

Wavelet–domain test for long–range dependence in the presence of a trend *

Agnieszka Jach and Piotr Kokoszka
Utah State University

July 24, 2007

Abstract

We propose a test to distinguish a weakly–dependent time series with a trend component, from a long–memory process, possibly with a trend. The test uses a generalized likelihood ratio statistic based on wavelet domain likelihoods. The trend is assumed to be a polynomial whose order does not exceed a known value. The test is robust to trends which are piecewise polynomials. We study the empirical size and power by means of simulations and find that they are good and do not depend on specific choices of wavelet functions and models for the wavelet coefficients. The test is applied to annual minima of the Nile River and confirms the presence of long–range dependence in this time series.

AMS 2000 Subject classification. 62M10 62M15 60G18 42C40.

Keywords and phrases: Generalized likelihood ratio, long–range dependence, polynomial trend, wavelets.

1 Introduction

Long–range dependent (LRD) processes, also known as long–memory processes, have been extensively used and studied in the past few decades. The relevant literature that has accumulated is too extensive to attempt even a limited review here, so we instead refer the reader to the collection Doukhan *et al.* (2002) which covers not only the most recent developments in the theory and applications of LRD processes, but also contains a number of review chapters tracing the historical development of important facets of these processes.

Realizations of stationary LRD processes exhibit long, non–periodic cycles which, in finite samples, resemble non–stationary behavior with trends and/or level shifts. In

*Research supported by NSF grants DMS-0413653 and INT-0223262 and NATO grant PST.EAP.CLG 980599 .

modeling geophysical phenomena, e.g. climatic variations, it is of importance to know if a given geophysical record is better described by a stationary model with long-range dependence or by a weakly dependent model with long term trends or abrupt changes. If the former model is more suitable, the observed record does not indicate a broadly understood shift in a geophysical pattern, e.g. a climatic change; what is observed is merely a manifestation of the long non-periodic cycles present in a long-range dependent model. Such a conclusion can be reached if a weakly dependent time series model, possibly with a trend or change-points, can be rejected.

The goal of the present paper is to explore a test which allows to reject the null hypothesis of weak dependence even if the time series contains a smooth trend or a few change-points. Roughly speaking, under the null hypothesis the observed time series X_0, X_1, \dots, X_{N-1} follows the model $X_t = Y_t + m_t$, where the process $\{Y_t\}$ is stationary and weakly dependent and m_t is a deterministic function, and under the alternative the X_t follow an LRD model. The testing problem is formulated precisely later in this section where the relevant background and research are also reviewed. The test is constructed in the wavelet domain and is motivated by the recent work of Craigmile *et al.* (2005). The main reason why the wavelet domain is suitable for such a test is that (non-boundary) wavelet coefficients are invariant with respect to an additive polynomial trend, i.e. the wavelet coefficients of $\{Y_t + m_t\}$ and $\{Y_t\}$ are the same, provided that m_t is a polynomial of sufficiently low order. Since wavelet coefficients are localized in time, only very few of them will be influenced by discontinuities in the function m_t or its derivatives, provided the number of discontinuities is small relative to the length of the realization. Thus, in the wavelet domain, the process $\{Y_t + m_t\}$ looks very much the same as the process $\{Y_t\}$, so the deterministic trend is effectively eliminated from the testing problem. The next step is to find a test statistic based on the wavelet coefficients which has a known distribution, at least in an asymptotic sense, if the underlying process is weakly dependent, and which diverges, if it is LRD. In this paper, we focus on the generalized likelihood ratio statistic which can be easily computed because of special properties of the wavelet coefficients, which will be discussed in the following, and whose asymptotic distribution is known to be chi-squared.

It has been argued for some time that the observed manifestations of long-range dependence can be explained assuming that the observations are weakly dependent, but follow a non-stationary model, for example the model $X_t = Y_t + m_t$, introduced above, see Bhattacharya *et al.* (1983) and Giraitis *et al.* (2001). Diebold and Inoue (2001) argued that the appearance of long-memory can be explained by models whose parameters change or evolve with time. Karagiannis *et al.* (2002) demonstrated that many estimators of the memory parameter can be “fooled” in the presence of periodicity or a trend. There is by now ample evidence that, in finite samples, standard tools like ACF plots and periodogram-based spectral estimates behave in a very similar way for LRD processes and for certain types of nonstationarities. Most long-memory tests reject in the presence of a trend or change-points. There is often a controversy which of the two modeling approaches is more appropriate for a specific time series.

There has however not been much research that produced effective tools for dis-

tinguishing between long-range dependence and a trend with weakly dependent noise. Künsch (1986) developed theoretical foundations for a periodogram-based procedure to discriminate between a LRD process and the process $\{X_t = Y_t + m_t\}$ with a “small” monotonic function m_t . Heyde and Dai (1996) showed that procedures for detecting long-memory which are based on a smoothed periodogram are robust in the presence of “small” trends. These ideas were recently developed by Sibbertsen and Venetis (2003) who proposed a test based on a difference between the Geweke and Porter-Hudak (1983) estimator and its version based on the tapered periodogram. In the latter test, the observations are LRD under the null, so it is not comparable with the test proposed in this paper. Berkes *et al.* (2006) proposed a test for discriminating between long-range dependence and weak dependence with a change-point in mean, which is a time domain procedure based on a CUSUM statistic for the partial sums of observations. The test of Berkes *et al.* (2006) is however not suitable if the mean changes smoothly under the null. Teyssiere and Abry (2006) used a wavelet domain least-squares estimator of LRD, and showed that when applied to financial time series, this estimator shows a lower intensity of LRD than implied by other estimators. This is due to the robustness of this estimator to trends and change-points.

