Statistics and Probability Letters 80 (2010) 390–402

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Distribution-free tests for no effect of treatment in heteroscedastic functional data under both weak and long range dependence

Haiyan Wang [a,*], James Higgins [a], Dale Blasi [b]

[a] Department of Statistics, Kansas State University, United States

[b] Department of Animal Sciences & Industry, Kansas State University, United States

ARTICLE INFO

Article history:
Received 3 September 2009
Received in revised form 9 October 2009
Accepted 19 November 2009
Available online 26 November 2009

MSC:
Primary 62G10
62G35

ABSTRACT

In this paper, we present distribution-free tests to evaluate the effect of multiple treatments when there are a large number of repeated measurements from each subject nested in a treatment. We formulate new test statistics to account for heteroscedasticity and unbalanced designs. The asymptotic distributions for the test statistics are obtained when the repeated measurements from the same subject have long range dependence and weak dependence, respectively. The asymptotic results hold under the nonclassical setting in which the number of repeated measurements is large while the number of subjects per treatment may be small. A real application to compare cattle ear temperature profiles under different antibiotic treatments is given for illustration. Simulation studies are undertaken to compare the empirical performance of the proposed tests to commonly used methods.

1. Introduction

Functional data involve many repeated measurements from the same experimental unit such as subjects or clusters. Such data may be collected in traditional long-term agricultural experiments to compare the effects of different treatments (Aulakh et al., 1991; Barber, 1979; McCollum, 1991; Bailey et al., 1996) or from modern advanced technology such as automatic temperature measurements through cattle ear tags, brain image data through functional magnetic resonance imaging etc. A major interest is to assess the effects of different treatments using the collected functional data.

Common methods to evaluate the effect of treatment with longitudinal data include linear and generalized linear mixed models (Liang and Zeger, 1986; Breslow and Clayton, 1993), and generalized additive mixed models (Lin and Zhang, 1999; Wood, 2008 and the references therein). The inference of these models were mainly based on (penalized quasi-) likelihood for data from a Gaussian or an exponential family through a known link function. Apart from the required distributional assumptions, the inference procedures often need an estimation of the inverse of unknown large covariance matrices that is a challenge when there are only a small number of subjects per treatment. In practice, it is typically advised that various serial correlation structures be fitted to the model and a 'best' correlation structure be chosen based on some model selection criterion such as Akaike Information Criterion. This assumes that the 'best' structure selected is an approximately correct covariance structure so that using the Kenward–Roger adjustment for the degrees of freedom works well (Guerin and Stroup, 2000; Littell, 2002; Loughin, 2006). However, there is often a non-trivial percentage of non-convergence and extensive computational time issues when the number of repeated measurements is large. As a consequence, the 'best' is often selected out of those that can be fitted. In addition, estimation of a large covariance matrix with a small number of subjects per treatment remains an unsolved challenge even though some efforts with additional simplifying assumptions have been pursued by some authors (cf. Fan et al., 2008; Schäfer and Strimmer, 2005; Zhu and Hero, 2007).

* Corresponding author. Tel.: +1 785 532 0519.
E-mail addresses: hwang@ksu.edu (H. Wang), jhiggins@ksu.edu (J. Higgins), dblasi@ksu.edu (D. Blasi).

0167-7152/$ – see front matter
doi:10.1016/j.spl.2009.11.016

In this paper, we consider the inference for testing hypotheses on the treatment for functional data when the data have unknown heteroscedastic distributions to allow more generality, i.e., we assume the observations for a randomly selected subject have unknown marginal distribution function that depends on the treatment and time. Cochran (1939) recognized that serial correlations between measurements taken relatively close together in time are likely to be greater than those taken far apart in time. He therefore did not recommend analysis of repeated measures using an equal correlation structure. Particularly, the data example that we will consider has lagged correlations that converge to zero as the time lag increases (cf. Fig. 2). Such correlations can be expressed as strong mixing or α-mixing (Rosenblatt, 1956; Billingsley, 1995). We will restrict our attention to α-mixing processes without imposing stationary assumption. The test procedures to be developed will adopt the asymptotic setting that the number of time points is large and the number of treatment levels is small. The sample sizes n_i are allowed to be small.

2. The model without parametric distributional assumption

Consider subjects nested within a total of a factor levels and each subject is measured at b time points. Suppose we have n_i randomly selected subjects, $S_{k(i)}, k = 1, \ldots, n_i$, nested in level i. Each randomly selected subject $S_{k(i)}$ generates a time series

$$\mathbf{X}_{ik} = (X_{i1k}, \ldots, X_{ibk})', \quad i = 1, \ldots, a.$$

Wang and Akritas (2009) considered nonparametric tests for no main effect of time and no treatment by time interaction effect in the following model

$$X_{ijk} = \mu + \alpha_i + \beta_j + B_{ik} + \gamma_{ij} + e_{ijk}, \tag{2.1}$$

where the parameters α_i are the main treatment effects, β_j are the main time effects, γ_{ij} are the treatment by time interaction effects, and B_{ik} are the subject specific random intercepts. The parameters were defined through a decomposition:

$$\mu_{ijS_{k(i)}} = E(X_{ijS_{k(i)}}|S_{k(i)}), \qquad \mu = (ab)^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b} E(\mu_{ijS_{k(i)}}), \qquad \alpha_i = b^{-1}\sum_{j=1}^{b} E(\mu_{ijS_{k(i)}}) - \mu,$$

$$B_{ik} = b^{-1}\sum_{j=1}^{b}\mu_{ijS_{k(i)}} - \mu - \alpha_i, \qquad \beta_j = a^{-1}\sum_{i=1}^{a} E(\mu_{ijS_{(i)}}) - \mu, \qquad \gamma_{ij} = E(\mu_{ijS_{k(i)}}) - \mu - \alpha_i - \beta_j,$$

and e_{ijk} is a composite random term that includes the measurement error and some random subject by time interactions $D_{ijS_{k(i)}} = \mu_{ijS_{k(i)}} - \mu - \alpha_i - \beta_j - B_{ik} - \gamma_{ij}$. The following constraints are a result of the decomposition: $\sum_i \alpha_i = \sum_j \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = E(B_{ik}) = E(e_{ijk}) = 0$.

It is typical to assume that the time series $\mathbf{X}_{ik}$ are independent for different k and different i. However, the subject specific random effect B_{ik} and the composite error e_{ijk} are not independent due to the random interactions included in e_{ijk}. The hidden random interactions can be easily justified from the Bayes regression model in which both the data and the model parameters come from some unknown stochastic process (cf. Morris, 1983), or from a multi-level mixed model (hierarchical linear model) commonly used in educational research where the response is related to a set of subject level predictors via a linear model for each subject at the first stage, and the model parameters from the first stage are used as the response variables in a second stage model (Raudenbush and Bryk, 2002)). When the models from the multi-levels are combined together, a random interaction term exists to account for cross-level interaction effects between variables located at different levels (for example, school-level funding may positively affect individual-level student performance by way of recruiting superior teachers, made possible by superior financial incentives).

In this paper, we consider inference to evaluate the effect **related to the treatment** as the number of repeated measurements per subject $b \to \infty$ under the heteroscedastic setting when the number of subjects per treatment may be small. We will test the hypotheses of no simple effect of treatment $H_0(\phi)$: all $\phi_{ij} = \alpha_i + \gamma_{ij} = 0$, and no main effect of treatment $H_0(\alpha)$: all $\alpha_i = 0$, or, more generally, $\widetilde{H}_0(\alpha)$: $\mathbf{C}_a\boldsymbol{\alpha} = \mathbf{0}$, where $\mathbf{C}_a$ is a contrast matrix and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_a)'$, when there are no distributional assumptions for the response and the covariance matrices $\mathrm{Var}(\mathbf{X}_{ik}) = \Sigma_i = \left(\sigma_{ijj'}\right)_{b\times b}$ may be different for subjects in different treatments. The hypothesis $\widetilde{H}_0(\alpha)$ is convenient to evaluate the effect of a single factor when the a groups are factor level combinations of several factors. The assumption of heteroscedastic covariance matrices is more general and realistic when the data exhibit different autocorrelation patterns for different treatments (cf. Fig. 2). The test of interaction effect $H_0(\gamma)$: $\gamma_{ij} = 0$ under weak dependence was considered in Wang and Akritas (2009). Here, we will give the test for interaction effect under strong dependence. The fact that the model parameters of model (2.1) lie in an infinite-dimensional space as $b \to \infty$ and there are no distributional assumptions for the data poses major differences between the present model with the Bayes regression model or multi-level mixed models.

