Detecting and estimating changepoints in nonlinear autoregressive models using simulated data

Echarif Elharfaoui [1] & Mohamed Salah Eddine Arrouch [2]

[1,2] *Université Chouaib Doukkali, Faculté des Sciences, 24000 El Jadida, Maroc*
elharfaoui.e@ucd.ac.ma, arrouch.m@ucd.ac.ma

Résumé. Cet article présente une étude par simulation de la détection de ruptures des modèles paramétriques conditionnels hétéroscédastiques non linéaires autorégressifs (CHARN). Une expérience de simulation est réalisée et appliquée à des ensembles de données réelles.

Mots-clés. modèles CHARN, moindres carrés conditionnels, ruptures, tests

Abstract. This paper presents a simulation study of change-point detection of parametric conditional heteroscedastic autoregressive nonlinear (CHARN) models. A simulation experiment is carried and applied to the sets of real data.

Keywords. CHARN models, conditional least-squares, change-points, tests

1 Introduction

Volatility models play a crucial role in the analysis of time series because they provide an overview of the degree of uncertainty and risk associated with a set of data. It refers to the phenomenon in which the variance of the terms of error in a time series model is not constant in time, which can complicate our statistical analyzes.

Detection of heteroscedasticity is the first crucial step to remedy it. Various methods to detect and discuss this volatility model. The study of conditional variations in financial and economic data receives particular attention as a result of its interest in hedging strategies and risk management. These models are of the most famous and significant ones in finance, which include many financial time series models. We suggest a hybrid estimation procedure, which combines CLS and non-parametric methods to estimate the change location. Indeed, conditional least-squares estimators own a computational advantage and require no knowledge of the innovation process.

The literature on change-points is vast. A popular alternative to using the likelihood ratio test was employed in Hinkley (1970,1972). Yao and Davis (1986) reviewed the asymptotic behavior of likelihood ratio statistics for testing a change in the mean in a series of iid Gaussian random variables. Csörgő and Horváth (1987) came up with statistics based on linear rank statistical processes with quantum scores. Chen and Gupta (1999) looked at detection and autoregressive parameters of a p-order autoregressive model.

There are several types of changes depending on the temporal behavior of the series studied. The usual ones are abrupt change, gradual change, and intermittent change. In this paper, we focus on abrupt change in the conditional variance of *off-line* data issue from a class of CHARN models (see Härdle and Tsybakov (1997) and Härdle et al. (1998)). These models are of the most famous and significant ones in finance, which include many financial time series models.

The rest of the paper is organized as follows. Section 2 presents the definitions and concepts, the main assumptions on the CLS estimators of the parameters. Section 3 presents the simulation results from a few simple time series models. They are applied on real data sets.

2 Definitions and concepts

2.1 Model and Conditional Least-Squares Estimation

We place ourselves in the framework where the observations at hand are assumed to be issued from the following CHARN (p,p) model :

$$X_t = m\left(\rho; Z_{t-1}\right) + \sigma\left(\theta; Z_{t-1}\right)\varepsilon_t,\ t \in \mathbb{Z}, \tag{1}$$

where $p \in \mathbb{N}^* \bigcup \{\infty\}$; $m(\cdot)$ and $\sigma(\cdot)$ are two real-valued functions of known forms depending on unknown parameters ρ and θ, respectively ; for all $t \in \mathbb{Z}$, $Z_{t-1} = (X_{t-1}, X_{t-2}, \ldots, X_{t-p})^\intercal$; $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a sequence of stationary random variables with $\mathbb{E}\left(\varepsilon_t \mid Z_{t-1}\right) = 0$ and $\mathbb{V}\left(\varepsilon_t \mid Z_{t-1}\right) = 1$ such that ε_t is independent of the σ-algebra $\mathcal{F}_{t-1} = \sigma\left(Z_k,\ k < t\right)$. The case $p = \infty$ is treated in Bardet and Kengne (2014) where the stationarity and the ergodicity of the process $(X_t)_{t\in\mathbb{Z}}$ is studied. Although we restrict to $p < \infty$, all the results stated here also hold for $p = \infty$.

Let $\psi = (\rho^\intercal, \theta^\intercal)^\intercal \in \Psi = \text{int}(\Theta) \times \text{int}(\widetilde{\Theta}) \subset \mathbb{R}^r \times \mathbb{R}^l$, the vector of the parameters of the model (1) and $\psi_0 = (\rho_0^\intercal, \theta_0^\intercal)^\intercal$ the true parameter vector. Denote by $\|M\|$ an appropriate norm of a vector or a matrix M. We assume that all the random variables in the whole text are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, assuming, the common fourth order moment of the ε_t is finite and regularity conditions for $m(\cdot)$ and $\sigma(\cdot)$.
The conditional mean and the conditional variance of X_t are given, respectively, by

$$\mathbb{E}\left(X_t \mid \mathcal{F}_{t-1}\right) = m\left(\rho; Z_{t-1}\right) \text{ and } \mathbb{V}\left(X_t \mid \mathcal{F}_{t-1}\right) = \sigma^2\left(\theta; Z_{t-1}\right).$$

