	.9310
10	.5	.0367	.0385(.0016)	.9553	.0651	.0561(.0021)	.9345	.0511	.0387(.0023)	.9187
	.9	.0424	.0337(.0018)	.9459	.0673	.0508(.0022)	.9282	.0532	.0422(.0024)	.9349
	0	.0364	.0269(.0016)	.9214	.0599	.0508(.0067)	.9285	.0471	.0413(.0068)	.9273
25	.5	.0372	.0285(.0016)	.9593	.0638	.0549(.0022)	.9309	.0494	.0378(.0024)	.9350
	.9	.0437	.0339(.0017)	.9537	.0664	.0502(.0022)	.9237	.0523	.0414(.0026)	.9305
	0	.0349	.0269(.0016)	.9218	.0600	.0512(.0068)	.9298	.0482	.0422(.0069)	.9373
40	.5	.0369	.0284(.0016)	.9350	.0646	.0566(.0022)	.9228	.0506	.0487(.0023)	.9187
	.9	.0440	.0337(.0018)	.9421	.0672	.0589(.0022)	.9505	.0533	.0423(.0025)	.9459

Table 4.17. Finite sample performance of Indices with 20% removed observations for each occasion

$$\sigma_\varepsilon^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right), \quad \sigma_u^2 = 1$$

FCCC		n = 50						n = 100					
		Intra-		Inter-		Total-		Intra-		Inter-		Total-	
t	ρ	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
THEORETICAL		.349		.511		.346		.349		.511		.346	
	0	.349	.0032	.506	.0040	.345	.0032	.347	.0015	.508	.0018	.344	.0015
10	.5	.348	.0032	.501	.0039	.345	.0030	.347	.0016	.505	.0020	.344	.0016
	.9	.345	.0033	.498	.0039	.339	.0031	.350	.0018	.504	.0023	.341	.0018
THEORETICAL		.333		.494		.331		.333		.494		.331	
	0	.333	.0031	.490	.0041	.330	.0031	.333	.0014	.492	.0017	.330	.0014
25	.5	.333	.0028	.491	.0038	.327	.0028	.332	.0015	.490	.0019	.328	.0015
	.9	.328	.0031	.489	.0041	.327	.0032	.331	.0017	.488	.0023	.328	.0018
THEORETICAL		.349		.511		.346		.349		.511		.346	
	0	.348	.0033	.506	.0041	.345	.0033	.347	.0015	.507	.0018	.344	.0014
40	.5	.346	.0029	.502	.0027	.342	.0029	.346	.0015	.505	.0019	.344	.0015
	.9	.343	.0032	.505	.0041	.339	.0039	.350	.0016	.512	.0019	.348	.0016
THEORETICAL		.345		.959		.342		.345		.959		.342	
	0	.0019	.1178	.151	.6605	.0019	.1159	.0018	.1179	.148	.6613	.0018	.1160
UCCC	.5	.0019	.1178	.150	.6619	.0019	.1159	.0019	.1178	.151	.6578	.0019	.1160
	.9	.0023	.1176	.174	.6251	.0023	.1157	.0022	.1176	.176	.6168	.0022	.1157

Table 4.18. SDs, SEs and 95% CPs of indices with 20% removed observations for each occasion

$$\sigma_\varepsilon^2(t) = 2\left(0.9 + 0.1\left(\sin(2\pi t/100)\right)^2\right), \quad \sigma_u^2 = 1$$

FCCC, $n = 50$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0568	.0525(.0040)	.9254	.0635	.0606(.0020)	.9297	.0566	.0542(.0042)	.9354
10	.5	.0551	.0524(.0038)	.9441	.0617	.0598(.0019)	.9405	.0544	.0529(.0039)	.9447
	.9	.0641	.0593(.0041)	.9292	.0722	.0694(.0023)	.9225	.0638	.0603(.0042)	.9272
	0	.0559	.0534(.0042)	.9393	.0636	.0603(.0021)	.9259	.0567	.0505(.0043)	.9316
25	.5	.0533	.0524(.0042)	.9414	.0611	.0596(.0020)	.9389	.0530	.0512(.0041)	.9432
	.9	.0642	.0594(.0042)	.9452	.0710	.0685(.0022)	.9469	.0620	.0609(.0043)	.9451
	0	.0573	.0528(.0041)	.9258	.0641	.0593(.0020)	.9220	.0571	.0512(.0042)	.9340
40	.5	.0540	.0519(.0039)	.9398	.0604	.0587(.0019)	.9326	.0536	.0511(.0039)	.9305
	.9	.0655	.0594(.0042)	.9221	.0736	.0685(.0024)	.9182	.0641	.0576(.0041)	.9204

FCCC, $n = 100$

t	ρ	Intra- SD	Intra- SE(std)	Intra- CP	Inter- SD	Inter- SE(std)	Inter- CP	Total- SD	Total- SE(std)	Total- CP
	0	.0367	.0329(.0020)	.9303	.0424	.0386(.0012)	.9217	.0381	.0357(.0020)	.9276
	.5	.0395	.0357(.0021)	.9421	.0441	.0393(.0013)	.9305	.0393	.0351(.0021)	.9286
10	.9	.0418	.0383(.0020)	.9631	.0476	.454(.0015)	.9438	.0426	.0391(.0020)	.9594
	0	.0369	.0342(.0020)	.9396	.0416	.0386(.0013)	.9396	.0367	.0331(.0021)	.9396
	.5	.0382	.0353(.0022)	.9440	.0435	.0402(.0013)	.9382	.0380	.0368(.0021)	.9479
25	.9	.0424	.0386(.0024)	.9562	.0479	.0455(.0014)	.9512	.0419	.0399(.0024)	.9501
	0	.0382	.0339(.0021)	.9299	.0421	.0382(.0020)	.9258	.0379	.0338(.0019)	.9286
	.5	.0391	.0367(.0021)	.9402	.0438	.0406(.0022)	.9392	.0391	.0351(.0021)	.9363
40	.9	.0432	.0394(.0023)	.9531	.0480	.0465(.0016)	.9469	.0430	.0424(.0024)	.9500