To allow for asymptotics, a certain control over the lagged correlation is necessary. We assume that the time series $\{X_{ijk}, j = 1, \ldots\}$ satisfy an α-mixing condition with decay rate $\alpha_m \to 0$. This condition does not imply that $\mathrm{Var}(B_{ik}) = 0$ for all i, k. In fact, $\mathrm{Var}(B_{ik})$ can be any non-negative values and can be different for different i as long as the α-mixing condition for the observations from the same subject is satisfied. However, it is cumbersome to work with B_{ik} and e_{ijk} in the model while the assumption of α-mixing is imposed on the observations. Therefore, we combine the two terms into a single term $u_{ijk} = B_{ik} + e_{ijk}$. Then the α-mixing condition for the observations implies that $\{u_{ijk}, j = 1, \ldots\}$ is an α-mixing time series. When only working with the fixed effects and u_{ijk}, the model is equivalent to the marginal model used in generalized estimating equation approach (GEE), in which all randomness is attributed to a composite random term containing all possible random subject effect, serial correlations, and measurement errors. However, GEE fail to provide reliable type I error and power in this setting due to violation of requirement for a large number of subjects and a small number of time points. The majority of inferences with α-mixing processes in the literature are focused on stationary processes (cf. Billingsley, 1995; Doukhan et al., 2003). In practice, however, the stationary assumption is often violated. We will proceed **without** the stationary assumption to allow for wide applications.

The following notation will be used in this article: $n = \sum_{i=1}^{a} n_i$, $N = nb$, $\overline{X}_{ij.} = \frac{1}{n_i}\sum_{k=1}^{n_i} X_{ijk}$, $\widetilde{X}_{i..} = \overline{X}_{i..} = \frac{1}{bn_i}\sum_{j=1}^{b}\sum_{k=1}^{n_i} X_{ijk}$, $\widetilde{X}_{...} = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_i}\sum_{k=1}^{n_i} X_{ijk}$, $\widetilde{X}_{.j.} = \frac{1}{a}\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i} X_{ijk}$, $\overline{X}_{i.k} = \frac{1}{b}\sum_{j=1}^{b} X_{ijk}$.

3. Main results

The decay rate α_m of α-mixing has a crucial control over the convergence rate of a test statistic. When the order of α_m is greater than $O(m^{-1})$, the α-mixing process has long memory dependence (or strong dependence) in that the summation of the lagged correlation diverges; when $\alpha_m = O(m^{-\nu})$, for some $\nu > 1$, the process has short memory or weak dependence. We separate the results when the data are weakly dependent with those when the data have strong dependence.

3.1. Result under weak dependence

For the test of no simple effect of treatment $H_0(\phi)$, consider the following adjusted mean squares

$$AM\phi = \frac{1}{(a-1)b}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{X}_{ij.} - \widetilde{X}_{.j.}\right)^2 \quad \text{and} \quad AME = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i}\frac{(X_{ijk} - \overline{X}_{ij.})^2}{n_i(n_i-1)}.$$

Theorem 3.1. *Assume that for each group i and subject k,* $X_{ijk}, j = 1, 2, \ldots$, *is* α*-mixing with* $\alpha_m = O\left(m^{-5}\right)$. *In addition, assume that* $\limsup_j E[(X_{ijk} - E(X_{ijk}))^{32}] < \infty$. *Let* $n(a) = \min_i\{n_i\}$ *and*

$$\zeta_1 = \frac{2}{a^2 b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\sum_{i=1}^{a}\frac{\sigma_{ijj'}^2}{n_i(n_i-1)}, \qquad \zeta_2 = \frac{2}{a^2 b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\sum_{i\neq i'}^{a}\frac{\sigma_{ijj'}\sigma_{i'jj'}}{n_i n_{i'}}.$$

Then under $H_0(\phi)$,

(1) *for* $n_i \geq 4$ *bounded,* $\sqrt{b}(AM\phi - AME) \xrightarrow{d} N\left(0, \tau_\gamma^2\right)$, *where* $\tau_\gamma^2 = \lim_{b\to\infty}\left(\zeta_1 + \frac{\zeta_2}{(a-1)^2}\right)$;

(2) *if* $n_i \to \infty$ *as* $b \to \infty$, *under the additional assumption* $\max_i\{n_i\}/n(a) = O(1)$, *we have under* $H_0(\phi)$, $\sqrt{b}n(a)(AM\phi - AME) \xrightarrow{d} N\left(0, \tau_{\gamma*}^2\right)$, *where* $\tau_{\gamma*}^2 = \lim_{b\to\infty} n(a)^2(\zeta_1 + \zeta_2/(a-1)^2)$.

The existence of the limits of ζ_1 and ζ_2 along with their consistent estimators were given in Wang and Akritas (2009) based on the thresholding idea in which a tapered covariance matrix was used. For convenience, the estimators of ζ_1 and ζ_2 are restated below:

For each $j = 1, \ldots, b$, let $C_u(j, h) = [\min\{b, j + b^h\}]$, $C_l(j, h) = [\max\{0, j - b^h\}]$, for some $0 < h < 1$, where $[x]$ denotes the largest integer less than or equal to x. Let

$$\widehat{\zeta}_1 = \frac{2}{a^2 b}\sum_{j=1}^{b}\sum_{j'=C_l(j,h)}^{C_u(j,h)}\sum_{i=1}^{a}\frac{\widehat{\sigma_{ijj'}^2}}{n_i(n_i-1)}, \qquad \widehat{\zeta}_2 = \frac{2}{a^2 b}\sum_{j=1}^{b}\sum_{j'=C_l(j,h)}^{C_u(j,h)}\sum_{i\neq i'}^{a}\frac{\widehat{\sigma}_{ijj'}\widehat{\sigma}_{i'jj'}}{n_i n_{i'}},$$

where

$$\widehat{\sigma}_{ijj'} = \sum_{k=1}^{n_i}(X_{ijk} - \overline{X}_{ij.})(X_{ij'k} - \overline{X}_{ij'.})/(n_i - 1), \tag{3.1}$$

$$\widehat{\sigma_{ijj'}^2} = \sum_{k_1\neq k_2\neq k_3\neq k_4}^{n_i}\frac{(X_{ijk_1} - X_{ijk_2})(X_{ij'k_1} - X_{ij'k_2})(X_{ijk_3} - X_{ijk_4})(X_{ij'k_3} - X_{ij'k_4})}{4n_i(n_i-1)(n_i-2)(n_i-3)}. \tag{3.2}$$

Let $\widehat{\tau}_\gamma^2 = \widehat{\zeta}_1 + \widehat{\zeta}_2/((a-1)^2)$ and $\widehat{\tau}_{\gamma*}^2 = n(a)^2\widehat{\tau}_\gamma^2$. Then under the assumptions of Theorem 3.1, $\widehat{\tau}_\gamma^2$ is a consistent estimator of τ_γ^2 if $b \to \infty$ and n_i stay bounded; if both b and n_i go to ∞, $\widehat{\tau}_{\gamma*}^2$ is a consistent estimator of $\tau_{\gamma*}^2$.

To allow for heteroscedasticity, we propose a Wald-type statistic for testing $H_0(\alpha)$ or $\widetilde{H}_0(\alpha)$. The exact form of this statistic and its asymptotic distribution are given in Theorem 3.2.

Theorem 3.2. *Let* $\mathbf{W} = (\overline{X}_{1..}, \ldots, \overline{X}_{a..})'$ *and* $\mathbf{C}_a$ *be a contrast matrix with rank r. If* $X_{i1k}, X_{i2k}, \ldots,$ *is* α*-mixing with* $\alpha_m = O\left(m^{-5}\right)$ *for all i, k, and* $\limsup_j E[(X_{ijk} - E(X_{ijk}))^{16}] < \infty$*, let*

$$\widehat{\eta}_i = \frac{n}{bn_i(n_i-1)} \sum_{j=1}^{b} \sum_{j'=C_l(j,h)}^{C_u(j,h)} \sum_{k=1}^{n_i} \left(X_{ijk} - \overline{X}_{ij.}\right)\left(X_{ij'k} - \overline{X}_{ij'.}\right).$$

Then under $\widetilde{H}_0(\alpha)$*,*

$$N\mathbf{W}'\mathbf{C}_a'\left[\mathbf{C}_a \operatorname{diag}(\hat{\eta}_1, \ldots, \hat{\eta}_a)\mathbf{C}_a'\right]^{-1}\mathbf{C}_a\mathbf{W} \xrightarrow{d} \chi_r^2,$$

as $b \to \infty$ *regardless of whether the* n_i *remain bounded or tend to* ∞ *with b provided that* $\max_i n_i/n(a) = O(1)$ *and* $h < 0.5$.

Remark 1. For each j, the covariance between X_{ijk} and neighboring observations from the same subject are calculated and accumulated for the calculation of $\widehat{\zeta}_1$, $\widehat{\zeta}_2$ and $\widehat{\eta}_i$. Therefore, the overall calculation is equivalent to collecting correlations between every pair of big overlapping moving blocks.

Remark 2. The finite higher order moment conditions in Theorems 3.1 and 3.2 are necessary since they are required for the Central Limit Theorem (CLT) for weak dependent processes (cf. Wang and Akritas, 2009). Similar moment condition (finite 12th moment) was also imposed for the CLT for stationary α-mixing processes on page 364 in Billingsley (1995). Such conditions can be relaxed by considering (mid-)rank statistics. However, this is beyond the scope of this paper.

3.2. Asymptotic distribution under long memory dependence

When there are long memory dependence among the observations from the same subject, the behavior of the test statistics driven by a large number of repeated observations from a small number of subjects requires different asymptotic tools to study from those that are driven by a large number of independent subjects. Specifically, in the event that there are a large number of subjects available, the test statistics based on the difference of two quadratic forms can be written as a clean quadratic form of infinitely many independent terms whose distribution can be obtained through the CLT for clean quadratic forms (de Jong, 1987) when there are certain control over the correlations. For the case that there are only a small number of subjects, we will apply the CLT of Rosenblatt (1956) on long memory dependent process that allows for a more general decay rate.

