Switching regime and ARFIMA processes

Luisa Bisaglia, Margherita Gerolimetto *

April 28, 2005

Abstract

The phenomena of long range dependence and regime switching are intimely related and it is very difficult to separate the two effects. In this paper we consider the problem of inference on the order of integration in presence of infrequent structural breaks. For this purpose we consider the most important estimators of long memory parameters used in literature and compare their performances in finite samples via Monte Carlo experiments.

key words: Long memory, Occasional structural breaks, Regime switching, Markov switching

1 Introduction

Since the seminal papers of Granger and Joyeux (1980) and Hosking (1981) there has been a considerable interest in modelling strong persistence in time series. This interest is motivated by the analysis of many empirical time series, such as, for instance, the hydrological time series of Nile River minima (Hurst, 1951) and financial time series (Granger and Ding, 1995, 1996). For all these series the autocorrelation function decreases to zero like a power function rather than exponentially and the spectral density diverges as the frequencies tend to zero.

However, recently, it has been shown that inference on the long memory parameter and persistence tests are severely compromised in series which display occasional breaks, since these processes give the impression of persistence. In other words, neglecting structural breaks causes an over estimation of the long memory parameter, leading the reseacher to believe in a long memory data generating process. Granger and Terasvirta (1999), Granger and Hyung (2004), Diebold and Inoue (2001), Mikosch and Starica (1999) and Gourieroux and Jasiak (2001) provide both theoretical justification and Monte Carlo evidence that models with structural breaks

*Department of Statistics, University of Padova, via C. Battisti 241, 35121 Padova, Italy, bisaglia@stat.unipd.it, gerolime@stat.unipd.it

may exhibit spurious long memory properties and it is troublesome in practice to distinguish between the occasional break process and the long memory process. Indeed, if an occasional break process (for instance a STOBREAK process) is generate and the long memory parameter, d, is estimated, it is likely that the estimation method does not recognize the process as a short memory one with breaks in mean, but it produces instead a value of $\hat{d}$ bigger than zero.

On the other hand, as formally demonstrated by Bai (1998), when the disturbances of a regression model follow an $I(1)$ process there is a tendency to estimate a break point in the middle of the sample, even though a break point does not actually exist. The long memory properties of the DGP might cause many breaks to be detected spuriously by standard estimation methods and unlike $I(0)$ processes, caution should be exercised in estimating break points in presence of $d > 0$.

In this paper, we consider the problem of distinguishing between long memory and swithing regime processes and in particular we are interested in evaluating and comparing the performance of some estimation methods often used in the literature. We orient our interest to the semiparametric GPH method, the parametric pseudo-likelihood Whittle method and some non parametric methods like the Higuchi method, the aggregate variance method and the riscaled range method. We want to find out whether the bad performance exhibited by the long memory tests considered so far in the literature (in particular the GPH) is typical also of the other estimation methods or some of them manage to be somehow more robust in case of structural breaks.

The plan of the paper is as follows. In Section 2, we briefly introduce long memory ARFIMA models and some simple linear models with occasional or infrequent breaks in mean. Section 3 is devoted to Monte Carlo simulations to compare the performance of the various estimation methods used to distinguish between long memory and structural breaks and Section 4 concludes.

2 Long memory ARFIMA and Occasional Structural Break processes

In this Section we want to recall some basic notions about long memory ARFIMA processes and three simple models of stochastic regime switching that we will use in our experiments.

There exists different definitions of long memory processes. In particular, long memory could be expressed either in the time domain or in the frequency domain. In the time domain, a stationary discrete time series is said to be long memory if its autocorrelation function decays to zero like a power function. This definition implies that the dependence between successive observations decays slowly as the number of lags tends to infinity. On the other hand, in the frequency domain, a stationary discrete time series is said to be long memory if its spectral density is unbounded at low frequencies.

Alternatively, the memory of a process can be expressed in terms of the rate of growth of variances of partial sums, $\mathrm{var}(S_T) = O(T^{2d+1})$, where $(S_t) = \Sigma_{t=1}^{t} x_t$ and d is the long memory parameter. These definitions are not completly equivalent but there is a tight connection between them (Beran, 1994).

In this paper we consider one of the most popular long memory processes that is the Autoregressive Fractionally Integrated Moving Average process, ARFIMA(p, d, q) in the following, independently introduced by Granger and Joyeux (1980) and Hosking (1981). This process simply generalizes the usual ARIMA(p, d, q) process by assuming d to be fractional.