Chapter 5

Core body temperature data

5.1 Introduction

We illustrate the proposed indices via an analysis of a subset of data collected in the Noll Physiological Research Center at Pennsylvania State University. Data were collected on 14 subjects, 7 obese and 7 lean. They exercised for 20 minutes on a motor driven treadmill, 20 minutes on a motor driven cycle and 20 minutes on a motor driven treadmill with 5 minutes of rest between exercises. Each of the 6 completed trials consisted of one exercise sequence followed by a recovery period. Core body temperature was measured 81 times for each trial by automatic portable measures of human physiology, and there are no missing data. The temperature and humidity levels did not change throughout the protocol. Of interest in this study is whether or not trial 6 is ineffective for burning body fat. It is known that prolonged exercise is effective for burning body fat because the body has adapts itself and starts to conserve energy. Since fat burning is related to body temperature, higher body temperature burns more body fat. Figure 5.1 presented actual core body temperature profiles for trials 5 and 6 for each subject with green

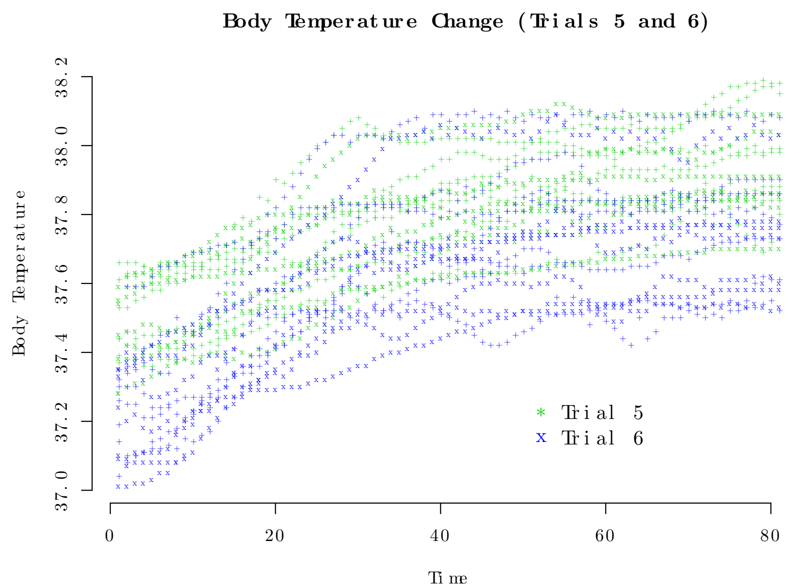

Figure 5.1. Plots for core body temperature against time. Observed individual trajectories are presented with green and blue color denoting trial 5 and 6 respectively.

and blue, respectively. The 14 green curves of trial 5 are generally higher than the 14 blue curves of trial 6. Thus it becomes important to measure how close the body temperatures of trials 5 and 6 are. If the body temperature of trial 6 is similar to the body temperature of trial 5, then trial 6 is as effective as trial 5 for losing body fat.

5.2 Functional mixed models

The 12 curves in Figure 5.2 are population means of 7 obese boys and 7 lean boys of 6 trials, and each point indicates readings for individuals. The 12 profiles of core body temperature is curved, and the curve feature is different by weight group and trial. Dots show the particular derivatives from population mean temperature, and the difference varies by weight group; lean boys have more variability then obese boys. Mixed effects models can be applied to take individual variety into account,

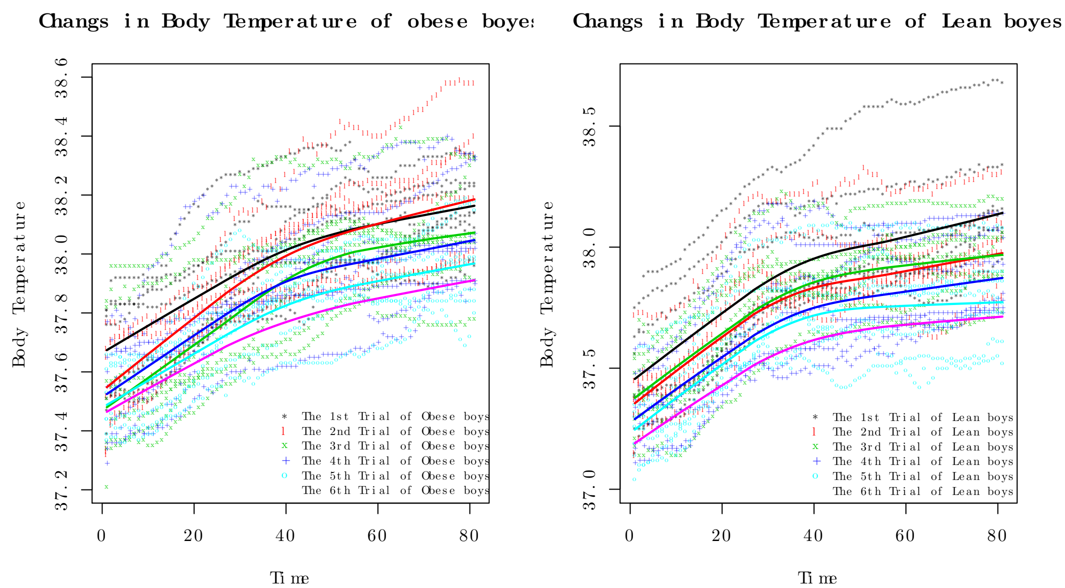

Figure 5.2. Plots for core body temperature against time. In the left panel, observed individual trajectories (dot) and the mean of six trials (solid line) of seven obese boys. In the right panel, actual individual trajectories (dot) and the mean of six trials (solid line) of seven lean boys.

but the effects are not linear over time. Thus, we fit the body temperature using the functional mixed model with a random effect of subject and fixed effects of weight group, trial, and the interaction between group and trial.

