

A wavelet-based approach for detecting changes in second order structure within nonstationary time series*

R. Killick, I. A. Eckley

Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK
e-mail: r.killick@lancs.ac.uk; i.eckley@lancs.ac.uk

and

P. Jonathan

Shell Technology Centre Thornton, P.O. Box 1, Chester, UK
e-mail: philip.jonathan@shell.com

Abstract: This article proposes a test to detect changes in general autocovariance structure in nonstationary time series. Our approach is founded on the locally stationary wavelet (LSW) process model for time series which has previously been used for classification and segmentation of time series. Using this framework we form a likelihood-based hypothesis test and demonstrate its performance against existing methods on various simulated examples as well as applying it to a problem arising from ocean engineering.

Keywords and phrases: segmentation, wavelets, local stationarity, significant wave height.

Received December 2012.

Contents

1	Introduction	1168
2	Background to locally stationary wavelet processes	1169
3	Detecting changes in second order structure	1170
3.1	Detecting a single change in autocovariance structure	1171
3.1.1	Likelihood of a LSW process with Gaussian innovations	1172
3.1.2	Likelihood-based test statistic	1172
3.2	Multiple changes in second order structure	1173
3.3	Penalty choice	1173
4	Simulation study	1175
5	Storm season identification	1179
6	Concluding remarks	1181
	References	1182

*R. Killick gratefully acknowledges financial support from Shell Research Limited and EPSRC. P. Jonathan was financially supported by Shell International Exploration Production. The authors are grateful to Richard Davis and Haeran Cho for providing the code for Auto-PARM and CF algorithms respectively, and to an Associate Editor and anonymous referee for helpful comments.

1. Introduction

This article considers the problem of detecting changepoints in the second order (i.e. autocovariance) structure of time series. The ability to detect such structural changes in temporal dependence retrospectively is of value in a number of applications, for example in renewable energy and biomedical research. It is therefore surprising that more attention has not been devoted to this topic within the literature to date, particularly when compared to the numerous contributions which already exist for the change in mean or change in variance problems (see Chen and Gupta (2000), Eckley et al. (2011) and references therein). Below we review the current state of the art in the area of autocovariance changepoint detection before outlining our approach and the structure of this paper.

Previous work. As observed, comparatively little attention has been given to the problem of autocovariance changepoint detection. A likelihood approach was taken by Davis et al. (2006) and Gombay (2008) to detect changes in autoregressive (AR) models. More specifically, Davis et al. (2006) introduce the Auto-PARM algorithm for segmenting time series with a piecewise AR structure. The test statistic used in Auto-PARM is a penalised likelihood ratio, the penalty being the Minimum Description Length (MDL). The multiple changepoint solution space is then traversed using a genetic algorithm to identify both the number and location of changepoints. In practice Auto-PARM is powerful in detecting changepoints due to the likelihood-based test statistic but has varied results when the assumption of piecewise AR structure is violated.

Other, non-parametric, approaches to this retrospective changepoint problem include the work of Ombao et al. (2001). Unfortunately this approach suffers when the change occurs at non-dyadic time points. An alternative method has been suggested by Ahamada et al. (2004) who consider the problem in a local Fourier setting, whereby changepoints are identified by considering the mean of the log-spectrum across a range of frequencies. As such, some changepoints may not be detectable due to frequency cancellation. Most recently, Cho and Fryzlewicz (2012) proposed a nonparametric test for autocovariance changepoint detection by considering the problem in a locally stationary wavelet paradigm. The results reported in their simulation study appear to be competitive with those of Davis et al. (2006), though as we shall see later (Section 4) their method can lack precision when one considers the locations of the detected changepoint locations in certain circumstances. In addition, their approach assumes that the variance of the summary statistics is constant across proposed changepoint locations, a feature which is very difficult to validate in practice.

Our approach. The changepoint method which we propose is founded on the Locally Stationary Wavelet time series model of Nason et al. (2000), a brief review of which is provided in the next section. In contrast to the approach proposed in Cho and Fryzlewicz (2012), we cast our changepoint problem in a parametric framework. The novelty of our approach comes by constructing a likelihood-based test for changes in the autocovariance structure in the time

domain which is equivalent to performing a test on the evolutionary wavelet spectral structure in the wavelet domain. Consequently, unlike the approach adopted by Davis et al. (2006) which assumes the process follows an AR model, the LSW model allows more general time series forms.

The article is structured as follows: Section 2 briefly reviews LSW processes before we proceed to describe our method for detecting changes in autocovariance structure in Section 3. The performance of the method is compared against Auto-PARM and the method from Cho and Fryzlewicz (2012) using simulated data in Section 4, and for detection of storm seasons from oceanographic time series in Section 5.

2. Background to locally stationary wavelet processes

We begin by providing a brief review of the LSW modelling paradigm. Recall that a wavelet is a (compactly supported) oscillating function with several properties that allow efficient location-scale decompositions; see Percival and Walden (2006) and Nason (2008) for accessible introductions.

Below we give a brief review of the LSW approach and the properties which we will utilise in Section 3.

