Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 451–464
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Change detection in autoregressive time series

Edit Gombay

Department of Mathematical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1

Received 5 April 2006
Available online 13 January 2007

Abstract

Autoregressive time series models of order p have $p+2$ parameters, the mean, the variance of the white noise and the p autoregressive parameters. Change in any of these over time is a sign of disturbance that is important to detect. The methods of this paper can test for change in any one of these $p+2$ parameters separately, or in any collection of them. They are available in forms that make one-sided tests possible, furthermore, they can be used to test for a temporary change. The test statistics are based on the efficient score vector. The large sample properties of the change-point estimator are also explored.

AMS 2000 subject classification: primary 62G20; secondary 60F17;62M10

Keywords: Time series; Efficient score vector; Strong approximation; Invariance; Principles; Brownian bridge

1. Introduction and results

Let observations $\{Y_i\}$ have structure

$$Y_i - \mu = \varphi_1(Y_{i-1} - \mu) + \ldots + \varphi_p(Y_{i-p} - \mu) + \varepsilon_i, \quad i \geqslant p+1 \tag{1.1}$$

with $\{\varepsilon_i\}$ i.i.d. white noise sequence, $E(\varepsilon_1) = \sigma^2$, and $\boldsymbol{\xi} = (\mu, \sigma^2, \varphi_1, \ldots, \varphi_p)^t$ the vector of parameters. Instability in the value of any component of $\boldsymbol{\xi}$ will lead to wrong forecasts and data analysis, so detecting change in them is a statistical problem of great importance. Once change has been detected, the time of change has to be estimated.

The available methods in the literature for detecting change in the parameters of the autoregressive time series model are mainly based on the estimated white noise sequence, the residuals, and the statistics process is a quadratic form created from them. (See [1,2,11,12,10,3], and references therein.) The use of quadratic forms does not allow the creation of one-sided tests, or

E-mail address: egombay@ualberta.ca.

doi:10.1016/j.jmva.2007.01.003

testing for change in a specific component of the $\xi = (\mu, \sigma^2, \varphi_1, \ldots, \varphi_p)^t$ vector. These limitations can be overcome if the efficient score vector is used as test statistic. Its structure is that of a partial sums process, which can be well approximated by a Brownian motion. This allows us to define one-sided-tests for any component of ξ, or the creation of tests for temporary change (the epidemic alternative). Davies et al. [1] use the generalized likelihood ratio, which is also leading to quadratic forms, so their method gives tests of change by which one-sided-tests cannot be done, and where we cannot separate the components of the parameter vector. So if the presence of change is indicated, the user has no information about which parameter has changed.

Gombay and Serban [8] used the efficient score vector in the sequential framework, when change from an initially given parameter value had to be detected, while all other components of ξ were nuisance parameters. The topic of the current paper is retrospective change detection where the full sequence $\{Y_1, \ldots, Y_n\}$ is available and no initial parameter value is specified but all parameters have to be estimated.

To be able to use the efficient score vector, we assume that the white noise $\{\varepsilon_i\}$ is Gaussian with identically distributed uncorrelated components. By the results of Gombay and Horváth [7] the assumption of normality can be replaced by appropriate moment conditions on the innovations. In such cases the likelihood function becomes quasi-likelihood, but the proposed tests do not lose their validity. For the sake of brevity and focus we do not give the details here.

Under the Gaussian assumption the components of the efficient score vector $\nabla_\xi \ell_k(Y_1, \ldots, Y_k; \xi) = \nabla_\xi \ell_k(\xi)$ are

$$\frac{\partial}{\partial\mu}\,\ell_k(\xi) = \frac{1-\sum_{j=1}^{p}\varphi_j}{\sigma^2}\sum_{i=1}^{k}\Big[Y_i - \mu - \sum_{j=1}^{p}\varphi_j(Y_{i-j}-\mu)\Big],$$

$$\frac{\partial}{\partial\sigma^2}\,\ell_k(\xi) = -\frac{k}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{k}\Big[Y_i - \mu - \sum_{j=1}^{p}\varphi_j(Y_{i-j}-\mu)\Big]^2,$$

$$\frac{\partial}{\partial\varphi_s}\,\ell_k(\xi) = \frac{1}{\sigma^2}\sum_{i=1}^{k}\Big[Y_i - \mu - \sum_{j=1}^{p}\varphi_j(Y_{i-j}-\mu)\Big](Y_{i-s}-\mu), \quad s = 1, \ldots, p,$$

where ℓ_k denotes the log-likelihood function based on observations $Y_1, \ldots, Y_k$. As our methods are based on large sample approximations, for $Y_0, Y_{-1}, \ldots, Y_{-p}$ we can substitute zero, or any other finite value. The $(p+2) \times (p+2)$ information matrix is

$$I(\xi) = I(\mu, \sigma^2, \varphi_1, \ldots, \varphi_p) = \begin{pmatrix} \frac{\left(1-\sum_{j=1}^{p}\varphi_j\right)^2}{\sigma^2} & 0 & 0 \\ 0 & \frac{1}{2\sigma^4} & 0 \\ 0 & 0 & \frac{1}{\sigma^2}\Gamma \end{pmatrix},$$

where Γ is the covariance matrix of vector $(Y_1, \ldots, Y_p)$.

Let $\widehat{\mu}_n, \widehat{\sigma}_n^2, \widehat{\boldsymbol{\varphi}}_n$ be the maximum likelihood estimators of the parameters, that is, the simultaneous solutions of equations $\dfrac{\partial}{\partial \xi_j}\,\ell_n(\boldsymbol{\xi}) = 0,\ j = 1, \ldots, p+2$, and let

$$\widehat{\mathbf{B}}(u) = n^{-1/2} I^{-1/2}(\widehat{\boldsymbol{\xi}}_n) \begin{pmatrix} \dfrac{\partial}{\partial \mu}\,\ell_{[nu]}(\widehat{\boldsymbol{\xi}}_n) \\ \dfrac{\partial}{\partial \sigma^2}\,\ell_{[nu]}(\widehat{\boldsymbol{\xi}}_n) \\ \nabla_{\varphi}\ell_{[nu]}(\widehat{\boldsymbol{\xi}}_n) \end{pmatrix}.$$