Let ϵ_t be a white noise process having $E[\epsilon_t^2] = \sigma^2$. The process $\{X_t,\ t \in \mathbf{Z}\}$ is said to be an ARFIMA(p, d, q) process with $d \in (-1/2, 1/2)$, if it is stationary and satisfies the difference equation

$$\Phi(B)\,\Delta(B)\,(X_t - \mu) = \Theta(B)\,\epsilon_t,$$

where $\Phi(\cdot)$ and $\Theta(\cdot)$ are polynomials in the backward shift operator B of degree p and q, respectively, $\Delta(B) = (1-B)^d = \sum_{j=0}^{\infty} \pi_j B^j$ with $\pi_j = \Gamma(j-d)/[\Gamma(j+1)\Gamma(-d)]$, and $\Gamma(\cdot)$ is the gamma function.

If $p = q = 0$ the process $\{X_t,\ t \in \mathbf{Z}\}$ is called Fractionally Integrated Noise and denoted by $I(d)$. When $d \in (0, 1/2)$ the ARFIMA(p, d, q) process is stationary and the autocorrelation function decays to zero hyperbolically at a rate $O(k^{2d-1})$, where k denotes the lag. In this case we say that the process has a long-memory behavior. When $d \in (-1/2, 0)$ the ARFIMA(p, d, q) process is a stationary process with intermediate memory. In the following we will concentrate on $I(d)$ processes with $d \in (0, 1/2)$: for this range of values the process is stationary, invertible and possesses long-range dependence.

Now we describe some well known structural break processes that we will use in our finite sample experiments.

In all the models we are describing there are only occasional breaks in mean, which means that the number of breaks that can occur in a specific period of time is somehow bounded.

More formally, we assume that the probabiliy of breaks, p, converges to zero slowly as the sample size increases, i.e. $p \to 0$ as $T \to \infty$, yet $\lim_{T\to\infty} Tp$ is a non-zero finite constant.

The first model we consider and describe is the so called mean plus noise or occasional break model (Chen and Tiao, 1990; Engle and Smith, 1999)

$$y_t = m_t + \epsilon_t, \qquad t = 1, ..., T, \tag{1}$$

where ϵ_t is a noise variable and occasional level shifts, m_t, are controlled by two variables q_t (date of breaks) and η_t (size of jump), as

$$m_t = m_{t-1} + q_t \eta_t \tag{2}$$

where η_t is i.i.d. $(0, \sigma_\eta^2)$. In the following sections the distribution of η_t has been taken to be normal although this distribution has no particular relevance. We assume that q_t follows an i.i.d. binomial distribution, that is,

$$q_t = \begin{cases} 0, & \text{with probability } 1-p \\ 1, & \text{with probability } p \end{cases} \tag{3}$$

The binomial model we described is so characterized by sudden changes only. It might be also the case that structural changes occur gradually. In this case a Markov Switching Model (Hamilton, 1989), a simple generalization of Eq. 3, is more adapt. Suppose s_t is a latent random variable that can assume only the two discrete values 0 or 1. Each value of s_t represents a different state in the length of memory of shock. s_t is assumed to be governed by the following Markov probability law: $p_{ij} = Pr\left(s_t = j/s_{t-1} = i\right)$. Then, it is possible to use a switching model of q_t such that $q_t = 0$ when $s_t = 0$ and $q_t = 1$ when $s_t = 1$. In this specification, called Markov switching model, a regime with $s_t = 1$ represents a period of structural change. Therefore, there is structural change everytime $s_t = 1$, both if $s_{t-1} = 1$ and if $s_{t-1} = 0$.

Finally we present the so called Stocastic Permanent Break Model (STOPBREAK model) formulated by Engle and Smith (1999) to bridge the gap between transience and permanence of the shocks. The STOPBREAK is a stocastic process in which the long run impact of each obervation is time varying and stochastic.

The formulation is as follows:

$$y_t = m_t + \epsilon_t \tag{4}$$

$$m_t = m_{t-1} + q_{t-1}\epsilon_{t-1} \tag{5}$$

where $q_t = q\left(|\epsilon_t|\right)$ is non decreasing in $|\epsilon_t|$ and bounded by zero and one, so that bigger innovations have more permanent effects, and ϵ_t are i.i.d $N\left(0, \sigma_\epsilon^2\right)$, moreover $q_t = \frac{\epsilon_t^2}{\left(\gamma+\epsilon_t^2\right)}$ for $\gamma > 0$. Therefore in the STOPBREAK process permanent shocks can be indentified by their larger magnitude. In this approach the effects of shocks can fluctuate between transitory and permanent and typically such data exhibit periods of apparent stationarity punctuated by occasional mean shifts.