The corresponding mixed effects model is

$$y_{ijk}(t_l) = \beta_0(t_l) + a_i(t_l) + \alpha_j(t_l) + \gamma_k(t_l) + \eta_{jk}(t_l) + \varepsilon_{ijk}(t_l), \tag{5.1}$$

where $y_{ijk}(t_l)$ is the body temperature of the j_{th} group and k_{th} trial at time t_l, and $a_i(t)$ is the subject specific variability at time t_l, normally distributed with mean zero and $\text{cov}(y_{ijk}(t_l), y_{i'j'k'}(t_{l'})) = \sigma_a^2(t) = \sigma_a^2 f(t_l) f(t_{l'})$ if $i = i'$. α_j is the effect of the j_{th} weight group, $j = 1, 2$, γ_k is the effect of trials, $k = 1, \cdots, 6$, and η_k is the interaction effect of the weight group and trial. We applied the estimating procedure described in Chapter 3 to estimate coefficients and variance components in model (5.5). Figure 5.5 depicts actual body temperatures and fitted values and it

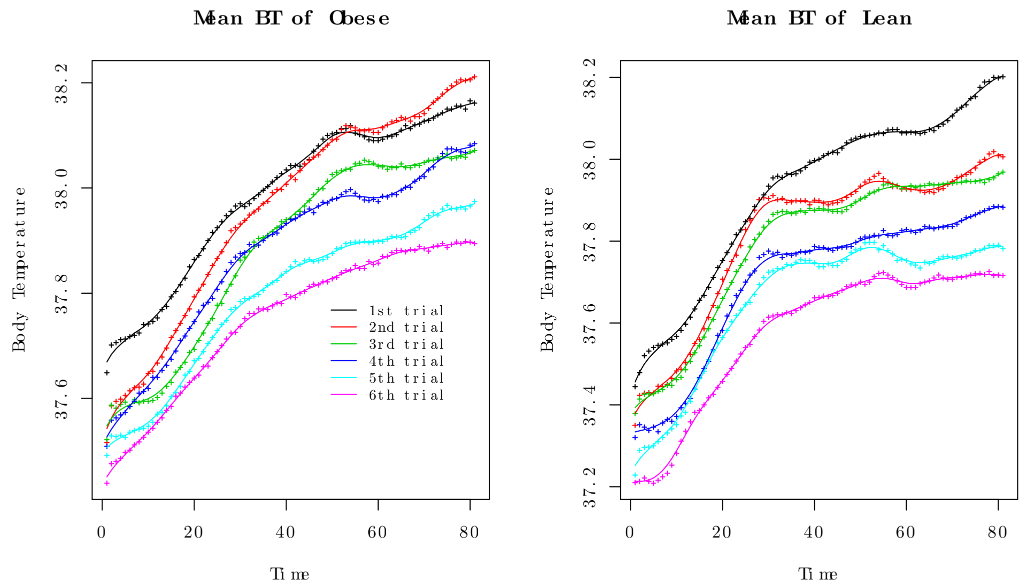

Figure 5.3. Plots for fitted curves of core body temperature against time. The points indicate the actual observations and solid curves refer to the fitted core body temperature by weight group: 6 trial mean curves of obese boys in the left panel and 6 trial mean curves of lean boys in the right panel.

shows that the model (5.5) explains the data well, and the time varying coefficients are estimated quite well by locally linear smoother with Epanechnikov kernel. Because weight group is a between-subject effect, FCCC has to be defined by the weight group. Define the lean group as a base group; then the obese boys have an additional term. The variance of trial effect by each group is redefined as

$$\begin{aligned} \sigma^2_{\gamma,o}(t) &= \frac{1}{30}\sum_{k=1}^{5}\sum_{k'=2}^{6}\left\{(\gamma_k(t)+\eta_k(t)) - (\gamma_{k'}(t)+\eta_{k'}(t))\right\}^2, \\ \sigma^2_{\gamma,l}(t) &= \frac{1}{30}\sum_{k=1}^{5}\sum_{k'=2}^{6}\left(\gamma_k(t)-\gamma_{k'}(t)\right)^2. \end{aligned} \tag{5.2}$$

The UCCC is

$$
\begin{aligned}
UCCC_{\text{intra}}(\text{Lean}) = UCCC_{\text{intra}}(\text{Obese}) &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\varepsilon^2}, \\
UCCC_{\text{inter}}(\text{Lean}) &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{\gamma,l}^2 + \sigma_\varepsilon^2/81}, \\
UCCC_{\text{inter}}(\text{Obese}) &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{\gamma,o}^2 + \sigma_\varepsilon^2/81}, \\
UCCC_{\text{total}}(\text{Lean}) &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{\gamma,l}^2 + \sigma_\varepsilon^2}, \\
UCCC_{\text{total}}(\text{Obese}) &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{\gamma,o}^2 + \sigma_\varepsilon^2}.
\end{aligned}
\tag{5.3}
$$

Because there is neither replication, intra-trial FCCC can not be evaluated, and inter-trial FCCC and total-trial FCCC are the same.

$$
\begin{aligned}
FCCC_{\text{total}}(\text{Lean}) &= \frac{\sigma_a^2(t)}{\sigma_a^2(t) + \sigma_{\gamma,l}^2(t) + \sigma_\varepsilon^2(t)}, \\
FCCC_{\text{total}}(\text{Obese}) &= \frac{\sigma_a^2(t)}{\sigma_a^2(t) + \sigma_{\gamma,o}^2(t) + \sigma_\varepsilon^2(t)}.
\end{aligned}
\tag{5.4}
$$