From these, one has that for all $z \in \mathbb{R}^p$,

$$\mathbb{E}\left(X_1 \mid Z_0 = z\right) = m\left(\rho; z\right) \quad \text{and} \quad \mathbb{E}\left(\left(X_1 - m\left(\rho; Z_0\right)\right)^2 \mid Z_0 = z\right) = \sigma^2\left(\theta; z\right).$$

Therefore, for any bounded measurable functions $g(\cdot)$ and $k(\cdot)$, we have

$$\mathbb{E}\left\{\left[X_1 - m\left(\rho; Z_0\right)\right] g(Z_0)\right\} = 0 \quad \text{and} \quad \mathbb{E}\left(\left\{\left[X_1 - m\left(\rho; Z_0\right)\right]^2 - \sigma^2\left(\theta; Z_0\right)\right\} k(Z_0)\right) = 0.$$

Without a loss of generality, in the following we take, for all $z \in \mathbb{R}^p$, $g(z) = k(z) = 1$.
Now, given $X_{-p+1}, \ldots, X_{-1}, X_0, X_1, \ldots, X_n$ with $n \gg p$, we let $\mathbb{X}_n = (X_{-p+1}, \ldots, X_{-1}, X_0, X_1, \ldots, X_n)$ and consider the sequences of random functions

$$\begin{aligned}
Q_n(\rho) = Q_n(\rho; \mathbb{X}_n) &= \sum_{t=1}^{n} \left(X_t - \mathbb{E}\left(X_t \mid \mathcal{F}_{t-1}\right)\right)^2 = \sum_{t=1}^{n} \left(X_t - m\left(\rho; Z_{t-1}\right)\right)^2 \\
S_n(\rho, \theta) = S_n(\rho, \theta; \mathbb{X}_n) &= \sum_{t=1}^{n} \left(\left(X_t - m\left(\rho; Z_{t-1}\right)\right)^2 - \sigma^2\left(\theta; Z_{t-1}\right)\right)^2.
\end{aligned}$$

Under regularity conditions, we have

- $\dfrac{\partial Q_n}{\partial \rho}(\widehat{\rho}_n;\mathbb{X}_n)=0$ and $Q_n(\rho;\mathbb{X}_n)$ attains a relative minimum at $\rho=\widehat{\rho}_n$;
- assuming $\widehat{\rho}_n$ fixed, $\dfrac{\partial S_n}{\partial \theta}\left(\widehat{\psi}_n;\mathbb{X}_n\right)=0$ and $S_n\left((\widehat{\rho}_n,\theta);\mathbb{X}_n\right)$ attains a relative minimum at $\theta=\widehat{\theta}_n$.

2.2 Change-Point Test and Change Location Estimation

We essentially use the techniques of Bai (1994), who studied the estimation of the shift in the mean of a linear process by a LS method. We first consider the model (1) for known ρ, and $\sigma(\theta;Z_{t-1})=\underline{\theta}\delta_0(Z_{t-1})$, for some known positive real-valued function $\delta_0(\cdot)$ defined on $\mathbb{R}^p$ and for an unknown positive real number $\underline{\theta}$. We wish to test $H_0:\underline{\theta}=\vartheta_1=\vartheta_2$ over $t\leqslant n$ against

$$H_1:\underline{\theta}=\begin{cases}\vartheta_1, & t=1,\ldots,t^*\\ \vartheta_2, & t=t^*+1,\ldots,n.\end{cases}\qquad(\vartheta_1\neq\vartheta_2)$$

where $\vartheta_1,\ \vartheta_2$ and t^* are unknown parameters.

We are also interested in estimating $\vartheta_1,\ \vartheta_2$ and the change location t^*, when H_0 is rejected. It is assumed that $t^*=[n\tau]$ for some $\tau\in(0,1)$, with $[x]$ standing for the integer part of any real number x. From (1), one can easily check that

$$\left(X_t-m(\rho;Z_{t-1})\right)^2=\delta^2(Z_{t-1})+\delta^2(Z_{t-1})\left(\varepsilon_t^2-1\right),\ t\in\mathbb{Z}\tag{2}$$

from which we define the LS estimator $\widehat{t^*}$ of t^* as follows :

$$\widehat{t^*}:=\underset{1\leqslant k<n}{\arg\min}\left[\min_{\vartheta_1,\vartheta_2}\left\{\sum_{t=1}^{k}\left(W_t^2-\vartheta_1^2\right)^2+\sum_{t=k+1}^{n}\left(W_t^2-\vartheta_2^2\right)^2\right\}\right],\tag{3}$$

where $W_t=\left(X_t-m(\rho;Z_{t-1})\right)/\delta_0(Z_{t-1})$. Thus, the change location is estimated by minimizing the sum of squares of residuals among all possible sample slits. Letting

$$\overline{W}_k=\frac{1}{k}\sum_{t=1}^{k}W_t^2,\ \overline{W}_{n-k}=\frac{1}{n-k}\sum_{t=k+1}^{n}W_t^2\text{ and }\overline{W}=\frac{1}{n}\sum_{t=1}^{n}W_t^2,$$

