Change-point Problems
IMS Lecture Notes – Monograph Series (Volume 23, 1994)

THE ANALYSIS OF CHANGE-POINT DATA WITH DEPENDENT ERRORS

BY F. LOMBARD AND J. D. HART
Rand Afrikaans University and Texas A&M University

We consider abrupt mean-change models for data with dependent, stationary, errors. No specific distributional assumptions, other than the existence and summability of cumulants, are made. A consistency property of the least squares estimator of the change-point is derived. This leads to the construction of consistent, asymptotically normal and efficient estimators of the error spectral density function and covariances. The application of these results in testing for the existence of a change is discussed. A test for uncorrelatedness of the errors is also given. An application is made to the detection of changes in the period of a variable star. The relationship between cusum charts used in statistics and O-C diagrams used in astronomy is pointed out.

1. Introduction. In the last decade the statistical literature has seen a steady increase in the number of published results dealing with so-called change-point problems. In their simplest form these are problems concerning the detection of sudden, often discontinuous, changes in the mean of a series of observations $X_1, \ldots, X_n$. To be more specific, a model of the form

$$X_t = f(t/n) + \epsilon_t; \quad t = 1, \ldots, n \tag{1.1}$$

is assumed to hold and the question concerns the possibility that f suffers jumps or sudden changes of slope, for example, at one or more of the points $t/n; \quad t = 1, \ldots, n$. A parametric form is often assumed for f, as in the simple abrupt (– or step) change model

$$f(t/n) = \mu + \Delta I(m < t \leq n) \tag{1.2}$$

with m an integer, $1 \leq m \leq n-1$, and I denoting the indicator function. The error series $\{\epsilon_t\}$ in (1.1) is usually assumed to consist of independent and identically distributed (i.i.d.) random variables with zero means and finite variances.

AMS 1991 Subject Classification: Primary 62G20; Secondary 62M10, 62M15

Key words and phrases: Change-point, dependent data, estimation of spectral density, O-C diagram, tests for dependence.

The primary purpose of this paper is to consider certain aspects of the problem when the $\{\epsilon_t\}$ are a stationary, not necessarily independent, sequence. In particular, we consider the estimation of the change-point m in (1.2) and of some characteristics of the error series $\{\epsilon_t\}$ in (1.1). The main results are formulated and discussed in Section 2. In particular, a consistency property of the least squares estimator (l.s.e.) of m is established. Estimation of the spectral density function (s.d.f.) of the error series is considered together with a test for uncorrelatedness of the errors. The role of these results in the construction of tests of the null hypothesis $H_0 : \Delta = 0$ is discussed. In Section 3 the results are illustrated in an analysis of some data from the field of variable star astronomy. A brief discussion is also given of the "O-C diagram" used by astronomers and its relationship to the cusum chart is pointed out. Proofs of the main results are given in Section 4.

Our proofs require the summability of cumulants of various orders. Cumulants are defined in Section 2.3 of Brillinger (1975) to whom we refer the reader for a survey of relevant results and for further details. The joint cumulant of random variables $Y_1, \ldots, Y_r$ is denoted by $\mathrm{cum}(Y_1, \ldots, Y_r)$ and if $Y_1 = \cdots = Y_r = Y$ by $\mathrm{cum}_r(Y)$. We have $\mathrm{cum}_2(Y) = \mathrm{var}(Y)$ and $\mathrm{cum}(Y_1, Y_2) = \mathrm{cov}(Y_1, Y_2)$. The error series $\{\epsilon_t\}$ is called weakly stationary to order $k(\geq 1)$ if

$$\mathrm{cum}(\epsilon_{t_1} \ldots, \epsilon_{t_r}) = \mathrm{cum}(\epsilon_0, \epsilon_{t_2 - t_1}, \ldots, \epsilon_{t_r - t_1})$$

for all $-\infty < t_1, \ldots, t_r < \infty$ and $1 \leq r \leq k$. In this case we shall write

$$\mathrm{cum}(\epsilon_0, \epsilon_{u_1}, \ldots, \epsilon_{u_r}) = c_\epsilon(u_1, \ldots, u_r).$$

2. Statement of Main Results.

2.1. Point Estimation of the Change-Point. For every fixed m, $1 \leq m \leq n-1$, the l.s.e.'s of μ and $\mu + \Delta$ in the model (1.2) are

$$\bar{X}_{m,1} = m^{-1} \sum_{t=1}^{m} X_t \quad \text{and} \quad \bar{X}_{m,2} = (n-m)^{-1} \sum_{t=m+1}^{n} X_t,$$

so that $\widehat{m}$, the l.s.e. of m, is the minimizer over k, $1 \leq k \leq n-1$, of

$$SS_W(k) = \sum_{t=1}^{k} (X_t - \bar{X}_{k,1})^2 + \sum_{t=k+1}^{n} (X_t - \bar{X}_{k,2})^2.$$

The well-known ANOVA decomposition

$$\sum_{t=1}^{n} (X_t - \bar{X})^2 = SS_W(k) + SS_B(k)$$

where $\bar{X} = n^{-1} \sum_{t=1}^{n} X_t$ and

$$SS_B(k) = k(\bar{X}_{k,1} - \bar{X})^2 + (n-k)(\bar{X}_{k,2} - \bar{X})^2,$$

shows that $\widehat{m}$ is the maximizer of $SS_B(k)$. Simple algebra yields

$$SS_B(k) = n(\bar{X}_{k,1} - \bar{X})^2 k/(n-k) = \left\{ n^{\frac{1}{2}} \sum_{t=1}^{k} (X_t - \bar{X}) / \left(k(n-k)\right)^{\frac{1}{2}} \right\}^2 . \quad (2.1)$$

A plot against k of the quantity in curly brackets at the extreme right in (2.1) is commonly referred to as a standardized cusum plot.