We now formulate precisely the testing problem and describe the testing procedure in greater detail.

We assume that the observations follow a Gaussian process both under the null and the alternative. If the weakly dependent process $\{Y_t\}$ has absolutely summable autocovariance function and its spectral density is positive at every frequency, then it admits both autoregressive and moving average representation of infinite order with absolutely summable coefficients. Such a process can thus be approximated in mean square by a causal and invertible ARMA(p, q) process. We therefore postulate that under the null $\{Y_t\}$ is an ARMA(p, q) process and m_t is a polynomial. As we will see in Section 5, we may in practice assume that m_t is a piecewise polynomial, but the theoretical arguments are available only if m_t is a polynomial. As a model for the LRD process under the alternative we use the fractional ARIMA model with the differencing parameter $\delta > 0$ and the same order p, q as for the ARMA process $\{Y_t\}$. We denote this model as ARFIMA(p, δ, q). The testing problem is thus formulated as follows:

Null Hypothesis. The observations X_0, X_1, \dots, X_{N-1} follow the model

$$(1.1) \quad X_t = Y_t + m_t, \quad 0 \leq t \leq N - 1,$$

where $\{Y_t\}$ is a causal and invertible Gaussian ARMA(p, q) process and m_t is a polynomial.

Alternative Hypothesis. The observations X_0, X_1, \dots, X_{N-1} follow model (1.1), where $\{Y_t\}$ is a Gaussian ARFIMA(p, δ, q) process with $\delta > 0$.

The trend function m_t need not be the same under the null and the alternative. The assumption of Gaussianity is used to construct the likelihoods in Section 3. Its practical relevance remains to be assessed.

The test is based on the approximate decorrelation property of the discrete wavelet transform (DWT) which asserts that the DWT coefficients, especially of an LRD process,

exhibit very small correlations within each level and between levels. The decorrelation property has been established through simulation and theoretical studies, see e.g. Section 9.1 of Percival and Walden (2000) and Abry *et al.* (2000, 2002), and references therein. The decorrelation property holds to a particularly good approximation for the non-boundary DWT (NBDWT) coefficients which are also not influenced by a polynomial of a sufficiently low order, see Section 2 for further details. Modeling the NBDWT coefficients within each level as either white noise or an AR(1) process and assuming that the coefficients at different levels are uncorrelated, as proposed in Craigmile *et al.* (2005), we can write down the likelihood function under both the null and alternative hypotheses and construct the generalized likelihood ratio (GLR) statistic. Since one more parameter, δ , is estimated under the alternative, the $-2 \log$ GLR has asymptotic $\chi^2(1)$ distribution.

The paper is organized as follows: We first review in Section 2 properties of the NBDWT coefficients of an ARFIMA process and the so-called white noise model for these coefficients. In Section 3 we introduce the GLR test procedure based on the white noise model. Section 4 contains a simulation study. We further investigate our procedure by applying it to a time series of Nile River yearly minimum water levels in Section 5. In Section 6, we summarize our findings and provide a broader perspective on the proposed procedure.

2 Discrete wavelet transform of the ARFIMA process

In this section we discuss the relevant properties of the NBDWT coefficients of the ARFIMA process. We use the notation and terminology introduced in Percival and Walden (2000).

Stationary causal and invertible ARFIMA(p, δ, q) process $\{X_t\}$ is defined by the difference equation

$$(1 - B)^\delta \Phi(B)X_t = \Theta(B)Z_t, \quad |\delta| < 0.5,$$

where

$$\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p, \quad \Theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$$

satisfy $\Phi(z) \neq 0$ and $\Theta(z) \neq 0$ for all $|z| \leq 1$, B is the backward shift operator, and $\{Z_t\}$ is a white noise (WN) sequence with mean 0 and variance σ^2 . If $\delta = 0$, the X_t follow the ARMA(p, q) model.

The spectral density of $\{X_t\}$ is

$$(2.1) \quad S_X(f) = \sigma^2 \frac{|\Theta(e^{-i2\pi f})|^2}{|\Phi(e^{-i2\pi f})|^2} |2\sin(\pi f)|^{-2\delta}, \quad |f| \leq 1/2.$$

If $\delta > 0$, the spectral density diverges at the origin and the process is seen to be LRD.

The GLR test described in Section 3 is based on the wavelet domain maximum likelihood estimation of the parameter vector $\boldsymbol{\beta}$ of the ARFIMA(p, δ, q) process defined as

$$(2.2) \quad \boldsymbol{\beta} = (\delta, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2), \quad \boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_p), \quad \boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_q).$$

The estimation is based on the autocovariance sequences of the NBDWT coefficients introduced below. Similar approach for the ARFIMA(0,δ,0) process is described in Sections 9.1–9.4 of Percival and Walden (2000) as well as in Craigmile *et al.* (2005).

Given a realization of a time series X_0, X_1, \dots, X_{N-1} , ($N = 2^J$, J – positive integer), the $N_j = 2^{J-j}$ DWT coefficients for the j -th level are obtained (theoretically) by filtering the data with a level j wavelet filter $\{h_{j,l} : l = 0, 1, \dots, L_j - 1\}$, where $L_j = (2^j - 1)(L - 1) + 1$ and L denotes the length of the corresponding wavelet filter (for example, Daubechies D(L) or least asymmetric LA(L) filter, see Section 4.8 of Percival and Walden (2000)). The transfer function for the level j filter is

$$H_{j,L}(f) = e^{-i2\pi(L_{j-1}-1)f} H_{1,L}(2^{j-1}f) \prod_{k=0}^{j-2} H_{1,L}(1/2 - 2^k f)$$

and its squared gain function is $\mathcal{H}_{j,L}(f) = |H_{j,L}(f)|^2$. Here $H_{1,L}(\cdot)$ is the Fourier transform of the wavelet filter.