The appareance of long memory may be a genuine feature of time series or it could occur when several processes are aggregated (Granger, 1980) or because of the presence of structural breaks and regime switching (Granger and Hyung, 2004; Diebold and Inoue, 2001). In order to discuss briefly how the recent literature has dealt with this issue, we point focus our attention on the paper of Granger and Hyung (2004) who study the sample autocorrelation function of an occasional break model. The authors prove that the sample autocorrelation function of the model (1) can be approximated by the following expression:

$$\hat{\rho}_T(k) \approx \frac{\frac{Tp\sigma_\eta^2}{6}\left(1-\frac{k}{T}\right)\left(1-2\frac{k}{T}+4\left(\frac{k}{T}\right)^2\right)}{\frac{Tp\sigma_\eta^2}{6}+\sigma_\epsilon^2} \tag{6}$$

where $\approx$ denotes approximate equality for any k such that $k/T \to 0$ as T increases. The value of Tp implies an expected number of structural breaks with the sample size T, and σ_η^2 is related to the size of breaks. Moreover, the authors show that for this process, if the probability of breaks p converges to zero slowly as the sample size increases, i.e., $p \to 0$ as $T \to \infty$, then

$$\hat{\rho}_T(k) \to \left(1+\frac{6\sigma_\epsilon^2}{Tp\sigma_\eta^2}\right)^{-1}$$

where $0 < \left(1+\frac{6\sigma_\epsilon^2}{Tp\sigma_\eta^2}\right)^{-1} < 1$. That is the $k-th$ sample autocorrelation converges to nonzero value for any k such that $k/T \to 0$ as T increases. Obviously, when $0 < p < 1$ and p is fixed the break process is $I(1)$ and $\hat{\rho}_T(k) \to 1$ as $T \to \infty$, while when $p = 0$ then there are not break, the process is $I(0)$ and $\hat{\rho}_T(k) \to 0$ as $T \to \infty$.

Note that, although the autocorrelations in eq. (6) decay very slowly as k increases, there is a fondamental difference with the theoretical autocorrelation function of an ARFIMA process.

The autocorrelation function for the occasional break process depends on the sample size T because the process is non stationary, while the ARFIMA process is stationary for the interest value of the parameter, $0 < d < 1/2$. Thus we make a mistake if we calculate the autocorrelation function for an occasional break process because we are considering just linear properties of the data.

Engle and Smith (1999) show that the STOPBREAK model they propose is an approximation to the mean-plus-noise model considered by Granger and Hyung (2004). So, we suppose, the autocorrelation function of this model behaves like eq. (6).

3 Monte Carlo study

In this Section we describe the Monte Carlo study we carried out to show and compare the performance of the different long memory estimation methods when the real DGP is a structural break process. The methods we consider are the R/S range, Higuchi, aggregate variance, GPH, Whittle.

The functions we use are written in R language (Ihaka and Gentlemen, 1996) and are available upon request by the authors.

Simulations were conducted for the following models:

1. DGP1: mean plus noise model (1), with $p = 0.0025$, 0.005, 0.01, 0.05, 0.1 and $\sigma_\eta^2 = 0.005$, 0.01, 0.05, 0.1;

2. DGP2: Markov switching model, with $(p, q) = (0.95, 0.95; 0.95, 0.99; 0.99, 0.95; 0.99, 0.99; 0.999, 0.999)$ and $\sigma_\eta^2 = 0.005$, 0.01, 0.05, 0.1. In this case the initial state s_1 is generated by a Bernoulli random variable with $p = 0.5$;

3. DGP3: STOPBREAK model (4), with $\gamma = (10^{-5}, 10^{-3}, 10^{-1}, 1, 10, 10^2, 10^3)$, following Diebold and Inuoe (2001), in order to make comparisons.

For each model we have considered $\sigma_\epsilon^2 = 1$ and $s = 1000$ independent realizations. Thus for a given estimation method we obtain $s = 1000$ estimated values for d. Moreover, to evaluate the effects of different sample size, we have considered $T = 500$, 1000, 2000. All series are generated with 200 additional values in order to obtain random starting values.

