Journal of Statistical Computation and Simulation
Vol. 76, No. 4, April 2006, 317–329

Testing and estimating for change in long memory parameter

LIHONG WANG* and JINDE WANG

Department of Mathematics, Institute of Mathematical Science, Nanjing University, Nanjing 210093, P.R. China

(Revised 20 May 2004; in final form 20 October 2004)

We consider a long memory time series where the long-memory parameter H appears to change with time. We are interested in detecting and estimating the change-point and studying the asymptotic properties of the estimator and the test statistic. We apply the method to two practical data series to investigate the presence of change in H. Simulations and data examples confirm the validity of the test and estimation.

Keywords: Asymptotic property; Change-point; Estimation; Long memory

1. Introduction

Let $\{\varepsilon_j, -\infty < j < \infty\}$ be a sequence of independent identically distributed random variables with mean zero, finite variance τ^2. We consider a long memory linear process x_t

$$x_t = \mu + \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} a_j^2 < \infty, \quad t = 0, \pm 1, \ldots,$$

which has mean μ, lag-j-autocovariance γ_j and spectrum $f(\lambda) = (2\pi)^{-1} \sum_{j=-\infty}^{\infty} \gamma_j \cos j\lambda$ which satisfies the following condition,

$$f(\lambda) \sim L(\lambda^{-1})\lambda^{1-2H} \quad \text{as} \quad \lambda \to 0_+, \ \frac{1}{2} < H < 1, \tag{1}$$

where $\sim$ indicates asymptotic equivalence and $L(\cdot)$ is a slowly varying function at infinity.

H is often called long-memory parameter, it characterizes the long-term dependent structure of the sequence x_t. As pointed out in Beran and Terrin [1] and Ray and Tsay [2], for some long memory time series, H seems to vary over time and even small changes of H would have a strong impact on the statistical inference for such long memory process. It is practically

*Corresponding author. Email: lhwang@nju.edu.cn

Journal of Statistical Computation and Simulation
ISSN 0094-9655 print/ISSN 1563-5163 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/10629360500107766

important to decide whether H has changed or not over the whole observational period and to estimate the time of change, if H indeed changed with time.

Beran and Terrin [1] proposed an approach based on the comparison of Whittle estimator to the detection of a change in H, but their asymptotic results are incorrect. Horváth and Shao [3] obtained the correct limit distribution of their test statistics. Horváth [4] suggested another test based on quadratic forms of Whittle estimator of H. Ray and Tsay [2] used a Bayesian method for detecting changes in H, which is also likelihood-based in the Gaussian setting.

Whittle estimator is based on the periodogram at all frequencies. However, in some situations, it is difficult or unnecessary to model the whole spectral space and a misspecified model may lead to inconsistent, biased estimates of H and then result in failure to detect structural changes, see, *e.g.* refs. [1, 2, 5]. Therefore a semiparametric estimate of H which relies mainly on low periodogram ordinates is preferable. The semiparametric estimate also has some computational advantage over Whittle estimates in models such as fractional *ARIMA* (*FARIMA*) model, see ref. [6].

In this article, we suggest a method based on the comparison of a semi-parametric estimate introduced by Robinson [6] for testing the no-change null hypothesis H_0 against the alternative that H changes somewhere in the observed time period. The test statistic and its asymptotic distribution are given in section 2. If H_0 is rejected, we need to know the time of change. In section 2, we also describe the estimator of the change-point and its weak consistency. We apply the method to investigate the change in persistence of two data series, the yearly Nile river minima and the VBR data in section 3. The proofs of the theorems are presented in Appendix A.

2. Main results

Our change-point problem can be formulated as

$$f(\lambda) = \begin{cases} f_1(\lambda) \sim L(\lambda^{-1})\lambda^{1-2H}, & \text{for the process } x_i, \quad 1 \le i \le k^*, \\ f_2(\lambda) \sim L(\lambda^{-1})\lambda^{1-2H'}, & \text{for the process } x_i, \quad k^* < i \le n, \end{cases}$$

as $\lambda \to 0_+$, where $H \neq H'$. The null hypothesis is

$$H_0\colon k^* \ge n.$$

Let

$$F(\lambda) = \int_0^\lambda f(t)\,\mathrm{d}t.$$

First we assume that H_0 holds. Then assumption (1) and Yong [7, p. 20] imply

$$F(\lambda) \sim \frac{L(\lambda^{-1})}{2-2H}\lambda^{2-2H} \quad \text{as } \lambda \to 0_+. \tag{2}$$

This yields, for any $0 < q < 1$,

$$\frac{F(\lambda q)}{F(\lambda)} \sim q^{2-2H} \quad \text{as } \lambda \to 0_+.$$

Hence, in order to estimate H, we should first estimate $F(\lambda)$. Robinson [6] provided an estimator of $F(\lambda)$ which is invariant to the mean μ and also leads itself to more direct computation,

$$\widehat{F}_n(\lambda) = \frac{2\pi}{n} \sum_{j=1}^{n\lambda/2\pi} I_n(\lambda_j),$$

where $\lambda_j = 2\pi j/n$ and

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} x_t \mathrm{e}^{it\lambda} \right|^2.$$