Let $\sum_{j=1}^{b} \epsilon_j$ be the sum of the α-mixing process with decay rate $\alpha_m \to 0$. The Rosenblatt CLT was established by partitioning the process into alternating big and small blocks (of size p_b and q_b, respectively) such that the big blocks are nearly independent. Let K be the number of big blocks. Suppose $E(\sum_{j=b_1}^{b_2} \epsilon_j)^2 = O(h(b_2 - b_1))$ for some function $h(\cdot)$, as $b_2 - b_1 \to \infty$. The Rosenblatt (1956) CLT states that $\sum_{j=1}^{b} \epsilon_j/\sqrt{Kh(p_b)}$ is asymptotically normally distributed if $E|\sum_{j=b_1}^{b_2} \epsilon_j|^{2+\delta} = o([h(b_2 - b_1)]^{1+\delta/2})$ as $b_2 - b_1 \to \infty$ for some $\delta > 0$. The p_b, q_b and K are required to go to infinity and need to satisfy conditions:

$$K(p_b + q_b) = b; \qquad q_b/p_b \to 0 \quad \text{as } b \to \infty; \qquad h(q_b) = o(h(b)/K^3); \qquad K \le [-\log \alpha_{q_b}]^{1/2}. \tag{3.3}$$

In this subsection, we also let p_b, q_b and K satisfy conditions in (3.3). In addition, we assume that for $b_2 - b_1 \to \infty$ and all i, k,

$$E\left|\sum_{j=b_1}^{b_2}(X_{ijk} - E(X_{ijk}))\right|^2 = O(h(b_2 - b_1)), \qquad E\left|\sum_{j=b_1}^{b_2}(X_{ijk} - E(X_{ijk}))\right|^{2+\delta} = o(h(b_2 - b_1)^{1+\delta/2}), \tag{3.4}$$

$$E\left|\sum_{j=b_1}^{b_2}(X_{ijk} - E(X_{ijk}))(X_{i'jk'} - E(X_{i'jk'}))\right|^{2+\delta} = o(h(b_2 - b_1)^{1+\delta/2}), \quad \text{for } i \neq i' \text{ or } k \neq k'. \tag{3.5}$$

For the test of interaction effect $H_0(\gamma) : \gamma_{ij} = 0$, Wang and Akritas (2009) gave a test under weak dependence. Here, we give a test of interaction effect under long memory dependence. The test statistic for testing $H_0(\gamma)$ is $b(AM\gamma - AME_\gamma)$, where

$$AM\gamma = (a-1)^{-1}(b-1)^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{ij.} - \widetilde{X}_{i..} - \widetilde{X}_{.j.} + \widetilde{X}_{...})^2,$$

$$AME_\gamma = a^{-1}(b-1)^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b} n_i^{-1}(n_i-1)^{-1}\sum_{k=1}^{n_i}\left(X_{ijk} - \overline{X}_{ij.} - \overline{X}_{i.k} + \widetilde{X}_{i..}\right)^2.$$

Denote

$$L_{lb} = (l-1)(p_b+q_b)+1, \quad \text{and} \quad U_{lb} = (l-1)(p_b+q_b)+p_b. \tag{3.6}$$

The following theorem states the asymptotic result for testing no simple effect of treatment and no treatment by time interactions.

Theorem 3.3 (*Simple Effect of Treatment and Treatment by Time Interaction*). *Assume that for each group i and subject k, X_{ijk}, $j = 1, 2, \ldots$, is α-mixing with $\alpha_m \to 0$ as $m \to \infty$. Let $\tau_L = \zeta_1^L + \zeta_2^L/(a-1)^2$, where*

$$\zeta_1^L = \frac{2}{a^2}\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=L_{lb}}^{U_{lb}}\sum_{i=1}^{a}\frac{\sigma_{ijj'}^2}{n_i(n_i-1)}, \qquad \zeta_2^L = \frac{2}{a^2}\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=L_{lb}}^{U_{lb}}\sum_{i\neq i'}^{a}\frac{\sigma_{ijj'}\sigma_{i'jj'}}{n_i n_{i'}},$$

where U_{lb} and L_{lb} are defined in (3.6). Then under the assumptions in (3.3)–(3.5),

(1) $\mathrm{Var}(b(AM\phi - AME)) - \tau_L = o(Kh(p_b)n^{-2}(a))$ *as $b \to \infty$ for all $n_i \geq 2$ provided that $\max_i\{n_i\}/n(a) = O(1)$ if n_i go to infinity;*

(2) *for $n_i \geq 2$ that are bounded, the following statements hold:*
   (i) *Under $H_0(\phi)$, $b(AM\phi - AME)/\sqrt{\tau_L} \xrightarrow{d} N(0, 1)$ as $b \to \infty$;*
   (ii) *Under $H_0(\gamma)$, $b(AM\gamma - AME_\gamma)/\sqrt{\tau_L} \xrightarrow{d} N(0, 1)$ as $b \to \infty$;*

(3) *if n_i go to ∞ and $\max_i\{n_i\}/n(a) = O(1)$, then the results in item (2) still hold provided that $\sum_{j_1=1}^{b}\sum_{j_2=1}^{b}\sum_{j_3=1}^{b}\sum_{j_4=1}^{b}\sigma_{ij_1j_2}\sigma_{ij_1j_3}\sigma_{ij_2j_4}\sigma_{ij_3j_4} = o(K^2h(p_b)^2)$.*

(4) *Let $\widehat{\tau}_L = \widehat{\zeta}_1^L + \widehat{\zeta}_2^L/(a-1)^2$, where $\widehat{\zeta}_1^L$ and $\widehat{\zeta}_2^L$ are similarly defined as ζ_1^L and ζ_2^L but with $\sigma_{ijj'}$ and $\sigma_{ijj'}^2$ replaced by $\widehat{\sigma}_{ijj'}$ in (3.1) and $\widehat{\sigma_{ijj'}^2}$ in (3.2), respectively. Then $n^2(a)(\widehat{\tau}_L - \tau_L)/(Kh(p_b)) \xrightarrow{p} 0$ as $b \to \infty$ for $n_i \geq 2$ no matter n_i stay bounded or go to infinity provided that $\max_i\{n_i\}/n(a) = O(1)$.*

With the constraints in (3.3), the order of $Kh(p_b)$ is between b and b^2. Therefore, the standardizing rate for the test statistics is $b/\sqrt{Kh(p_b)} = O(b^\nu)$ for some $\nu < 1/2$. This tells us that allowing slower rate for the mixing coefficient leads to slower convergence rate for the test statistics. In addition, under long memory dependence, the calculation of the asymptotic variances need to be restricted to those among pairwise non-overlapping big blocks only.

Theorem 3.4 (*Main Effect of Treatment*). *Assume that for each group i and subject k, X_{ijk}, $j = 1, 2, \ldots$, is α-mixing with $\alpha_m \to 0$ for all i, k as $m \to \infty$. Let $\mathbf{C}_a$ be a contrast matrix with rank r and $\mathbf{Z} = (\sqrt{n}\, b\overline{X}_{1..}, \ldots, \sqrt{n}\, b\overline{X}_{a..})'$. Let U_{lb}, L_{lb} be as defined in (3.6) and*

$$\widehat{V}_i^L = \frac{n}{n_i(n_i-1)}\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=L_{lb}}^{U_{lb}}\sum_{k=1}^{n_i}\left(X_{ijk} - \overline{X}_{ij.}\right)\left(X_{ij'k} - \overline{X}_{ij'.}\right).$$

Then under $\widetilde{H}_0(\alpha)$ and assumptions in (3.3)–(3.5),

$$\mathbf{Z}'\mathbf{C}_a'\left[\mathbf{C}_a\mathrm{diag}(\widehat{V}_1^L, \ldots, \widehat{V}_a^L)\mathbf{C}_a'\right]^{-1}\mathbf{C}_a\mathbf{Z} \xrightarrow{d} \chi_r^2,$$

as $b \to \infty$ regardless of whether the n_i remain bounded or tend to ∞.

Remark 3. The results of Theorems 3.3 and 3.4 also hold under weak dependence. However, we expect that the test statistics in Theorems 3.1 and 3.2 have better power than those in Theorems 3.3 and 3.4 for the same b. This is because overlapping blocks are used in calculation of the test statistics in Theorems 3.1 and 3.2 whereas non-overlapping blocks are used for Theorems 3.3 and 3.4.

4. Numerical studies

4.1. Analysis of Bovine calf ear temperature profiles

Normal body temperatures for stocker cattle typically range between 100.5 and 102.5F. Elevated temperature can be a symptom of Bovine Respiratory Disease. However, during the summer months, body temperatures may climb above 103F during the heat of the day in the absence of disease for cattle on feed. In a study conducted at Kansas State University (Nickell et al., 2008), the efficacy of two broad spectrum antibiotics were compared on health and feed performance when administered metaphylactically on highly stressed beef stocker calves. Both antibiotics were administered subcutaneous according to label directions. The study would like to compare the effects of the two antibiotics in controlling body temperature.