Theorem 1. *Assume the sequence of observations* $\{Y_i\}$ *satisfy model* (1.1) *with Gaussian* i.i.d. *white noise* $\{\varepsilon_i\}$, Var $(\varepsilon_i) = \sigma^2$, *and characteristic polynomial* $\varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p$ *with roots outside the unit circle. Then there exists a* $p+2$*-dimensional Gaussian process* $\mathbf{B}(u)$ *with independent Brownian bridge components* $B^{(j)}(u),\ j = 1, \ldots, p+2$, *such that*:

$$\max_{1 \leqslant j \leqslant p+2}\ \sup_{0 \leqslant u \leqslant 1} |\widehat{B}^{(j)}(u) - B^{(j)}(u)| = o_p(1).$$

It is an easy exercise to show that $\widehat{\mathbf{B}}(u)$ has uncorrelated components, from which we get the asymptotic independence of the components of the transformed efficient score vector. Hence we need to define the test component-wise only. For simultaneous test-for-change in d parameters, to have an overall level of significance α, we use $\alpha^* = 1 - (1-\alpha)^{1/d}$ for each component.

For a one-sided test we have

Test 1: If

$$\sup_{0 \leqslant u \leqslant 1} \widehat{B}^{(j)}(u) \geqslant C_1(\alpha^*),$$

then conclude, that there was a change in parameter ξ_j $(1 \leqslant j \leqslant p+2)$ along the sequence $Y_1, \ldots, Y_n$.

Critical value $C_1(\alpha^*)$ is obtained from relationship

$$P\Big(\sup_{0 \leqslant u \leqslant 1} B^{(1)}(u) \geqslant x\Big) = e^{-2x^2}.$$

For two sided-test we have

Test 2: If

$$\sup_{0 \leqslant u \leqslant 1} |\widehat{B}^{(j)}(u)| \geqslant C_2(\alpha^*),$$

then conclude that there was a change in parameter ξ_j $(1 \leqslant j \leqslant p+2)$ along the sequence $Y_1, \ldots, Y_n$.

Critical value $C_2(\alpha^*)$ is obtained from relationship

$$P\Big(\sup_{0 \leqslant u \leqslant 1} |B^{(1)}(u)| > x\Big) = \sum_{k \neq 0} (-1)^{k+1} e^{-2k^2 x^2}.$$

When change in the parameter is temporary only, and the initial values are restored after a period of disturbance (epidemic), then it is more efficient to use a test designed with this alternative in mind. For example, if the initial mean μ_1 changed to μ_2 $(> \mu_1)$, and then μ_1 was restored, then the standardized partial sums $\max_{1\leqslant i<j\leqslant n}\sum_{k=i}^{j} x_i$ of the observations is a more efficient test statistic, than the $\max_{1\leqslant i\leqslant n}\sum_{k=i}^{n} x_i$ version. For such temporary change in the jth component of $\boldsymbol{\xi}$ we have the following:

Test 3: If

$$\max_{0\leqslant u\leqslant 1} \widehat{B}^{(j)}(u) - \min_{0\leqslant u\leqslant 1} \widehat{B}^{(j)}(u) \geqslant C_3(\alpha^*),$$

then conclude that there was a change in parameter ξ_j $(1\leqslant j\leqslant p+2)$ along the sequence $Y_1,\ldots,Y_n$.

Critical value $C_3(\alpha^*)$ is obtained from the relation of Kuiper [9]

$$P\Big(\max_{0\leqslant u\leqslant 1} B^{(j)}(u) - \min_{0\leqslant u\leqslant 1} B^{(j)}(u)\leqslant x\Big) = 1-\sum_{i=1}^{\infty} 2(4i^2x^2-1)e^{-2i^2x^2}.$$

Remark. If $\boldsymbol{\varphi}\equiv\mathbf{0}$, then the observations are independent, and the process for detecting change in the mean is

$$\widehat{\sigma}_n^{-1}\Big(\sum_{i=1}^{k} Y_i - \frac{k}{n}\sum_{i=1}^{n} Y_i\Big), \quad 1<k\leqslant n,$$

for change in variance, the process is

$$2^{-1/2}\widehat{\sigma}_n^{-2}\sum_{i=1}^{k}[(Y_i-\overline{Y}_n)^2-\widehat{\sigma}_n^2], \quad 1<k\leqslant n.$$

In case independent observations $X_1, X_2,\ldots,X_n$ are from density $f(x;\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ is the vector of the unknown parameters, then the statistic testing for change in the jth component of $\boldsymbol{\xi}$ is the jth component of the vector process

$$n^{-1/2}I^{-1/2}(\widehat{\boldsymbol{\xi}}_n)\sum_{i=1}^{k}\nabla_{\boldsymbol{\xi}}\log\ f(X_i;\widehat{\boldsymbol{\xi}}_n), \quad 1<k\leqslant n.$$

The difference between this type of change detection, and the ones found in the literature is that only one parameter estimation, based on the full sequence $Y_1,\ldots,Y_n$, is required. Most algorithms (see e.g. [1]) calculate estimators of the before and after change parameter value at all possible change-point k, $1<k<n$. They are then either compared to see if their difference is significant at some k, or used in the likelihood function. The new algorithms above are computationally much more simple, and avoid estimations based on a few observations only.

2. Change-point estimation

We assume that there is a change in ξ_j, $1\leqslant j\leqslant p+2$, at $\tau=[\rho n]$, $0<\rho<1$. First note, that estimators $\hat{\mu}_n$, $\widehat{\sigma}_n^2$, and $\widehat{\boldsymbol{\varphi}}_n$, that are the solutions of the $p+2$ equations $\frac{\partial}{\partial\xi_j}\ \ell_n(\boldsymbol{\xi})=0$, $j=1,\ldots,p+2$, although not estimators of true parameters, are still finite values. The estimator of τ is

$$\widehat{\tau}_n=\min_k\left\{\frac{\partial}{\partial\xi_j}\ell_k(\widehat{\boldsymbol{\xi}}_n)=\max_{1<m\leqslant n}\frac{\partial}{\partial\xi_j}\ell_m(\widehat{\boldsymbol{\xi}}_n)\right\}$$

or

$$\widehat{\tau}_n = \min_n \left\{ \frac{\partial}{\partial \xi_j} \ell_k(\widehat{\boldsymbol{\xi}}_n) = \min_{1<m\leqslant n} \frac{\partial}{\partial \xi_j} \ell_m(\widehat{\boldsymbol{\xi}}_n) \right\}$$

for the two one-sided tests, and

$$\widehat{\tau}_n = \min_k \left\{ \left| \frac{\partial}{\partial \xi_j} \ell_k(\widehat{\boldsymbol{\xi}}_n) \right| = \max_{1<m\leqslant n} \left| \frac{\partial}{\partial \xi_j} \ell_m(\widehat{\boldsymbol{\xi}}_n) \right| \right\}$$

for the two-sided test.