In Figure 5.4, FCCCs delineate agreement trends over time. FCCCs decrease along with time and are consistent with the profile feature of core body temperature in Figure 5.5; the gap between trials widens over time and it is reflected in the decreasing FCCCs. The difference for obese boys is smaller than lean boys; hence the FCCC of obese boys is larger than the FCCC of lean boys, but the difference is not significant because the 95% confidence interval of the FCCCs overlap. Both FCCCs of trials are between .4 and .6, which indicates moderate agreement. Inter UCCC gives a large number because 81 repetitions of each trial lead to a small denominator in equation (5.3). The FCCC between trials 5 and 6 is a slightly different trend from the overall inter-trial index. For lean boys, FCCC is small at time between 20 and 40 because the distance between the profile of trials 5 and 6

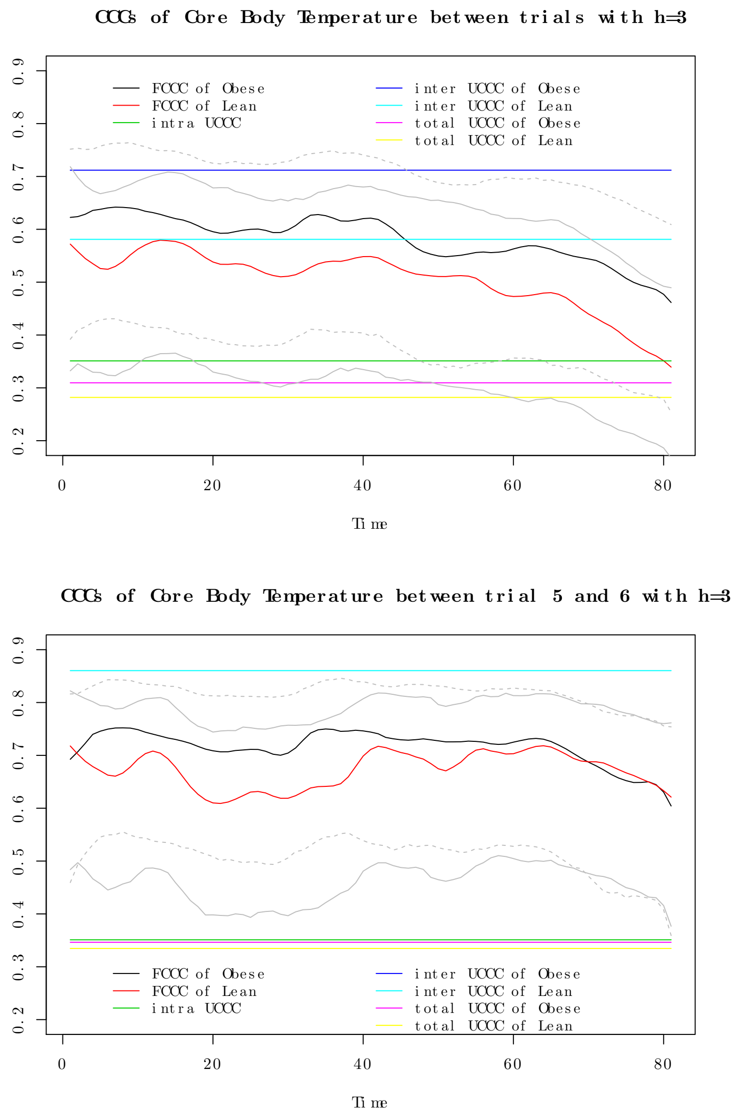

Figure 5.4. Plots for agreement indices of core body temperature over time. Black and red curves shows the inter-trial FCCC of two groups with 95% confidence interval and other five straight lines indicate the intra-, inter-, total-trial UCCCs. The upper panel shows the indices for 6 trials. The bottom panel depicts the indices for trial 5 and trial 6.

is wider than other time intervals (smaller accuracy) and has different incremental rate (smaller precision). Assuming equal weight over time, the FCCCs are summarized to one number in Table 5.2. Mean and standard deviation of each index were obtained via 1000 parametric bootstrap samples. The accuracy of the obese

Table 5.1. Comparison of agreement indexes

	UCCC		FCCC					
			$h=1$		$h=3$		h=6	
	Obese	Lean	Obese	Lean	Obese	Lean	Obese	Lean
$\text{CCC}_{\text{total}}$	.309	.282	.496 (.04)	.433 (.05)	.552 (.04)	.477 (.05)	.562 (.04)	.481 (.05)
Accuracy	.881	.803	.759 (.04)	.658 (.06)	.772 (.05)	.665 (.07)	.774 (.05)	.668 (.07)
Precision	.351		.648		.709		.720	

boys is higher than that of the lean boys because the body temperature difference between trials 5 and 6 of the lean boys is wider than the difference of obese boys. The agreement between trials 5 and 6 is decided by bootstrap confidence intervals; whether or not the intervals of obese and lean boys include 1 or a pre-specified tolerable value. Because the upper limit of the confidence intervals are around .85, we can't say that the body temperature of trial 5 and trial 6 are similar unless pre-specified tolerance is less than .85. Therefore trial 6 is not as effective as trial 5.

Based on our simulation results in Chapter 3, functional mixed models provide optimistic results for small samples when the data are highly correlated. Because core body temperatures are highly correlated over time, $\rho = .92$, with small sample size 14, the estimated confidence interval of FCCCs can be wider than actual. Thus, we investigate FCCCs based on the partially linear mixed model, which is relatively robust against the degree of strong serial correlation in the next section.

5.3 Partially linear mixed models

To construct the partially linear mixed model, we explore the data first. In Figure 5.2, the population change in body temperature over time is not linear. The gap between trials are different by weight group: obese boys generally have higher temperature than lean boys. Therefore, we consider weight group, trial, and the interaction between weight group and trial.