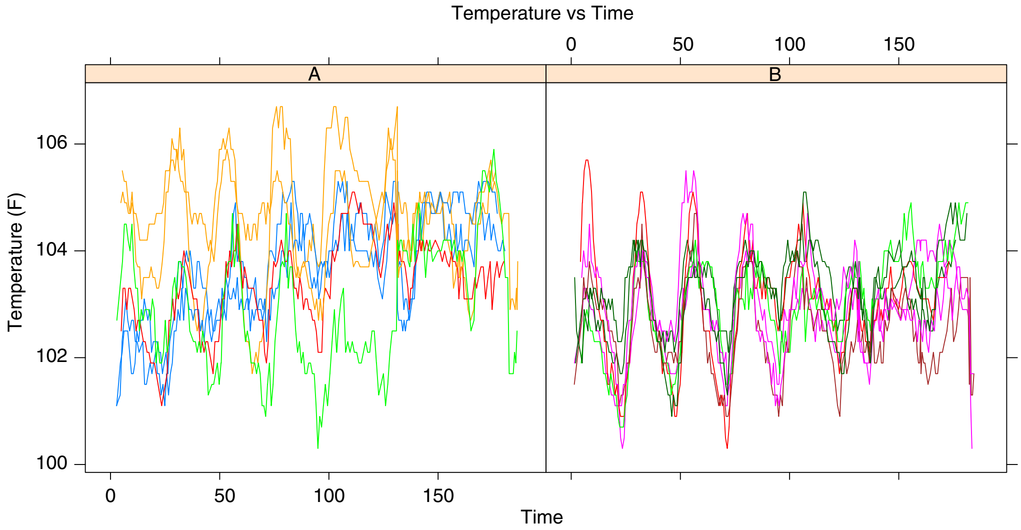

Fig. 1. Calf ear temperature profile plot. Each colored curve in the same panel represents the temperatures from the same calf over time. The left panel is for calves in treatment A and the right panel is for calves in treatment B. The temperature variations at each time point are larger in treatment A than those in B suggesting heteroscedastic variances. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

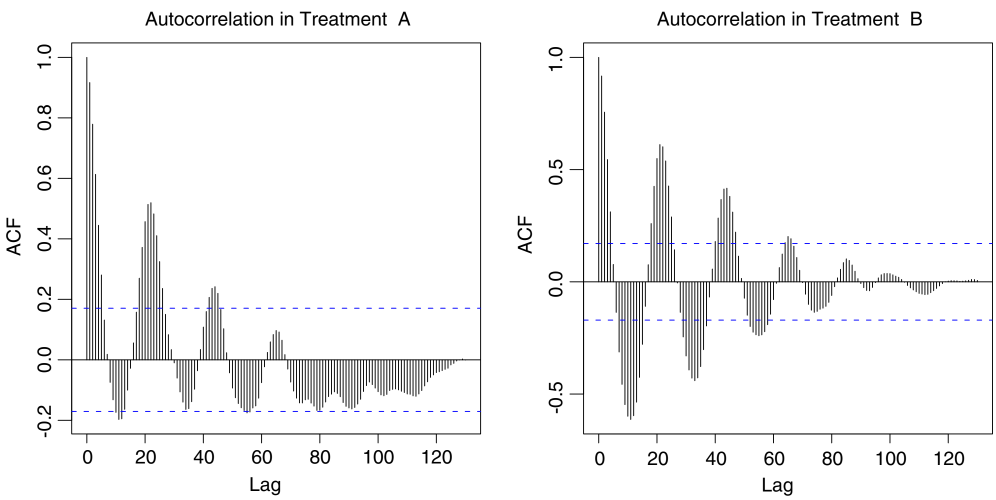

Fig. 2. Autocorrelation function versus time lag. The magnitude of correlation declines in both treatments as the time lag increases.

In this subsection, we analyze a data set collected in a recent clinical trial performed at Kansas State University (KSU) for six stockers calves randomly assigned to an antibiotic treatment A and seven stockers calves randomly assigned to another antibiotic treatment B. The body temperature of these calves was measured on an hourly basis utilizing a novel technology housed within an ear tag equipped with a thermistor that rests within the calf's internal ear canal. There were 132 repeated measurements per calf taken during the experiment. All 13 calves remained alive throughout the experiment.

From the temperature profile plot in Fig. 1, it can be seen that the temperatures of all calves under treatment B follow similar pattern with well controlled variations over time. On the other hand, the temperatures of the calves under treatment A exhibit very different variations over time. The autocorrelation function (Fig. 2) for average temperature over time in treatment A also exhibits a different pattern from that in treatment B. These give us evidence that there are heteroscedastic covariance structures that exist in the data. In addition, the magnitude of the autocorrelations in both treatments declines as the time lag increases suggesting that the α-mixing condition is appropriate for the data.

For the traditional methods, we consider linear mixed effect models with various covariance structures fitted with *lme* command in R package *nlme*, generalized least square method with several covariance structures fitted with *gls* command in package *nlme*, and generalized estimating equation approach fitted with R package *gee*. The *gls* command fits a linear model using generalized least squares. The errors are allowed to be correlated with certain correlation structures. No random effects are necessary in a model fitted with *gls*. The *lme* command fits a linear mixed-effects model using restricted maximum

Table 1
Obtained *p*-values from various tests.

	NP.S	NP.L	LME			GLS			GEE
			Ran	RanAR1	RanGaus	AR1	ARMA(2, 2)	ARMA(1, 1)	
Simple treatment	1.1×10^{-4}	8.4×10^{-7}							
Treatment:Time	0.858	0.822	6.3×10^{-7}	0.503	3.0×10^{-4}	0.020	7.1×10^{-3}	1.7×10^{-3}	0.129
Treatment	1.0×10^{-12}	2.7×10^{-4}	1.1×10^{-6}	0.154	5.8×10^{-5}	1.1×10^{-3}	1.7×10^{-4}	8.9×10^{-4}	0.333

NP.S — proposed nonparametric test assuming weak dependence; NP.L — proposed nonparametric test assuming long range dependence; Ran — linear mixed effect model with a random intercept; RanAR1 — linear mixed effect model with a random intercept plus an AR(1) serial correlation; RanGaus — linear mixed effect model with a random intercept plus a Gaussian serial correlation; AR1 — generalized least square fit with an AR(1) structure; ARMA(2, 2) — generalized least square fit with an ARMA(2, 2) structure; ARMA(1, 1) — generalized least square fit with an ARMA(1, 1) structure; GEE — generalized estimating equation approach with a working independence correlation.

likelihood or maximum likelihood in the formulation described in Laird and Ware (1982) but allowing for nested random effects. The within-group errors are allowed to be correlated. Random effects specified with grouping variables must be included in a model fitted with *lme*. These traditional methods require specification of the covariance structure. We did the following to select a best structure. We first fitted ARMA models with various orders for the pooled average temperature profile. The best ARMA model selected using AIC through *autoarmafit* function in R package *timesac* (Akaike et al., 1975) is ARMA(2, 2). Alternatively, we first fitted a periodic curve using spline for the pooled data. Then the residual curves were obtained to choose the best ARMA model with AIC. The selected model is ARMA(1, 1) for the residual curves. Since the order selection is among ARMA models only, so we also consider a few other covariance structures to compare.

We applied the proposed tests and common traditional methods to the data. For the *gls* fit with ARMA(1, 1), we use the fitted periodic spline curve instead of original time as the covariate. To apply Theorems 3.1 and 3.2, we take b^h to be the smallest integer that satisfies $h > 1/4$. This gives us $b^h = 4$. In calculation of the test statistics in Theorems 3.3 and 3.4, $q_b = b^{3/8}$, and $K = [-log(q_b^{-3/4}) + 0.5]$, where $[x]$ is the integer closest to x, and $p_b = [b/K - q_b]$. The proposed tests of no simple effect and no main effect of treatment yielded highly significant *p*-values. On the other hand, the linear mixed effect model (LME) with a random intercept plus an AR(1) correlation (RanAR1) and generalized estimating equations (GEE) with independent working correlation and Gaussian family produced insignificant results for both the treatment by time interaction effect and the main treatment effect. The generalized least squares method with an AR(1) serial correlation found significant treatment effect but medium evidence of treatment by time interaction. The LME with constant correlation structure (Ran), LME with a random intercept plus a Gaussian serial correlation (RanGaus), and the two *gls* tests with ARMA structures had highly significant *p*-values for both the interaction effect and main treatment effect. We know from the autocorrelation plot, however, the constant correlation structure is not a reasonable assumption for this data set. In addition, the simulation study in next section indicates that these tests have highly inflated type I error.

We also fitted a generalized additive mixed model (GAMM) Wood (2006, 2008) including the additive effect of treatment and a penalized spline of time. The interaction between the spline of time and treatment could not be included because the model fitting and parameter estimation does not allow us to include the interaction term. The *p*-value for the test of no treatment effect is 2.07×10^{-4} when an AR(1) correlation structure was used, and $p < 2 \times 10^{-16}$ when a compound symmetry correlation structure was used. These are only for reference since the interaction term cannot be included in the model.