it is easily seen that for some k, the LS estimator of $\vartheta_1^2\,(t\leqslant k)$ and $\vartheta_2^2\,(t>k)$ are $\overline{W}_k$ and $\overline{W}_{n-k}$, respectively, and that (3) can be written as

$$\widehat{t^*}=\underset{1\leqslant k<n}{\arg\min}\left\{\sum_{t=1}^{k}\left(W_t^2-\overline{W}_k\right)^2+\sum_{t=k+1}^{n}\left(W_t^2-\overline{W}_{n-k}\right)^2\right\}=\underset{1\leqslant k<n}{\arg\min}\,S_k^2.\tag{4}$$

Let $S^2=\sum_{t=1}^{n}\left(W_t^2-\overline{W}\right)^2$. A simple algebra gives

$$S^2=S_k^2+U_k,\tag{5}$$

where

$$U_k=k\left(\overline{W}_k-\overline{W}\right)^2+(n-k)\left(\overline{W}_{n-k}-\overline{W}\right)^2.\tag{6}$$

From (4) and (5), we have

$$\widehat{t^*} = \underset{1\leqslant k<n}{\arg\min}\ \left(S^2 - U_k\right) = \underset{1\leqslant k<n}{\arg\max}\, U_k. \tag{7}$$

From (6), a simple algebraic computation gives the following alternative expression for U_k :

$$U_k = \frac{n}{k(n-k)}\left(\sum_{t=1}^{k}\left(W_t^2 - \overline{W}\right)\right)^2 = \left(\sqrt{\frac{n}{k(n-k)}}\sum_{t=1}^{k}\left(W_t^2 - \overline{W}\right)\right)^2 = T_k^2. \tag{8}$$

It results from (7) and (8) that

$$\widehat{t^*} = \underset{1\leqslant k<n}{\arg\max}\, T_k^2 = \underset{1\leqslant k<n}{\arg\max}\, |T_k|. \tag{9}$$

Writing $T_k^2 = n\Delta_k^2$, it is immediate that

$$\Delta_k^2 = \frac{1}{k(n-k)}\left(\sum_{t=1}^{k}\left(W_t^2 - \overline{W}\right)\right)^2 = \frac{k}{(n-k)}\left(\overline{W}_k - \overline{W}\right)^2.$$

Simple computations give $\Delta_k^2 = \dfrac{k(n-k)}{n^2}\left(\overline{W}_{n-k} - \overline{W}_k\right)^2$, from which we have

$$\widehat{t^*} = \underset{1\leqslant k<n}{\arg\max}\, \Delta_k^2 = \underset{1\leqslant k<n}{\arg\max}\ |\Delta_k|\,. \tag{10}$$

The test statistic we use for testing H_0 against H_1 is a scale version of $\max\limits_{1\leqslant k\leqslant n-1}|T_k|$.

One can observe that under some conditions (e.g., ε_t i.i.d. with $\varepsilon_t \sim \mathcal{N}(0,1)$), this statistic is the equivalent likelihood based test statistic for testing H_0 against H_1 (see, e.g., Hawkins (1977)). Let

$$C_k = \sum_{t=1}^{k} W_t^2,\quad C_{n-k} = \sum_{t=k+1}^{n} W_t^2 \ \text{ and }\ C_n = \sum_{t=1}^{n} W_t^2. \tag{11}$$

By simple calculations, we obtain

$$T_k = \sqrt{\frac{n}{k(n-k)}}\sum_{t=1}^{k}\left(W_t^2 - \overline{W}\right) = \left(q\left(\frac{k}{n}\right)\right)^{-1}\left(\frac{1}{\sqrt{n}}\left(C_k - \frac{k}{n}C_n\right)\right), \tag{12}$$

where $q(\cdot)$ is a positive weight function defined for any $x \in (0,1)$ by $q\,(x) = \sqrt{x\,(1-x)}$.

The study of the asymptotic distribution of the test statistic under H_0, is based on that of the the process $\xi_n(\cdot)$ defined for any $s \in [0,1]$ by $\xi_n(s) = C_n(s) - sC_n(1)$,

For $\delta \in (0,1/2)$, we denote by $D_\delta \equiv D\left([\delta, 1-\delta]\right)$ the space of all right continuous functions with left limits on $[\delta, 1-\delta]$ endowed with the Skorohod metric. It is clear that $C_n(\cdot), \xi_n(\cdot) \in D_0$ and $T_n(\cdot) \in D_\delta$.
For the study of the estimator $\widehat{t^*}$, we let $\kappa = \kappa_n = \vartheta_2^2 - \vartheta_1^2$ and assume without loss of generality that $\kappa_n > 0$ $(\vartheta_2 > \vartheta_1)$, $\kappa_n \longrightarrow 0$ as $n \longrightarrow \infty$ and that the unknown change point t^* depends on the sample size n. We have the following result :