Following Fryzlewicz and Nason (2006), a triangular stochastic array $\{X_{t,N}\}_{t=0}^{N-1}$ for $N = 1, 2, \dots$, is in a class of Locally Stationary Wavelet (LSW) processes if there exists a mean-square representation

$$X_{t,N} = \sum_{j=1}^{\infty} \sum_k W_j\left(\frac{k}{n}\right) \psi_{j,k-t} \xi_{jk}. \tag{1}$$

with j in $\{1, 2, \dots\}$ and $k \in \mathbb{Z}$ as scale and location parameters, respectively. The $\psi_j = (\psi_{j,0}, \dots, \psi_{i,\mathcal{L}_j-1})$ are discrete, real-valued, compactly supported, non-decimated wavelet vectors with support lengths $\mathcal{L}_j = \mathcal{O}(2^j)$, and the $\xi_{j,k}$ are zero-mean, orthonormal, identically distributed random variables. For each $j \geq 1$, the time-varying amplitudes, $W_j(z) : [0, 1] \rightarrow \mathbb{R}$ are real valued, piecewise constant functions with a finite (but unknown) number of jumps. Let \mathcal{P}_j denote the total magnitude of jumps in $W_j^2(z)$, then the variability of functions $W_j(z)$ is controlled so that

- $\sum_{j=1}^{\infty} W_j^2(z) < \infty$ uniformly in z ,
- $\sum_{j=1}^J 2^j \mathcal{P}_j = \mathcal{O}(\log n)$ where $J = \lfloor \log_2 N \rfloor$.

This modification of the constraints on the $W_j(z)$ extends the LSW model class to include time series with piecewise stationary second order structure.

Recall from classical time series theory that the second order structure of a time series is characterised by the spectrum. The spectrum for a LSW process is defined as $S_j(k/N) = |W_j(k/N)|^2$ and changes over time as well as scale. Nason et al. (2000) show that if X_t is a stationary process, i.e. does not change over time, then $W_j^2(k/N)$ is constant across each scale j . Thus if the second order structure of a time series is piecewise stationary, this manifests itself as piecewise

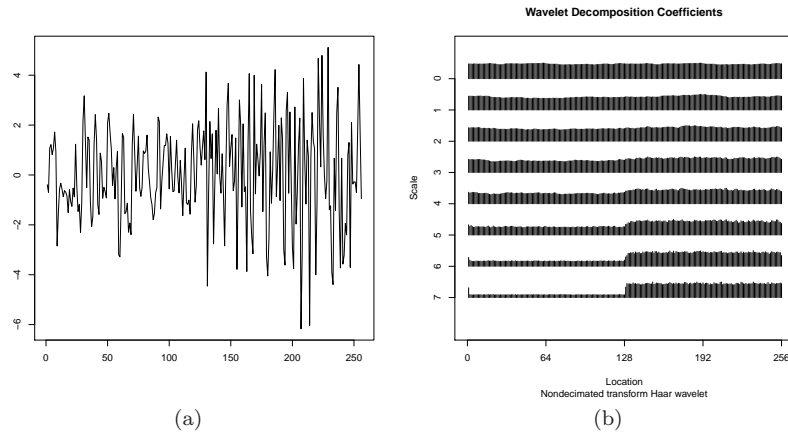


FIG 1. Example of a piecewise stationary MA process (model H) with changepoint at 128 (a) Original Time Series, (b) Average Evolutionary Wavelet Spectrum estimate over 100 realisations.

constant $W_j^2(k/N)$. Figure 1 provides an example of this where (a) shows a realisation from a piecewise MA model with a changepoint at time 128 and (b) shows the spectral estimate which, in the finer scales, clearly demonstrates a step change at 128.

The specific relationship between the autocovariance structure of X_t and $W_j^2(k/N)$ is such that it is easily shown that

$$\text{cov}(X_t, X_{t+v}) = \sum_{j,k} W_j^2\left(\frac{k}{N}\right) \psi_{j,k-t} \psi_{j,k-t-v}. \quad (2)$$

In the next section we utilise this decomposition to construct a test for a change in second order structure.

3. Detecting changes in second order structure

The Locally Stationary Wavelet (LSW) process can capture many dependence structures. Of particular interest in changepoint applications is the fact that a piecewise second order time series will have its structure encoded as piecewise constant sequences in the local wavelet periodogram - a feature noted by Cho and Fryzlewicz (2012). Consequently when using the LSW framework for changepoint methods we do not need to be prescriptive about the structure of the dependence beyond the requirements of the LSW definition.

In this article we restrict our consideration to LSW processes where the innovations $\{\xi_{j,k}\}$ in (1) are Gaussian although extensions to other distributions are possible.

Below we seek to identify a finite number of changes in autocovariance structure within a time series of length n . This is achieved by first expressing our

hypothesis for a single autocovariance change in the traditional setting. We then re-formulate the hypothesis in terms of the $W_j^2(k/n)$ parameters from the LSW model. Following this, we derive the form of the likelihood of a Gaussian LSW process under the null and alternative hypotheses which allows us to construct a test statistic in the usual way. We then conclude this section by extending the single changepoint model to multiple changes.

3.1. Detecting a single change in autocovariance structure

Our formulation of the hypothesis test for a single autocovariance changepoint follows that of Berkes et al. (2009):

Definition 3.1. The hypothesis to test for a single change in second order structure at any lag $v \geq 0$ is

$$\begin{aligned} \mathbf{H}_0 &: \text{cov}(X_0, X_{0-v}) = \text{cov}(X_1, X_{1-v}) = \dots = \text{cov}(X_{n-1}, X_{n-1-v}) = c_{0,v} \\ \mathbf{H}_1 &: c_{1,v} = \text{cov}(X_0, X_{0-v}) = \dots = \text{cov}(X_\tau, X_{\tau-v}) \\ &\quad \neq \text{cov}(X_{\tau+1}, X_{\tau+1-v}) = \dots = \text{cov}(X_{n-1}, X_{n-1-v}) = c_{n,v}. \end{aligned}$$

For notational simplicity we assume that X_{-v}, \dots, X_{-1} are observed.