Recall that $L'_j = \min(\lceil(L-2)(1-2^{-j})\rceil, N_j)$ DWT coefficients at level j are influenced by circular filtering (for more details see Comments and Extensions to Section 4.11 of Percival and Walden (2000)). Let

$$d_{j,k}, \quad j = 1, 2, \dots, J, \quad k = 0, 1, \dots, M_j - 1, \quad M_j = N_j - L'_j$$

denote the NBDWT coefficients. The exclusion of the boundary coefficients has two important consequences. Firstly, the NBDWT are “blind” to polynomials of order K , $K \leq L/2 - 1$. Secondly, it allows us to view the NBDWT coefficients as a sequence following approximately a white noise WN model, i.e. to a good approximation we may assume that the $d_{j,k}$ are uncorrelated. For a fixed j , the autocovariance sequence of the $d_{j,k}$ is, in fact, given by

$$(2.3) \quad s_{j,\tau}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta}) = \int_{-1/2}^{1/2} e^{i2\pi\tau f} S_j(f) df,$$

where

$$S_j(f) = 2^{-j} \sum_{k=0}^{2^j-1} \mathcal{H}_{j,L}(2^{-j}(f+k)) S_X(2^{-j}(f+k)),$$

and where $S_X(\cdot)$ is defined by (2.1). Under the WN model all autocovariances except at lag $\tau = 0$ are assumed to vanish. In the sequel, it is convenient to work with the quantities $c_{j,\tau}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})$ defined by

$$s_{j,\tau}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta}) = \sigma^2 c_{j,\tau}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta}).$$

Thus, under the WN model

$$(2.4) \quad d_{j,k} \sim \text{i.i.d. } N(0, c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})\sigma^2).$$

The quantities $c_{j,\tau}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})$ are explicitly given as

$$(2.5) \quad c_{j,\tau}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})$$

$$= \int_{-1/2}^{1/2} e^{i2\pi\tau f} 2^{-j} \sum_{k=0}^{2^j-1} \mathcal{H}_{j,L}(2^{-j}(f+k)) \frac{|\Theta(e^{-i2\pi 2^{-j}(f+k)})|^2}{|\Phi(e^{-i2\pi 2^{-j}(f+k)})|^2} |2\sin(\pi 2^{-j}(f+k))|^{-2\delta} df.$$

To speed up the calculations, what is particularly important for a simulation study, an approximation for $c_{j,\tau}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})$ is needed. We used so-called bandpass approximation, see Craigmile *et al.* (2005) for the details. In a few test cases, the results obtained by using this approximation were barely distinguishable from those relying on the exact formula (2.5).

In finite samples the wavelet coefficients $d_{j,k}$ exhibit a small lag-1 autocorrelation. To soak-up this dependence, it is often assumed that for any fixed j the $d_{j,k}$ follow an AR(1) model. In the next section we present the calculations for the WN model (2.4), similar formulas for the AR(1) can be readily obtained, and we used them in our simulation study.

3 The test procedure

Since the NBDWT coefficients do not depend on the polynomial m_t in (1.1), any test based on a statistic which is a function of the NBDWT coefficients can be reformulated as testing

$$(3.1) \quad H_0 : \delta = 0 \quad \text{against} \quad H_A : \delta > 0.$$

In fact, we could consider a broader class of alternatives in which the observations follow an ARFIMA model with a polynomial trend. Thus, a rejection of H_0 means that the data exhibit long-range dependence, possibly with a polynomial trend, but they are not weakly dependent with a polynomial trend.

Our task is thus to propose a test statistic corresponding to the testing problem (3.1). To do so, we combine the techniques described by Percival and Walden (2000) and Craigmile *et al.* (2005) with the generalized likelihood ratio (GLR) test, see e.g. Section 18.1 of Bain and Engelhardt (1991) for an elementary introduction to the GLR test.

Assuming the WN model for the NBDWT coefficients, the likelihood function takes the form

$$(3.2) \quad f(\underline{d}; \delta, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = \prod_{j=1}^J \prod_{k=0}^{M_j-1} (2\pi c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})\sigma^2)^{-1/2} \exp\left(-\frac{d_{j,k}^2}{2c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})\sigma^2}\right),$$

where the array

$$\underline{d} = \{d_{j,k} : j = 1, 2, \dots, J, k = 0, 1, \dots, M_j - 1\}$$

plays the role analogous to the vector of observations in the classical maximum likelihood estimation.

Twice the negative log likelihood, $-2 \log f(\underline{d}; \delta, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2)$ ($\log(\cdot)$ denotes the natural logarithm), is given as

$$-2 \log f(\underline{d}; \delta, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = M \log(2\pi\sigma^2) + \sum_{j=1}^J M_j \log(c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})) + \sum_{j=1}^J \sum_{k=0}^{M_j-1} \frac{d_{j,k}^2}{c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})\sigma^2}$$

$$= M \log(2\pi\sigma^2) + \sum_{j=1}^J M_j \log(c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})) + \sum_{j=1}^J \frac{R_j}{c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})\sigma^2},$$

where

$$R_j = \sum_{k=0}^{M_j-1} d_{j,k}^2 \quad \text{and} \quad M = \sum_{j=1}^J M_j.$$

Minimizing this expression with respect to σ^2 yields the maximum likelihood estimate of σ^2 as a function of the remaining parameters:

$$(3.3) \quad \hat{\sigma}^2(\delta, \boldsymbol{\phi}, \boldsymbol{\theta}) = \frac{1}{M} \sum_{j=1}^J \frac{R_j}{c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})}.$$