Theorem 1 *Assume that mixing is satisfied,* $t^*/n \in (a, 1-a)$ *for some* $0 < a < 1/2$, $t^* = [n\tau]$ *for some* $\tau \in (0,1)$ *and as* $n \longrightarrow \infty$, $\kappa_n \longrightarrow 0$ *and* $\dfrac{\kappa_n\sqrt{n}}{\sqrt{\ln n}} \longrightarrow \infty$. *Then, we have*

$$\widehat{t^*} - t^* = O_{\mathbb{P}}\Big(\frac{1}{\kappa_n^2}\Big),$$

3 Practical Consideration

In this section we perform numerical simulations to evaluate the performances of our methods and these are applied to two sets of real data. We start with the presentation of the results of numerical simulations found with the software R, version 4.1.1. The trials are based on 1000 replications of observations of lengths $n = 500$, 1000, 5000 and 10000 generated from the model (1) for $\rho = (\rho_0, \rho_1, \rho_2)^\intercal$; $\theta = (\theta_0, \theta_1, \theta_2)^\intercal$; $m(\rho; x) = (\rho_0 + \rho_1 \exp(-\rho_2 x^2))\, x$; $\sigma(\theta; x) = \underline{\theta}\delta_0(x)$ with $\delta_0(x) = \sqrt{\theta_0^2 + \theta_1^2 x^2 \exp(-\theta_2 x^2)}$; $\rho_2 > 0$, $\rho_0\rho_1 \geqslant 0$, $\theta_2 \geqslant 0$ and $0 < \underline{\theta}^2\theta_1^2 < 1$; $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a white noise with density function f. We also assume the sufficient condition $|\rho_0| + |\rho_1| + |\underline{\theta}\theta_1| + 2\rho_0\rho_1 < 1$, to ensure the strict stationarity and ergodicity of the process $(X_t)_{t\in\mathbb{Z}}$ (see, e.g., Theorem 3.2.11 of Taniguchi and Kakizawa (2000), p. 86 and Ngatchou-Wandji (2005), p. 5). The noise densities f that we employed were Gaussian.
The change-point location is estimated using the following algorithm :

Algorithm 1 Change-point location estimation

for $i = 1, \ldots, 1000$ **do**
  for $t = 1, \ldots, n$ **do** $W_t = (X_t - m(\rho; X_{t-1}))\,/\delta_0(X_{t-1})$
  end for
  $\overline{W} = \frac{1}{n}\sum_{t=1}^{n} W_t^2$
  for $k = 1, \ldots, n-1$ **do** $T_k = \sqrt{\frac{n}{k(n-k)}}\sum_{t=1}^{k}\left(W_t^2 - \overline{W}\right)$
  end for
  Compute $\widehat{t}_i^* = \arg\max_{1\leqslant k<n} |T_k|$ (that is the value of k for which $|T_k|$ is the largest)
end for
Compute $L = \left(\widehat{t}_1^* + \widehat{t}_2^* + \ldots + \widehat{t}_{1000}^*\right)/1000$
Change-point location estimation is given by $\widehat{t}^* = [L]$, the integer part of L

Example. We consider the model (1) for $\rho_0 = \rho_1 = 0$, $\theta_2 = 0$, $\delta_0(X_{t-1}) = \sqrt{0.04 + 0.36X_{t-1}^2}$, $\vartheta_1 = 1$, $\vartheta_2 = 1 + \phi$ and $\varepsilon_t \sim \mathcal{N}(0, 1)$. The resulting model is an ARCH(1). The change location estimators are calculated for $\phi = 0.3$, 0.8 and 1.5 at the locations $t^* = \tau \times n$ for $\tau = 0.25$, 0.5 and 0.75. In each case, we compute the bias and the standard error SE (SE=SD/$\sqrt{n}$), where SD denotes the standard deviation) of the change location estimator. Table 1 shows that the bias declines rapidly as ϕ increases. Also, as the sample size n increases, the bias and the SE decrease. This tends to show the consistency of $\widehat{t}^*$, as expected from the asymptotic results.
We also consider the case $\varepsilon_t = \beta\varepsilon_{t-1} + \gamma_t$, where $|\beta| < 1$ and $\gamma_t \sim \mathcal{N}(0, \sqrt{1-\beta^2})$. It is easy to check that with this $(\varepsilon_t)_{t\in\mathbb{Z}}$ is stationary and strongly mixing, and that $\mathbb{E}(\varepsilon_t) = 0$ and $\mathbb{V}(\varepsilon_t) = 1$. In this case, we only study the SE for $n = 5000$, 10000 and the results are compared to those obtained for $\varepsilon_t \sim \mathcal{N}(0, 1)$, for the same values of ϕ as above but for $\tau = 0.25$ and 0.75. These results listed in Table 2 show that for $\varepsilon_t \sim \mathcal{N}(0, 1)$, the location estimator is more accurate and the SE decreases slightly compared to the case $\varepsilon_t \sim$ AR(1). It seems from these results that the nature of the white noise ε_t does not much affect the location estimator for larger values of n and ϕ.