4.2. Simulation study

In this subsection, we compare the proposed tests with commonly used traditional tests on type I error and power estimates via simulation studies. For the proposed test under long range dependence, we take the same b^h, p_b, q_b and K as in Section 4.1. For the *gls* test with ARMA(1, 1) structure, the periodic spline fit of the mean curve over time is used as the covariate to replace time in the model. For all other tests, time was used directly in the model. The *gls* test with ARMA(2, 2) covariance structure took about 2 hours to conduct a single test using Intel(R) Pentium(R) M Processor 1.86 GHz, 1G of RAM. This makes it impossible to conduct simulation study with it. The *gls* test using ARMA(1, 1) structure took about 17 minutes for a single test using the same computer. Therefore, all methods used for the calf data in Section 4.1 were considered for comparison except for the *gls* with ARMA(2, 2) structure. All results reported here are based on 3000 runs with the exception of the *gls* test with ARMA(1, 1) that is based on 1000 runs.

Under the null hypothesis that there is no effect of the treatment, average ear temperature ($\boldsymbol{\mu}$) from all calves in both treatments at each time point were used as the common mean profile to generate simulated data. The standard deviation of the calf ear temperature at each time point and the autocorrelation of the average temperature are used to specify the common covariance matrix Σ under H_0. The average temperature, standard deviation, and autocorrelation were plotted in Fig. 3.

We simulated 132 hourly temperature measurements for each calf from multivariate normal with mean $\boldsymbol{\mu}$ and covariance matrix Σ. Data for six calves in one treatment group and seven calves in the other treatment group were generated. In the *gls* test with ARMA(1, 1) structure, we used the periodic spline fit of the mean curve as the covariate in the model. For all other tests, time was used directly. The type I error estimates at level 0.05 and 0.10 for all the tests are given in Table 2.

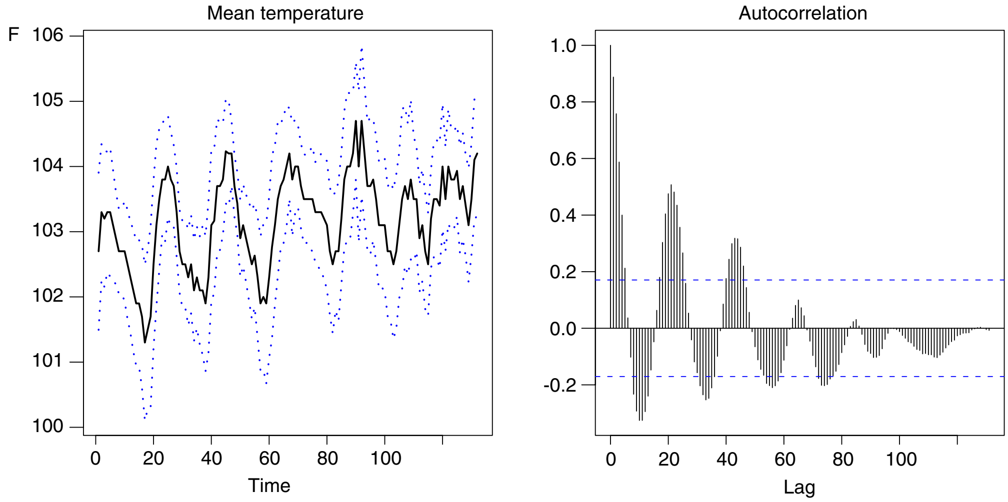

Fig. 3. Left panel: Average hourly calf ear temperature $\pm$ one standard deviation; right panel: autocorrelation of the calf ear temperature to be used to generate the correlation matrix for simulated data under H_0.

Table 2
Type I error estimates at levels 0.05 and 0.10. Table legend is the same as that for Table 1.

Level	Effect			LME			GLS		GEE
		NP.L	NP.S	Ran	RanAR1	RanGaus	AR1	ARMA(1,1)	
$\alpha = 0.05$	Simple treatment	0.069	0.051						
	Interaction	0.070	0.048	0.415	0.011	0.274	0.250	0.282	0.099
	Treatment	0.051	0.041	0.441	0.007	0.294	0.431	0.284	0.100
$\alpha = 0.10$	Simple treatment	0.111	0.075						
	Interaction	0.108	0.072	0.492	0.034	0.362	0.345	0.364	0.160
	Treatment	0.095	0.086	0.525	0.022	0.385	0.504	0.364	0.155

The proposed test assuming weak dependence has the best type I error estimate at level 0.05 while the proposed method with long range dependence has best type I error estimate at level 0.10. The traditional methods all have highly inflated type I error with the exception of LME with a random intercept plus an AR(1) serial correlation (which is conservative). Even though the ARMA(1, 1) structure has the closest covariance to the generated data, the *gls* test with ARMA(1, 1) structure does not perform well because of the small number of calves. Due to the elevated type I error of most traditional methods, we only consider the two proposed methods, LME with a random intercept plus an AR(1) serial correlation and GEE in power comparisons.

For power comparison, we generated data from multivariate normal as below with mean and covariance matrices produced from calf ear temperature data:

$$\mathbf{X}_{1k} \sim MVN(\boldsymbol{\mu} + \theta(\bar{\mathbf{x}}_{1\cdot} - \boldsymbol{\mu}), \Sigma_1), \quad \mathbf{X}_{2k} \sim MVN(\boldsymbol{\mu} + \theta(\boldsymbol{\mu} - \bar{\mathbf{x}}_{2\cdot}), \Sigma_2),$$

where $\bar{\mathbf{x}}_{1\cdot} = (\bar{x}_{11\cdot}, \ldots, \bar{x}_{1b\cdot})'$ and $\bar{\mathbf{x}}_{2\cdot} = (\bar{x}_{21\cdot}, \ldots, \bar{x}_{2b\cdot})'$ are the average temperature from all calves in treatments A and B, respectively. Σ_1 and Σ_2 are the covariance matrices generated using the sample covariance at eat time point and autocorrelation from treatment A and B of the calf data, respectively. The data are heteroscedastic since $\Sigma_1 \neq \Sigma_2$ that can be seen from the difference in variations over time and autocorrelation plots. Value $\theta = 0$ corresponds to the null hypotheses. The power estimates for $\theta = 0.25, 0.5, 0.75, 1$ are presented in Fig. 4.

The power plot shows that the proposed tests under long range dependence and weak dependence assumptions have similar performance. The current data generation setting is under the alternatives for the treatment effect, the proposed tests of no main treatment effect have better power than the tests of no simple treatment effect. Note that this will not always be the case. When strong interaction effects exist, we expect the tests of no simple treatment effect to be more powerful to detect the difference among the treatments than the tests of no main treatment effect. The proposed tests have much better power than the traditional methods in that the power of the traditional methods are still below 0.5 when that of the proposed tests approaches 1. The generalized least square method with AR(1) serial correlation has inflated type I error and also rejects more under the alternatives compared to GEE and linear mixed effect model with a random intercept plus AR(1) serial correlation.

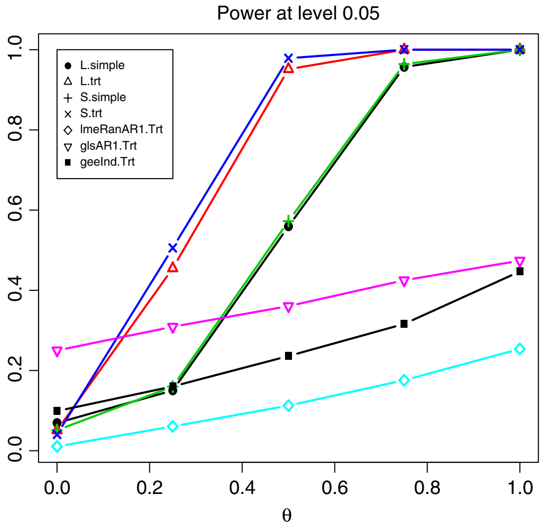

Fig. 4. Power of simple and main effects of treatment at level 0.05. L.simple — proposed test of no simple effect of treatment under long range dependence; L.trt — proposed test of no main treatment effect under long range dependence; S.simple — proposed test of no simple effect of treatment under weak dependence; S.trt — proposed test of no main treatment effect under weak dependence; lmeRanAR1.Trt — test of no main treatment effect based on linear mixed effect model with a random intercept plus an AR(1) serial correlation; glsAR1.Trt — test of no main treatment effect using generalized least squares method with an AR(1) serial correlation; geeInd.Trt — test of no main treatment effect based on generalized estimating equation approach with independent working correlation.