Replacing σ^2 by $\hat{\sigma}^2(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})$ in $-2 \log f(\underline{d}; \delta, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2)$ yields a function of δ , $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$ only, namely

$$\begin{aligned} & -2 \log f(\underline{d}; \delta, \boldsymbol{\phi}, \boldsymbol{\theta}, \hat{\sigma}^2(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})) \\ &= M(\log(2\pi) + 1) + M \log(\hat{\sigma}^2(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})) + \sum_{j=1}^J M_j \log(c_{j,0}(\delta, \boldsymbol{\phi}, \boldsymbol{\theta})). \end{aligned}$$

Recall now the definition (2.2) of the parameter vector $\boldsymbol{\beta}$ and introduce the parameter spaces

$$\Omega_0 = \{\boldsymbol{\beta} : \delta = 0\}, \quad \Omega = \{\boldsymbol{\beta} : \delta \geq 0\}.$$

Minimizing $-2 \log f(\underline{d}; \delta, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2)$ over Ω_0 leads to the estimators $\hat{\boldsymbol{\phi}}_0$ and $\hat{\boldsymbol{\theta}}_0$ whereas minimizing over Ω yields estimators $\hat{\delta}$, $\hat{\boldsymbol{\phi}}$, and $\hat{\boldsymbol{\theta}}$. We thus obtain two estimators of the parameter vector $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}}_0 = (0, \hat{\boldsymbol{\phi}}_0, \hat{\boldsymbol{\theta}}_0, \hat{\sigma}^2(0, \hat{\boldsymbol{\phi}}_0, \hat{\boldsymbol{\theta}}_0)), \quad \hat{\boldsymbol{\beta}} = (\hat{\delta}, \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2(\hat{\delta}, \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}})),$$

from which we can construct the GLR statistic

$$(3.4) \quad \lambda(\underline{d}) = \frac{\max\{f(\underline{d}; \boldsymbol{\beta}) : \boldsymbol{\beta} \in \Omega_0\}}{\max\{f(\underline{d}; \boldsymbol{\beta}) : \boldsymbol{\beta} \in \Omega\}} = \frac{f(\underline{d}; \hat{\boldsymbol{\beta}}_0)}{f(\underline{d}; \hat{\boldsymbol{\beta}})}.$$

Under $H_0 : \delta = 0$, $-2 \log \lambda(\underline{d})$ converges to the $\chi^2(1)$ distribution, see e.g. Theorem 6.3.2 of Bickel and Doksum (2001). Therefore, the size α asymptotic GLR test rejects H_0 if $-2 \log \lambda(\underline{d}) > \chi_{1-\alpha}^2(1)$, where $\chi_q^2(r)$ denotes q -th quantile of the chi-square distribution with r degrees of freedom.

4 Simulation study

Design of the study. For our simulation study, we implemented the test procedure assuming the ARMA(1,0) model under the null and the ARFIMA(1, δ , 0) under the alternative. We assessed its finite sample performance based on at least $R = 500$ replications of data

generating processes of lengths $N = 512$ and $N = 1024$. We analyzed the simulated realizations with wavelet filters D(6) and LA(8), sufficiently long to use the bandpass approximation. We focused on nominal sizes $\alpha = 0.05$ and $\alpha = 0.10$ and considered both WN and AR(1) models for the NBDWT coefficients. According to the theory explained in Sections 2 and 3, the size and power of the test do not depend on a polynomial trend, provided the order of the polynomial does not exceed $L/2 - 1$, where L is the length of the wavelet filter. This was confirmed in several test cases we considered: by adding different polynomials of degree 0, 1 or 2, we obtained essentially the same sizes, with differences only slightly greater than machine precision. We report here sizes obtained with the polynomial

$$(4.1) \quad m_t = 0.25t^2/N$$

under the null hypothesis and no polynomial trend under the alternative.

When evaluating the empirical size, the process $\{Y_t\}$ in (1.1) was generated according to the AR(1) model with nine choices of ϕ , $\phi \in \{0.1, 0.2, \dots, 0.9\}$ and common variance $\sigma^2 = 1$. Only positive values of the autocorrelation coefficient were considered because realizations with $\phi \leq 0$ (and possibly a trend) do not resemble realizations of LRD processes and can be told apart from them by eye. To compute the empirical power of the test, we used realizations of ARFIMA(0, δ , 0) processes, also with nine different values of the parameter δ , $\delta \in \{0.10, 0.15, \dots, 0.50\}$, and common variance $\sigma^2 = 1$. For fixed ϕ , δ and N , the same replications were used to better assess the effect of the model assumed for the NBDWT coefficients and the type of the wavelet used.

All numerical experiments reported here were performed in R. We also implemented the procedure in Matlab and obtained very similar results, but the Matlab implementation was slower. R's optimization routines, "optimize" (minimization with respect to one variable) and "optim" (minimization with respect to two variables), were used to estimate the parameters. Both of them allow to specify the range of minimization. Due to the constraints on δ dictated by the testing problem, this flexibility plays an important role in the practical implementation of the test procedure. We minimized over $\delta \in [0, 0.99]$ under the alternative and over $\phi \in [-0.99, 0.99]$ under both null and the alternative, c.f. (3.4).

Discussion of the results. Empirical sizes and powers, for $N = 512$, and different choices of wavelet models, wavelet filters, and nominal sizes are presented in Figures 5.1–5.2. Results for $N = 1024$, additional results of a more comprehensive simulation study which includes some empirical calibration of the tests, as well as all numerical values and the R code are available upon request.

We begin with the discussion of the empirical sizes. Figure 5.1 includes the 95 percent asymptotic confidence bounds $\pm 1.96\sqrt{\hat{\alpha}(1 - \hat{\alpha})/R}$, where $\hat{\alpha}$ is the empirical size.