We assume that m varies with n in such a way that the ratio m/n remains bounded away from 0 and 1 for all n.

PROPOSITION 1. *Assume that the error process $\{\epsilon_t\}$ is weakly stationary to order 4. Then*

$$\widehat{m} = m + O_p(1) \quad \text{as } n \to \infty. \qquad \blacksquare$$

This is possibly the strongest form of "consistency" one can hope for. From a theoretical point of view one important consequence of Proposition 1 is that in asymptotic inference regarding the other parameters in the model (1.1), (1.2) one can effectively proceed as if m were known. This is illustrated by Propositions 2 and 3 below.

It should be noted that $\widehat{m}$ can be rather sensitive to outliers occurring among observations with indices close to 1 or n. The example in Section 3 illustrates this well. Thus, it will often be sensible to trim the data series by omitting the first and/or last few observations.

Proposition 1 is not true when $\Delta = 0$. In this case it can be shown that

$$\widehat{m}/n \to V \quad \text{in distribution where} \quad Pr[V=0] = Pr[V=1] = 1/2. \quad (2.2)$$

2.2. Estimation of Dependence Characteristics. In the case $\Delta = 0$ it is well known how to construct consistent and asymptotically normal estimates of $S_\epsilon(\lambda)$ and $c_\epsilon(\ell)$, the sdf and covariance at lag ℓ of the error series; see e.g. Brillinger (1975, Theorems 5.6.3 and 5.10.1).

Proposition 1 suggests that ϵ_t will be closely approximated by

$$e_t = X_t - \widehat{X}_t,$$

where

$$\widehat{X}_t = \bar{X}_{\widehat{m},1} + (\bar{X}_{\widehat{m},2} - \bar{X}_{\widehat{m},1}) I(\widehat{m} < t \leq n).$$

It is therefore natural to expect consistent and asymptotically normal estimates of $S_\epsilon(\lambda)$ and $c_\epsilon(\ell)$ to result upon replacing ϵ_t by e_t in the aforementioned estimation formulas. Proposition 2 justifies this procedure. To state the result we define the periodogram

$$I_\epsilon(\lambda) = n^{-1}\left|\sum_{t=1}^{n} \epsilon_t \exp(-i\lambda t)\right|^2 \tag{2.3}$$

and estimated covariances

$$\widehat{c}_\epsilon(\ell) = n^{-1}\sum_{t=1}^{n-\ell} \epsilon_t\epsilon_{t+\ell} \quad ; \quad \ell \geq 0 \tag{2.4}$$

of the error series. An estimate of the sdf $S_\epsilon(\lambda)$ is defined in terms of a bounded weight function $W(x)$ which is even, vanishes for $|x| > \pi$ and satisfies $\int_{-\pi}^{\pi} W(x)dx = 2\pi$, and in terms of a sequence, $\{b_n\}$, of bandwidths which satisfy $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$. The estimate is

$$\widehat{S}_\epsilon(\lambda) = n^{-1}\sum_{j=0}^{n-1} W^{(n)}(\lambda - w_j) I_\epsilon(w_j) \tag{2.5}$$

where $w_j = 2\pi j/n$ and $W^{(n)}(x)$ is the periodic extension of the function $b_n^{-1}W(x/b_n)$ to $(-\infty, \infty)$. The quantities $I_e(\lambda), c_e(\ell)$ and $\widehat{S}_e(\lambda)$ are defined by replacing ϵ by e in (2.3), (2.4) and (2.5).

PROPOSITION 2. *Suppose the error series $\{\epsilon_t\}$ is strictly stationary with cumulants satisfying*

$$\sum_{u_1,\dots,u_k} |c_\epsilon(u_1,\dots,u_k)| < \infty \quad \textit{for all } k \geq 1.$$

Then

$$(i) \qquad (nb_n)^{\frac{1}{2}}(\widehat{S}_e(\lambda) - \widehat{S}_\epsilon(\lambda)) = o_p(1) \quad \textit{uniformly in} \quad \lambda$$

and

$$(ii) \qquad n^{\frac{1}{2}}(\widehat{c}_e(\ell) - \widehat{c}_\epsilon(\ell)) = o_p(1) \ \ \textit{uniformly in} \ \ \ell \qquad \blacksquare$$

The proof is given in Section 4.2.

