Communications for Statistical Applications and Methods
2016, Vol. 23, No. 1, 33–45

http://dx.doi.org/10.5351/CSAM.2016.23.1.033
Print ISSN 2287-7843 / Online ISSN 2383-4757

Variance components estimation in the presence of drift

Jaehee Kim[1,a], Todd Ogden[b]

[a]Department of Statistics, Duksung Women's University, Korea
[b]Department of Biostatistics, Columbia University, USA

Abstract

Variance components should be estimated based on mean change when the mean of the observations drift gradually over time. Consistent estimators for the variance components are studied for a particular modeling situation with some underlying functions or drift. We propose a new variance estimator with Fourier estimation of variations. The consistency of the proposed estimator is proved asymptotically. The proposed procedures are studied and compared empirically with the variance estimators removing trends. The result shows that our variance estimator has a smaller mean square error and depends on drift patterns. We estimate and apply the variance to Nile River flow data and resting state fMRI data.

Keywords: fMRI data, fourier series, pseudoresiduals, m-dependent data, slowly varying functions

1. Introduction

Variance parameters should be estimated accounting for the mean drift after the assumption of constant means is violated in which the mean drifts. This paper presents methodology for variance estimation to handle possible drift in the data. The motivating example involves a common psychological experiment on human rhythmic and motor control. Ogden and Collier (2002) proposed a variance estimator for tapping data with drift. The methods developed in this paper can be adapted to other situations in which observations from a specified stationary time series model are suspected to demonstrate drift in the mean, especially with fMRI data. In addition, the methods developed in this paper can be applied to astronomy data in which the time between the maximum (or minimum) brightness of stars is measured (see, Eddington and Plakidis (1929)). Astronomers model the data with the exact model (as shown in Section 2); in addition, they are interested in testing for drift in intervals between maximum brightness.

The problem in the classical approach to the estimation of variance parameters is the tendency for subjects to drift from the starting tempo. For example, the simple sample variance estimator is unbiased when there is no drift, but can be badly biased when there is drift present and the drift is not considered. Drift is present in many of these experiments; therefore, some method is necessary to account for drift when making an inference on these parameters.

This paper presents a new "drift-free" method for variance component estimation. Methodology is established in this paper with the tapping experiment in mind; however, the techniques developed can be adapted to any situation such as an explicit model for time series-type data.

Our approach can be applied to derive a variance estimator effective for a wide class of functions with some type of drift.

[1] Corresponding author: Department of Statistics, Duksung Women's University, Seoul 10369, South Korea.
E-mail: jaehee@duksung.ac.kr

Published 31 January 2016 / journal homepage: http://csam.or.kr

2. Residual variance estimation with the drift effect

Consider a nonparametric regression model

$$y_i = f(x_i) + \epsilon_i, \tag{2.1}$$

where y_i's are observations, f is an unknown mean function and ϵ_i's are independent and identically distributed random errors with mean zero and variance σ^2. The estimation of σ^2 is an important problem since it is essential to make inferences about the underlying function.

In this section, we review some variance estimators with nonparametric regression models.

Reinsh (1967) proposed choosing the curve that minimized $\sum(y_i - f(x_i))^2$ subject to

$$\int_0^1 \{f''(x)\}^2 dx \le C.$$

This justification is to try to obtain the best possible fit to data subject to the curve with minimal local variation as measured by its integrated squared second derivative.

One possible approach is to use the idea of differencing to remove trend. Rice (1984) proposed the first-order difference-based estimator by

$$\hat{\sigma}_R^2 = \frac{1}{2(n-1)} \sum_{i=2}^{n} (y_i - y_{i-1})^2. \tag{2.2}$$

Gasser *et al.* (1986) used a similar idea to remove the local trend and proposed a second-order difference-based estimator

$$\hat{\sigma}_{GSJ}^2 = \frac{1}{n-2} \sum_{i=2}^{n-1} c_i^2 \hat{\epsilon}_i^2, \tag{2.3}$$

where $\hat{\epsilon}_i$ is the difference between y_i and the value at x_i of the line joining (x_{i-1}, y_{i-1}) and (x_{i+1}, y_{i+1}) such as

$$\hat{\epsilon}_i = \frac{x_{i+1} - x_i}{x_{i+1} - x_{i-1}} y_{i-1} + \frac{x_i - x_{i-1}}{x_{i+1} - x_{i-1}} y_{i+1} - y_i.$$

The coefficients c_i are chosen such that $E(c_i^2 \hat{\epsilon}_i^2) = \sigma^2$ for all i when f is linear. For equidistance design points, $\hat{\sigma}_{GSJ}^2$ is reduced to

$$\hat{\sigma}_{GSJ}^2 = \frac{2}{3(n-2)} \sum_{i=2}^{n-1} \left(\frac{1}{2} y_{i-1} - y_i + \frac{1}{2} y_{i+1} \right)^2. \tag{2.4}$$

Hall *et al.* (1990) introduced the estimator

$$\hat{\sigma}_{HKT}^2(m) = \frac{1}{m-n} \sum_{i=m_1+1}^{n-m_2} \left(\sum_{k=-m_1}^{m_2} d_i y_{k+i} \right)^2, \tag{2.5}$$

where m_1 and m_2 are nonnegative integers, $m = m_1 + m_2$ is referred to as the order, and the difference sequence $\{d_i\}_{i=-m_1,\ldots,m_2}$ satisfies $\sum_{-m_1}^{m_2} d_j = 0$, $\sum_{-m_1}^{m_2} d_j^2 = 1$ and $d_{-m_1} d_{m_1} \neq 0$.