(i) Assume that change is in the mean μ only.

In the Appendix it is shown that

$$|\widehat{\tau}_n - \tau| = O_p(1), \tag{2.1}$$

and $\widehat{\tau}_n$ is distributed approximately as the place of maximum/minimum of a two-sided random walk with mean increasing/decreasing up to a point and then decreasing/increasing. Such change-point estimators were first discussed by Hinkley [5].

(ii) Now consider the test for change in φ_s, a component of $\boldsymbol{\varphi}$, $s = 1, \ldots, p$.

As Cov $(Y_1, Y_{1+s}) = \gamma(s) = \varphi_1\gamma(s-1) + \varphi_2\gamma(s-2) + \cdots + \varphi_s\gamma(0) + \cdots + \varphi_p\gamma(s-p)$, and $\gamma(0) \geqslant \gamma(h)$, $h \geqslant 1$, if change is in φ_s only, then, although all $\gamma(h)$, $h \geqslant 1$, change, it is $\gamma(s)$ that changes to the largest degree:

$$\gamma^{(1)}(s) - \gamma^{(2)}(s) = \left(\boldsymbol{\varphi}^{(1)}\right)^t \left(\boldsymbol{\gamma}^{(1)} - \boldsymbol{\gamma}^{(2)}\right) + \left(\varphi_s^{(1)} - \varphi_s^{(2)}\right)\gamma^{(2)}(0),$$

where $\boldsymbol{\varphi}^{(1)}$ is the vector $\boldsymbol{\varphi}$ before change, $\boldsymbol{\varphi}^{(2)}$ is the vector $\boldsymbol{\varphi}$ after change, $\boldsymbol{\gamma}^{(1)}$, $\boldsymbol{\gamma}^{(2)}$ are the before and after change values of vector $\boldsymbol{\gamma} = \left(\gamma(0), \gamma(1), \ldots, \gamma(p-1)\right)^t$. Note that in the expression of $\gamma^{(1)}(s) - \gamma^{(2)}(s)$ above the first term is the same for all lags s, $1 \leqslant s \leqslant p$.

By using the strong law of large numbers, it is easy to see that

$$\widehat{\boldsymbol{\varphi}}_n \to \widehat{\boldsymbol{\varphi}}^A = \left(\rho\Gamma^{(1)} + (1-\rho)\Gamma^{(2)}\right)^{-1} \begin{pmatrix} \rho E(Y_1Y_2) \ + (1-\rho)E(Y_nY_{n-1}) \\ \rho E(Y_1Y_{1+s}) + (1-\rho)E(Y_nY_{n-s}) \\ \rho E(Y_1Y_{1+p}) + (1-\rho)E(Y_nY_{n-p}) \end{pmatrix}, \quad \text{a.s.}, \tag{2.2}$$

where $\Gamma^{(1)}$, $\Gamma^{(2)}$ are the before and after change $p \times p$ covariance matrices of $(Y_i, \ldots, Y_{i+p-1})$. So estimator $\widehat{\boldsymbol{\varphi}}_n$ converges to $\widehat{\boldsymbol{\varphi}}^A$, the right-hand side of (2.2), and $\widehat{\boldsymbol{\varphi}}^A \neq \boldsymbol{\varphi}^{(1)}$, $\widehat{\boldsymbol{\varphi}}^A \neq \varphi^{(2)}$.

If these two values are not the same the proof of (2.1) is the same as in the case of the change in μ only.

(iii) Test for change in σ^2 uses the process

$$\sum_{i=1}^{k} \widehat{\varepsilon}_i^2 - \frac{k}{n} \sum_{i=1}^{n} \widehat{\varepsilon}_i^2, \quad 1 < k \leqslant n, \tag{2.3}$$

where $\widehat{\varepsilon}_i = Y_i - \overline{Y}_n - \widehat{\varphi}_{n1}(Y_{i-1} - \overline{Y}_n) - \cdots - \widehat{\varphi}_n(Y_{i-p} - \overline{Y}_n)$.

The $\widehat{\varepsilon}_i$ terms are identically distributed for $i \leqslant \tau$ and for $i > \tau + p$, so for (2.1) we need to show only that there is a change in the mean of the process in (2.3). Taking $\widehat{\varphi}_{n0} = 1$

$$\widehat{\varepsilon}_i^2 = \sum_{\ell=0}^{p} \sum_{m=0}^{p} \widehat{\varphi}_{n\ell}\widehat{\varphi}_{nm}(Y_{i-\ell} - \overline{Y}_n)(Y_{i-m} - \overline{Y}_n),$$

and using, again, the convergence of $\widehat{\boldsymbol{\varphi}}_n$ to a value $\boldsymbol{\varphi}^A$, for large n $E\widehat{\varepsilon}_i^2$ is approximately $(\boldsymbol{\varphi}^A)^t \Gamma^{(j)} \boldsymbol{\varphi}^A$, $j = 1$ before change and $j = 2$ after change, where $\Gamma^{(j)}$, $j = 1, 2$, is defined in (2.2), so change in mean for process (2.3) follows.