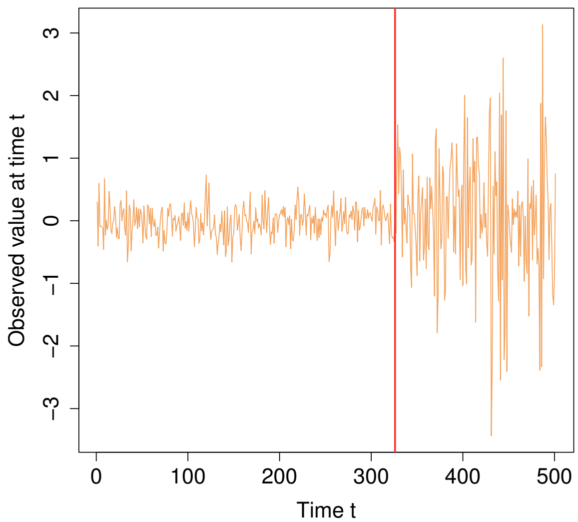
TABLE 1 – Change location estimation, its bias and SE for several values of ϕ, n and τ for iid $\varepsilon_t \sim \mathcal{N}(0,1)$.

ϕ	n	$t^* = 0.25n$ ($\tau = 0.25$)			$t^* = 0.5n$ ($\tau = 0.5$)			$t^* = 0.75n$ ($\tau = 0.75$)		
		$\widehat{t^*}$	SE	Bias	$\widehat{t^*}$	SE	Bias	$\widehat{t^*}$	SE	Bias
0.3	500	181	4.9667	0.1120	277	3.4284	0.0540	384	3.3993	0.0180
	1000	287	3.8961	0.0370	522	2.4946	0.0220	767	2.8024	0.0170
	5000	1264	0.6270	0.0028	2516	0.6260	0.0032	3765	0.8172	0.0030
	10000	2517	0.4398	0.0017	5015	0.4131	0.0015	7515	0.4701	0.0015
0.8	500	137	1.8286	0.0240	258	0.9659	0.0160	383	1.0514	0.0160
	1000	257	0.5079	0.0070	507	0.6687	0.0070	757	0.6378	0.0070
	5000	1256	0.1874	0.0012	2506	0.1750	0.0012	3755	0.1602	0.0010
	10000	2506	0.1230	0.0006	5006	0.1388	0.0006	7505	0.1169	0.0005
1.5	500	130	0.8538	0.0100	254	0.4724	0.0080	379	0.4611	0.0080
	1000	253	0.2662	0.0030	504	0.2884	0.0040	753	0.2562	0.0030
	5000	1254	0.1344	0.0008	2503	0.1053	0.0006	3754	0.1174	0.0008
	10000	2504	0.0880	0.0004	5004	0.0842	0.0004	7504	0.0753	0.0004

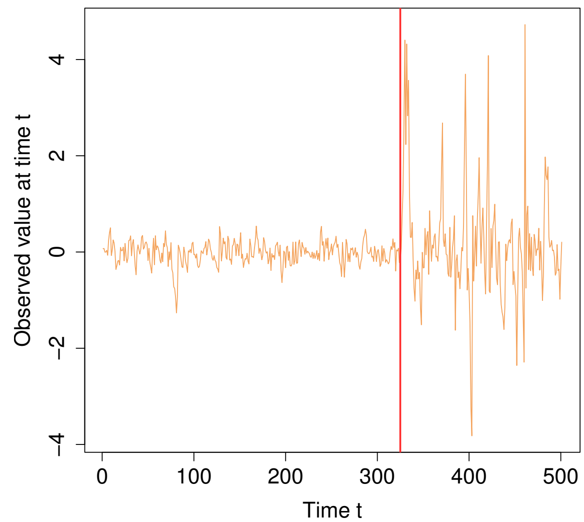
TABLE 2 – Change location estimation, its bias and SE for several values of ϕ, n and τ for iid $\varepsilon_t \sim \mathcal{N}(0,1)$ and for $\varepsilon_t \sim$ AR(1).

ϕ	n	$t^* = 0.25n$ ($\tau = 0.25$)						$t^* = 0.75n$ ($\tau = 0.75$)					
		$\varepsilon_t \sim \mathcal{N}(0,1)$			$\varepsilon_t \sim$ AR(1)			$\varepsilon_t \sim \mathcal{N}(0,1)$			$\varepsilon_t \sim$ AR(1)		
		$\widehat{t^*}$	SE	Bias	$\widehat{t^*}$	SE	Bias	$\widehat{t^*}$	SE	Bias	$\widehat{t^*}$	SE	Bias
0.3	5000	1265	0.6091	0.0030	1286	2.6225	0.0072	3766	0.6596	0.0032	3776	1.2243	0.0052
	10000	2516	0.4471	0.0016	2525	0.7136	0.0025	7515	0.4182	0.0015	7523	0.6563	0.0023
0.8	5000	1256	0.1701	0.0012	1260	0.2760	0.0020	3756	0.1818	0.0012	3760	0.2870	0.0020
	10000	2506	0.1482	0.0006	2510	0.1907	0.0010	7506	0.1338	0.0006	7509	0.1835	0.0009
1.5	5000	1254	0.1165	0.0008	1256	0.1835	0.0012	3754	0.1154	0.0008	3756	0.1784	0.0012
	10000	2503	0.0807	0.0003	2506	0.1284	0.0006	7504	0.0776	0.0004	7506	0.1263	0.0006