2.3. Hypothesis Tests. There exists a rather staggering variety of large-sample tests of the hypotheses $H_0 : \Delta = 0$ in the case of independent error series $\{\epsilon_t\}$. Unfortunately, an up-to-date review is not available. We therefore refer the reader to Brillinger (1989), Csörgo and Horváth (1988), Lombard (1987) and MacNeill (1974) for a sampling of the tests in question. One feature these tests have in common is that they are based on functions T_n of the cusum process $(k/n, S_k)$, where

$$S_k = n^{-\frac{1}{2}} \sum_{t=1}^{k} (X_t - \bar{X}) \quad \text{and} \quad \bar{X} = n^{-1} \sum_{t=1}^{n} X_t,$$

and reject H_0 whenever $T_n/\widehat{\sigma}$ is too large. Here $\widehat{\sigma}^2$ is a consistent estimator of $\sigma^2 = \text{var}(\epsilon_t)$. The asymptotic theory often relies on the fact that the scaled process $(k/n, \sigma^{-1}S_k)$ approximates to a Brownian bridge B. Now this approximation also holds for dependent error processes satisfying our moment conditions provided σ^2 is replaced by $S_\epsilon(0)$. It follows that the aforementioned tests remain valid for these error processes if $\widehat{\sigma}$ is replaced by $\widehat{S}_\epsilon(0)^{\frac{1}{2}}$ in the expression for the test statistic.

One well known statistic is

$$T_n = \max_{1 \le k < n} |S_k|. \tag{2.6}$$

Here $T_n/\widehat{S}_\epsilon(0)^{\frac{1}{2}} \longrightarrow \sup_{0 \le s \le 1} |B(s)|$ in distribution under $H_0 : \Delta = 0$ and

$$P(\sup_{0 \le s \le 1} |B(s)| > u) \approx 2\exp(-2u^2) \quad \text{for large } u. \tag{2.7}$$

A set of tests with some attractive features are those based on Fourier analysis ideas (Eubank and Hart (1992), Lombard (1988)). With independent data one such test, proposed by Eubank and Hart (1992), rejects $H_0 : \Delta = 0$ when

$$\max_{1 \le k \le n} \frac{1}{k} \sum_{j=1}^{k} n\widehat{\phi}_j^2/\widehat{\sigma}^2$$

is large, where $\widehat{\phi}_j = n^{-1} \sum_{r=1}^{n} X_r \cos(\pi j r/n), \quad j = 1, \ldots, n$. Eubank and Hart (1992) provide evidence that their test is unusually powerful in detecting multiple change-points.

2.4. Detecting Autocorrelation in the Error Process. One can use Proposition 2(ii) together with known results regarding the asymptotic distribution of $\widehat{c}_\epsilon(\ell)$ to test the hypothesis $H_0 : c_\epsilon(\ell) = 0$ for each individual $\ell \ge 1$ or, for any given k, to test the joint hypotheses $H_0 : c_\epsilon(1) = \cdots = c_\epsilon(k) = 0$. We close this section by considering global tests of uncorrelatedness. The basis of these tests is the fact that the errors are uncorrelated if and only if their sdf $S_\epsilon(\lambda)$ is constant. Also, it follows from theorem 4.3.1 of Brillinger (1975), under the conditions of Proposition 2 above, that

$$\max_{0 \le \omega_j \le \pi} |E(I_\epsilon(w_j)) - S_\epsilon(w_j)| = o(1)$$

as $n \to \infty$. Asymptotically then, testing uncorrelatedness is equivalent to testing constancy of the means of the periodogram ordinates. In Section 2.3 it was pointed out that many such tests of constancy are based on a cusum process. If we use I_e as surrogate for I_ϵ the relevant cusum process is

$$S_n^*(\lambda) = \left(\frac{2}{n}\right)^{\frac{1}{2}} \sum_{0 \le w_j \le \pi\lambda} (I_e(w_j) - \bar{I}_e), \quad 0 \le \lambda \le 1$$

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

$$\bar{I}_e = \frac{2}{n} \sum_{0 \le w_j \le \pi} I_e(w_j).$$

The requirement that this process approximates to a Brownian bridge is also fulfilled.

PROPOSITION 3. *If the error series $\{\epsilon_t\}$ is stationary and uncorrelated with $var(\epsilon_t) = \sigma^2$, then $(S_{n,e}^*(\lambda)/\bar{I}_e, \quad 0 \le \lambda \le 1)$ converges in distribution to a standard Brownian bridge (in the uniform topology of the space $D[0,1]$ – see Billingsley (1968, page 150)).* ▮

The proof is given in Section 4.3.

3. An Example. The circles in Figure 1 are a plot of the $n = 172$ times (in days) between 173 successive maxima on the light curve of the variable star T Centaurus. A question of interest to astronomers is whether the mean times between maxima are changing in a systematic fashion. The standardized cusum chart is shown in Figure 2. The sensitivity to the "outlier" at the extreme right in Figure 1 is clear from Figure 2 – strictly according to definition we would have $\widehat{m} = n = 172$ for the l.s.e. of m in a simple step-change model. If the last observation is omitted, $\widehat{m} = 94$ results! The solid lines in Figure 1 represent the fitted model: $\widehat{\mu} = 90.2$, $\widehat{\Delta} = 1.6$, $\widehat{m} = 94$. The periodogram of the residuals together with an estimate using the simple "boxcar" weight function $W(x) = 1$ for $|x| \le 1$; $= 0$ otherwise, and bandwidth $b = 4/172$ is shown in Figure 3. This yields $\widehat{S}_e(0) = 8.3$. We also used bandwidths $b = 8/172$, $b = 2/172$ and obtained $\widehat{S}_e(0) = 5.7$, $\widehat{S}_e(0) = 10.52$ respectively.