They provide optimal difference sequences for $1 \leq m \leq 10$. Four decimals entries for each m are:

- $m = 1$ $(-0.7071, -0.7071)$
- $m = 2$ $(0.8090, -0.5, -0.3090)$
- $m = 3$ $(0.1942, 0.2809, 0.3832, -0.8582)$
- $m = 4$ $(0.2708, -0.0142, 0.6909, -0.4858, -0.4617)$.

None of the above difference-based estimators achieves the asymptotic optimal rate for the mean squared error (Dette *et al.*, 1998) such as

$$\mathrm{MSE}\left(\hat{\sigma}^2\right) = \frac{1}{n}\mathrm{Var}\left(\epsilon^2\right) + O\left(n^{-1}\right).$$

Buckley *et al.* (1988) considered the residual variance estimator as the minimax mean squared error estimator and provided the optimal estimator which has the form

$$\hat{\sigma}^2_{BES} = \frac{\mathbf{y}^t\mathbf{D}\mathbf{y}}{\mathrm{tr}(\mathbf{D})}, \tag{2.6}$$

where $\mathbf{D}$ is a symmetric $n \times n$ matrix nonnegative-definite matrix satisfying the mean squared error and $\mathbf{y}$ is a vector of y_i's. These estimators usually the residual sum of squares from some nonparametric fit to f (Wahba, 1990). Or with the linear smoother A, $\hat{\mathbf{y}} = \mathbf{A}\mathbf{y}$ and $\mathbf{D} = (\mathbf{I} - \mathbf{A})^t(\mathbf{I} - \mathbf{A})$ (Hastie and Tibshirani, 1990).

Chaudhuri (1992) compared Sarndal *et al.* (1989) and Kott (1990) variance estimators of a finite population mean based on a simple random sample without replacement using regression estimator.

Müller *et al.* (2003) proposed the class of difference-based estimators

$$\hat{\sigma}^2_{MSW} = \frac{1}{2\sum_{i \neq j} W_{ij}} \sum_{i \neq j} W_{ij}(y_i - y_j)^2, \tag{2.7}$$

where the weights W_{ij} depend on the covariates only, but not on errors such as

$$\begin{aligned} &W_{ij} \geq 0, \quad i, j = 1, \ldots, n, \; i \neq j, \\ &W_{ij} = W_{ji}, \quad i, j = 1, \ldots, n, \; i \neq j, \\ &\frac{1}{n(n-1)} \sum \sum_{i \neq j} W_{ij} = 1. \end{aligned}$$

$\hat{\sigma}^2_{MSW}$ achieves the asymptotic optimal rate under certain assumptions for weights and constructed weights based on a kernel density estimate.

Tong and Wang (2005) proposed the variance estimator as the intercept in a simple linear regression model with squared differences of paired observations as the dependent variable and squared distances between paired covariates as the regressor, which achieves the optimal rate. Considering the simple linear model

$$s_k = \alpha + \beta d_k + e_k,$$

where $d_k = k^2/n^2$ and for $1 \leq k \leq m < n$ $s_k = \sum_{i=k+1}^{n}(y_i - y_{i-k})^2/\{2(n-k)\}$ is the average of $(n-k)$ lag-k differences. Assign weight $w_k = (n-k)/N$ to the observation s_k, where $N = (n-1) + (n-2) + \cdots + (n-m) = nm - m(m+1)/2$. Tong and Wang (2005) estimator is

$$\hat{\sigma}^2_{TW} = \hat{\alpha} = \bar{s}_w - \hat{\beta}\bar{d}_w \tag{2.8}$$

which is unbiased when f is linear, where

$$\hat{\beta} = \frac{\sum_{k=1}^{m} w_k s_k \left(d_k - \overline{d}_w\right)}{\sum_{k=1}^{m} w_k \left(d_k - \overline{d}_w\right)^2}.$$

Recently Brown and Levine (2007) proposed a difference-based kernel estimators for the variance function for both unknown mean function and unknown variance function.

3. Proposed variance estimator

In this section, we propose a variance estimator useful for underlying functions with some drift including abrupt, smooth changes. Consider the variance function estimation with Fourier series with difference estimates in model (2.1) where $\mathrm{var}(\epsilon_i) = \sigma_i^2$. The difference estimated are defined as

$$r_i = \frac{1}{2}(y_{i+1} - y_i)^2, \quad i = 1, 2, \ldots, n-1. \tag{3.1}$$

Fourier representation can be applied to s_i's as

$$E[r_i] = s_i = \phi_0 + \sum_{j=1}^{\infty} \phi_j b_j(i), \quad i = 1, 2, \ldots, n-1,$$

where $b_j(i)$ are the orthogonal bases such as trigonometric basis functions. Therefore s_i can be approximated with the sample Fourier coefficients as

$$\hat{s}_i = \hat{\phi}_0 + \sum_{j=1}^{K} \hat{\phi}_j b_j(i), \quad i = 1, 2, \ldots, n-1, \tag{3.2}$$

where

$$\hat{\phi}_j = \frac{1}{n-1} \sum_{t=1}^{n-1} r_t b_j(t), \quad j = 1, 2, \ldots, K. \tag{3.3}$$