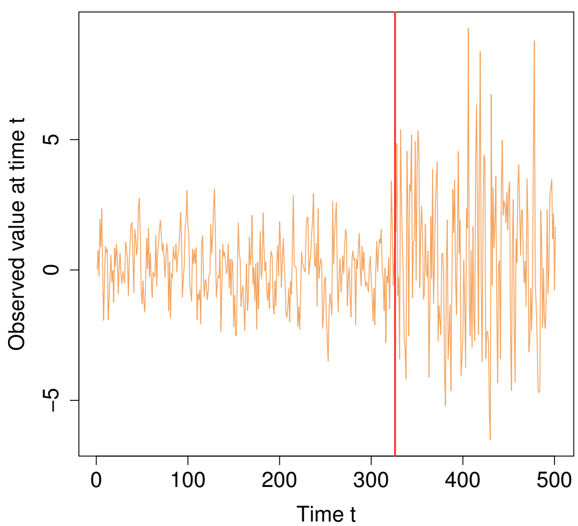
We present two graphs showing a change in volatility at a time $\widehat{t^*}$. This is indicated by a vertical red line on both graphics where one can easily see the evolution of the time series before and after the change location estimator $\widehat{t^*}$. The series in both figures are obtained for $m(\rho; x) = 0$, $\delta_0(x) = \sqrt{1 + 0.036x^2}$, $n = 500$, $\tau = 0.65$, and $\phi = 0.8$. That in Figure 1a is obtained for standard iid Gaussian ε_ts. In this case, using our method, the change location $t^* = 0.65 \times 500 = 325$ is estimated by $\widehat{t^*} = 326$. The time series in Figure 1b is obtained with $\varepsilon_t \sim$ AR(1). In this case, t^* is estimated by $\widehat{t^*} = 325$.



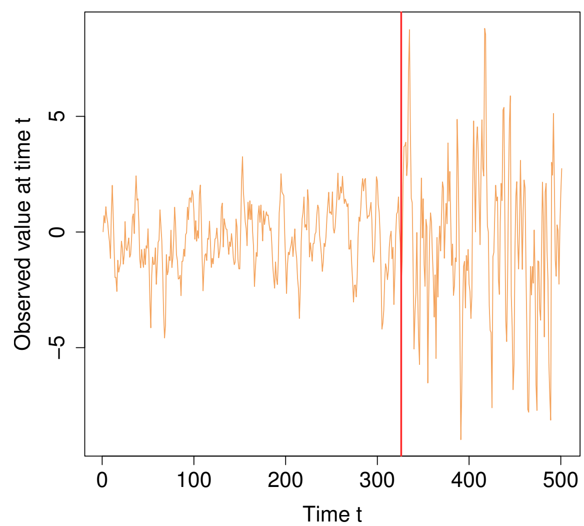
($\mathbf{a}$) ARCH(1) model, $\widehat{t}^* = 326$



($\mathbf{b}$) ARCH(1) model, $\widehat{t}^* = 325$



($\mathbf{c}$) CHARN model, $\widehat{t}^* = 326$



($\mathbf{d}$) CHARN model, $\widehat{t}^* = 326$

FIGURE 1 – Estimation of change-point in volatility for 500 observations. ($\mathbf{a}$) ARCH(1) model with change point at $\widehat{t}^* = 326$; ($\mathbf{b}$) ARCH(1) model with change point at $\widehat{t}^* = 325$; ($\mathbf{c}$) CHARN model with change point at $\widehat{t}^* = 326$; ($\mathbf{d}$) CHARN model with change point at $\widehat{t}^* = 326$.

3.1 Comparison with Some Recent Algorithms

We compare our method, referred to as LS, with the Wild Binary Segmentation (WBS) method studied in Fryzlewicz (2014), and one of its variants, called Narrowest-Over-Threshold (NOT), proposed by Baranowski et al. (2019), as well as, the Iterative Cumulative Sum of Squares (ICSS) algorithm suggested by Inclan and Tiao (1994). All these methods are implemented under R software, and can, respectively, be found in the packages *wbs*, *not* and *ICSS*. Our comparison is based on n observations simulated from (1) for $\rho_0 = \rho_1 = 0$, $\theta_2 = 0$, $\delta_0(X_{t-1}) = \sqrt{0.04 + 0.36X_{t-1}^2}$, $\vartheta_1 = 1$, $\vartheta_2 = 1 + \phi$ and $\varepsilon_t \sim \mathcal{N}(0, 1)$. The change location estimators are calculated for $\phi = 0.3$, 0.8 and 1.5 at the locations $t^* = \tau \times n$ for $\tau = 0.25$ and 0.75.