The corresponding partially linear mixed effects model is

$$y_{ijk}(t_l) = \mu(t_l) + a_i + \alpha_j + \gamma_k + \eta_{ij} + \varepsilon_{ijk}(t_l), \tag{5.5}$$

where $y_{ijk}(t_l)$ is the body temperature of j_{th} group and k_{th} trial at time t_l, and a_i is the subject specific variability, normally distributed with mean zero and variance σ_a^2. α_j is the effect of the j_{th} weight group, $j = 1, 2$, γ_k is the effect of trials, $k = 1, \ldots, 6$, and η_{jk} is the interaction effect of the weight group and trial. $\varepsilon_{ijk}(t_l)$ is assumed Gaussian process with mean zero, variance $\sigma_\varepsilon^2(t)$ and AR serial covariance. We estimate coefficients and variance components in model (5.5) by following the proposed estimating steps. In Figure 5.5, six fitted curves are near six observed

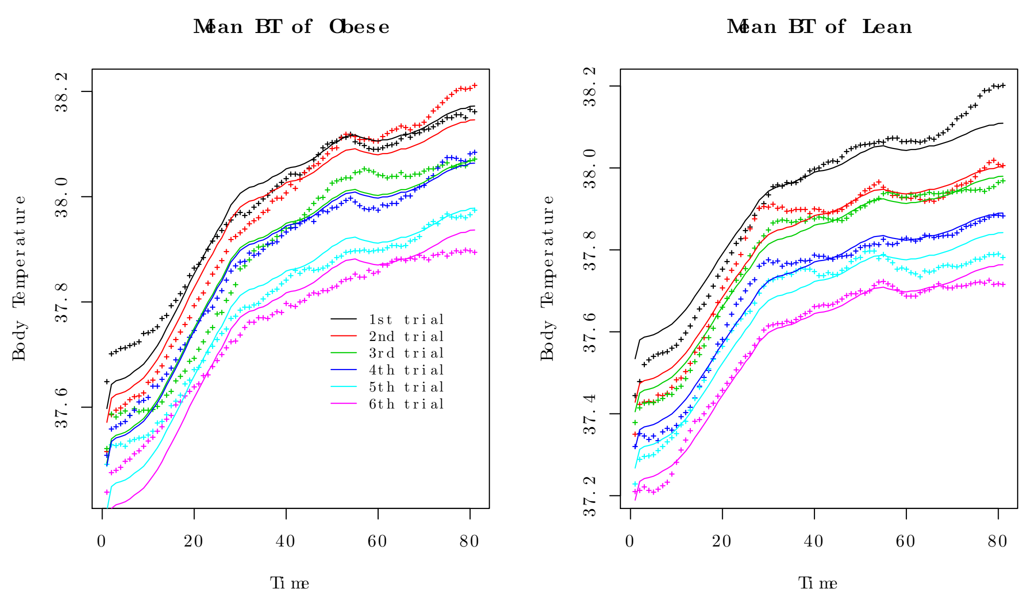

Figure 5.5. Plots for mean of core body temperature over time. The crosses indicate the actual observations and lines refer to the fitted values of obese boys in the left panel and lean boys in the right panel.

trial curves by obesity group. Because we consider a time-constant trial effect, six estimated curves are parallel over time which results a bias. Nevertheless, the 95%

bootstrap confidence interval for FCCC based on the partially linear mixed model may be a fairly conservative reference for strong serial correlation at small sample size.

There is neither replication nor other effects crossed with trial, so $FCCC_{\text{inter}}$ and $FCCC_{\text{total}}$ are the same and $FCCC_{\text{intra}}$ cannot be evaluated. UCCC and FCCC may be different because the curvature in a body temperature profile is considerable. Because there is a constant difference between weight groups in Figure 5.5, the index must be defined by two groups. Define the lean group as a base group; then obese boys have additional term. Since we include the interaction effect between trial and group, the trial effect by each group is redefined as

$$\begin{aligned}\sigma^2_{\gamma,o} &= \frac{1}{30}\sum_{k=1}^{5}\sum_{k'=2}^{6}\left\{(\gamma_k+\eta_k)-(\gamma_{k'}+\eta_{k'})\right\}^2, \\ \sigma^2_{\gamma,l} &= \frac{1}{30}\sum_{k=1}^{5}\sum_{k'=2}^{6}\left(\gamma_k-\gamma_{k'}\right)^2.\end{aligned}$$

The UCCCs are defined as follows:

$$\begin{aligned}UCCC^{\text{Lean}}_{\text{intra}} = UCCC^{\text{Obese}}_{\text{intra}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma_\varepsilon^2}, \\ UCCC^{\text{Lean}}_{\text{inter}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{\gamma,l}+\sigma_\varepsilon^2/81}, \\ UCCC^{\text{Obese}}_{\text{inter}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{\gamma,o}+\sigma_\varepsilon^2/81}, \\ UCCC^{\text{Lean}}_{\text{total}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{\gamma,l}+\sigma_\varepsilon^2}, \\ UCCC^{\text{Obese}}_{\text{total}} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{\gamma,o}+\sigma_\varepsilon^2}.\end{aligned} \tag{5.6}$$

The inter-trial FCCC is defined as follows:

$$FCCC_{\text{inter}}^{\text{Lean}} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{\gamma,l}^2 + \sigma_\varepsilon^2(t)},$$
$$FCCC_{\text{inter}}^{\text{Obese}} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{\gamma,o}^2 + \sigma_\varepsilon^2(t)}. \tag{5.7}$$