Typically an upper bound is placed on v to ensure that identifiability problems do not occur but this should not be so small that the model fails to capture the second order structure adequately (e.g. using $v = 0$ when there is temporal dependence). Auto-PARM and similar time domain methods require the practitioner to decide how many lags, v should be considered. However, if we assume that the time series X_t can be well modelled by a LSW process then (2) demonstrates that the autocovariance structure across all lags is defined in terms of the $W_j^2(k/n)$. Thus there is no requirement to choose a maximum lag to consider. In addition, if the time series X_t is stationary (H_0) then $W_j^2(k/n) = \gamma_j$ at each scale j . In other words, under the null hypothesis the spectrum is constant over time and hence only one parameter per scale need be estimated. We therefore propose an equivalent form of the hypothesis provided in Definition 3.1, given here:

Definition 3.2. The hypothesis to test for a single change in autocovariance structure for a LSW time series can be stated as follows:

$$\begin{aligned} \mathbf{H}_0 &: W_j^2\left(\frac{0}{n}\right) = W_j^2\left(\frac{1}{n}\right) = \dots = W_j^2\left(\frac{n-1}{n}\right) = \gamma_{0,j} \quad \forall j \\ \mathbf{H}_1 &: \gamma_{1,j} = W_j^2\left(\frac{0}{n}\right) = \dots = W_j^2\left(\frac{\tau}{n}\right) \neq W_j^2\left(\frac{\tau+1}{n}\right) = \dots = W_j^2\left(\frac{n-1}{n}\right) = \gamma_{n,j}, \end{aligned}$$

for some $j \in \{1, 2, \dots\}$.

Given this alternative definition we must now derive the form of the likelihood under the null and alternative hypotheses (Section 3.1.2). Before we consider this, let us first describe the likelihood of a general LSW process with Gaussian innovations.

3.1.1. Likelihood of a LSW process with Gaussian innovations

Let $\mathbf{x} = \{x_1, \dots, x_n\}$ be an observed time series which is a LSW process with Gaussian innovations $\{\xi_{j,k}\}$. Then the log-likelihood of the time series can be expressed as follows:

$$\ell(W|\mathbf{x}) = \frac{n}{2} \log 2\pi - \frac{1}{2} \log |\Sigma_W| - \frac{1}{2} \mathbf{x}' \Sigma_W^{-1} \mathbf{x}, \quad (3)$$

where the variance-covariance matrix, Σ_W , has elements;

$$\Sigma_W(k, k') = \text{cov}(X_k, X_{k'}) = \sum_{l,m} W_l^2 \left(\frac{m}{n}\right) \psi_{l,m-k} \psi_{l,m-k'}.$$

Here $\Sigma_W(k, k')$ denotes the element in the k -th row and k' -th column of the Σ_W matrix. Henceforth, for notational simplicity, we shall omit the indexing of Σ by W although its dependence on W will still be implicit.

We now turn to consider the formulation of the likelihood under the null and alternative hypotheses.

3.1.2. Likelihood-based test statistic

Recall our equivalent form of the single changepoint hypothesis (Definition 3.2). Under the null hypothesis the spectrum of a stationary process is constant, i.e. $W_j^2(k/n) = \gamma_j$ for all locations k and scales j . Similarly under the alternative hypothesis the spectrum is piecewise constant. Hence, given a time series $\mathbf{x} = \{x_t\}_{t=1}^n$ which can be modelled as a LSW process with Gaussian innovations the likelihood ratio test statistic for a change in autocovariance structure is given by

$$\lambda = \max_{J < \tau < n-J} \left\{ \log |\hat{\Sigma}_0| + \mathbf{x}' \hat{\Sigma}_0^{-1} \mathbf{x} - \log |\hat{\Sigma}_1| - \mathbf{x}' \hat{\Sigma}_1^{-1} \mathbf{x} \right\}$$

where $J = \lfloor \log_2 n \rfloor$. Here $\hat{\Sigma}_0$ and $\hat{\Sigma}_1$ are the maximum likelihood estimates of the null and alternative variance-covariance matrices respectively. The entries of these matrices are given by,

$$\hat{\Sigma}_0(k, k') = \sum_l \sum_m \hat{\gamma}_{0,l} \psi_{l,m-k} \psi_{l,m-k'}, \quad (4)$$

$$\text{and} \quad \hat{\Sigma}_1(k, k') = \sum_l \left[\sum_{m \leq \tau} \hat{\gamma}_{1,l} \psi_{l,m-k} \psi_{l,m-k'} + \sum_{m > \tau} \hat{\gamma}_{n,l} \psi_{l,m-k} \psi_{l,m-k'} \right], \quad (5)$$

where $\hat{\gamma}_{0,l}$, $\hat{\gamma}_{1,l}$ and $\hat{\gamma}_{n,l}$ are the maximum likelihood estimates of the null, pre-change and post-change transfer function, $W_l^2(m/n)$ respectively. Using the above test statistic a changepoint is detected when, $\lambda > c$, where c is a constant. If a changepoint is detected, the location $\hat{\tau}$ of the change is estimated as $\arg_{\tau} \lambda$. The choice of c is discussed in Section 3.3.

Remark In practice, the variance-covariance matrix estimates occasionally produce singular matrices thus when this occurs we regularize the estimates to obtain invertible matrices. The work reported in this article adopts the regularization approach proposed by Schnabel and Eskow (1999), one of many regularization methods available.

Unfortunately, due to the structure of Σ , the maximum likelihood estimates of $W_j^2(k/n)$ have no closed form. However several authors including (Nason et al., 2000; Fryzlewicz and Nason, 2006) have derived alternative means of estimating $\{W_j^2(k/n)\}$. We turn to these to provide plug-in estimates as an alternative for $\hat{\gamma}$. In what follows we estimate $W_j^2(k/n)$ by averaging the empirical estimates from Nason et al. (2000) within each segment and using these as plug-in estimates in equations (4) and (5). Due to the use of plug-in estimates the test statistic is thus a likelihood-based test rather than a formal likelihood ratio test.