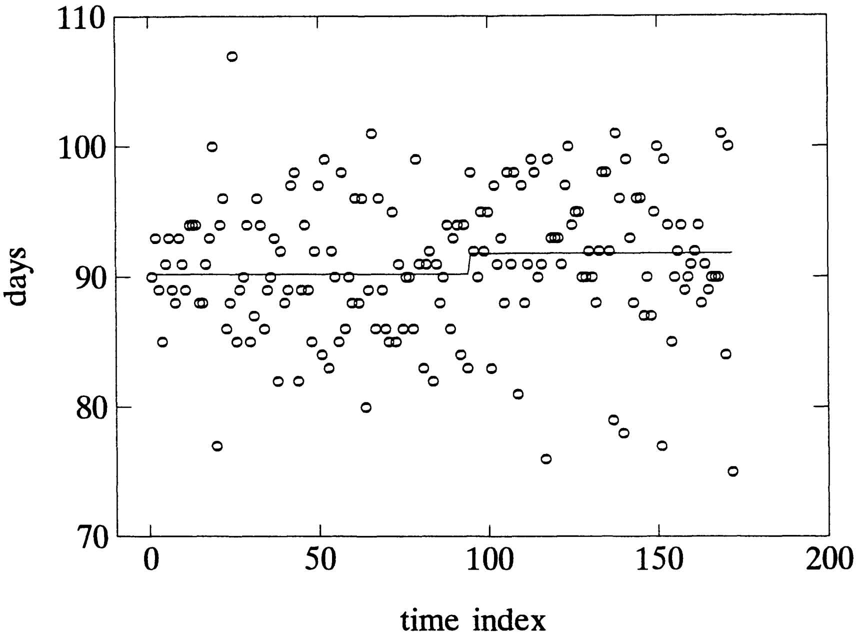

Figure 1: Times (in days) between light curve maxima of the variable star T CENTAURUS.

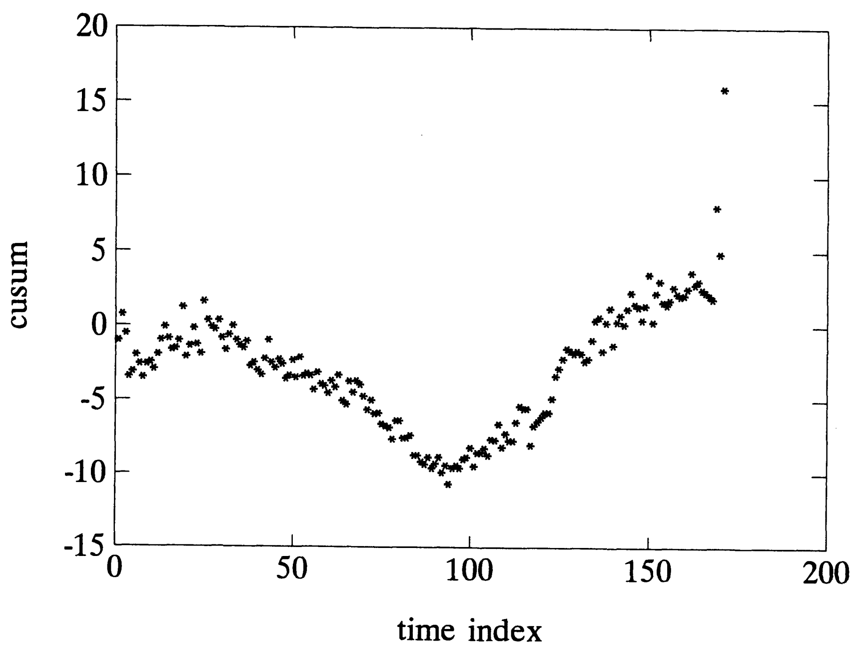

Figure 2: Standardized cusum of T CENTAURUS data.

It is pretty clear that constancy of the sdf is not a tenable hypothesis in this case, notwithstanding the large variability of the periodogram ordinates. For the data in question we get

$$\sup_{0 \leq \lambda \leq 1} |S^*_{n,e}(\lambda)|/\bar{I}_e = 2.713$$

with an asymptotic P-value of 0.0006 – see Proposition 3 and (2.7). The estimated autocorrelations at lags $1, \ldots, 5$ are $-0.38, -0.09, 0.09, 0.01$ and -0.07. To test $H_0 : \Delta = 0$ we use the T_n in (2.6) and $\hat{S}_e(0) = 10.52$. The test statistic is $T_n/\widehat{S}_e(0)^{\frac{1}{2}} = 1.63$ with an asymptotic P-value of 0.01.

Were the dependence in the data to be ignored, a startlingly different P-value results. For then one would use $n^{-1}SS_W(\widehat{m}) = 29.1$, the variance of the residuals, as scale factor in place of $\widehat{S}_e(0)$. Then $T_n/\widehat{S}_e(0)^{\frac{1}{2}} = 0.98$, and a P-value of 0.29 results. This type of phenomenon is, of course, well-known in statistics.