The method works well when the parameter ϕ of is between 0.1 and 0.8. The rejection probabilities become too large as ϕ approaches unity. This is not surprising, as for large values of ϕ the spectrum of an AR(1) process looks similar to the spectrum of an LRD process. In the following, we therefore focus our discussion on the cases with $0.1 \leq \phi \leq 0.8$.

By comparing the empirical sizes obtained under the assumption of the WN model

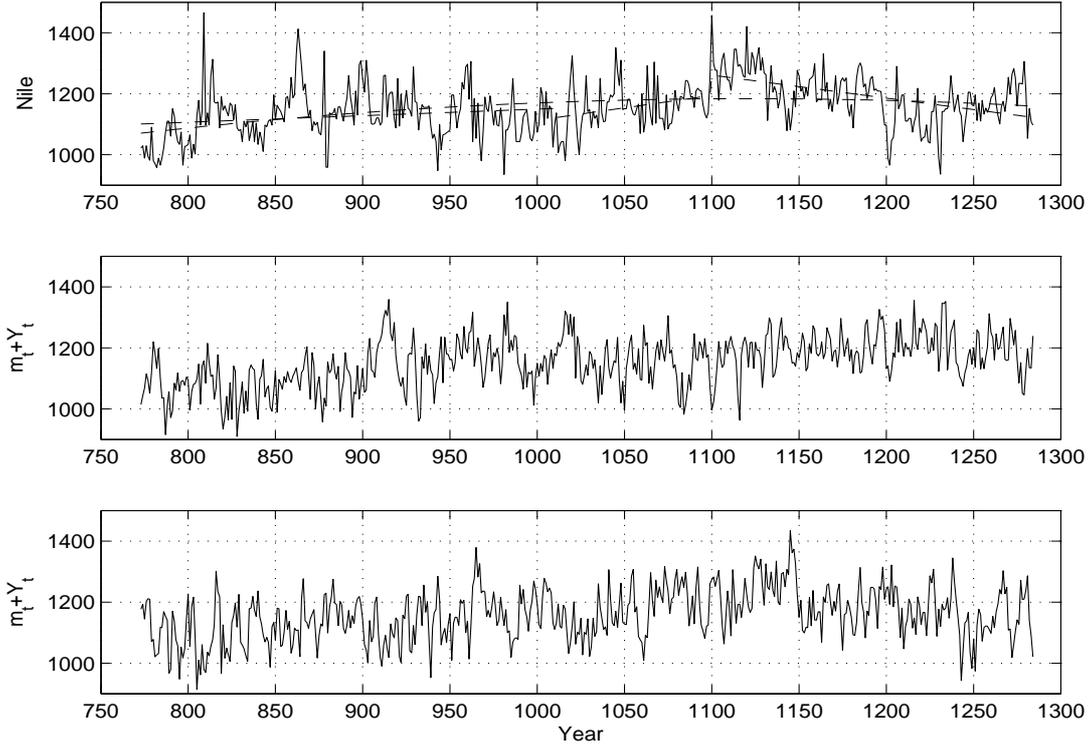


FIGURE 5.1 *Top:* Nile River annual minimum water levels for years 773 to 1284 with fitted second order polynomial trend and piecewise regression line. *Middle:* Simulated realization of $X_t = m_t + Y_t$, where m_t is the fitted polynomial and Y_t follows an AR(1) model fitted to the residuals. *Bottom:* Simulated realization of $X_t = m_t + Y_t$, where m_t is the fitted piecewise regression line and Y_t follows an AR(1) model fitted to the residuals.

for the NBDWT coefficients to those computed under the AR(1) model, we conclude that the former provides more accurate results. The choice of the wavelet filter, D(6) versus LA(8), does not substantially affect the results (compare top two panels to bottom two in Figure 5.1). The LA(8) filter should be used if the presence of a cubic polynomial trend is suspected ($K = 8/2 - 1 = 3$). Our test provides more accurate results for $N = 1024$ than for $N = 512$. This is in agreement with the asymptotic nature of the test.

We conclude this section with a brief discussion of the empirical power. The general shape of the power curve is very similar for the two significance levels considered and for fixed N (see Figure 5.2). The power is high and exceeds 80% for $N = 1024$ and $\delta \geq 0.2$. The power converges to 1 (with $\delta \rightarrow 1/2$) much faster for $N = 1024$ than for $N = 512$. The choice of the filter does not matter much.

5 Application to the annual minima of the Nile

We applied the test to Nile River yearly minimum water levels. The whole data set covers years 622 to 1284 and has been extensively studied in the long-memory literature. Here we focus on the last 512 observations plotted in the top panel of Figure 5.1. Visual examination of this time series suggests that rather than considering a long-memory model one might use an AR(1) model with a positive AR coefficient and with a smooth or discontinuous trend. Using the least squares method, we fitted a second order polynomial and a piecewise linear function to the Nile minima. The break points of the piecewise linear function were chosen somewhat arbitrarily to reflect the apparent level and slope shifts in the observations. We then estimated the AR(1) model on the two sets of the residuals. For comparison, we simulated realizations from the resulting two models and added them to the corresponding trends. Visual comparison of the two lower panels of Figure 5.1 with the top panel shows that an AR(1) model with a trend, especially with a piecewise linear trend, might be a reasonable alternative to an LRD model.

We applied our test procedure to the 512 observations assuming the WN and AR(1) models for the NBDWT coefficients and using both D(6) and LA(8) wavelet filters. The p-values presented in Table 5.1 are small, so we reject the null hypothesis and conclude that an AR(1) model with a trend is not suitable.