Theorem 2. *If there is change in the mean only from μ_1 to μ_2, $\mu_1 > \mu_2$, at $\tau = [\rho n]$, $0 < \rho < 1$, then for the asymptotic distribution of the test statistic we have*

$$n^{-1/2}\sigma_*^{-1}\left[\max_{1<k\leqslant n} \frac{\partial}{\partial\mu}\,\ell_k(\widehat{\boldsymbol{\xi}}_n) - \frac{1 - \sum_{j=1}^{p} \widehat{\varphi}_{jn}}{\widehat{\varphi}_n^2}\,\tau(1-\rho)(\mu_1 - \mu_2)\right] \to^{\mathcal{D}} N(0, 1),$$

$$\textit{as } n \to \infty, \quad \textit{where } \sigma_*^2 = \frac{\left(1 - \sum_{j=1}^{p} \widehat{\varphi}_{jn}\right)^2}{\widehat{\varphi}_n^2}\,\rho.$$

The proof of Theorem 2 is along the lines of the corresponding part of Theorem 1 of Gombay [6], and some calculations, hence it is omitted. From Theorem 2 the consistency of the tests follow. Consistency of tests, when parameters other than the mean change, can be deducted from the appropriate versions of Theorem 2.

3. Empirical study

In the AR(1) model we have three parameters. For mean μ, variance σ^2, and for the time series parameter φ the standardized efficient score vectors are

$$\frac{\partial \ell_k}{\partial \mu} = n^{-1/2}\widehat{\sigma}_n^{-1} \sum_{i=1}^{k} [(Y_i - \overline{Y}_n) - \widehat{\varphi}_n(Y_{i-1} - \overline{Y}_n)],$$

$$\frac{\partial \ell_k}{\partial \sigma^2} = n^{-1/2} 2^{-1/2} \widehat{\sigma}_n^{-2} \sum_{i=1}^{k} \{[Y_i - \overline{Y}_n - \widehat{\varphi}_n(Y_{i-1} - \overline{Y}_n)]^2 - \widehat{\sigma}_n\}$$

and

$$\frac{\partial \ell_k}{\partial \varphi} = n^{-1/2}(1 - \widehat{\varphi}_n^2)^{1/2} \widehat{\sigma}_n^{-2} \sum_{i=1}^{k} \{[(Y_i - \overline{Y}_n) - \widehat{\varphi}_n(Y_{i-1} - \overline{Y}_n)](Y_{i-1} - \overline{Y}_n)\},$$

respectively. We assume, that the change results is large positive values of the test statistic, and do a small simulation study for the one-sided tests.

Table 1 compares the theoretical percentiles, based on the asymptotic distribution (first row), to the empirical percentiles of the three tests. The sample size was $n = 200$, the mean $\mu = 0$, the variance $\sigma^2 = 1$, and $\varphi = 0.3$, all assumed to be unknown. In parenthesis we record the

Table 1
Comparison of percentiles for Test 1 ($n = 200$) if no change occurred

	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
Asymptotic	1.073	1.224	1.517
Test for change in μ	1.035 (0.086)	1.179 (0.039)	1.452 (0.006)
Test for change in σ^2	1.020 (0.076)	1.166 (0.036)	1.430 (0.006)
Test for change in ϕ	1.039 (0.084)	1.189 (0.042)	1.533 (0.012)

$\mu = 0,\ \sigma^2 = 1,\ \phi = 0.3$, all assumed unknown.

Table 2
Empirical power of Test 1 ($n = 200$) from unknown $\mu = 0$, $\sigma^2 = 1$, $\phi = 0.3$ at $\tau = 100$, $\alpha = 0.05$

Change in μ		Change in σ^2		Change in ϕ	
μ	Power	σ^2	Power	ϕ	Power
0.1	0.093	0.9	0.185	0.1	0.283
0.3	0.324	0.8	0.577	0.0	0.518
0.5	0.662	0.7	0.913	−0.1	0.757
0.7	0.899	0.6	0.997	−0.2	0.913
0.9	0.987	0.5	1.00	−0.3	0.974

empirical level, that is, the rejection rate if no change occurred. We can see that these tests are somewhat conservative. Table 2 summarizes the results of a small power study. The sample size was $n = 200$ again, and the above parameter values changed to the ones recorded at change-point $k = 100$. Only one parameter changed in each case, and tests for change in either of the other two parameters maintained an empirical level close to the nominal $\alpha = 0.05$. Figures in the tables are based on 2000 Monte Carlo experiments for each scenario. Increasing the number of repetitions did not change the overall picture.

Appendix. Proofs

Proof of Theorem 1. It has to be done separately for the different parameters.

(i) First, we consider the case of detecting change in μ. The component of the efficient score vector that we use for this purpose is

$$\frac{\partial \ell_k}{\partial \mu} = \frac{1 - \sum_{j=1}^{p} \widehat{\varphi}_{jn}}{\widehat{\sigma}_n^2} \sum_{i=1}^{k} \Big[(Y_i - \widehat{\mu}_n) - \sum_{j=1}^{p} \widehat{\varphi}_{jn} (Y_{i-j} - \widehat{\mu}_n) \Big]$$
$$\pm \frac{1 - \sum_{j=1}^{p} \varphi_j}{\sigma^2} \sum_{i=1}^{k} \Big[(Y_i - \overline{Y}_n) - \sum_{j=1}^{p} \varphi_j (Y_{i-j} - \overline{Y}_n) \Big]. \tag{A.1}$$

First we show the convergence of the dominant term to a Brownian bridge $B(u)$.

$$\begin{aligned}
&\frac{1-\sum\limits_{j=1}^{p}\varphi_j}{\sigma^2}\sum_{i=1}^{k}\Big[(Y_i-\overline{Y}_n)-\sum_{j=1}^{p}\varphi_j(Y_{i-j}-\overline{Y}_n)\Big]\\
&\quad=\frac{1-\sum\limits_{j=1}^{p}\varphi_j}{\sigma^2}\Big\{\sum_{i=1}^{k}(Y_i-\overline{Y}_n)-\sum_{j=1}^{p}\varphi_j\sum_{i=1}^{k}(Y_{i-j}-\overline{Y}_n)\Big\}. \qquad \text{(A.2)}
\end{aligned}$$

In Gombay and Serban [8] it is shown that

$$\Big|\sum_{i=1}^{[t]}(Y_i-\mu)-\sigma^* W(t)\Big| = o(t^{1/v}) \quad \text{a.s.} \qquad \text{(A.3)}$$

for some $v > 2$, where $W(t)$ is a Brownian motion, and $\sigma^* > 0$.