Then, Robinson [6] suggested the consistent estimator of H,

$$\widehat{H}_{mq,n} = 1 - \frac{\log\{\widehat{F}_n(\lambda_m q)/\widehat{F}_n(\lambda_m)\}}{2\log q}, \quad 1 \le m < n.$$

In this article, we use this semiparametric estimator to construct the test statistic

$$\phi_{n,q}(t, \lambda_m) = n^{1/2} t(1-t) 2\log q (\widehat{H}_{mq,[nt]} - \widehat{H}^*_{mq,n-[nt]}) \widehat{F}_{[nt]}(\lambda_m),$$

where

$$\widehat{H}_{mq,k} = 1 - \frac{\log\{\widehat{F}_k(\lambda_m q)/\widehat{F}_k(\lambda_m)\}}{2\log q}, \quad \widehat{H}^*_{mq,n-k} = 1 - \frac{\log\{\widehat{F}^*_{n-k}(\lambda_m q)/\widehat{F}^*_{n-k}(\lambda_m)\}}{2\log q},$$

here $\widehat{F}_k(\lambda_m)$ (respectively, $\widehat{F}^*_{n-k}(\lambda_m)$) is the estimator of $F(\lambda_m)$ based on observations $x_1, \ldots, x_k$ (respectively, on observations $x_{k+1}, \ldots, x_n$), *i.e.*

$$\widehat{F}_k(\lambda_m) = \frac{2\pi}{n} \sum_{j=1}^{m} I_k(\lambda_j), \quad \widehat{F}^*_{n-k}(\lambda_m) = \frac{2\pi}{n} \sum_{j=1}^{m} I^*_{n-k}(\lambda_j),$$

where

$$I_k(\lambda) = \frac{1}{2\pi k} \left| \sum_{t=1}^{k} x_t \, \mathrm{e}^{it\lambda} \right|^2, \quad I^*_{n-k}(\lambda) = \frac{1}{2\pi(n-k)} \left| \sum_{t=k+1}^{n} x_t \, \mathrm{e}^{it\lambda} \right|^2.$$

Assume the following assumption holds:

$$1/m + m/n \to 0, \quad \text{as } n \to \infty. \tag{3}$$

Let

$$b_k(m) = \frac{1}{n} \sum_{j=1}^{m} \mathrm{e}^{ik\lambda_j}, \quad c_k(m) = a_k b_0(m) + 2 \sum_{l=1}^{k} a_{k-l} b_l(m),$$

and

$$\sigma_m^2 = (\mathrm{E}\varepsilon_0^4 - \tau^4) \left(\sum_{j=0}^{\infty} a_j c_j(m) \right)^2 + \tau^4 \sum_{l=1}^{\infty} \left(\sum_{j=0}^{\infty} \{a_j c_{j+l}(m) + c_j(m) a_{j+l}\} \right)^2.$$

THEOREM 2.1 *Let H_0, assumptions* (1) *and* (3) *hold, and assume that* $1/2 < H < 3/4$ *and* $\mathrm{E}|\varepsilon_0|^{4+r} < \infty$ *for some* $r > 0$. *Then we have, for any* $0 < q < 1$,

$$\phi_{n,q}(t, \lambda_m) \xrightarrow{\mathcal{D}} \sigma(1 - q^{2H-1}) B(t)$$

in $D[0,1]$ *as* $n \to \infty$, *where* '$\xrightarrow{\mathcal{D}}$' *stands for convergence in distribution*, $\sigma^2 = \lim_{m\to\infty} \sigma_m^2$ *and* $\{B(t), 0 \le t \le 1\}$ *is a Brownian bridge.*

Let

$$\psi_{mq,n}(k) = (1 - q^{2H-1})^{-1} \phi_{n,q}(k/n, \lambda_m).$$

COROLLARY 2.1 *Under conditions of Theorem* 2.1 *we have*, *as* $n \to \infty$,

$$\max_{1\le k<n} |\psi_{mq,n}(k)| \xrightarrow{\mathcal{D}} \sigma \sup_{0<t<1} |B(t)|.$$

Now Corollary 2.1 solves our test problem. We reject H_0 for large values of $\max_{1\le k<n} |\psi_{mq,n}(k)|$. Quantiles can be obtained from the distribution of $\sup_{0<t<1} |B(t)|$ [cf. Theorem 1.5.1 of ref. 8]. The limit of $\phi_{n,q}(t, \lambda_m)$ for $3/4 < H < 1$ remains to be derived. It is likely to be a non-Gaussian process which can be represented by a multiple Wiener–Itô–Dobrushin integral [cf. refs. 9, 10].

In practice, we may have to fix m and replace H and σ by using any consistent estimates. However, σ is somewhat difficult to determine because of the complicated terms in σ_m^2. We provide the approximate critical values for the test with different values of H and q in section 3.