The consistency of likelihood methods for detecting changepoints has been shown in Csorgo and Horváth (1997). It is well known that plug-in estimates that are consistent possess the same properties as maximum likelihood estimates. Thus the proof of consistency for identifying changepoints using plug-in estimates follows directly provided the estimates used are consistent.

3.2. Multiple changes in second order structure

Thus far we have considered the problem of detecting a single changepoint. More generally we are interested in detecting (potentially) multiple changes. The likelihood ratio test statistic was shown to be consistent when estimating a single changepoint in the presence of multiple changes in Vostrikova (1981). Several algorithms have been proposed in the changepoint literature which extend single changepoint tests to multiple changes (e.g. Scott and Knott (1974); Bai and Perron (2003); Killick et al. (2012) and references therein). Typically the consistency of these search methods are proved for additive error whereas the LSW model has a multiplicative error structure. However, Cho and Fryzlewicz (2012) extend the consistency of the Binary Segmentation algorithm (Scott and Knott, 1974) to multiplicative errors and thus we implement this algorithm here.

In essence, the Binary Segmentation algorithm (Algorithm 1) applies the single changepoint test and upon identifying a change, iteratively implements the test statistic on sub-segments of the time series. In practice, we require an upper bound on the number of changepoints in order that the model remains identifiable. Thus we assume that there exists an upper bound, Q , on the number of changepoints in the series such that $J(Q+1) < n = 2^J$, which when rearranged gives $Q < (2^J - J)/J$.

3.3. Penalty choice

It is common in the changepoint literature to introduce a penalty to guard against overfitting (too many changepoints). When consideration extends to (potentially) multiple changepoints the penalty typically takes the form of a

Input: A set of data of the form, (x_1, x_2, \dots, x_n) .
 A test statistic $\lambda(\cdot)$ dependent on the data.
 An estimator of changepoint position $\hat{\tau}(\cdot)$.
 A rejection threshold (penalty) c .

Initialise: Let $\mathcal{C} = \emptyset$, and $\mathcal{S} = \{[1, n]\}$

Iterate while $\mathcal{S} \neq \emptyset$

1. Choose an element of \mathcal{S} ; denote this element as $[s, t]$.
2. If $\lambda(x_{s:t}) < c$, remove $[s, t]$ from \mathcal{S} .
3. If $\lambda(x_{s:t}) \geq c$ then:
 - (a) remove $[s, t]$ from \mathcal{S} ;
 - (b) calculate $r = \hat{\tau}(x_{s:t}) + s - 1$, and add r to \mathcal{C} ;
 - (c) if $r \neq s$ add $[s, r]$ to \mathcal{S} ;
 - (d) if $r \neq t - 1$ add $[r + 1, t]$ to \mathcal{S} .

Output the set of changepoints recorded \mathcal{C} .

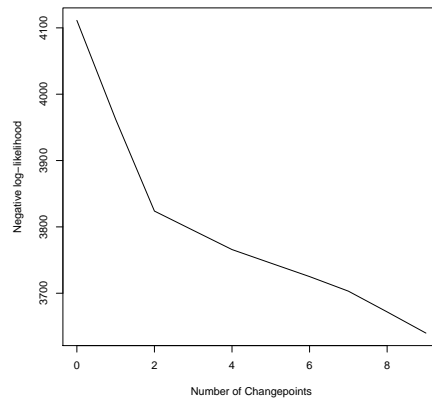
Algorithm 1: The Binary Segmentation Algorithm.

function dependent on the number of changepoints, e.g. $\beta f(m)$ where β is a constant and $f(\cdot)$ is a function of the number of changepoints m . See, for example Lavielle (2005); Harchaoui and Levy-Leduc (2010).

The approach we adopt follows the work of Lavielle (2005); a data driven method which is intuitive whilst also conveying the complexity of the decision through a graphical manner. The intuition behind the approach is the same as that of parameter selection in linear modelling i.e., when a true changepoint (informative parameter) is included the negative log-likelihood will decrease by a larger amount than when a false changepoint (uninformative parameter) is included. Hence, for a given sequence of data, we seek to identify the point at which the negative log-likelihood ceases to reduce significantly. More formally, let K_{max} be an upper bound on the number of segments; $1 \leq K \leq K_{max}$; τ be the ordered locations of changepoints; $J(\tau, \mathbf{x})$ be the negative log-likelihood for time series \mathbf{x} with changepoints τ ; and $\hat{\tau}_K$ the locations of K changepoints that minimise $J(\tau_K, \mathbf{x})$. Then, the approach can be summarised as follows:

1. For each $K \in \{0, \dots, K_{max}\}$ calculate $\hat{\tau}_K$ and $J(\hat{\tau}_K, \mathbf{x})$.
2. Determine, $l_K = J(\hat{\tau}_K, \mathbf{x}) - J(\hat{\tau}_{K+1}, \mathbf{x})$ for $K \in \{0, \dots, K_{max} - 1\}$.
3. The number of changepoints is the largest K such that $l_K \gg l_j$ for $j > K$.

A simple graphical method can also be implemented for the above which may prove more intuitive in practice. This involves producing a plot of $J(\tau_K, \mathbf{x})$ against K and visually identifying the point of maximum curvature. Figure 2(a) shows an example, from model B in Section 4, of the type of graphic produced by this method where the true number of changes is 2. For this example the values of l_k are 150, 138, 29, 29, 20, ... from which it is clear that $l_i \gg l_j$ for $i = 2$. This is the procedure which we implement in Section 4. As such this is reminiscent of using a scree plot in Principal Components Analysis to identify the number of principal components.



(a)

FIG 2. An example plot, from model B, of the negative log-likelihood against number of change-points for the purpose of identifying the number of changepoints.