5. Technical proofs

Proof of Theorem 3.1. Write $AM\phi - AME = P_{u1} - P_{u2}/(a-1)$, where

$$P_{u1} = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)}, \qquad P_{u2} = \frac{1}{ab}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\bar{u}_{ij.}\bar{u}_{i'j.}. \tag{5.1}$$

To show the result under $H_0(\phi)$, write $n(a)\sqrt{b}(AM\phi - AME) = b^{-1/2}\sum_{j=1}^{b}W_j(\mathbf{u})$, where

$$W_j(\mathbf{u}) = \frac{n(a)}{a}\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)} - \frac{n(a)}{a(a-1)}\sum_{i\neq i'}^{a}\bar{u}_{ij.}\bar{u}_{i'j.}.$$

Note that $W_1(\mathbf{u}), W_2(\mathbf{u}), \ldots,$ are functions of independent α-mixing processes with $\alpha_m = O(m^{-5})$ and $E(W_j(\mathbf{u})) = 0$. We can apply the CLT in Lemma 2.2 of Wang and Akritas (2009) to prove the asymptotic normality if we can show that $\limsup_j E(W_j(\mathbf{u})^{16}) < \infty$. Apply the inequality: for any $P \geq 2$, there exists a finite positive constant A_P (depending only on P), such that for any iid random variables $Z_1, \ldots, Z_m$ with $E(Z_i) = 0$,

$$E|Z_1 + \cdots + Z_m|^P \leq A_P m^{P/2} E(|Z_1|^P). \tag{5.2}$$

This inequality follows if we first use the Marcinkiewicz–Zygmund inequality

$$E|Z_1 + \cdots + Z_m|^P \leq A_P E(Z_1^2 + \cdots + Z_n^2)^{P/2}, \quad P \geq 1,$$

and then apply Hölder's Inequality to the last summation. Note that u_{ijk}, $k = 1, \ldots, n_i$ are iid, so

$$E\left(\sum_{k=1}^{n_i}u_{ijk}\right)^{32} \leq A_{32}n_i^{16}E(u_{ij1}^{32}); \qquad E\left(\bar{u}_{ij.}\right)^{16} \leq \frac{A_{16}}{n_i^8}E(u_{ij1}^{16}).$$

Therefore,

$$E(W_j^{16}(\mathbf{u})) \leq \frac{2^{15}}{a}\left[2^{15}\sum_{i=1}^{a}A_{32}E(u_{ij1}^{32}) + \frac{\sum_{k=1}^{n_i}E\left(u_{ijk}^{32}\right)}{n_i} + a^{15}\sum_{i\neq i'}^{a}A_{16}^2E(u_{ij1}^{16})E(u_{i'j1}^{16})\right],$$

where A_{16} and A_{32} are finite positive constants. In addition,

$$\frac{1}{b}E\left(\sum_{j=1}^{b}W_j(\mathbf{u})\right)^2 = \frac{1}{b}\text{Var}\left(\sum_{j=1}^{b}W_j(\mathbf{u})\right) = \frac{n^2(a)}{a^2b}\sum_{i=1}^{a}\text{Var}\left(\sum_{k\neq k'}^{n_i}\sum_{j=1}^{b}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)}\right)$$
$$+\frac{n^2(a)}{a^2(a-1)^2b}\text{Var}\left(\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\bar{u}_{ij.}\bar{u}_{i'j.}\right)$$
$$=\frac{2n^2(a)}{a^2b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\left[\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\frac{E(u_{ijk}u_{ij'k})E(u_{ijk'}u_{ij'k'})}{n_i^2(n_i-1)^2}+\frac{1}{(a-1)^2}\sum_{i\neq i'}^{a}E\left(\bar{u}_{ij.}\bar{u}_{ij'.}\right)E\left(\bar{u}_{i'j.}\bar{u}_{i'j'.}\right)\right]$$
$$=\frac{2n^2(a)}{a^2b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\left[\sum_{i=1}^{a}\frac{\sigma_{ijj'}^2}{n_i(n_i-1)}+\frac{1}{(a-1)^2}\sum_{i\neq i'}^{a}\frac{\sigma_{ijj'}}{n_i}\frac{\sigma_{i'jj'}}{n_{i'}}\right]\to\widetilde{\tau}_{\gamma}^2. \tag{5.3}$$

Applying the CLT in Wang and Akritas (2009), we have completed the proof. □

Proof of Theorem 3.2. Note that $\overline{X}_{i..} - E(\overline{X}_{i..}) = \bar{u}_{i..}$. By inequality (5.2) and the moments assumption, we have $E\left(\sqrt{n}n_i^{-1}\sum_{k=1}^{n_i}u_{ijk}\right)^{16} < \infty$. Thus, by Lemma 2.2 in Wang and Akritas (2009) we have $\sqrt{N}[\overline{X}_{i..} - E(\overline{X}_{i..})] \xrightarrow{d} N(0,\eta_i)$, where η_i is defined by $\text{Var}(\sqrt{N}[\overline{X}_{i..}-E(\overline{X}_{i..})]) = n(bn_i)^{-1}\sum_{j=1}^{b}\sum_{j'=1}^{b}E(u_{ijk}u_{ij'k}) \to \eta_i$. By independence, $\sqrt{N}(\mathbf{W}-E(\mathbf{W})) \xrightarrow{d} N_a(\mathbf{0},\mathbf{V})$, where $\mathbf{V} = diag(\eta_1,\ldots,\eta_a)$. Therefore, by Continuous Mapping Theorem, $\sqrt{N}\mathbf{C}_a\mathbf{W} \xrightarrow{d} N_r(\mathbf{0},\mathbf{C}_a\mathbf{V}\mathbf{C}_a')$, as $b\to\infty$ regardless of whether $n_i\to\infty$ or not. By Lemma 4.5 in Wang (2004), $\widehat{\eta}_i$ are consistent estimators of η_i, $i=1,\ldots,a$. The proof is completed by an application of Slutsky's Theorem. □

Proof of Theorem 3.3. (1). We write

$$E[b(AM\phi - AME)]^2 = n(a)^{-2}E\left(\sum_{j=1}^{b}W_j(\mathbf{u})\right)^2 = \frac{2}{a^2}\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}E\left(\sum_{j=1}^{b}u_{ijk}u_{ijk'}\right)^2$$
$$+\frac{2}{a^2(a-1)^2}\sum_{i\neq i'}^{a}E\left(\sum_{j=1}^{b}\bar{u}_{ij.}\bar{u}_{i'j.}\right)^2. \tag{5.4}$$

The first term on the right hand side of (5.4) can be decomposed into three components (for $k\neq k'$)

$$VW_1 = \frac{2}{a^2}\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}E\left(\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}u_{ijk}u_{ijk'}\right)^2,\qquad VW_2 = \frac{2}{a^2}\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}E\left(\sum_{l=1}^{K}\sum_{j=U_{lb}+1}^{U_{lb}+q_b}u_{ijk}u_{ijk'}\right)^2,$$
$$VW_3 = \frac{4}{a^2}\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}E\left(\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}u_{ijk}u_{ijk'}\sum_{l'=1}^{K}\sum_{j'=U_{l'b}+1}^{U_{l'b}+q_b}u_{ij'k}u_{ij'k'}\right).$$

VW_1 and VW_2 mainly contain variances from the big blocks and small blocks, respectively; and VW_3 consists of the covariances between big and small blocks. By Rosenblatt (1956) CLT assumptions,

$$VW_1 = \frac{2}{a^2}\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\sum_{l=1}^{K}E\left(\sum_{j=L_{lb}}^{U_{lb}}u_{ijk}u_{ijk'}\right)^2 + o(Kh(p_b)) = \zeta_1^L + o(Kh(p_b)n^{-2}(a)),$$
$$VW_2 = O(Kh(q_b)n^{-2}(a)) = o(Kh(p_b)n^{-2}(a)) \quad \text{(cf. Eq. (7) of Rosenblatt (1956))}.$$

Applying the Cauchy–Schwarz inequality, we have

$$VW_3 \le \frac{4}{a^2}\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\left[E\left(\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}u_{ijk}u_{ijk'}\right)^2 E\left(\sum_{l'=1}^{K}\sum_{j'=U_{l'b}+1}^{U_{l'b}+q_b}u_{ij'k}u_{ij'k'}\right)^2\right]^{1/2}$$
$$= O\left(\frac{Kh(p_b)^{1/2}h(q_b)^{1/2}}{n^2(a)}\right) = o\left(\frac{Kh(p_b)}{n^2(a)}\right).$$

Combining the three terms together, we know that the first term of (5.4) is equal to $\zeta_1^L + o(Kh(p_b)n^{-2}(a))$. Similarly, the second term of (5.4) is equal to $\zeta_2^L/(a-1)^2 + o(Kh(p_b)n^{-2}(a))$.