We continue with the estimation of the change-point k^*. The estimator of k^* is defined by

$$\hat{k}_{mq} = \min \left\{ k\colon |\psi_{mq,n}(k)| = \max_{1\le i<n} |\psi_{mq,n}(i)| \right\}.$$

One more assumption is needed to show the weak consistency of $\hat{k}_{mq}$.

$$k^* = [n\theta] \quad \text{for some } 0 < \theta < 1. \tag{4}$$

THEOREM 2.2 *Assume that the alternative and the assumptions* (1), (3) *and* (4) *hold*, *then we have*, *for any* $0 < q < 1$,

$$\hat{k}_{mq}/n \xrightarrow{\mathrm{P}} \theta$$

as $n \to \infty$, *where* '$\xrightarrow{\mathrm{P}}$' *denotes convergence in probability.*

3. Numerical examples

All simulations and statistical computations in this section have been done by using SPLUS functions.

As pointed out in section 2, it is not easy to get the critical values for our test due to the appearance of the constant σ. We suggest to use the simulated sampling distributions of $\max_{1\le k<n} |\psi_{mq,n}(k)|$, because they asymptotically approach the limit distribution according to Corollary 2.1.

The output shown in table 1 displays quantiles of the simulated distributions of $\max_{1\le k<n} |\psi_{mq,n}(k)|$ for each of five values of H, $H = 0.6$ (0.1) 0.9 and three values of q, $q = 0.25$ (0.25) 0.75. The results are obtained from 10,000 simulations based on $m = n^{0.5}$, $m = n^{0.45}$ and $m = n^{0.55}$. The simulated 10,000 sequences of length 10,000 follow the *FARIMA*$(0, H - 0.5, 0)$ model. Then, we can decide whether H_0 is true or not by comparing the test statistic $\max_{1\le k<n} |\psi_{mq,n}(k)|$ against the approximate critical values in table 1. As expected, the critical values increase in H for the same value of m, but they are very stable

Table 1. Simulated 10%-, 5%- and 1%-critical values for our test.

		α		
m	q	0.1	0.05	0.01
$H = 0.6$				
$n^{0.45}$	0.25	0.963	1.061	1.165
	0.5	0.964	1.060	1.164
	0.75	0.962	1.059	1.163
$n^{0.5}$	0.25	0.899	1.027	1.113
	0.5	0.900	1.028	1.112
	0.75	0.898	1.025	1.110
$n^{0.55}$	0.25	0.875	1.013	1.120
	0.5	0.874	1.014	1.119
	0.75	0.876	1.012	1.122
$H = 0.7$				
$n^{0.45}$	0.25	1.026	1.108	1.219
	0.5	1.024	1.109	1.217
	0.75	1.025	1.108	1.220
$n^{0.5}$	0.25	1.007	1.085	1.208
	0.5	1.008	1.085	1.207
	0.75	1.006	1.083	1.209
$n^{0.55}$	0.25	1.002	1.061	1.210
	0.5	0.988	1.060	1.212
	0.75	0.985	1.092	1.213
$H = 0.8$				
$n^{0.45}$	0.25	1.099	1.215	1.323
	0.5	1.095	1.209	1.328
	0.75	1.096	1.211	1.320
$n^{0.5}$	0.25	1.027	1.184	1.304
	0.5	1.023	1.187	1.304
	0.75	1.028	1.188	1.301
$n^{0.55}$	0.25	1.020	1.191	1.273
	0.5	1.021	1.190	1.272
	0.75	1.018	1.191	1.270
$H = 0.9$				
$n^{0.45}$	0.25	1.193	1.268	1.394
	0.5	1.190	1.269	1.391
	0.75	1.189	1.267	1.393
$n^{0.5}$	0.25	1.108	1.224	1.317
	0.5	1.107	1.223	1.315
	0.75	1.105	1.221	1.312
$n^{0.55}$	0.25	1.039	1.219	1.327
	0.5	1.035	1.218	1.325
	0.75	1.037	1.220	1.327

in q. This is because the limit random variable $\sigma \sup_{0<t<1} |B(t)|$ in Corollary 2.1 does not depend on q, but the constant σ increases as H increases. In the following simulations, we focus on $q = 0.5$ and remove q in all of the notations.

As an illustration and a check on our method, we first conduct a small simulation study using series containing a change in long-memory parameter. Secondly, we give the practical application of our method to the Nile river data and the VBR data (the logarithm of the number