4. Simulation study

In this section we compare the performance of the method presented in Section 3 against Auto-PARM (AP) proposed by Davis et al. (2006) and the work of Cho and Fryzlewicz (2012) (CF). We replicate a selection of simulations from these earlier papers to demonstrate how the proposed method, henceforth referred to as WL (wavelet-based likelihood) compares. In addition, new simulation examples are introduced to highlight the flexibility of the WL method. In all simulations we report results obtained for Haar wavelets though we note that similar results were obtained for other wavelet families. For comparison we used the default values specified in Davis et al. (2006) and Cho and Fryzlewicz (2012) for implementation of the AP and CF methods respectively. For the WL method, the number of changepoints were identified using the method described in Section 3.3 and the practitioner implementing the approach had no knowledge of the true number of changepoints.

As Eckley et al. (2011) observe, consideration of both the number *and* distribution of changepoint estimates is helpful in assessing the benefits of the different methods. Thus for each simulation case we consider (i) the number and (ii) the location of the identified changepoints. Tables 1, 2 and 3 report the number of changepoints detected by simulation scenario whilst Figure 3 displays the density of identified changepoints for scenarios where the methods are materially different. In each of the scenarios detailed below the innovations, ϵ_t have standard Normal distribution unless specified otherwise.

(A) Stationary AR(1) process with no changepoints This scenario is designed to assess the performance of the test when there are no changepoints. We fit an AR(1) model for a variety of parameter values, a (as used by Cho and

Fryzlewicz (2012)). More specifically we simulate from,

$$X_t = aX_{t-1} + \epsilon_t \quad \text{for } 1 \leq t \leq 1024.$$

Results from Table 1 show that as expected, Auto-PARM outperforms both CF and WL. However, the WL method has similar performance to Auto-PARM for all but the highest positive autocovariances.

(B) Piecewise stationary AR process with clearly observable changes Datasets are simulated from,

$$X_t = \begin{cases} 0.9X_{t-1} + \epsilon_t & \text{if } 1 \leq t \leq 512, \\ 1.68X_{t-1} - 0.81X_{t-2} + \epsilon_t & \text{if } 513 \leq t \leq 768, \\ 1.32X_{t-1} - 0.81X_{t-2} + \epsilon_t & \text{if } 769 \leq t \leq 1024, \end{cases}$$

i.e. the changes occur at dyadic points in the time series. This example highlights the increased power the WL method has over the traditional likelihood when the changepoints are at dyadic points in the time series. This is because there are more scales where the changepoint can be identified and thus more evidence for a change. Table 2 shows that we have comparable performance to Auto-PARM and substantially outperform CF in this type of scenario.

(C) Piecewise stationary AR process with less clearly observable changes Datasets are simulated from,

$$X_t = \begin{cases} 0.4X_{t-1} + \epsilon_t & \text{if } 1 \leq t \leq 400, \\ -0.6X_{t-1} + \epsilon_t & \text{if } 401 \leq t \leq 612, \\ 0.5X_{t-1} + \epsilon_t & \text{if } 613 \leq t \leq 1024. \end{cases}$$

In this example the changes no longer occur at dyadic time points and the changes are less clear. As with scenario B, our method has a similar performance to that of Auto-PARM whilst CF over estimates the number of changepoints 24% of the time.

(D) Piecewise stationary AR process with a short segment Datasets are simulated from,

$$X_t = \begin{cases} 0.75X_{t-1} + \epsilon_t & \text{if } 1 \leq t \leq 50, \\ -0.5X_{t-1} + \epsilon_t & \text{if } 51 \leq t \leq 1024. \end{cases}$$

The results for this example demonstrate that with relatively short segments the three methods perform well.

(E) Piecewise stationary AR process with high autocorrelation Datasets are simulated from,

$$X_t = \begin{cases} 1.399X_{t-1} - 0.4X_{t-2} + \epsilon_t, & \epsilon_t \sim \mathcal{N}(0, 0.8^2) & \text{if } 1 \leq t \leq 400, \\ 0.999X_{t-1} + \epsilon_t, & \epsilon_t \sim \mathcal{N}(0, 1.2^2) & \text{if } 401 \leq t \leq 750, \\ 0.699X_{t-1} + 0.3X_{t-2} + \epsilon_t & \epsilon_t \sim \mathcal{N}(0, 1) & \text{if } 751 \leq t \leq 1024. \end{cases}$$

The changepoints in this model are hard to identify because the autocorrelation function does not change very much between segments. For number of changepoints, the CF method outperforms both Auto-PARM and WL due to its variance stabilizing techniques. However, when we consider location of changepoints Figure 3(b) shows that the performance of our method is perhaps closer to CF than first appears. The distribution of changepoints for CF and our method are similar in shape and noticeably different from that of Auto-PARM which detects spurious changepoints towards the start of the data.

(F) Piecewise stationary ARMA(1,1) process Datasets are simulated from,

$$X_t = \begin{cases} 0.7X_{t-1} + \epsilon_t + 0.6\epsilon_{t-1} & \text{if } 1 \leq t \leq 125, \\ 0.3X_{t-1} + \epsilon_t + 0.3\epsilon_{t-1} & \text{if } 126 \leq t \leq 352, \\ 0.9X_{t-1} + \epsilon_t & \text{if } 353 \leq t \leq 704, \\ 0.1X_{t-1} + \epsilon_t - 0.5\epsilon_{t-1} & \text{if } 705 \leq t \leq 1024. \end{cases}$$

Figure 3(c) shows that all the methods identify the final changepoint more often than the first two. This is because the autocovariance function has greater difference between the third and fourth segments. For the remaining two changepoints, the WL method identifies both with similar probability whereas Auto-PARM prefers the second and CF the first changepoint. This again shows that the WL method is, in some sense, a compromise between the general structure of CF and the likelihood power of Auto-PARM.