We propose the variance estimator as

$$\hat{\sigma}^2_J = \frac{1}{n-1} \sum_{i=1}^{n-1} \hat{s}_i. \tag{3.4}$$

For smoothing and reducing the gap between discontinuous points, take an average of three r_i's such as

$$r_i^* = \frac{1}{3}\left(r_{i-1} + r_i + r_{i+1}\right), \quad i = 2, \ldots, n-1.$$

And $\hat{s}_i^*$ is Fourier estimates with r_i^*. Then the proposed variance estimator is

$$\hat{\sigma}_{J3}^2 = \frac{1}{n-3}\sum_{i=2}^{n-1}\hat{s}_i^*. \tag{3.5}$$

Kim and Hart (2011) used Fourier series for the mean change-point estimator and derived its asymptotics. They used $K = 1$ for Fourier estimation. The choice of K turns out to have little impact on the asymptotic results for $\hat{\sigma}_J^2$. For simplicity, proofs are given for the mean level change model with $K = 1$. For general K the results follow in a similar way. If the model allows for nonconstancy of f away from the discontinuity point, then the results continue to hold so long as f varies smoothly away from this point. It follows that there exists β such that

$$|f(x) - f(y)| \leq \beta|x - y|, \quad \text{for all } x, y.$$

Theorem 1. *Suppose $E[\epsilon_i^4] < \infty$, for any $a > 0$*

$$P\left(\left|\hat{\sigma}_J^2 - \sigma^2\right| > a\right) \to 0 \quad \text{as } n \to \infty \tag{3.6}$$

with

$$E\left[\hat{\sigma}_J^2\right] = \sigma^2 + O\left(\frac{1}{n^2}\right), \quad i = 1, 2, \ldots, n-1,$$
$$\mathrm{Var}\left[\hat{\sigma}_J^2\right] = \frac{1}{4n}\left(4E\left[\epsilon^4\right] + 11\sigma^4\right) + O\left(\frac{1}{n}\right).$$

Proof: We have

$$E\left[\hat{\sigma}_J^2\right] = \frac{1}{n-1}\sum_{i=1}^{n-1}E\left[\hat{s}_i\right],$$
$$E\left[\hat{s}_i\right] = E\left[\hat{\phi}_0\right] + E\left[\hat{\phi}_j\right]\sqrt{2}\cos\pi x_i, \quad i = 1, 2, \ldots, n-1.$$

When the underlying function is smooth with $|f'(x_i^*)| < B$ bounded,

$$\begin{aligned}(y_{i+1} - y_i) &= f(x_{i+1}) - f(x_i) + (\epsilon_{i+1} - \epsilon_i)\\ &= f'(x_i^*)(x_{i+1} - x_i) + (\epsilon_{i+1} - \epsilon_i).\end{aligned}$$

Since $\mathrm{Cov}(\epsilon_{i+1}, \epsilon_i) = 0$, $\mathrm{Var}(\epsilon_i) = \sigma^2$, and $x_{i+1} - x_i = 1/n$,

$$E[r_i] = \frac{1}{2}E\left[(y_{i+1} - y_i)^2\right] = \frac{1}{2n^2}\left|f'\left(x_i^*\right)\right|^2 + \sigma^2 = \sigma^2 + O\left(\frac{1}{n^2}\right).$$

Take the expectation as

$$E\left[\hat{\phi}_0\right] = E\left[\frac{1}{n}\sum_{i=1}^{n}r_i\right] = \frac{1}{n}\sum_{i=1}^{n}E[r_i] = \sigma^2 + O\left(\frac{1}{n^2}\right)$$

and

$$E\left[\hat{\phi}_1\right] = \frac{1}{n}\sum_{i=1}^{n} E[r_i]\sqrt{2}\cos\pi x_i$$
$$= \sigma^2\frac{1}{n}\sum_{i=1}^{n}\sqrt{2}\cos\pi x_i + O\left(\frac{1}{n^2}\right) \approx O\left(\frac{1}{n^2}\right)$$

since

$$\frac{1}{n}\sum_{i=1}^{n}\sqrt{2}\cos\pi x_i \approx \sqrt{2}\int_0^1 \cos\pi x dx = 0.$$

Therefore we have

$$E[\hat{s}_i] = \sigma^2 + O\left(\frac{1}{n^2}\right), \quad i = 1, 2, \ldots, n-1,$$

and

$$E\left[\hat{\sigma}_J^2\right] = \sigma^2 + O\left(\frac{1}{n^2}\right).$$

Consider the variance of $\hat{\sigma}_J^2$. Let $\epsilon = \epsilon_i$ for some i. First, calculate

$$\mathrm{Var}\left[\hat{\phi}_0\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}[r_i] + \frac{1}{n^2}\sum_{i<j}\mathrm{Cov}\left(r_i, r_j\right)$$
$$= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}[r_i] + \frac{2}{n^2}\sum_{j=i+1}\mathrm{Cov}(r_i, r_{i+1})$$
$$= \frac{1}{n}\left(E\left[\epsilon^4\right] + \frac{5}{2}\sigma^4\right) + O\left(\frac{1}{n^2}\right)$$

since

$$E\left[r_i^2\right] = \frac{1}{4}E\left[(y_{i+1} - y_i)^4\right]$$
$$= \frac{1}{4}\left(\frac{f'(x_i^*)^4}{n^4} + 2E\left[\epsilon^4\right] + 6\sigma^4\right)$$
$$= \frac{1}{2}\left(E\left[\epsilon^4\right] + 3\sigma^4\right) + O\left(\frac{1}{n^4}\right),$$

and

$$\mathrm{Var}[r_i] = E\left[r_i^2\right] - E[r_i]^2$$
$$= \frac{1}{2}\left(E\left[\epsilon^4\right] + \sigma^4\right) + O\left(\frac{1}{n^2}\right).$$

$$\operatorname{Cov}\left(r_i, r_j\right) = \frac{1}{4} E\left[(y_{i+1} - y_i)^2 (y_i - y_{i-1})^2\right]$$
$$= \begin{cases} \frac{1}{4} E\left[\epsilon^4\right] + \sigma^4 + O\left(\frac{1}{n^2}\right), & j = i-1,\ i = 1, \\ 0, & |j - i| > 1. \end{cases}$$