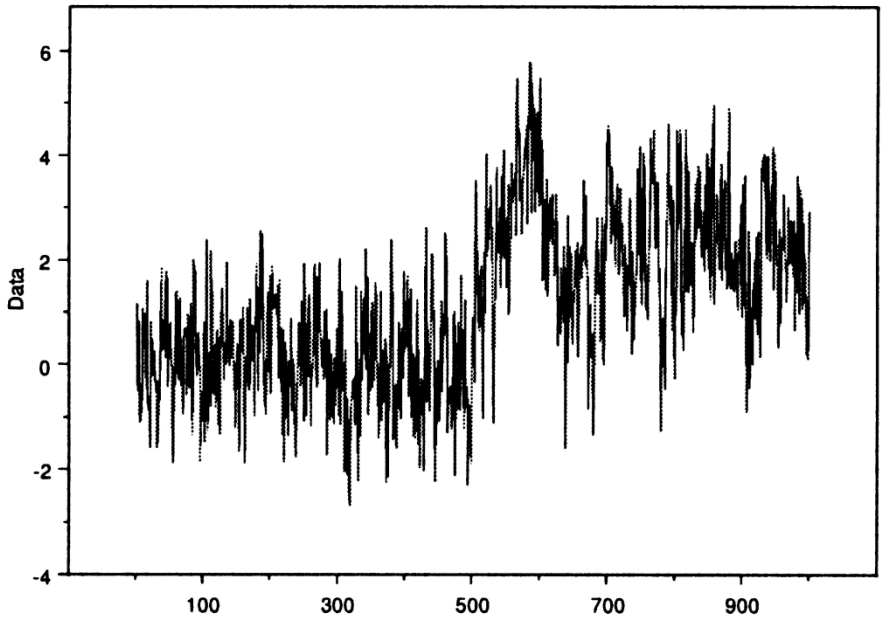


Figure 1. Simulated *FARIMA*$(0, H - 1/2, 0)$ model with $n = 1000$, $k^* = 500$, $H = 0.6$ and $H' = 0.85$.

of ATM cells per frame for a certain video scene), which are widely believed to exhibit long-range dependence. Historically, the analysis of the Nile river data led to the discovery of the Hurst effect. A characterization of the dependence structure of the VBR data is very useful for the assessment of the capacity a communication network has to have in order to guarantee reliable transmission [see, *e.g.* ref. 11].

We generate two *FARIMA*$(0, H - 1/2, 0)$ time series with $H = 0.6$ and $H = 0.85$, respectively. Then, we combine these data into one series of length $n = 1000$ with change in the long-memory parameter of size 0.25 at $k^* = 250$, 500 and 750, respectively. For such simulated time series, we use the estimator $\hat{k}_m$ to detect the change-points. The long memory

Table 2. Estimation results for the simulated FARIMA series.

m	$\widehat{H}_{m,n}$	$\hat{k}_m$	$\psi_{m,n}(\hat{k}_m)$	$\widehat{H}_{m,\hat{k}_m}$	$\widehat{H}^*_{m,n-\hat{k}_m}$
$k^* = 250$					
$n^{0.45}$	0.748	300	1.611	0.653	0.861
$n^{0.5}$	0.765	261	1.549	0.566	0.829
$n^{0.55}$	0.776	259	1.533	0.579	0.821
$k^* = 500$					
$n^{0.45}$	0.728	519	1.477	0.542	0.867
$n^{0.5}$	0.731	508	2.160	0.597	0.854
$n^{0.55}$	0.747	510	1.661	0.561	0.875
$k^* = 750$					
$n^{0.45}$	0.819	709	1.521	0.623	0.908
$n^{0.5}$	0.823	739	1.837	0.631	0.846
$n^{0.55}$	0.833	743	2.320	0.626	0.795

parameter H for the whole series should be estimated by its consistent estimator $\widehat{H}_{m,n}$. The series for $k^* = 500$ is plotted in figure 1.

The testing and estimating results are reported in table 2. In addition to the $\hat{k}_m$ that is based on $m = n^{0.5}$, which is commonly used for estimating long-memory parameter, we also calculate the $\hat{k}_m$'s based on $m = n^{0.45}$ and $n^{0.55}$ to check the sensitivity of our estimator to the choice of m. Plots of the $\psi_{m,n}(k)$'s with $k^* = 500$, $1 \le k < n$, $m = n^{0.45}, n^{0.5}, n^{0.55}$ are given in figure 2. The plots for $k^* = 250$ and 750 are similar, so we left them out. By checking the critical values in table 1, we should reject H_0 for all the three time series.