(G) Piecewise stationary MA process Datasets are simulated from,

$$X_t = \begin{cases} \epsilon_t + 0.8\epsilon_{t-1} & \text{if } 1 \leq t \leq 128, \\ \epsilon_t + 1.68\epsilon_{t-1} - 0.81\epsilon_{t-2} & \text{if } 129 \leq t \leq 256, \end{cases}$$

In this scenario the WL method outperforms both Auto-PARM and CF methods. Although Auto-PARM detects the largest number of correct changepoints, Figure 3(d) shows that the location of these changepoints is split across two modes; around 125 and 135. The CF method performs poorly in this example, particularly when we consider Figure 3(d) which demonstrates that CF can severely mislocate the changes. In contrast to both these methods, the WL method has its mode around 127 and retains good detection of the correct number of changepoints.

The simulation results can be summarized as follows. As expected, when the Auto-PARM assumptions of piecewise AR model and independent segments are valid, it tends to outperform both CF and the proposed method. However, when the underlying assumption of a piecewise AR model is not valid, there are cases where the proposed method outperforms both Auto-PARM and CF. Indeed in 7 out of 10 scenarios the WL method clearly outperforms that of its nonparametric counterpart CF. Of the 3 remaining cases, results also highlighted that when considering performance of a changepoint method, the location of the changes

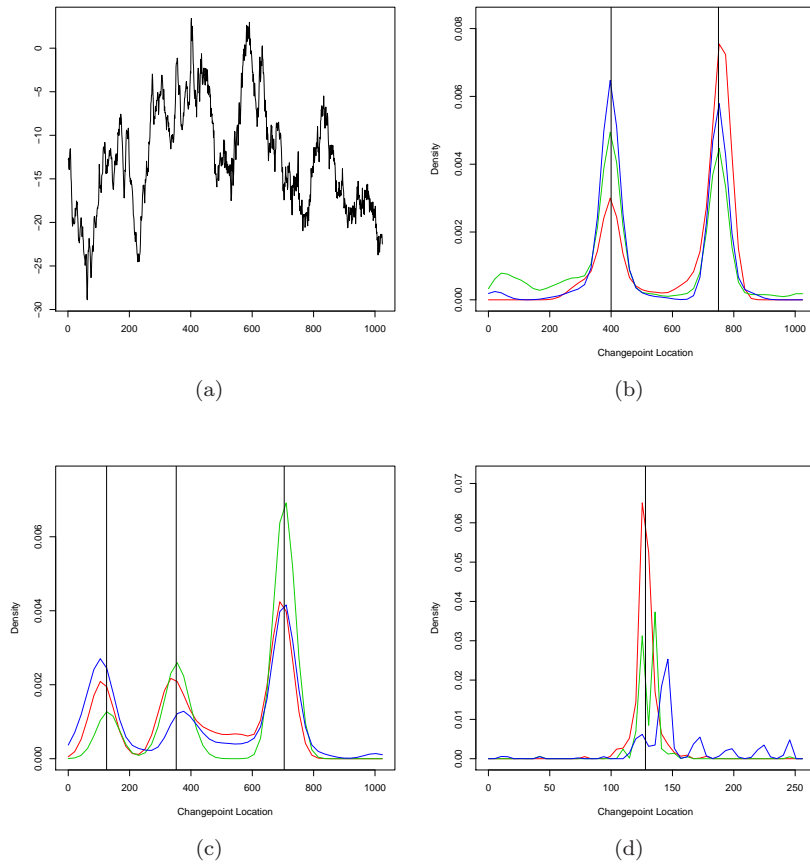


FIG 3. (a) Realisation from model E. (b)-(d) Density plots of the changepoints identified for (b) model E, (c) model F and (d) model G. In each plot the red line is the WL method; the blue line is the CF method and the green line is the AP method.

TABLE 1
Results for scenario A from our method (WL), the method from Cho and Fryzlewicz (2011) (CF) and Auto-PARM (AP). We report the percentage of repetitions that identified that number of changepoints. True number of changepoints is in bold.

a no. cpts	Model A											
	-0.7			-0.1			0.4			0.7		
	WL	CF	AP	WL	CF	AP	WL	CF	AP	WL	CF	AP
0	100	71	100	100	89	100	100	94	100	91	92	100
1	0	24	0	0	11	0	0	5	0	9	7	0
≥ 2	0	5	0	0	0	0	0	1	0	0	1	0

should be taken into account in addition to the number of changepoints detected. This is particularly apparent for Model G where the number of changepoints identified by CF was reasonable but Figure 3(d) showed many of these changes were far from the true location. The simulation study used the Haar wavelet.

TABLE 2

Results for scenarios B-D from our method (WL), the method from Cho and Fryzlewicz (2011) (CF) and Auto-PARM (AP). We report the percentage of repetitions that identified that number of changepoints. True number of changepoints is in bold.

no. of cpts	Model B			Model C			Model D		
	WL	CF	AP	WL	CF	AP	WL	CF	AP
0	0	0	0	0	0	0	4	2	0
1	0	0	0	0	0	0	94	83	100
2	98	70	94	94	76	100	2	15	0
3	2	27	6	6	22	0	0	0	0
4	0	3	0	1	1	0	0	0	0
≥ 5	0	0	0	0	1	0	0	0	0

TABLE 3

Results for scenarios E-G from our method (WL), the method from Cho and Fryzlewicz (2011) (CF) and Auto-PARM (AP). We report the percentage of repetitions that identified that number of changepoints. True number of changepoints is in bold.

no. of cpts	Model E			Model F			Model G		
	WL	CF	AP	WL	CF	AP	WL	CF	AP
0	0	0	0	0	0	0	0	0	0
1	26	9	9	20	12	51	99	85	100
2	45	75	33	22	36	33	1	15	0
3	26	15	31	35	45	16	0	0	0
4	3	1	15	22	6	0	0	0	0
≥ 5	0	0	12	1	1	0	0	0	0

Similar results were obtained for other wavelet families as expected, see Gott and Eckley (2013).