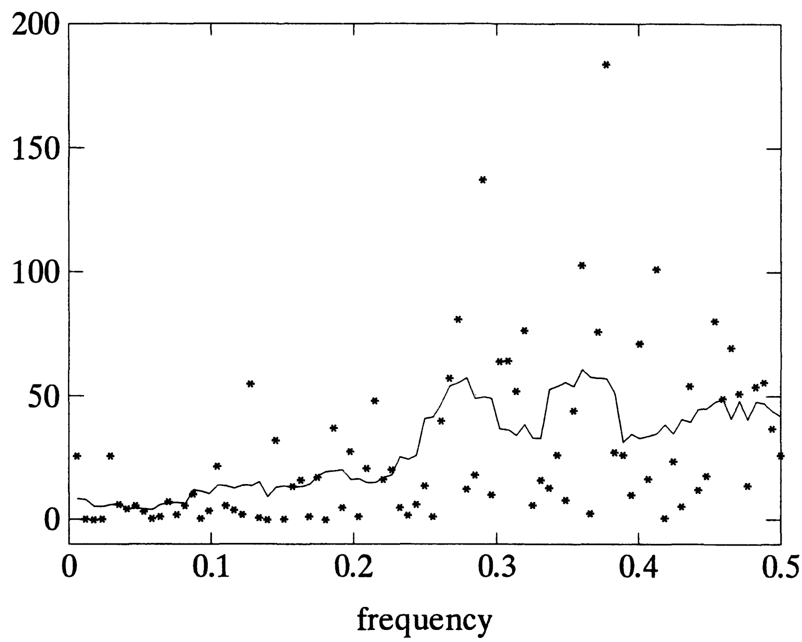

Figure 3: Raw (stars) and smoothed (solid line) periodogram of estimated residuals for T CENTAURUS data

We close this section by bringing to the reader's attention the fact that astronomers routinely make use of a device, known as an O-C chart, which is nothing but the cusum chart which is so well known to statisticians. As such, the use of cumulative sum charts by astronomers probably predates their use by statisticians. Sterne and Campbell (1937) give a brief review of

O-C methodology and refer to Sterne (1934) who gives a statistically sophisticated discussion of certain pitfalls in its application. Here we give only a brief explanation of the method, leaving it to the interested reader to pursue the matter further.

Let $Y_0 < Y_1 < \ldots < Y_n$ denote successive times of maxima observed on the light curve of a variable star. The times *between* maxima are then $X_t = Y_t - Y_{t-1}, 1 \leq t \leq n$. If the mean period is constant, the Y's should fluctuate around a line of slope β, the unknown period. A crude estimate,

$$\tilde{\beta} = (Y_n - Y_0)/n,$$

of β is used to obtain estimated values under the no-change assumption:

$$\widehat{Y}_i = Y_0 + \tilde{\beta} i \quad ; \quad 1 \leq i \leq n.$$

The differences, $Y_i - \widehat{Y}_i$, between the **O**bserved and **C**alculated values are plotted against i to give the O-C chart. Simple calculations now reveal that

$$Y_i - \widehat{Y}_i = \sum_{t=1}^{i} (X_t - \bar{X}), \tag{3.1}$$

the cusums of the residual times *between* maxima. Unfortunately, astronomers generally seem to have treated the residuals $Y_i - \widehat{Y}_i$ as independent random variables, a rather dubious assumption in view of the representation (3.1). This has led to a number of identifications of period changes which are unlikely to be justifiable by generally accepted statistical criteria. One such case is discussed by Lombard and Koen (1992).

4. Proofs.

4.1 Proof of Proposition 1. We need to show that

$$\lim_{r \to \infty} \limsup_{n \to \infty} P[|\widehat{m} - m| > r] = 0.$$

For this it is sufficient to show that

$$\lim_{r \to \infty} \limsup_{n \to \infty} P[\widehat{m} < m - r] = 0 \tag{4.1}$$

since the result involving the opposite inequality $\widehat{m} > m + r$ will follow from (4.1) by simply reversing the direction on the time axis. Define $\bar{X}_k = \bar{X}_{k,1}$ for each k, let $\alpha_n = \Delta(1 - m/n)$ and observe that

$$\bar{X} - \mu - \alpha_n = \bar{\epsilon}$$

while, for $1 \le k \le m$,

$$\bar{X}_k - \mu = \bar{\epsilon}_k.$$

Observe also that $\widehat{m} < m - r$ implies $SS_B(m) \le SS_B(k)$ for at least one $k \in \{1, \ldots, m-r\}$. Hence

$$P[\widehat{m} < m - r] \le \sum_{k=1}^{m-r} P[SS_B(m) \le SS_B(k)]. \tag{4.2}$$

Furthermore, from the definitions, $SS_B(m) \le SS_B(k)$ if and only if,

$$(\bar{X}_m - \bar{X})^2 m/(n-m) \le (\bar{X}_k - \bar{X})^2 k/(n-k),$$

that is, if and only if,

$$(\bar{\epsilon}_m - \bar{\epsilon} - \alpha_n)^2 m/(n-m) \le (\bar{\epsilon}_k - \bar{\epsilon} - \alpha_n)^2 k/(n-k).$$