From this we get

$$\sup_{0\leqslant u\leqslant 1}\Big|n^{-1/2}\Big(\sum_{i=1}^{[nu]}Y_i-\frac{[nu]}{n}\sum_{i=1}^{n}Y_i\Big)-\sigma^*\big(W(u)-uW(1)\big)\Big| = o_p(1). \qquad \text{(A.4)}$$

For the difference

$$n^{-1/2}\Big(\sum_{i=1}^{k}Y_{i-j}-\sum_{i=1}^{k}Y_i\Big)=n^{-1/2}\sum_{i=1}^{j}Y_{1-i}-n^{-1/2}\sum_{i=k-j+1}^{k}Y_i, \qquad \text{(A.5)}$$

we recall, that by Theorem 1.2.1 of Csörgő and Révész [4]

$$\limsup_{k\to\infty}\sup_{0\leqslant s\leqslant p}|W(k-s)-W(k)| = O\big((\log k)^{1/2}\big) \quad \text{a.s.,} \qquad \text{(A.6)}$$

so for the error term (A.5), using (A.3) and (A.6) we have as $n\to\infty$

$$\max_{1<k\leqslant n}\max_{1\leqslant j\leqslant p} n^{-1/2}\Big|\sum_{i=k-j+1}^{k}Y_i\Big| = o_p(1). \qquad \text{(A.7)}$$

Putting (A.1)–(A.7) together, after standardization we get

$$\sup_{0\leqslant u\leqslant 1}\Big|\Big(1-\sum_{j=1}^{p}\varphi_j\Big)^{-1}\sum_{i=1}^{[nu]}\big[(Y_i-\overline{Y}_n)-\sum_{j=1}^{p}\varphi_j(Y_{i-j}-\overline{Y}_n)\big]-\sigma^* B(u)\Big| = o_p(1).$$

In the error analysis we recall that for the maximum likelihood estimators we have by Lemmas 2.1–2.3 of Gombay and Serban [8]

$$|\widehat{\mu}_n-\mu| = O\Big(\sqrt{\frac{\log\log n}{n}}\Big) \quad \text{a.s.,} \qquad \text{(A.8)}$$

$$|\widehat{\sigma}_n^2-\sigma^2| = O\Big(\sqrt{\frac{\log\log n}{n}}\Big) \quad \text{a.s.} \qquad \text{(A.9)}$$

and for $j = 1, \dots, p$

$$|\widehat{\varphi}_{jn} - \varphi_j| = O\Big(\sqrt{\frac{\log\log n}{n}}\Big) \quad \text{a.s.} \tag{A.10}$$

By (A.9) and (A.10), the error committed by replacing $(1 - \sum \widehat{\varphi}_j)/\widehat{\sigma}^2$ by $(1 - \sum \varphi_j)/\sigma^2$ is $O(n^{-1/2}(\log\log n)^{1/2})$ a.s. We can write

$$\begin{aligned}
&\frac{1 - \sum\limits_{j=1}^{p} \varphi_j}{\sigma^2} \Big\{ \sum_{i=1}^{k} (Y_i - \overline{Y}_n) - \sum_{j=1}^{p} \varphi_j (Y_{i-j} - \overline{Y}_n) \Big\} \\
&\quad - \frac{1 - \sum\limits_{j=1}^{p} \varphi_j}{\sigma^2} \Big\{ \sum_{i=1}^{k} (Y_i - \overline{Y}_n) - \sum_{j=1}^{p} \widehat{\varphi}_{jn} (Y_{i-j} - \overline{Y}_n) \Big\} \\
&\qquad = \frac{1 - \sum\limits_{j=1}^{p} \varphi_j}{\sigma^2} \Big\{ \sum_{i=1}^{k} \sum_{j=1}^{p} (\widehat{\varphi}_{jn} - \varphi_j)(Y_{i-j} - \overline{Y}_n) \Big\} \\
&\qquad \leqslant \max_{1 \leqslant j \leqslant p} |\widehat{\varphi}_{jn} - \varphi_j| \, \Big| \sum_{i=1}^{k} (Y_{i-j} - \overline{Y}_n) \Big|.
\end{aligned}$$

So by (A.4), (A.7) and (A.10)

$$\sup_{0 \leqslant u \leqslant 1} \max_{1 \leqslant j \leqslant p} |\widehat{\varphi}_{jn} - \varphi_j| n^{-1/2} \Big| \sum_{i=1}^{[nu]} (Y_{i-j} - \overline{Y}_n) \Big| = O_p\Big(\sqrt{\frac{\log\log n}{n}}\Big).$$

Finally as

$$\widehat{\mu}_n = \frac{1}{n\big(1 - \sum\limits_{j=1}^{p} \varphi_j\big)} \sum_{i=1}^{n} \Big[Y_i - \sum_{j=1}^{p} \varphi_j Y_{i-j} \Big]$$

using (A.3) and (A.6) we get that

$$|\widehat{\mu}_n - \overline{Y}_n| = o(n^{1/\nu - 1}) \quad \text{a.s.,} \tag{A.11}$$

which contributes a negligible error, when we replace $\widehat{\mu}_n$ by $\overline{Y}_n$ in our test statistic.

(ii) For detecting change in φ_s, $s = 1, \dots, p$, we consider statistic

$$\frac{\partial}{\partial \varphi_s} \ell_k(\widehat{\mu}_n, \widehat{\sigma}^2_n, \widehat{\boldsymbol{\varphi}}_n) = \frac{1}{\widehat{\sigma}^2_n} \sum_{i=1}^{k} (Y_{i-s} - \widehat{\mu}_n) \Big[Y_i - \widehat{\mu}_n - \sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \widehat{\mu}_n) \Big].$$