5. Storm season identification

We now consider storm season identification using significant wave heights. There is evidence (see for example Hidgkins et al. (2002)) that the onset of storm seasons is shifting. This is of interest as short-term operations, such as inspection and maintenance of marine structures, are typically performed outside seasons when storms are prevalent to minimise risk. As such knowledge of the date of onset of the storm season, e.g. of winter in the North Sea, or of hurricane season in the Gulf of Mexico, is particularly important if this is changing. For example, within the oil and gas industry, the current popularity of liquefying natural gas on an offshore production platform and offloading it to a shuttle gas carrier, requires knowledge of environmental winds, waves and currents that excite vessel motions, limiting floating liquefied natural gas (FLNG) operations. If the time periods that allow production and offloading are sufficient, the system has the potential to work safely and efficiently, and the economic benefit of FLNG can be realised, maximising the time periods during which FLNG platforms can be offloaded. Similarly, for offshore wind turbines, environmental conditions frequently cause costly delays in the installation and maintenance process. A systematic approach to identify approximately homogeneous periods in the ocean environment, and transitions between calm and stormy seasons in particular, is therefore highly desirable.

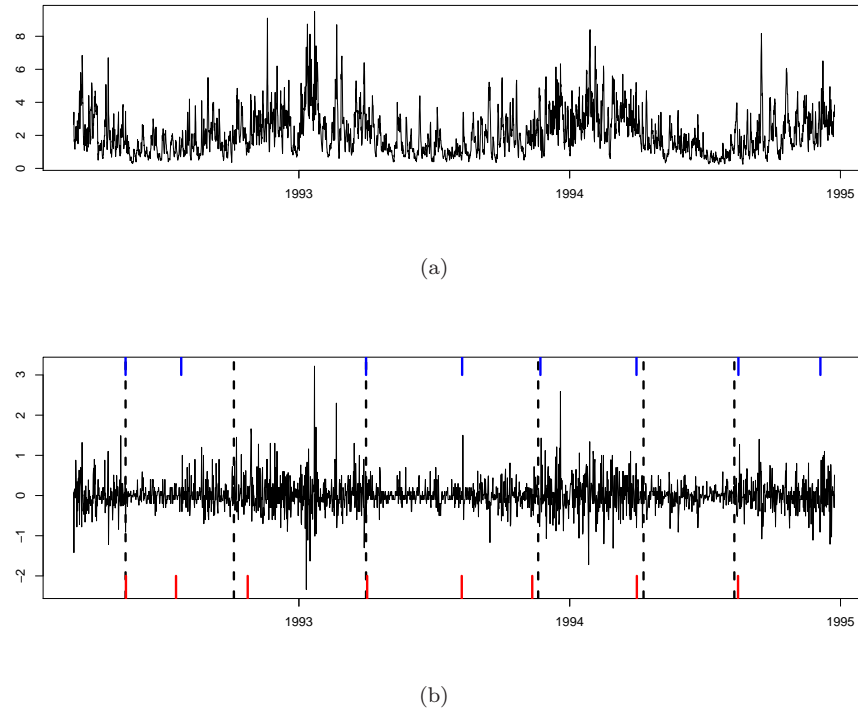


FIG 4. Central North sea application: (a) Significant wave heights (b) First difference of (a) with identified changepoints. Vertical dotted lines: WL method, blue short lines: CF method, red short lines: AP method.

Consider, by way of example, significant wave heights for a location in the central North Sea. Figure 4(a) depicts the measured wave height at 12 hour intervals for the period March 1992 - December 1994. Ocean engineers note that whilst certain features such as the cyclic nature of the series or the fact that wave heights tend to be lower during the summer than the winter are clear, other features are harder to discern. Specifically the points at which the storm season starts and ends are much more challenging to identify. For example methods of modelling the trend alone do not produce sufficiently accurate estimates of the storm season start and end. Use of established change in mean or change in regression methods such as those implemented in Killick and Eckley (2010) confirm this. However ocean engineers have suggested that looking at the second order structure could be more helpful since the variability of wave heights is expected to be small during calm and larger during storm seasons.

To identify whether a change in second order structure has occurred we must first remove the trend from the data. We choose to remove the trend by differencing but other methods of trend removal could be used. The resulting series is depicted in Figure 4(b). The change in second order structure is now more evi-

dent and the series appears to have a locally stationary structure. The question we therefore consider is how to detect the changes in autocovariance structure in this series and whether these changes correspond to the start and end of the storm season.

We apply the WL, AP and CF methods to detect changes in autocovariance of the differenced data shown in Figure 4(b). Applying the WL method from Section 3 we identify changepoints at 12th May 1992, 5th October 1992, 1st April 1993, 19th November 1993, 10th April 1994 and 10th August 1994. Each of these dates is indicated by a dashed vertical line in Figure 4(b). The changepoints identified by the CF and AP methods are indicated by vertical lines at the top and bottom of Figure 4(b) respectively. It is clear that the automatic penalty selection contained within the CF and AP methods results in more changepoints being identified than storm seasons. Both methods identify changes at approximately the same times as the WL method with additions between seasons. The exception to this is the CF method which does not identify a change around October 1992 where the WL and AP methods do. The results for the three methods were (blindly) presented to ocean engineers who selected the WL method as the set of changepoints which most closely matched their own assessment of the start and end of the storm season.