In Figure 5.6, FCCCs depict inter-trial agreement trend against time. The FCCC

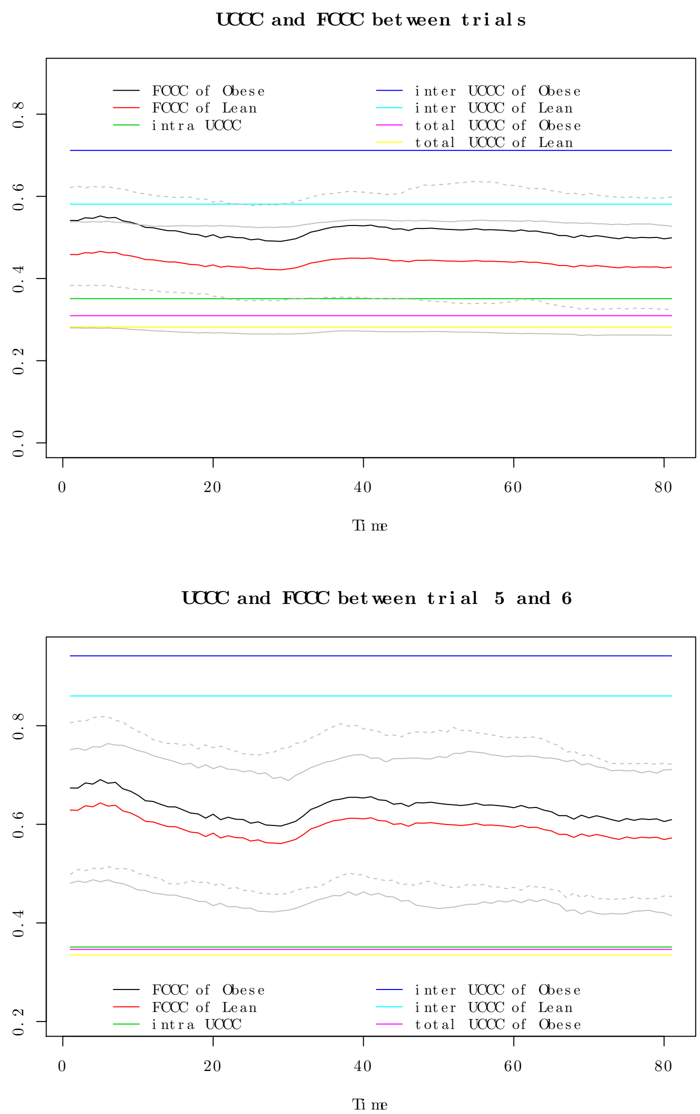

Figure 5.6. Plots for indices. Black and red curves shows the inter-trial FCCC of obese and lean boys and the other five straight lines indicate the intra-, inter-, total-trial UCCCs. The indices for 6 trials are shown in the upper panel and the indices between trials 5 and 6 are presented in the bottom panel.

is around .5 for overall trials and it increases to above .6 for trials 5 and 6. FCCCs decrease with time and are consistent with the trend of body temperature in Figure 5.2, whereas the gap between trials 5 and 6 widens over time. The difference in obese boys is smaller than lean boys; hence FCCCs of obese boys are larger than FCCCs of lean boys, but the difference between two parallel curves over time is not significant because the 95% confidence interval of FCCCs overlap. Inter UCCC is over .8 because 81 repetitions may cause overestimation by removing the influence of residual variance in equation (5.6). Assuming equal weight over time, FCCCs are summarized to one number in Table 5.2. Mean and standard deviation of each index were obtained via 1000 parametric bootstrap samples. Both FCCCs are around .6, which indicates moderate agreement. Whether or not the trials

Table 5.2. Functional Concordance Correlation Coefficient

	Obese			Lean		
	Intra-	Inter-	Total-	Intra-	Inter-	Total-
UCCC	.35	.94	.35	.35	.86	.34
FCCC	.63 (.19)			.60 (.20)		

5 and 6 have comparable core body temperature over time is decided based on the bootstrap confidence interval; body temperature of trials 5 and 6 are not comparable because bootstrap confidence intervals of both obese and lean boys between trials 5 and 6 does not include 1 unless pre-specified tolerable value is less than .8. Thus, trial 6 is not as effective as trial 5 for body fat burning.

Chapter 6

Future Research

6.1 Concluding Remarks

In this dissertation, we develop a time varying agreement index FCCC. The index is robust against functional form in models, and informative about the trend of measurement agreement over time period. We also propose a new estimation procedure for the purpose of estimating the FCCC for ILD. In Chapter 3, we found the use of semiparametric variance covariance of the random process in the functional mixed model increases the efficiency of the bootstrap variance of the random process by reducing the uncertainty via separating the source of variability as keeping the flexibility for the covariance function. The use of semiparametric variance also improves the efficiency of the estimates of FCCC. In Chapter 4, we found that if homogeneous covariance is incorrectly assumed, mixed models produce highly biased estimates of variance components for highly serially correlated data. We also found that biased estimates of variance components also create biased estimates of FCCC and its standard error. The use of serial correlation in the estimation and bootstrap inference improve the accuracy of the estimate of FCCC. FCCC based on partially linear mixed model is flexible enough to accommodate data structure even for highly correlated data with small sample $n = 50$.

6.2 Future work

6.2.1 Semiparametric varying coefficient partially linear mixed models

The functional mixed model and the partially linear mixed model are flexible for data structure and serial correlation for ILD. The functional mixed model fits ILD exactly, although the model requires sufficiently large sample size ($n \geq 100$) for strong serial correlation. The partially linear mixed model is robust against the degree of serial correlation for small sample size, but it has relatively large MSE by assuming the time-constant impact of covariates. The compromise between the two models is the semiparametric varying coefficient partially linear mixed model:

$$y(t) = x_1^T(t)\alpha(t) + x_2^T\beta + z_1(t)^T\nu(t) + z_2^T\eta + \varepsilon(t), \tag{6.1}$$

where $\alpha(t)$ comprises p_1 unknown smooth functions, β is a p_2 dimensional unknown parameter vector, $z_1(t)$ is q_1 vector of the individual subject contributions at time t considered as realizations of a mean 0 process, and z_2 is a q vector of random effects with mean 0.

We are interested in estimation of the variance components as well as the baseline function and regression coefficients. Based on our research, we will propose new estimating procedures for the semiparametric varying coefficient partially linear mixed model. For model (6.1), we are also interested in conducting variable selection for β and model selection for $\alpha(t)$. Furthermore, we would like to consider how to make proper statistical inference for variance components.