By (A.9) replacing σ^2 by $\widehat{\sigma}^2_n$ does not change the asymptotic distribution. Let

$$\widehat{\varepsilon}_i = (Y_i - \widehat{\mu}_n) - \sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \widehat{\mu}_n),$$

which estimates $\varepsilon_i = (Y_i - \mu) - \sum_{j=1}^{p} \varphi_j (Y_{i-j} - \mu)$. Note that $\sum_{i=1}^{n}(Y_{i-s} - \widehat{\mu}_n)\widehat{\varepsilon}_i \equiv 0$. By the invariance principle for the sequence $\sum_{i=1}^{k}(Y_{i-s} - \mu)\varepsilon_i$, proven in Gombay and Serban [8],

$$\sup_{0 \leqslant u \leqslant 1} \left| n^{-1/2}\Big[\sum_{i=1}^{[nu]}(Y_{i-s} - \mu)\varepsilon_i - \frac{[nu]}{n} \sum_{i=1}^{n}(Y_{i-s} - \mu)\varepsilon_i - \sigma_1 B(u)\Big]\right| = o_p(1) \tag{A.12}$$

with some $\sigma_1^2 > 0$. We now have to show that the error committed by replacing the parameters with their maximum likelihood estimators is negligible. This error can be written as a constant times

$$\begin{aligned}
&\Big[\sum_{i=1}^{k}(Y_{i-s} - \widehat{\mu}_n)\widehat{\varepsilon}_i - \sum_{i=1}^{k}(Y_{i-s} - \mu)\varepsilon_i\Big] \\
&\quad + \Big[-\frac{k}{n} \sum_{i=1}^{n}(Y_{i-s} - \widehat{\mu}_n)\widehat{\varepsilon}_i + \frac{k}{n} \sum_{i=1}^{n}(Y_{i-s} - \mu)\varepsilon_i\Big] = B_{nk} + D_{nk}.
\end{aligned}$$

We analyze B_{nk} as

$$\begin{aligned}
B_{nk} &= \sum_{i=1}^{k}(Y_{i-s} - \widehat{\mu}_n \pm \mu)(\widehat{\varepsilon}_i \pm \varepsilon_i) - \sum_{i=1}^{k}(Y_{i-s} - \mu)\varepsilon_i \\
&= (\mu - \widehat{\mu}_n)\sum_{i=1}^{k}\varepsilon_i + \sum_{i=1}^{k}(Y_{i-s} - \mu)(\widehat{\varepsilon}_i - \varepsilon_i) + (\mu - \widehat{\mu}_n)\sum_{i=1}^{k}(\widehat{\varepsilon}_i - \varepsilon_i) \\
&= B_{nk}^{(1)} + B_{nk}^{(2)} + B_{nk}^{(3)}.
\end{aligned} \tag{A.13}$$

By (A.8) and the law of iterated logarithm for the $\{\varepsilon_i\}$ sequence $B_{nk}^{(1)} = O(\log\log n)$ a.s. As

$$\begin{aligned}
\widehat{\varepsilon}_i - \varepsilon_i &= (\widehat{\mu}_n - \mu)\Big(1 - \sum_{j=1}^{p}\varphi_j\Big) + \sum_{j=1}^{p}(\widehat{\varphi}_{nj} - \varphi_j)(Y_{i-j} - \mu) \\
&\qquad + (\mu - \widehat{\mu}_n)\sum_{j=1}^{p}(\widehat{\varphi}_{nj} - \varphi_j),
\end{aligned}$$

we get

$$B_{nk}^{(3)} = O(\log\log n) \quad \text{a.s.}, \tag{A.14}$$

using the invariance principle and the law of iterated logarithm for the $\sum(Y_{i-j} - \mu)$ sequence, (A.8), and (A.10).

For $B_{nk}^{(2)}$ the analysis is similar, but now we invoke the invariance principle for the sequence $\sum_{i=1}^{k}\left[(Y_{i-s}-\mu)(Y_{i-j}-\mu)-E\big((Y_{i-s}-\mu)(Y_{i-j}-\mu)\big)\right]$ as well:

$$\begin{aligned}B_{nk}^{(2)} &= \sum_{i=1}^{k}(Y_{i-s}-\mu)\Big[(\widehat{\mu}_n-\mu)\Big(1-\sum_{j=1}^{p}\varphi_j\Big)+(\mu-\widehat{\mu}_n)\sum_{j=1}^{p}(\widehat{\varphi}_{nj}-\varphi_j)\Big]\\ &\quad+\sum_{i=1}^{k}(Y_{i-s}-\mu)\sum_{j=1}^{p}(Y_{i-j}-\mu)(\widehat{\varphi}_{nj}-\varphi_j)\\ &\quad\pm k\sum_{j=1}^{p}E\big((Y_{i-s}-\mu)(Y_{i-j}-\mu)\big)(\widehat{\varphi}_{nj}-\varphi_j)\\ &= k\sum_{j=1}^{p}E\big((Y_{i-s}-\mu)(Y_{i-j}-\mu)\big)(\widehat{\varphi}_{nj}-\varphi_j)+O(\log\log n)\quad\text{a.s.}\end{aligned}\tag{A.15}$$

The error term D_{nk} is of the same structure as B_{nk}, except for the multiplier k/n and that the summation is $\sum_{i=1}^{n}$. For this reason all terms in (A.14) and (A.15) that are not $O(\log\log n)$ are cancelled, hence we get $\sup_{1\leqslant k\leqslant n}|B_{nk}+D_{nk}|=O(\log\log n)$ a.s., which concludes the proof of the theorem for case (ii).

(iii) For change detection in the value of σ^2 consider

$$\begin{aligned}\frac{\partial}{\partial\sigma^2}\,\ell_k(\widehat{\mu}_n,\widehat{\sigma}_n^2,\widehat{\boldsymbol{\varphi}}_n) &= \frac{1}{2\widehat{\sigma}_n^4}\sum_{i=1}^{k}\widehat{\varepsilon}_i^{\,2}-\frac{1}{2\widehat{\sigma}_n^4}\frac{k}{n}\sum_{i=1}^{n}\widehat{\varepsilon}_i^{\,2}\\ &= \frac{1}{2\sigma^4}\Big(\sum_{i=1}^{k}\varepsilon_i^2-\frac{k}{n}\sum_{i=1}^{n}\varepsilon_i^2\Big)+\Big[\frac{1}{2\widehat{\sigma}_n^4}\sum_{i=1}^{k}\widehat{\varepsilon}_i^{\,2}-\frac{1}{2\sigma^4}\sum_{i=1}^{k}\varepsilon_i^2\Big]\\ &\quad-\frac{k}{n}\Big[\frac{1}{\widehat{\sigma}_n^4}\sum_{i=1}^{n}\widehat{\varepsilon}_i^{\,2}-\frac{1}{\sigma^4}\sum_{i=1}^{n}\varepsilon_i^2\Big]\\ &= A_{nk}+B_{nk}+D_{nk}.\end{aligned}$$