After some algebraic manipulation, the last inequality is seen to hold if and only if,

$$\begin{aligned}
\alpha_n^2 \le\ & (\bar{\epsilon}_k - \bar{\epsilon})^2 \cdot \{k(n-m)/n(m-k)\} - (\bar{\epsilon}_m - \bar{\epsilon})^2 \cdot \{m(n-k)/n(m-k)\} \\
& - (\bar{\epsilon}_k - \bar{\epsilon}) \\
& \quad \cdot \{2k(n-m)/n(m-k)\} \cdot \alpha_n + (\bar{\epsilon}_m - \bar{\epsilon}) \cdot \{2m(n-k)/n(m-k)\} \cdot \alpha_n \\
=\ & (\bar{\epsilon}_k - \bar{\epsilon})^2 (k/(m-k))(\alpha_n/\Delta) - (\bar{\epsilon}_m - \bar{\epsilon})^2((n-k)/(m-k))(1 - \alpha_n/\Delta) \\
& + (m-k)^{-1} \sum_{t=k+1}^{m} (\epsilon_t - \bar{\epsilon}) \cdot 2\alpha_n^2/\Delta + (\bar{\epsilon}_m - \bar{\epsilon}) \cdot 2\alpha_n(1 - \alpha_n/\Delta) \\
=\ & \sum_{i=1}^{4} Z_{i,k}, \quad \text{say.}
\end{aligned}$$

Now

$$\bigcap_{i=1}^{4} [|Z_{i,k}| < \alpha_n^2/4] \subseteq \left[\sum_{i=1}^{4} Z_{i,k} < \alpha_n^2\right] = [SS_B(m) > SS_B(k)],$$

so that, from (4.2),

$$P[\widehat{m} < m - r] \le \sum_{i=1}^{4} \sum_{k=1}^{m-r} P[|Z_{i,k}| \ge \alpha_n^2/4]. \tag{4.3}$$

The following Lemma pertaining to the moments of $Z_{i,k}$ is required in order to bound the probabilities on the right hand side of (4.3).

LEMMA 4.1. *Under the conditions of Proposition 1, $E(\bar{\epsilon}_k - \bar{\epsilon})^4 \le Ck^{-2}$ for some constant C not depending on k or n.*

PROOF. Set $a_{s,k} = I(s \le k) - k/n$ and note that $|a_{s,k}| \le 1$ for all s and $1 \le k \le n$. We have

$$\bar{\epsilon}_k - \bar{\epsilon} = k^{-1} \sum_{s=1}^{n} \epsilon_s a_{s,k},$$

whence

$$\begin{aligned} |\mathrm{cum}_4(\bar{\epsilon}_k - \bar{\epsilon})| &= k^{-4} \left| \sum_{s_1,\ldots,s_4=1}^{n} \mathrm{cum}(\epsilon_{s_1}, \epsilon_{s_2}, \epsilon_{s_3}, \epsilon_{s_4}) \prod_{i=1}^{4} a_{s_i,k} \right| \\ &= k^{-4} \left| \sum_{|u_1|,|u_2|,|u_3| \le n-1} c_\epsilon(u_1, u_2, u_3) \sum_{s_1=1}^{k} a_{s_1,k} \prod_{i=1}^{3} a_{u_i+s_1,k} \right| \\ &\le k^{-3} \sum_{|u_1|,|u_2|,|u_3|<\infty} |c_\epsilon(u_1, u_2, u_3)| = Ck^{-3} \end{aligned}$$

and, similarly,

$$|\mathrm{cum}_2(\bar{\epsilon}_k - \bar{\epsilon})| \le Ck^{-1}.$$

The Lemma now follows upon substitution of these inequalities into the well known formula

$$0 \le E(\bar{\epsilon}_k - \bar{\epsilon})^4 = \mathrm{cum}_4(\bar{\epsilon}_k - \bar{\epsilon}) + 3(\mathrm{cum}_2(\bar{\epsilon}_k - \bar{\epsilon}))^2. \quad \blacksquare$$

Let $0 < \delta < 1$ be given. With the help of Lemma 4.1 we will now show for $i = 1, \ldots, 4$ that

$$\limsup_{n\to\infty} \sum_{k=1}^{m-r} P[|Z_{i,k}| > \alpha_n^2/4] < \delta$$

for all sufficiently large r. Below, C denotes a generic finite, positive, constant. We have

$$\begin{aligned} \sum_{k=1}^{m-r} P[|Z_{1,k}| \ge \alpha_n^2/4] &\le C \sum_{k=1}^{m-r} E(Z_{1,k}^2) \\ &\le C \sum_{k=1}^{m-r} E\{(\bar{\epsilon}_k - \bar{\epsilon})^4\} \cdot k^2/(m-k)^2 \\ &\le C \sum_{k=1}^{m-r} (m-k)^{-2} \le C \sum_{j=r}^{\infty} j^{-2} < \delta \end{aligned}$$

for all sufficiently large r. The term involving $|Z_{2,k}|$ is handled in exactly the same way. Next

$$\sum_{k=1}^{m-r} P[|Z_{3,k}| \geq \alpha_n^2/4] \leq C \sum_{k=1}^{m-r} E(Z_{3,k}^4) \leq C \sum_{k=1}^{m-r} (m-k)^{-2}$$

and the result follows as in the case of $|Z_{1,k}|$. Finally

$$\sum_{k=1}^{m-r} P[|Z_{4,k}| \geq \alpha_n^2/4] \leq C \cdot (m-r) \cdot m^{-2} \leq Cm^{-1} \to 0 \text{ as } n \to \infty. \quad \blacksquare$$