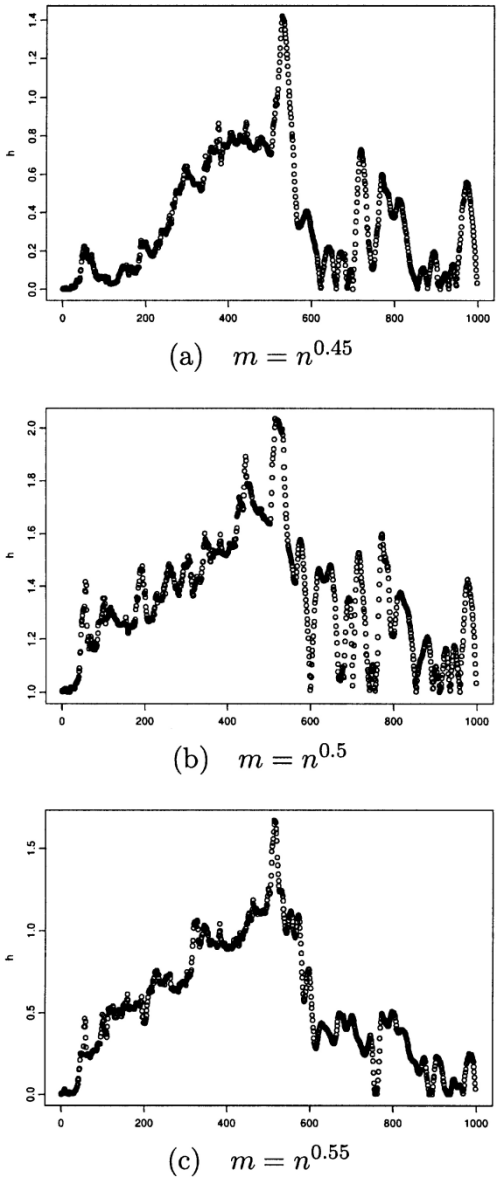


Figure 2. The graphs of $\psi_{m,n}(k)$ with $m = n^{0.45}, n^{0.5}, n^{0.55}$ for the simulated *FARIMA* series in figure 1.

Table 2 and figure 2 give a positive support to our change-point detection and estimation scheme. We also note that the estimates $\hat{k}_m$ and $\hat{H}_{m,\hat{k}_m}$, $\hat{H}^*_{m,n-\hat{k}_m}$ are qualitatively the same across different choices of m. But the estimates corresponding to $m = n^{0.45}$ have larger biases than the other two.

The series of the Nile river data and the VBR data are plotted in figure 3. A *FARIMA*(0, $H - 1/2, 0$) model with $H = 0.837$ turns out to be a good model for the Nile river data [12]. Table 3 and figure 4 report the results of applying our method. Our results indicate that H_0

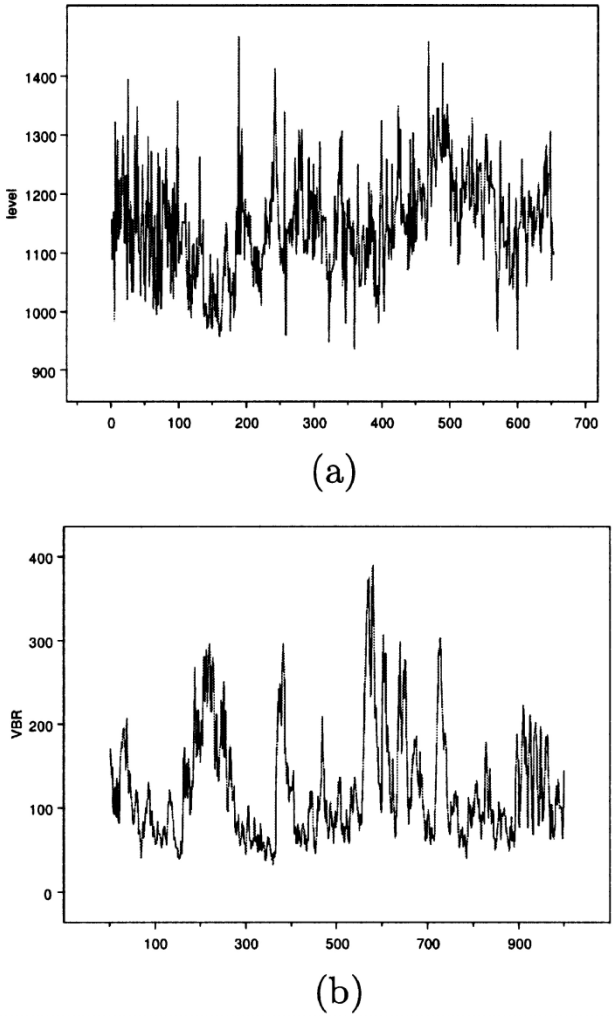
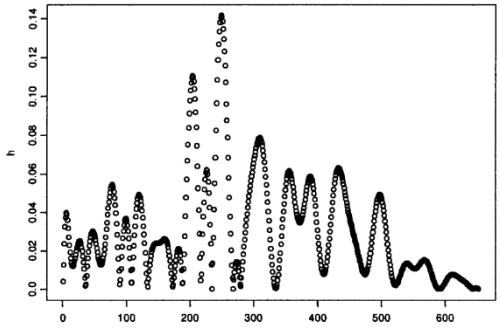


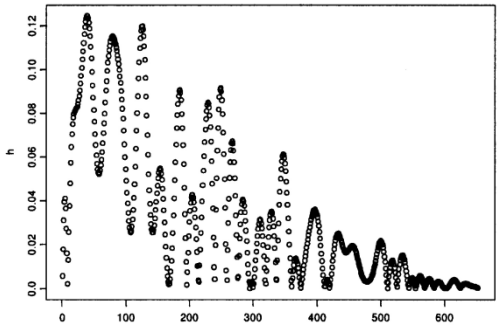
Figure 3. (a) Yearly minimum water level at the Roda gauge for the year 622–1274. (b) VBR data: logarithm of the number of ATM cells per frame. The data sets can be found in the Appendix of ref. [11].

Table 3. Estimation results for the Nile river data.