For A_{nk} we get by standard invariance principle arguments, that

$$\sup_{0\leqslant u\leqslant 1}\left|n^{-1/2}A_{[nu]}-\sigma_2 B(u)\right|=o_p(1)$$

for some $\sigma_2^2>0$, and a Brownian bridge $B(u)$. As by (A.9) the estimation of σ^2 by $\widehat{\sigma}^2$ does not change the asymptotic distribution, we will consider

$$\begin{aligned}\sum_{i=1}^{k}(\widehat{\varepsilon}_i^{\,2}-\varepsilon_i^2) &= \sum_{i=1}^{k}\Big[Y_i-\widehat{\mu}_n-\sum_{j=1}^{p}\widehat{\varphi}_j(Y_{i-j}-\widehat{\mu})\Big]^2\\ &\quad-\sum_{i=1}^{k}\Big[Y_i-\mu-\sum_{j=1}^{p}\varphi_j(Y_{i-j}-\mu)\Big]^2\\ &\quad\pm\sum_{i=1}^{k}\Big[Y_i-\mu-\sum_{j=1}^{p}\widehat{\varphi}_j(Y_{i-j}-\mu)\Big]^2.\end{aligned}$$

We have

$$\sum_{i=1}^{i} \varepsilon_i^2 - \sum_{i=1}^{k} \Big[Y_i - \mu - \sum_{j=1}^{p} \widehat{\varphi}_j (Y_{i-j} - \mu) \Big]^2$$
$$= \sum_{i=1}^{k} \Big\{ -2(Y_i - \mu) \sum_{j=1}^{p} (\varphi_j - \widehat{\varphi}_{nj})(Y_{i-j} - \mu)$$
$$+ \Big[\sum_{j=1}^{p} \varphi_j (Y_{i-j} - \mu) \Big]^2 - \Big[\sum_{j=1}^{p} \widehat{\varphi}_j (Y_{i-j} - \mu) \Big]^2 \Big\}.$$

We recall invariance principles for the $\sum_{i=1}^{k} (Y_i - \mu)$ and $\sum_{i=1}^{k} (Y_i - \mu)(Y_{i-j} - \mu)$ sequences, the law of iterated logarithm, (A.10), and get that terms of the type:

$$(\varphi_j - \widehat{\varphi}_{nj}) \sum_{i=1}^{k} (Y_i - \mu)(Y_{i-j} - \mu) = (\varphi_j - \widehat{\varphi}_{nj}) k E\big((Y_i - \mu)(Y_{i-j} - \mu)\big)$$
$$+ O(\log \log n) \quad \text{a.s.}, \quad 0 \leqslant j \leqslant p.$$

Next, consider

$$\sum_{i=1}^{k} \Big\{ \Big[Y_i - \widehat{\mu}_n - \sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \widehat{\mu}_n) \Big]^2 - \Big[Y_i - \mu - \sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \mu) \Big]^2 \Big\}$$
$$= \sum_{i=1}^{k} \Big\{ 2(Y_i - \mu)(\mu - \widehat{\mu}_n) + (\mu - \widehat{\mu}_n)^2 - 2(Y_i - \mu) \sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \widehat{\mu}_n)$$
$$- 2(\mu - \widehat{\mu}_n) \sum_{i=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \widehat{\mu}_n) + \Big(\sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \widehat{\mu}_n) \Big)^2$$
$$+ 2(Y_i - \mu) \sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \mu) - \Big(\sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \mu) \Big)^2 \Big\}.$$

We have terms of the following types:

$$(\mu - \widehat{\mu}_n) \sum_{i=1}^{k} (Y_i - \mu) = O(\log \log n) \quad \text{a.s.}$$
$$k(\mu - \widehat{\mu}_n)^2 = O(\log \log n) \quad \text{a.s.}$$

and

$$\sum_{i=1}^{k} \Big[\Big(\sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \widehat{\mu}_n) \Big)^2 - \Big(\sum_{j=1}^{p} \widehat{\varphi}_{nj} (Y_{i-j} - \mu) \Big)^2 \Big]$$
$$= \sum_{i=1}^{k} \Big[\sum_{j=1}^{p} \sum_{\ell=1}^{p} \widehat{\varphi}_{nj} \widehat{\varphi}_{n\ell} (Y_{i-j} - \widehat{\mu}_n)(Y_{i-\ell} - \widehat{\mu}_n)$$
$$- \sum_{j=1}^{p} \sum_{\ell=1}^{p} \widehat{\varphi}_{nj} \widehat{\varphi}_{n\ell} (Y_{i-j} - \mu)(Y_{i-\ell} - \mu) \Big]$$

$$= \sum_{i=1}^{k} \left\{ \sum_{j=1}^{p} \sum_{\ell=1}^{p} \widehat{\varphi}_{nj} \widehat{\varphi}_{n\ell} \left[(\widehat{\mu}_n - \mu)\big((Y_{i-\ell} - \mu) + (Y_{i-j} - \mu)\big) + (\widehat{\mu}_n - \mu_n)^2 \right] \right\}$$
$$= O(\log\log n) \quad \text{a.s.}$$

As terms $(\varphi_j - \widehat{\varphi}_{nj})kE\big((Y_i - \mu)(Y_{i-j} - \mu)\big)$ in B_{nk} and D_{nk} cancel each other, $0 \leqslant j \leqslant p$, and the rest of the terms are $O(\log\log n)$ a.s. part (iii) of the theorem is proven. □