Wavelet filter	WN model	AR(1) model
D(6)	0.0033	0.0026
LA(8)	0.0155	0.0045

TABLE 5.1 P-values for the GLR test applied to $N = 512$ annual Nile minima.

We now provide a more detailed validation of our conclusion. The polynomial fitted to the data is

$$(5.1) \quad m_t = -0.0009t^2 + 2.0796t + 16.7069.$$

Fitting an AR(1) model to the residuals, we obtained the autoregressive coefficient $\phi = 0.5820$ and the WN variance $\sigma^2 = 4394$. The simulations in Section 4 show that for $\phi = 0.6$ our test has about the correct size. This is confirmed by additional simulation results presented in Table 5.2. For example, when the test with the WN model and the D(6) filter is applied to the estimated model, the empirical sizes of 10 and 5 percent level tests are, respectively, 10.8 and 5.5 percent.

Filter \ α	WN model		AR(1) model	
	0.10	0.05	0.10	0.05
D(6)	0.108	0.055	0.081	0.045
LA(8)	0.077	0.044	0.063	0.026

TABLE 5.2 Empirical size of the GLR test based on $R = 1000$ replications of $X_t = Y_t + m_t$ of length $N = 512$, where $\{Y_t\}$ follows AR(1) model with parameters $\phi = 0.5820$ and $\sigma^2 = 4333$ and m_t is given by (5.1).

The theory and simulations presented earlier in the paper do not apply to piecewise polynomial trends. It can however be expected that if there are relatively few break points compared to the length of the series, only few DWT coefficients will be affected by these breaks and the test will continue to have correct size. This is indeed confirmed by our simulations. The piecewise linear function we considered has constant slope over the following three periods: 773–1009, 1010–1099, 1100–1284. If $t = 0$ corresponds to year 773, it can be written as

$$(5.2) \quad m_t = \begin{cases} 0.2t + 934.9, & t = 0, 1, \dots, 236, \\ 0.8t + 299.4, & t = 237, 238, \dots, 326, \\ -0.8t + 2088.8, & t = 327, 328, \dots, 511. \end{cases}$$

The AR(1) model for the residuals has parameters $\phi = 0.5628$ and $\sigma^2 = 4333$. The empirical sizes are shown in Table 5.3. For the test with the WN model and the D(6) filter, the empirical sizes of 10 and 5 percent level tests are, respectively, 7.8 and 4.8 percent.

Filter \ α	WN model		AR(1) model	
	0.10	0.05	0.10	0.05
D(6)	0.078	0.048	0.067	0.036
LA(8)	0.083	0.044	0.071	0.033

TABLE 5.3 Empirical size of the GLR test based on $R = 1000$ replications of $X_t = Y_t + m_t$ of length $N = 512$, where $\{Y_t\}$ follows AR(1) model with parameters $\phi = 0.5628$ and $\sigma^2 = 4394$ and m_t is given by (5.2).

The ARFIMA(1, δ ,0) model estimated on the observations has parameters $\phi = 0.0660$, $\delta = 0.4013$, and the WN variance $\sigma^2 = 0.0028$. The empirical power of the test for this model is presented in Table 5.4. For all but one combinations of nominal size, wavelet filter, and wavelet model, the power exceeds 70 percent and is about 80 percent for the D(6) filter and 5 percent nominal significance level.

Filter \ α	WN model		AR(1) model	
	0.10	0.05	0.10	0.05
D(6)	0.838	0.773	0.851	0.817
LA(8)	0.769	0.693	0.801	0.711

TABLE 5.4 Empirical power of the GLR test based on $R = 1000$ replications of ARFIMA(1, δ ,0) process of length $N = 512$ with parameters $\phi = 0.0660$, $\delta = 0.4013$ and $\sigma^2 = 0.0028$.

6 Summary and discussion

Motivated by the work of Craigmile *et al.* (2005), we developed and investigated a wavelet domain test in which under the null the time series is weakly dependent with a polynomial

trend and under the alternative it is LRD, possibly with a polynomial trend. We assumed that the Y_t in (1.1) follow an ARMA(p, q) model under the null and ARFIMA(p, δ, q), $\delta > 0$ under the alternative. We investigated the finite sample performance of the test for $p = 1$ and $q = 0$. Our findings can be summarized as follows:

1. The test has about correct size for moderate weak dependence which can be quantified by the condition $0.1 \leq \phi \leq 0.8$. It performs noticeably better for $N = 1024$ than for $N = 512$. For $N = 512$ the test is somewhat conservative for $0.1 \leq \phi \leq 0.6$.
2. The test has very good power.
3. The WN model for the NBDWT coefficients is slightly better than the AR(1) model, as far as the empirical size is concerned.
4. Both wavelet filters, D(6) and LA(8), yield similar results.
5. The test can be applied if a piecewise polynomial trend is suspected, provided there are few break points relative to the length of the series.

The approach proposed here can be extended to different settings. The key requirement is that the spectral densities under the null and the alternative must be specified by a parametric model such that weak dependence corresponds to a fixed value of a memory parameter and long-range dependence to a range of values. A very natural alternative to the ARFIMA specification considered here is the fractional exponential model (FEXP) recently studied by Moulines and Soulier (1999), among others.

The optimal choice of the order p, q requires further investigation. From the practical point of view, AR(1) and ARMA(1,1) models can be used as good approximations to the autocovariance structure of a weakly dependent linear process.