For the variance of the first sample Fourier coefficient,

$$\begin{aligned}
\operatorname{Var}\left[\hat{\phi}_1\right] &= \frac{2}{n^2} \sum_{i=1}^{n} \operatorname{Var}[r_i] \cos^2 \pi x_i + 4 \sum_{i<j} \cos^2 \pi x_i \cos^2 \pi x_j \operatorname{Cov}\left(r_i, r_j\right) \\
&= \frac{1}{2n}\left(E\left[\epsilon^4\right] + \sigma^4\right) + 4\frac{2}{n^2} \sum_{j=i+1} \cos^2 \pi x_i \cos^2 \pi x_j + O\left(\frac{1}{n}\right) + O\left(\frac{1}{n^2}\right) \\
&= \frac{1}{2n}\left(E\left[\epsilon^4\right] + \sigma^4\right) + O\left(\frac{1}{n}\right) + O\left(\frac{1}{n^2}\right),
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{Var}\left[\hat{s}_i\right] &= \operatorname{Var}\left[\hat{\phi}_0\right] + 2\cos^2 \pi x_i \operatorname{Var}\left[\hat{\phi}_1\right] + \sqrt{2} \cos \pi x_i \operatorname{Cov}\left(\hat{\phi}_0, \hat{\phi}_1\right) \\
&= \frac{1}{n}\left(E\left[\epsilon^4\right] + \frac{5}{2}\sigma^4\right) + 2\cos^2 \pi x_i \frac{1}{2n}\left(E\left[\epsilon^4\right] + \sigma^4\right) O\left(\frac{1}{n}\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\operatorname{Var}\left[\hat{\sigma}_J^2\right] &= E\left[\left(\frac{1}{n-1} \sum_{i=1}^{n-1} \hat{s}_i\right)^2\right] \\
&= \frac{1}{(n-1)^2} \sum_{i=1}^{n-1} \operatorname{Var}\left(\hat{s}_i\right) + \frac{1}{(n-1)^2} \sum_{i<j}^{n-1} 2\operatorname{Cov}\left(\hat{s}_i, \hat{s}_j\right) \\
&= \frac{\sigma^2}{n-1} + O\left(\frac{1}{n}\right) \\
&= O\left(\frac{1}{n}\right).
\end{aligned}$$

For any $a > 0$, use Markov inequality such as

$$\begin{aligned}
P\left(\left|\hat{\sigma}_J^2 - \sigma^2\right| \geq a\right) &\leq \frac{\operatorname{Var}\left(\hat{\sigma}_J^2\right)}{a^2} \\
&= O\left(\frac{1}{n}\right).
\end{aligned}$$

Therefore $\hat{\sigma}_J^2$ is a consistent estimator of σ^2 as $n \to \infty$. For consistency, a similar procedure can be done for $\hat{\sigma}_{J3}^2$. □

4. Simulation

We conducted simulations to investigate the behavior of the proposed change-point estimator. Data were generated from various change-point models with i.i.d. normal errors having mean 0 and variance $\sigma^2 = 1$, and the design points $x_i = i/n$, $i = 1, \ldots, n$. The models considered are in:

$$y_i = f(x_i) + \epsilon_i, \quad i = 1, \ldots, n,$$

(1) One step up function

$$f(t) = \begin{cases} 0, & 0 \le x \le 0.5, \\ 2, & 0.5 < x \le 1. \end{cases}$$

(2) Linear function

$$f(t) = 5x.$$

(3) Quadratic function

$$f(t) = 3x^2.$$

(4) Quadratic and linear function

$$f(t) = \begin{cases} 5x^2, & 0 \le x \le 0.5, \\ 2 + x^2, & 0.5 < x \le 1. \end{cases}$$

(5) Cyclic smooth cosine function

$$f(t) = \cos(8\pi(0.5 - x)), \quad 0 \le x \le 1.$$

(6) Cyclic smooth sine function

$$f(t) = 5\sin(2\pi x).$$

(7) Exponential function

$$f(t) = 2\exp\left(-4(4x-1)^2\right) + 6\exp\left(-16(4x-3)^2\right), \quad 0 \le x \le 1.$$

(8) Two step change function

$$f(t) = \begin{cases} 5x, & 0 \le x \le 0.3, \\ 1.5 - 5x, & 0.3 < x \le 0.7. \\ 4x, & 0.7 < x \le 1. \end{cases}$$

A sample size of $n = 100$ was used in all cases, and 1,000 repetitions were performed for each set. We computed variance estimators with their variances and MSEs. Table 1 shows that the proposed variance estimator has smaller MSEs when there is drift including cycle or two step changes for underlying functions and provides similar MSEs in other models. Therefore our estimator can be an alternative for the variance estimation for nonparametric regression models that include functions with some drift or trend.