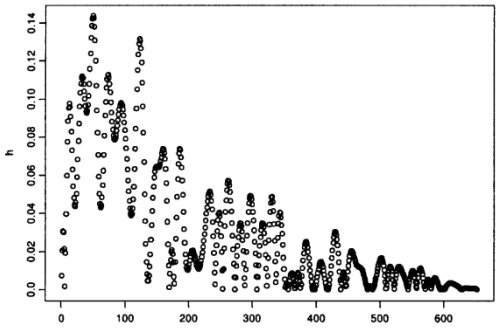
m	$\hat{H}_{m,n}$	$\hat{k}_m$	$\psi_{m,n}(\hat{k}_m)$	$\hat{H}_{m,\hat{k}_m}$	$\hat{H}^*_{m,n-\hat{k}_m}$
$n^{0.45}$	0.825	249	0.142	0.979	0.987
$n^{0.5}$	0.826	39	0.125	0.964	0.948
$n^{0.55}$	0.836	50	0.144	0.960	0.946



(a) $m = n^{0.45}$



(b) $m = n^{0.5}$



(c) $m = n^{0.55}$

Figure 4. The graphs of $\psi_{m,n}(k)$ with $m = n^{0.45}, n^{0.5}, n^{0.55}$ for the Nile river data.

should be accepted. The values of $\psi_{m,n}(k)$ are very small and there are not much differences among the $\psi_{m,n}(k)$'s, the $\hat{k}_m$'s are very unstable in m and the estimated size of the change in H is very small. This suggests that, for the Nile river data, H remains constant during the whole observational period. The estimators $\widehat{H}_{m,n}$ are stable in m, lying between 0.825 and 0.836. However, Ray and Tsay [2] showed that there is a change in the value of H around the 100th observation for the Nile river data. This leads us to conjecture that our method may not be efficient for detecting the changes at the beginning or the end of the data.

As pointed out in chapter 1 of ref. [11], the dependent structure of the VBR data is much more complicated than that of the Nile river data. A closer look at figure 3(b) suggests that the

Table 4. Estimation results for the VBR data.

m	$\hat{H}_{m,n}$	$\hat{k}_m$	$\psi_{m,n}(\hat{k}_m)$	$\hat{H}_{m,\hat{k}_m}$	$\hat{H}^*_{m,n-\hat{k}_m}$
$n^{0.45}$	0.811	301	1.333	0.988	0.907
$n^{0.5}$	0.813	251	1.933	0.987	0.882
$n^{0.55}$	0.828	245	1.565	0.981	0.892

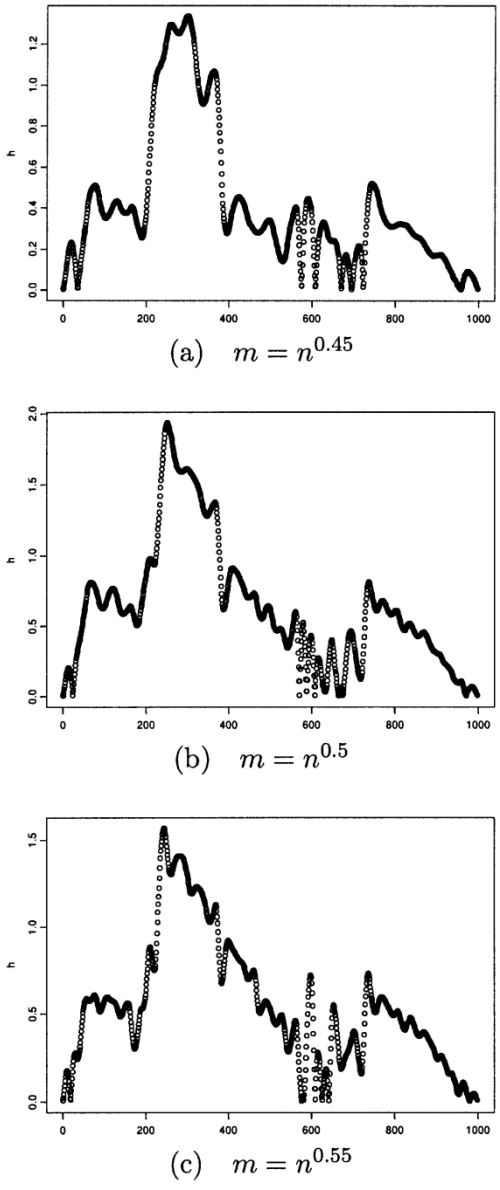


Figure 5. The graphs of $\psi_{m,n}(k)$ with $m = n^{0.45}, n^{0.5}, n^{0.55}$ for the VBR data.

value of H for the first 300 or so observations might be different from that for the subsequent data. Table 4 reports the change-point estimators of the VBR data by applying our estimation procedure for different choices of m. Figure 5 shows the graphs of the $\psi_{m,n}(k)$'s. The results confirm the visual impression in figure 3(b). We conclude that, based on $m = n^{0.5}$, the long-memory parameter H of the VBR data changes at the 251th observation. For the first 251 observations, the estimated H is 0.987, thereafter the long-memory parameter reduces approximately to 0.882.