References

- Abry, P., Flandrin, P., Taqqu, M. S. and Veitch, D. (2000). Wavelets for the analysis, estimation and synthesis of scaling data. In *Self-Similar Network Traffic and Performance Evaluation, New York* (eds K. Park and W. Willinger), pp. 39–88. Wiley (Interscience Division).
- Abry, P., Flandrin, P., Taqqu, M. S. and Veitch, D. (2002). Self-similarity and long-range dependence through the wavelet lens. In *Theory and Applications of Long Range Dependence, Boston* (eds P. Doukhan, G. Oppenheim and M. S. Taqqu). Birkhäuser.
- Bain, L. J. and Engelhardt, M. (1991). *Introduction to Probability and Mathematical Statistics*. Duxbury Press, Belmont, California.
- Berkes, I., Horváth, L., Kokoszka, P. and Shao, Q-M. (2006). On discriminating between long-range dependence and changes in mean. *The Annals of Statistics*, **34**, 1140–1165.
- Bhattacharya, R. N., Gupta, V. K. and Waymire, E. (1983). The Hurst effect under trend. *Journal of Applied Probability*, **20**, 649–662.

- Bickel, P. J. and Doksum, K. A. (2001). *Mathematical Statistics*, Second edn, volume I. Prentice Hall.
- Craigmile, P. F., Guttorp, P. and Percival, D. B. (2005). Wavelet-Based Parameter Estimation for Polynomial Contaminated Fractionally Differenced Process. *IEEE Transactions on Signal Processing*, **53**, 3151–3161.
- Diebold, F. and Inoue, A. (2001). Long memory and structural change. *Journal of Econometrics*, **105**, 131–159.
- Doukhan, P., Oppenheim, G. and Taqqu, M. S. (2002) (eds). *Theory and Applications of Long Range Dependence*. Birkhäuser, Boston.
- Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, 221–238.
- Giraitis, L., Kokoszka, P. S. and Leipus, R. (2001). Testing for long memory in the presence of a general trend. *Journal of Applied Probability*, **38**, 1033–1054.
- Heyde, C. C. and Dai, W. (1996). On the robustness to small trends of estimation based on the smoothed periodogram. *Journal of Time Series Analysis*, **17**, 141–150.
- Karagiannis, T., Faloutsos, M. and Riedi, R. (2002). Long-Range Dependence: Now you see it, now you don't! In *Global Internet*. IEEE.
- Künsch, H. (1986). Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability*, **23**, 1025–1030.
- Moulines, E. and Soulier, P. (1999). Broad band log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, **27**, 1415–1439.
- Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.
- Sibbertsen, P. and Venetis, I. (2003). Distinguishing between long-range dependence and deterministic trends. Technical Report SFB 475 TR 16-03. Universität Dortmund. Available at <http://www.sfb475.uni-dortmund.de/dienst/de/index.html>.
- Teyssiere, G. and Abry, P. (2006). Wavelet analysis of nonlinear long-range dependent processes. Applications to financial time series. In *Long memory in Economics* (eds G. Teyssiere and A. Kirman), pp. 173–238. Springer.