Table 1: Comparison of variance estimators in 1,000 repetitions with the sample size $n = 100$

Statistic			Model (1)					Model (2)		
Estimator	Mean	Sd	MSE	Lower	Upper	Mean	Sd	MSE	Lower	Upper
$\hat{\sigma}^2_R$	1.029	0.179	0.032	0.702	1.390	0.998	0.174	0.030	0.691	1.357
$\hat{\sigma}^2_{GSJ}$	1.022	0.203	0.041	0.669	1.450	0.997	0.198	0.039	0.656	1.411
$\hat{\sigma}^2_{HKT}(m = 2)$	1.040	0.167	0.028	0.751	1.375	0.998	0.161	0.026	0.715	1.328
$\hat{\sigma}^2_{HKT}(m = 3)$	1.051	0.172	0.030	0.759	1.389	0.995	0.152	0.023	0.713	1.312
$\hat{\sigma}^2_{TW}(m = \sqrt{n})$	1.051	0.165	0.027	0.779	1.371	0.996	0.162	0.026	0.703	1.343
$\hat{\sigma}^2_{TW}(m = n^{1/3})$	1.034	0.168	0.028	0.740	1.384	0.998	0.197	0.039	0.658	1.405
$\hat{\sigma}^2_{TW}(m = 2)$	1.021	0.202	0.041	0.661	1.440	0.977	0.174	0.030	0.672	1.333
$\hat{\sigma}^2_{J,K=1}$	1.009	0.176	0.031	0.689	1.380	0.977	0.174	0.030	0.672	1.333
$\hat{\sigma}^2_{J3,K=1}$	1.002	0.175	0.030	0.685	1.377	0.970	0.174	0.030	0.672	1.325
Statistic			Model (3)					Model (4)		
Estimator	Mean	Sd	MSE	Lower	Upper	Mean	Sd	MSE	Lower	Upper
$\hat{\sigma}^2_R$	1.004	0.181	0.033	0.688	1.369	0.999	0.176	0.031	0.690	1.375
$\hat{\sigma}^2_{GSJ}$	1.002	0.205	0.042	0.647	1.417	0.996	0.204	0.042	0.653	1.452
$\hat{\sigma}^2_{HKT}(m = 2)$	1.006	0.167	0.028	0.713	1.375	1.003	0.159	0.025	0.718	1.352
$\hat{\sigma}^2_{HKT}(m = 3)$	1.016	0.169	0.029	0.721	1.380	1.014	0.161	0.026	0.725	1.366
$\hat{\sigma}^2_{TW}(m = \sqrt{n})$	1.002	0.157	0.025	0.723	1.330	1.006	0.152	0.023	0.730	1.325
$\hat{\sigma}^2_{TW}(m = n^{1/3})$	1.003	0.166	0.028	0.711	1.362	1.001	0.161	0.026	0.709	1.347
$\hat{\sigma}^2_{TW}(m = 2)$	1.003	0.205	0.042	0.646	1.413	0.996	0.203	0.041	0.654	1.444
$\hat{\sigma}^2_{J,K=1}$	0.983	0.180	0.032	0.667	1.351	0.978	0.176	0.031	0.671	1.363
$\hat{\sigma}^2_{J3,K=1}$	0.977	0.180	0.032	0.662	1.339	0.972	0.177	0.031	0.669	1.355
Statistic			Model (5)					Model (6)		
Estimator	Mean	Sd	MSE	Lower	Upper	Mean	Sd	MSE	Lower	Upper
$\hat{\sigma}^2_R$	1.011	0.172	0.030	0.689	1.355	1.021	0.173	0.030	0.716	1.388
$\hat{\sigma}^2_{GSJ}$	0.995	0.198	0.039	0.635	1.402	0.997	0.197	0.039	0.655	1.438
$\hat{\sigma}^2_{HKT}(m = 2)$	1.034	0.160	0.026	0.742	1.347	1.055	0.165	0.027	0.768	1.383
$\hat{\sigma}^2_{HKT}(m = 3)$	1.045	0.164	0.027	0.750	1.361	1.066	0.170	0.029	0.776	1.398
$\hat{\sigma}^2_{TW}(m = \sqrt{n})$	1.070	0.167	0.028	0.774	1.366	1.024	0.154	0.024	0.741	1.350
$\hat{\sigma}^2_{TW}(m = n^{1/3})$	1.002	0.156	0.024	0.709	1.318	0.999	0.158	0.025	0.699	1.338
$\hat{\sigma}^2_{TW}(m = 2)$	0.995	0.197	0.039	0.635	1.389	0.998	0.197	0.039	0.656	1.432
$\hat{\sigma}^2_{J,K=1}$	0.991	0.170	0.029	0.676	1.342	0.999	0.169	0.028	0.687	1.355
$\hat{\sigma}^2_{J3,K=1}$	0.984	0.169	0.029	0.670	1.328	0.992	0.168	0.028	0.684	1.349
Statistic			Model (7)					Model (8)		
Estimator	Mean	Sd	MSE	Lower	Upper	Mean	Sd	MSE	Lower	Upper
$\hat{\sigma}^2_R$	1.039	0.176	0.031	0.717	1.403	1.136	0.239	0.057	0.795	1.552
$\hat{\sigma}^2_{GSJ}$	1.002	0.196	0.038	0.641	1.417	1.092	0.240	0.058	0.696	1.589
$\hat{\sigma}^2_{HKT}(m = 2)$	1.096	0.185	0.034	0.806	1.431	1.206	0.278	0.077	0.890	1.612
$\hat{\sigma}^2_{HKT}(m = 3)$	1.107	0.192	0.037	0.814	1.445	1.219	0.289	0.083	0.900	1.629
$\hat{\sigma}^2_{TW}(m = \sqrt{n})$	1.151	0.214	0.046	0.874	1.455	1.287	0.345	0.119	0.944	1.691
$\hat{\sigma}^2_{TW}(m = n^{1/3})$	1.017	0.158	0.025	0.724	1.359	1.168	0.253	0.064	0.849	1.570
$\hat{\sigma}^2_{TW}(m = 2)$	1.001	0.196	0.038	0.650	1.412	1.090	0.238	0.057	0.700	1.572
$\hat{\sigma}^2_{J,K=1}$	1.018	0.170	0.029	0.702	1.380	1.116	0.226	0.051	0.788	1.539
$\hat{\sigma}^2_{J3,K=1}$	1.012	0.168	0.028	0.696	1.374	1.110	0.222	0.050	0.774	1.534