PROOF OF (2.2). We give the proof for i.i.d. errors $\{\epsilon_t\}$. Set

$$Z_k = \sum_{t=1}^{k} (X_t - \bar{X})$$

and observe that

$$Z_k = n^{-1} \sum_{t=1}^{k} \sum_{s=k+1}^{n} (X_t - X_s)$$

so that (1.1) of Csörgo and Horváth (1988) holds with the antisymmetric function $h(x,y) = x - y$. We can therefore apply the results in their Section 4. Careful scrutiny of the proof of their Theorem 4.3 (see also their Theorem 2.3) reveals that $P[A_n] + P[B_n] \to 1$ as $n \to \infty$ where

$A_n = [(\log n)^3 \leq \widehat{m} \leq n/(\log n)^2]$ and $B_n = [n - n/(\log n)^2 \leq \widehat{m} \leq n - (\log n)^3]$.

Reversing the direction on the time axis makes it clear that $P[A_n] = P[B_n]$; hence the result. $\blacksquare$

4.2. Proof of Proposition 2. The residuals are

$$e_t = \epsilon_t + r_t$$

where

$$\begin{aligned} r_t = & -\bar{\epsilon}_{1,\widehat{m}} - \Delta(1 - (m \wedge \widehat{m})/\widehat{m})) \cdot I(1 \leq t \leq m \wedge \widehat{m}) \\ & - \bar{\epsilon}_{2,\widehat{m}} - \Delta(\widehat{m} - (m \vee \widehat{m}))/(n - \widehat{m})) \cdot I(m \vee \widehat{m} < t \leq n) \\ & - \bar{\epsilon}_{1,\widehat{m}} - \Delta(n-m)/(n - \widehat{m}) \cdot I(\widehat{m} + 1 \leq t \leq m) \\ & - \bar{\epsilon}_{2,\widehat{m}} + \Delta m/\widehat{m} \cdot I(m + 1 \leq t \leq \widehat{m}), \end{aligned}$$

and we have

$$|I_e(w_j) - I_\epsilon(w_j)| \leq I_r(w_j) + 2\{I_\epsilon(w_j) \cdot I_r(w_j)\}^{\frac{1}{2}} \tag{4.4}$$

for $j = 1, \ldots, [n/2]$. The following lemma will be used.

LEMMA 4.2.

$$\max_{1 \le j \le [n/2]} I_r(w_j)/(j^{-\frac{3}{2}} + n^{-1}) = O_p(1).$$

PROOF. From Proposition 1 and a central limit theorem for stationary sequences (Brillinger (1975, page 94)) we see that both $\bar{\epsilon}_{1,\widehat{m}}$ and $\bar{\epsilon}_{2,\widehat{m}}$ are $O_p(n^{-\frac{1}{2}})$. Thus, we can write

$$r_t = \beta_{1n} I(1 \le t \le m \wedge \widehat{m}) + \beta_{2n} I(m \vee \widehat{m} < t \le n) + (\beta_{3n} + \Delta) I(m \wedge \widehat{m} < t \le m \vee \widehat{m})$$

where, for $i = 1, 2, 3$, $\beta_{in} = O_p(n^{-\frac{1}{2}})$ and does not involve t. Then,

$$\begin{aligned} \sum_{t=1}^{n} r_t \exp(-iw_j t) &= (\beta_{1n} - \beta_{2n}) \sum_{t=1}^{m \wedge \widehat{m}} \exp(-iw_j t) + O_p(1) \\ &= O_p(n^{-\frac{1}{2}}) \cdot \sum_{t=1}^{m \wedge \widehat{m}} \exp(-iw_j t) + O_p(1). \end{aligned}$$

Furthermore,

$$\left| \sum_{t=1}^{m \wedge \widehat{m}} \exp(-iw_j t) \right|^2 \le \sin^{-2}(w_j/2) \le (n/4j)^2$$

whence, using the inequality

$$I_{a+b}(w) \le I_a(w) + I_b(w) + 2(I_a(w) I_b(w))^{\frac{1}{2}},$$

we see that

$$\begin{aligned} I_r(w_j) &\le O_p(1) \cdot j^{-2} + O_p(n^{-1}) + O_p(n^{-\frac{1}{2}}) \cdot j^{-1} \\ &= O_p(1)(j^{-\frac{3}{2}} + n^{-1}) \end{aligned}$$

with the O_p-terms all independent of j. ∎

(i) Returning to (4.4), set $L = nb_n$ and let $\ell \equiv \ell_n \to \infty$ in such a way that $\ell/L \to 0$. We have, using the inequality $(a+b)^2 \le 2a^2 + 2b^2$, the Schwarz-inequality and Lemma 4.2,

$$
\begin{aligned}
&\left\{L^{-\frac{1}{2}}\sum_{j=1}^{L}(I_\epsilon(w_j)I_r(w_j))^{\frac{1}{2}}\right\}^2 \\
&\le 2L^{-1}\sum_{j=1}^{\ell} I_\epsilon(w_j)\cdot\sum_{j=1}^{\ell} I_r(w_j) + 2L^{-1}\sum_{j=\ell+1}^{L} I_\epsilon(w_j)\cdot\sum_{j=\ell+1}^{L} I_r(w_j) \\
&\le 2(\ell/L)\cdot\ell^{-1}\sum_{j=1}^{\ell} I_\epsilon(w_j)\cdot\left\{\sum_{j=1}^{\ell} j^{-\frac{3}{2}} + \ell/n\right\}\cdot O_p(1) \\
&+ 2L^{-1}\sum_{j=\ell+1}^{L} I_\epsilon(w_j)\cdot\left\{\sum_{j=\ell+1}^{L} j^{-\frac{3}{2}} + (L-\ell)/n\right\}\cdot O_p(1) \\
&= o(1)\cdot O_p(1)\cdot\{O(1)+o(1)\}\cdot O_p(1) + O_p(1)\{o(1)+o(1)\}\cdot O_p(1) = o_p(1).
\end{aligned}
$$