In the following tables we present some of the results we obtained. For briefness' sake we report here only tables for $T = 2000$, but other results for $T = 500, 100$ that are not different from the case $T = 2000$, are available upon request by the authors.

From a general point of view, the perfomance of all methods becomes poor as the sample size increases. This could be justified by thinking that as the series become longer, the non stationary behaviours are more emphasized, therefore a structural break process is more similar to a long memory one.

It is also worth noticing that the Whittle estimator exhibits always the best performance. Indeed, if compared to the other techniques, the estimates obtained with the Whittle method are often much closer to zero over all the DGP's we considered. On the other hand, it is interesting to point out that the GPH technique has often a bad perfomance. This is because the GPH procedure focuses on the elasticity of the spectral density close to zero frequency: when $p = 0$ (no break), the elasticity is 0 which implies no long memory in the frequency domain, but when $0 < p < 1$ and p is fixed, then the elasticity tends to one as the frequency tends to zero (for details, see Granger and Hyung, 2004).

Going into further details, if we read table 1 about DGP1 we can notice that for all methods the bias tend to grow with the increase both of the number of expected breaks Tp, and the size of jump σ_η. In those conditions the level of similarity between long memory and structural break processes is much higher and because of this it is more difficult for the estimation methods to distinguish the two different typologies.

As we already pointed out, the Whittle method is the one which performs best. Indeed, for low values of σ_η the Whittle estimation manages to approach zero quite well even when p reaches the higher value. At the same time when the expected number of breaks is not too high the Whittle method produces an estimate close to zero also when σ_η is big. This good performance of the Whittle method is very interesting and we are still working on a justification for it.

The Higuchi and GPH methods are those which perform worst. Indeed, only for the lowest value of p and σ_η the two estimates manage to be reasonably close to zero.

If we read table 2 about DGP2, first of all we notice that the general performance of all methods is worse than in case of DGP1. Even the Whittle method does not manage to produce an estimate close to zero, although it still performs better than the other procedures.

When $p = q$ the unconditional probability of having a break is always $1/2$. If p and q increase, the bias reduces, but the variabiliy of the estimates grows because the probability of remaining in the same initial state is bigger and so the estimate can be evaluated from very low values of d to high ones. For instance, if the initial state is 1 and $p = 0.999$ it is very likely that the process will remain in state 1, so we would end up having a $I(1)$ process, on the contrary, but following the same logic, if the initial state is 0, and $q = 0.999$, we would end up with an $I(0)$ process.

When $p \neq q$ the performances are definitely not symmetric and that is because the unconditional probabiliy are different. The latter can be computed following Hamilton (1989) $\tilde{p} = \frac{1-q}{2-p-q}$. So, as expected, when $p = 0.95$ and $q = 0.99$ the performance is generally better since there is a higher probability of remaining in the state 0 and so the process has a more stationary appearance.

As far as tables 3 about DGP3 are concerned, first of all we can notice that with the increase of γ the performance of the estimator improves since the size of the jump becomes smaller. Also we observe that the GPH and the Whittle methods seem to be more sensitive to the variations of γ, than the other methods. However the standard error of the Whittle method is always the smallest.

For all the DGP's we considered the results we obtained are consistent with previous results presented in literature (Granger and Hyung, 2004; Diebold and Inoue, 2001), where, in fact, only the GPH (in test version) has been used.

4 Conclusions

In this article we have presented further Monte Carlo evidence about the issue of distinguishing between a time series process exhibiting long-range dependence and one with short memory but suffering from structural shifts. We considered several different estimation methods of the long memory parameter d, in particular we focused our attention on the GPH method, Whittle, Higuchi, aggregate variance and riscaled range.

We considered three different DGP's among the most used in recent literature on this topic, the mean plus noise, the Markov switching and the STOPBREAK model.

We found out that almost all the estimation methods tend to be biased and this is because the process with an occasional breaks looks very similar to a long memory one, especially when the size of the jumps is relatively big and also it is high the expected number of breaks.