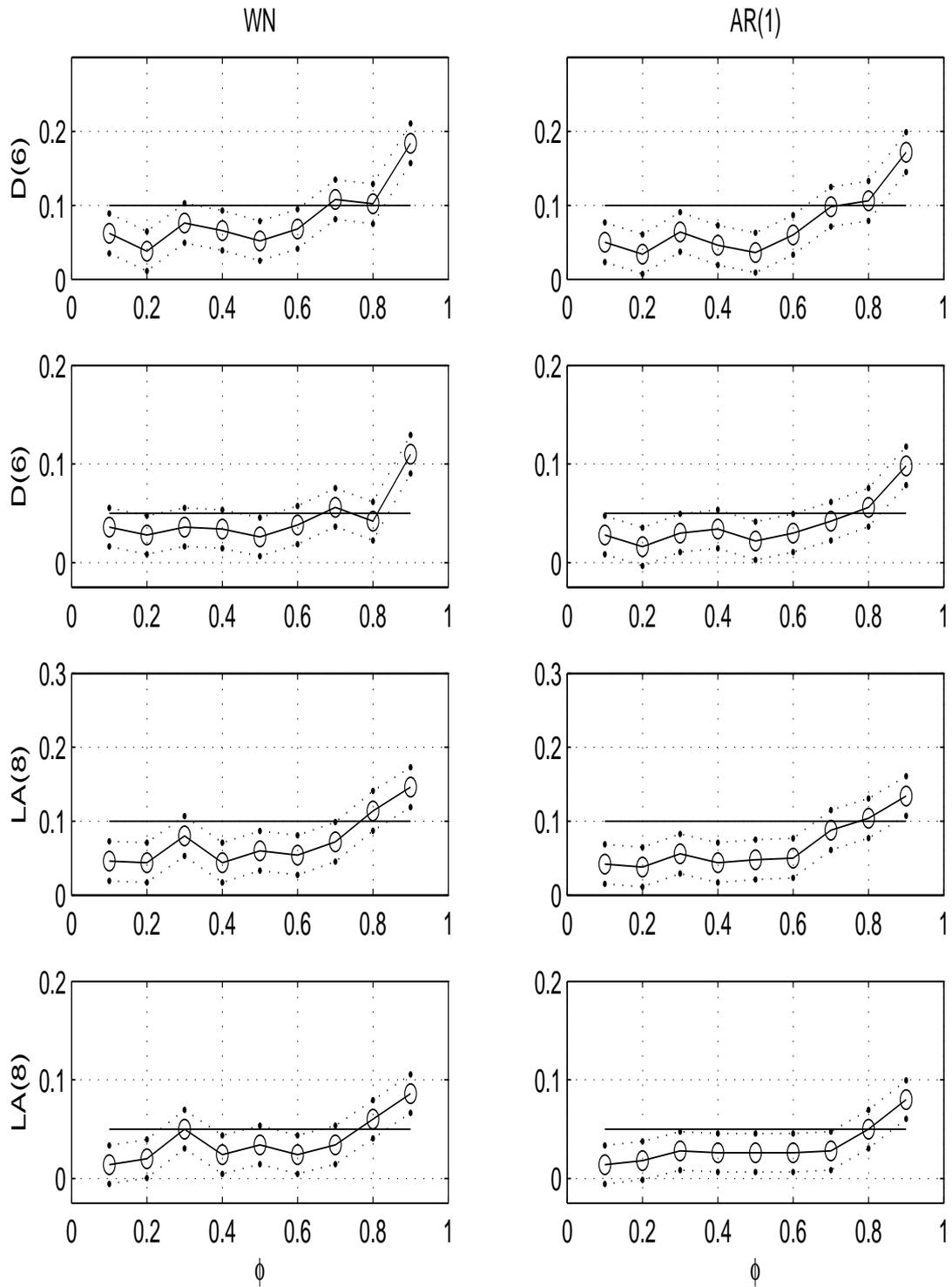


FIGURE 5.1 Empirical size of the GLR test based on $R = 500$ replications of $X_t = Y_t + m_t$ of length $N = 512$, where $\{Y_t\}$ follows AR(1) model with given ϕ and $m_t = 0.25t^2/N$. Wavelet filter: D(6), LA(8); model for the wavelet coefficients: WN (left panel) and AR(1) (right panel); nominal size indicated by the solid horizontal line.

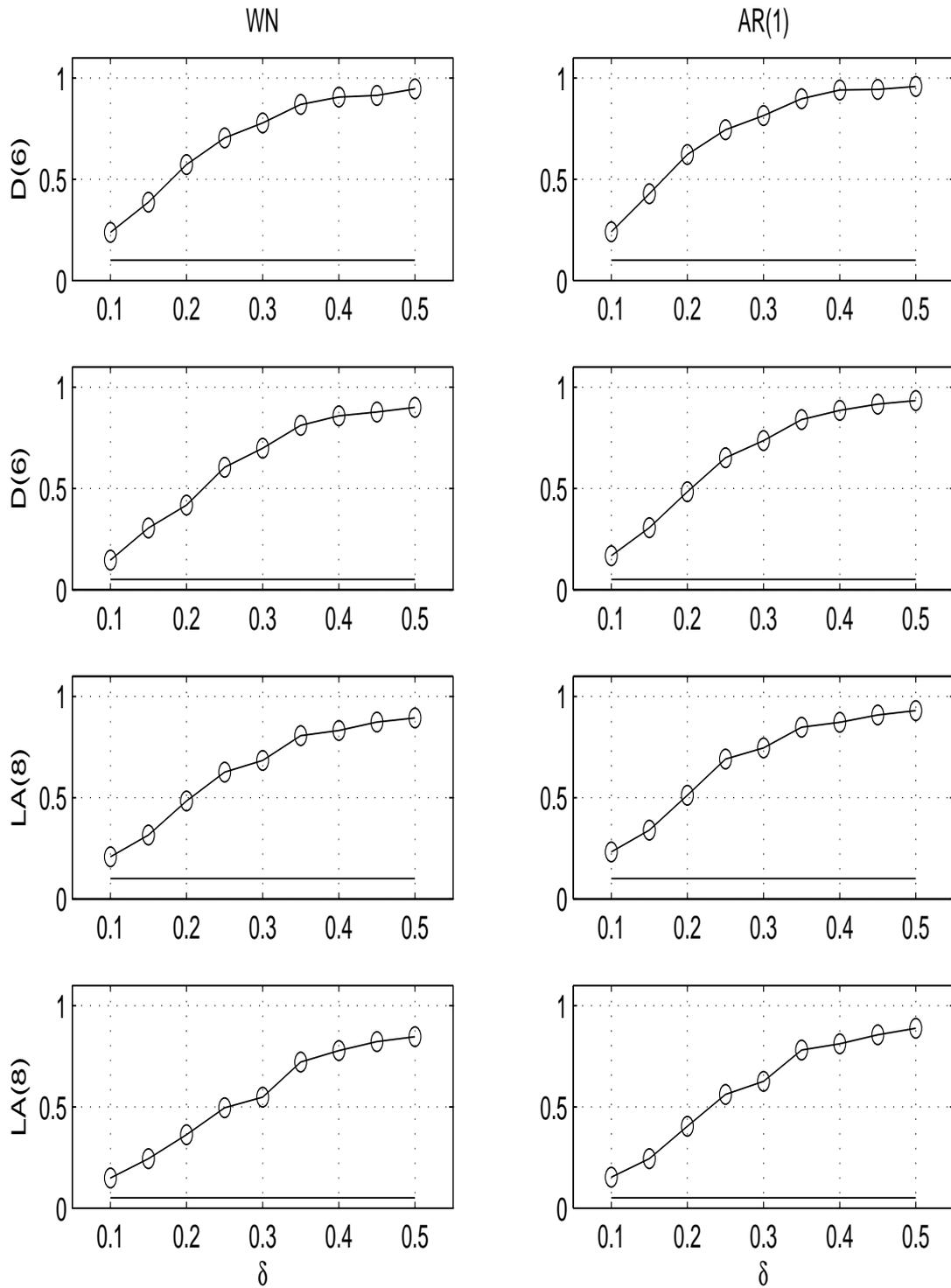


FIGURE 5.2 Empirical power of the GLR test based on $R = 500$ replications of X_t of length $N = 512$, where $\{X_t\}$ follows ARFIMA(0, δ ,0) model with given δ . Wavelet filter: D(6), LA(8); model for the wavelet coefficients: WN (left panel) and AR(1) (right panel); nominal size indicated by the solid horizontal line.