6.2.2 Bandwidth selection for ILD

In the estimation of the functional mixed model, we used specified bandwidths to explore the impact of bandwidth on the performance of FCCCs. A bandwidth controls the model complexity of the local linear fit and controls the bias and variance trade-off. Thus the performance of locally linear smoothers for time-varying coefficients largely depends on choice of bandwidth h. Bandwidth can be subjectively chosen by a trial-error approach via visually inspecting estimates in the graph, or chosen by data-driven rules via minimizing theoretical risk. Cross-validation and plug-in methods are two popular data-driven rules. Plug-in bandwidth is an asymptotically chosen optimal constant derived by minimizing the conditional weighted MISE (or other loss function). It provides a reliable bandwidth theoretically and practically for independent data. In the case of correlated data, however, the plug-in approach provides wide bandwidth in our simulation result. Another approach is cross-validation (CV). The basic idea of CV is to prevent overfitting induced by data reuse on estimating coefficients and measuring residuals; it holds out part of the sample with which to evaluate the performance of a predictor. It is conceptually simple. Wu & Zhang (2002) proposed four types of bandwidth criterion: leave-one-point-out cross-validation (CVP), leave-one-subject-out cross-validation (CVS), hybrid bandwidth (CVH) and bias-corrected hybrid bandwidth (CVBH). Hart & Wehrly (1993) show that the CVS is consistent. Based on the fact that individuals are independent, the CVS criterion is

$$CVS(h) = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\left[y_{ij} - x_i\hat{a}^{(-i)}(t_{ij}) - z_i\hat{c}_i^{(-i)}(t_{ij})\right]^2 /(nn_i) \tag{6.2}$$

where $\hat{a}^{(-i)}(t)$ represents the estimator of a based on the data with the measurements on subject i excluded, and the weights $1/(nn_i)$ take into account the number of measurements from individual subjects. The optimal bandwidth h_s is defined as the minimizer of CVS(h), appropriate for a population curve because CVS(h) does not rely on a structure of within-subject correlation.

Observations between time points are not related if random effects are given under the i.i.d. assumption of measurement errors. Assume that $t_1, \cdots, t_L$ is the set of distinct time points among an occasion set. For a given time t_l, there are M_{il} observations for subject i_l, $i_l = 1, \cdots, n_l$, $p = 1, \cdots, M_{il}$.

$$CVP(h) = \sum_{l=1}^{L}\sum_{i_l=1}^{n_l}\sum_{p=1}^{M_{il}}\left[y_i(t_p) - x_{ip}^T\hat{a}^{(-l)}(t_l) - z_{ip}^T\hat{c}_i^{(-l)(t_p)}\right]^2 /(M_{il}n_lL) \tag{6.3}$$

where $\hat{a}^{(-l)}(t_p)$ and $c_i^{(-l)}(t_p)$ represent the estimator of $a(t_p)$ and $c_i(t_p)$ based on the data after excluding all of the measurements obtained at time point l, and the weights $1/M_{il}n_lL$ take into account the number of measurements from an occasion set. The optimal bandwidth h_p obtained by minimizing CVP(h) is appropriate for estimating individual derivative curves.

Hybrid bandwidth is the combination of CVS bandwidth for a population curve and CVP bandwidth for individual derivations. Fan & Zhang (2000) showed that choosing different bandwidths for each functional coefficient performs better than using the one best bandwidth and suggested a two-step estimation method for coefficient-wise bandwidth. The two-step method consists of a raw estimation step and a smoothing step; raw estimators are obtained via least squares at every distinct time point, then refined estimator is obtained through existing smoothing techniques allowing different smoothnesses in each functional coefficient. The two-step method is better for setting different bandwidth for fixed effects. The method is flexible in the sense that it allows different smoothing techniques for different functional coefficients as well as different smoothnesses, and the estimates of the two-step approach outperform one-step estimation under independent observations. The two-step method has difficulty obtaining raw estimates for sparse and unequally measured data.

It is difficult to establish a systematic procedure for different bandwidths of fixed and random effects when they have different smoothnesses. We are interested in the bandwidth selection of fixed effects and random effects for ILD.

6.2.3 Interval approach for assessment of measurement agreement

Indices for assessment of measurement agreement have been widely used. However, indices depend on the covariance between subject variance; large between-subject variability would imply a large value of CCC even if the individual difference between measurements by the two methods remains the same. Another consequence of this, covariance based index might be overestimated if potential confounding variables are not taken into account. On the other hand, covariance based index is that the index might underestimate measurement agreement by including confounding covariates which reduce the range of the variable in the study. FCCC can provide the degree of agreement and make an inference at each time point. We are interested in the unscaled agreement measuring tool which can provide simultaneous agreement over time. Lin et al. (2007) proposed the tolerance deviation index (TDI) and coverage probability (CP) approach for repeated measures taken by each of the multiple methods with multiple readings being compared. Because the estimation and inference are based on the assumption of time constant mean and variance, we will propose simultaneous tolerance band of agreement over time for ILD.

Bibliography

Altman, N. (1990). Kernel smoothing of data with correlated errors. *Journal of the American Statistical Association*, *85*(411), 749–759.

Barnhart, H., Haber, M., & Song, J. (2002). Overall concordance correlation coefficient for evaluating agreement among multiple observers. *Biometrics*, *58*(4), 1020–1027.

Barnhart, H., Kosinski, A., & Haber, M. (2007). Assessing individual agreement. *Journal of Biopharmaceutical Statistics*, *17*(4), 697–719.

Barnhart, H., Song, J., & Haber, M. (2005). Assessing intra, inter and total agreement with replicated readings. *Statistics in Medicine*, *24*(9), 1371–1384.

Carrasco, J., & Jover, L. (2003). Estimating the generalized concordance correlation coefficient through variance components. *Biometrics*, *59*(4), 849–858.

Chinchilli, V., Martel, J., Kumanyika, S., & Lloyd, T. (1996). A weighted concordance correlation coefficient for repeated measurement designs. *Biometrics*, *52*(1), 341–353.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220.

Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, *87*(420), 998–1004.

Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. *Annals of Statistics*, *21*, 196–196.