The Markov switching model is the one where it is too difficult to detect its structural breaks

Table 1: Estimation results for d: DGP1, T=2000

σ_η^2		$p = 0.0025$	$p = 0.005$	$p = 0.01$	$p = 0.05$	$p = 0.1$
	rs	0.096 (0.101)	0.143 (0.111)	0.206 (0.124)	0.408 (0.149)	0.493 (0.149)
	av	0.051 (0.088)	0.101 (0.094)	0.153 (0.104)	0.310 (0.091)	0.368 (0.073)
0.005	hi	0.189 (0.155)	0.265 (0.152)	0.323 (0.131)	0.416 (0.084)	0.440 (0.074)
	gph	0.071 (0.124)	0.124 (0.122)	0.174 (0.132)	0.373 (0.144)	0.486 (0.147)
	wh	0.011 (0.014)	0.019 (0.019)	0.030 (0.026)	0.091 (0.042)	0.130 (0.047)
	rs	0.143 (0.119)	0.226 (0.127)	0.303 (0.135)	0.515 (0.154)	0.580 (0.151)
	av	0.097 (0.102)	0.166 (0.104)	0.235 (0.096)	0.371 (0.072)	0.407 (0.060)
0.01	hi	0.250 (0.152)	0.324 (0.131)	0.375 (0.111)	0.439 (0.077)	0.461 (0.062)
	gph	0.118 (0.130)	0.182 (0.132)	0.268 (0.124)	0.487 (0.140)	0.596 (0.142)
	wh	0.019 (0.020)	0.034 (0.028)	0.054 (0.035)	0.130 (0.048)	0.170 (0.052)
	rs	0.304 (0.180)	0.408 (0.167)	0.501 (0.157)	0.681 (0.155)	0.706 (0.157)
	av	0.223 (0.126)	0.302 (0.103)	0.362 (0.080)	0.445 (0.045)	0.458 (0.039)
0.05	hi	0.361 (0.133)	0.412 (0.102)	0.438 (0.084)	0.466 (0.056)	0.471 (0.056)
	gph	0.263 (0.168)	0.378 (0.166)	0.482 (0.145)	0.736 (0.142)	0.816 (0.134)
	wh	0.059 (0.041)	0.089 (0.048)	0.127 (0.051)	0.242 (0.059)	0.292 (0.063)
	rs	0.385 (0.192)	0.493 (0.170)	0.579 (0.164)	0.710 (0.151)	0.729 (0.156)
	av	0.281 (0.122)	0.350 (0.088)	0.402 (0.067)	0.461 (0.038)	0.470 (0.031)
0.1	hi	0.393 (0.121)	0.423 (0.091)	0.454 (0.071)	0.475 (0.052)	0.474 (0.052)
	gph	0.347 (0.183)	0.470 (0.163)	0.584 (0.155)	0.826 (0.126)	0.895 (0.121)
	wh	0.083 (0.051)	0.122 (0.055)	0.169 (0.059)	0.352 (0.059)	0.352 (0.059)

Table 2: Estimation results for d: DGP2, T=2000

σ_η^2		$p = 0.95$ $q = 0.95$	$p = 0.95$ $q = 0.99$	$p = 0.99$ $q = 0.95$	$p = 0.99$ $q = 0.99$	$p = 0.999$ $q = 0.999$
	rs	0.669 (0.148)	0.552 (0.165)	0.714 (0.151)	0.655 (0.168)	0.559 (0.265)
	av	0.446 (0.045)	0.384 (0.079)	0.461 (0.038)	0.434 (0.059)	0.376 (0.151)
0.005	hi	0.471 (0.055)	0.441 (0.084)	0.474 (0.049)	0.462 (0.062)	0.409 (0.152)
	gph	0.736 (0.133)	0.533 (0.159)	0.826 (0.123)	0.715 (0.148)	0.637 (0.264)
	wh	0.240 (0.055)	0.151 (0.055)	0.293 (0.058)	0.231 (0.064)	0.204 (0.103)
	rs	0.713 (0.148)	0.636 (0.158)	0.728 (0.165)	0.724 (0.157)	0.601 (0.304)
	av	0.462 (0.037)	0.424 (0.059)	0.469 (0.032)	0.458 (0.041)	0.381 (0.169)
0.01	hi	0.476 (0.051)	0.455 (0.069)	0.471 (0.058)	0.473 (0.054)	0.409 (0.165)
	gph	0.824 (0.120)	0.670 (0.169)	0.881 (0.114)	0.816 (0.142)	0.702 (0.314)
	wh	0.295 (0.057)	0.205 (0.064)	0.353 (0.062)	0.291 (0.065)	0.252 (0.128)
	rs	0.741 (0.213)	0.742 (0.147)	0.755 (0.208)	0.768 (0.184)	0.683 (0.315)
	av	0.476 (0.028)	0.464 (0.038)	0.478 (0.028)	0.473 (0.032)	0.410 (0.159)
0.05	hi	0.475 (0.055)	0.475 (0.052)	0.476 (0.055)	0.476 (0.050)	0.419 (0.163)
	gph	0.965 (0.104)	0.870 (0.135)	0.964 (0.097)	0.947 (0.109)	0.837 (0.303)
	wh	0.496 (0.061)	0.330 (0.073)	0.495 (0.059)	0.431 (0.067)	0.384 (0.157)
	rs	0.744 (0.224)	0.779 (0.169)	0.739 (0.243)	0.788 (0.214)	0.695 (0.298)
	av	0.478 (0.026)	0.467 (0.039)	0.481 (0.025)	0.477 (0.030)	0.420 (0.141)
0.1	hi	0.481 (0.047)	0.472 (0.059)	0.479 (0.048)	0.480 (0.049)	0.426 (0.142)
	gph	0.968 (0.096)	0.870 (0.135)	0.987 (0.096)	0.975 (0.109)	0.837 (0.272)
	wh	0.502 (0.061)	0.330 (0.073)	0.573 (0.062)	0.501 (0.069)	0.448 (0.162)