(2). For n_i bounded, we write $n(a)(AM\phi - AME) = b^{-1}\sum_{j=1}^{b} W_j(\mathbf{u})$ as in the proof of Theorem 3.1. It can be shown that $W_1(\mathbf{u}), W_2(\mathbf{u}), \ldots,$ is an α-mixing process with decay rate α_m. We will apply the CLT for α-mixing processes in Rosenblatt (1956) to obtain the asymptotic distribution. Note that $E(W_j(\mathbf{u})) = 0$. To show Theorem 3.3, We only need to show that $E\left(\sum_{j=b_1}^{b_2} W_j(\mathbf{u})\right)^2 = O(h(b_2 - b_1))$ and $E\left(\sum_{j=b_1}^{b_2} W_j(\mathbf{u})\right)^{2+\delta} = o(h(b_2 - b_1)^{1+\delta/2})$ for some $\delta > 0$. The first one is obvious since we know from (5.3) and assumption (3.4) that

$$E\left(\sum_{j=b_1}^{b_2} W_j(\mathbf{u})\right)^2 = \frac{2n^2(a)}{a^2}\sum_{j=b_1}^{b_2}\sum_{j'=b_1}^{b_2}\left[\sum_{i=1}^{a}\frac{\sigma_{ijj'}^2}{n_i(n_i-1)} + \frac{1}{(a-1)^2}\sum_{i\neq i'}^{a}\frac{\sigma_{ijj'}}{n_i}\frac{\sigma_{i'jj'}}{n_{i'}}\right] = O(h(b_2 - b_1)).$$

Additionally, applying Holder's Inequality and assumption in (3.5), we have

$$\begin{aligned}
E\left(\sum_{j=b_1}^{b_2} W_j(\mathbf{u})\right)^{2+\delta} &= \frac{n^{2+\delta}(a)}{a^{2+\delta}}E\left\{\sum_{j=b_1}^{b_2}\left[\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)} + \frac{1}{(a-1)}\sum_{i\neq i'}^{a}\bar{u}_{ij.}\bar{u}_{i'j.}\right]\right\}^{2+\delta} \\
&\le \frac{n^{2+\delta}(a)}{a^{2+\delta}}2^{1+\delta}\left\{E\left(\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\sum_{j=b_1}^{b_2}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)}\right)^{2+\delta} + (a-1)^{-2-\delta}E\left(\sum_{i\neq i'}^{a}\sum_{j=b_1}^{b_2}\bar{u}_{ij.}\bar{u}_{i'j.}\right)^{2+\delta}\right\} \\
&\le \frac{n^{2+\delta}(a)}{a}2^{1+\delta}\left\{\sum_{i=1}^{a}E\left(\sum_{j=b_1}^{b_2}u_{ij1}u_{ij2}\right)^{2+\delta} + \frac{1}{(a-1)}\sum_{i\neq i'}^{a}E\left(\sum_{j=b_1}^{b_2}u_{ij1}u_{i'j1}\right)^{2+\delta}\right\} \\
&= o(h(b_2-b_1)^{1+\delta/2}).
\end{aligned}$$

(3). If n_i also go to infinity when $b \to \infty$, we let $l_{1k} = k$, $l_{ik} = \sum_{i_1=1}^{i-1} n_{i_1} + k$ for $i = 2, \ldots, a$, and write $bn(a)(AM\phi - AME) = \sum_{i=1}^{a}\sum_{i'=1}^{a}\sum_{k=1}^{n_i}\sum_{k=1}^{n_{i'}} H(l_{ik}, l_{i'k'})$, where

$$H(l_{ik}, l_{i'k'}) = \begin{cases} n(a)a^{-1}n_i^{-1}(n_{i'}-1)^{-1}\sum_{j=1}^{b} u_{ijk}u_{i'jk'} & \text{if } i = i' \text{ and } k \neq k'; \\ n(a)a^{-1}(a-1)^{-1}n_i^{-1}n_{i'}^{-1}\sum_{j=1}^{b} u_{ijk}u_{i'jk'} & \text{if } i \neq i'; \\ 0 & \text{otherwise.}\end{cases}$$

Treating l_{ik} as a single index instead of double indices by both i and k, $bn(a)(AM\phi - AME)$ is a clean quadratic form defined by de Jong (1987) since $H(l, l) = 0$ and $H(l_{ik}, l_{i'k'})$ is a quadratic form of independent vectors $\mathbf{u}_{ik} = (u_{i1k}, \ldots, u_{ibk})'$ and $\mathbf{u}_{i'k'}$. Hence, the asymptotic normality of $bn(a)(AM\phi - AME)$ can be shown by applying the CLT for clean quadratic forms in de Jong (1987) that is equivalent to showing that the following terms are of smaller order than $(\mathrm{Var}(bn(a)(AM\phi - AME)))^2 = O(K^2h(p_b)^2)$ as n_i and $b \to \infty$:

$$\begin{aligned}
G_I &= \sum_{1\le l<l_1\le n} E(H(l, l_1)^4), \qquad G_{II} = \sum_{1\le l<l_1<l_2\le n}\{E[H(l, l_1)^2H(l, l_2)^2] + E[H(l_1, l)^2H(l_1, l_2)^2] + E[H(l_2, l)^2H(l_2, l_1)^2]\}, \\
G_{IV} &= \sum_{1\le l<l_1<l_2<l_3\le n}\{E[H(l, l_1)H(l, l_2)H(l_3, l_1)H(l_3, l_2)] \\
&\quad + E[H(l, l_1)H(l, l_3)H(l_2, l_1)H(l_2, l_3)] + E[H(l, l_2)H(l, l_3)H(l_1, l_2)H(l_1, l_3)]\}.
\end{aligned}$$

Note that for $k \neq k'$,

$E(H(l_{ik}, l_{ik'})^4) = n^4(a)a^{-4}n_i^{-4}(n_{i'}-1)^{-4}E\left(\sum_{j=1}^{b} u_{ijk}u_{i'jk'}\right)^4 = O(K^2h(p_b)^2n^{-4}(a)),$

$E[H(l, l_1)^2H(l, l_2)^2] \le E(H(l, l_1)^4 + H(l, l_2)^4)/2 = O(K^2h(p_b)^2n^{-4}(a)),$

$$\begin{aligned}
E[H(l, l_1)H(l, l_2)H(l_3, l_1)H(l_3, l_2)] &= O\left\{n^{-4}(a)E\left(\sum_{j_1=1}^{b}u_{ij_1k}u_{i_1j_1k_1}\sum_{j_2=1}^{b}u_{ij_2k}u_{i_2j_2k_2}\sum_{j_3=1}^{b}u_{i_3j_3k_3}u_{i_1j_3k_1}\sum_{j_4=1}^{b}u_{i_3j_4k_3}u_{i_2j_4k_2}\right)\right\} \\
&= O\left\{n^{-4}(a)\sum_{j_1=1}^{b}\sum_{j_2=1}^{b}\sum_{j_3=1}^{b}\sum_{j_4=1}^{b}\sigma_{ij_1j_2}\sigma_{ij_1j_3}\sigma_{ij_2j_4}\sigma_{ij_3j_4}\right) = o(K^2h(p_b)^2n^{-4}(a)).
\end{aligned}$$

Therefore, G_I, G_{II} and G_{IV} are all of smaller order than $K^2h(p_b)^2$ as $b \to \infty$ and n_i go to infinity.

(4). Note that

$$\widehat{\zeta}_1^L - \zeta_1^L = \frac{2}{a^2}\sum_{i=1}^{a}\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=L_{lb}}^{U_{lb}}\frac{\widehat{\sigma^2_{ijj'}} - \sigma^2_{ijj'}}{n_i(n_i-1)}.$$

Hence,

$$\begin{aligned} E(\widehat{\zeta}_1^L - \zeta_1^L)^2 &= \frac{4}{a^4}\sum_{i=1}^{a}\sum_{l=1}^{K} E\left[\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=L_{lb}}^{U_{lb}}\frac{\widehat{\sigma^2_{ijj'}} - \sigma^2_{ijj'}}{n_i(n_i-1)}\right]^2 + o(Kh(p_b)^2 n^{-4}(a)) \\ &\le \frac{4}{a^4}\sum_{i=1}^{a}\sum_{l=1}^{K} n_i^{-2}(n_i-1)^{-2}\left[\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=L_{lb}}^{U_{lb}} E^{1/2}(\widehat{\sigma^2_{ijj'}} - \sigma^2_{ijj'})^2\right]^2 + o(Kh(p_b)^2 n^{-4}(a)) \\ &= O(Kh(p_b)^2 n^{-4}(a)) = o(K^2 h(p_b)^2 n^{-4}(a)), \end{aligned}$$

where the inequality is due to Cauchy–Schwarz Inequality. Similarly, $E(\widehat{\zeta}_1^L - \zeta_1^L)^2$ can be shown to be of the same order. The proof is then completed. □

Proof of Theorem 3.4. Denote $G_{ij} = \sqrt{n}\bar{u}_{ij\cdot}$, $G_i = \sqrt{n}\, b(\overline{X}_{i\cdot\cdot} - E(\overline{X}_{i\cdot\cdot})) = \sqrt{n}\sum_{j=1}^{b}\bar{u}_{ij\cdot} = \sum_{j=1}^{b} G_{ij}$, and $\mathbf{G} = (G_1, \ldots, G_a)'$. Under $\widetilde{H}_0(\alpha)$, $E(\overline{X}_{i\cdot\cdot}) = \mu$ and $\mathbf{G} = \mathbf{Z} - \mu\mathbf{1}_a$, where $\mathbf{1}_a$ is an a-dimensional vector of ones. Hence $\mathbf{C}_a\mathbf{Z} = \mathbf{C}_a\mathbf{G}$ since $\mathbf{C}_a$ is a contrast matrix. Define $V_i = n/n_i\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=L_{lb}}^{U_{lb}}\sigma_{ijj'}$. The components of $\mathbf{G}$ are independent implying that the proof can be achieved by showing $\frac{G_i/\sqrt{Kh(p_b)}}{\sqrt{V_i/(Kh(p_b))}} \to N(0,1)$, $(\widehat{V}_i - V_i)/(Kh(p_b)) \xrightarrow{p} 0$ followed by an application of the Continuous Mapping Theorem.