6. Concluding remarks

As discussed in Section 1, oceanographic applications require methods for segmenting time series where the observations cannot be assumed to be independent. The work presented in this paper has addressed this issue by developing a test for a change in autocovariance structure using the LSW framework and assuming no prior knowledge of the structure of the autocovariance. A benefit of this approach is that it does not assume a specific structure for the autocovariance.

The proposed approach was compared with both the Auto-PARM method of Davis et al. (2006) and a wavelet-based method recently proposed by Cho and Fryzlewicz (2012). The simulation study in Section 4 showed that the WL method provides an extra degree of refinement between a restrictive model assumption (Auto-PARM) and a nonparametric approach (Cho and Fryzlewicz, 2012). The refinement comes by using a more flexible model structure, namely the LSW model within a likelihood based framework. Intuitively since it is founded on the likelihood framework, when the LSW model assumptions hold the WL method should outperform a nonparametric approach.

Whilst all the changes in the simulations were detected, situations could arise where a change is too small in magnitude or too close to another change to be identified. This is a challenge for all the methods detailed in this paper. For changes in mean and variance there are clear guidelines on how small a change can be detected by specific methods. For changes in autocovariance there are no guidelines due to the complexities of possible pre and post change structures; as such it would be an interesting avenue for future research.

The proposed method was also applied to oceanographic data encountered by an industrial collaborator. The results demonstrate that the proposed method gives an automatic approach for segmenting series into seasons when storms are prevalent for further analysis or planning of maintenance. An added advantage of the WL method is that a specific number of changepoints can be identified whereas the CF and AP methods would require postprocessing procedures for selecting the changes that correspond to storm seasons.

References

- AHAMADA, I., JOUINI, J., BOUTAHAR, M., 2004. Detecting multiple breaks in time series covariance structure: a non-parametric approach based on the evolutionary spectral density. *Applied Economics* 36 (10), 1095–1101.
- BAI, J., PERRON, P., 2003. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18 (1), 1–22.
- BERKES, I., GOMBAY, E., HORVATH, L., 2009. Testing for changes in the covariance structure of linear processes. *Journal of Statistical Planning and Inference* 139 (6), 2044 – 2063. [MR2497559](#)
- CHEN, J., GUPTA, A. K., 2000. Parametric statistical change point analysis. Birkhauser. [MR1761850](#)
- CHO, H., FRYZLEWICZ, P., 2012. Multiscale and multilevel technique for consistent segmentation of nonstationary time series. *Statistica Sinica* 22, 207–229. [MR2933173](#)
- CSORGO, M., HORVÁTH, L., 1997. Limit theorems in change-point analysis. *Wiley series in probability and statistics*. Wiley. [MR2743035](#)
- DAVIS, R. A., LEE, T. C. M., RODRIGUEZ-YAM, G. A., 2006. Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association* 101 (473), 223–239. [MR2268041](#)
- ECKLEY, I. A., FEARNHEAD, P., KILLICK, R., 2011. Analysis of changepoint models. In: Barber, D., Cemgil, T., Chiappa, S. (Eds.), *Bayesian Time Series Models*. Cambridge University Press. [MR2894230](#)
- FRYZLEWICZ, P., NASON, G. P., 2006. Haar-fisz estimation of evolutionary wavelet spectra. *Journal of the Royal Statistical Society (Series B)* 68 (4), 611–634. [MR2301011](#)
- GOMBAY, E., 2008. Change detection in autoregressive time series. *Journal of Multivariate Analysis* 99 (3), 451–464. [MR2396974](#)
- GOTT, A., ECKLEY, I., 2013. A note on the effect of wavelet choice on the estimation of the evolutionary wavelet spectrum. *Communications in statistics - simulation and computation* 42, 393–406.
- HARCHAOUI, Z., LEVY-LEDUC, C., 2010. Multiple change-point estimation with a total variation penalty. *Journal of the American Statistical Association* 105 (492), 1480–1493. [MR2796565](#)
- HIDGKINS, G. A., JAMES, I. C., HUNTINGTON, T. G., 2002. Historical changes in lake ice-out dates as indicators of climate change in new england, 1850–2000. *International Journal Climatology* 22, 1819–1827.

- KILLICK, R., ECKLEY, I. A., 2010. *changeoint: Analysis of Changepoint Models*. Lancaster University, Lancaster, UK.
URL <http://CRAN.R-project.org/package=changeoint>
- KILLICK, R., FEARNHEAD, P., ECKLEY, I. A., 2012. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association* 107 (500), 1590–1598.
- LAVIELLE, M., 2005. Using penalized contrasts for the change-point problem. *Signal Processing* 85, 1501–1510.
- NASON, G. P., 2008. *Wavelet Methods in Statistics with R*. Springer, New York. [MR2445580](#)
- NASON, G. P., VON SACHS, R., KROISANDT, G., 2000. Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *Journal of the Royal Statistical Society (Series B)* 62 (2), 271–292. [MR1749539](#)
- OMBAO, H. C., RAZ, J. A., VON SACHS, R., MALOW, B. A., 2001. Automatic Statistical Analysis of Bivariate Nonstationary Time Series. In *Memory of Jonathan A. Raz*. *Journal of the American Statistical Association* 96 (454), 543–560. [MR1946424](#)
- PERCIVAL, D. B., WALDEN, A. T., 2006. *Wavelet Methods for Time Series Analysis*. Cambridge University Press. [MR2218866](#)
- SCHNABEL, R. B., ESKOW, E., 1999. A revised modified cholesky factorization algorithm. *SIAM J. Optim* 9, 1135–1148. [MR1724780](#)
- SCOTT, A. J., KNOTT, M., 1974. A cluster analysis method for grouping means in the analysis of variance. *Biometrics* 30 (3), 507–512.
- VOSTRIKOVA, L. J., 1981. Detecting disorder in multidimensional random processes. *Soviet Mathematics Doklady* 24, 55–59.