Acknowledgement

The project is sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

References

[1] Beran, J. and Terrin, N., 1996, Testing for a change of the long-memory parameter. *Biometrika*, **83**, 627–638.
[2] Ray, B.K. and Tsay, R.S., 2002, Bayesian methods for change-point detection in long-range dependent processes. *Journal of Time Series Analysis*, **23**(6), 687–706.
[3] Horváth, L. and Shao, Q.M., 1999, Limit theorems for quadratic forms with applications to Whittle's estimate. *The Annals of Applied Probability*, **9**, 146–187.
[4] Horváth, L., 2001, Change-point detection in long-memory processes. *Journal of Multivariate Analysis*, **78**, 218–234.
[5] Geweke, J. and Porter-Hudak, S., 1983, The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, 221–238.
[6] Robinson, P.M., 1994, Semiparametric analysis of long-memory time series. *The Annals of Statistics*, **22**, 515–539.
[7] Yong, C.H., 1974, *Asymptotic Behavior of Trigonometric Series* (Hong Kong: Chinese University).
[8] Csörgö, M. and Révész, P., 1981, *Strong Approximations in Probability and Statistics* (New York: Academic Press).
[9] Fox, R. and Taqqu, M.S., 1985, Non-central limit theorems for quadratic forms in random variables having long-range dependence. *The Annals of Probability*, **13**, 428–446.
[10] Ho, H.C. and Hsing, T., 1997, Limit theorems for functionals of moving averages. *The Annals of Probability*, **25**, 1636–1669.
[11] Beran, J., 1994, *Statistics for Long-Memory Processes* (New York: Chapman and Hall).
[12] Beran, J., 1992, A goodness-of-fit test for time series with long-range dependence. *Journal of the Royal Statistical Society Series B*, **54**, 749–760.

Appendix A

Proof of Theorem 2.1 By the definition of $\phi_{n,q}(t, \lambda_m)$, we have

$$\phi_{n,q}(t, \lambda_m) = n^{1/2} t(1-t)\widehat{F}_{[nt]}(\lambda_m)\left\{\log\left\{\frac{\widehat{F}^*_{n-[nt]}(\lambda_m q)}{\widehat{F}_{[nt]}(\lambda_m q)}\right\} - \log\left\{\frac{\widehat{F}^*_{n-[nt]}(\lambda_m)}{\widehat{F}_{[nt]}(\lambda_m)}\right\}\right\}.$$

Theorem 1 of Robinson [6] asserts that, for any $t \in (0, 1)$,

$$\widehat{F}_{[nt]}(\lambda_m)/F(\lambda_m) \xrightarrow{\mathrm{P}} 1, \quad \widehat{F}_{[nt]}(\lambda_m q)/F(\lambda_m q) \xrightarrow{\mathrm{P}} 1$$

as $n \to \infty$. Then using equation (2) and Taylor expansion, we obtain

$$\begin{aligned}\phi_{n,q}(t, \lambda_m) = n^{1/2} t(1-t)(1 + o_P(1))\{&\widehat{F}_{[nt]}(\lambda_m) - \widehat{F}^*_{n-[nt]}(\lambda_m) \\ &- q^{2H-2}(\widehat{F}_{[nt]}(\lambda_m q) - \widehat{F}^*_{n-[nt]}(\lambda_m q))\}\end{aligned} \tag{A1}$$

Theorem 1 and Proposition 2 of Robinson [6] imply that

$$\mathrm{E}(\widehat{F}_{[nt]}(\lambda_m)) \sim F(\lambda_m), \quad \mathrm{E}(\widehat{F}^*_{n-[nt]}(\lambda_m)) \sim F(\lambda_m)$$

for any $0 < t < 1$ as $n \to \infty$. Moreover, we note that equation (1) implies

$$a_k \sim L_1(k)k^{H-3/2} \quad \text{as } k \to \infty$$

for some L_1 slowly varying at infinity. And Lemma 7 of Robinson [6] yields that, for any $0 < \eta < 3/2 - 2H$, as $n \to \infty$,

$$|b_k(m)| = O(k^{\eta-1})$$

uniformly in k such that $k > 2/m$. Hence we have, as $n \to \infty$,

$$\begin{aligned} |c_k(m)| &= O(L_1(k)k^{H-3/2}) + O(L_1(k)k^{H-3/2+\eta}\beta(\eta, H-1/2)), \\ &= O(L_1(k)k^{H-3/2+\eta}) \quad \text{as } k \to \infty, \end{aligned} \tag{A2}$$

where $\beta(\eta, H-1/2) = \int_0^1 x^{\eta-1}(1-x)^{H-3/2}\,\mathrm{d}x$ is the beta-function.