To show the normality, we will check the condition for Rosenblatt (1956) CLT. Note that $E(G_i) = 0$ and $\mathrm{Var}(G_i) - V_i \to 0$ as $b \to \infty$. In addition,

$$E\left(\sum_{j=b_1}^{b_2} G_{ij}\right)^2 = E\left(\frac{\sqrt{n}}{n_i}\sum_{j=b_1}^{b_2}\sum_{k=1}^{n_i} u_{ijk}\right)^2 = \frac{n}{n_i^2}\sum_{k=1}^{n_i} E\left(\sum_{j=b_1}^{b_2} u_{ijk}\right)^2 = O(h(b_2 - b_1)),$$

$$E\left(\sum_{j=b_1}^{b_2} G_{ij}\right)^{2+\delta} = \frac{n^{1+\delta/2}}{n_i^{2+\delta}} E\left(\sum_{j=b_1}^{b_2}\sum_{k=1}^{n_i} u_{ijk}\right)^{2+\delta} = \frac{n^{1+\delta/2}}{n_i^{2+\delta}}\sum_{k=1}^{n_i} E\left(\sum_{j=b_1}^{b_2} u_{ijk}\right)^{2+\delta} = o(h(b_2 - b_1)^{1+\delta/2}).$$

Applying the CLT by Rosenblatt (1956), we know that $G_i/(Kh(p_b))^{1/2}$ is asymptotically normal.

To show $(\widehat{V}_i - V_i)/(Kh(p_b)) \xrightarrow{p} 0$, we only need to show $E(\widehat{V}_i - V_i)^2/(Kh(p_b))^2 \to 0$ since $\widehat{V}_i$ is an unbiased estimator of V_i. Note that

$$\widehat{V}_i - V_i = \sqrt{n}\sum_{l=1}^{K}\sum_{j=(l-1)(p_b+q_b)+1}^{(l-1)(p_b+q_b)+p_b}\sum_{j'=(l-1)(p_b+q_b)+1}^{(l-1)(p_b+q_b)+p_b}(\widehat{\sigma}_{ijj'} - \sigma_{ijj'}).$$

The correlation between two zero mean random variables with finite fourth moments that belong to the σ-field generated by two different big blocks is at most $\alpha(q_b)$ since the distance between any two big blocks is at least q_b. Hence, we have

$$\begin{aligned} E[(\widehat{V}_i - V_i)^2] &= n\sum_{l=1}^{K}\sum_{j=L_{lb}}^{U_{lb}}\sum_{j'=C_l}^{U_l}\sum_{l'=1}^{K}\sum_{j_1=C_{l'}}^{U_{l'}}\sum_{j_1'=C_{l'}}^{U_{l'}}\sum_{k=1}^{n_i} E\left\{\left[\frac{(u_{ijk} - \bar{u}_{ij.})(u_{ij'k} - \bar{u}_{ij'.})}{(n_i-1)} - \frac{\sigma_{ijj'}}{n_i}\right]\right. \\ &\quad \left.\times\left[\frac{(u_{ij_1k} - \bar{u}_{ij_1.})(u_{ij_1'k} - \bar{u}_{ij_1'.})}{(n_i-1)} - \frac{\sigma_{ij_1j_1'}}{n_i}\right]\right\} \\ &\le \frac{n}{(n_i-1)^2}\left(1 + \frac{c_1}{n_i} + \frac{c_2}{n_i^2} + \frac{c_3}{n_i^3}\right)\sum_{l=1}^{K} E\left(\sum_{j=L_{lb}}^{U_{lb}} u_{ijk}\right)^4 \qquad (5.5) \\ &\quad + \frac{n}{n_i-1}\left(\frac{d_1}{n_i^2} + \frac{d_2}{n_i^3}\right)\sum_{l=1}^{K}\left[E\left(\sum_{j=L_{lb}}^{U_{lb}} u_{ijk}\right)^2\right]^2 + \frac{n}{(n_i-1)^2}\sum_{l\ne l'}^{K}\alpha(q_b), \qquad (5.6) \end{aligned}$$

for some finite constants c_1, c_2, c_3, d_1 and d_2 that do not depend on n_i or b. The term in (5.5) and the first term in (5.6) are of order no more than $Kh(p_b)^2$. The second term in (5.6) is of order no more than $K^2\alpha(q_b) = o(Kh(p_b))$ (cf. equation (7) of Rosenblatt (1956)) since $K \to \infty$ as $b \to \infty$. Therefore, $E[(\widehat{V}_i - V_i)^2]/(Kh(p_b))^2 \to 0$. The proof is completed by applying Slutsky's Theorem and the Continuous Mapping Theorem. □

Acknowledgements

The authors would like to thank Prof. Michael Akritas at Pennsylvania State University for his insightful discussion on the results under weak dependence. In addition, the authors would like to thank Prof. Suojin Wang at Texas A&M University for his careful proof reading of the manuscript.

References

Akaike, H., Arahata, E., Ozaki, T., 1975. Computer Science Monograph, No.5, Timsac74, A Time Series Analysis and Control Program Package (1). The Institute of Statistical Mathematics.
Aulakh, M., Pasricha, N., Baddesa, H., Bahl, G., 1991. Long-term effects of rate and frequency of applied p on crop yields, plant available p, and recovery of fertilizer p in a peanutwheat rotation. Soil Sci. 151, 317–322.
Bailey, T., Swan, J., Higgs, R., Paulson, W., 1996. Long-term tillage effects on continuous corn yields. In: Proc. of the 8th Annual Kansas State Univ. Conf. on Applied Statistics in Agriculture. Kansas State Univ., Manhattan, pp. 18–32.
Barber, S., 1979. Soil phosphorus after 25 tears of cropping with five rates of phosphorus application. Commun. Soil Sci. Plant Anal. 10, 1459–1468.
Billingsley, P., 1995. Probability and Measure, 3rd ed. Wiley, New-York.
Breslow, N., Clayton, D., 1993. Approximate inference in generalized linear mixed models. J. Amer. Statist. Assoc. 88, 9–25.
Cochran, W., 1939. Long-term agricultural experiments. J. R. Stat. Soc. [Ser. A] 6 (Suppl.), 104–148.
de Jong, P., 1987. A central limit theorem for generalized quadratic forms. Probability Theory 75, 261–277.
Doukhan, P., Oppenheim, G., Taqqu, M.S., 2003. Theory and Applications of Long-Range Dependence. Birkhauser Boston, c/o Springer-Verlag, New York.
Fan, J., Fan, Y., Lv, J., 2008. High dimensional covariance matrix estimation using a factor model. J. Econometrics 147, 186–197.
Guerin, L., Stroup, W., 2000. A simulation study to evaluate proc mixed analysis of repeated measures data. In: Proc. 12th Kansas State Univ. Conf. on Applied Statistics in Agriculture. Kansas State Univ., Manhattan, pp. 170–203.
Laird, N., Ware, J., 1982. Random-effects models for longitudinal data. Biometrics 38, 963–974.
Liang, K.-Y., Zeger, S., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73, 13–22.
Lin, X., Zhang, D., 1999. Inference in generalized additive mixed models by using smoothing splines. J. Roy. Statist. Soc. Ser. B 61, 381–400.
Littell, R., 2002. Analysis of unbalanced mixed model data: A case study comparison of anova versus reml/gs. J. Agric. Biol. Environ. Stat. 7, 472–490.
Loughin, T., 2006. Improved experimental design and analysis for long-term experiments. Crop Sci. 46, 2492–2502.
McCollum, R., 1991. Buildup and decline in soil phosphorus: 30-year trends on a typic umbraquult. Agron. J. 83, 77–85.
Morris, H., 1983. Parametric empirical bayes inference: Theory and applications. J. Amer. Statist. Assoc. 78, 47–65.
Nickell, J., White, B., Larson, R., Blasi, D., Renter, D., 2008. Comparison of short-term health and performance effects related to prophylactic administration of tulathromycin versus tilmicosin in long-hauled, highly stressed beef stocker calves. Vet. Therapeutics 9, 147–156.
Raudenbush, S., Bryk, A., 2002. Hierarchial Linear Models, 2nd ed. Thousand Oaks, CA, Sage.
Rosenblatt, M., 1956. A central limit theorem and a strong mixing condition. Proc. Natl. Acad. Sci. USA 42, 43–47.
Schäfer, J., Strimmer, K., 2005. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. Statistical Applications in Genetics and Molecular Biology 4, Article 32.
Wang, H., 2004. Testing in Multifactor Heteroscedastic ANOVA and Repeated Measures Designs with Large Number of Levels. Ph.D. thesis, Pennsulvania State University.
Wang, H., Akritas, M., 2009. Inference from heteroscedastic functional data. J. Nonparametr. Stat., in press (doi:10.1080/10485250903171621).
Wood, S., 2006. Low rank scale invariant tensor product smooths for generalized additive mixed models. Biometrics 62, 1025–1036.
Wood, S., 2008. Fast stable direct fitting and smoothness selection for generalized additive models. J. Roy. Statist. Soc. Ser. B 70, 495–518.
Zhu, D., Hero, A., 2007. Bayesian hierarchical model for large-scale covariance matrix estimation. J. Comput. Biol. 14, 1311–1326.