From the results obtained (see Table 3), WBS, NOT and ICSS generate sequences of change-point location estimates, some of which have values close to the true locations. For $n = 100$ and $n = 200$, LS generally provides more accurate estimates $\widehat{t^*}$ of the true change-point location than WBS, NOT and ICSS, for different ϕ and locations t^*. Among all, it is generally the best, especially for larger values of n and ϕ.

Table 3 – Estimates of change location derived from LS, WBS, NOT and ICSS for a sample with a single break.

		$\phi = 0.3$		$\phi = 0.8$		$\phi = 1.5$	
		$t^* = 0.25n$	$t^* = 0.75n$	$t^* = 0.25n$	$t^* = 0.75n$	$t^* = 0.25n$	$t^* = 0.75n$
Methods	n	$\widehat{t^*}$	$\widehat{t^*}$	$\widehat{t^*}$	$\widehat{t^*}$	$\widehat{t^*}$	$\widehat{t^*}$
LS	100	39	70	31	79	28	78
WBS		46|43	87|90	38|39	82|80	29|30	78|76
NOT		51|55	96	38|42	79|82|91|94	38|35	78|82
ICSS		42|48	68	31|43	78|81	33|41	78|97
LS	200	57	145	54	157	52	155
WBS		70|66	175|177	62|59	171|167	58|57	159|158
NOT		66|70|77	175|190|194	67|70	158|161	64|68	155|162
ICSS		66|87	80|159	61|76|87	159	59|66|170|187	156|173

3.2 Application to Real Data

In this section, we apply our procedure to two sets of genuine time series, namely, the USA stock market prices. These data of length 2022 from the American stock market were recorded daily from 2 January 1992 to 31 December 1999. They represent the daily stock prices of the S&P 500 stock market (SPX). They are among the most closely followed stock market indices in the world and are considered as an indicator of the USA economy. They have also been recently examined by Kouamo et al. (2010) and can be found at `www.investing.com`. In Figure 2, we observe that the trend of the SPX daily stock price series is not constant over time. We also observe that stock prices have fallen sharply, especially in the time interval between 1997 and 1998.

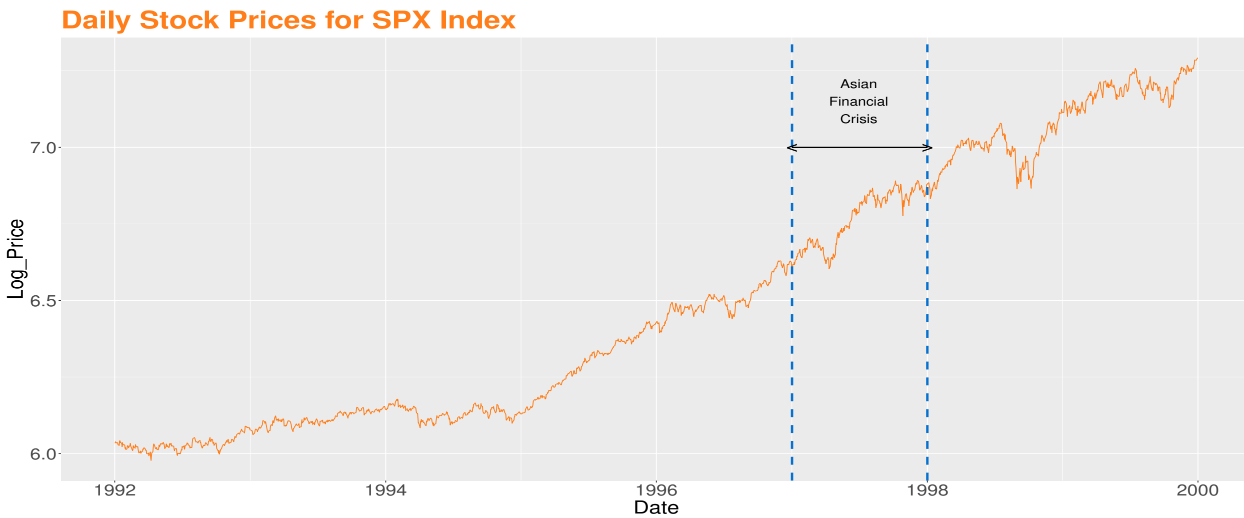


FIGURE 2 – Logarithmic series of S&P 500 stock prices from January 1992 to December 1999.

Denote by D_t the value of the stock price for the SPX index at day t, and the first difference of the logarithm of stock price, X_t as : $X_t = \log(D_t) - \log(D_{t-1}) = \log\left(\frac{D_t}{D_{t-1}}\right)$.
X_t is the logarithmic return of stock price for the SPX index at day t.
The series (X_t) is approximately piece-wise stationary on two segments and symmetric around zero (see Figure 3). This brought us to consider a CHARN model with $m(\rho; x) = 0$, $\delta_0(x) = \sqrt{\theta_0 + \theta_1 x^2}$ for $\theta_0 = 1,\ \theta_1 = 0$, ϑ_1 and ϑ_2 estimated by CLS described in Section 2.1. Using our procedure, we found an important change point in stock price volatility on 26 March 1997, which is consistent with the date found by Kouamo et al. (2010) (see Figure 3).