Fan, J., & Gijbels, I. (1992). Variable bandwidth and local linear regression smoothers. *Annals of Statistics*, *20*, 2008–2008.

Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications*. London: Chapman and Hall.

Fan, J., Huang, T., & Li, R. (2007). Analysis of longitudinal data with semiparametric estimation of covariance function. *Journal of the American Statistical Association*, *102*(478), 632–641.

Fan, J., & Li, R. (2004). New estimation and model selection procedures for semiparametric modeling in longitudinal data analysis. *Journal of the American Statistical Association*, *99*, 710–723.

Fan, J., & Li, R. (2006). An overview on nonparametric and semiparametric techniques for longitudinal data. In J. Fan, & H. Koul (Eds.) *Frontiers in statistics*, (pp. 277 –303). London: Imperial College Press.

Fan, J., & Zhang, J. (2000). Two-step estimation of functional linear models with applications to longitudinal data. *Journal of the Royal Statistical Society. Series B*, *62*(2), 303–322.

Fleiss, J. (1986). Reliability of measurement. In *The Design and Analysis of Clinical Experiments*, (pp. 1–32). New York: John Wiley & Sons.

Fleiss, J., Cohen, J., & Everitt, B. (1969). Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, *72*(5), 323–327.

Guo, W. (2002). Functional mixed effects models. *Biometrics*, *58*(1), 121–128.

Härdle, W., Liang, H., & Gao, J. (2000). *Partially linear models*. Heidelberg: Physica Verlag.

Hart, J., & Wehrly, T. (1993). Consistency of cross-validation when the data are curves. *Stochastic Processes and their Applications*, *45*(2), 351–361.

Heckman, N. (1986). Spline smoothing in a partly linear model. *Journal of the Royal Statistical Society. Series B*, *48*, 244–248.

King, T. (2007). A class of repeated measures concordance correlation coefficients. *Journal of Biopharmaceutical Statistics*, *17*(4), 653–672.

King, T., & Chinchilli, V. (2001a). A generalized concordance correlation coefficient for continuous and categorical data. *Statistics in Medicine*, *20*(14), 2131–2147.

King, T., & Chinchilli, V. (2001b). Robust estimators of the concordance correlation coefficient. *Journal of Biopharmaceutical Statistics*, *11*(3), 83–105.

King, T., Chinchilli, V., & Carrasco, J. (2007). A repeated measures concordance correlation coefficient. *Statistics in Medicine*, *26*(16), 3095–3113.

Li, R., & Chow, M. (2005). Evaluation of reproducibility for paired functional data. *Journal of Multivariate Analysis*, *93*(1), 81–101.

Lin, L., AS, H., & W., W. (2007). A Unified Approach for Assessing Agreement for Continuous and Categorical Data. *Journal of Biopharmaceutical Statistics*, *17*(4), 629–652.

Lin, L., Hedayat, A., Sinha, B., & Yang, M. (2002). Statistical Methods in Assessing Agreement: Models, Issues, and Tools. *Journal of the American Statistical Association*, *97*(457), 257–271.

Lin, L. I. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, *45*(1), 255–268.

Lin, X., & Carroll, R. (2001a). Semiparametric regression for clustered data. *Biometrika*, *88*(4), 1179–1185.

Lin, X., & Carroll, R. (2001b). Semiparametric Regression for Clustered Data Using Generalized Estimating Equations. *Journal of the American Statistical Association*, *96*(455), 1045–1056.

Morris, J. (2002). The blups are not "best" when it comes to bootstrapping. *Statistics & Probability Letters*, *56*(4), 425–430.

Moyeed, R., & Diggle, P. (1994). Rates of converence in semi-parametric modeling of longitudinal data. *Australian Journal of Statistics*, *36*(1), 75–93.

Park, J., & Wu, H. (2006). Backfitting and local likelihood methods for nonparametric mixed-effects models with longitudinal data. *Journal of Statistical Planning and Inference*, *136*(11), 3760–3782.

Pinheiro, J., & Bates, D. (2009). *Mixed-effects models in S and S-PLUS*. New York: Springer Verlag.

Shrout, P., & Fleiss, J. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*, *86*(2), 420–428.

Speckman, P. (1988). Kernel smoothing in partial linear models. *Journal of the Royal Statistical Society. Series B*, *50*(3), 413–436.

Wu, H., & Liang, H. (2004). Backfitting Random Varying-Coefficient Models with Time-Dependent Smoothing Covariates. *Scandinavian Journal of Statistics*, *31*(1), 3–19.

Wu, H., & Zhang, J. (2002). Local polynomial mixed-effects models for longitudinal data. *Journal of the American Statistical Association*, *97*(459), 883–898.

Wu, H., & Zhang, J. (2006). *Nonparametric regression methods for longitudinal data analysis*. Hoboken, New Jersey: John Wiley & Sons.

Zeger, S., & Diggle, P. (1994). Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. *Biometrics*, *50*(3), 689–699.

Zhang, D., Lin, X., Raz, J., & Sowers, M. (1998). Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association*, *93*(442), 710–719.

Vita

Mina Yoo

Education

Ph.D. in Statistics August 2011
The Pennsylvania State University, USA
M.S. in Statistics February 2004
Ewha Womans University, Korea
B.S. in Statistics February 2002
B.A. in Business Administration February 2002
Ewha Womans University, Korea

Professional Experience

Research Assistant *05/2006-08/2006, 05/2007-08/2007 and 05/2009-Present*
Advisor: Prof. Runze Li, Department of Statistics, PSU
Research Summer Assistant *06/2008 - 08/2008*
Advisor: Dr. Jinglin Zhong, Office of Biostatistics, CDER, FDA
Research Assistant *March 2004 - May 2005*
Advisor: Dr. Youngdai Kim, Statistical Research Center for Complex Systems, Seoul National University

Teaching Experience

Instructor *Fall 2008*
Department of Statistics, The Pennsylvania State University, USA
Teaching Assistant *2005 - 2008*
Department of Statistics, The Pennsylvania State University, USA