Now the conditions of Theorem 1.2 of Horváth and Shao [3] are satisfied, then we can define a sequence of Brownian bridges $\{B_n(t), 0 \le t \le 1\}$ such that, as $n \to \infty$,

$$\sup_{0<t<1} |n^{1/2}t(1-t)(\widehat{F}_{[nt]}(\lambda_m) - \widehat{F}^*_{n-[nt]}(\lambda_m)) - \sigma_m B_n(t)| = O_P(n^{-1/2}) \tag{A3}$$

and

$$\sup_{0<t<1} |n^{1/2}t(1-t)(\widehat{F}_{[nt]}(\lambda_m q) - \widehat{F}^*_{n-[nt]}(\lambda_m q)) - \sigma_{mq} B_n(t)| = O_P(n^{-1/2}). \tag{A4}$$

On the other hand, it is easy to check that, when $0 < [mq] \le [m/2]$, for every $k = 0, 1, 2, \ldots,$

$$\begin{aligned} b_k(m) &= \{1 + \exp(ik2\pi[mq]/n) + \cdots + \exp(ik2\pi[mq](m/[mq]-1))\}b_k(mq) \\ &\sim q^{-1}b_k(mq) \end{aligned}$$

as $n \to \infty$. In a similar way, when $m > [mq] > [m/2]$,

$$\begin{aligned} b_k(m) &\sim \frac{m}{[mq]-[m/2]}(b_k(mq) - b_k(m/2)) \\ &\sim q^{-1}b_k(mq) \end{aligned}$$

Then we get

$$\sigma_m^2 \sim q^{-2}\sigma_{mq}^2$$

as $n \to \infty$. This, together with (A1), (A3) and (A4), implies that

$$\sup_{0<t<1} |\phi_{n,q}(t, \lambda_m) - \sigma_m(1 - q^{2H-1})B_n(t)| = O_P(n^{-1/2}).$$

It follows from equation (A2) that the limit of σ_m^2 exists as $m \to \infty$, then let $\sigma^2 = \lim_{m\to\infty}, \sigma_m^2$ we get

$$\sup_{0<t<1} |\phi_{n,q}(t, \lambda_m) - \sigma(1 - q^{2H-1})B_n(t)| = o_P(1).$$

This completes the proof. ■

Proof of Theorem 2.2 Let

$$F_i(\lambda) = \int_0^{\lambda} f_i(t)\,\mathrm{d}t, \quad i = 1, 2.$$

By Theorem 1 of Robinson [6], we arrive at

$$nt(1-t)\widehat{F}_{[nt]}(\lambda_m) = nt(1-t)F_1(\lambda_m) + o_P(nt(1-t)F_1(\lambda_m))$$

and

$$\begin{aligned} nt(1-t)\widehat{F}^*_{n-[nt]}(\lambda_m) &= nt(\theta - t)F_1(\lambda_m) + nt(1-\theta)F_2(\lambda_m) \\ &\quad + o_P(nt(\theta - t)F_1(\lambda_m)) + o_P(nt(1-\theta)F_2(\lambda_m)) \end{aligned}$$

for any $0 < t \le \theta$ as $n \to \infty$.

Thus, if $1 \le k \le k^*$, by equation (A1), we have

$$\begin{aligned} n^{1/2}\phi_{n,q}(k/n, \lambda_m) &= \frac{k}{n}(n-k^*)\{F_1(\lambda_m) - F_2(\lambda_m) - q^{2H-2}\{F_1(\lambda_m q) - F_2(\lambda_m q)\}\} \\ &\quad + o_P\left(\frac{k}{n}(n-k^*)\{F_1(\lambda_m) - F_2(\lambda_m) - q^{2H-2}\{F_1(\lambda_m q) - F_2(\lambda_m q)\}\}\right). \end{aligned}$$

This implies that, as $n \to \infty$,

$$\max_{1\le k\le k^*} \frac{n^{-1/2}}{F_1(\lambda_m) - F_2(\lambda_m) - q^{2H-2}\{F_1(\lambda_m q) - F_2(\lambda_m q)\}}\phi_{n,q}(k/n, \lambda_m) \xrightarrow{\mathrm{P}} \theta(1-\theta). \quad \text{(A5)}$$

Let $0 < \varepsilon < \theta$, similar arguments give

$$\begin{aligned} &\max_{1\le k\le k^*-n\varepsilon} \frac{n^{-1/2}}{F_1(\lambda_m) - F_2(\lambda_m) - q^{2H-2}\{F_1(\lambda_m q) - F_2(\lambda_m q)\}}\phi_{n,q}(k/n, \lambda_m) \\ &\xrightarrow{\mathrm{P}} (\theta - \varepsilon)(1-\theta). \end{aligned} \quad \text{(A6)}$$

Equations (A5) and (A6) yield that

$$\lim_{n\to\infty} P\{\hat{k}_{mq} < k^* - n\varepsilon\} = 0$$

for all $\varepsilon > 0$. In a similar fashion, we obtain,

$$\lim_{n\to\infty} P\{\hat{k}_{mq} > k^* + n\varepsilon\} = 0$$

for all $\varepsilon > 0$. This implies the result of Theorem 2.2. ■

Copyright of Journal of Statistical Computation & Simulation is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.