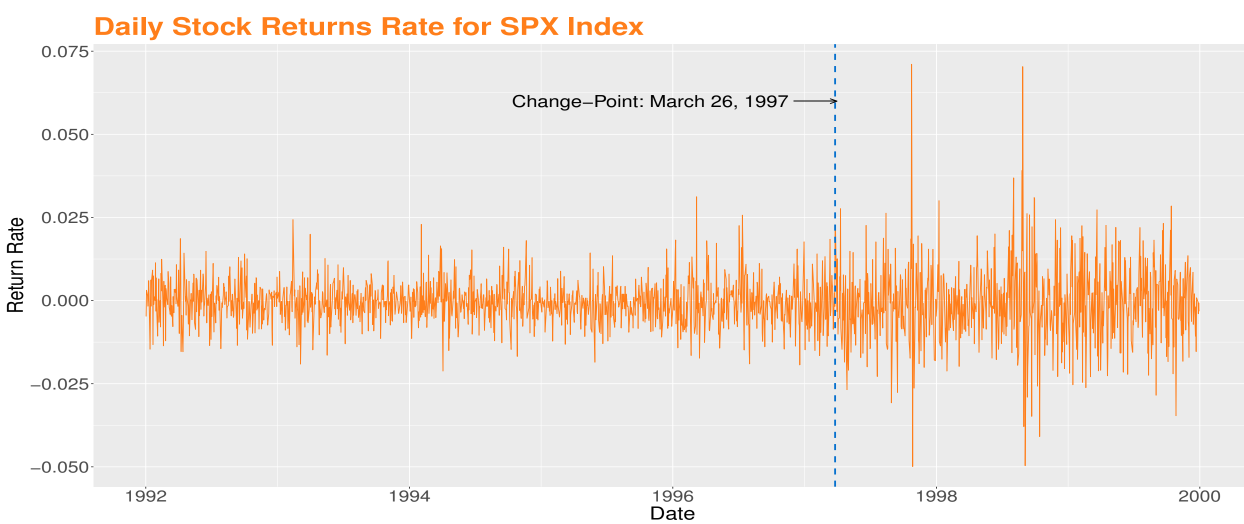


FIGURE 3 – Location of the change point in the volatility of the logarithmic stock price return series of the SPX Index from January 1992 to December 1999.

It should be noted that the change in volatility coincides with the Asian crisis in 1997 when

Thailand devalued its currency, the baht, against the US dollar. This decision led to a fall in the currencies and financial markets of several countries in its surroundings. The crisis then spread to other emerging countries with important social and political consequences and repercussions on the world economy.

Bibliography

Bai, J. (1994), Least squares estimation of a shift in linear processes, *Journal of Time Series Analysis*, 15, pp. 453-472.

Baranowski, R., Chen, Y. and Fryzlewicz, P. (2019), *Narrowest-over-Threshold Change-Point Detection*, R package version 1, CRAN : Vienna, Austria.

Chen, J. and Gupta, A. (1999), Change point analysis of a Gaussian model, *Statistical Papers*, 40, pp. 323-333.

Csörgő, M. and Horváth, L. (1987), Nonparametric tests for the changepoint problem, *Journal of Statistical Planning and Inference*, 17, pp. 1-9.

Fryzlewicz, P. (2014), Wild Binary Segmentation for multiple change-point detection, *Annals of Statistics*, 42, pp. 2243-2281.

Hawkins, D.M. (1977), Testing a sequence of observations for a shift in location, *Journal of American Statistical Association*, 72, pp. 180-186.

Härdle, W. and Tsybakov, A. (1997), Local polynomial estimators of the volatility function in nonparametric autoregression, *Journal of Econometrics*, 81, pp. 223-242.

Härdle, W., Tsybakov, A. and Yang, L. (1998), Nonparametric vector autoregression, *Journal of Statistical Planning and Inference*, 68, pp. 221-245.

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

Hinkley, D.V. (1972), Time-ordered classification, *Biometrika*, 59, pp. 509-523.

Inclan, C. and Tiao, G.C. (1994), Use of cumulative sums of squares for retrospective detection of changes of variance, *Journal of American Statistical Association*, 89, pp. 913-923.

Kouamo, O., Moulines, E. and Roueff, F. (2010), Testing for homogeneity of variance in the wavelet domain, *Dependence in Probability and Statistics*, 200, pp. 175-205.

Ngatchou-Wandji, J. (2005), Checking nonlinear heteroscedastic time series models, *Journal of Statistical Planning and Inference*, 133, pp. 33-68.

Taniguchi, M. and Kakizawa, Y. (2000), *Asymptotic theory of estimation and testing for stochastic processes*, In Asymptotic Theory of Statistical Inference for Time Series, Springer : New York, NY, USA, pp. 51-165.

Yao, Y.C. and Davis, R.A. (1986), The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates, *Sankhyā : The Indian Journal of Statistics, Series A*, pp. 339-353.