5. Applications

We applied the variance component methods to some real data sets. Table 1 provides the estimation results. The techniques derived in this paper were applied to Nile River flow data from 1871 to 1970 (Figure 2). The sample variance that does not account for drift, is $\hat{\sigma}^2_{NileRiver} = 28637.95$ which seems overestimated. Table 2 gives the variance estimates of Nile River data and shows that the variance considering trend is necessary.

Table 2: Comparison of variance estimators for real data

Data	Nile River flow	fMRI: brain R1	fMRI: brain R10
Sample size	$n = 100$	$n = 2048$	$n = 2048$
$\hat{\sigma}^2$	28637.95	59.067	200.479
$\hat{\sigma}^2_R$	13998.77	33.748	120.164
$\hat{\sigma}^2_{GSJ}$	13206.36	29.434	110.023
$\hat{\sigma}^2_{HKT}(m = 2)$	15531.97	39.434	135.490
$\hat{\sigma}^2_{HKT}(m = 3)$	15692.10	39.454	135.556
$\hat{\sigma}^2_{TW}\ (m = \sqrt{n})$	17404.00	51.943	174.491
$\hat{\sigma}^2_{TW}\ (m = n^{1/3})$	15228.11	48.340	155.452
$\hat{\sigma}^2_{TW}\ (m = 2)$	13023.64	29.943	109.957
$\hat{\sigma}^2_{J,K=1}$	13998.77	33.670	120.145
$\hat{\sigma}^2_{J3,K=1}$	13921.91	33.669	120.031

fMRI = functional magnetic resonance imaging

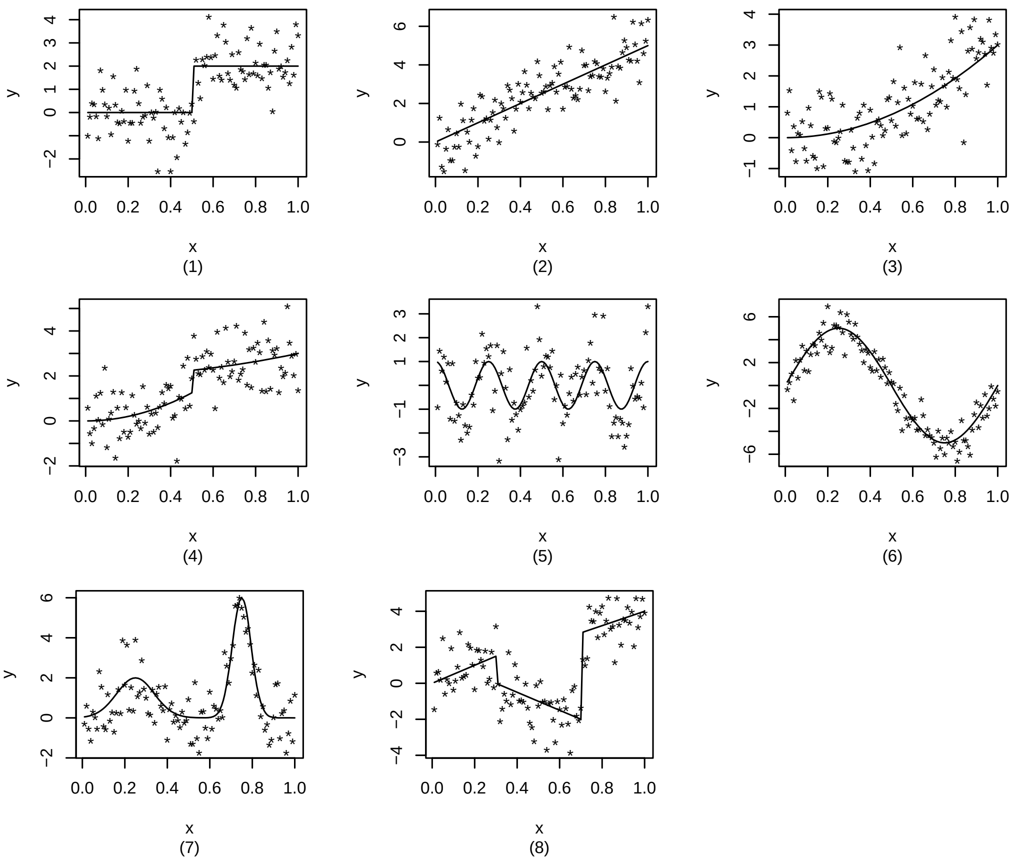

Figure 1: *The true underlying functions (solid lines) with a typical simulated data set of size* $n = 100$ *and variance* $\sigma^2 = 1$.