The proof of (2.1). If the mean changes from μ_1 to μ_2, then $E(\overline{Y}_n) = \rho\mu_1 + (1-\rho)\mu_2$. Assuming $\mu_1 > \mu_2$

$$E\left\{ \sum_{i=1}^{k} (Y_i - \overline{Y}_n) \right\} = k(1-\rho)(\mu_1 - \mu_2)$$

increases as $k = 1, 2, \ldots, \tau$, then $E\left\{ \sum_{i=1}^{k} (Y_i - \overline{Y}_n) \right\} = \tau(\mu_1 - \mu_2) - k\rho(\mu_1 - \mu_2)$ decreases as $k = \tau + 1, \ldots, n$, so we take

$$\widehat{\tau}_n = \min_k \left\{ \frac{\partial}{\partial\mu} \ell_k(\overline{Y}_n, \widehat{\sigma}_n, \widehat{\boldsymbol{\varphi}}_n) = \max_{1 < m \leqslant n} \frac{\partial}{\partial\mu} \ell_m(\overline{Y}_n, \widehat{\sigma}_n^2, \overline{\boldsymbol{\varphi}}_n) \right\}$$

as the change-point estimator. With $Y_0 = Y_{-1} = \cdots = Y_{-p+1} = 0$

$$\frac{\partial}{\partial\mu} \ell_k(\widehat{\boldsymbol{\xi}}_n) = \frac{\left(1 - \sum_{j=1}^{p} \widehat{\varphi}_{nj}\right)^2}{\widehat{\sigma}_n^2} \sum_{i=1}^{k} (Y_i - \overline{Y}_n)$$
$$+ \frac{1 - \sum_{j=1}^{p} \widehat{\varphi}_{nj}}{\widehat{\sigma}_n^2} \sum_{\ell=k+1-p}^{k} \left(Y_\ell \sum_{m=k+1-\ell}^{p} \widehat{\varphi}_{nm} \right)$$
$$= \frac{\left(1 - \sum_{j=1}^{p} \widehat{\varphi}_{nj}\right)^2}{\widehat{\sigma}_n^2} \sum_{i=1}^{k} (Y_i - \overline{Y}_n) + R_{nk}.$$

By the same argument that lead to (A.7)

$$\sup_{1<k\leqslant n} |R_{nk}| \leqslant C \left| \sum_{\ell=k+1-p}^{k} Y_\ell \right| = o_p(n^{1/2})$$

with some constant C, hence to prove (2.1) it is enough to show that

$$\lim_{K\to\infty} \limsup_{n\to\infty} P\left\{ \max_{1<k\leqslant \tau-K} f_k \geqslant \max_{\tau-K<k<\tau+K} f_k \right\} = 0, \tag{A.16}$$

and

$$\lim_{K\to\infty} \limsup_{n\to\infty} P\left\{ \max_{\tau+K\leqslant k\leqslant n} f_k \geqslant \max_{\tau-K<k<\tau+K} f_k \right\} = 0, \tag{A.17}$$

where $f_k = \sum_{i=1}^{k} (Y_i - \overline{Y}_n)$. For (A.16) we consider with a constant K, $K < \tau$,

$$
\begin{aligned}
&P\Big\{ \max_{1<k\leqslant \tau-K} \sum_{i=1}^{k} (Y_i - \overline{Y}_n) > \max_{\tau-K<k<\tau+K} \sum_{i=1}^{k} (Y_i - \overline{Y}_n)\Big\} \\
&\leqslant P\Big\{ \max_{1<k\leqslant \tau-K} \sum_{i=1}^{k} (Y_i - \overline{Y}_n) > \max_{\tau-K<k\leqslant\tau} \sum_{i=1}^{k} (Y_i - \overline{Y}_n)\Big\} \\
&= P\Big\{ \exists k : 1 < k \leqslant \tau - K,\ \max_{1\leqslant \ell \leqslant \tau-k} \sum_{i=1}^{\ell} (Y_i - \overline{Y}_n) \leqslant 0 \Big\} \\
&\leqslant P\Big\{ \sum_{i=1}^{K} (Y_i - \overline{Y}_n) < 0 \Big\} \\
&= P\Big\{ K^{-1/2} \sum_{i=1}^{K} (Y_i - \mu_1) + K^{1/2}[(1-\rho)(\mu_1 - \mu_2) - (\overline{Y}_n - E\overline{Y}_n)] < 0 \Big\}.
\end{aligned}
$$

The strong law of large numbers holds separately for the $\{Y_1, \ldots, Y_{\tau(n)}\}$ and $\{Y_{\tau(n)+1}, \ldots, Y_n\}$ sequences, so $\limsup_{n\to\infty} |\overline{Y}_n - E\overline{Y}_n| = 0$ a.s. The central limit theorem holds for $\sum_{i=1}^{K} (Y_i - \mu_1)$, so choosing $\delta > 0$ arbitrarily

$$
\limsup_{n\to\infty} P\Big\{ K^{-1/2} \sum_{i=1}^{K} (Y_i - \mu_1) + K[(1-\rho)(\mu_1 - \mu_2) - (\overline{Y}_n - E\overline{Y}_n) < 0 \Big\} < \delta
$$

if K is large enough, so (A.16) is proven. The proof of (A.17) is the same by symmetry, hence it is omitted. □

References

[1] R.A. Davies, D. Huang, Y.-C. Yao, Testing for a change in the parameter value and order of an autoregressive model, Ann. Statist. 23 (1995) 282–304.

[2] M. Bagshaw, R.A. Johnson, Sequential procedures for detecting parameter changes in time-series model, J. Amer. Statist. Assoc. 72 (1977) 593–597.

[3] M. Csörgő, L. Horváth, Limit Theorems in Change-point Analysis, Wiley, Chichester, 1997.

[4] M. Csörgő, P. Révész, Strong Approximations in Probability and Statistics, Academic Press, New York, 1981.

[5] D.V. Hinkley, Inference about the change-point in a sequence of random variables, Biometrika 57 (1970) 1–17.

[6] E. Gombay, U-statistics for change under alternatives, J. Multivariate Anal. 78 (2000) 139–158.

[7] E. Gombay, L. Horváth, Strong invariance principles for stationary time series, Manuscript, 2006.

[8] E. Gombay, D. Serban, Monitoring parameter change in $AR(p)$ time series models, Statistics Centre Technical Reports 05.04, The University of Alberta, Edmonton, Canada, 2005.

[9] N.H. Kuiper, Tests concerning random points on a circle, Proc. Nederl. Akad. Wetensch. Indag. Math. (A) 63 (1960) 38–47.

[10] A.M. Lurie, N.K. Neerchal, Bayes type tests for constancy of autoregressive parameters, Environmetrics 10 (1999) 737–752.

[11] D. Picard, Testing and estimating change-points in time series, Adv. Appl. Probab. 17 (1985) 841–867.

[12] R.S. Tsay, Outliers, level shifts, and variance changes in time series, J. Forecast. 7 (1988) 1–20.