Table 3: Estimation results for d: DGP3, T=2000

γ	10^{-5}	10^{-3}	10^{-1}	1	10	10^{2}	10^{3}
rs	0.725 (0.384)	0.702 (0.369)	0.692 (0.362)	0.708 (0.327)	0.741 (0.225)	0.611 (0.156)	0.119 (0.098)
av	0.482 (0.026)	0.479 (0.026)	0.480 (0.026)	0.481 (0.026)	0.478 (0.028)	0.421 (0.056)	0.070 (0.083)
hi	0.474 (0.049)	0.478 (0.051)	0.479 (0.049)	0.479 (0.052)	0.476 (0.053)	0.458 (0.067)	0.244 (0.139)
gph	1.002 (0.104)	1.003 (0.101)	1.004 (0.108)	1.002 (0.103)	0.976 (0.097)	0.634 (0.139)	0.089 (0.117)
wh	0.992 (0.010)	0.992 (0.010)	0.968 (0.019)	0.841 (0.049)	0.542 (0.062)	0.191 (0.052)	0.014 (0.015)

nature, probably because in this case the shifts take place gradually, giving the impression of a local trend more than a break point.

Among our future research lines we can mention:

1. the analysis of the other side of the problem, that is to understand the performance of the tests to detect structural breaks in hypothesis of long memory data generating processes;
2. to develop an appropriate test procedure to distinguish real long memory from spurious one generate by infrequent breaks;
3. the problem of forecasting with long memory processes when the real DGP is a switching regime process and viceversa.

These future developments are currently under way by the authors with some first encouraging results. They are not completely available by now because of the computational burden of the simulation design.

References

Bai, J. (1998). A note on spurious break *Econometric Theory*, **14**, 663–669.

Beran, J. (1994). *Statistics for Long-Memory Processes.* Chapman and Hall, London.

Chen, C. and Tiao, G.C. (1990). Random level-shift time series models, ARIMA approximations, and level-shift detection *Journal of Business and Economics Statistics*, **8**, 83–97.

Diebold, F.X. and Inoue, A. (2001). Long memory and regime switching *Journal of Econometrics*, **105**, 131–159.

Engle, R.F. and Smith, A.D. (1999). Stochastic permanent breaks *The Review of Economics and Statistics*, **81**, 553–574.

Gourieux, C. and Jasiak, J. (2001). Memory and infrequent breaks. *Economics Letters*, **70**, 29–41.

Granger, C.W.J. (1980). Long-memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, **14**, 227–238.

Granger, C.W.J. and Ding, Z. (1995). Some properties of absolute returns. An alternative measure of risk. *Annales d'Economie et de Statistique*, **40**, 67–91.

Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-range time series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15–30.

Granger, C.W.J. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S$P500 absolute stock returns. *Journal of Empirical Finance*, **11**, 399–421.

Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**, 357–384.

Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, **68**, 165–176.

Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **116**, 770–799.

Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics, **5**, 299–314.*

Mikosh, T. and Starica, C. (1999). Change of structure in financial time series, long range dependence and the GARCH model. Working Paper *The Wharton School, Chalmers University of Technology*

Robinson P. M. (1995) Log-periodogram regression of time series with long range dependence The Annals of Statistics*, **23**, (1995) 1048-1072.*