Recent attention has been devoted to investigating fMRI in neuroimaging for brain activity and connectivity patterns over time. Logothetis *et al.* (2001) provide a neurophysiological investigation of the basis of the fMRI signal for brain function. The fMRI data has some fluctuation according to

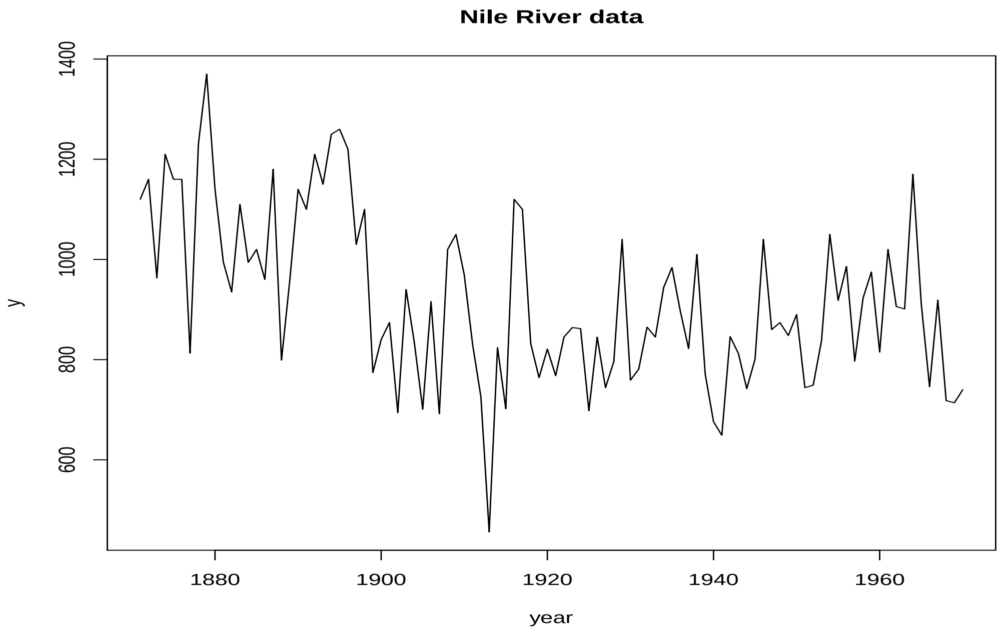

Figure 2: *Nile River flow data.*

stimuli or disease. Several change point methods have been proposed for fMRI signals (Lindquist *et al.*, 2007, 2014). Were acquired functional magnetic resonance imaging (fMRI) time series data from five healthy volunteers in the resting state to estimate functional connectivity between 90 cortical and subcortical regions. Volunteers had no personal history of neurologic or psychiatric disorders and were not abusing alcohol or illicit drugs.

Each participant was scanned on a single occasion, lying quietly at rest with eyes closed for 37 min, 44 s. Gradient-echo echoplanar imaging (EPI) data depicting blood oxygen level-dependent (BOLD) contrast were acquired using a Med Spec S300 scanner (Bruker Medical, Ettlingen, Germany) operating at 3.0T in the Wolfson Brain Imaging Centre (Cambridge, UK).

We selected region 1 (R1; precuneus) and region 10 (R10; dorsal cingulate gyrus) for variance estimation to give an example. Achard *et al.* (2006) previously studied functional connectivity with these data. Table 2 shows that the variance with trend correction is smaller than simple sample variance. The variance of R10 is also bigger than R1 since R10 has more variability due to regional characteristics; in addition, it might be more biased since the sample variance (without considering the underlying function) is bigger than other variance estimates.

Our proposed method can reflect the underlying function and can be used for variance component especially without little information about the underlying functions.

6. Concluding remarks

This paper develops one possible solution to the inference problem considered when estimating variance components. The proposed estimator is based on some method of nonparametric regression or “smoothing” with Fourier series estimation. The primary advantage of taking such an approach is the simplicity that can be quickly coded and a straightforward explanation to scientists. It can also be