Also,

$$L^{-\frac{1}{2}}\sum_{j=1}^{L} I_r(w_j) \le L^{-\frac{1}{2}}\sum_{j=1}^{L}\{j^{-\frac{3}{2}} + n^{-1}\}\cdot O_p(1) = L^{-\frac{1}{2}}O_p(1) = o_p(1).$$

Thus

$$L^{\frac{1}{2}}\sum_{j=1}^{L}|I_e(w_j) - I_\epsilon(w_j)| = o_p(1).$$

This completes the proof since

$$L^{\frac{1}{2}}|\widehat{S}(0) - \widehat{S}_\epsilon(0)| \le L^{-\frac{1}{2}}\sum_{j=1}^{L}|I_e(w_j) - I_\epsilon(w_j)|. \qquad ∎$$

(ii) Using essentially the same technique as in (i), it can be shown that

$$n^{-\frac{1}{2}}\sum_{j=1}^{n}|I_e(w_j) - I_\epsilon(w_j)| = o_p(1). \tag{4.5}$$

The result is now a direct consequence of (4.5) and of the fact that

$$\widehat{c}_x(k) = \int_{-\pi}^{\pi} I_x(w)\exp(-2\pi ikw)\,dw$$

for $x = e, \epsilon$. ∎

4.3. Proof of Proposition 3. Set

$$Y_n(\lambda) = n^{-\frac{1}{2}} \sum_{0 \le w_j \le \lambda} (I_e(w_j) - \sigma^2) \qquad 0 \le \lambda \le \pi.$$

It is an easy consequence of (4.5) above and of (the univariate version of) Theorem 7.6.3 of Brillinger (1975) that

$$Y_n(\cdot) \to_{\mathcal{L}} Y(\cdot)$$

where the latter process is Gaussian with zero mean and covariance function

$$\mathrm{cov}(Y(\lambda), Y(\mu)) = \sigma^2 \min(\lambda, \mu) + c_4(\epsilon)\lambda\mu.$$

Since

$$n^{-\frac{1}{2}} \sum_{0 \le w_j \le \lambda} (I_e(w_j) - \bar{I}) = Y_n(\lambda) - Y_n(\pi),$$

the continuous mapping theorem implies that $Y_n(\cdot) - Y_n(\pi) \to_{\mathcal{L}} Y(\cdot) - Y(\pi)$. The latter process is Gaussian with zero mean and covariance function $\sigma^2 \cdot (\min(\lambda, \mu) - \lambda\mu)$. The proof is completed by observing that $\bar{I}_e$ is a consistent estimator of σ^2. ∎

Acknowledgements. Lombard's work was supported by a grant from the Foundation for Research Development. The data used in Section 3 were provided by the American Association of Variable Star Observers.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* New York: John Wiley & Sons.

BRILLINGER, D. (1975). *Time Series: Data Analysis and Theory.* New York: Holt, Rinehart and Winston, Inc.

BRILLINGER, D. (1989). Consistent detection of a monotonic trend superposed on a stationary time series. *Biometrika*, **76**, 23–30.

CSÖRGŐ, M. and HORVÁTH, L. (1988). Invariance principles for changepoint problems. *J. Mult. Anal.*, **27**, 151–168.

EUBANK, R. and HART, J. D. (1992). Testing goodness of fit in regression via order selection criteria. *Ann. Statist.*, **20**, 1412–1425.

LOMBARD F. (1987). Rank tests for change-point problems. *Biometrika*, **74**, 615–624.

LOMBARD, F. (1988). Detecting change-points by Fourier Analysis. *Technometrics*, **30**, 305–310.

LOMBARD, F. and KOEN, C. (1993). The analysis of indexed astronomical time series II: The O-C (observed-calculated) technique reconsidered. *Mon. Not. R. Astron. Soc.* **263**, 309–313.

MACNEILL, I. B. (1974). Test for change of parameter at unknown time and distributions of some related functionals on Brownian motion. *Ann. Statist.*, **2**, 950–962.

STERNE, T. E. and CAMPBELL, L. (1937). Changes of period in long period variables. *The Harvard Annals*, **105**, 459–489.

STERNE, T. E. (1934). The errors of period of variable stars. - Part I: The general theory, illustrated by RR Scorpii. Harvard College Observatory, Circular **386**, 1–36.

DEPARTMENT OF STATISTICS
RAND AFRIKAANS UNIVERSITY
JOHANNESBURG 2000, SOUTH AFRICA

DEPARTMENT OF STATISTICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 78743-3143