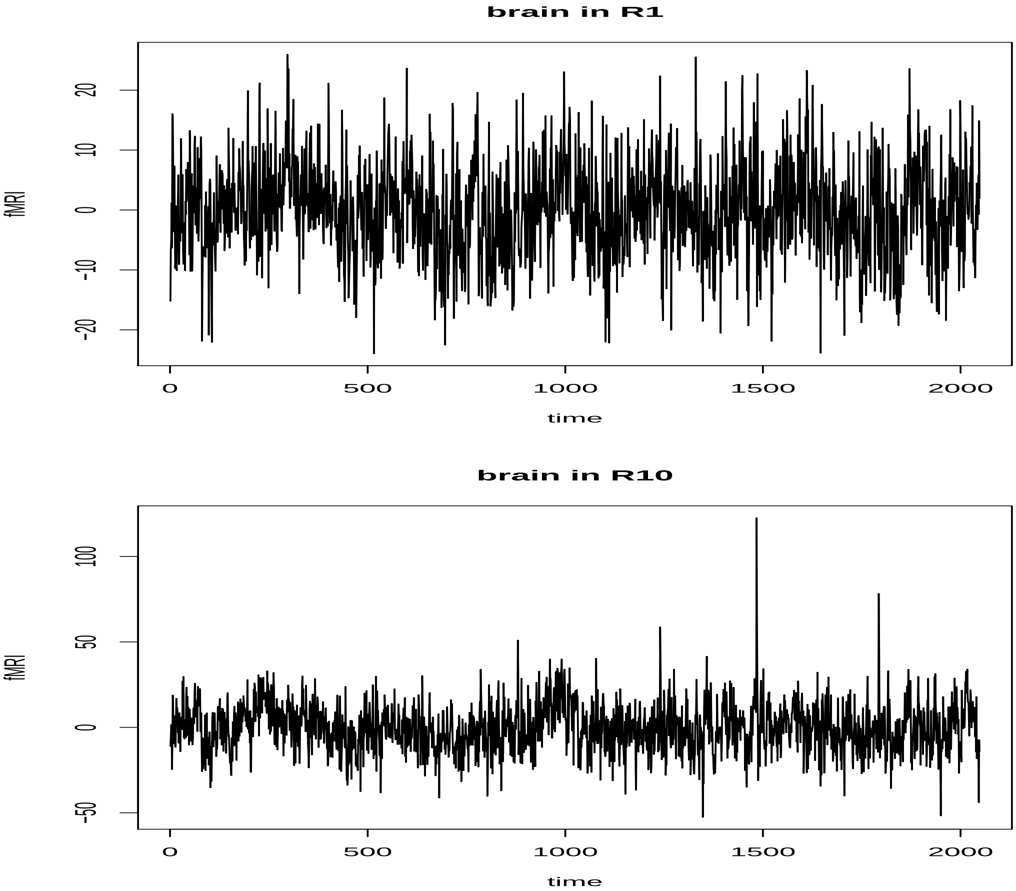

Figure 3: *Functional magnetic resonance imaging (fMRI) data from brain region R1 (upper) and fMRI data from brain R10 (lower).*

applied to similar models such as regressions. We expect further research on variance and covariance with the time series data incorporation the dependency.

Acknowledgement

This research was supported by the LG Yonam Foundation Overseas Research Professor Program in 2013.

References

Achard S, Raymond S, Whitcher B, Suckling J, and Bullmore E (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs, *Journal of Neuroscience*, **26**, 63–72.

Brown LD and Levine M (2007). Variance estimation in nonparametric regression via the difference sequence method, *Annals of Statistics*, **35**, 2219–2232.

Buckley MJ, Eagleson GK, and Silverman BW (1988). The estimation of residual variance in nonparametric regression, *Biometrika*, **75**, 189–199.

Chaudhuri A (1992). A note on estimating the variance of the regression estimator, *Biometrika*, **79**, 217–218.

Dette H, Munk A, and Wagner T (1998). Estimating the variance in nonparametric regression-what is a reasonable choice?, *Journal of Royal Statistical Society B*, **60**, 751–764.

Eddington, A. S. and Plakidis, S. (1929). Irregularities of period of long-period variable stars, *Monthly Notices of the Royal Astronomical Society*, **90**, 65–71.

Gasser T, Sroka L, and Jennen-Steinmetz C (1986). Residual variance and residual pattern in nonlinear regression, *Biometrika*, **73**, 625–633.

Hall P, Kay JW, and Titterington DM (1990). On variance estimation in nonparametric regression, *Biometrika*, **77**, 515–419.

Hastie T and Tibshirani R (1990). *Generalized Additive Models*, Chapman and Hall, London.

Kim J and Hart J (2011). A change-point estimator using local Fourier series, *Journal of Nonparametric Statistics*, **23**, 83–98.

Kott PS (1990). Estimating the conditional variance of a design consistent regression estimator, *Journal of Statistical Planning and Inference*, **24**, 287–296.

Lindquist MA, Waugh C, and Wager TD (2007). Modeling state-related fMRI activity using change-point theory, *NeuroImage*, **35**, 1125–1141.

Lindquist MA, Xu Y, Nebel MB, and Caffo BS (2014). Evaluating dynamic bivariate correlations in resting-state fMRI: A comparison study and a new approach, *NeuroImage*, **101**, 531–546.

Logothetis NK, Pauls J, Augath M, Trinath T, and Oeltermann A (2001). Neurophysiological investigation of the basis of the fMRI signal, *Nature*, **412**, 150–157.

Müller UU, Schick A, and Wefelmeyer W (2003). Estimating the error variance in nonparametric regression by a covariate-matched U-statistic, *Statistics*, **37**, 179–188.

Ogden RT and Collier GL (2002). Inference on variance components of autocorrelated sequences in the presence of drift, *Journal of Nonparametric Statistics*, **14**, 409–420.

Reinsch C (1967). Smoothing by spline functions, *Numerical Mathematics*, **24**, 375–382.

Rice JA (1984). Bandwidth choice for nonparametric regression, *Annals of Statistics*, **12**, 1215–1230.

Sarndal CE, Swensson B, and Wretman JH (1989). The weighted residual technique for estimating the variance of the general regression estimator, *Biometrika*, **76**, 527–537.

Tong T and Wang Y (2005). Estimating residual variance in nonparametric regression using least squares, *Biometrika*, **92**, 821–830.

Wahba G (1990). Spline models for observational data, *CBMS-NSF Regional Conference Series in Applied Mathematics*, **59**, PA:SIAM.

Received September 9, 2015; Revised November 22, 2015; Accepted November 22, 2015